Loss of Setd1b methyltransferase in the murine forebrain as a novel model for human intellectual disability

Dissertation

for the award of the degree

"Doctor rerum naturalium"

of the Georg-August-Universität Göttingen

within the doctoral program "Genome Science"

of the International Max-Planck Research School

submitted by

Alexandra Michurina

from Moscow, Russia

Göttingen 2020

Thesis Committee

Prof. Dr. André Fischer, Epigenetics and Systems medicine in Neurodegenerative Diseases, German Center for Neurodegenerative Diseases, Göttingen

Prof. Dr. Bernd Wollnik, Institute of Human Genetics, University Medical Center, Göttingen

Prof. Dr. Tiago Outeiro, Experimental Neurodegeneration Department, University Medical Center, Göttingen

Members of the Examination Board

1st Referee: **Prof. Dr. André Fischer**, Epigenetics and Systems medicine in Neurodegenerative Diseases, German Center for Neurodegenerative Diseases, Göttingen

2nd Referee: **Prof. Dr. Bernd Wollnik**, Institute of Human Genetics, University Medical Center, Göttingen

Prof. Dr. Tiago Outeiro, Experimental Neurodegeneration Department, University Medical Center, Göttingen

Prof. Dr. Hubertus Jarry, Department of Research Animal Welfare Officer, University Medical Center, Göttingen

Prof. Dr. André Fiala, Behavioral Molecular Neurobiology, Georg August University, Göttingen

Dr. Ufuk Günesdogan, Department of Developmental Biology, Göttingen Center for Molecular Sciences

Date of oral examination: 26th of November 2020

Abstract

Methylation of histone 3 lysine 4 (H3K4me3) is an epigenetic mark for active transcription, which is mediated in mammals by six enzymes of the SET1/MLL family of histone methyltransferases (KMTs). All of them have been linked to schizophrenia, neurodevelopmental disorders and intellectual disabilities in humans, but the exact molecular functions of these enzymes in the adult brain remain unclear. In this study we have assessed the consequences of the conditional loss of function of one of the KMTs, the Setd1b histone methyltransferase, from excitatory neurons of the adult murine forebrain. We have evaluated the cognitive abilities of the conditional knock-out (cKO) mice through behavioral tests, analyzed changes in the chromatin landscape and gene expression profiles by neuron-specific ChIP sequencing, RNA sequencing and single cell RNA sequencing. We also made an attempt to restore the severe behavioral phenotype we observed for Setd1b cKO mice with the histone deacetylase inhibitor Vorinostat. In previous studies of our group, we have shown the importance of two other SET1 members for memory consolidation and epigenetic programming, *Kmt2a* and *Kmt2b*. However, by comparing all three cKO mice for the respective KMT, we found out that the loss Setd1b impacts neuronal identity genes most significantly and leads to a more drastic impairment of cognitive performance in mice. Our study suggests that Setd1b-related gene-expression programs could be targeted for the treatment of patients suffering from intellectual disabilities.

To my daughters.

Table of Contents

1. Introduction

The onset of big data science can be regarded as a paradigm change in our approach to understand diseases. Being able to read the genomics, epigenetics, transcriptomics, proteomics and posttranslational modifications of proteomics, gives us a unique opportunity to understand the underlying causes of disease in their molecular details. Of special interest in the epigenetics field are the so-called chromatin readers, writers and erasers, who are maintaining the sensitive equilibrium of the epigenome globally[1]. The ability to extract high resolution information about what these chromatin modifiers are doing in a specific patient due to increased efficacy and reduced costs of the new technologies, enable us to study rare diseases in a personalized medicine approach. It remains a captivating thought, that some diseases might find their origin in a single chromatin modifier mutation. If such a gene could be identified and the protein dysfunction efficiently treated through pharmacological agents, such diseases could be treated causally.

We are interested in neurological disorders, and the focus of the first part of my thesis lays on the Setd1b (SET Domain Containing 1B), a histone lysine methyltransferase. Being a global histone methyltransferase, the gene is implicated in multiple categories of biological functions, such as development, hematopoiesis, cancer and also neurological disorders[2]. Setd1b has been identified as one of the key affected genes in the very rare 12q24.31 syndrome, which is defined by intellectual disability, autism, epilepsy and craniofacial anomalies[3,4]. With the 12q24.31 deletion being the smallest known microdeletion syndrome up to date, it is tempting to assume that the affected genes also play a role in many other neurological disorders. As an example, 2-3% of the general population are affected by intellectual disability, of which one fourth can be attributed to a genetic disorder[5]. As for the year of 2013, 95 million cases of intellectual disability do

not have a known cause[6]. It is therefore highly likely, that the genes associated with the 12q24.31 syndrome are also affected in these unknown origin cases of intellectual disorder. If pharmacological treatment becomes available, diagnostic tests for such genes could enter the neurological clinical routine. For such neuro-drug development, basic research has to provide the knowledge foundation.

For Setd1b loss of function mutations in neurological disorder patients, DNA methylation analyses have shown hypermethylation of active promoters and enhancer regions[7], this might be due to Setd1b mediated H3K4 methylation loss at specific loci. However, it has not yet been shown whether Setd1b is interacting with these loci under physiological conditions. To answer this question, we have used a Cre-inducible Setd1b knockout mouse as the central model of our studies. We have assessed the cognitive abilities of the knockout mice through behavioral studies and the epigenetic loci affiliation of Setd1b through ChIP-seq and single-cell RNA-seq. Utilizing this powerful model with cutting edge sequencing technology, we wish to unravel the physiological functions of Setd1b on chromatin and behavior *in vivo*.

1.1. Chromatin modification

The chromatin structure in the eukaryotic nucleus is mainly determined by histone proteins. DNA is coiled around these histones, forming nucleosomes, which make up the basic structure of chromatin. Nucleosomes consist of histone octamers, a combination of each two H2A/H2B and two H3/H4 histone dimers, with 147 bp of DNA wrapped around them. This interaction is facilitated by the DNA negative and the histone positive charge. Posttranslational modifications (PTM), which are enzymatic alterations of specific histone residues, are able to modulate this interaction, which alters the density of chromatin packaging from loose euchromatin to highly condensed heterochromatin and vice versa[8].

Relaxation of chromatin enables transcription factors (TF) to access the DNA, which can either lead to the activation or repression of RNA transcription. The most prominent PTM include histone-methylation, acetylation, phosphorylation, ubiquitination and SUMOylation[9]. Histone-methylation will be the main focus of our work.

1.2.   Chromatin writers: Histone methyltransferases

Chromatin can be in an active, poised or silent condition, these states are mediated by histone PTMs[10] and appear to be primarily determined by histone methylation, creating a central histone methylation code. Epigenetic control via histone methylation is regulated by the opposing actions of histone methyltransferases (KMT) and lysine demethylases (KDM)[11]. Histone methylations dominantly occur on the sidechains of lysine and arginine on histones H3 and H4. Unlike acetylation and phosphorylation, methylation does not change the histone charge but the hydrophobicity. Sequential addition of methyl-groups results in histone mono-, di- or trimethylation, which leads to differential histone conformations[10]. Specific histone lysine residues can be subjected to methylation, which then impacts transcriptional initiation, elongation and chromatin silencing. Histone methylation has been confirmed at H3K4, H3K9, H3K27, H3K79, H4K20 histone lysine sites[12,13]. In our study, we focus the H3K4 methylation, which is associated with global transcription regulation and represents an epigenetic tag of promoter and enhancer regions. ChIP-seq studies of the human brain have revealed H3K4me3 marks to be organized in over thirty thousand peaks throughout the genome, they mostly reside on transcription start sites (TSS) in promoter proximal regions and correlate with RNA polymerase II occupancy at active gene expression sites[13]. In our work, we have utilized the chromatin immunoprecipitation technique with the next-generation sequencing technology (ChIP-seq) for the H3K4me3 histone modification, which enables us to detect

and characterize H3K4me3 mediated epigenetic functions in the Setd1b deficient genome in a base pair motif resolution.
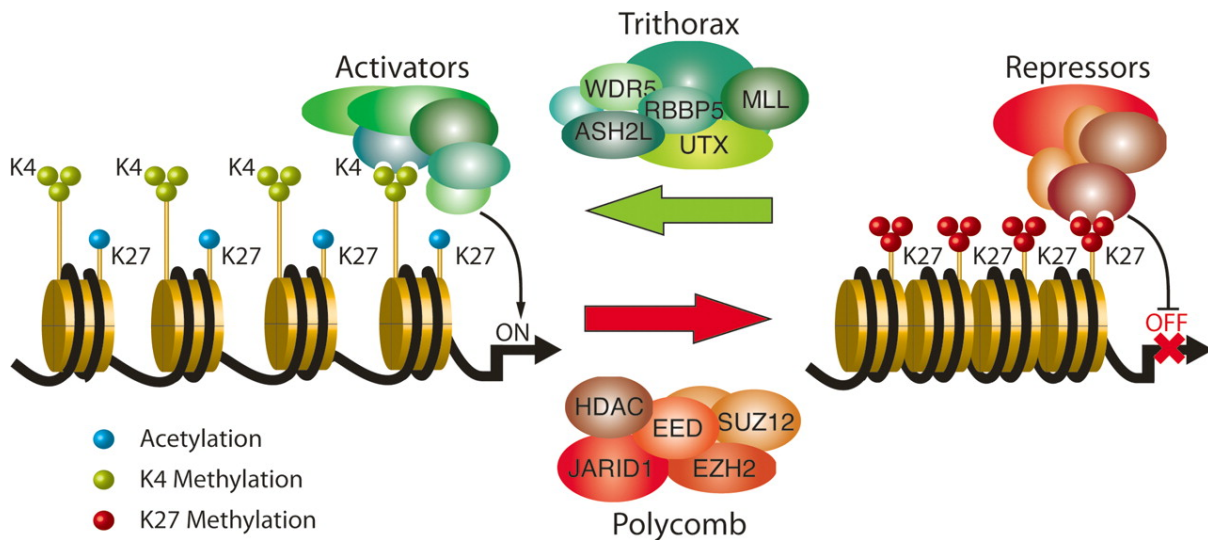
Figure 1-2-1: Post-translational histone modifications in activated and silenced states of chromatin[14]. The Trithorax group of proteins transfers the chromatin activating marks H3K4me3 and H3K27ac, leading to relaxation of the chromatin, whereas the Polycomb group of proteins promote the H3K27me3 chromatin repressing marks, which lead to the condensation of chromatin.

H3K4 methylation is mediated by the KMT2 (Lysine-N-methyltransferase 2) family of proteins, which is called the MLL (Mixed-lineage 10 leukemia) protein family in humans. The functional methyl group transferring domain of these proteins is the so called SET (Su(var)3-9, Enhancer(zeste) and Trithorax) domain, with the exception of Dot1, which methylates H3K79 without a functional SET domain[8]. The SET1 domain is conserved from yeast to humans, with H3K4 methylation being an important epigenetic modification in eukaryotic gene expression, playing important roles in transcription and chromatin dynamics. Initially discovered in yeast (*Saccharomyces cerevisiae*), which only possess one H3K4 methyltransferase, it was purified as the COMPASS (Complex Of Proteins Associated with SET1) complex, which has seven other components[15]. In fruit flies

(*Drosophila melanogaster*), the gene family already diversifies into three groups, the trithorax (Trx), trithorax-related (Trr) and Set1 proteins. In mammals, including humans, six SET-related methyltransferases can be found: Trx-related Mll1 (Kmt2a) and Mll2 (Kmt2b); Trr-related Mll3 (Kmt2c) and Mll4 (Kmt2d); and Set1-related Setd1a and Setd1b[16].
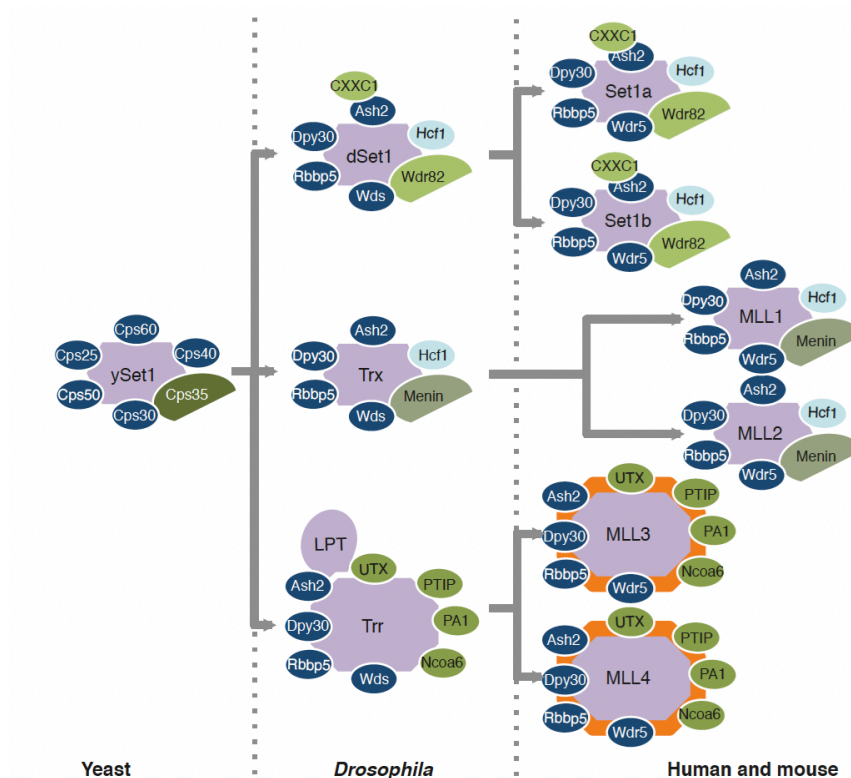
Figure 1-2-2: Taxonomy and evolution of Set1 proteins[17]. Set proteins are conserved among eukaryotes from yeast to humans. Diversification among homologues occurs with evolution, conservation of interaction partners can be observed. Mll1=Kmt2a, Mll2=Kmt2b, Mll3=Kmt2c, Mll4=Kmt2d.

SET1/MLL proteins alone are enzymatically inactive, therefore they are part of a larger complex, which consists of several sub-complexes (WRAD: WD-40 repeat protein 5 (WDR5), Retinoblastoma binding protein 5 (RBBP5), Absent small homeotic 2-like (ASH2L) and Dumpy-30 (DPY-30)), for specific methyltransferase activity towards H3K4[17]. WRAD proteins are part of the functional complex in all SET1/MLL transferases.

Further factors are Menin, Hcf1 (host cell factor 1) and Cxxc1 (CXXC-type zinc finger protein 1), which are only associated with a subgroup of family members. Evidence for the necessity of complex integrity comes from the observation that deletion of any mammalian complex subunit leads to a drastic reduction of H3K4 methylation[18], whereas the deletion of single SET1/MLL methyltransferases does so only minimally, most likely due to their redundancy[19].
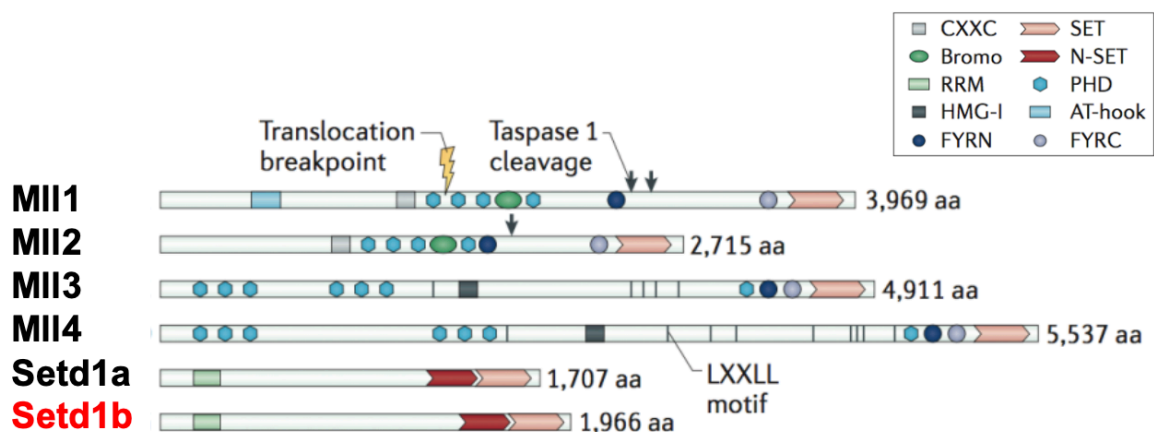
Figure 1-2-3: Protein domain structures of human histone methyltransferases: CXXC - finger protein 1, Bromo - bromodomains, RRM - RNA recognition motif, HMGI - high mobility group I, FYRN, PHD - plant homeotic domain, FYRC-FY-rich N-terminal, FYRC-FY-rich N-terminal[16]. Mll1=Kmt2a, Mll2=Kmt2b, Mll3=Kmt2c, Mll4=Kmt2d.

1.3. Chromatin erasers: Histone demethylases

To maintain chromatin homeostasis, histone methylation can be removed. This is achieved by a group of proteins called histone demethylases. Histone methylation was supposedly irreversible, as the half-life of histone methylation was as long as the half-life of histones themselves[20]. This perception was changed by the discovery of the first histone demethylating enzyme, Lysine-specific histone demethylase 1A (Lsd1) or Lysine (K)-specific demethylase 1A (Kdm1A), a homologue of nuclear amine oxidases[21]. Subsequently, two families of histone demethylases have been characterized, a flavin adenine dinucleotide

(FAD)-dependent amine oxidase group, and an α-ketoglutarate-dependent hydroxylase group, which furthermore needs an Fe(II) metal catalyst[22]. The molecular mechanism in both families is the hydroxylation of the target methyl group, followed by the dissociation of a formaldehyde molecule. Histone demethylases are large proteins with various domains, each with a specific function, which are histone, DNA or nucleosome binding, substrate recognition, enzyme function and cofactor binding. Biochemical cofactors include FAD, α-ketoglutarate and Fe(II)[23]. Important domains include zinc-finger domains for DNA binding, PHD-finger domains for substrate recognition, amino oxidase domains for the catalytic function and the Jumonji domain as a cofactor binding site[24].

A subclassification of histone demethylases can be made by distinguishing substrates specificity, this is done by the subdivision into the KDM1-KDM6 families. KDM1A recognizes mono- and di-methylated H3K4 and H3K9 residues, whereas KDM1B acts only on mono- and di-methylated H3K4 modifications. The members of the KDM1 family play crucial roles in embryogenesis, cellular differentiation and oogenesis[20]. Deletion of either KDM1A and KDM1B leads to embryonic lethality in mice[25–27].

The KDM2 family consists of KDM2A and KDM2B, in which KDM2A recognizes di-methylated H3K36 and tri-methylated H3K4 histone modifications and KDM2B only acts on mono- and di-methylated H3K36. They have been associated with functions in oncogenesis and tumor maintenance[20]. The KDM3 family consist of KDM3A-KDM3C, which can all act on mono- and di-methylated H3K9 modified lysine. Interestingly, KDM3A knockdown in mice leads to male infertility and adult onset-obesity[28], while the mutation has also been detected in two infertile human males[29].

The KDM4 family consists of the protein members KDM4A-KDM4D, they can all act on di- and tri-methylated H3K9, H3K36, H1K26 histone lysines, while their functions have been linked to tumorigenesis[20]. The KDM5 protein family of histone demethylases consists of KDM5A-KDM5D, which can all target di- and tri-methylated H3K4 methylated lysine.

KDM5 family demethylases have been associated with developmental functions, for instance, KDM5B is known to interact with PcG proteins[30], which are involved in transcriptional repression, and KDM5C loss of function, which is situated on the X-chromosome, has been associated with X-chromosome-linked mental retardation[31]. Finally, the KDM6 family of histone demethylases consists of KDM6A and KDM6B, which act on di- and tri-methylated H3K27 histone modifications. Their function has been implied in development, for instance, mutations of KDM6A has been associated with the Kabuki syndrome[32,33].

1.4. Histone acetyltransferases (HATs) and Histone deacetylases (HDACs)

Besides methylation, lysine residues of histone tails can also be subjected to acetylation. So-called histone acetylation plays an activating role in transcription[34]. Mechanistically, the addition of a negatively charged acetyl group counteracts the protonated lysine residue and therefore reduces its interaction with negatively charged DNA. Subsequently, DNA binding to nucleosomes is reduced, allowing a better access for the transcriptional machinery. Furthermore, acetylated lysine residues function as binding targets for transcriptional activators, most importantly for proteins that harbor a bromodomain, which can bind acetylated lysine residues specifically[35–38]. The equilibrium of histone acetylation on chromatin is maintained by the balancing activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs), with HATs stimulating and HDACs inhibiting active transcription[39,40]. Histone acetylation is mainly associated with active promoters, however, also coding regions and even telomeres have been found to exhibit a certain degree of acetylation. The functions of acetylated histones outside of active promoters are yet to be elucidated[41,42].

HATs consist of a large group of proteins, which are also referred to as Lysine (K) acetyltransferases (KATs). The nomenclature differentiates between type A nuclear HATs and type B cytoplasmatic HATs. Type B HATs consist of KAT1 and HAT4, they are located in the cytoplasm and modify histones before their transport and integration into chromatin[43], whereas type A HATs are mainly localized to the nucleus. Based on their sequence homology, they are categorized in the GNAT (Gcn5-related N-acetyltransferases) family, which comprises of KAT2A and KAT2B, the MYST family, consisting of KAT5-KAT8, and the p300/CBP family, which includes KAT3A and KAT3B[44]. The single HAT proteins exhibit different substrate specificities, for instance KAT3A and KAT3B are able to acetylate all H2A, H2B, H3 and H4 histone subtypes, whereas KAT6A only interacts with histone H3, and KAT8 only specifically acetylates H4K16[45,46]. This substrate specificity can be modulated by combining HATs with other factors in large protein complexes. As an example, the incorporation of KAT2A into the SAGA (Spt-Ada-Gcn5 acetyltransferase) complex significantly changes its substrate specificity and catalytic activity[47]. Their conserved sequence homology, differential substrate specificity and activity in various protein complexes render the clinical development of HAT pharmacological inhibitors a real challenge.

Similar to HATs, HDACs are also a large and heterogenous group of genes, with 11 mammalian genome encoded proteins subdivided into the class I, IIa and IIb, and class IV family of HDACs. The class I of HDACs includes HDAC1-HDAC3 and HDAC8, they are ubiquitously expressed and situated mainly in the nucleus[48]. The class IIa of HDACs consists of HDAC4,5,7 and HDAC9, they exhibit an N-terminal extension, which serves as a docking site for MEF2 (myocyte enhancer factor 2) and the 14-3-3 protein, which can bind to the HDAC after the latter has been phosphorylated by either the CaMK (Calcium/Calmodulin-dependent protein Kinase) or the PKD (Protein Kinase D) kinases[49,50]. The binding of HDAC to 14-3-3 leads to protein shuttling from the nucleus to

the cytoplasm[51]. The dissociation of MEF2 allows the binding of the HAT p300, which leads to the conversion of MEF2 function from transcriptional repression to activation[52,53]. This transactivation mechanism is able to implement extracellular cues in the regulation of transcription, and can be found in various tissues for different functions in development and disease[54]. However, the expression patterns of HDAC class IIa proteins are somehow more restricted than of the other HDAC families. HDAC5 and HDAC9 can be found to be highly enriched in muscle and heart[55,56], while HDAC4 is abundant in the brain and in skeletal tissues[57]. The HDAC class IIb consists of HDAC6 and HDAC10, with HDAC6 being the main cytoplasmatic deacetylase in mammalian cells[58]. It is also structurally unique, as it exhibits two domains with deacetylation function and a zinc finger at its C-terminal end.

1.5. Memory formation

One of the mysteries and ongoing frontiers of biomedical studies is the exact mechanism, how memory formation, consolidation and storage works in the brain. We know that it is highly complex and that it involves protein synthesis, differential gene expression, synapses and neuronal structure. A central brain structure of memory functions is the hippocampus, which has been experimentally validated in rodents and humans[59], furthermore, it is one of the first brain regions to be affected in dementia[60]. Anatomically, the hippocampus consists of two parts, the dentate gyrus (DG) and the cornu ammonis (CA). The CA has been subdivided in three zones, CA1-3.
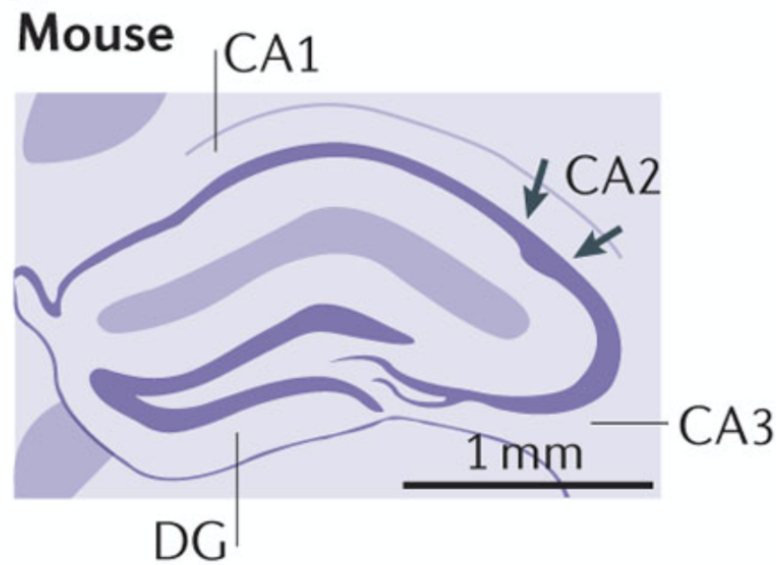
Figure 1-2-4: The anatomical hippocampal structure in the mouse[61]. CA – cornu ammonis, DG – dentate gyrus.

Neurophysiologically, long-term potentiation (LTP) has been identified as a main mechanism in memory formation. Calcium influx into the synaptic cleft triggers long-term plasticity, which in turn leads to gene expression and protein synthesis to promote long-lasting changes in neuronal structure[62]. New dendritic spines appear through de novo cytoskeletal protein biosynthesis and preexisting ones enlarge with enhanced AMPA receptor expression, which render the synapses more glutamatergic. In addition, cytoplasmatic calcium influx stimulates cAMP production, which activates the transcription factor CREB (cAMP-responsive element-binding protein) via the PKA/MAPK kinase pathways. CREB can be further activated through phosphorylation through other kinases such as CaMKII and RAS signaling, which activates the transcription of plasticity associated genes. CREB is thus a central activator in LTP and memory enhancement[63].
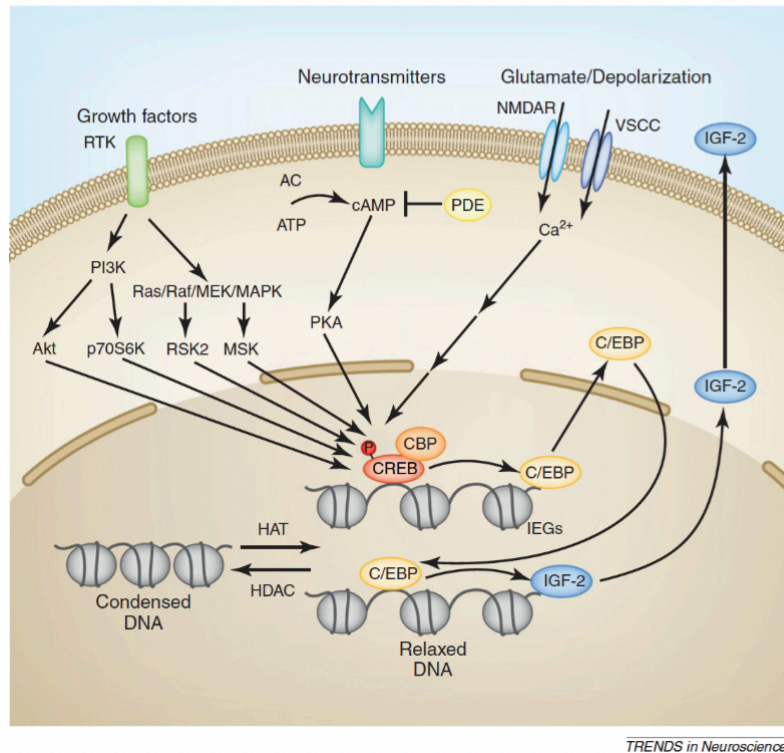
Figure 1-2-5: Molecular mechanism of LTP induction and gene transcription via CREB- activating pathways[63]. Extracellular signals are transduced via mitogen activated protein kinase (MAPK) signaling cascades, cAMP signaling and ion channel mediated calcium influx to activate CREB mediated transcription of plasticity associated genes.

1.6.   Transcription and gene expression in the brain

The epigenetic modification of H3K4me3 can render chromatin more accessible for transcription factors and RNAPII (RNA polymerase II), which is the main RNA polymerase of eukaryotic RNA transcription . Prior to RNAPII association with promoter regions downstream of transcription start sites (TSS), the chromatin is loosened up by transcription factors, which act as nucleosome remodelers[65]. Nucleosomes make space for the unphosphorylated RNAPII to recognize the promoter sequence, which then initiates transcription. RNAPII is subsequently phosphorylated at its carboxy terminal domain on Ser5, whereas Ser2 is only phosphorylated when the transcription reaches the gene body with elongation factor recruitment[66]. The basal transcription factor TFIID can recognize

H3K4me3 through its subunit TAFIII, which carries a conserved plant homeodomain (PHD) finger[67]. TFIID acts in the RNAPII complex, indicating active chromatin reading for histone methylation during transcription[68]. In the brain, rapid transcription of immediate-early genes (IEG), such as c-Fos, Egr1, Arc, plays a central role in memory formation. They have been shown to be plasticity evoked, stimuli induced and instantly upregulated upon neuronal activation or behavioral tasks. Rapid IEG has been shown to occur in the hippocampus during hippocampal specific learning tasks, such as the Morris Water Maze (MWM), novel environment exposure and contextual fear conditioning, but also exists in other brain regions[69].

1.7. Histone methyltransferases in development and neurogenesis

Histone methyltransferases have been shown to play a crucial role in embryogenesis and cell differentiation in different species[70,71]. They play a crucial role in the maintenance of stem cells, which is regulated by keeping chromatin transcriptionally permissive. Activity of bivalent promotors of developmental genes are balanced through Polycomb-group proteins (PcG), which transfer H3K27 repressing marks, and Trithorax-group proteins (TrxG), which deposit H3K4me3 activating marks on nucleosomes. The histone methyltransferase complex subunit Wdr5 is able to interact with the pluripotency transcription factor Oct4 and associated pluripotency maintenance factors Sox2 and Nanog through TrxG, which promotes the transcription of key self-renewing genes[72]. As an example, the promoters of Oct3 or Nanog in undifferentiated embryonic stem cells are enriched with H3K4me3 epigenetic modifications, while the majority of genes responsible for cell differentiation are not active or poorly expressed[73,74]. Furthermore, neurogenesis in the adult hippocampus has been shown to be under the control of epigenetic modifications[75].

1.8. Neurophysiological significance of lysine methylation and associated pathologies

Global and gene-specific elevation of H3K4me3 can be observed after contextual fear conditioning experiments in the murine hippocampus[76]. Conditional knockout mice for the Mll1 and Mll2 methyltransferases exhibit impairments in hippocampus-related learning and memory functions[77,78]. Furthermore, the loss of the H3K9 dimethylating enzyme G9a/G9a-like protein hinders memory formation in the hippocampus and impairs mouse normal cognitive development[79]. Pharmacological modulation of histone readers and writers are able to influence memory formation by hampering epigenetic regulation, which has been shown for histone deacetylase (HDAC) inhibitors[80] and the Bromodomain and Extra-Terminal motif (BET) inhibitor JQ1[81]. Also, the pharmacological disruption of lysine methylation leads to the impairment of hippocampal memory formation[82]. Taken together, memory formation in the hippocampus strongly depends on different epigenetic modifying proteins.

Dysregulated histone methylation patterns have been associated with a growing number of psychiatric and neurodevelopmental disorders[13,83]. To date mutations in all six human SET1 family members have been associated with neurodevelopmental disorders and syndromes. Kmt2a was identified as a causative gene for Wiedemann-Steiner syndrome characterized by intellectual disability, microcephaly, hypertrichosis and short stature[84]. The protein itself contains both transcriptional activator activity, linked to C-terminal, and repressor activity at N-terminal, enabling the balance of both activities[85]. Monoallelic variants in Kmt2b are associated with Kmt2b-related childhood-onset dystonia. Disease is associated with patterned, tremulous and repetitive movements and postures. Sometimes patients have additional clinical manifestations, such as facial dysmorphia, microcephaly and seizures[86]. De novo variants in Kmt2c (facilitating H3K4me1 and

H3K4me3) are associated with Kleefstra syndrome 2 and is characterized by intellectual disabilities (ID), autism spectrum disorders (ASD), aggressive/hyperactive behaviors and craniofacial dysmorphisms[87]. This syndrome is also associated with haploinsufficiency of euchromatic histone methyltransferase 1 (Ehmt1) (facilitating transcriptional repression via H3K9me1 and H3K9me2 marks)[88]. Kmt2d H3K4 trimethylating enzyme heterozygous loss-of-function is associated with Kabuki syndrome together with mutations in the complementary function enzyme, lysine-specific demethylase 6A (Kdm6a) responsible for removing H3K27me3 mark for closed chromatin[89]. Disease is accompanied by striking facial features, mild to moderate intellectual disability and postnatal growth deficiency[90]. Loss of function variants in Setd1a were identified as a risk factor for schizophrenia[91] including genome-wide association studies (GWAS)[92]. Schizophrenia is a severe neuropsychiatric disorder with the positive core symptoms (delusions and hallucinations associated with the loss of contact with a reality) and negative symptoms (social avoidance, lack of motivation and reduced speech)[93]. The role of Setd1b will be described in the coming section.

1.9. Histone acetylation in learning and memory

Histone acetylation of the CA1 hippocampal region was the first histone modification found to be important for memory formation[94,95]. Subsequently, also histone phosphorylation[96], DNA methylation[97,98] and histone methylation[76,77]. The main mechanism of histone acetylation in memory formation seems to be the process of memory consolidation[99]. The activating effects of histone acetylation on transcription, specifically on learning and memory associated genes, such as those for synaptic plasticity, have been proposed to play a mechanistic role in memory formation[100].

The pivotal role of histone acetylation in learning and memory has been shown for numerous diseases, including the Rubinstein-Taybi-Syndrome (RSTS), which is a rare genetic disease exhibiting skeletal abnormalities and intellectual disabilities with low levels of intelligence, associated with the loss of function of the HATs CBP or p300[101]. Specifically, impaired histone acetylation of H2A and H2B have been correlated with the syndrome[102]. CBP knockout mice resemble the human disease[103], and interestingly, these mice are proficient in short term memory learning tasks, but exhibit specific deficits in long term memory experiments, suggesting a role of histone acetylation in long term memory consolidation[104]. CBP loss of function has also been described for the Huntington disease (HD), an autosomal dominant genetic disease correlated with chorea, cognitive disorders and motor impairments[105], the pathophysiology is based on the expansion of an N-terminal polyglutamine stretch of the Huntingtin protein (Htt). In the HD, CBP is inactivated by different mechanisms, one being the sequestration into mutated Htt aggregates[106], direct inhibition by soluble huntingtin protein or proteasomal degradation[107]. As a result, a decrease in histone acetylation of H3 and H4 can be observed in cells expressing the mutant Htt[106]. The lack of soluble CBP protein leads to impaired transcriptional regulation of neuronal genes, which in turn results neuron dysfunction. Another important disease correlated with altered histone acetylation is the Alzheimer's disease (AD). AD etiology is multifactorial, factors contributing to its onset are aging, levels of Apolipoprotein E (ApoE) and the environment. For characterization, the toxic protein aggregates of beta amyloid and tau can be found in patient brains[108]. In AD, the balance of site directed actions between HATs and HDACs seems to be perturbed, chromatin regions of synaptic-plasticity related genes are hypo-acetylated, whereas genes involved in the generation of beta amyloids are hyper-acetylated[109].

The convincing evidence of altered histone acetylation in neuronal disease has prompted a strong interest in the development of pharmacological inhibitors against histone

acetylation modifying enzymes. A number of HDAC inhibitors have been designed to increase histone acetylation upon treatment, which do not exhibit a strong specificity for single HDAC isoforms[124]. Still, a clear clinical benefit can be seen. In the mouse model of RSTS, CBP heterozygous mice treated with Suberoylanilide hydroxamic acid (SAHA), also known as Vorinostat[111,112], a small molecule HDAC inhibitor that crosses the blood-brain-barrier, greatly improve their performance in cognition behavioral experiments[102]. It has also been shown, that administration of HDAC inhibitors can reduce the toxicity of polyglutamines as in the HD[113,114,115,116,117]. And finally, also for the AD a neuroprotective function of HDAC inhibitors has been shown in AD experimental mouse models[118,119].

1.10. Setd1b protein function in physiology and disease

In our study, we have investigated the functions of Setd1b in the mouse brain. Together with Setd1a, they are the smallest proteins of the KMT2 family. Their structure contains an N-terminal RNA recognition motif (RRM) and C-terminal N-SET plus SET/postSET domains. The Wdr82 protein only complexes with Setd1a and Setd1b proteins and is required for the interaction with RNAPII (RNA-polymerase II) via the RRM domain. Knockdown of Wdr82 leads to a decrease of H3K4me3 levels in human cell lines[120].

Being able to induce global changes in the epigenome, the Setd1b methyltransferase has distinct functions in physiology and disease. In normal physiology, Setd1b functions have been implicated in development, fertility and hematopoiesis. Consequently, diseases associated with Setd1b dysfunction include hematologic malignancies and cancer, but also and most interestingly, neurological disorders[2]. The multitude of Setd1b functions for the organism has certain implications for the establishment of a Setd1b mouse model: Setd1b knockout mice have been reported to die after gastrulation between E10.5 and E11.5[121], and also conditional knockout adult mice die after 30 weeks[2]. To study the function of Setd1b in the brain thus prompted the need for the establishment of a brain specific

conditional knockout mouse. To accomplish this goal, we have used a previously published transgenic mouse[122] in which Setd1b was brought under the control of the Cre recombinase. We have then crossed this strain with another mouse line in which the Cre is induced by the forebrain specific CaMKII promoter[123]. This ensured the birth of viable Setd1b homozygous forebrain knockout mice, which could be used for our behavioral experiments.

Setd1b has been reported to play a role in neurological disorders. Clinical representation of Setd1b loss of function are intellectual disability, autism, epilepsy, and craniofacial anomalies. Human genetic analysis of this rare condition reveal different mutational constellations, for one the 12q24.31 microdeletion, which involves a total of six genes, with the most important ones being Setd1b and the histone demethylase Kdm2b[3,4,7]. Further reported mutations are missense variants either in the SET[124] or RRM[125] domain, and a SET domain frameshift mutation[126]. As previously mentioned, Setd1b loss of function has been associated with DNA hypermethylation[7], which is an interesting observation and could relate to the pathological outcome of the mutation. Although a rare condition, Setd1b loss of function in patients could teach us a lot about the functions of chromatin writers in neurological diseases, as the 12q24.31 microdeletion is the smallest microdeletion syndrome we know up to date and the other patients seem to be single gene mutation cases.

1.11. The impact of sugar metabolism on the Alzheimer disease

In the second project, we have aimed to investigate the influence of sugar metabolism on the development of the Alzheimer disease (AD). As previously mentioned, AD onset is correlated with multiple risk factors and the regulation of epigenetic modifications is severely perturbed in AD patient brains[109]. One of the risk factors is altered insulin signaling, as patients suffering from diabetes type II have an increased risk to develop

AD[127]. Therefore, our hypothesis was to test whether the epigenetic signatures in the adult brain are dysregulated upon altered insulin signaling, which contributes to age-associated memory decline and AD. One of the pathways implicated in diabetes and AD is the insulin-like growth factor (IGFs) signaling pathway[128]. Most interestingly, while IGFs are known to be produced and act in the adult brain, there is also evidence that IGFs from the periphery impact on brain plasticity and memory functions[129,130]. One of the mechanisms to regulate the IGF signaling pathway is by IGF-binding proteins (IGFBPs), which prevent IGF action via binding and thus reducing their bioavailability[131]. Previous findings identified one member of the IGFBP family, IGFBP7, as a key factor in post-traumatic stress disorder (PTSD)[132] and as a risk factor for AD, IGFBP7 is up-regulated in AD mouse models, as well as in AD patient post-mortem brain tissues in an DNA-methylation dependent manner[133]. Unpublished data also showed that IGFBP7 levels are increased in blood samples from AD patients.

Insulin-like growth factor binding protein 7 (IGFBP7), also called IGFBP-related protein-1 (IGFBP-rP1) or Mac25, belongs to the family of insulin-like growth factor binding proteins (IGFBPs) in mammals. IGFBPs are secreted soluble proteins regulating the bioavailability of Insulin, Insulin-like growth factors I and II (IGF-I, IGF-II). Mammals have six proteins that can bind with high affinity to the two insulin-like growth factors (IGFs), IGF-I and IGF-II. These proteins are called IGF binding proteins 1-6 (IGFBP1 - 6) and have diverse roles in mammals[134]. For instance, IGFBPs prevent IGFs from binding to the insulin receptor and increase the half-life of IGFs[135]. The seventh member, insulin-like growth factor binding protein 7 (IGFBP7) however has a 100x lower binding affinity to IGF-I and IGF-II[136], but its affinity an estimated 500x higher to insulin[137]. Still , it remains an important regulating factor for the availability of IGFs in body physiology[138]. In the last decades studies of IGFBP7 reveal its suppressing effect on various cancer diseases like breast[139], prostate[140], colorectal cancer[141], hepatocellular carcinoma[142], acute

myeloid leukemia[143] and multiple myeloma[144]. Furthermore, IGFBP7 regulates physiological processes independent of IGFs or insulin, by inhibiting the BRAF/MEK/ERK signaling cascade, it is able to induce senescence and apoptosis[145]. IGFBP7 has also been reported to play a role in the formation and homeostasis of memory in the brain[132,146].

Alzheimer's Disease (AD) is one of the most common neurodegenerative disease worldwide[147]. AD is characterized by impairments in memory functions, personality / behavioral changes and a decrease of ability to execute known or learned activities. Initially, neuritic plaques and neurofibrillary tangles are observed in the temporal lobes of the brain with an eventual extension throughout the cortex[148]. The leading hypothesis of the pathogenesis of AD is the amyloid cascade hypothesis, which implies a central role for Aβ being responsible for oxidative injuries, synaptic / neuronal dysfunction and eventual neurodegeneration[149]. AD is caused by multiple genetic and environmental risk factors[150]. For instance, people with type 2 diabetes have a higher risk of developing AD in old age[151], suggesting a link between insulin resistance, IGFs and IGFBPs and AD. To understand these connections in more detail, the mouse is a suitable model organism. For instance, one report has shown an increased level of IGFBP7 in the hippocampus of an AD mouse model[133], these findings were linked to altered DNA-methylation in the promoter region of the IGFBP7 gene. Furthermore, an artificially increase of IGFBP7 in wildtype mice impaired their memory function. In another report, downregulation of IGFBP7 rescued impairment of Insulin Like Peptide (ILP) signaling, which takes part in neuronal activity, memory formation and learning[152], leading to the recovery of memory function. Taken together, IGFBP7 is a potential biomarker and therapeutical target for AD early diagnosis and treatment.

We wish to mimic the IGFBP7 situation observed in human AD patients in mouse and use the Cre-loxP-system to produce inducible IGFBP7 overexpressing mouse line.

Experimental evidence suggests an influence of peripheral IGFs on neuronal plasticity and memory formation in mice[153,129]. Therefore, mouse strain was designed to have increased levels of IGFBP7 in the blood by overexpressing the protein in the liver. The aim is to test our hypotheses, whether altered brain plasticity in response to increased peripheral IGFBP7 levels is linked to cognitive phenotype.

2. Study objectives

Part 1: The role of Setd1b histone methyltransferase in the mouse brain

In the first part of this PhD project we wanted to elucidate the molecular and behavioral consequences of Setd1b histone methyltransferase knock-out in the murine brain. Interfering with methylation machinery in the mouse brain has been already reported for four out of six members of SET1 family of histone methyltransferases (Kmt2a, Kmt2b, Kmt2d and Setd1a), but not for Setd1b methyltransferase. Loss-of-function of histone methyltransferases is strongly associated with neurodevelopmental disorders and intellectual disabilities in humans but their functions and targets in the adult brain is poorly explored. Mechanisms underlying complex behavioral disorders can be revealed by studying interaction of chromatin and expression machinery. Many of mutant mouse strains modelling loss-of-functions exhibit behavioral abnormalities similar to human neurodevelopmental and psychiatric disorders. Therefore, our aims were (1) to generate and validate the model of Setd1b conditional knock-out from the excitatory neurons of the mouse forebrain, (2) to perform profound behavioral characterization of animals upon the enzyme loss, (3) to study transcriptome and chromatin landscape associated changes and identify potential target genes, (4) in case of pronounced behavioral deficits, try to reverse the phenotype using a pharmacological approach.

Part 2: IGFBP7 overexpression as a mouse model of Alzheimer's disease

In the second part of my PhD project I have focused on one of the mechanisms potentially underlying Alzheimer's Disease (AD), one of the most common neurodegenerative diseases worldwide. From the unpublished data it is known that plasma levels of IGFBP7 protein is increased in AD patients. Therefore, we aimed at (1) mimicking the increase of IGFBP7 in plasma by overexpressing the protein in the liver, (2) evaluating whether this will result in an AD-like phenotype in mice. Taken together, we wish to unravel the impact of peripherally overexpressed IGFBP7 on the mouse brain on a cognitive level.

3. Materials and methods

3.1. Reagents

Complete Protease Inhibitor Cocktail	(Roche applied science,)
DirectPCR tail-lysis buffer	(Viagen Biotech)
DPBS	(Invitrogen)
Dynabeads	(Invitrogen)
GeneRuler™ 100 bp	(ThermoFischer)
LPA	(ThermoFischer Scientific)
Midori Green	(Nippon Genetics)
PageRuler™ Plus prestained ladder	(ThermoFischer)
RNAZap	(Sigma-Aldrich Co)
SYBR Green	(Roche applied science)
SureClean Plus	(Bioline)
1x Histoacryl	(Braun)

3.2. Technical equipment

-80°C Freezer	(ThermoFischer Scientific)
-20°C Freezer	(Liebherr)
Benchmark Stereotaxic Instrument	(Myneurolab)
Bioanalyzer High sensitivity chips	(Agilent technologies)
Bioruptor Plus sonication device	(Diagenode SA)
Centrifuge (5418)	(Eppendorf)
Centrifuge 4°C (5418R)	(Eppendorf)
Cryotome	(Leica)
Illumina HiSeq2000	(Illumina)
Illumina NextSeq550	(Illumina)
FACSaria III	(BD Biosciences)
FLUOstar Omega platereader	(BMG Labtech)
Homogenizer IKA T19 Ultra	(Ika)
Microcentrifuge	(Roth)
Microsyringe pump	(World Precision Instruments)
Multi-rotator RS-24	(Biosan)
Nanodrop 2000	(ThermoFischer Scientific)
Nanoliter 2000 Inkector	(World Precision Instruments)
Odyssey Licor CLX	(Li-COR Biosciences)
Peristaltic perfusion pump	(Heidolph)
Pipettes ResearchPlus	(Eppendorf)
Pipette filler S1	(ThermoFischer Scientific)

Refrigerator	(Liebherr)
Roche LightCycler® 480 System	(Roche applied science)
pH meter peqMIX plus	(Peqlab)
Scales Practum	(Sartorius)
Shaker Mini-Rocker	(Peqlab)
Shaker Unimax1010	(Heidolph)
Thermal Cycler T100™	(Bio-Rad Laboratories)
ThermoMixer C	(Eppendorf)
TransBlot® Turbo™ Transfer system	(Bio-Rad Laboratories)
Vortex Genie2	(Electro Scientific Industries)

3.3. Consumable materials

DNA LoBinding tubes 1,5 ml	(Eppendorf)
15 ml; 50 ml tubes	(Greiner bio-one)
Glass capillaries	(World Precision Instruments)
Microlex gloves	(Ansell)
Disposal bags	(Roth)
SafeSeal 1,5 ml, 2 ml microtubes	(Sarstedt)
4-15% mini-PROTEAN® SDS-acrylamide gels	(Bio-Rad Laboratories)
TransBlot®Turbo™ Midi-nitrocellulose transfer packs	(Bio-Rad Laboratories)
Serological pipettes	(Sarstedt)
Pipette filter tips TipOne	(StarLab)
96-Well PCR plate, non-skirted	(StarLab)
96-Well PCR microplate, LightCycler-type, white	(StarLab)

3.4. Chemicals

Agarose	(Roth)
Ammonium persulphate	(Roth)
BSA	(Roth/Serva)
$CaCl_2xH_2O$	(Roth)
DTT	(Roth)
DAPI	(Vectashield)
EDTA	(VWR)
EGTA	(Fermentas)
EZ prep lysis buffer	(Sigma)
Ethanol	(Roth)
Formaldegyde	(Sigma-Aldrich Co)
Glycine	(Applichem)

HEPES	(Roth)
KCl	(VWR)
LiCl	(Applichem)
Methanol	(Roth)
$MgAc_2$	(Applichem)
$MgCl_2$	(VWR)
NaCl	(Applichem)
Nonident P-40	(Fluka)
Parafolmaldehyde	(Roth)
PBS	(Roth, Invitorgen)
PonceauS	(Roth)
Protease inhibitor (EDTA free)	(Roche)
Proteinase K	(Roth)
RNaseIN plus inhibitor	(Promega)
SDS	(Applichem)
Sodium azide	(Applichem)
Sodium deoxychalate	(Roth)
Sucrose	(Applichem)
TEMED	(Roth)
Tris	(Applichem)
Triton-X-100	(Roth)
Tween20	(Roth)
Water treated with DEPC	(Sigma-Aldrich Co)

3.5. Kits and reagents

NucleoSpin® TriPrep kit	(Macherey-Nagel)
Pierce BCA Protein Assay Kit	(Thermo Fischer Scientific)
Transcriptor First Strand cDNA Synthesis Kit	(Roche applied science)
ENZO corticosterone ELISA kit	(Enzo)
Qubit dsDNA High Sensitivity Kit	(Thermo Fischer Scientific)
Qubit RNA High Sensitivity Kit	(Thermo Fischer Scientific)
Zymo RNA clean & concentrator-5 kit	(Zymo Research)
TruSeq RNA Sample Preparation v2 Kit	(Illumina)
Takara SMART-Seq v4 Ultra Low Input RNA Kit	(Takara)
Illumina TruSeq RNA Library Prep Kit	(Illumina)
10X Chromium Single Cell 3' reagents(v3)	(10xGenomics)

3.6. Software

Image Studio Lite 5.2	(Li-COR Biosciences)
GraphPad Prism 7.0	(GraphPad)
LightCycler® 480 Software	(Roche)
The Integrative Genomics Viewer (IGV)	(Broad Institute)
VideoMot2 system	(TSE Systems)
Video Freeze®	(Med Associates)

3.7. Buffers and media

TAE buffer: 40 mM Tris-acetate, 1 mM EDTA

TBS 10X (for 1 L): 24 g Tris-base, 88 g NaCl, pH=7.6

TBS-T (for 1 L): 100 ml 10X TBS with 1 ml Tween 20

Citrate buffer: 10 mM Tris-sodium citrate, 0,1% Triton-X

SDS solution (10%): 10 g SDS in H_2O

Blocking solution: 1x TBS-T, 5% (w/v) BSA

Running buffer 5X (for 5 L): 75.5 g Tris/HCl (pH=8.3), 360 ml Glycine, 25 g SDS

Low Sucrose Buffer: 5mM $CaCl_2$, 5mM $MgAc_2$, 0,1mM EDTA, 10mM HEPES (pH=8), 1mM DTT, 0.1% Triton X-100, PI (1:100), 320 mM sucrose

High Sucrose Buffer: 3mM $MgAc_2$, 0,1mM EDTA, 10mM HEPES (pH=8), 1mM DTT, PI (1:100), 1 M sucrose

PBTB buffer: PBS, 1% BSA, 0.2% Tween-20, PI (1:100)

RIPA buffer: 140 mM NaCl, 1 mM EDTA, 0.1% sodium deoxychelate, Triton X-100, 10mM Tris pH=8, 1% SDS, PI (1:100)

Wienmann Buffer 2X: 20mM EDTA, 100 mM Tris pH=8, 2% SDS

IP Buffer: 140 mM NaCl, 1% NP-40, 0.5% sodium deoxychelate, 50 mM Tris pH=8, 20 mM EDTA, 0,1% SDS, PI (1:100)

Wash Buffer: 100 mM Tris (pH=8), 1% NP-40, 1% sodium deoxychelate, 20 mM EDTA, 500 mM LiCl, PI (1:100)

Nuclear storage buffer: 1x PBS (Invitrogen), 0.5% RNase free BSA, 1:200 RNaseIN plus inhibitor; 1x EDTA-free protease inhibitor

3.8. Animals

3.8.1. Animal welfare

All mice were kept in the individually ventilated cages and were exposed to 12/12 h light/dark cycle. Food and water were provided *ad libitum.* Housing and breeding was done at the certified animal facility in Deutsches Zentrum für Neurodegenerative Erkrankungen by professional animal caretakers. All animal experiments were planned and conducted in full compliance with the German Federal Act on the Protection of Animals and were approved by the responsible ministry of Lower Saxony ("Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit") under protocol G17/2695.

3.8.2. *Setd1b-CamKIICre* transgenic mice

Generation of mutant mice of C57B/6J background with floxed allele for *Setd1b* KO has been described previously[121]. Animals with exon 5 of *Setd1b* gene flanked with loxP sites were kindly provided by the group of Franz Stewart and Andrea Kranz (Biotechnology center, Technische Universität Dresden). They were crossed with mice expressing cre-recombinase under the control of the forebrain-specific Ca^{2+}/CaMKII-dependent protein kinase IIα (CaMKIIα) promoter, generating conditional KO (cKO) line with selective ablation of *Setd1b.* In all experiments animals containing loxP site and not expressing cre-recombinase were used as a control and are referred to as "WT (wild-type)" in the text.

3.8.3. *IGFBP7-AlbCre* transgenic mice

IGFBP7 mice containing mouse IGFBP7 allele in ROSA26 locus all over the body were acquired from the company "Taconic" and have C57BL/6NTac background. For the experiments animals were continuously crossed with AlbCre animals, expressing cre-recombinase under liver-specific alb promoter (mice were provided by Dr. Elisabeth

Hessmann (Gastroenterologie, Universitats Medizin Göttingen)), which would cut out STOP-codon with loxP sites located in front of conditional IGFBP7 allele and leading to the overexpression. In the text, WTA is designated to mice containing only cre-recombinase expressed under liver promoter, WTB to mice containing non-activated extra IGFBP7 allele and KI to mice with both, cre-recombinase liver expression and, therefore, activated extra IGFBP7 allele.

3.8.4. Genotyping

Genomic DNA was isolated from mice tails with 200 µl of DirectPCR® Lysis Reagent (Viagen, USA) and 40 µg of proteinase K (Roth, Germany). Biopsies were lysed for a minimum of 3 hours at 55°C at 400 rpm followed by 45 min incubation at 85°C in Thermomixer Comfort (Eppendorf, Germany).

PCR master mix composed is presented in the Tabl. 1. Primers were supplemented according to mouse line: Setd1b, IGFBP7, IGFBP7-albCre, albCre and provided in Tabl. 2 together with expected fragment size. Cycling conditions are given in Tabl.3 and Tabl.4. All primers were ordered from Sigma (Germany) and diluted to 10 µM.

Samples were run in 1,5% agarose gel with 0,06 µl/ml Midori Green dye (Takara CloneTech Japan) for 1 hour at 120V in 1xTAE buffer. As molecular weight standard ThermoFischer GeneRuler™ 100 bp was used. Fragments were visualized with FAS V Gel documentation System (Nippon genetics, Japan).

Component	Amount per reaction (µl)
Genomic DNA soluction	1
DreamTaq Buffer 10x	2.5
dNTP Mix (10 mM)	2
Setd1b_F/ Igfbp7_oligo3	0.25
Setd1b_R/ Igfbp7_oligo4	0.25
CamK_Cre_F	0.13 µl
CamK Cre_R	0.13
DreamTaq polymerase	0.2
ddH$_2$O	18.54

Table 1: Reaction mixes for genotyping PCRs.

Name	Sequence	Fragment size
Setd1b_F	gaaactcgcatgcgcttctac	696 bp Setd1b loxP allele
Setd1b_R	agttcatactgtggctgaatgg	507 bp WT allele
CamK2a_Cre_F	tctcacgtactgacggtgg	~400 bp for cre+
CamK2a Cre_R	accagcttgcatgatctcc	
Igfbp7_oligo3	tggcaggcttgagatctgg	~500 KI allele
Igfbp7_oligo4	cccaaggcacacaaaaaacc	
DNA_1	ggggcaatcaattgaggg	333 bp for the presence of DNA in the sample
DNA_2	caacctctgcttggttctgg	

Table 2: Primer sequences for genotyping PCRs.

Step	Temperature	Duration	Number of cycles
Denaturation	95°C	5 min	1
Amplification	95°C	30 sec	X35
	58°C	30 sec	
	68°C	1 min	
Elongation	68 °C	7 min	1
Cooling	4 °C	∞	1

Table 3: Amplification protocol for genotyping Setd1b, AlbCre lines

Step	Temperature	Duration	Number of cycles
Denaturation	95°C	5 min	1
Amplification	95°C	30 sec	X35
	60°C	30 sec	
	72°C	1 min	
Elongation	72 °C	10 min	1
Cooling	4 °C	∞	1

Table 4: Amplification protocol for genotyping IGFBP7, IGFBP7-albCre lines

3.8.5. Behavioral studies

Group housed *Setd1b* transgenic and their wild-type littermates animals at the age 8 weeks onwards were housed in the individual cages for at least 1 week prior to testing in the animal room. Mice of both genders were subjected to behavioral tests. Tracking experiments were performed using VideoMot2 system (TSE Systems, Bad Homburg, Germany). If plastic surfaces were used, they were cleaned with 70% EtOH to remove olfactory cues after every tested animal. Tests were performed in following order: open field, novel object recognition, elevated plus maze, contextual fear conditioning, nesting, Morris water maze, rotarod, Y-maze, social novelty test, tail suspension test, prepulse inhibition.

3.8.5.1. *Open field*

For general locomotor ability, basal anxiety and exploratory behavior animals were subjected to open field test[154]. Animals were individually placed in a square uniform gray plastic arena (50 x 50 x 40 cm) and allowed to explore the area for 5 min. Virtually arena was divided into 16 quadrants and time spent in the 4 central quadrants was measured. Total distance travelled and average speed were quantified.

3.8.5.2. *Novel object recognition*

For the next consecutive days after the open field test, the mice were habituated with two identical objects for 5 min each time (training day). 24 h after, one of the objects was replaced by new non-identical object and mice were left to explore it for 5 min (test day). The relative exploration time of the novel object was measured.

$$\% \, Time \, near \, novel \, object = \frac{Time \, spent \, near \, novel \, object}{Time \, spent \, near \, both \, objects} * 100\%$$

3.8.5.3. *Elevated plus maze*

Mice were placed into the center of plus-shaped maze with two open and two closed arms (10 x 40 cm each, walls 40 cm high) for 5 min facing the open arms. Time spent in the arms of both types was measured.

3.8.5.4. *Fear conditioning*

Associative learning and memory can be accessed with classic Pavlovian fear conditioning test. Experiment was controlled and recorded via NIR Video Fear Conditioning System (Med Associates, USA). Test was carried out in plexiglass training chamber located inside plastic box. Mouse was placed on a removable metal electrifiable grid floor in a sound-proof chamber with white noise generator and 12 Watt light. On the first day (training day) mice were exploring the chamber for 3 min and after they received an electric foot shock for 2 sec (constant current, 0.5 mA) while having a white noise on the background. 24 hours later (test day) mice were subjected to the same conditions for 3 min without experiencing foot-shock. The percentage freezing during training and test days were tracked with VideoFreeze® software provided with a system (Med Associates, USA). Freezing was considered if motion index was lower than threshold of 50 for at least 30 frames per second. % Freezing was calculated by the number of measurements when mouse showed freezing behavior divided to a total number of measurements.

3.8.5.5. *Y-maze*

Working memory was accessed with a Y-shaped plastic runway (10 x 40 cm arms, walls 40 cm high). Mouse was placed into the Y-maze in the triangle-shaped central platform and left to freely explore the maze for 5 min. % of spontaneous alterations (choice of "novel" arm: when animal goes into different arm then before) was scored and analyzed.

3.8.5.6. *Social novelty test*

Sociability and social memory were tested in the same arena as used for open field test. Two mice of the same age and gender as a tested mouse were placed inside metal grid cups, so that tested mouse could smell both of them through the grid (training day). On the test day, one mouse in the cup was replaced with a new one. Time spent near both mice was analyzed in order to evaluate the preference for a new mouse.

$$\% \, Time \, near \, new \, mouse = \frac{Time \, spent \, near \, new \, mouse}{Time \, spent \, near \, both \, mice} * 100\%$$

3.8.5.7. *Morris water maze*

Mouse was placed into a swimming basin (1.2 m diameter) with opaque water of room temperature containing a submerged platform and four visual cues (star, triangle, square and circle, see Fig.3A). On each training day, mice were swimming for 4 times, each time starting from a different visual cue, for 60 sec. Escape latency of mice finding the platform was measured. If platform was not found during 60 sec set, mouse was guided to the platform gently. On the day of probe test platform was removed and mice were allowed to swim for 1 min. The percentage of the time spent in the target quadrant and in the platform area as well as number of crossings of platform area were measured. Strategy analysis was performed based on modified MUST-C algorithm[155]. Direct, corrected, focused and short chaining strategies represent higher cognitive performance, while circling, long chaining, scanning, random, thigmotaxis and passivity with moderate or poor hippocampal spatial memory.

3.8.5.8. *Nest-building behavior*

Impaired nest building has been associated with autistic spectrum disorder (ASD) phenotype in mice[155]. Animals were kept in the transparent plastic cages with autoclaved dust-free sawdust bedding and one cotton fiber nestlet (Ancare, USA) as nesting material, provided 3-5 hours prior to the onset of the dark phase of the lighting cycles. Nest scores were accessed after two nights, since females from our experience sometimes fail to tear the material over one night. Scoring was done according to previously described methodology)[156]: "1" nestlet is 90% intact; "2" nestlet is partially torn (50-90% remain intact); "3" Shredded nestlet but nestside is not identifiable; "4" Identifiable but flat nests: more than 90% of the nestlet is torn; "5" A (near) perfect nest: more than 90% of the nestlet is torn and the nest is a crater. After nest-building test, nestlets were replaced with the fresh paper towels.

3.8.5.9. *Rotarod*

Rotarod is a behavioral test aimed at accessing motor functions, balance and coordination in rodents which can indicate neurological deficits[157]. Mice were placed on individual chambers of rotating drum and habituated for 4 trials when bar rotated for 180s at 5 rpm (training phase). If mouse was falling down, it was returned at the rod for the time remaining. During the test phase, mice were subjected for two trials in the same daily schedule with intervals 4 hours between the trials for 2 consecutive days (4 trials in total). During the trial that lasted 180 s rod was accelerating from 5 to 40 rpm. TSE RotaRod Advanced system was used for the measurements of latency to fall.

3.8.5.10. *Prepulse Inhibition*

Sensorimotor gating was accessed with the prepulse inhibition of startle response (PPI), which is very often impaired in schizophrenia mouse models[158]. Loudspeakers delivering acoustic stimuli were located above the small metal grid cage sized to restrict mouse major

movements. Cage was standing on a sensitive transducer platform sensoring vertical movements of the floor. Whole system was placed inside plastic boxes with sound isolation (TSE Systems). On the test session, first 3 min mouse was habituated to the 65 dB background noise followed by 2 min baseline recording. Afterwards, mouse was exposed to 6 pulse-alone trials with 120 dB startle stimuli for 40 ms were applied in order to decrease within session habituation. For test of PPI the startle pulse was applied by itself or after a prepulse stimulus (70 dB, 75 dB, 80 dB, 85 dB, 90 dB intensity). An interval of 100 ms with background noise was applied between prepulse and pulse stimulus. The trials were presented in pseudorandom order with 8-22 s interval.

The average of the maximum force (MaxG) was measured as reaction to the 120 dB startle stimulus and comprised startle response amplitude.

$$\%PPI = 100 - \frac{Startle\ amplitude\ after\ prepulse}{Startle\ amplitude\ after\ pulse\ alone} * 100$$

3.8.5.11. *Tail suspension*

Depressive behavior in rodents can be measured by tail suspension test[159]. Mouse tail was manually attached with a tape to a wooden rod. Light plastic tube was placed on the tail to avoid climbing attempts. Mouse behavior was recorded for 5 min on a video (iPhone7, Apple) and time of immobility was subsequently analyzed manually from the video record with a stopwatch.

3.8.6. SAHA oral administration

Suberoylanilide hydroxamic acid (SAHA), also known as Vorinostat, is a small-molecule histone deacetylase inhibitor (HDAC) can cross blood-brain barrier and clinically approved for cancer therapy under the name Zolinza®. SAHA was ordered Cayman Chemical Co in the amount of 0.67 g was dissolved in 1 l of drinking water according to the publication[80] together with 18 g of β-cyclodextrin (Sigma-Aldrich) that increased the solubility of SAHA. Vehicle group received β-cyclodextrin solution only. In order to

dissolve SAHA was stirred at 95°C. Dosage was calculated as equivalent of 375 mg/day for humans, which is within the tolerable zone. Fresh solution was provided every week to the animals 4 weeks prior to the onset of the behavioral testing and during the testing.

3.8.7. Brain samples preparation for molecular analysis

Animals were sacrificed by cervical dislocation. Whole brain was quickly isolated on ice. Dissection was undergone in 1X PBS supplemented with EDTA-free protease inhibitor cocktail (Roche). Regions of interest (DG, CA, cortex, cerebellum) were isolated and snap frozen in the liquid nitrogen. Tissue samples were stored at -80°C until processing for RNA and protein extraction, RNA-seq and ChIP-seq. Tissues of both genders were mixed for the analysis.

3.8.8. Intracardial perfusion

Setd1b mutant and control mice were dispatched using ketamine/xylazine (Ketamin 20 µl, Xylarem 150 µl, NaCl 0,9% 30 µl, 100 µl per 10 g body weight). If animal did not response to toe-pinch reflex, transcardial perfusion with 4% PFA in PBS using peristaltic pump (Heidolph, Pumpdrive 5201) followed. Brains were extracted and incubated overnight at 4% PFA at 4°C, then switched to 10%, 20% and 30% sucrose solution subsequently till soaking.

3.8.9. Immunohistochemistry

Coronal brain sections (30 µm) were obtained with Leica cryostat at -20° C after embedding whole brain into Jung freezing medium (Leica, Germany) with the posterior plane facing down. Free-floating cryosections were washed 3 times for 5 min with PBS, permeabilized with 0,3% Triton X-100 in PBS for 10 min, followed by 1 h blocking in 5% goat serum with 0,3% Triton-X and then incubated with primary antibody overnight at 4°C. List of antibodies used is presented in the Tabl. 5. Next day after washing with PBS,

sections were incubated with corresponding ThermoFischer secondary antibodies: Alexa-488, Alexa-555 and Cy-3 in the dilution 1:1000, as well as DAPI (1:10000 in PBS), and mounted in Fluoromount-G (Invitrogen) mounting medium. Images were taken on Nikon Eclipse Ti microscope. ImageJ was used to crop images, merge channels and perform quantifications.

Immunostaining of Setd1b (Bethyl) was done with prior antigen retrieval with citrate buffer (10 mM Tris-sodium citrate, 0,1% Triton-X in water) at 95% for 10 min with cryosections located on SuperFrost® Plus (Thermofischer) before washing and blocking. Of note, we have tested several Setd1b antibodies that were not specific to detect reduction of Setd1b in the cryosections of cKO mice and therefore considered to be unspecific.

Target protein	Dilution for IHC	Source	Product number
Anti-Setd1b	1:100	MBS	MBS126717 Did not work
Anti-Setd1b	1:200	ThermoFischer	PA5-66255 Did not work
Anti-Setd1b	1:500	Bethyl	A302-281A Did not work
anti-NeuN	1:100	Merck	MAB377
anti-MAP2	1:1000	Biosensis	C-1382-50
anti-Iba-1	1:500	Wako	019-19741
anti-GFAP	1:1000	Abcam	ab4674

Table 5: Antibody list for immunohistochemistry.

3.8.10. Intracranial AAV-virus injections

Viral particles for the Cre expression under control of the hSyn promoter were obtained from Penn Vector Core (AAV2/1.hsynapsin.hGHitron.GFP-Cre.WPRE.SV40, University of Pennsylvania). For Cre-AAV, stock was diluted 1:100, what resulted in 1.76×10^8 genome copies per µl. Animals were anaesthetized with xylazine/ketamine mixture (10 µl per 1 g mouse weight), Analgesic (50 µl of buprex) was preadministered and an incision was made to expose the skull. Eyes were covered with bepanthen creme. Skull was drilled with using a surgical driller (Fordom). Injections were done in both hemispheres using digital stereotaxic manipulator (Leica Microsystems), coordinates for the injection (relative to bregma): anteroposterior -1.75 mm, lateral ±1 mm, dorsoventral 2 mm to target DG and anteroposterior -1.75 mm, lateral ±1 mm, dorsoventral 1.3 mm, to target the dorsal CA1 region. Glass capillaries delivered the virus with a speed 350 nl/min, resulting in 900 µl of virus injected in each hemisphere. After surgery wound was glued with tissue glue (Histoacryl®, Braun) and animals were left on the heating pad till awakening.

3.8.11.Basal corticosterone measurement

To measure basal corticosterone level, cKO and WT males (n/genotype=5) were anesthetized with CO_2 and once the animals did not respond to the toe pinch anymore, blood was collected via cardiac puncture. Corticosterone level was measured from isolated plasma with ENZO ELISA kit according to manufacturer's protocol.

2.9. Molecular Biology

2.9.1. Nucleic acids isolation

DNA and RNA isolations were performed according to NucleoSpin® TriPrep kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instruction. Shortly, brain samples were homogenized with a mortar in 350 µl of lysis buffer containing 14 µl of TCEP reducing agent (with a final concentration 500 mM). Then lysate was filtered through centrifugation for 1 min at 11,000 g, mixed with 350 µl of EtOH (70%) and centrifuged again for 30 s at 11,000g on the column. Flow-through was used for further protein extraction. After, 500 µl of DNA wash buffer were added on a column and centrifuged for 1 min at 11,000g, this procedure was performed twice. Afterwards flow-through was discarded and membrane was dehydrated for 3 min. Next step 100 µl of DNA Elute buffer were applied onto the membrane, incubated for 1 min and centrifuged for 1 min at 11,000g. Eluted liquid contained DNA, it was stored at +4°C. To digest residual DNA on a column, 95 µl of rDNase reaction mixture was applied on a membrane and incubated at RT for 15 min. Then membrane with RNA was washed with 200 µl of RA2 buffer for 30 s at 11,000g and 600 µl of RA3 buffer for 30 s at 11,000g. Last wash was with 250 µl of RA3 buffer with 2 min centrifugation at 11,000g. RNA was eluted in 60 µl of RNase-free H_2O at 11,000g for 1 min. The RNA concentration was measured with NanoDrop Spectrophotometer at the wavelength 280 nm. RNA was stored at -80°C. RNA extracted this was also used for whole-tissue of RNA-seq. RNA degradation was analyzed

at Bioanalyzer 2100 with NanoReagents RNA kit (Agilent). Only samples with RNA integrity number(RIN) values higher than 8.0 were used for whole tissue RNA-seq.

2.9.2. Protein extraction

700 µl of flow-through containing protein obtained with NucleoSpin® TriPrep kit was mixed with 700 µl of Protein Precipitator buffer, incubated at RT for 10 min and centrifuged for 5 min at 11,000 g. Supernatant was removed, 500 µl of 50% EtOH was added to the residual pellet and the mixture was centrifuged for 1 min at 11,000g. After removing supernatant, precipitate was dried for 10 min at RT with an open lid. Then, based on the size of the pellet, it was disaggregated in 20-100 µl of PSB-TCEP (Protein Solving Buffer) and incubated for 5 min at 95°C for complete protein dissolving and denaturation. After samples were cooled down, tubes were centrifuged for 1 min at 11,000g to pellet residual insolvable material. Recovered supernatant was utilized for protein quantification and western-blot. Proteins were stored at -80°C.

2.9.3. Western Blotting

Protein isolation was performed with NucleoSpin® TriPrep kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instruction. Protein quantification was done with Pierce BCA Protein Assay Kit according to manufacturer's instruction. 30 µg of protein were loaded on 4-15% mini-PROTEAN® SDS-acrylamide gels (Bio-Rad, USA). Transfer was performed with TransBlot® Turbo™ Transfer system using TransBlot® Turbo™ Midi-nitrocellulose transfer packs (Bio-Rad). Afterwards membrane was blocked with 5% BSA and incubated at 4°C overnight with target antibodies, see Tabl. 6. The next day membrane was incubated with corresponding secondary antibody (IRDye 800, Li-COR, 1:5000), followed by visualization and quantification of immunoblots with Odyssey Imager Li-COR. Values in *Setd1b* transgenic mice were normalized to values in

control littermates. Of note, we have tested several Setd1b antibodies that detected a band in the wild-type and *Setd1b* cKO mice and therefore considered to be unspecific.

Target protein	Dilution for WB	Source	Product number
anti-Setd1b	1:200	ThermoFischer	PA5-40072 Did not work
anti-hSET1B-C-term	1:100	Abcam	Ab170645 Did not work
anti-Setd1b	1:100	MBS	MBS126717 Did not work
anti-Setd1b	1:500	Bethyl	A302-281A
anti-β-actin	1:1000	Santa-Cruz	sc-69879
anti-H3K4me3	1:1000	Abcam	ab8580
anti-HSC70	1:1000	Abcam	19136
anti-IGFBP7	1:1000	R&D Systems	MAB2120 Did not work
anti-IGFBP7	1:300	ProteinTech	19961-1-AP Did not work

Table 6: Antibodies used for western blot.

2.9.4. cDNA synthesis

Reverse transcription was done with Transcriptor First Strand cDNA Synthesis Roche Kit (Roche Applied Science, Germany). 1 μg of RNA was mixed with 2μl of random hexamer primers and filled up to 13 μl with PCR-grade water. The mixture was incubated for 10 min at 65°C in a Bio-Rad T100 Thermal cycler (Bio-Rad Laboratories, USA) and cooled on the ice. After, each reaction was supplemented with 4 μl of Transcriptor Reverse Transcriptase Reaction buffer, 0.5 μl of RNase Inhibitor, 2 μl of deoxynucleotide mix and 0.5 μl of Transcriptor Reverse Transcriptase (total volume of the mixture was 20 μl). Tubes were incubated for 10 min at 25°C, 30 min at 55°C and 5 min at 85°C and then placed on ice. For further qRT-PCR, 15 μl of cDNA was diluted 10-times with nuclease free water, the rest 5 μl from all the samples were used to create the primers standard

curve with for serial dilutions of 1:1, 1:2, 1:4, 1:8, 1:16, 1:32 in the qRT-PCR machine. 1:32 dilution was used as a positive calibrator. cDNA was stored at -20°C.

2.9.5. qRT-PCR

RNA isolation was performed with NucleoSpin® TriPrep kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instruction. 1 µg of total RNA was used for reverse transcription with Transcriptor First Strand cDNA Synthesis Roche Kit (Roche, Germany). Quantitative real-time PCR (qRT-PCR) with SYBR Green PCR Master Mix (Roche) was performed in a Roche 480 II Light Cycler (Roche). Primers were designed with online tool (https://lifescience.roche.com/en_de/brands/universal-probe-library.html). Data were normalized by *Hprt* housekeeping gene. For graphical presentation and statistical analysis, the mRNA values in Setd1b transgenic mice were normalized to the values in control littermates. Primers used for Setd1b deletion evaluation, *Hprt* housekeeping gene, IGFBP7 knock-in evaluation primers and expression of cre-recombinase and GFP are shown in the Tabl. 7. Primers used for qPCR verification of whole tissue RNA-seq are shown in the Supp.Tabl.17. Amplification protocol is described in Tabl.8 and reaction composition in the Tabl.9.

Name	Sequence
Setd1b_exon5	ctgttggtgagctggatgcta
Setd1b_exon6	ctggagtaagctgtgtcttgg
Hprt Mm right	cctggttcatcatcgctaatc
Hprt Mm left	tcctcctcagaccgctttt
eGFP_F	ctgctgcccgacaaccac
eGFP_R	tgtgatcgcgcttctcgtt
Cre_recombinase_F	ggcgcgagttgatagctg
Cre_recombinase_R	gttttgccgggtcagaaaa
Igfbp7_trvar1_R	tgccctccatgaaataccac
Igfbp7_trvar1_L	ggctgtctgagagcaccttt
Igfbp7_trvar2_R	tgccctccatgaaataccac
Igfbp7_trvar2_L	ggctgtctgagagcaccttt

Table 7: qRT-PCR primers.

Step	Temperature	Duration	Number of cycles
Denaturation	95°C	5 min	1
Amplification	95°C	10 sec	X45
	60°C	30 sec	
	72°C	10 sec	
Cooling	40 °C	30 sec	1

Table 8: Amplification protocol for qRT-PCR.

Component	Amount per reaction (µl)
H_2O	3
Forward primer (20 mM)	0.75
Reverse primer 20 mM)	0.75
SYBR Green Mix	7.5
UPL probe	0.15
cDNA	3

Table 9: Reaction mixes for qRT-PCRs with SYBR Green.

2.9.6. Statistical analysis

Statistical analysis for behavioral and molecular experiments was performed with GraphPad Prism software. Bars represent mean±S.E.M and significance of group differences was evaluated by t-test for independent samples or, when indicated, by one-way analysis of variance (ANOVA) with repeated measures. p-values < 0.05 were considered significant.

2.10. Sequencing and bioinformatics analysis

2.10.1. RNA sequencing of the whole tissue

RNA was isolated as described in the section 2.9.1. cDNA library was prepared with TruSeq RNA Sample Preparation v2 Kit (Illumina) and single-end 50 bp sequencing was done using HiSeq2000 (Illumina). Quantity and quality of library were checked with Qubit dsRNA HS Assay kit and Agilent 2100 Bioanalyzer subsequently.

2.10.2. Cell-type specific RNA isolation and RNA sequencing

Each sample comprised of two pulled frozen CA regions from left and right hemisphere of two different mice. Tissues were homogenized on ice with 30 strokes plastic pestle in 1,5 ml Eppendorf tube containing 500 µl of EZ prep lysis buffer (Sigma). This was followed by homogenate transfer into 2 ml tubes, adding up lysis buffer to 2 ml and incubating on ice for 7 min. Samples were centrifuged at 500g, supernatant discarded and the nuclear pellet was mixed with 2 ml of lysis buffer and incubated on ice for 7 min again. Centrifugation step was repeated and nuclear pellet was mixed with 500 µl of nuclei storage buffer and filtered with 40 µm filter (BD falcon) with extra 100 µl of nuclei storage buffer. Collect nuclei were incubated for 45 min with anti-NeuN-Alexa488 antibody (1:1000) and washed with nuclei storage buffer. Next we performed FACS sorting of the stained nuclei with 85 µm nozzle on FACSaria III machine. Gating was done by the size of the nuclei, doublets were excluded. Separation of neuronal nuclei from non-neuronal was based on NeuN-Alexa488 fluorescent signal. Sorted nuclei were collected into 15 ml falcon covered with nuclei storage buffer. RNA was extracted using Trizol LS protocol: After incubating nuclei with 750 µl of Trizol for 5 min, 200 µl of chloroform was added and incubated for 5 min. This was followed by centrifugation at 12000g at 4°C for 15 min. Aqueous phase with RNA was collected and processed with Zymo RNA clean & concentrator-5 kit with DNAse treatment according to manufacturer's protocol. RNA concentration was measured with Qubit RNA HS Assay. 100 ng of RNA was used for

RNA-seq with Illumina TruSeq RNA Library Prep Kit and sequenced single-end 75 bp in Illumina Nextseq 550. Glial nuclei were prepared for sequencing with Takara SMART-Seq v4 Ultra Low Input RNA Kit and Illumina Nextera XT.

2.10.3 List of bioinformatics tools used for RNA-seq analysis

Tool	Purpose	Notes
Customized Illumina pipeline	Base calling	
Illumina bcl2fastq 1.8.4	Bcl to FASTQ conversion	
FastQC 0.11.5	FASTQ quality control	
STAR aligner 2.3.0	Reads mapping to mouse reference genome (mm10)	
FeatureCount package of Bioconductor	Reads counting	
DESeq2 package of Bioconductor	Differential gene expression	
EdgeR package of Bioconductor	RPKM values calculation	
topGO package of Bioconductor	Gene Ontology (GO) enrichment with weighted analysis option	categories being significantly enriched with $p<0.05$ and false discovery rate (FDR)<0.05
Eulerr package of Bioconductor	Gene sets overlaps	
DeepTools 2.4.0	Creation of bigwig file for genome browser visualization	

Table 10: Tools used for RNA-seq analysis.

2.10.4. Cell-type specific chromatin isolation and ChIP sequencing

Neuron-specific chromatin isolation and ChIP were done as described previously with modifications[160]. Each sample comprised two pulled frozen CA regions from left and right hemisphere of two different mice (we utilized the remaining hemispheres from the tissues

used for NeuN+ RNA-seq sorting). All steps were performed on ice unless stated differently. Tissues were homogenized with a plastic pestle in 1.5 ml Eppendorf with 500 µl of low sucrose buffer and fixed with 1% formaldehyde for 10 min followed by 5 min quenching with 125 mM glycine on rotating wheel. Suspension was centrifuged at 2000g for 3 min, the supernatant was discarded and the pellet resuspended in 1 ml of low sucrose buffer. The pellet was homogenized with a mechanical homogenizer (IKA Ultraturax T10). Oak Ridge centrifuge tubes were filled with high sucrose buffer and homogenized pellet was layered onto it, followed by centrifugation at 3148g for 10 min. Upper phase containing myelin was discarded and the nuclear pellet was resuspended in PBTB buffer. The resulting solution was stained with 1:1000 dilution Alexa Fluor 488-conjugated anti-NeuN antibody in PBTB on a rotating wheel for 25 min, then centrifuged for 3 min at 2000g and washed 3 times with PBTB. Stained nuclei were passed through 40 µm filters into FACS tubes and sorted with 85 µm nozzle on FACSaria III machine with the same gating parameters as described in section 2.3.2. After sorting samples were centrifuged for 15 min at 3148g. Supernatant was discarded, and remaining solution centrifuged for 2 min at 10000g in DNA low binding tube. Nuclear pellet was snap-frozen in liquid nitrogen and stored at -80°C. For chromatin shearing the pellet was mixed with 100 µl of RIPA buffer and incubated on the wheel for 10 min at 4°C. Then the samples were sonicated in Bioruptor (Diagenode, Belgium) 30s ON/30s OFF, high power, for 25 cycles. Next, step samples were centrifuged at 18.000g for 5 min and the supernatant was transferred in the low-binding tube. Shearing check was done by decrosslinking 5 µl of sample with 20 µl of EB buffer and 0,1 µg/µl of RNAse A. Mixture was incubated for 30 min at 37°C, shaking at 450 rpm. After 1 µl of Proteinase K and 20 µl of WB(2x) buffer was added followed by incubation for 2h at 65°C, shaking at 1300 rpm. Samples were mixed with 3 µl of LPA and 46 µl of SureClean, incubated for 10 min at RT and centrifuged for 20 min at 20,000g at RT. Pellet was washed twice with 400 µl of 80% EtOH by centrifugation for 5 min at 20,000g, dried for 5 min and re-suspended in 20 µl of EB. Measurement of chromatin

concentration was analyzed with Qubit 2.0 fluorometer (DNA HS kit) and fragmentation was checked with Bioanalyzer2100 (DNA HS kit). Beads preparation was done as following: necessary amount of protein A or IgG coated beads (Invitrogen) were washed once with 1 ml of IP buffer, re-suspended in 0.5 ml of IP buffer with 0.5% BSA and incubated for 1-2 hours at 4°C on the rotating wheel. Then beads were washed with 1 ml of IP buffer and re-suspended in the initially taken volume. Original samples were diluted 10 times to reduce SDS concentration throughout the precipitation and precleared with 20 µl of activated beads by 1h incubation at 4°C on a rotating wheel. Then tubes were placed on a magnetic rack, supernatant was collected, and the necessary amount of antibody was added and incubated overnight at 4°C on the rotating wheel. All antibodies used are mentioned in the Tabl. 11. Next day each sample was incubated for 1.5h with 15 µl of the activated beads at 4°C on the rotating wheel. After placing tubes at magnetic rack and discarding supernatant, beads were washed twice with 1 ml of IP buffer, three times with 1 ml of wash buffer (last time with 5 min rotation at 4°C), and the last time with 1 ml of IP buffer. To isolate precipitated DNA, samples were incubated with 20 µl of EB with 0.1 µg/µl of RNAse A for 30 min at 37°C shaking at 450 rpm. Then 20 µl of WB buffer and 1 µl of Proteinase K were added to samples and mixture was incubated overnight at 65°C at 800 rpm. Next morning tubes were placed at magnetic rack, supernatant was collected, beads washed once more with 20 µl of EB and incubated for 10 min at 65°C shaking at 800 rpm and supernatant was added to already collected one. After, 3 µl of LPA were added to samples and inputs, vortexed. Then 60 µl of SureClean were added to samples and 90 µl to inputs, vortexed and incubated for 10 min at RT, centrifuged for 20 min at 20000g at RT. Supernatant was removed and washed twice with 80% EtOH. Pellet was dried for 5 min at RT and re-suspended in 30 µl of EB. ChIP DNA was used for library preparation using NEB Next Ultra II DNA library preparation kit and sequenced for single end 50bp at Illumina HiSeq 2000.

Target protein	ChIP dilution	Source	Product number
Anti-NeuN-Alexa488	1:1000 (FACS)	Merck	Mab377x
Anti-H3K4me3	1 µg for 0.3µg chromatin	Abcam	ab8580
Anti-H3K4me1	1 µg for 0.3µg chromatin	Abcam	ab8895
Anti-H3K27ac	1 µg for 0.3µg chromatin	Abcam	ab4729
Anti-H3K9ac	1 µg for 0.3µg chromatin	Millipore	07-352

Table 11: Antibody list for ChIP

2.10.5. List of bioinformatics tools used for ChIP-seq analysis

Tool	Purpose	Notes
Customized Illumina pipeline	Base calling	
Illumina bcl2fastq 1.8.4	Bcl to FASTQ conversion (demultiplexing)	
FastQC 0.11.5	FASTQ quality control	
Bowtie2 2.0.2	Mapping FASTQ to mouse reference genome (mm10)	
SAMTOOLS 1.9.0	- PCR duplicates removal (*rmdup -s*) - BAM files generation - Merging samples of the same group (*merge*)	
NGSPlot 2.61.0	Profile plots from merged BAM files	
MACS2/2.1.1 Bedtools 2.25.0	Peak calling	q value < 0.1
Diffbind package of Bioconductor	Differential binding analysis	
TopGO package of Bioconductor	Pathway analysis	categories being significantly enriched with p<0.05 and false discovery rate (FDR)<0.05
HOMER 4.10	Genomic region annotation	

DeepTools 2.4.0	Creation of bigwig files for genome browser visualization	

Table 12: Tools used for ChIP-seq analysis.

2.10.6. Single-nucleus RNA-Seq

WT mice CA tissue was used to isolate the unfixed NeuN+ nuclei same was as it was described in the section 2.3.2. Sorted nuclei counting was performed in Neubauer chamber with 10% trypan blue in PBS and resuspended to concentration of 1000 nuclei/µl. Library was prepared with 10X Chromium Single Cell 3' reagents(v3). According to manufacturer's protocol barcoding, GEM formation, reverse transcription, cDNA synthesis and library preparation were performed, followed by sequencing in Illumina NextSeq550.

Tool	Purpose	Notes
CellRanger software (v.3.0.2)	- Reads alignment to hg38 genome (GRCm38.p4) - mapping reads to the pre-mRNA	Dataset of 3841 cells (3801 after excluding low quality cells)
SCANPY package	- Prefiltering - Normalization - Clustering	DG and inhibitory neurons clusters were excluded

Table 13: Tools used for single nucleus RNA-seq analysis.

4. Results

Note: some of the figure legends may be identical with a corresponding paper ("SETD1B controls cognitive function via cell type specific regulation of neuronal identity genes", Michurina et al) that has been submitted for publication and is already available via BioRxiv preprint server (https://doi.org/10.1101/2020.08.07.240853).

4.1. The Setd1b – CaMKIICre conditional knockout mouse

Our lab has previously established the essential functions of Kmt2a and Kmt2b in learning and memory formation in the murine hippocampus[77,78]. We therefore hypothesized, that the genetic manipulation of Setd1b protein might also impact mouse cognitive functions. To evaluate this idea, we have used a previously published transgenic mouse model in which Setd1b is brought under the control of the Cre-recombinase by flanking Setd1b exon 5 with loxP sites[121,122], and crossed it with another mouse line in which the Cre is induced by the forebrain specific CaMKII promoter[123] (see Fig.1A). The resulting transgenic mouse is designed to have lost Setd1b protein in excitatory neurons of the 3 weeks old postnatal forebrain (cKO mice). To evaluate the efficacy of region-specific Cre mediated Setd1b knockout, we have conducted a qPCR analysis targeting exon 5 of Setd1b, and indeed, we could observe a strong reduction of Setd1b in CA, DG and in the cortex, but not in the cerebellum (see Fig.1B). These results imply a CaMKII promotor region specific Cre-recombinase activity and loss of Setd1b. As the CaMKII promoter is specifically active in excitatory neurons, residual Setd1b expression can be explained by other cell types being unaffected by the knockout. Additionally, we have conducted whole tissue RNA-seq with CA samples, which showed a specific loss of Setd1b exon 5 in cKO (see Fig.1C), further confirming the functionality of the genetic mouse model. To evaluate the impact of Setd1b knockout on the protein level, we have conducted Western Blot analysis with cortical samples, and indeed we could observe a significant reduction of Setd1b protein in cKO hippocampal CA region brain samples (see Fig.1D).

Morphological assessment did not show any differences between wildtype and Setd1b knockout animals, neither in males, nor in females, although Setd1b cKO mice are slightly lighter in bodyweight (see Fig.2A). Similarly, the brains of wildtype and Setd1b cKO animals did not exhibit any obvious morphological differences, but Setd1B cKO brains were slightly lighter in weight (see Fig.2B). Immunohistochemical analysis of mouse brain tissue sections staining for the neuronal nuclei specific Neuronal N protein (NeuN), Microtubule-83 Associated Protein 2 (MAP2) as well as the microglial marker Ionized calcium-Binding Adapter molecule 1 (IBA1) and the astrocytic marker Glial Fibrillary Acidic Protein (GFAP) did not show any obvious differences between wildtype and Setd1b cKO animals (see Fig.2C and 2D). Interestingly, there was a trend of increased GFAP signal, which might be originating from astrogliosis, however, the difference was not significant. Taken together, the conditional knockout mouse model for Setd1b is functional both on mRNA and on protein level in CaMKII promoter specific brain regions, without inducing any gross anatomical brain aberrations. This verification enabled us to move on to behavioral experiments to evaluate the cognitive functions o Setd1b cKO mice in learning and memory formation.

Fig.1A

B
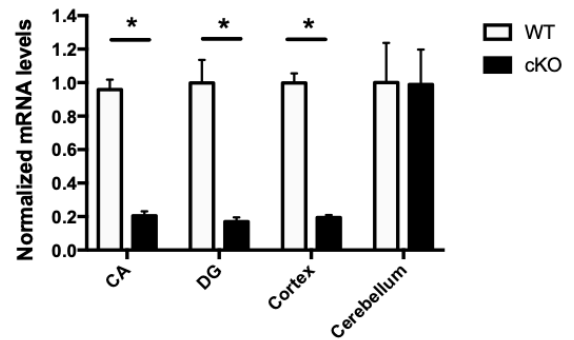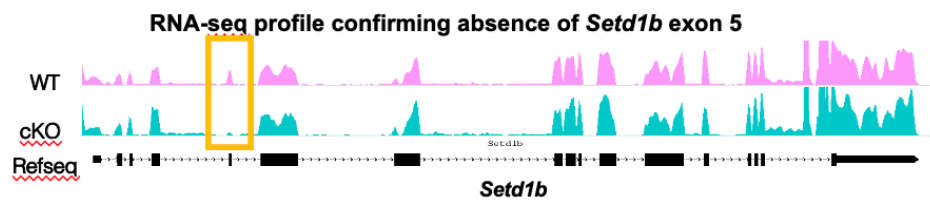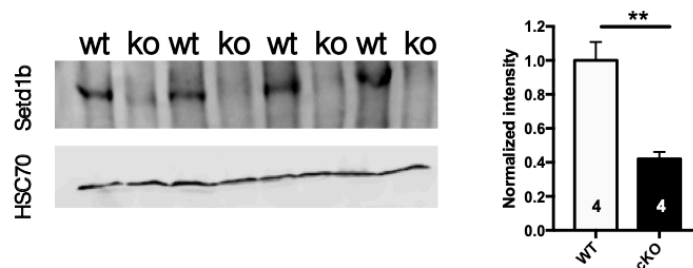
Figure 2-1 Design and validation of the Setd1b loss-of-function mouse model. (A) Breeding scheme for the Setd1b cKO mouse. (B) mRNA levels of Setd1b are decreased in cortical and hippocampal (CA and DG) regions but not in the cerebellum (WT: n=6, cKO: n=6; * Student t-test < 0.05) (C) RNA-seq profiles of wildtype and Setd1b cKO mice, confirming the loss of Setd1b exon 5 in cKO mice. (D) Western Blot analysis of Setd1b protein levels in wildtype and Setd1b cKO mice. Setd1b protein is significantly reduced in cKO mice (WT: n=4, cKO: n=4; ** Student t-test < 0.01).
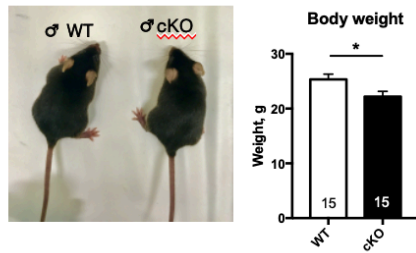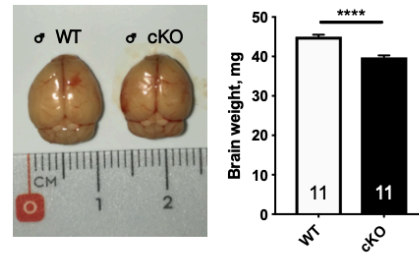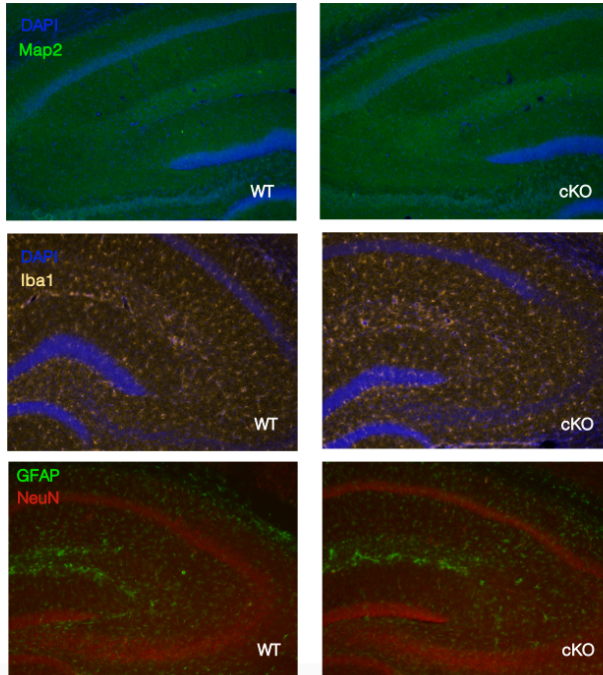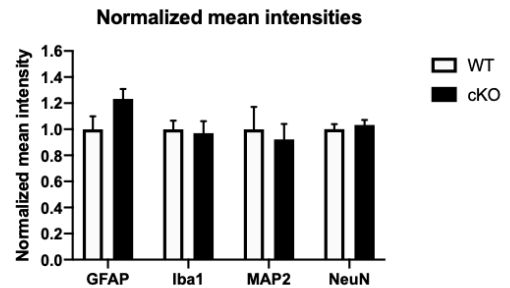
Figure 2-2 Setd1b cKO mice do not exhibit gross morphology changes. (A) Body weight is slightly decreased in cKO when compared to control animals (WT: n=15, cKO: n=15; * Student t-test < 0.05). (B) Brain weight is also slightly decreased, the brain does not exhibit any obvious anatomical differences (WT: n=11, cKO: n=11; **** Student t-test < 0.0001). (C) Also, immunostainings of the hippocampal region do not reveal any histological alterations. Scale bar 100 µm. (D) Quantification of marker proteins shows no difference between WT and cKO brain neuronal structures. Normalized mean intensities for immunostaining (WT: n=6, cKO: n=6; Student t-test). Error bars indicate SEM.

4.2. Setd1b Loss-of-function impairs hippocampus-dependent spatial memory

To evaluate hippocampus-dependent spatial memory formation in Setd1b cKO mice, we have subjected three months old animals to the most widely used paradigm of **Morris water maze (MWM)** test[161] (see Fig.3A). Test consisted of the training phase, during which we trained the animals to learn where the hidden platform is located based on the visual cues. While WT animals showed an everyday improvement in the time they were

59

managing to successfully find a platform location and reached the plateau of less than 20s escape latency after the 3rd day, cKO animals stopped any improvement in the latent time of finding hidden platform after the 5th day of training (see Fig. 3B). We also observed that on the very 1st day and the very first trial (when mice were for the first time introduced to water) groups did not show any differences in the swimming speed (data not shown). To exclude the effect of motor impairments, we compared the average swim speed for each day during learning phase. First few days speed between groups did not differ, however starting from the 4th day, cKO group started to differ in the swimming speed, resulting in approximately 10 cm/s more slowly speed than WT group on the last day before probe test (see Fig. 3D). To address this issue, we further performed rotarod test to evaluate whether Setd1b cKO exhibit motor abnormality and tail suspension depression test, as described in the next chapter. After 10 days of training we have carried out the **probe test (PT)** when the platform is being removed from the setup to access memory retrieval. Data evaluation showed that Setd1b cKO mice performed poorly in remembering the site of the platform: they sent significantly less time in the area of target quadrant (see Fig. 3E), less time in the target itself (see Fig. 3F) and the number of the crossings of the area where platform used to be located is strongly reduced (see Fig. 3G). Next we investigated the search strategies in the maze during learning phase for the both groups. A modified version of the MUST-C algorithm gave the hippocampus-dependent readout[155]. We observed growing preference to select cognitively challenging strategies during learning among WT animal group, such as "direct", "corrected", "focused" and "short-chaining" strategies. Setd1b cKO mice had longitudinal increase "circling", "long-chaining", "focused false", "random", "thigmotaxis" which were barely selected by WT group after the 3rd day of training. These strategies are considered to be less challenging and rather indicate poor cognitive abilities (see Fig.3C). Taken together, we may conclude that Setd1b plays a role in the modulating spatial memory and its deletion from the excitatory neurons of the postnatal forebrain leads to significant retardation in

hippocampus-dependent learning and memory abilities already from the 3rd month after birth. Since the tested mice were of both genders, we analyzed whether females and males performed differently (see Supp.Fig.3) and did not discover any sex-related differences. We made an attempt to deliver Cre-containing adeno-associated virus to the DG region of WT animals (Setd1b$^{fl/fl}$) in order to exclude the developmental defects upon Setd1b loss in our system and to evaluate purely hippocampus-dependent function. We collected DG region 14 days after the injection, since this timepoint has been previously shown in our laboratory as sufficient for viral expression. Virus expression was sufficient based on the amount of Cre-recombinase mRNA after injection, but we did not observe the reduction of Setd1b (see Supp.Fig.2A,B). To test whether Setd1b was not knocked down due to the increased half-life of the protein, we performed the same experiment three and four weeks after viral injection, but still the knock down did not succeed (data not shown).
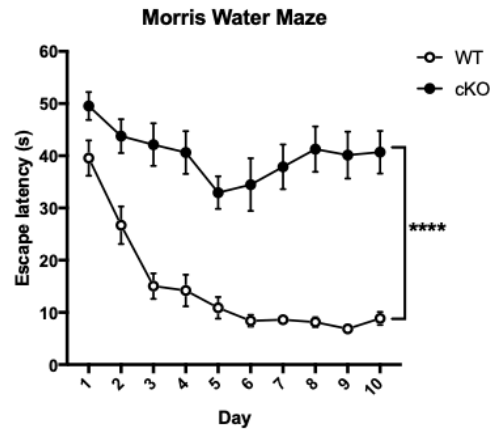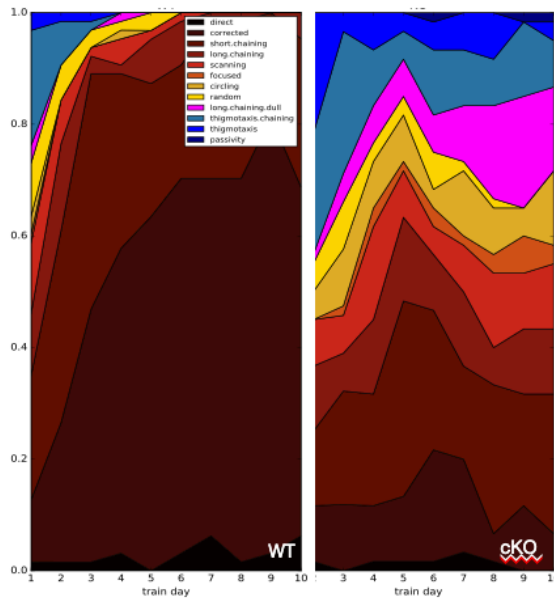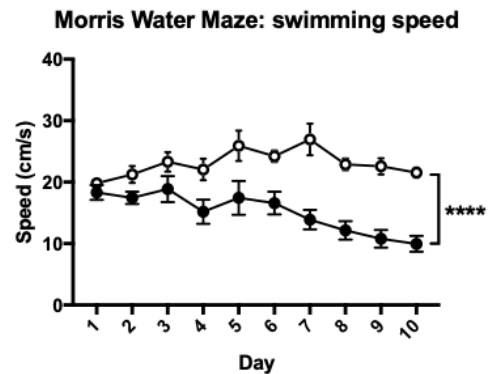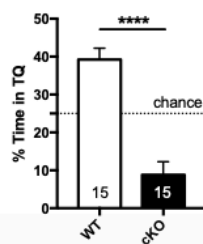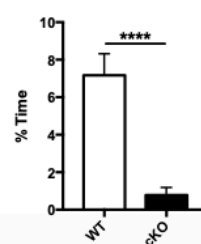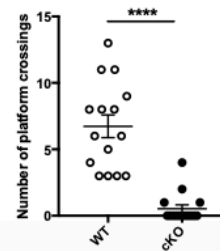
Figure 2-3 Spatial memory is severely impaired in Setd1b cKO mice in the Morris water maze test. (A) Scheme of the test setup. (B) Escape latency during MWM shows poor learning performance of cKO mice (WT: n=15, cKO: n=15; two-way-ANOVA, genotype effect: $F_{(1,28)}=82.34$; **** p-value < 0.0001). (C) Strategy analysis: Setd1b cKO lack cognitive challenging searching strategies. (D) Average speed during 10 days of training (WT: n=15, cKO: n=15; two-way-ANOVA, genotype effect: $F_{(1,14)}=35,53$; **** p-value < 0.0001. (E) % Time spent in the target quadrant during the probe test is reduced in Setd1b cKO mice (WT: n=15, cKO: n=15; Student t-test: **** p-value < 0.0001). (F) % of time spent in the target zone on the probe test day is reduced (WT: n=15, cKO: n=15; Student t-test: **** p-value < 0.0001). (G) Number of platform crossings on the probe test day is decreased (WT: n=15, cKO: n=15; Student t-test: **** p-value < 0.0001). Error bars indicate SEM.

4.3. Setd1b knock-out affects multiple cognitive functions

We assessed potential cognitive deficits of postnatal Setd1b loss-of-function in the number of other behavioral tests. In the **open field test (OF)** both cKO and WT showed the same levels of exploratory behavior and had no differences in the time spent in the center (see Fig.6A) or the speed (see Fig.6C) and distance travelled (see Fig.6B) indicating that basal anxiety is not affected and there is no major motor impairment. Working memory, very often affected in the mouse models of schizophrenia[162], was accessed with the **Y-maze test**, where we counted successful alterations as visits to arms that were not visited in the trial before, was not affected among cKO (see Fig.6D). Loss of Setd1b did not affect a **novel object recognition (NOR)** paradigm for hippocampal long-term memory consolidation (see Fig.6E). Both genotypes were spending similar time near the new object introduced 24h after the training day. Importantly, cKO mice showed higher than 50% preference for the new object on the test day, showing that they could distinguish the new object from the old. Absence of Setd1b in the postnatal forebrain did not induce pro-depressant behavior as can be seen from **tail suspension test (TST)**, paradigm used to access despair traits in rodents and often performed to evaluate antidepressant efficiency[163]. Both genotypes exhibited similar immobility time (see Fig. 6F), therefore it is unlikely that poor performance in the MWM was due to depressive phenotype and the lack of motivation. In **elevated plus maze (EPM)**, paradigm to test anxiety-like behavior, mice of both groups spent most of the time in the closed arms (see Fig.4B), however we observed that cKO spent significantly more time in the open arms (see Fig.4A). These results could be an indication of decreased anxiety among Setd1b cKO animals or, spoken differently, avoidance of the closed arms. Time in the center region did not differ among groups (data not shown). Total distance travelled in the arms of EPM did not differ between groups (see Fig.4C) confirming that there is no major motor deficit in cKO group.
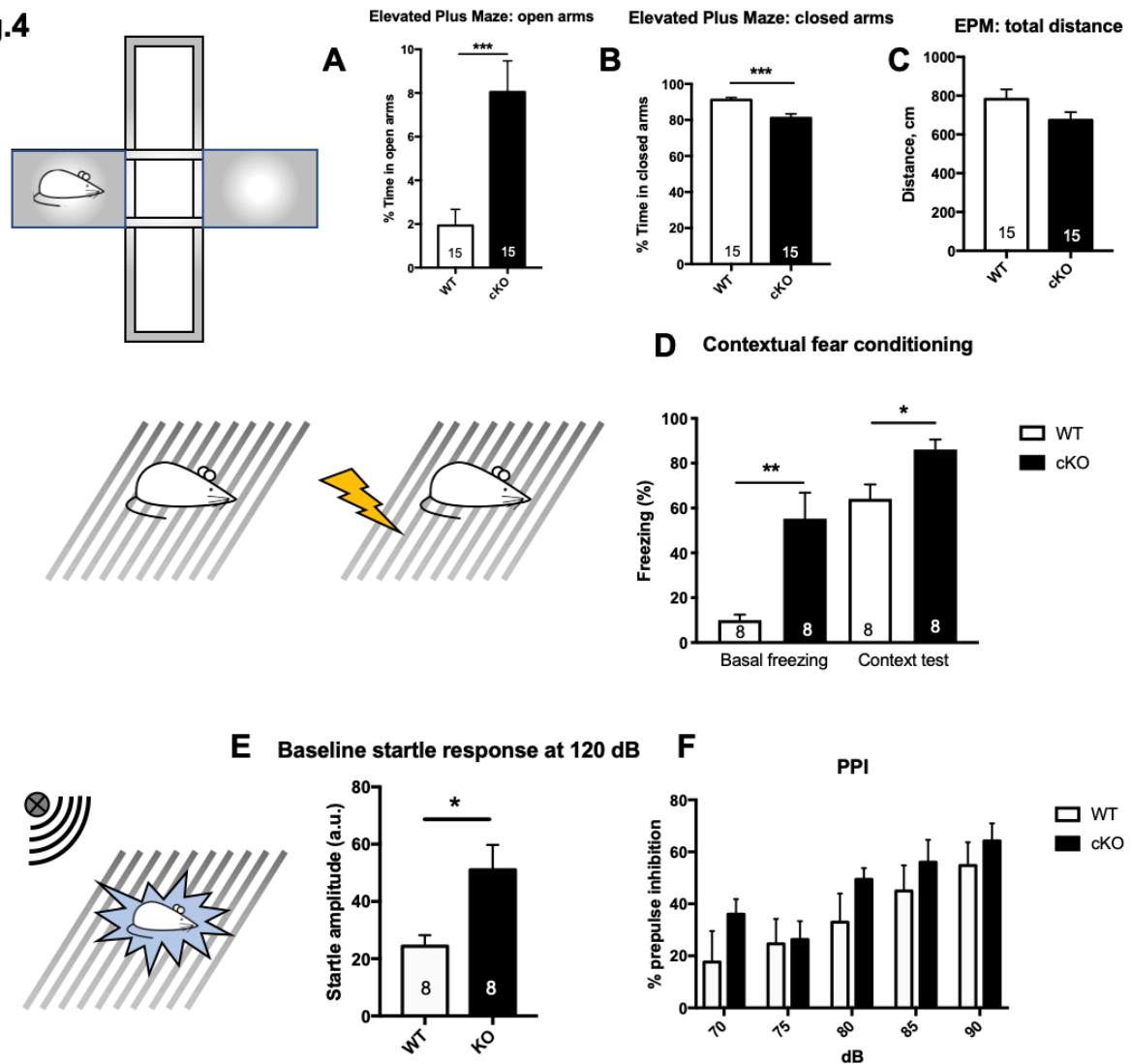
Fig.4

A Elevated Plus Maze: open arms

B Elevated Plus Maze: closed arms

C EPM: total distance

D Contextual fear conditioning

E Baseline startle response at 120 dB

F PPI

Figure 2-4 Loss of Setd1b affects different behaviors. (A) cKO mice spent more time in open arms of elevated plus maze (WT: n=15, cKO: n=15; Student t-test: *** p-value < 0.001). (B) cKO mice showed avoidance of closed arms maze (WT: n=15, cKO: n=15; Student t-test: *** p-value < 0.001). (C) Total distance travelled in EPM did not differ among two groups maze (WT: n=15, cKO: n=15; Student t-test: p-value=0.108). (D) Freezing of cKO mice occurred already during preliminary stage of habituation in the shock-chamber before electric shock during fear conditioning test (WT: n=8, cKO: n=8; Student t-test: ** p-value<0.01; * p<0.05). (E) Baseline startle response to 120 dB was increased in cKO mice (WT: n=8, cKO: n=8; Student t-test: * p-value<0.05). (F) Prepulse inhibition was not affected (WT: n=8, cKO: n=8; two-way-ANOVA, pulse intensity effect: $F_{(1,14)}=1.1$; p-value=0.3127). Prepulse intensity still had an effect mice (WT: n=8, cKO: n=8; two-way-ANOVA, pulse intensity effect: $F_{(2.2, 30.87)}=21.68$; **** p-value < 0.0001). Error bars indicate SEM.

We subjected animals to **rotarod** to test motor coordination, learning and locomotion impairment. Since locomotor activity has been shown to be not impaired in the open field, elevated plus maze and tail suspension test, results of rotarod test indicate rather absent

differences in the coordination and balancing, since there was no difference between two genotypes on the very first learning trial. However, further trials show increasing difference in the latency to fall off the rod between two groups (see Fig.5C), where cKO do not improve in the latent time. Therefore, we concluded that motor learning is mildly deregulated. Since we have shown that Setd1b expression levels are not altered in cerebellum (see Fig.1B), the phenotype we observe in this test might arise from the reduction of Setd1b in striatum, where CamKII is expressed[164]. This region has been shown to be involved in the motor learning on rotarod after knocking-out glutamate ionotropic receptor NMDA type subunit 1 (Grin1)[165]. Interestingly, we noticed by a casual observation that cKO animals failed to use paper tissues normally provided to every cage by the animal facility, while WT shredded provided tissues building proper nests. Therefore, we decided to perform **nest building test** based on commonly used protocol to assess this behavior[156]. WT male mice managed to shred the nestlet material and build an adequate nest overnight (for some WT females it took two nights), whereas only few cKO mice slightly shredded the provided nestlet and in the majority cages of other cKO material remained mainly untouched (see Fig.5A,B).

Fig.5

A Nesting behaviour

B

C Rotarod test
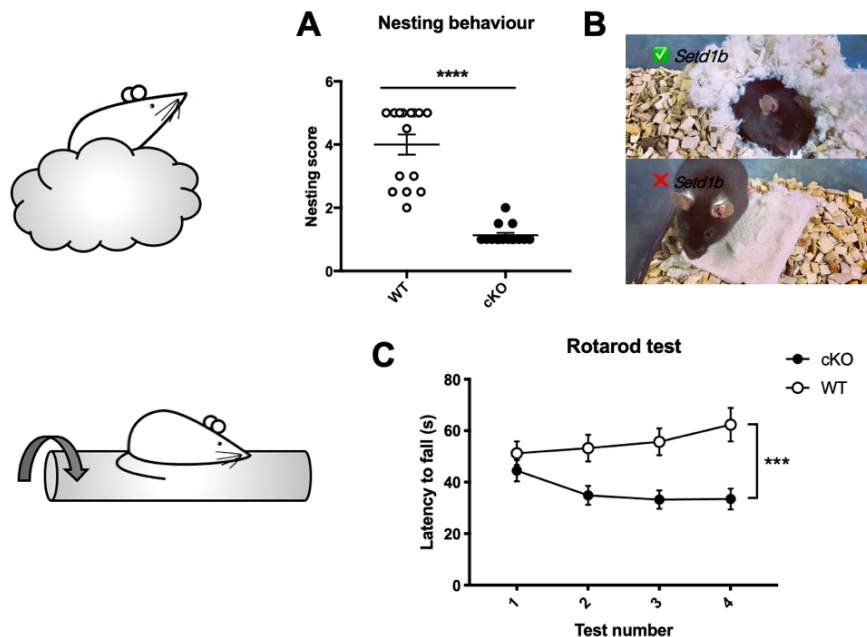
Figure 2-5 Setd1b cKO mice do not build nests and have moderate motorical learning impairment. (A) cKO mice do not make nests as checked in nest building test (WT: n=15, cKO: n=15; Student t-test: **** p-value < 0.0001). (B) Photo of WT mouse in its nest and cKO on the unshredded nestlet. (C) In rotarod test cKO had slightly decreased motor learning (WT: n=15, cKO: n=15; two-way-ANOVA, genotype effect: $F_{(1,28)}$=14.02; *** p-value < 0.001). Error bars indicate SEM.

Since nesting is thought to be connected to social skills in mice[166], we next performed **social novelty test** to see if the communication of male and female mice is affected (see Fig.6G). We did not observe significant difference in the time spent with novel mouse between cKO and WT group. A **contextual fear conditioning (CFC)** test was supposed to indicate the potential of cKO mice to consolidate long-term memories and it is known to be hippocampus dependent. However, we observed the very interesting phenomenon of cKO mice showing very low level of motility in the chamber even before the electric current was applied (see Fig.4D). This made it impossible to make proper conclusions about the freezing response 24h after subjecting animals to electric shocks and therefore access memory impairment, but it was an interesting observation per se which we have never observed before in our laboratory. The **prepulse inhibition (PPI)** of acoustic startle response is homologous in humans and rodents and is known to be impaired in many

schizophrenic patients and mouse models of schizophrenia[167]. The PPI concept is that weak sensory stimulus ("prepulse") applied shortly before the strong stimulus ("pulse") suppresses a startle reflex. It is considered to indicate sensorimotor gating, a way of inhibiting irrelevant sensory signals in CNS during the early processing steps, so that the organism can direct its attention towards more important signals. We did not reveal significant difference between cKO and WT group in PPI task, indicating that attenuation of motor response to a startling sound was not deregulated (see Fig.4F). However, the startle response itself among cKO was twice as high as in WT animals (see Fig.4E). Sensory gating defects and, as an outcome, hypersensitivity to a pre-stimulus has been also observed in individuals with ASD[168]. Increased startle amplitude has been previously observed in the autism mouse model with the loss-of-function in the gene coding glutamate ionotropic receptor NMDA type subunit 1 (Grin1)[169] and reduced levels of nSR100 splicing regulator[170]. Most probably, in our case increased baseline startle is observed due to startle response modulation from the cortex, where Setd1b is gone from the excitatory neurons. In all above-mentioned tests Setd1b WT and cKO male and female mice were tested together and showed the same trends and no sex-specific effects or sex-genotype interactions were detected.
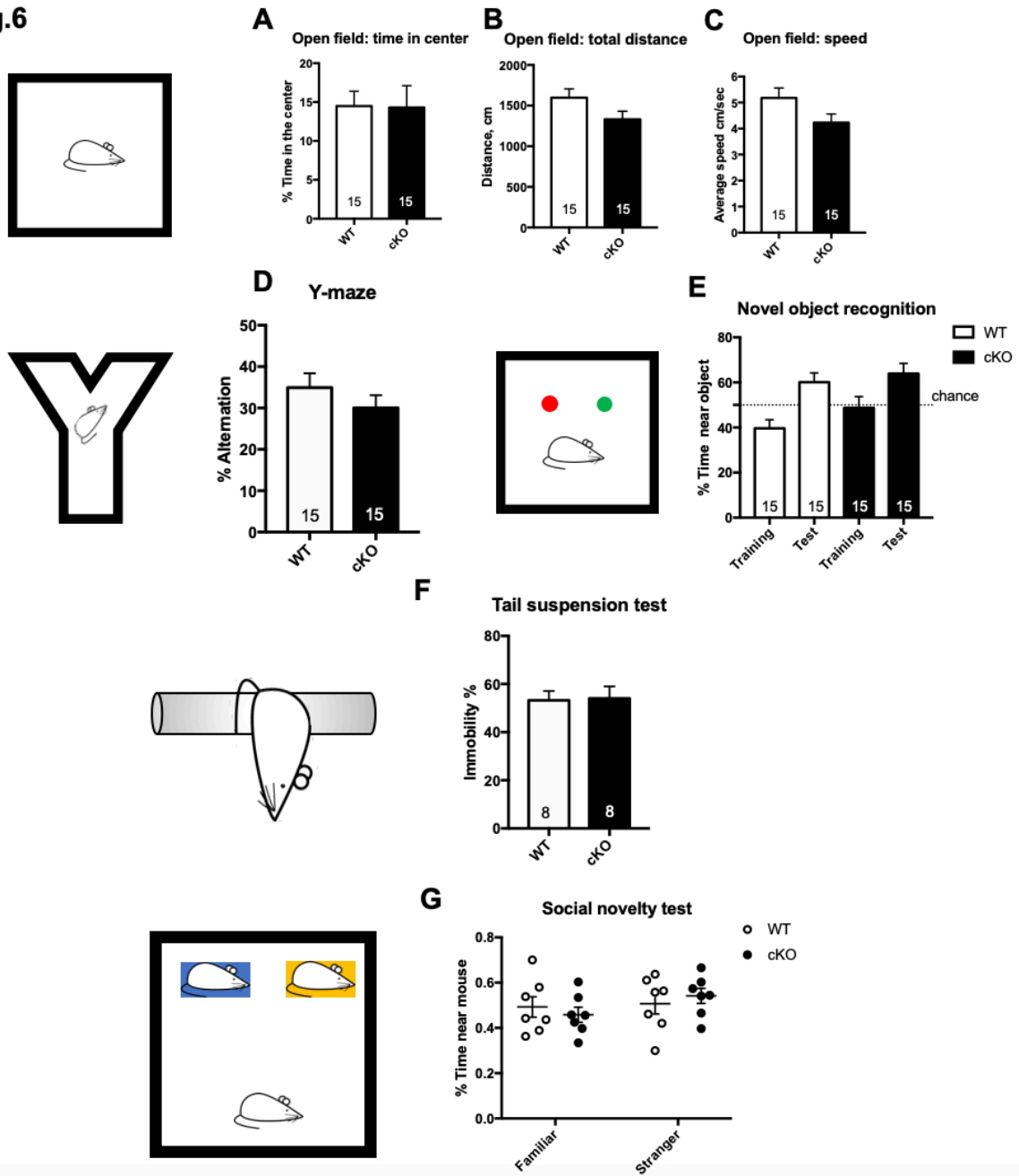
Figure 2-6 Non-affected behaviors upon Setd1b loss. (A) Time spent in the center of open field did not differ among two groups (WT: n=15, cKO: n=15; Student t-test: p-value=0.9574). (B) Total distance travelled in OF was similar between WT and cKO (WT: n=15, cKO: n=15; Student t-test: p-value=0.0763). (C) Speed in the OF was similar between two groups (WT: n=15, cKO: n=15; Student t-test: p-value=0.0749). (D) Working memory is not changed in Y-maze test maze in cKO group (WT: n=15, cKO: n=15; Student t-test: p-value=0.299). (E) No changes in short-term memory test of novel object recognition (WT: n=15, cKO: n=15; Student t-test: p-value=0.766). (F) Depression phenotype is absent according to tail suspension test (WT: n=8, cKO: n=8; Student t-test: p-value=0.897). (G) Setd1b cKO do not differ in time spent near stranger mouse in social novelty test (WT: n=15, cKO: n=15; Student t-test: p-value=0.0701). Error bars indicate SEM.

To exclude the effect of cre-insertion under CamKII promoter in our cKO group to the above-mentioned behaviors we have subjected cre+ (n=8) and cre- (n=8) animals of transgenic CamKII line. We have shown that CamKII-promoter controlled Cre-expressing mice did not have reduction of body weight and did not show any anomalies in behavioral testing when compared to control animals (see Supp.Fig.1), including the most severe Setd1b phenotypes such as basal freezing in fear conditioning (see Supp.Fig.1H) and MWM (see Supp.Fig.1D,E,F,G). Cre+ mice also did proper nests (data not shown).

In summary, we performed profound behavioral characterization of mice with postnatal neuronal loss of Setd1b histone methyltransferase from the forebrain excitatory neurons. Setd1b deficient mice and their wild-type littermates showed no differences in the open field test, Y-maze, novel object recognition, tail suspension and social novelty test. However, cKO animals showed aberrant cognitive phenotype in Morris water maze, basal freezing in contextual fear conditioning, acoustic startle response, rotarod motor learning, nest building and elevated-plus maze. Behavioral alterations observed are most probably not related to gross morphology changes, since we did not observe any (see Fig.2C,D).

4.4. Setd1b epigenetic and RNA regulatory neuronal functions

After the identification of impaired hippocampal learning, memory formation and cognitive disabilities in Setd1b cKO mice, we wanted to unravel the processes by which Setd1b influences memory formation in molecular detail. For this endeavor, we have investigated the epigenetic influence on gene expression in hippocampal neurons. Prior to this study, we have accessed transcriptomics of CA and DG of Setd1b cKO animals with the whole tissue RNA-seq (see Supp.Fig.5). Later, in our laboratory we have developed a more precise way of analyzing gene expression changes in neuronal cells only, which we utilized in this study for the first time. Hippocampal tissue of the CA region was obtained via microdissection from Setd1b cKO and WT animals and neuronal cells were isolated through anti-NeuN antibody mediated FACS (see Fig.7A-C). Nuclei were then prepared of the sorted cells, which were then used to perform neuron-specific chromatin-immunoprecipitation (ChIP) and RNA-seq, generating two different datasets from the same mouse brain tissue (courtesy of Sakib Sadman and Dr. Cemil Kerimoglu). Specificity of the anti-NeuN antibody for neuronal cells was validated through specifically enriched genes in the RNA-seq and GO-term analysis (see Fig.7D,E). Genes specific for neuronal (see Fig.7F) and glial cells (see Fig.7G) were further quantified for validation of the neuronal specificity of NeuN+ mediated FACS.

For the ChIP experiments, we have analyzed the histone modifications H3K4me3, H3K4me1, H3K9ac and H3K27ac. H3K4me3, a marker for active genes and transcription start sites (TSS), was chosen as Setd1b is a H3K4 methyltransferase, with H3K4me1 being an intermediate transitional state[171]. H3K9ac has been reported to be associated with H3K4me3[77,78], and H3K27ac modification is enriched at active gene TSS and enhancer regions[172], with no direct correlation to H3K4me3. As a result, H3K4me3 levels are significantly decreased in neuronal nuclei of Setd1b cKO mice throughout the genome,

and the majority of changed H3K4me3 modifications occur at TSS regions (see Fig.8A,B). The same could be observed for the H3K9ac and H3K27ac epigenetic modifications, however to a lesser extent than H3K4me3 (see Fig.8C). H3K4me1 levels were also significantly changed in Setd1b cKO neurons, but in contrast to the H3K27ac modifications, it was increased in the proximity of TSS (see Fig.8A, B). Further analysis showed, that of the four epigenetic marks, H3K4me3 is the most affected histone modification in the proximity of TSS (see Fig.8D). In addition and as previously reported, H3K4me3 can be functionally linked to H3K9ac[77,78], as the genes affected exhibit a major overlap (see Fig.8E), while the H3K27ac epigenetic mark does not. Total numbers of differentially methylated and acetylated regions for each mark are shown in Supp.Fig.6.
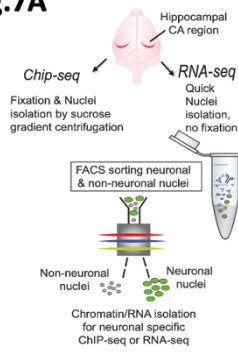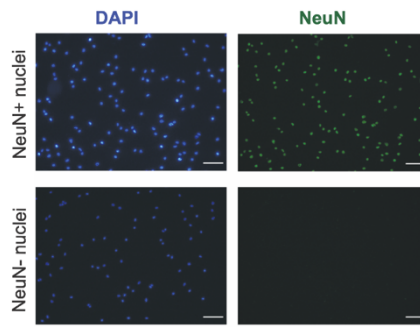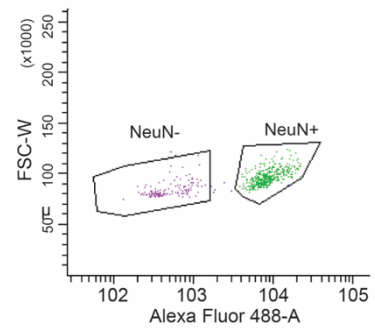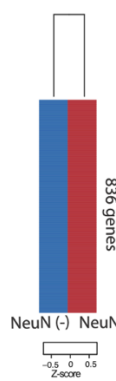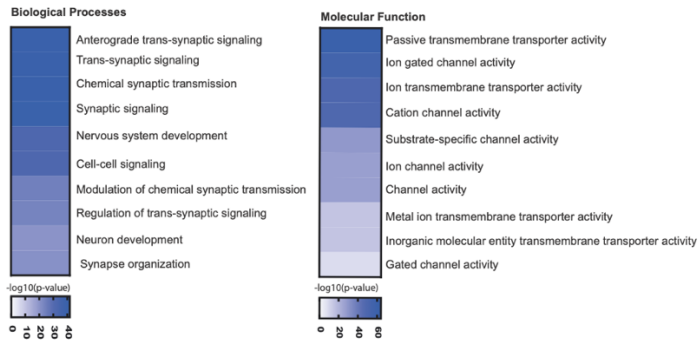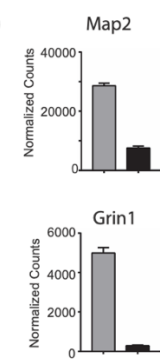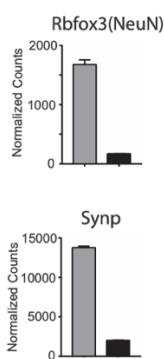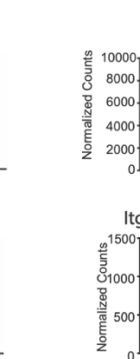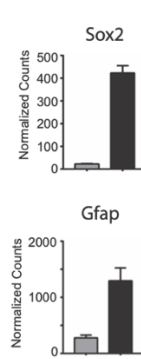
Figure 2-7 Neuronal specific nuclei sorting for ChIP and RNA-seq. (A) Experimental pipeline for cell-type specific ChIP-seq and RNA-seq. For Chip-seq we sacrificed 4 control and 4 Setd1b cKO mice. (B) Representative images showing nuclei that were sorted using the neuronal marker NeuN. Scale bar: 50μm. (C) Gating strategy for sorted NeuN(+) and NeuN(-) nuclei. (D) RNA-seq (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, base mean > 150, fold change > 5. (E) 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. (F) Normalized expression values obtained from the RNA-seq experiment showing the expression of selected genes known to be enriched in neurons. (G) Normalized expression values of genes that are known to be enriched in non-neuronal cells including glia cells. Error bars indicate SEM. (A-G) are courtesy of Sakib Sadman.

Most interestingly, H3K4me3 decrease in Setd1b cKO neurons is the most prominent downstream of TSS, effectively influencing the width of the epigenetic mark at the site (see Fig.8F). Further analysis revealed, that this reduced peak width was most prominent in genes which exhibited both a decrease in H3K4me3 and an increase in H3K4me1 (see Fig.8G), when compared to genes where only H3K4me3 was decreased (see Fig.8H). Broad H3K4me3 peaks are correlated with very stable gene expression levels and cellular identity[173]. We have therefore checked genes sorted by H3K4me3 peak width at TSS for increased H3K4me1, and indeed, genes with already broad H3K4me3 peaks exhibit increased H3K4me1 levels in the proximity of TSS in Setd1b cKO mice (see Fig.8I). Furthermore, the genes that have a broad H3K4me3 peak are also expressed at a higher rate by default (see Fig.8J). Functional analysis revealed that the genes with elevated H3K4me3 levels at the TSS in Setd1b cKO mice are indeed genes associated with specific neuronal identity and functions in hippocampal excitatory neurons (see Supp.Fig.8). Taken together, Setd1b seems to play a crucial role for neuronal identity, learning and memory formation, through its histone methylation function, in hippocampal neurons of the mouse.

Analysis of the RNA-seq data of Setd1b cKO mouse sorted nuclei of the CA region revealed, that indeed the majority of affected genes is downregulated (see Fig.9A), and that downregulated genes are indeed associated with decreased H3K4me3 peak width and elevated H3K4me1 around the TSS (see Fig.9B,C). As a control, genes unaffected by Setd1b exhibited normal H3K4me3 peak width and no enrichment of the H3K4me1 epigenetic modification, when compared to control animal RNA-seq results (see Fig.9D). As described above, genes affected by the Setd1b cKO normally have a broader H3K4me3 peak width and are expressed at higher rates than genes not impacted by Setd1b (see Fig.9E). A GO functional analysis confirmed the previous finding, that Setd1b regulated genes are active in synaptic plasticity and cognitive function pathways (see Fig.9F). In

summary, our RNA-seq data also suggests, that Setd1b regulates a subset of genes, which are characterized by broad H3K4me3 peak width, are highly expressed at basal levels, and function in neuronal plasticity, memory formation and learning pathways.
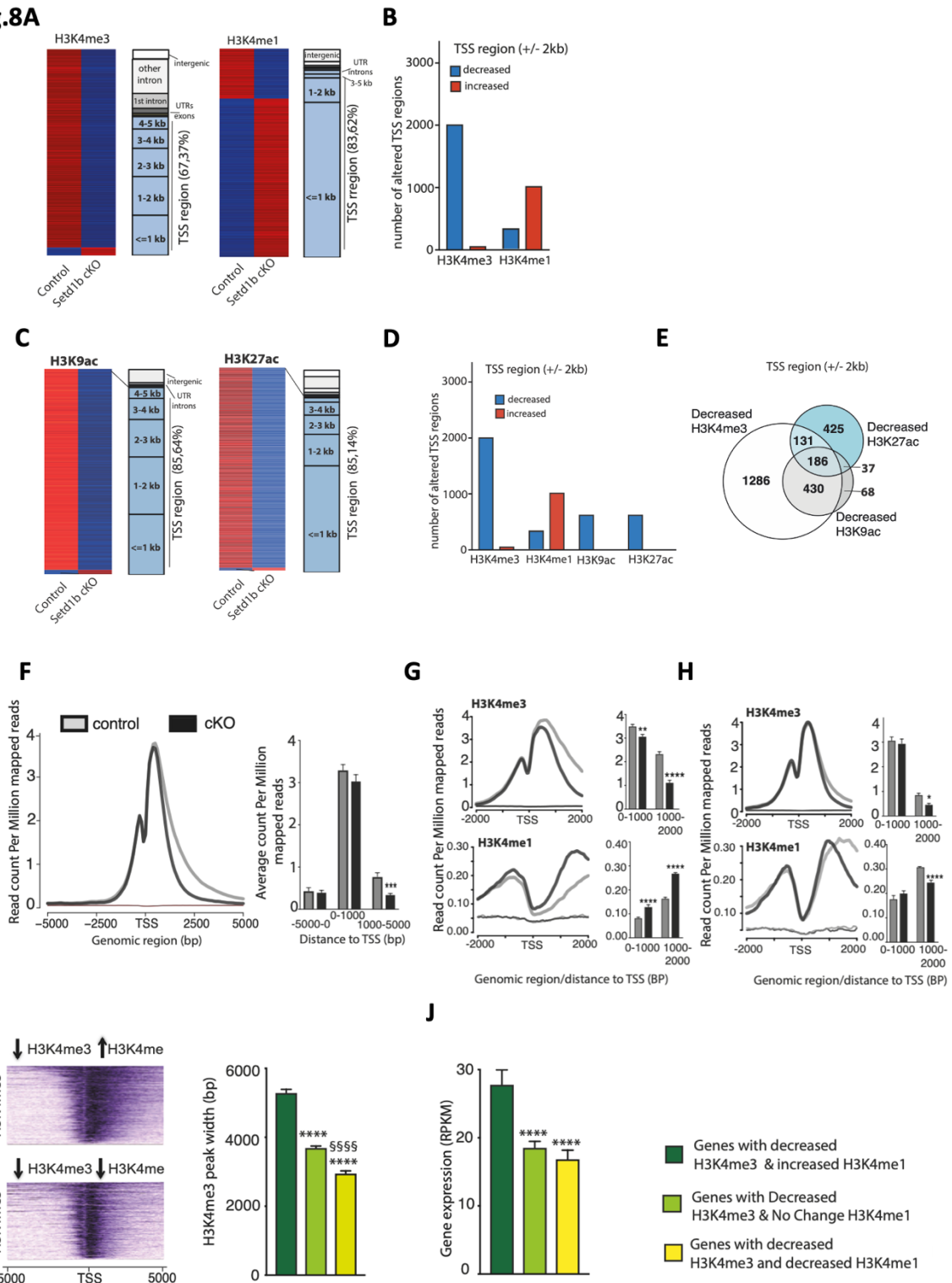
Figure 2-8 Setd1b controls histone-methylation and H3K4me3 peak width. (A) Left: Genes with significant differential H3K4me3 levels at TSS (+/-2kb) in Setd1b cKO compared to wildtype mice and overall loci with altered H3K4me3. Right: Similar analysis for H3K4me1 (FDR<0.05 & |fold change|>1.5). (B) Number of genes with decreased and increased H3K4me3 and H3K4me1 levels at TSS in Setd1b cKO mice (FDR<0.05 & |fold change|>1.5). (C) Left: Genes with differential H3K9ac at TSS in Setd1b cKO mice. Right: Similar analysis for H3K27ac levels. (D) Genes with decreased and increased H3K4me3, H3K4me1, H3K9ac and H3K27ac levels at TSS. (E) Decreased H3K9ac, but not H3K27ac at the TSS correlates with reduced H3K4me3 levels. (F) NGS plot showing reduced H3K4me3 peak width in Setd1b cKO mice at TSS. Right: Reduced H3K4me3 in Setd1b cKO mice is mainly occurring downstream of TSS (*** Student t-test p-value < 0.001). (G) NGS plots showing H3K4me3 and H3K4me1 at TSS of genes which show significantly reduced H3K4me3 and increased H3K4me1 in Setd1b cKO mice. Right panel shows quantification (Student t-test: ** p-value < 0.01, **** p-value < 0.0001). (H) NGS plot showing H3K4me3 and H3K4me1 at TSS of genes that show both reduced H3K4me3 and H3K4me1 in Setd1b cKO mice. Right panel shows quantification (Student t-test: * p-value < 0.05, **** p-value < 0.0001). (I) Left: Basal H3K4me3 peak width for genes characterized by decreased H3K4me3 in combination with either increased or decreased H3K4me1 in Setd1b cKO mice. Right: Peak width quantification of 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). (J) Basal wild type expression levels of genes that display altered H3K4me3 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=0.6967). (A-J) are courtesy of Dr. Cemil Kerimoglu.
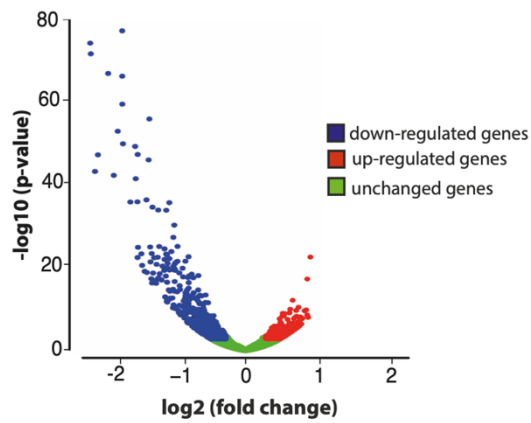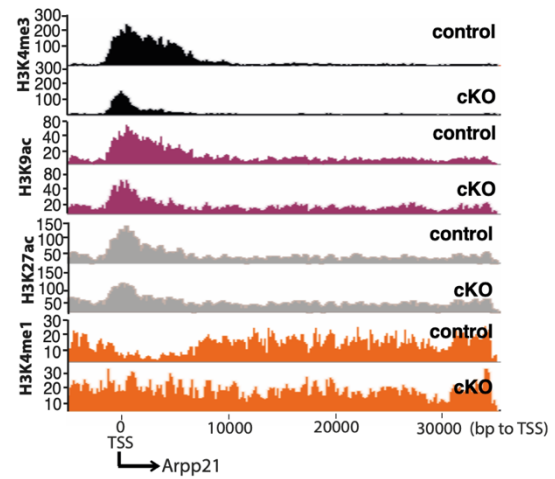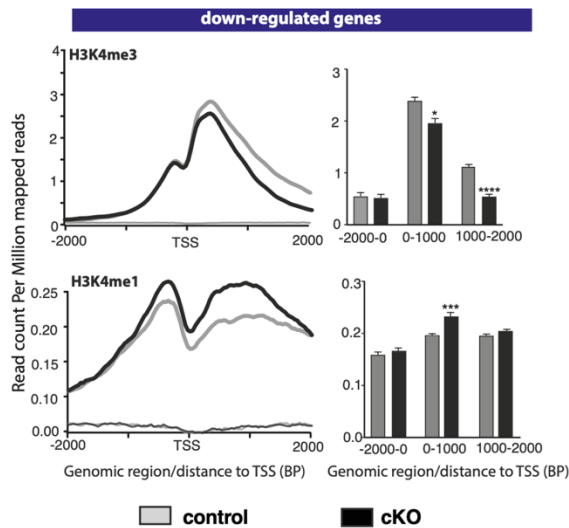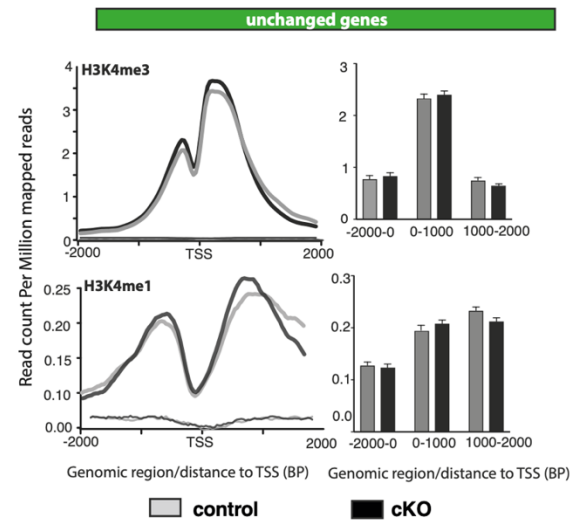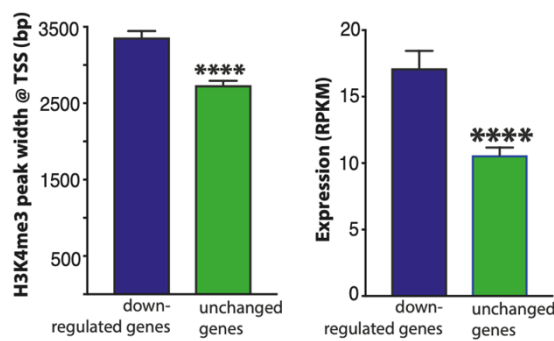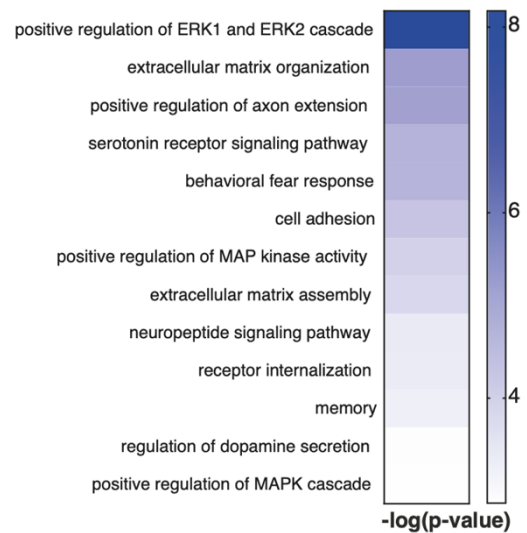
Figure 2-9 Setd1b controls highly expressed genes correlated with learning and memory which exhibit broad H3K4me3 peaks. (A) Genes differentially expressed in hippocampal neurons of Setd1b cKO mice.

n=3. (B) H3K4me3 peak width is significantly decreased in Setd1b cKO mice, but there is a marked increase of H3K4me1 at the TSS of the example gene Arpp21 in Setd1b cKO mice. (C) NGS plots showing H3K4me3 and H3K4me1 at TSS of downregulated genes in Setd1b cKO mice. Right panel shows quantification. (D) NGS plots showing H3K4me3 and H3K4me1 at TSS of genes that were not altered in Setd1b cKO mice. Right panel shows quantification. (E) Left: H3K4me3 peak width is significantly broader in downregulated genes, when compared to genes that were unaffected in Setd1b cKO mice. Right: Genes downregulated in Setd1b cKO mice exhibit a higher baseline expression when compared to genes that were unaffected. (F) Heat map showing functional pathways affected by downregulated genes in Setd1b cKO mice. Error bars indicate SEM. Student t-test: * p-value < 0.05, **** p-value < 0.0001. (A-F) are courtesy of Dr. Cemil Kerimoglu.

4.5. Setd1b specifically regulates highly expressed genes with neuronal functions, which exhibit broad H3K4me3 peak width

To evaluate the specificity of Setd1b function among other H3K4 methyltransferases, we have compared our Setd1b cKO ChIP datasets with H3K4me3 and H3K4me1 ChIP-seq data of cKO mice, in which Kmt2a or Kmt2b have been conditionally knocked out in excitatory neurons of the forebrain[78]. As expected, cKO of the different H3K4 methyltransferases decreased H3K4me3 at TSS in a number of genes in all three datasets (see Fig.10A), but interestingly, only Setd1b cKO induced H3K4me1 accumulation at TSS in a significant subset of genes (see Fig.10B). Furthermore, the subsets of genes with decreased H3K4me3 did not overlap among the three different datasets to a greater extent, suggesting that Kmt2a, Kmt2b and Setd1b each have specific regulatory functions in excitatory neurons of the mouse forebrain (see Fig.10C). Interestingly, among these three enzymes, Setd1b was the only one who regulated only 25% of trimethylation in the TSS (+/-1kb) regions, while Kmt2a and Kmt2b regulated 84% and 61% correspondingly (see Supp.Fig.7). To further investigate this finding, we have compared gene expression changes in whole tissue RNA-seq datasets of the hippocampal CA region of Kmt2a, Kmt2b and Setd1b cKO mice. There was again not a large overlap of downregulated genes, further suggesting specific functions in each of the three H3K4 methyltransferases in the mouse forebrain (see Fig.10D and Fig.12B). In more details this will be discussed in the Results 4.7 section. Also, GO term analysis again confirmed genes impacted by Setd1b cKO to be

involved in neuronal identity pathways, which was not the case for the Kmt2a and Kmt2b datasets (see Fig.10E). Comparative analysis with the ChIP data revealed, that genes decreased in expression in each of the H3K4 methyltransferases also displayed decreased H3K4me3 at TSS, but only Setd1b exhibited additional accumulation of H3K4me1 at these TSS (see Fig.10F). Further investigations revealed genes affected by Setd1b to have broader H3K4me3 peaks at TSS (see Fig.10G) and higher gene expression levels (see Fig.10H) in normal conditions, when compared to genes impacted by either Kmt2a or Kmt2b cKO. Similar to the NeuN+ RNA-seq, GO term analysis showed Setd1b regulated genes to be associated with learning, memory formation and hippocampal neuronal functions (see Fig.10I). Taken together, Setd1b is important for hippocampal neuronal identity and synaptic plasticity in the mouse forebrain.
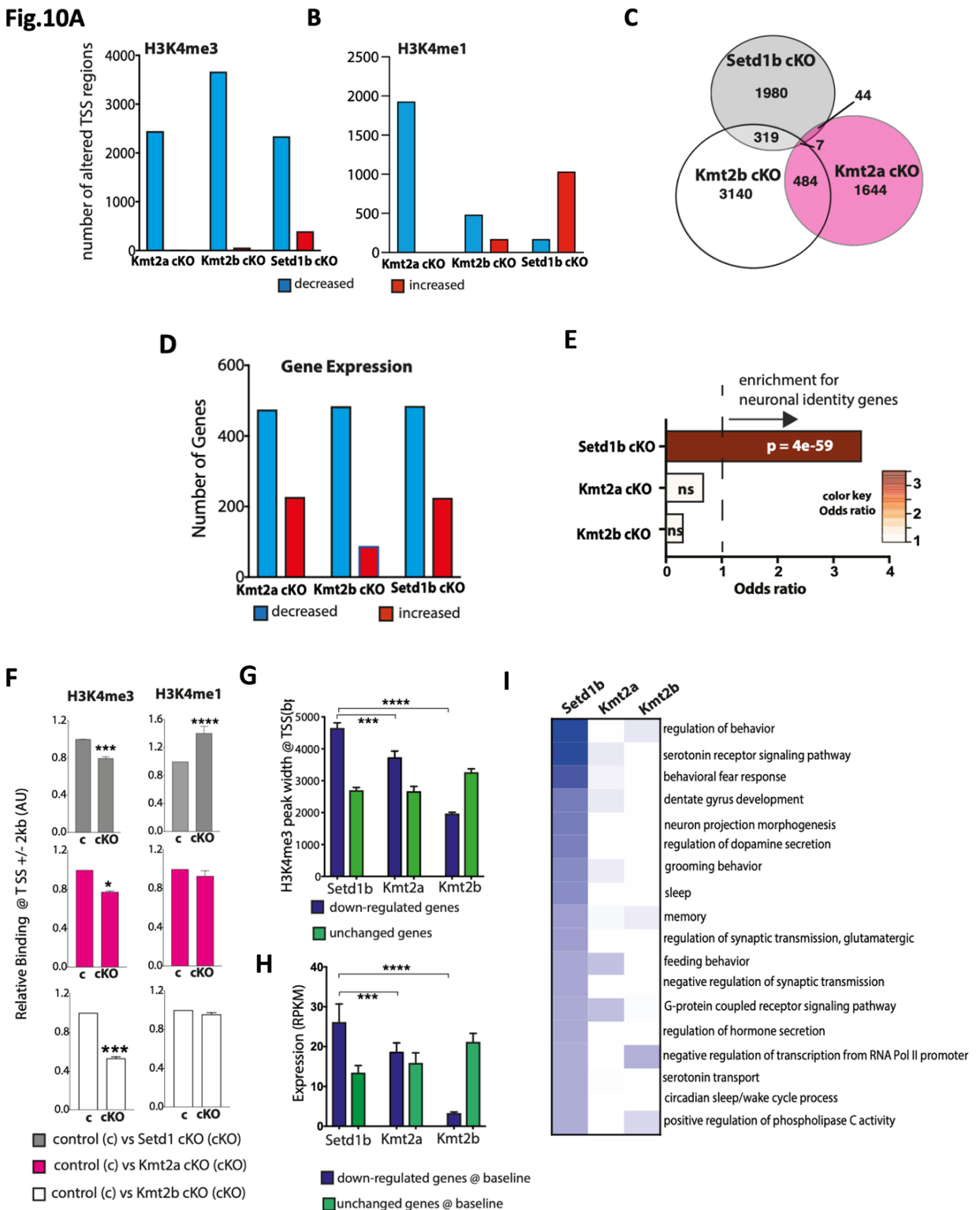
Figure 2-10 Comparative analysis of the hippocampal transcriptomes in Setd1b, Kmt2a and Kmt2b cKO mice. (A) Number of genes with significantly altered H3K4me3 at TSS (Kmt2a: control, n = 5; cKO, n = 3. Kmt2b: control, n = 6; cKO, n = 5. Setd1b: control, n = 4; cKO, n = 4). (B) Number of genes with significantly altered H3K4me1 at TSS. (C) Overlap of genes with significantly decreased H3K4me3 at TSS among Kmt2a, Kmt2b and Setd1b cKO mice. (D) Number of differentially expressed genes from RNA-seq data of the three different cKO mice (Kmt2a: control, n = 5; cKO, n = 6. Kmt2b: control, n = 8; cKO, n

= 11. Setd1b: control, n = 6; cKO, n = 6). (E) Downregulated genes with decreased H3K4me3 were tested for overlap with 836 neuronal identity genes defined earlier for the hippocampal CA region (see Fig.7D). (F) Left: H3K4me3 binding around TSS of downregulated genes exhibiting significantly decreased H3K4me3 in the three respective cKO mice (Two-way ANOVA: * p-value < 0.05, *** p-value < 0.001). Right: H3K4me1 for the same TSS regions as for H3K4me3 (Two-way ANOVA **** p-value < 0.0001). (G) Genes exhibiting decreased H3K4me3 and reduced expression in Kmt2a, Kmt2b or Setd1b cKO mice were analyzed for H3K4me3 peak-width at TSS. (H) Basal expression levels of genes downregulated with decreased H3K4me3 at the TSS in Kmt2a, Kmt2b or Setd1b cKO mice. (I) Heat map showing functional pathways of genes affected in Kmt2a, Kmt2b or Setd1b cKO mice. Error bars indicate SEM. (A-I) are courtesy of Dr. Cemil Kerimoglu.

4.6. Setd1b conveys specific functions in a subset of neurons as revealed by single cell sequencing analysis

Among the H3K4 methyltransferases in the mouse forebrain, according to our NeuN+ RNA-seq data, Setd1b is significantly less expressed in comparison to Kmt2a or Kmt2b (see Fig.11A). Setd1b is therefore either uniformly expressed at lower levels in all neurons, or not all cells express Setd1b, a hypothesis that cannot be addressed by our NeuN+ RNA-seq data performed on total brain tissue of the CA region. To address this question, we have performed single cell sequencing on NeuN+ FACS sorted cell nuclei (see Fig.7A). As a result, we obtained sequencing data of CA region specific excitatory neurons, neuronal cells specific to the dentate gyrus and also inhibitory neurons (see Fig.11B). The dataset of excitatory neurons was taken and expression levels of Kmt2a, Kmt2b and Setd1b were analyzed. As a result, Kmt2a was uniformly expressed among all sequenced nuclei (see Fig.11C), whereas Kmt2b and Setd1b, with the expression of all three methyltransferases being at a comparable level (see Fig.11D), were only expressed in a subset of cells (see Fig.11E). This finding explains the discrepancy found in the bulk RNA-seq data obtained from CA tissue, and confirms the hypothesis, that Setd1b is expressed only in a subset of selected neurons in the CA region. Most interestingly, when comparing the datasets of nuclei expressing exclusively Kmt2a, Kmt2b and Setd1b with each other, Kmt2a seems to specifically increase expression of 897 genes, whereas 432 genes were specifically upregulated by Kmt2b and 214 genes by Setd1b (see Fig.11F). Although

Setd1b is expressed in the smallest faction of nuclei and upregulates the smallest subset of genes, when compared to Kmt2a and Kmt2b, but its impact on the upregulation of gene expression is the strongest with 66% and 1,5x higher upregulation vs. Kmt2a 7% and Kmt2b 41% (see Fig.11F). GO term analysis of the dataset revealed that genes upregulated by Setd1b function in pathways linked to hippocampal function and histone acetylation, whereas Km2ta and Kmt2b impact these pathways to a much lesser extent (see Fig.11G). In summary, Setd1b function in the hippocampus seems to intimately be linked to neuronal identity, and the subset of Setd1b expressing neurons, when compared to Kmt2a and Kmt2b expressing cells, might have a pivotal impact on neuronal plasticity through a stronger impact on gene expression and specific activation of hippocampal and histone acetylation pathways.
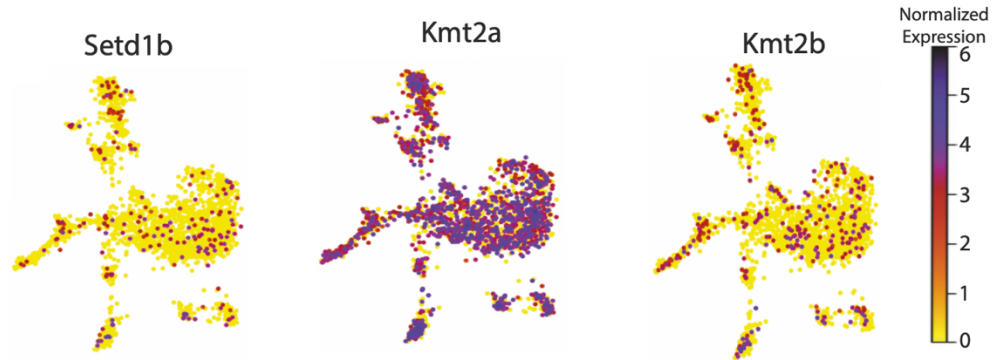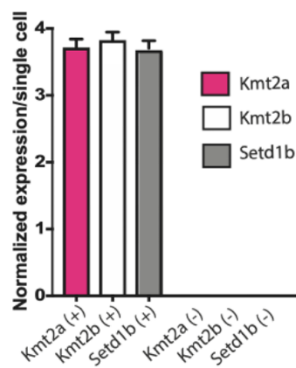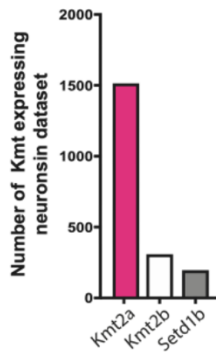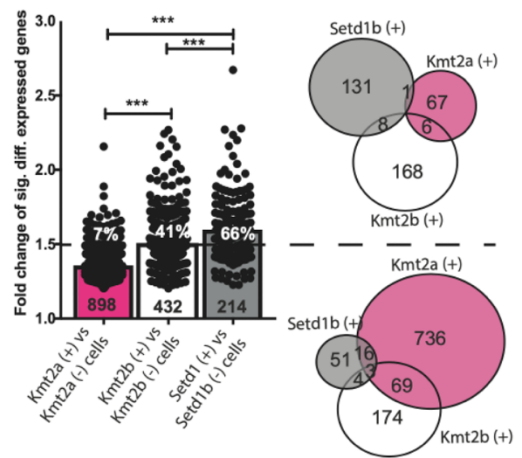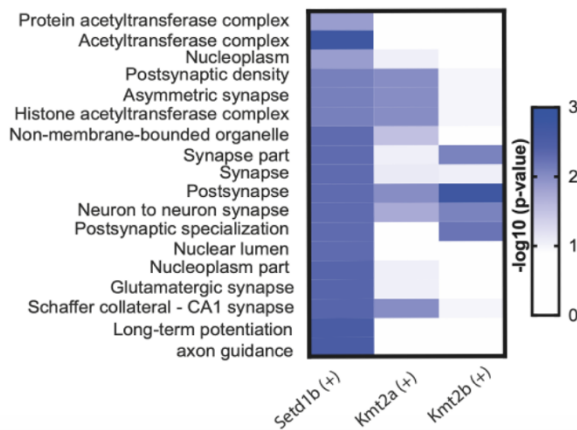
Fig.11A
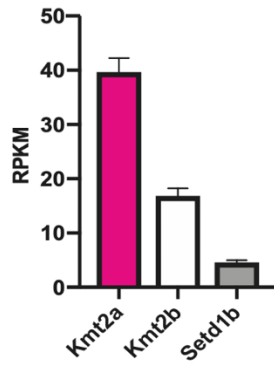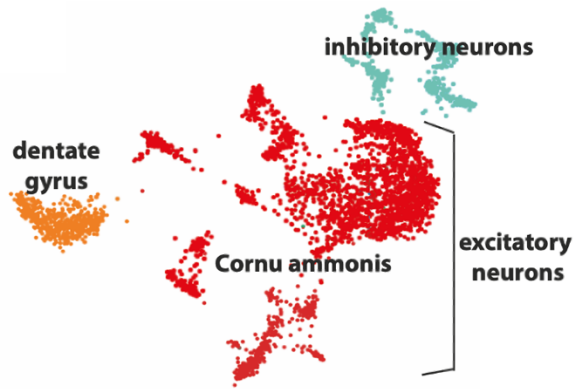
Figure 2-11 Setd1b functions are distinct to Kmt2a and Kmt2b as revealed by single cell sequencing data analysis. (A) Expression of Kmt2a, Kmt2b and Setd1b in neuronal nuclei from the hippocampal CA region (n = 3798). (B) UMAP plot showing the data from 3798 neuronal nuclei. (C) UMAP plot showing the clustering of 2619 nuclei from hippocampal excitatory neurons from the CA region indicating the normalized expression of Setd1b, Kmt2a and Kmt2b. (D) Normalized expression of Kmt2a, Kmt2b and Setd1b in respective positive cells. (E) Number of nuclei positive for Kmt2a, Kmt2b or Setd1b. (F) Differential expression analysis for Kmt2a, Kmt2b and Setd1b positive nuclei vs. negative nuclei. Left: Fold change of genes significantly increased in either Kmt2a, Kmt2b or Setd1b positive nuclei. Right: Comparison of 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). (G) Top GO and Kegg pathways representing the genes increased in Setd1b positive nuclei. (A-G) are courtesy of Dr. Cemil Kerimoglu and Dr. Dennis Krüger.

4.7. Differences and similarities between different KMT cKOs

We have previously studied Kmt2a and Kmt2b cKO mice, where the ablation of methyltransferases was generated in a similar manner to Setd1b as previously shown. We are now able to analyze the similarities and differences between behaviors (see Tabl. 14) and try to correspond them to the altered transcriptome and trimethylation data for all three enzyme datasets.

Kmt2a conditional knock-out mice also failed to build nests, had impaired spatial memory and increased anxiety in the open field test and light-dark box test[174]. Furthermore, as revealed in a study of our group, Kmt2a cKO had poor performance in the contextual fear conditioning, Morris water maze and decreased anxiety behavior in the elevated plus maze[78]. Of note, viral ablation of Kmt2a from the excitatory neurons of the striatal area induced pro-depressant (tail suspension test and forced swimming test) and anxiogenic behavior (elevated plus maze)[175]. Kmt2b loss affected hippocampus-dependent memory based on contextual fear conditioning and novel object recognition and performance during the probe test of MWM[77].

Kmt2a and Setd1b similarity	Kmt2a cKO only	Setd1b only
- Absent nesting behavior*[174] - More time in the open arms of EPM[78] - Spatial memory impairment in MWM[78]	- Impaired learning in CFC[78] - Increased anxiety in OF and light-dark box*[174] - Spatial memory impairment in T-maze and eight-arm radial maze*[174]	- Increased basal freezing in CFC - Increased acoustic startle response - Avoidance of closed arms in EPM/more time in the open arms - Disturbed learning in rotarod (was tested in Kmt2a cKO and not affected*)[174]
Kmt2b and Setd1b similarity	Kmt2b cKO only	
- Spatial memory impairment in MWM during probe test[77]	- Increased anxiety in OF and EPM[175] - Impaired learning in NOR - Impaired learning in CFC[77]	

Table 14: Comparative analysis of behaviors observed among Kmt2a, Kmt2b and Setd1b cKO (*marked results are originating from observation of another scientific group who used the same mouse model but accessed behaviors through different tests).

Disturbed performance in MWM test is the only uniting behavioral phenotype among Kmt2a, Kmt2b and Setd1b observed in our group. We have plotted MWM data previously generated in our laboratory with the same experimental protocol for Kmt2a and Kmt2b against newly obtained Setd1b data as normalized to the corresponding control escape latency (See Fig.12A). Setd1b cKO animals have the most severe learning impairment among three KMT cKOs. The degree of spatial memory impairment seems to be corresponding with the obtained gene expression data, in which Setd1b appeared to be the most crucial regulator for neuronal identity, learning and memory formation in comparison of all three KMTs, while Kmt2a loss affected more general cellular pathways together with neuronal function genes. Kmt2b mice, who had the mildest impairment of spatial learning, did have deficits during the probe test[77] and correspondingly affected genes not specific to brain function.

Since all three KMT cKOs are associated with a widely deregulated hippocampal transcriptome, which was accessed by the whole tissue CA-region RNA-seq, we have plotted the Venn-diagrams of down- and upregulated genes (see Fig.12B,C). Even though Kmt2a and Kmt2b are orthologues closely related to each other, we see that genes deregulated upon the respective loss are very different and barely overlap, which is the same for Setd1b cKO data (see Table 15). The only transcript differentially expressed for all three enzymes was the Oxidation Resistance 1 gene (Oxr1).

Group	Total	Downregulated genes	Total	Upregulated genes
Kmt2a/Setd1b	43	*Thsd4, Spink8, Serpina3n, Zbtb20, Glt8d2, Fam84b, Amigo2, Skida1, Nrp1, Synpr, Lefty1, Map4, Apln, Ppapdc2, Hdac9, Klhl13, Myo5b, Dnah9, Acan, Wfs1, Epha6, Strip2, Nr3c2, Gm2115, Dmd, Cbfb, Gabra5, Mctp1, Lix1, Aqp4, Ccbe1, Ugt8a, Fgfr1, Dock11, Ednrb, Robo1, Zdhhc23, Bcl6, Tmem229a, Zeb2, Nbl1, Sema3e, Oxr1*	1	*Grasp*
Kmt2b/Setd1b	19	*Kl, Col11a1, Grem1, Drd1a, Tmem255a, Gpr165, Gdf10, Tusc1, Rab27a, Marcksl1, Gm12462, Hcrtr1, Doc2a, Tuft1, Pmaip1, Rimklb, Slc29a4, 1110032F04Rik, Oxr1*	3	*Obscn, Gm3693, 8030453O22Rik*
Kmt2a/Kmt2b/Setd1b	1	*Oxr1*	0	

Table 15: Overlapping differentially expressed genes for *Kmt2a*, *Kmt2b* and *Setd1b* cKO (whole tissue RNA-seq of CA region, FC>1.2, FDR<0.1).
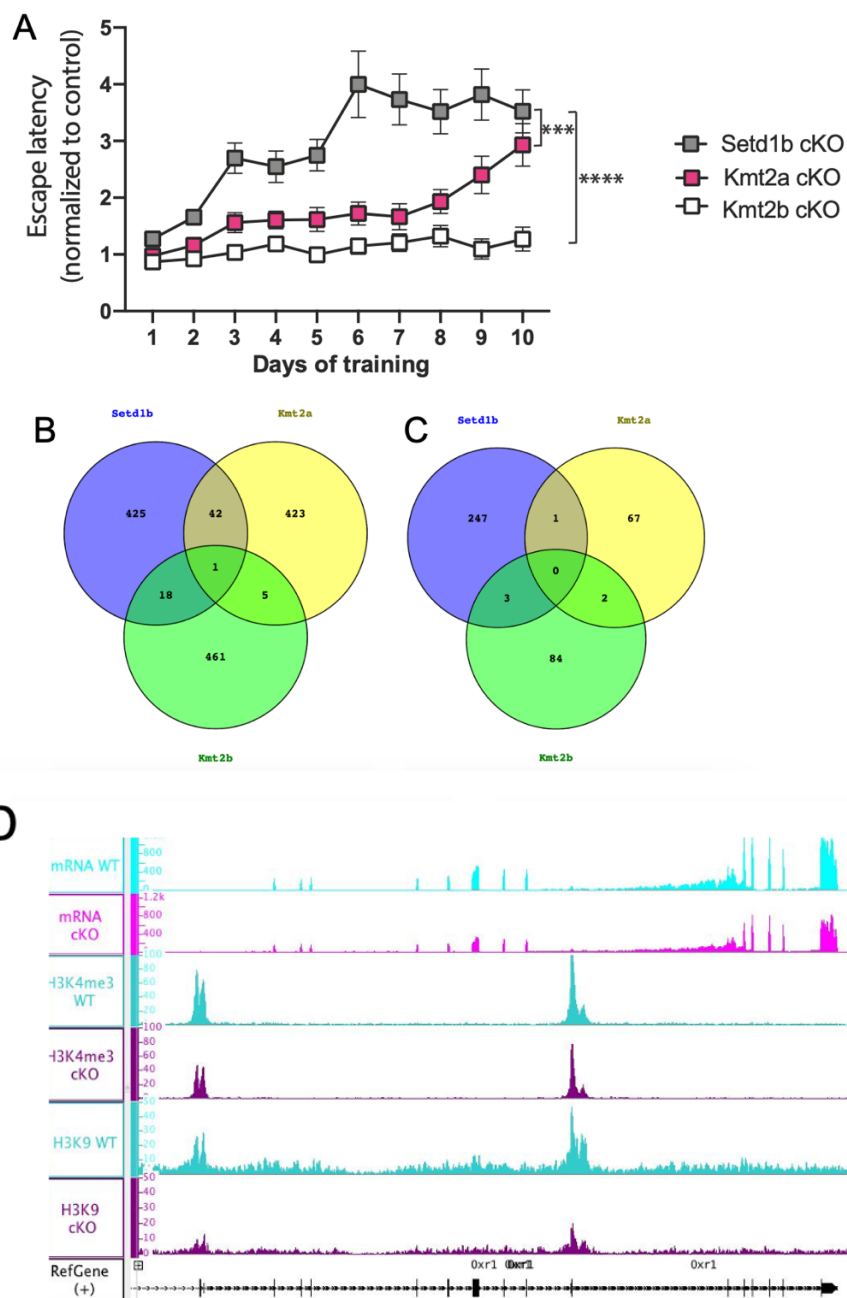
Fig. 12

Figure 2-12 Comparison of performance in MWM and gene expression of Kmt2a, Kmt2b and Setd1b cKO. (A) Escape latencies of Kmta, Kmt2b and Setd1b cKO mice in Morris water maze revealing differences in spatial learning between KMT knock-outs. Time required to reach the platform for KO animals was normalized to corresponding control group. Increase of normalized escape latency indicates shows the difference between cKO and corresponding control group, thus showing that Setd1b cKO have the poorest performance in spatial learning. (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). Error bars indicate SEM. (B) Genes downregulated and upregulated (C) in the whole-tissue CA-region RNA-seq of Kmta, Kmt2b and Setd1b cKO, fold change>1.2, padj<0.1. (D) Decreased expression in NeuN+ RNA-seq and decreased H3K4me3 and H3K9ac peaks of the only overlapping gene between all three cKO, *Oxr1* (IGB profile).

The role in neuronal functions of the closest homologue of Setd1b, Setd1a, has been investigated due to the connection of its loss of function with schizophrenia[176]. In the study a transcriptomic dataset from the PFC region of a Setd1a heterozygous loss-of-function mouse model was analyzed. Mice did not exhibit any abnormalities in the open field test, novel object recognition, fear conditioning and social memory, but had impairment in the working memory according to a T-maze delayed non-match to place. We found that only 10 genes were overlapping between the two Setd1 cKO models, and Oxr1 was not found among them (see Supp.Fig.9). We therefore conclude that the Setd1 orthologues regulate distinct sets of target genes.

Mice with heterozygous mutations in the Kmt2d have deficits in novel object recognition, probe test of Morris water maze and contextual fear conditioning, indicating hippocampal memory disfunction[89]. They also had deficiency in the bulk levels of H4K4me3 in the hippocampus and reduced neurogenesis levels. There are no studies showing the role of Kmt2c in the mouse brain and corresponding phenotype, even though a loss of function in humans is associated with intellectual disability and the Kleefstra syndrome[87]. Taken together, we show for the first time a transciptopathy and mental retardation associated with the loss of Setd1b in the adult murine brain, which has not been previously described in the mouse model. Our data and the published literature indicate KMTs to play non-redundant roles in the gene expression and behavior.

4.8.   The HDAC inhibitor SAHA does not restore cognitive functions in Setd1b cKO mice

It has been previously reported, that the inhibition of HDAC enzymes can improve learning and memory formation in cognitively impaired mice[102]. As we also observed a loss of histone acetylation in the Setd1b cKO mice, we have tried to rescue their cognitive performance in behavioral experiments by the application of the HDAC inhibitor

Vorinostat[111,112] into the animal drinking water. We have treated four groups of animals in total: WT and cKO receiving SAHA, and WT and cKO treated with vehicles as corresponding controls. After four weeks of treatment, Setd1b cKO mice were subjected to the MWM behavioral experiment (see Fig.13). Interestingly, after 5 days of training we observed the improvement of learning speed among the SAHA treated cKO group compared to vehicle treated (see Fig.13A), therefore, we performed the first probe test. We clearly saw the trend of increased number of visits to the control region and to the target quadrant, as well as in the % of time spent in those areas (see Fig.13B-E) among SAHA treated cKO group, but none of them we significant according to Student's t-test. Surprisingly, after conducting the probe test, improvement of learning among SAHA treated cKO mice was diminished (see Fig.13A). After 4 more days of learning, we subjected animals to the second PT. None of the read-outs for memory retrieval showed anymore difference between SAHA and vehicle treated cKO animals and the trend we observed during the PT1 was also gone (see Fig.13F-I). Search strategies used by two groups of cKO did not reveal any conclusive insights about the performance improvement (see Fig.13K).
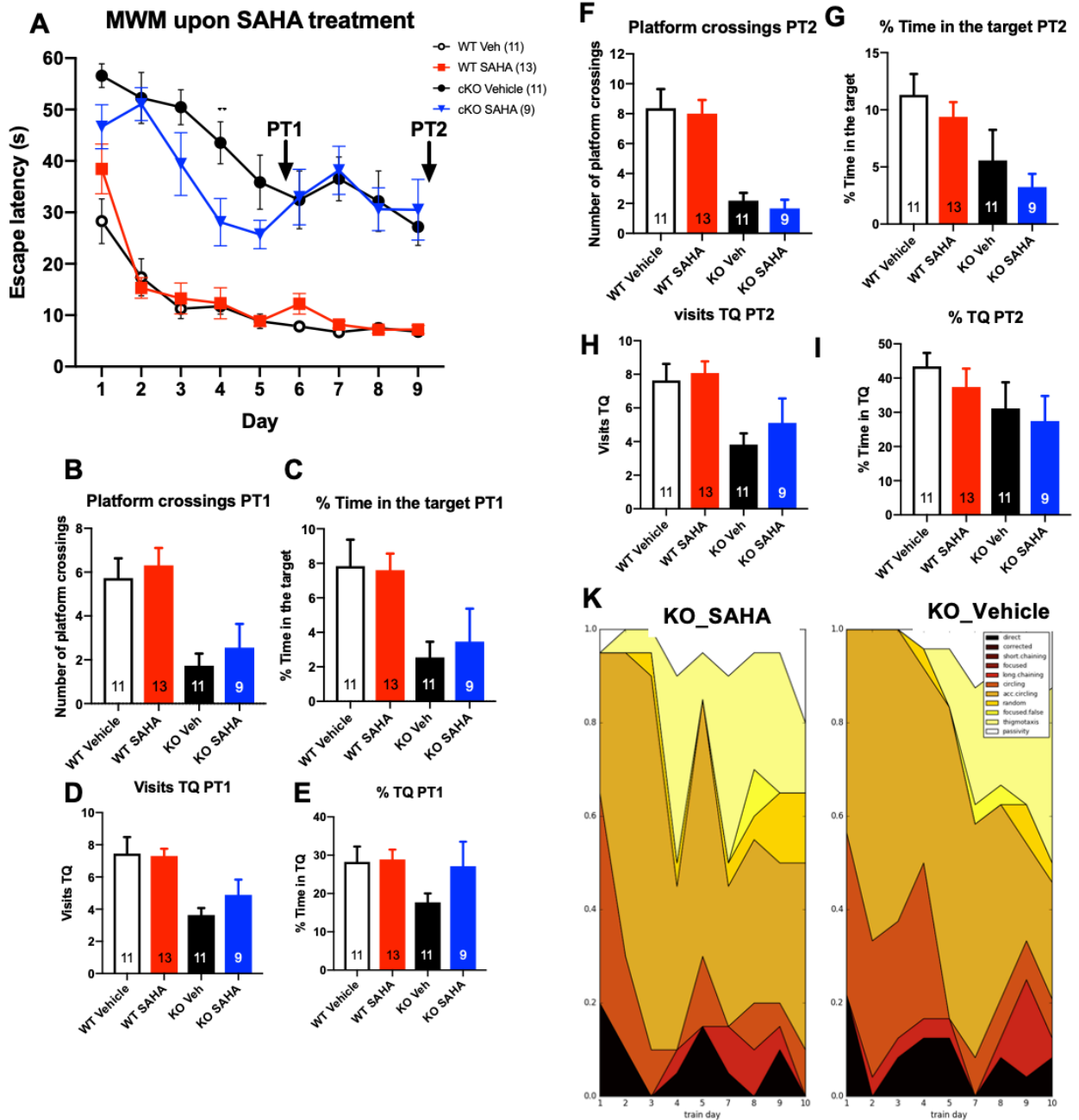
Figure 2-13 SAHA effect on the spatial memory impairment of Setd1b cKO in Morris water maze. (A) Escape latency during MWM (two-way-ANOVA, treatment effect for KO_S vs KO_V: F $(1,18)=1.965$, p-value=0.1780). (B) Visits to target zone on PT1 (p-value=0.4817). (C) % of time spent in the target zone on the PT1 (p-value=0.6524). (D) Visits to target quadrant on PT1 (p-value=0.2171). (E) % Time spent in the target quadrant during the PT1 (p-value=0.1513). (F) Visits to target zone on PT2 (p-value=0.5151). (G) % of time spent in the target zone on the PT2 (p-value=0.4684). (H) Visits to target quadrant on PT2 (p-value=0.4001). (I) % Time spent in the target quadrant during the PT2 (p-value=0.7352). (K) Search strategies of cKO animals treated with SAHA (n=5) vs vehicle (n=6). For all graphs WT_SAHA: n=13, cKO_SAHA: n=9; WT_Vehicle: n=11; cKO_Vehicle n=11 and Student's t-test p-values are indicated for cKO_SAHA vs cKO_Vehicle unless other is indicated. Error bars indicate SEM.

4.9. Mouse model of IGFBP7 liver overexpression

To evaluate whether increased blood levels of IGFBP7 would affect cognition and cause AD-like phenotype, we have crossed mice containing with stop-codon surrounded by loxP sites in front extra IGFBP7 allele in Rosa26 locus with animals expressing cre-recombinase under liver-specific albumin promoter (see Fig.14A). Prerequisite of this study was the data generated in our group showing increased levels of IGFBP7 in the plasma of AD patients (see Fig.14B). Mice did not differ in their weight (see Fig.14C) or physical appearance (data not shown). We have observed 3-fold increase of IGFBP7 mRNA levels in liver tissue, verifying the validity of our model, while hippocampal levels in CA and DG were not affected (see Fig.14D,E,F). We have tested two different antibodies to evaluate protein levels in the liver, but they were not specific (data not shown).
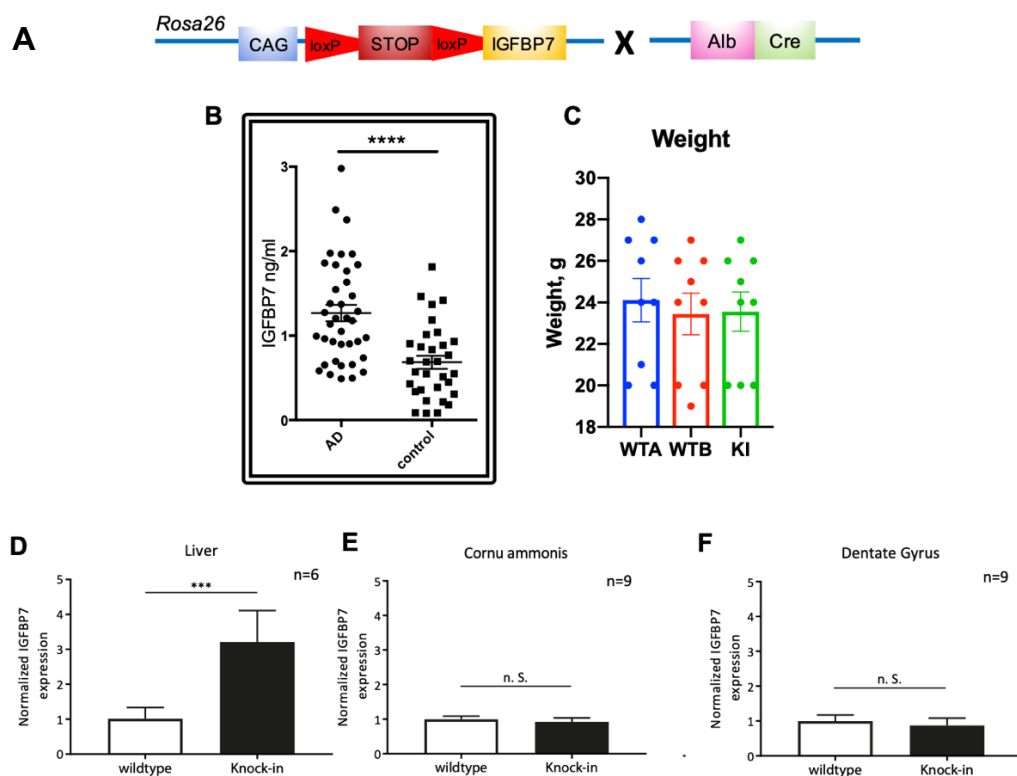
Fig. 14

Figure 2-14 Modelling human plasma elevation of IGFBP7 in mice. (A) Principle of generation liver-originating IGFBP7 overexpression mouse line. (B) Elevated concentrations of IGFBP7 in plasma of AD patients. (C) Weight of IGFBP7-albCre mice is not changed (WTA: n=9, WTB: n=9; KI: n=9; Student t-test). (D) mRNA level of IGFBP7 is decreased in the liver tissue of KI mice but not in the wild-types (WTA) (WT: n=6, KI: n=6; Student t-test: *** p-value < 0.001). (E) mRNA levels of IGFBP7 do not differ in CA region and (F) DG region of animals with liver overexpression of IGFBP7 (WT: n=6, KI: n=6; Student t-test: p-value=0.5943 and 0.6434 accordingly). Error bars indicate SEM.

5. Peripheral increase of IGFBP7 does not cause memory deficits

To test the behavioral effects of IGFBP7 overexpression originating from liver, we subjected three months old animals to the standard behavioral battery. In order to exclude the effect of genetic constructions of two lines used to generate knock-in line, we subject to testing all three groups, WTA (mice with cre-recombinase under albumin promoter in liver), WTB (mice with inactivated IGFBP7 allele in ROSA26 locus) and KI (extra IGFBP7 gene is activated in ROSA26 locus in albumin expressing liver cells).

Groups did not differ in the speed and distance travelled in the **open field test**, indicating the absence of major motor disability (see Fig.15A,B). Explorative behavior was similar between WT groups and KI based on the time spent in the center of the open field (see Fig.15C).

We did not observe genotype effect in the **elevated plus maze**, aimed at recognizing anxiety levels and exploratory behavior (see Fig.15D,E). Both groups spent comparable amount of time in the closed arms or center region. Subsequently, no difference was observed in the time spent in open arms.

Novel object recognition paradigm was used to evaluate deficits in long-term recognition memory. All three groups spent comparable amount of time near new object in the novel object recognition test (see Fig.15F).

Since Alzheimer's disease is characterized by gradual loss of memory, we have accessed spatial memory performance in Morris water maze test, highly dependent on hippocampus. Many AD mouse models exhibit memory impairment at the different age, including spatial memory impairment[177,178,179]. All three groups reached the learning plateau on the 4th and 5th day, indicating fast memory acquisition (see Fig. 16A), therefore, we conducted the first probe test after 5th day. No significant differences between groups was detected neither in the numbers of visits to the removed platform area and quadrant, where the platform used to be located (see Fig.16C,D), nor in the amount of time they spent in the

area of platform (see Fig.16E,F). Same was observed for the second PT conducted after three additional days of learning. Groups did not show any difference in the swim speed on the very first trying (see Fig.16K), additionally proving the absence of major motor impairment. Strategies used for all three groups to find a hidden platform we similar and progressed with the same trend for WTs and KI (see Fig.16B)

In order to see whether cognitive phenotype develops negatively with a time, we tested the animals three months later, when they were six months old. We used the reversal learning concept in MWM, when the location of hidden platform is changed comparing to the former MWM testing. Already on the 1st day all three groups were identifying platform location within 20 seconds (see Fig.17A) and during PT conducted after 5 days of learning we did not observe any abnormal memory retrieval phenotype among tested animals (see Fig.17B-E).
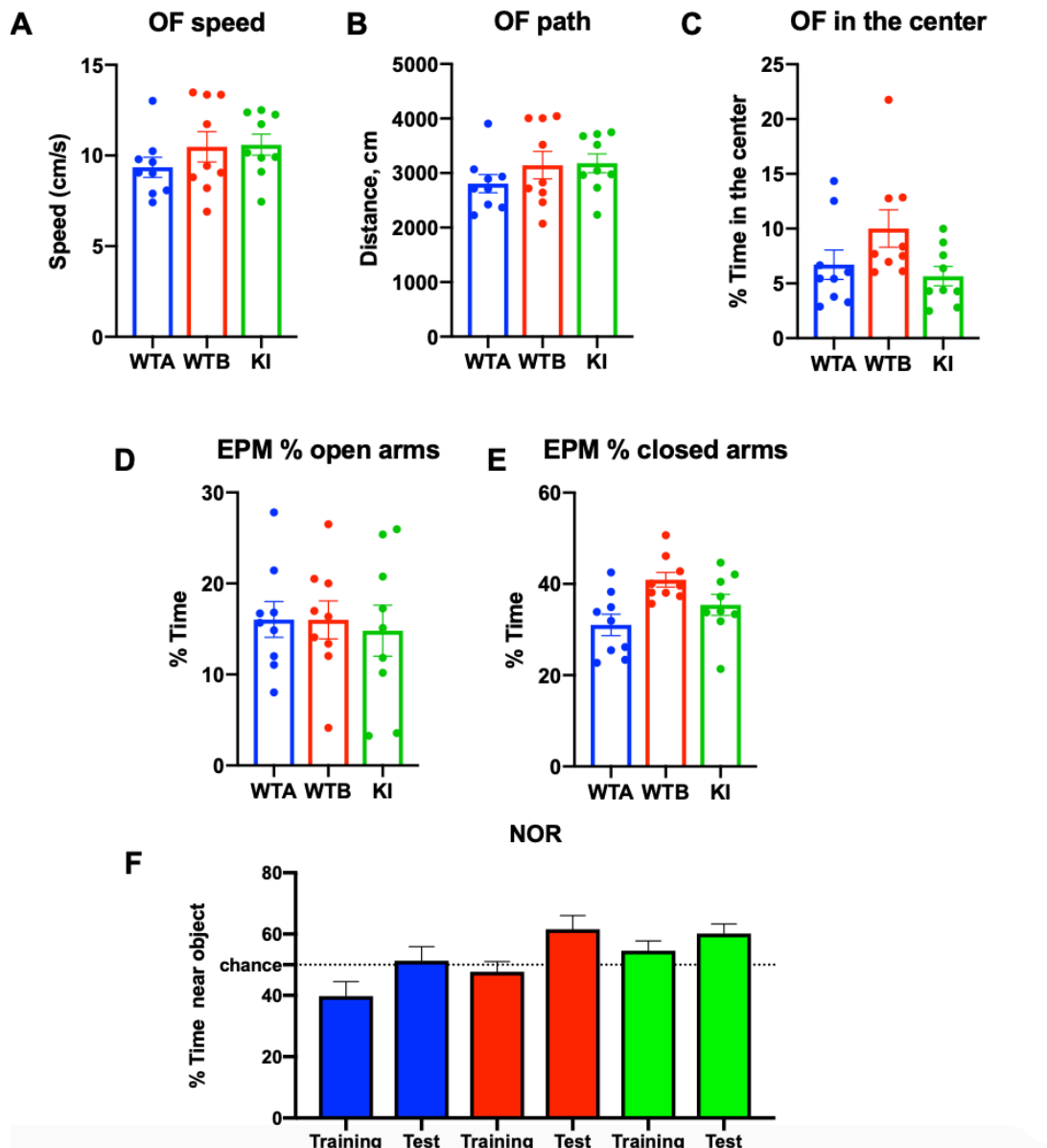
Fig. 15

Figure 2-15 **Increase of liver-originated IGFBP7 does not affect basic behaviors in mice.** (A) Speed in the open field, (B) total distance travelled and (C) % of time in the center did not differ between the groups. (D) % of time spent in the open arms as well as closed arms (E) of elevated plus maze was not affected by KI of IGFBP7. (F) Novel object recognition test did not reveal any differences between the groups. In all experiments WTA: n=9, WTB: n=9; KI: n=9; Student t-test; Error bars indicate SEM.
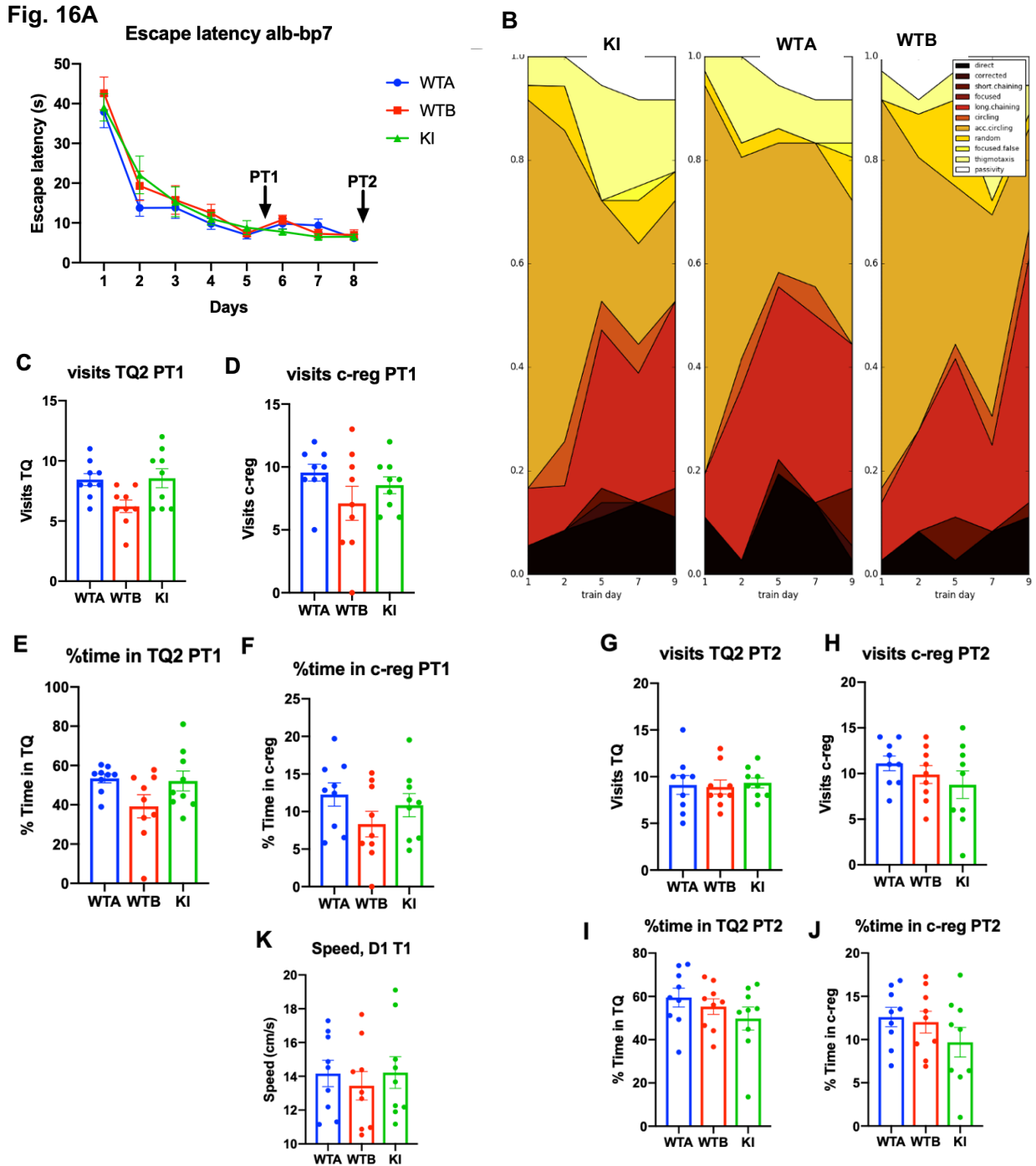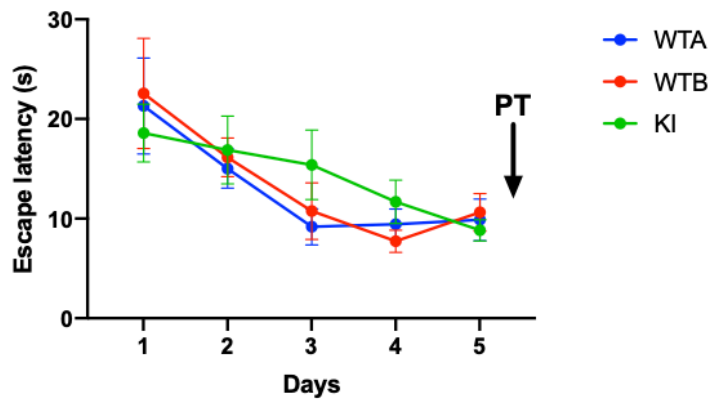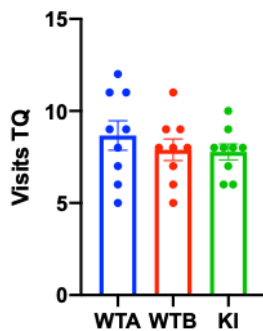
Figure 2-16 Spatial memory is not impaired in 3 months old mice with IGFBP7 overexpression in Morris water maze test. (A) Platform learning curves were similar between KI and WT groups (two-way-ANOVA, genotype effect: WTA vs KI $F_{(1,16)}=1.19$; p-value=0.2897; WTB vs KI $F_{(1,16)}=0.16$; p-value=0.6887). (B) We did not observe any differences between the strategy selected and distribution of strategies corresponded to normal hippocampal functions. (C) First probe test performed after day 5 of training did not reveal any changes in the number of visits in target quadrant (B), in platform area (C) as well as % of time spent in both of those areas (E, F). Same trend we observed for the second probe test, performed after day 8 of training (G, H, I, J). (K) Swimming speed on the very first trial of the first day of learning was the same among groups. In all experiments WTA: n=9, WTB: n=9; KI: n=9; Student t-test if not indicated; Error bars indicate SEM.
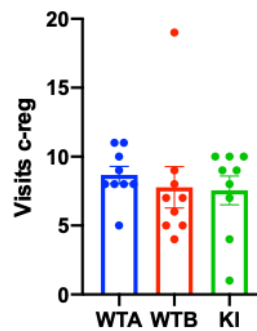
Fig. 17A
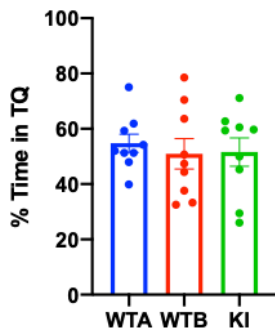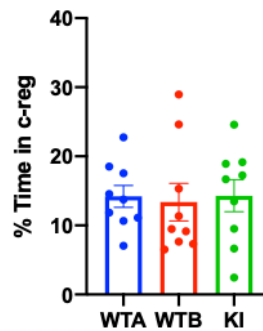
Figure 2-17 Spatial memory is not impaired in 6 months old mice with IGFBP7 overexpression in longitudinal Morris water maze test. (A) Escape latencies did not differ between KI and both WT groups (two-way-ANOVA, genotype effect: WTA vs KI $F_{(1,16)}=0.39$; p-value=0.5435; WTB vs KI $F_{(1,16)}=0.11$; p-value=0.7354). (B) Neither number of visits in target quadrant, nor in platform area (C) as well as % of time spent in both of those areas (D, E) had differences. In all experiments WTA: n=9, WTB: n=9; KI: n=9; Student t-test if not indicated; Error bars indicate SEM.

6. Discussion

In this work, we have shown that Setd1b loss of function in excitatory neurons of the mouse forebrain results in severe learning and memory formation impairments and few other behavioral deficits. Subsequent molecular workup with RNA-seq and ChIP-seq approaches revealed, that Setd1b indeed regulates highly expressed genes for neuronal identity and cognitive functions on the epigenetic and on gene expression levels, in a distinct manner to the other H3K4 methyltransferases Kmt2a and Kmt2b.

6.1. Benefits and drawbacks of the Setd1b-CaMKII-Cre mouse model

One obvious benefit of the Setd1b-CaMKII-Cre mouse is its completion of embryonic development and survival in adulthood, as Setd1b unconditional knockout mice die after gastrulation[122] and conditional knockout mice die after 30 weeks[2]. This Setd1b induced lethality was successfully circumvented by creating the Setd1b-CaMKII-Cre mouse with a Cre-controlled Setd1b[122] and a Cre-recombinase restricted to the mouse forebrain[123]. Furthermore, cognitive impairments cannot be attributed to developmental alterations, as Setd1b is only deleted 3 weeks after the animals were born, and no structural brain abnormalities have been observed (see Fig.2C,D). The results of our behavioral experiments support the observation that Setd1b mutations in humans are correlated with intellectual disability[4,124,125]. Phenotypically, the Setd1b cKO mice appear to be slightly smaller, with a reduced body and brain weight, when compared to the control animals (see Fig.2A,B). This has to be taken into consideration when interpreting the behavioral results, as the Setd1b cKO mice might be in a reduced physiological state, maybe due to reduced food intake, making it more difficult for them to feed, nest, swim or hold on the rod. Due to these concerns, we wanted to reduce the area in the mouse brain, in which Setd1b is inactivated, to the DG region. For this endeavor, using stereo-tactical injections,

we have injected Adeno-Associated-Virus (AAV) bearing Cre-GFP into the hippocampus of Setd1b mice. In the fluorescence microscopy, parts of the hippocampus and the DG region were indeed GFP positive (see Supp.Fig.5A,B). However, qPCR analysis revealed, that Cre and GFP mRNA levels were indeed upregulated, but Setd1b mRNA levels remained unchanged (see Supp.Fig.5C,D). Thus, the injections were negative and injected mice were not subjected to behavioral experiments. This was a puzzling result for us, but it could be explained by the findings that we made with the single cell sequencing approach, which showed that Setd1b is only expressed in a subset of neurons. If the injected virus did not transfect into these Sedt1b positive cells, the total Setd1b mRNA in the qPCR analysis might remain unchanged. Our experimental approaches must therefore be designed similar to the Setd1b cKO mouse model, and only target the subset of Setd1b expressing neurons, to obtain a similar phenotype. And finally, as our pharmacological intervention with the HDAC inhibitor SAHA did not rescue the cognitive impairment phenotype in Setd1b cKO mice (see Fig.13), it can also be assumed, that the global upregulation of histone acetylation does not rescue the phenotype. It is therefore possible, that only the subset of Setd1b expressing neurons should be targeted by the drug, and a global upregulation of histone acetylation might actually impede the reconstitution of synaptic plasticity. The finding, that only a subset of neurons in the mouse hippocampus express Setd1b, is a crucial observation, which would have heavily influenced experimental decisions if it was discovered earlier in the project. Taken together, Setd1b expressing neurons might function as nucleating factors to promote neuronal identity and synaptic plasticity in the mouse forebrain.

6.2. Dramatic deficits in nesting behavior of Setd1b cKO

Nesting is an evolutionarily ancient behavior and is common for many animals, including fish, birds, rodents and apes. Nests provide to mammals a secure environment, maintaining the body thermoregulation, protection from predators and are crucial for

parental behavior[180]. The inability to build nests is found to be correlated with impaired social behavior[181]. Normally, a piece of cotton introduced to a mouse cage is torn by the animal and piled up to form a single structure. The striking observation we made with Setd1b cKO mice was that they do not build nests even when supplied with good nesting material (see Fig.5A,B). Interestingly, not only did Setd1b cKO mice not build functional nests, but the nesting material remained mostly untouched. One could think of impaired motoric in Setd1b cKO affecting the nesting process, since it requires physical performance to rip a nestlet. Our results of other tests, such as immobility time in tail suspension test, speed in the open field, elevated plus maze and Morris water maze are arguing against it, since cKO and WT animals did not differ in performance and only motor learning was moderately decreased in rotarod test. We also cannot claim that animals failed to build nests due to apathetic behavior, since they did not have signs of depression according to the tail suspension test. Kmt2a-CaMKIIa-Cre cKO mice also show impairments in nesting[174], indicating that H3K4 trimethylation in the mouse forebrain seems to play an important role in nesting behavior. We have attempted to rescue the phenotype with the HDAC inhibitor SAHA[111], as HDAC inhibition has been shown to rescue Kmt2d deficiency induced learning impairments[89]. H3K9ac was significantly decreased in Setd1b cKO cells (see Supp.Fig.6), however, we did not see a rescue of nesting behavior in SAHA treated Setd1b cKO mice.

6.2.1. Neurological disorders and impairment of nesting behavior

Impaired nesting has been reported in multiple mouse models mimicking neurodevelopmental, neurodegeneration and neuropsychiatric disorders of genetic origin. For instance, the Fragile X Mental Retardation 1 (FMR1) gene in mice, linked to **Fragile X syndrome** in humans and characterized by mental retardation and autism, encoding the translational repressor protein FMRP also results in impaired nesting[182].

Rett syndrome is a neuropsychiatric disorder associated with sensory, cognitive and motor deficits and derives from gain of function or loss of function of methyl CpG-binding protein 2 (Mecp2). Mice with truncated Mecp2 allele show multiple symptomatic phenotypes, including compromised nest building, Morris water maze spatial learning difficulties[183] and mice with 95% knock-out of Mecp2 exhibit impaired nesting, while reactivation of Mecp2 in the CNS restores the phenotype[184].

Autism spectrum disorders (ASD) unite etiologically diverse neurodevelopmental conditions with multiple cognitive symptoms, such as compromised verbal and social communication, stereotyped behaviors and limited interests[245]. Many autism mouse models have shown the absence of nest-building behavior. Animals with mutated presynaptic cell adhesion protein neurexin1a (Nrxn1), phosphatase and tensin homologue (Pten), gamma-aminobutyric acid type A receptor subunit beta 3 (Gabrb3), tuberous sclerosis-1 (Tsc1) exhibit impaired nest building. All of these genes have been reported in the patients with autism[185]. Mutations in synaptic cell adhesion molecules neuroligin-1 (Nlgn1) and neuroligin-4 (Nlgn4), the most frequent monogenic cause of autism[186], have been documented in ASD, and mice with dysfunction of the corresponding protein are unable to build nests. Another autism-associated loss of function mutations in the SH3 and multiple ankyrin repeat domains 2 gene (Shank2) and SH3 and multiple ankyrin repeat domains 3 gene (Shank3)[187] also produce poor nest builders. ASD-related contactin associated protein-like2 (Cntnap2) knockout mice fail to make nests and the FDA approved antagonist of dopamine D2 receptor and serotonin 5-HT2a receptor risperidone was able to improve the nesting score[188]. Among the above-mentioned genes, we have observed significantly decreased H3K4me3 at the **Nrxn1**, **Pten**, **Shank2** and **Cntnap2** loci, but only **Nrxn1** was significantly downregulated in the whole tissue RNA-seq data and not in the neuronal RNA-seq (see Supp.Tabl.19).

Decline in nesting activity was suggested to mirror negative symptoms of **schizophrenia** and treatment of mice with the NMDA receptor antagonist phencyclidine (PCP, drug used to model schizophrenia) caused a significant nest building deficit[189]. Postnatal loss of glutamate ionotropic receptor NMDA subunit 1 (**Grin1**) from corticolimbic interneurons affects nest formation and results in other behaviors associated with schizophrenia, such as impaired PPI[190]. The gene has also been found to be affected among schizophrenia patients[191,192]. As previously mentioned, the Grin1 gene is responsible for striatal motor learning and a normal startle response amplitude, both of the behaviors are impaired in Setd1b cKO. In our dataset, we have discovered a reduction of Grin1 trimethylation levels, but this effect was not observed on the gene expression level. However, the gene expression levels of glutamate ionotropic receptor NMDA type subunit associated protein 1 (**Grina**) and glutamate ionotropic receptor NMDA type subunit 3A (**Grin3a**) are decreased in the NeuN+ RNA-seq dataset (see Supp.Tabl.19). Therefore, glutamatergic signaling could be compromised among Setd1b cKO animals.

A triple transgenic mouse model of the **Alzheimer's disease**, 3xTg-AD, mimicking familial AD mutations in PS1/M146V, AßAPPSwe and tauP301L showed age-dependent regression of nest building[193]. Offspring of hAPP and htau mouse strains, possessing the amyloid and tau features of AD, were also unable to build functional nests[194]. The same refers to the Tg2576 mouse model, which overexpresses the amyloid-beta precursor APP[195] and the APPswe/PS1 bigenic model of AD[196].

Hypothalamus is responsible for parental-related hormone release, such as **oxytocin** and **vasopressin**, and has been shown to be responsible for nest building behavior in mice[197,198]. We cannot exclude that the absence of nesting behavior among Setd1b cKO animals is due to deregulated gene expression within the hypothalamus region since the driving promotor of our line, CamKIIa, is still moderately expressed within the hypothalamic

region (Allen Brain Atlas). However, lesions of the hippocampus modeled with intrahippocampal injections of scarpie (prionic protein) resulted in worse nest building scores, whereas cytotoxic lesions of the prefrontal cortex did not result in a similar phenotype[156]. This might indicate a crucial role of the hippocampus in recognition of nest-like structures[181]. It was also shown that there are nest responsive CA1 hippocampal cells firing during mouse engagement with beds and nests[199].

6.2.2. Neurotransmitter metabolism and impairment of nesting behavior

In our dataset, the **serotonin receptor signaling pathway** appeared among one of the most downregulated pathways in the NeuN+ RNA-seq (see Table 16) and we could observe multiple serotonin receptor subunits to be strongly deregulated.

Gene name	Encoded protein	Log2FoldChange in NeuN+ RNA-seq(padj<0.1)	log2FoldChange in NeuN+ H3K4me3 ChIP-seq (padj<0.05)
Htr1a	5-Hydroxytryptamine receptor 1a	-1,3102	-0.39
Htr4	5-Hydroxytryptamine receptor 4	-0,3161	-2.17
Htr2a	5-Hydroxytryptamine receptor 2a	-0,6787	-0.71
Htr1b	5-Hydroxytryptamine receptor 1b	-0,5096	-3.6
Htr5b	5-Hydroxytryptamine receptor 5b	-1,0506	-0.46

Table 16: Genes from the downregulated GO-pathway "Serotonin receptor signaling pathway" from neuronal RNA-seq

Knockout of the 5-Hydroxytryptamine receptor 4 (**Htr4**) has previously been linked to the impaired nesting behavior[200]. The mouse line Ts65Dn, carrying three copies of murine chromosome 16, 17 and a partially homologous human chromosome 21, is mimicking Down syndrome, fails to build proper nests in a novel environment[201]. These mice had increased levels of Htr2a and serotonin in the frontal cortex, but not in the hippocampus. **Htr2a** blockade with risperidone, an antagonist of the serotonin receptor, caused restoration of

context-dependent nest building, while Htr2a activation disrupted it. **Htr1a** was one of the overlapping genes between the Setd1a KO model of schizophrenia[176] and our Setd1b cKO mice (see Supp.Fig.9A,B), serotonin signaling is therefore impaired in the both Setd1a and Setd1b.

Serotonin availability can be influenced not only by receptors, but also by serotonin producing enzymes and serotonin transporters. Tryptophan hydroxylase 1 (Tph1) and tryptophan hydroxylase 2 (Tph2) are encoding proteins responsible for the first and rate limiting step of serotonin biosynthesis; solute carrier family 6 member 4 (Slc6a4), is coding for a serotonin transporter. None of them were affected by Setd1b cKO on trimethylation or gene expression level. Slc6a4 and Tph2 have been associated with altered serotonin levels and polymorphisms in autism patients[202–204]. Tph2 null mutant female mice exhibit decreased maternal care skills and are poor at nest building[205]. The selective serotonin reuptake inhibitor (SSRI) fluoxetine provided to pregnant female mice showed that prenatal exposure to SSRI altered maternal nest building among the offspring[206] and a number of SSRI (clomipramine, citalopram, fluoxetine and venlafaxine) has inhibited nestlet shredding in a dose-dependent manner[207].

Even though it is widely accepted that nesting is an innate behavior, there are certain incidences when it can be disrupted in adult rodents. For instance, it has been shown that amphetamine, mescaline (can bind to Htr2a) and lysergic acid diethylamide can affect nest building in adult mice[208]. Setd1b cKO mice also exhibit changes in **dopamine signaling**. It has been reported that dopamine deficient mice lacking tyrosine hydroxylase (Th), encoding a tyrosine to dopamine conversion enzyme, have impaired nesting, but could be restored by restoration of dopamine production via viral delivery into the caudate putamen[209]. Dopamine receptor D2 (Drd2) mutant mice have decreased goal-oriented behavior and lack of motivation[210]. In our Setd1b cKO animals we discovered

downregulation of the dopamine receptor D1 (**Drd1**), indicating that Setd1b directly regulates Drd1 expression via promoter methylation (see Supp.Fig.10).

We suspect the nesting circuit, mediated by **oxytocin and vasopressin**, to be compromised in Setd1b mice. Oxytocin is a posterior pituitary peptide hormone and it is important for the induction of labor and breastfeeding in mammals[211]. It is also facilitating nesting behavior[212] and even can cause "false-nesting" (nesting behavior without egg laying) in turtles after intramuscular injections[213].

The three main components of oxytocin action are the oxytocin structural gene Oxt, the oxytocin receptor Oxtr and the CD38 antigen, which is responsible for oxytocin secretion. Since oxytocin plasma levels are lower in ASD patients[214], it is an interesting question to ask, whether Setd1b cKO mice exhibit decreased amounts of oxytocin. The **Oxtr** transcript was downregulated in our NeuN+ RNA-seq of Setd1b cKO mice (see Supp.Fig.10), an absence of sufficient receptor amounts consequently could have caused nesting behavioral impairment. CamKIIa cKO mice lacking Oxtr exhibit impaired social memory[215], the gene has also been reported to be involved in ASD[216]. **CD38** deficient mice also exhibit deficits in social behavior[217], and it was strongly downregulated in our NeuN+ RNA-seq data of the Setd1b cKO. Downstream oxytocin physiology involves MAPK signaling[218]. Interestingly, the positive regulation of **ERK1 and ERK2 cascade** was the most downregulated pathway from our Setd1b cKO neuronal RNA-seq (see Supp.Table 18), which might be an indication that oxytocin signaling is indeed disrupted. This goes together with mitogen-activated protein kinase 1 (Mapk1, old name Erk2) deficiency known to cause ASD-like phenotype in mice including impaired nest formation[219]. Taken together, we present strong evidence for deregulated oxytocin-related aberrant nesting behavior among Setd1b cKO mice.

Increased levels of vasopressin were shown to inhibit nest building, both on the genetic and pharmacological level[197]. The vasopressin gene (Avp) does not appear as differentially expressed in our data. However, the vasopressin 1b receptor (**Avpr1b**) was significantly downregulated in the whole tissue RNA-seq of the CA region. Avpr1b KO in mice results in the reduction of social recognition[220] and reduced ultrasonic vocalization in pups, abnormalities associated with the autistic spectrum disorders. There is no published data on nesting behavior upon Avpr1b loss in mice.

In conclusion, the nesting phenotype observed in our case can been seen in different mouse models of autism, mental retardations and intellectual disabilities, schizophrenia and the Alzheimer's disease. Nesting can be affected by genetic, biological and environmental factors. Aberrant serotonin and dopamine signaling, oxytocin and vasopressin regulation might be the underlying causes of the nesting impairment phenotype in Setd1b cKO mice.

6.3.   Setd1b exerts distinct functions from other H3K4 methyltransferases in the mouse forebrain

H3K4 methylation has been shown to increase during mouse behavioral training[76] and has also been found to be reduced in AD mouse models[221] and post mortem tissues of patient brains with cognitive disorders[222]. Similarly to our findings in Setd1b (see Fig.3), Kmt2a[78] and Kmt2b[77] forebrain cKO mice also exhibited impairment in hippocampal memory functions in behavioral experiments. Importantly, mice heterozygous for the Setd1b homologue Setd1a are not impaired in hippocampal memory formation, but rather exhibit working memory defects and schizophrenia-like phenotype[176], although a forebrain specific cKO could perform differently. Together with our ChIP and RNA sequencing data, we provide proof that the four H3K4 methyltransferases exert distinct functions in the mouse forebrain. We show differentially decreased H3K4me3 at TSS of distinct genes in Setd1b,

Kmt2a and Kmt2b cKO mice (see Fig.10A). Alterations of H3K4me3 levels seem to impact H3K9ac levels as well (see Fig.8C), this finding is supported by previous studies, which link H3K9ac downstream to H3K4me3[77,223]. One key finding of our screens was that a number of genes with decreased H3K4me3 in the Setd1b cKO samples in turn showed increased H3K4me1 levels at TSS (see Fig.10B), which was not observed for Kmt2a and Kmt2b cKO mouse samples. These genes were further investigated and turned out to exhibit a reduced H3K4me3 peak width when compared with the control (see Fig.10H). Furthermore, these genes are highly expressed at basal levels (see Fig.10I) and function in pathways promoting neuronal identity, synaptic plasticity, learning and memory functions (see Fig,10J). The association between reduced H3K4me3 peak width, enhanced H3K4me1 and high basal gene expression has been reported previously with the Setd1b yeast homologue Set1 . A possible explanation for the increased H3K4me1 levels around TSS in Setd1b regulated genes might be the subsequence of H3K4me1 to H3K4me3[171]. Together with our single cell sequencing data, which showed that only a subset of neurons expressed Setd1b (see Fig.11E) and that Setd1b specifically and strongly regulates a subset of genes in these cells (see Fig.11F), we hypothesize H3K4me1 to be the incompletely methylated H3K4me3 transitional state at Setd1b regulated TSS, as Setd1b specifically regulates these genes while having little overlap with the other KMTs. Therefore, the specific dependency on Setd1b will accumulate H3K4me1 at TSS without further methylation processing by other H3K4 methyltransferases, this situation might be compensated in the Kmt2a and Kmt2b cKO mice by other KMTs. A possible experiment to prove this hypothesis is a Setd1b knockdown or knockout in primary neuronal cells with subsequent ChIP analysis of H3K4me1 levels at TSS. This short term Setd1b loss of function should expose the H3K4me3 transient state H3K4me1, which could be rescued with Setd1b overexpression. Furthermore, it would be interesting, whether a Kmt2a and Kmt2b double cKO would also expose genes with increased H3K4me1 levels at TSS.

The discovery that Setd1b is only active in a subset of neurons is intriguing. For one, it suggests that NeuN+ sorted cells do not resemble a homogenous population of neurons (see Fig.7), there are subgroups of neuronal cells, which can be, in our case, distinguished by the activity of different H3K4 methyltransferases in the single cell sequencing data (see Fig.11F). It would therefore be interesting, to further characterize these subgroups by in depth analysis of the differences in gene expression. Furthermore, these different populations might play specific roles in the neuronal network and even have specific positioning in the architecture of the brain. Evidence for such heterogeneity in the neuronal population has been provided by other research groups utilizing the single cell sequencing technology. In one study, it was shown in single cell data that cells expressing Pvalb (Parvalbumin) were actually a distinct cell population among interneurons and that Vip (Vasoactive intestinal peptide) expression defines a subset of Calb2 (Calbindin 2) expressing cells[225]. The authors were able to further subdivide the Pvalb and Vip populations by differential gene expression in the single cell datasets. Furthermore, they identified a subpopulation of NeuN positive neurons which express Cplx3 (Complexin 3). The hitherto known group of neuronal cells is therefore highly heterogenous and we are just beginning to understand their diversity. Technologies such as single cell sequencing or sorted cell sequencing will rapidly push this frontier of understanding and categorizing these distinct neuronal cell populations. Large scale screens for such distinct neuronal subtypes of the brain have been conducted and need to be thoroughly evaluated[226], the same is true for neuronal disease datasets, such as from AD patients[227]. With our work, we provide evidence that in the group of NeuN+ cells Kmt2a, Kmt2b and Setd1b expression is differential (see Fig.11E,F). As epigenetic regulators such as H3K4 methyltransferases play a key role in regulating the overall cell fate and thus gene expression, their differential expression might play a key role in regulating the overall cellular organization and functions of the brain.

Another interesting question raised by the observation of the complex diversity of neurons is their positioning within the architecture of the neuronal network. Setd1b is expressed in a minority of NeuN+ cells, which is distinct to Kmt2a and Kmt2b (see Fig.11E,F). It would therefore not be plausible, if Setd1b cells would cluster in a distinct location in the hippocampus, it would rather have a dispersed distribution to promote neuronal fate and synaptic plasticity all over the brain structure. Therefore, it would be really interesting to investigate the positioning of different neurons in their smallest functional unit. Molecular mapping of neuronal distribution in the mouse brain has been conducted and yielded a really impressive database[228], but due to the sheer number of neurons in the mammalian brain, the data is insufficient to create a map in single neuron interaction resolution. It would therefore be feasible to investigate neuron positioning in the brain network in simpler organisms, such as the zebrafish embryo. The zebrafish *Danio rerio* is an established model organism in neuroscience[229], besides cheap housing and short generation times, it can be engineered to be transgenic and most importantly, it has a transparent embryo[230]. For the investigation of Setd1b in the zebrafish *in vivo*, one could genetically engineer a strain in which GFP is behind the Setd1b promoter. Imaging of neurogenesis and the adult brain could reveal the specific positioning of Setd1b cells within the fish brain. Setd1b is conserved from zebrafish to human and has already been shown in in situ hybridization experiments in the fish, in which a strong staining of the brain could be observed[4]. However, these results are unable to resolve the fish brain in a three-dimensional neuronal network with single cell interaction resolution. Therefore, a Setd1b::GFP transgenic fish might reveal the spatial positioning of Setd1b expressing cells within the neuronal network architecture. Introduction of further promoter driven fluorescent proteins could visualize the smallest interaction unit of Setd1b expressing cells with other neurons, for instance with Kmt2a or Kmt2b expressing cells. Such data could help understand the general function of Setd1b expressing neurons and H3K4 methyltransferase differential distribution in the vertebrate brain.

6.4. Spatial memory deficits in KMT cKOs and the Oxr1 gene

We have observed a spatial memory impairment with deficits in learning the hidden platform location and compromised memory retention on the probe test day in the MWM experiment with Setd1b cKO animals (see Fig.3). Spatial navigation is strongly dependent on hippocampal function[231], which appears to be compromised in Setd1b cKO according to our behavioral and sequencing data. Multiple origins of impaired spatial learning have been reported, including ASD[232,233], AD[234,235] and schizophrenia[236].

Kmt2a, Kmt2b and Setd1b cKO animals all exhibit impairment of spatial learning in the MWM experiment, and the analysis of sequencing data reveals only one common downregulated gene, **Oxr1** (see Fig.12). Oxr1 is known to protect against oxidative stress-induced damage, it was initially discovered in E.coli as a DNA oxidative stress protector[237]. It has also been shown to be an essential protector against oxidative stress in the mammalian central nervous system (CNS) and has been linked to cerebellar neurodegeneration[238]. Neuronal Oxr1 overexpression leads to motor improvement and increases lifespan in the amyotrophic lateral sclerosis (ALS) SODG93A mouse model[239]. Lentiviral overexpression of Oxr1 was able to ameliorate a Parkinson's disease phenotype in the corresponding mouse model, including trembling and rigidity of limbs. Depletion of miR-137, a negative regulator of Oxr1, reduced oxidative stress from neurons[240]. A very recent study identified the biallelic loss of Oxr1 in individual with intellectual disability, cerebellar atrophy, developmental delay and seizures[241]. Interestingly, Oxr1 transcripts are colocalized with CamKIIa transcripts and not with Glutamic Acid Decarboxylase (GAD), indicating expression only in excitatory neurons[242]. Interestingly, ROS are required for learning and memory, participating in their signal transduction cascades[243] and at the same time has been shown to be involved in pathological conditions, such as the Parkinson disease, AD, depression, schizophrenia and ASD[244]. Taken together, downregulation of

Oxr1 might have compromised the attenuated spatial learning and memory retrieval in MWM of Kmt2a, Kmt2b and Setd1b cKO mice through an imbalance of oxidation levels.

6.5. Setd1b regulates the claustrophobia-mediating gene Gpm6a

First, we had difficulties interpreting the data of fear conditioning test, since such behavior has been never observed before in our laboratory. While trying to find any explanation for such behavior we came across the paper dedicated to the study of claustrophobia, the fear of being trapped in the closed spaces. Researchers for the first time discovered that mice lacking the neuronal tetraspan membrane glycoprotein (Gpm6a) gene exhibit "claustrophobia-like" phenotype, including avoidance of the closed arms in EPM[245]. We notice the same trend among our cKO animals (see Fig.4A,B). Gpm6a were selecting wide areas in the wide/narrow box and avoiding dark spaces in the light/dark box. This is not very natural for rodents, who usually prefer to hide from predators in the small dark places. Researches also confirmed mouse data by analyzing human samples and finding rare sequence variants in the Gpm6a. But the most striking, however, was their observation that Gpm6a mutants had higher baseline freezing in the fear conditioning chamber, same as Setd1b cKO. Gpm6a mutants had increased basal levels of corticosterone in urine, while in our knock-out animals we did not observe elevated levels of the stress hormone in plasma (see Supp.Fig.4).

Trimethylation of H3K4 as well as H3K9 acetylation of **Gpm6a** in the hippocampi of Setd1b cKO is significantly decreased and it is also strongly deregulated on mRNA level in both, whole tissue and neuronal RNA-seq from our data (see Supp.Fig.11). This indicates that Setd1b can be direct regulator of Gpm6a by mediating promoter methylation and, therefore, affecting the expression levels. We cannot be sure, whether the Gpm6a function is hippocampus-specific, but even for the case of other brain regions we can assume that Setd1b regulates the expression of Gpm6a in those areas.

## 6.6. SAHA is not efficient to compensate methylation losses

Even though the treatment of neurodevelopmental disorders is majorly performed via behavioral therapy and symptoms management, agents that promote chromatin opening can have therapeutic potential and histone deacetylase inhibitors are one of them. Oral administration of HDACi AR-42 was able to rescue Morris water maze performance and H3K4me3 levels in Kabuki syndrome mouse model with the absence of Kmt2d allele[89]. Water administration of SAHA to Setd1b cKO mice did not reverse nest building behavior and did not improve memory retrieval during probe tests of MWM, however we observed a trend on the first probe test as well as in the speed of learning for before the first probe test among SAHA treated cKO animals. It might be that SAHA was indeed improving learning and memory for Setd1b cKO mice but the total number of animals was not sufficient to make proper statistical conclusions. It is further possible that the oral SAHA delivery was not successful. Also, the pharmacological rescue of the Setd1b cognitive impairment phenotype might not be easily achieved, as the specific targeting of Setd1b expressing neurons is difficult to obtain with systematically applied HDAC inhibitors.

6.7. IGFBP7 overexpression in the murine liver does not impact learning and memory formation in transgenic animals

Previous work in our lab has identified elevated levels of IGFBP7 in the blood serum of AD patients (see Fig.14B), which gave rise to the hypothesis, that artificial enhancement of IGFBP7 in the mouse could lead to the impairment of cognitive functions in behavioral experiments. For this endeavor, we have designed a transgenic mouse, that overexpresses the IGFBP7 protein in the liver, which has been verified by qPCR analysis (see Fig.14D). However, these transgenic animals did not show cognitive impairments or memory deficits in the behavioral studies when compared to wildtype animals, prompting the conclusion, that overexpression of IGFBP7 is not sufficient to induce a cognitive impairment

phenotype in mice. We are currently validating in cooperation with DZNE Bonn, whether our transgenic animals do indeed possess elevated plasma levels of IGFBP7. The result will determine, whether the transgenic model is to be questioned or whether IGFBP7 plasma levels is a secondary effect of AD progression and not the cause.
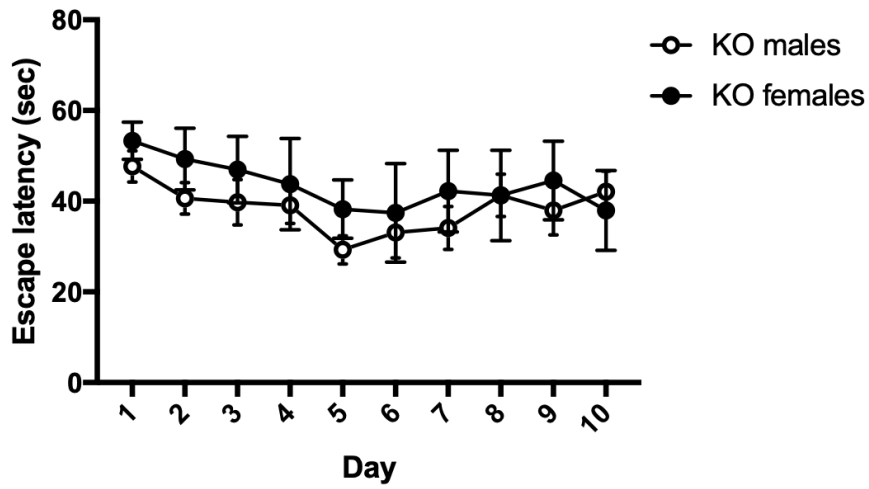
7. Supplementary
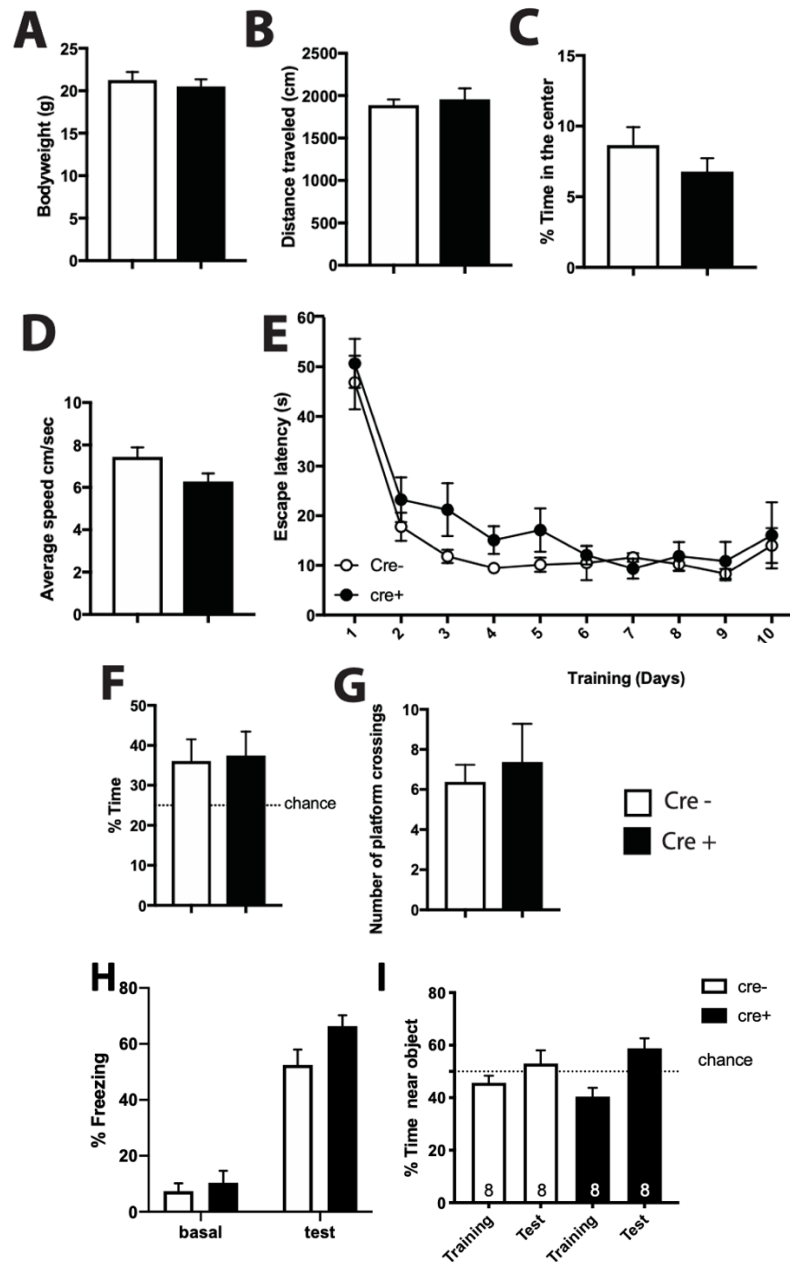
Supplementary Figure 1. (A) Behavioral analysis of mice expressing CamKII-driven Cre recombinase. Body weight did not differ between cre+ and cre- mice. (B) The distance traveled in the open filed test and (C) the time spent in the center of OF was comparable among the groups. (D) In the Morris water maze mice had the same speed on the first trial of the first day. (E) Escape latency was similar between cre+ and cre- mice and both grouped learned the platform location. (F) % of time in the target quadrant and (G) the number of platform crossings did not differ among two groups at the probe test conducted after 10 days of training. (H) In contextual fear conditioning test cre+ mice did not exhibit freezing before foot-shock and had comparable time freezing after electric shock exposure (two-way ANOVA, p-value=0.1145). (I) Cre+ mice spend the same amount of time near new object in novel object recognition test. 8 mice of each group were tested, Student's t-test was used unless other test is indicated, error bars indicate SEM.
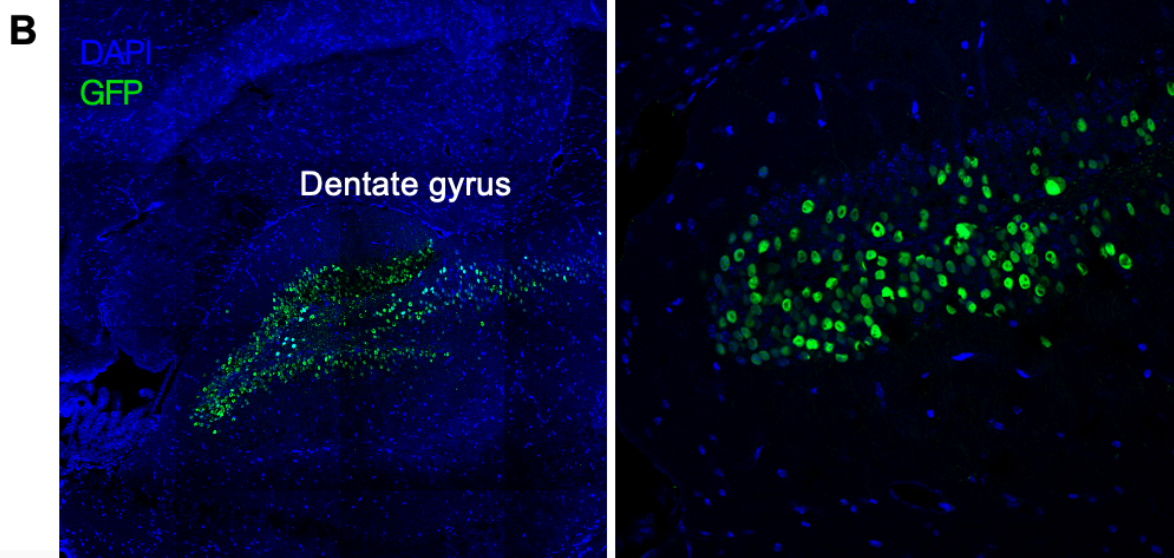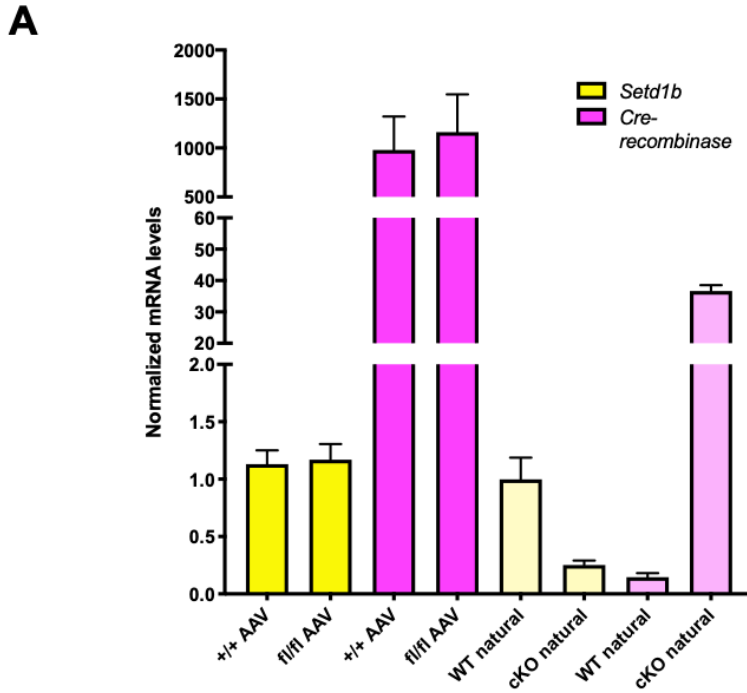
Supplementary Figure 2. Attempt to initiate AAV-mediated DG-specific Setd1b loss did not succeed. (A) Analyzing expression of Setd1b in viral-injected DG measured with qPCR we did not observe the loss of Setd1b in AAV-treated Setd1b$^{fl/fl}$ samples comparing to cKO samples. Cre-recombinase expression upon virus delivery was sufficient. For convenience, data on Setd1b and cre-recombinase expression in cKO and WT mice used for this study is also presented. (+/+ AAV, n=10; fl/fl AAV, n=10; WT, n=8; cKO, n=8; Student's t-test, error bars indicate SEM). (B) Immunofluorescent staining of GFP signal within AAV-cre virus construct delivered to DG region of Setd1b$^{fl/fl}$ mice.

Morris Water Maze: gender difference

Supplementary Figure 3. Spatial memory impairment in Morris water maze test is not gender-related. Escape latency during MWM (males: n=10, females: n=5; two-way-ANOVA, gender effect: F (1, 13)=0.7701; p-value=0.3951).

plasma corticosterone
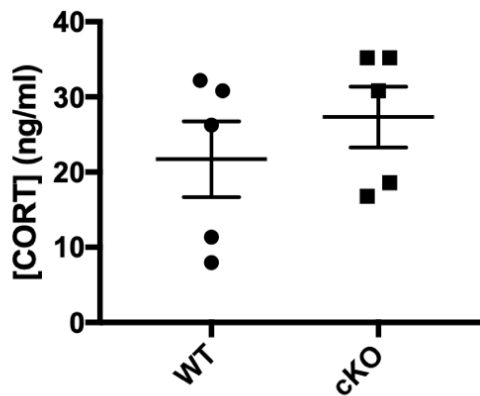
Supplementary Figure 4. Stress hormone corticosterone is not elevated in the blood of Setd1b cKO mice. 5 mice of each group were tested, Student's t-test, error bars indicate SEM.
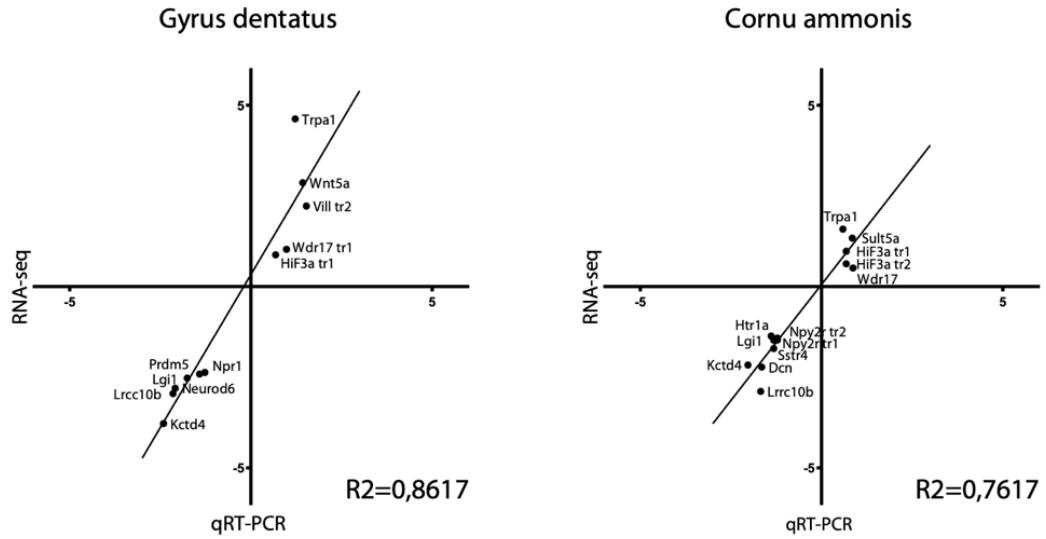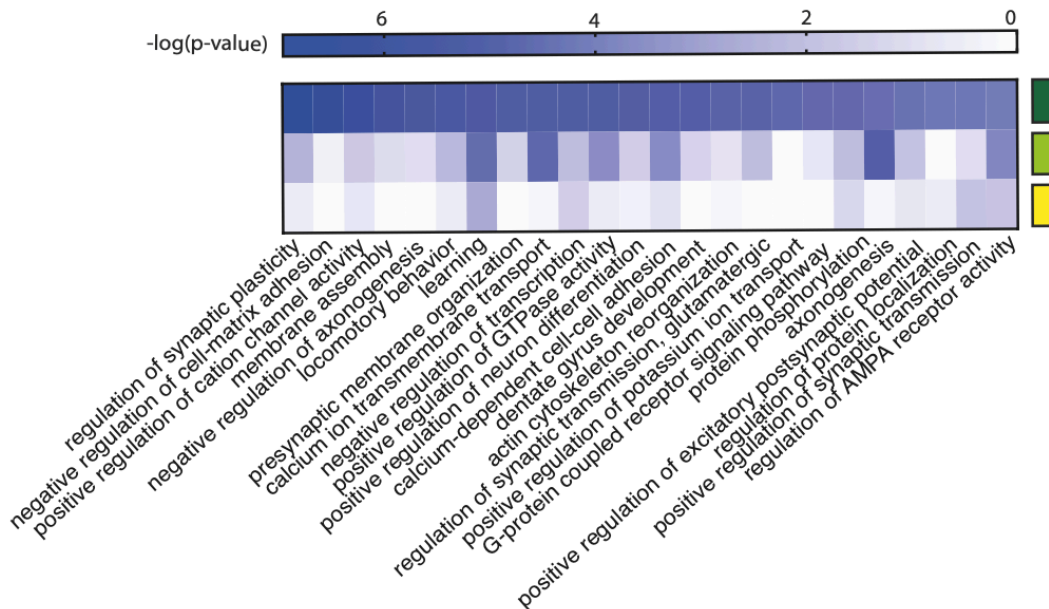
Supplementary Figure 5. The most downregulated genes from whole-tissue DG (left) and CA region (right) of Setd1b cKO mice RNA-seq correlate with mRNA expression levels validated with qPCR.
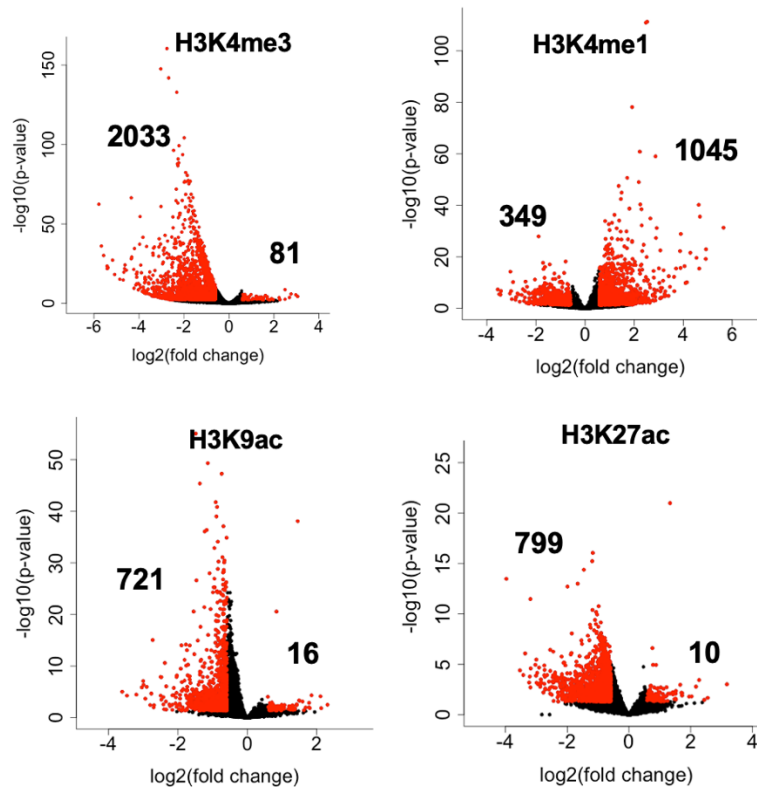

Supplementary Figure 6. Total number of differentially methylated and acetylated regions in H3K4me3, H3K4me1, H3K9ac and H3K27ac ChIP-seq. |Fold change|>1.5, padj<0.05.
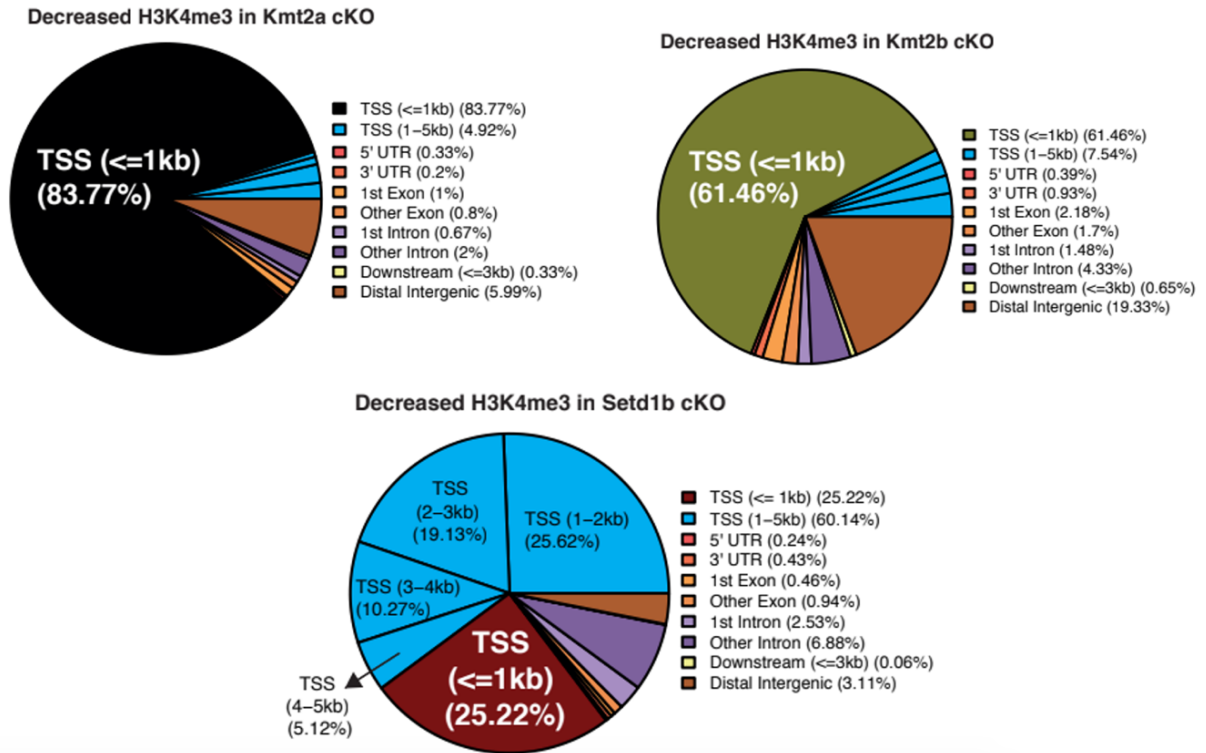
Supplementary Figure 7. Genome-wide distribution statistics of H3K4me3 in Kmt2a, Kmt2b and Setd1b cKO. Setd1b, unlike Kmt2a and Kmt2b, regulates trimethylation of H3K4 in distal promoter regions.


Supplementary figure 8. Heat map showing functional pathways for the 3 categories of genes affected by reduced H3K4me3 in Setd1b cKO mice.
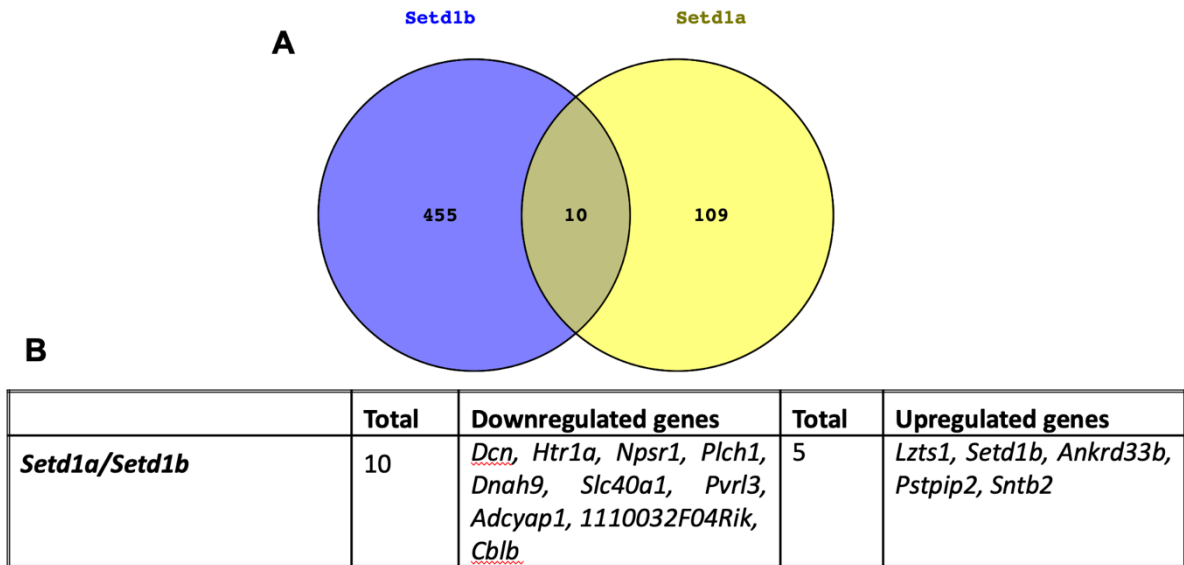
Supplementary figure 9. Genes common for downregulated and upregulated genes for Setd1b cKO (CA-region whole tissue RNA-seq) and Setd1a loss-of-function allele KO (PFC whole tissue RNA seq from published study[176]. (A) Venn diagram for overlapping downregulated genes. (B) List of overlapping up- and downregulated genes. Fold change>1.2, padj<0.1.
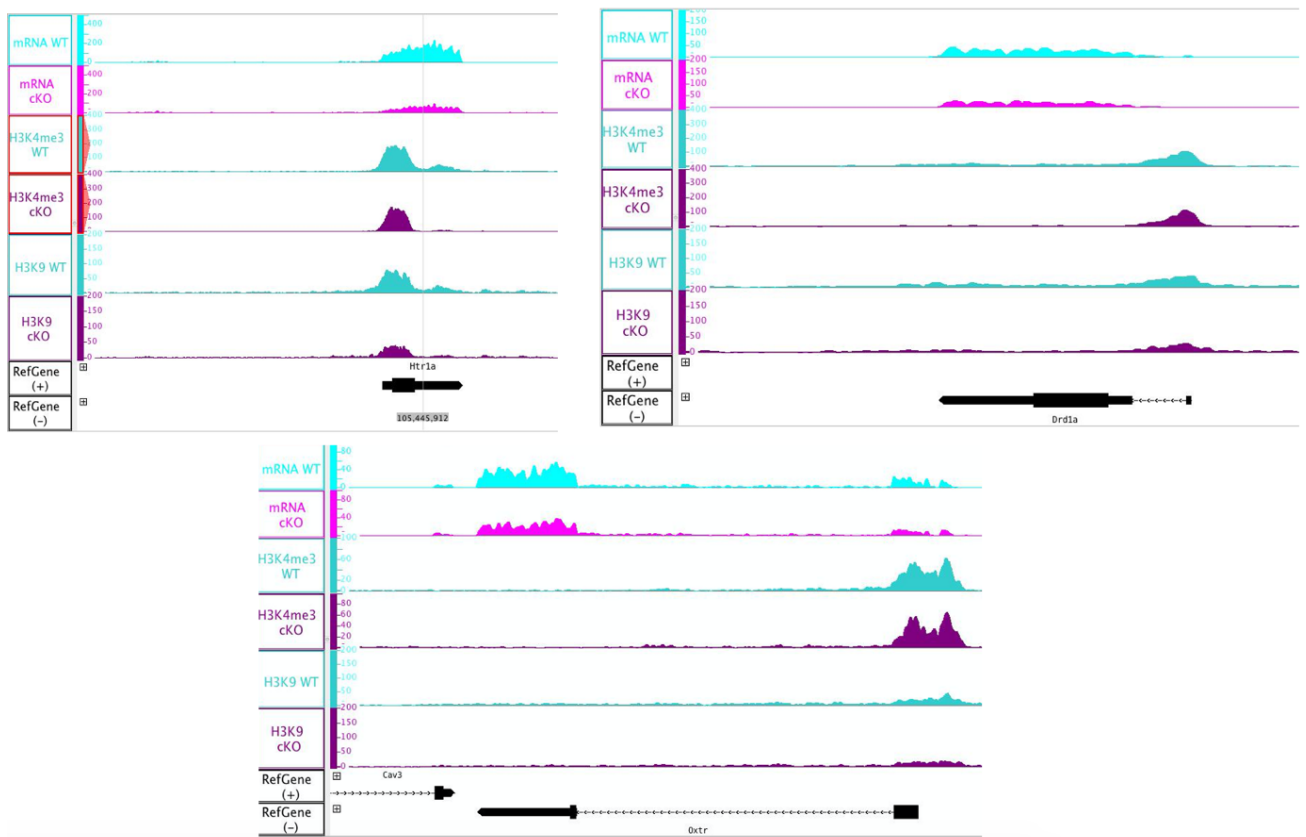
Supplementary Figure 10. IGB profiles for target genes causing nesting defect in Setd1b cKO animals: Htr1a, Drd1, Oxtr.
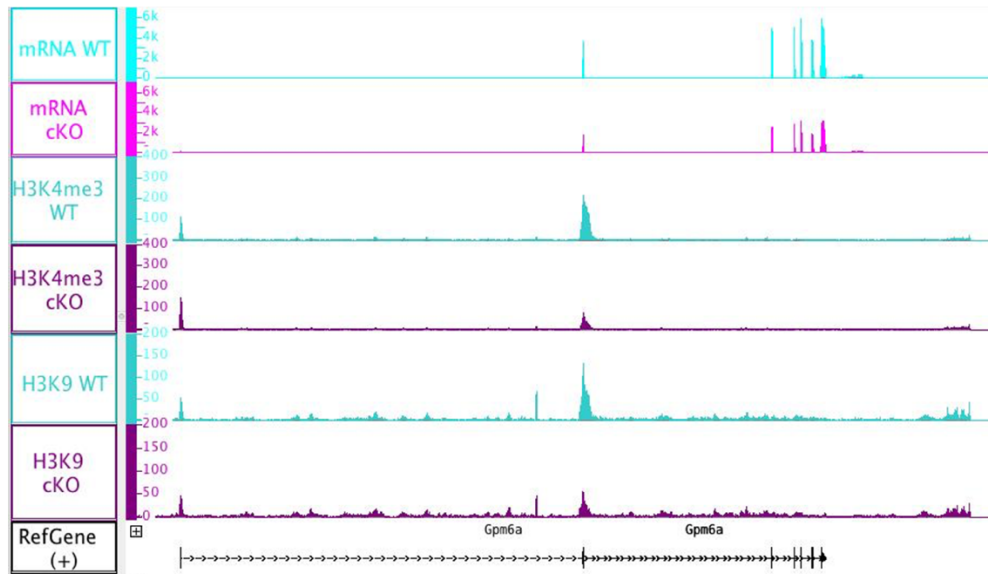
Supplementary Figure 11. IGB profile for Gpm6a gene causing claustrophobia-like behavior in Setd1b cKO mice. H3K4me3 peak is strongly decreased in Gpm6a gene (Log2FoldChange=-1.93, padj<0.05) as well as expression levels in NeuN+ RNA-seq (Log2FoldChange=-0.5927, padj<0.1), indicating that Setd1b is a direct regulator of Gpm6a gene expression.

Name	Gene name	Sequence
Kctd4_F	Potassium channel tetramerization domain containing 4	agggttactatgtgaggctctatctc
Kctd4_R		ttggattatagttgttagtttccatga
Npy2r_tr1_F	Neuropeptide Y receptor Y2	taagctgaggaaccacgtca
Npy2r_tr1_R		gtgccttcgctgatggtaat
Lrcc10b_F	Leucine-rich repeat containing protein 10b	ggaagctcctccaggattg
Lrcc10b_R		tccagagtagccaacccatc
Npy2r_tr2_F	Neuropeptide Y receptor Y2	taagctgaggaaccacgtca
Npy2r_tr2_R		gtgccttcgctgatggtaat
Dcn_F	Decorin	cttctttccacacctgcaaac
Dcn_R		ccctcaaagctgcgatgta
Hif3a_tr2_F	Hypoxia inducible factor 3 subunit alpha	ggccttaggcacacaaccta
Hif3a_tr2_R		cttggtcccagtccatgtg
Lgi_F	Leicine-rich glioma inactivated	ctgcccaggataagtctggt
Lgi_R		tgcttctttggagactgtgtg
Hif3a_tr1_F	Hypoxia inducible factor 3 subunit alpha	ccttccacgatggtgctact
Hif3a_tr1_R		ccttccaaagaggggtgaat
Htr1a_F	5-Hydroxytryptamine receptor 1A	agggcaacaacaccacaac
Htr1a_R		ttggagaggccagtatcgtt
Sstr4_F	Somatostatin receptor 4	cttctgcctcaccgtgct
Sstr4_R		cacactcccaggttgattagc
Wdr17_F	WD repeat domain 17	agtgtctgtgcagtgcttgg
Wdr17_R		aaaacgcccacttgagaatc

Name	Gene name	sequence
Trpa1_F	Transient receptor potential cation channel subfamily a member 1	ctgagatcgaccggagtgtt
Trpa1_R		aacataaaagctgaggccaaaa
Sult5a_F	Sulfotransferase 5	agcctgtcctcctactgtgc
Sult5a_R		tagaccaggggcgtagca
Npr1_F	Natriuretic peptide receptor 6	aaggagctggtctcagagttgt
Npr1_R		ctccgaaggtgcctctcc
Neurod6_F	Neuronal differentiation 6	aacaactattcctcttcaagcatttt
Neurod6_R		ataacctcccccaaataccg
Prdm5_F	PR/SET domain 5	caccttcagtgcttaggagtca
Prdm5_R		gccacaatatgggcagactt
Vill_F	Villin like	cctggggccttgagtacc
Vill_R		ttgcttgaccacaaagatgg

Supplemental Table 17. Table with qPCR (whole tissue RNA-seq validation) primers (all designed in this study)

Gene name	Encoded protein	log2FoldChange	adjusted p-value
Pdgfb	Platelet Derived Growth Factor Subunit B	-0,47619	0,00277
Ctgf	Connective tissue growth factor	-0,61095	0,00360
Vegfa	Vascular endothelial growth factor	-0,40517	0,00017
Hras	Harvey rat sarcoma virus oncogene	-0,30559	0,02860
Nrp1	Neuropilin 1	-0,58547	0,00000
Pdgfa	Platelet derived growth factor alpha	-0,34224	0,00945
Fbxw7	F-box and WD-40 domain protein 7	-0,37384	0,00187
Mt3	Metallothionein 3	-0,29980	0,01809
Htr2a	5-hydroxytryptamine receptor 2A	-0,67873	0,00000
Abca7	ATP-binding cassette, sub-family A, member 7	-0,32750	0,03331
Rara	Retinoic acid receptor alpha	-0,57218	0,00012
Nod1	Nucleotide-binding oligomerization domain containing 1	-0,53256	0,01173
Sema7a	Semaphorin 7A	-0,28726	0,07704
Npnt	Nephronectin	-0,42466	0,00616
Ackr3	Atypical chemokine receptor 3	-0,69140	0,00000
Arrb2	Arrestin beta 2	-0,33221	0,02610
Kdr	Kinase insert domain receptor	-0,43525	0,01888

Supplemental Table 18. Genes from the most downregulated GO-pathway "Positive regulation of ERK1 and ERK2 cascade" from neuronal RNA-seq

Gene name	Encoded protein	log2FoldChange in NeuN+/whole tissue RNA-seq (padj<0.1)	log2FoldChange in NeuN+ H3K4me3 ChIP-seq (padj<0.05)
Nrxn1	neurexin1a	-2.3411 (whole tissue)	-2.21
Pten	phostphatase and tensin homologue	Not significant	-0.95
Shank2	SH3 and multiple ankyrin repeat domains 2	Not significant	-0.92
Cntnap2	contactin associated protein-like2	Not significant	-3.13
Grina	glutamate ionotropic receptor NMDA type subunit associated protein 1	-0.2775 (NeuN+)	Not significant
Grin3a	glutamate ionotropic receptor NMDA type subunit 3A	-0.308 (NeuN+)	Not significant
Grin2a	glutamate ionotropic receptor NMDA type subunit 2A	Not significant	-0.99
Grin1	glutamate ionotropic receptor NMDA type subunit 1	Not significant	-0.84
Grin2d	glutamate ionotropic receptor NMDA type subunit 2D	Not significant	-0.73
Drd1	dopamine receptor D1	-0.4323 (NeuN+)	-2.69
Oxtr	oxytocin receptor	-0.4938 (NeuN+)	Not significant
Cd38	CD38 antigen	-1.2815 (NeuN+)	Not significant
Avpr1b	vasopressin 1b receptor	-0.6114 (whole tissue)	Not significant

Supplemental Table 19. Target genes identified in this study as the causes of impaired nesting phenotype of Setd1b cKO.

8. Bibliography

1. Biswas, S. & Rao, C. M. Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *Eur. J. Pharmacol.* **837**, 8–24 (2018).
2. Kranz, A. & Anastassiadis, K. The role of SETD1A and SETD1B in development and disease. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**, 194578 (2020).
3. Palumbo, O. *et al.* Microdeletion of 12q24.31: report of a girl with intellectual disability, stereotypies, seizures and facial dysmorphisms. *Am. J. Med. Genet. A.* **167A**, 438–444 (2015).
4. Labonne, J. D. J. *et al.* 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**, 757–771 (2016).
5. Daily, D. K., Ardinger, H. H. & Holmes, G. E. Identification and evaluation of mental retardation. *Am. Fam. Physician* **61**, 1059–1067, 1070 (2000).
6. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet Lond. Engl.* **386**, 743–800 (2015).
7. Krzyzewska, I. M. *et al.* A genome-wide DNA methylation signature for SETD1B-related syndrome. *Clin. Epigenetics* **11**, 156 (2019).
8. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
9. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17**, 487–500 (2016).
10. Filipp, F. V. Crosstalk between epigenetics and metabolism-Yin and Yang of histone demethylases and methyltransferases in cancer. *Brief. Funct. Genomics* **16**, 320–325 (2017).
11. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* **49**, e324 (2017).
12. Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer* **7**, 823–833 (2007).
13. Shen, E., Shulha, H., Weng, Z. & Akbarian, S. Regulation of histone H3K4 methylation in brain development and disease. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **369**, (2014).
14. Cloos, P. A. C., Christensen, J., Agger, K. & Helin, K. Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes Dev.* **22**, 1115–1140 (2008).
15. Nislow, C., Ray, E. & Pillus, L. SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell* **8**, 2421–2436 (1997).
16. Rao, R. C. & Dou, Y. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat. Rev. Cancer* **15**, 334–346 (2015).
17. Sze, C. C. & Shilatifard, A. MLL3/MLL4/COMPASS Family on Epigenetic Regulation of Enhancer Function and Cancer. *Cold Spring Harb. Perspect. Med.* **6**, (2016).
18. Gu, B. & Lee, M. G. Histone H3 lysine 4 methyltransferases and demethylases in self-renewal and differentiation of stem cells. *Cell Biosci.* **3**, 39 (2013).
19. Wang, P. *et al.* Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol. Cell. Biol.* **29**, 6074–6085 (2009).
20. Pedersen, M. T. & Helin, K. Histone demethylases in development and disease. *Trends Cell Biol.* **20**, 662–671 (2010).

21. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).

22. D'Oto, A., Tian, Q.-W., Davidoff, A. M. & Yang, J. Histone demethylases and their roles in cancer epigenetics. *J. Med. Oncol. Ther.* **1**, 34–40 (2016).

23. Mosammaparast, N. & Shi, Y. Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu. Rev. Biochem.* **79**, 155–179 (2010).

24. Aprelikova, O. *et al.* The epigenetic modifier JMJD6 is amplified in mammary tumors and cooperates with c-Myc to enhance cellular transformation, tumor progression, and metastasis. *Clin. Epigenetics* **8**, 38 (2016).

25. Wang, J. *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat. Genet.* **41**, 125–129 (2009).

26. Wang, J. *et al.* Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature* **446**, 882–887 (2007).

27. Ciccone, D. N. *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* **461**, 415–418 (2009).

28. Kasioulis, I. *et al.* Kdm3a lysine demethylase is an Hsp90 client required for cytoskeletal rearrangements during spermatogenesis. *Mol. Biol. Cell* **25**, 1216–1233 (2014).

29. Hojati, Z., Soleimanpour, E., Javadirad, S.-M. & Nasr-Esfahani, M. H. Identification of Two Novel Mutations in KDM3A Regulatory Gene in Iranian Infertile Males. *Iran. Biomed. J.* **23**, 220–227 (2019).

30. Zhou, W., Chen, H. & Zhang, L. The PcG protein hPc2 interacts with the N-terminus of histone demethylase JARID1B and acts as a transcriptional co-repressor. *BMB Rep.* **42**, 154–159 (2009).

31. Jensen, L. R. *et al.* Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. *Am. J. Hum. Genet.* **76**, 227–236 (2005).

32. Lederer, D. *et al.* Deletion of KDM6A, a histone demethylase interacting with MLL2, in three patients with Kabuki syndrome. *Am. J. Hum. Genet.* **90**, 119–124 (2012).

33. Miyake, N. *et al.* KDM6A point mutations cause Kabuki syndrome. *Hum. Mutat.* **34**, 108–110 (2013).

34. Dion, M. F., Altschuler, S. J., Wu, L. F. & Rando, O. J. Genomic characterization reveals a simple histone H4 acetylation code. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5501–5506 (2005).

35. Hassan, A. H. *et al.* Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. *Biochem. J.* **402**, 125–133 (2007).

36. LeRoy, G., Rickards, B. & Flint, S. J. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol. Cell* **30**, 51–60 (2008).

37. Kanno, T. *et al.* Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol. Cell* **13**, 33–43 (2004).

38. Winston, F. & Allis, C. D. The bromodomain: a chromatin-targeting module? *Nat. Struct. Biol.* **6**, 601–604 (1999).

39. Carrozza, M. J., Utley, R. T., Workman, J. L. & Côté, J. The diverse functions of histone acetyltransferase complexes. *Trends Genet. TIG* **19**, 321–329 (2003).

40. Ekwall, K. Genome-wide analysis of HDAC function. *Trends Genet. TIG* **21**, 608–615 (2005).

41. Ginsburg, D. S., Govind, C. K. & Hinnebusch, A. G. NuA4 lysine acetyltransferase Esa1 is targeted to coding regions and stimulates transcription elongation with Gcn5. *Mol. Cell. Biol.* **29**, 6473–6487 (2009).

42. Johnsson, A. E. & Wright, A. P. H. The role of specific HAT-HDAC interactions in transcriptional elongation. *Cell Cycle Georget. Tex* **9**, 467–471 (2010).

43. Parthun, M. R., Widom, J. & Gottschling, D. E. The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell* **87**, 85–94 (1996).

44. Wapenaar, H. & Dekker, F. J. Histone acetyltransferases: challenges in targeting bi-substrate enzymes. *Clin. Epigenetics* **8**, 59 (2016).

45. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).

46. Smith, E. R. *et al.* The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. *Mol. Cell. Biol.* **20**, 312–318 (2000).

47. Grant, P. A. *et al.* Expanded lysine acetylation specificity of Gcn5 in native complexes. *J. Biol. Chem.* **274**, 5895–5900 (1999).

48. Yang, X.-J. & Seto, E. The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat. Rev. Mol. Cell Biol.* **9**, 206–218 (2008).

49. Passier, R. *et al.* CaM kinase signaling induces cardiac hypertrophy and activates the MEF2 transcription factor in vivo. *J. Clin. Invest.* **105**, 1395–1406 (2000).

50. Vega, R. B. *et al.* Protein kinases C and D mediate agonist-dependent cardiac hypertrophy through nuclear export of histone deacetylase 5. *Mol. Cell. Biol.* **24**, 8374–8385 (2004).

51. McKinsey, T. A., Zhang, C. L., Lu, J. & Olson, E. N. Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* **408**, 106–111 (2000).

52. Lu, J., McKinsey, T. A., Nicol, R. L. & Olson, E. N. Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4070–4075 (2000).

53. Lu, J., McKinsey, T. A., Zhang, C. L. & Olson, E. N. Regulation of skeletal myogenesis by association of the MEF2 transcription factor with class II histone deacetylases. *Mol. Cell* **6**, 233–244 (2000).

54. Haberland, M., Montgomery, R. L. & Olson, E. N. The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat. Rev. Genet.* **10**, 32–42 (2009).

55. Chang, S. *et al.* Histone deacetylases 5 and 9 govern responsiveness of the heart to a subset of stress signals and play redundant roles in heart development. *Mol. Cell. Biol.* **24**, 8467–8476 (2004).

56. Zhang, C. L. *et al.* Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell* **110**, 479–488 (2002).

57. Vega, R. B. *et al.* Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell* **119**, 555–566 (2004).

58. Zhang, Y. *et al.* Mice lacking histone deacetylase 6 have hyperacetylated tubulin but are viable and develop normally. *Mol. Cell. Biol.* **28**, 1688–1701 (2008).

59. Voss, J. L., Bridge, D. J., Cohen, N. J. & Walker, J. A. A Closer Look at the Hippocampus and Memory. *Trends Cogn. Sci.* **21**, 577–588 (2017).

60. Tromp, D., Dufour, A., Lithfous, S., Pebayle, T. & Després, O. Episodic memory in normal aging and Alzheimer disease: Insights from imaging and behavioral studies. *Ageing Res. Rev.* **24**, 232–262 (2015).

61. Strange, B. A., Witter, M. P., Lein, E. S. & Moser, E. I. Functional organization of the hippocampal longitudinal axis. *Nat. Rev. Neurosci.* **15**, 655–669 (2014).

62. Caruso, V., Lagerström, M. C., Olszewski, P. K., Fredriksson, R. & Schiöth, H. B. Synaptic changes induced by melanocortin signalling. *Nat. Rev. Neurosci.* **15**, 98–110 (2014).

63. Alberini, C. M. & Chen, D. Y. Memory enhancement: consolidation, reconsolidation and insulin-like growth factor 2. *Trends Neurosci.* **35**, 274–283 (2012).

64. D'Urso, A. *et al.* Set1/COMPASS and Mediator are repurposed to promote epigenetic transcriptional memory. *eLife* **5**, (2016).

65. Cole, M. D. & Cowling, V. H. Transcription-independent functions of MYC: regulation of translation and DNA replication. *Nat. Rev. Mol. Cell Biol.* **9**, 810–815 (2008).

66. Jonkers, I. & Lis, J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 167–177 (2015).

67. van Ingen, H. *et al.* Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Struct. Lond. Engl. 1993* **16**, 1245–1256 (2008).

68. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343–357 (2012).

69. Minatohara, K., Akiyoshi, M. & Okuno, H. Role of Immediate-Early Genes in Synaptic Plasticity and Neuronal Ensembles Underlying the Memory Trace. *Front. Mol. Neurosci.* **8**, 78 (2015).

70. Lin, H., Min, Z. & Tao, Q. The MLL/Setd1b methyltransferase is required for the Spemann's organizer gene activation in Xenopus. *Mech. Dev.* **142**, 1–9 (2016).

71. Crump, N. T. & Milne, T. A. Why are so many MLL lysine methyltransferases required for normal mammalian development? *Cell. Mol. Life Sci. CMLS* **76**, 2885–2898 (2019).

72. Ang, Y.-S. *et al.* Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183–197 (2011).

73. Shimomura, A. & Hashino, E. Epigenetic Regulation of Neural Differentiation from Embryonic Stem Cells. in *Trends in Cell Signaling Pathways in Neuronal Fate Decision* (ed. Wislet-Gendebien, S.) (InTech, 2013). doi:10.5772/53650.

74. Denissov, S. *et al.* Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Dev. Camb. Engl.* **141**, 526–537 (2014).

75. Covic, M., Karaca, E. & Lie, D. C. Epigenetic regulation of neurogenesis in the adult hippocampus. *Heredity* **105**, 122–134 (2010).

76. Gupta, S. *et al.* Histone methylation regulates memory formation. *J. Neurosci. Off. J. Soc. Neurosci.* **30**, 3589–3599 (2010).

77. Kerimoglu, C. *et al.* Histone-methyltransferase MLL2 (KMT2B) is required for memory formation in mice. *J. Neurosci. Off. J. Soc. Neurosci.* **33**, 3452–3464 (2013).

78. Kerimoglu, C. *et al.* KMT2A and KMT2B Mediate Memory Function by Affecting Distinct Genomic Regions. *Cell Rep.* **20**, 538–548 (2017).

79. Gupta-Agarwal, S. *et al.* G9a/GLP histone lysine dimethyltransferase complex activity in the hippocampus and the entorhinal cortex is required for gene activation and silencing during memory consolidation. *J. Neurosci. Off. J. Soc. Neurosci.* **32**, 5440–5453 (2012).

80. Benito, E. *et al.* HDAC inhibitor-dependent transcriptome and memory reinstatement in cognitive decline models. *J. Clin. Invest.* **125**, 3572–3584 (2015).

81. Benito, E. *et al.* The BET/BRD inhibitor JQ1 improves brain plasticity in WT and APP mice. *Transl. Psychiatry* **7**, e1239 (2017).

82. Morse, S. J., Butler, A. A., Davis, R. L., Soller, I. J. & Lubin, F. D. Environmental enrichment reverses histone methylation changes in the aged hippocampus and restores age-related memory deficits. *Biology* **4**, 298–313 (2015).

83. Jakovcevski, M. & Akbarian, S. Epigenetic mechanisms in neurological disease. *Nat. Med.* **18**, 1194–1204 (2012).

84. Mietton, L. *et al.* RNA Sequencing and Pathway Analysis Identify Important Pathways Involved in Hypertrichosis and Intellectual Disability in Patients with Wiedemann-Steiner Syndrome. *Neuromolecular Med.* **20**, 409–417 (2018).

85. Vallianatos, C. N. & Iwase, S. Disrupted intricacy of histone H3K4 methylation in neurodevelopmental disorders. *Epigenomics* **7**, 503–519 (2015).

86. Meyer, E. *et al.* Mutations in the histone methyltransferase gene KMT2B cause complex early-onset dystonia. *Nat. Genet.* **49**, 223–237 (2017).

87. Koemans, T. S. *et al.* Functional convergence of histone methyltransferases EHMT1 and KMT2C involved in intellectual disability and autism spectrum disorder. *PLoS Genet.* **13**, e1006864 (2017).

88. Lavery, W. J., Barski, A., Wiley, S., Schorry, E. K. & Lindsley, A. W. KMT2C/D COMPASS complex-associated diseases [KCDCOM-ADs]: an emerging class of congenital regulopathies. *Clin. Epigenetics* **12**, 10 (2020).

89. Bjornsson, H. T. *et al.* Histone deacetylase inhibition rescues structural and functional brain deficits in a mouse model of Kabuki syndrome. *Sci. Transl. Med.* **6**, 256ra135 (2014).

90. Adam, M. P., Hudgins, L. & Hannibal, M. Kabuki Syndrome. in *GeneReviews®* (eds. Adam, M. P. et al.) (University of Washington, Seattle, 1993).

91. Takata, A. *et al.* Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron* **82**, 773–780 (2014).

92. Singh, T. *et al.* Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nat. Neurosci.* **19**, 571–577 (2016).

93. Owen, M. J., Sawa, A. & Mortensen, P. B. Schizophrenia. *Lancet Lond. Engl.* **388**, 86–97 (2016).

94. Levenson, J. M. *et al.* Regulation of histone acetylation during memory formation in the hippocampus. *J. Biol. Chem.* **279**, 40545–40559 (2004).

95. Swank, M. W. & Sweatt, J. D. Increased histone acetyltransferase and lysine acetyltransferase activity and biphasic activation of the ERK/RSK cascade in insular cortex during novel taste learning. *J. Neurosci. Off. J. Soc. Neurosci.* **21**, 3383–3391 (2001).

96. Chwang, W. B., O'Riordan, K. J., Levenson, J. M. & Sweatt, J. D. ERK/MAPK regulates hippocampal histone phosphorylation following contextual fear conditioning. *Learn. Mem. Cold Spring Harb. N* **13**, 322–328 (2006).

97. Miller, C. A. & Sweatt, J. D. Covalent modification of DNA regulates memory formation. *Neuron* **53**, 857–869 (2007).

98. Miller, C. A., Campbell, S. L. & Sweatt, J. D. DNA methylation and histone acetylation work in concert to regulate memory formation and synaptic plasticity. *Neurobiol. Learn. Mem.* **89**, 599–603 (2008).

99. Gräff, J. & Tsai, L.-H. Histone acetylation: molecular mnemonics on the chromatin. *Nat. Rev. Neurosci.* **14**, 97–111 (2013).

100. McQuown, S. C. & Wood, M. A. HDAC3 and the molecular brake pad hypothesis. *Neurobiol. Learn. Mem.* **96**, 27–34 (2011).

101. Petrij, F. *et al.* Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **376**, 348–351 (1995).

102. Alarcón, J. M. *et al.* Chromatin acetylation, memory, and LTP are impaired in CBP+/- mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. *Neuron* **42**, 947–959 (2004).

103. Chen, G., Zou, X., Watanabe, H., van Deursen, J. M. & Shen, J. CREB binding protein is required for both short-term and long-term memory formation. *J. Neurosci. Off. J. Soc. Neurosci.* **30**, 13066–13077 (2010).

104. Wood, M. A. *et al.* Transgenic mice expressing a truncated form of CREB-binding protein (CBP) exhibit deficits in hippocampal synaptic plasticity and memory storage. *Learn. Mem. Cold Spring Harb. N* **12**, 111–119 (2005).

105. Valor, L. M. & Guiretti, D. What's wrong with epigenetics in Huntington's disease? *Neuropharmacology* **80**, 103–114 (2014).

106. Steffan, J. S. *et al.* The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6763–6768 (2000).

107. Jiang, H., Nucifora, F. C., Ross, C. A. & DeFranco, D. B. Cell death triggered by polyglutamine-expanded huntingtin in a neuronal cell line is associated with degradation of CREB-binding protein. *Hum. Mol. Genet.* **12**, 1–12 (2003).

108. Huang, Y. & Mucke, L. Alzheimer mechanisms and therapeutic strategies. *Cell* **148**, 1204–1222 (2012).

109. Lu, X. *et al.* Histone acetyltransferase p300 mediates histone acetylation of PS1 and BACE1 in a cellular model of Alzheimer's disease. *PloS One* **9**, e103067 (2014).

110. Fischer, A., Sananbenesi, F., Mungenast, A. & Tsai, L.-H. Targeting the correct HDAC(s) to treat cognitive disorders. *Trends Pharmacol. Sci.* **31**, 605–617 (2010).

111. Mann, B. S., Johnson, J. R., Cohen, M. H., Justice, R. & Pazdur, R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *The Oncologist* **12**, 1247–1252 (2007).

112. Stowell, J. C., Huot, R. I. & Van Voast, L. The synthesis of N-hydroxy-N'-phenyloctanediamide and its inhibitory effect on proliferation of AXC rat prostate cancer cells. *J. Med. Chem.* **38**, 1411–1413 (1995).

113. McCampbell, A. *et al.* Histone deacetylase inhibitors reduce polyglutamine toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 15179–15184 (2001).

114. Steffan, J. S. *et al.* Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. *Nature* **413**, 739–743 (2001).

115. Ferrante, R. J. *et al.* Histone deacetylase inhibition by sodium butyrate chemotherapy ameliorates the neurodegenerative phenotype in Huntington's disease mice. *J. Neurosci. Off. J. Soc. Neurosci.* **23**, 9418–9427 (2003).

116. Hockly, E. *et al.* Suberoylanilide hydroxamic acid, a histone deacetylase inhibitor, ameliorates motor deficits in a mouse model of Huntington's disease. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2041–2046 (2003).

117. Thomas, E. A. *et al.* The HDAC inhibitor 4b ameliorates the disease phenotype and transcriptional abnormalities in Huntington's disease transgenic mice. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15564–15569 (2008).

118. Ricobaraza, A. *et al.* Phenylbutyrate ameliorates cognitive deficit and reduces tau pathology in an Alzheimer's disease mouse model. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **34**, 1721–1732 (2009).

119. Kilgore, M. *et al.* Inhibitors of class 1 histone deacetylases reverse contextual memory deficits in a mouse model of Alzheimer's disease. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **35**, 870–880 (2010).

120. Wu, M. *et al.* Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.* **28**, 7337–7344 (2008).

121. Bledau, A. S. *et al.* The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Dev. Camb. Engl.* **141**, 1022–1035 (2014).

122. Brici, D. *et al.* Setd1b, encoding a histone 3 lysine 4 methyltransferase, is a maternal effect gene required for the oogenic gene expression program. *Dev. Camb. Engl.* **144**, 2606–2617 (2017).

123. Wang, X., Zhang, C., Szábo, G. & Sun, Q.-Q. Distribution of CaMKIIα expression in the brain in vivo, studied by CaMKIIα-GFP mice. *Brain Res.* **1518**, 9–25 (2013).

124. Hiraide, T. *et al.* De novo variants in SETD1B cause intellectual disability, autism spectrum disorder, and epilepsy with myoclonic absences. *Epilepsia Open* **4**, 476–481 (2019).

125. Hiraide, T. *et al.* De novo variants in SETD1B are associated with intellectual disability, epilepsy and autism. *Hum. Genet.* **137**, 95–104 (2018).

126. Den, K. *et al.* A novel de novo frameshift variant in SETD1B causes epilepsy. *J. Hum. Genet.* **64**, 821–827 (2019).

127. Ferreira, L. S. S., Fernandes, C. S., Vieira, M. N. N. & De Felice, F. G. Insulin Resistance in Alzheimer's Disease. *Front. Neurosci.* **12**, 830 (2018).

128. Westwood, A. J. *et al.* Insulin-like growth factor-1 and risk of Alzheimer dementia and brain atrophy. *Neurology* **82**, 1613–1619 (2014).

129. Llorens-Martín, M., Torres-Alemán, I. & Trejo, J. L. Exercise modulates insulin-like growth factor 1-dependent and -independent effects on adult hippocampal neurogenesis and behaviour. *Mol. Cell. Neurosci.* **44**, 109–117 (2010).

130. Trejo, J. L., Carro, E. & Torres-Aleman, I. Circulating insulin-like growth factor I mediates exercise-induced increases in the number of new neurons in the adult hippocampus. *J. Neurosci. Off. J. Soc. Neurosci.* **21**, 1628–1634 (2001).

131. Allard, J. B. & Duan, C. IGF-Binding Proteins: Why Do They Exist and Why Are There So Many? *Front. Endocrinol.* **9**, 117 (2018).

132. Agis-Balboa, R. C. *et al.* A hippocampal insulin-growth factor 2 pathway regulates the extinction of fear memories. *EMBO J.* **30**, 4071–4083 (2011).

133. Agbemenyah, H. Y., Agis-Balboa, R. C., Burkhardt, S., Delalle, I. & Fischer, A. Insulin growth factor binding protein 7 is a novel target to treat dementia. *Neurobiol. Dis.* **62**, 135–143 (2014).

134. Oh, Y. *et al.* Synthesis and characterization of insulin-like growth factor-binding protein (IGFBP)-7. Recombinant human mac25 protein specifically binds IGF-I and -II. *J. Biol. Chem.* **271**, 30322–30325 (1996).

135. Clemmons, D. R. Modifying IGF1 activity: an approach to treat endocrine disorders, atherosclerosis and cancer. *Nat. Rev. Drug Discov.* **6**, 821–833 (2007).

136. Kim, H.-S., Rosenfeld, R. G. & Oh, Y. Biological roles of insulin-like growth factor binding proteins (IGFBPs). *Exp. Mol. Med.* **29**, 85–96 (1997).

137. Collet, C. & Candy, J. How many insulin-like growth factor binding proteins? *Mol. Cell. Endocrinol.* **139**, 1–6 (1998).

138. Evdokimova, V. *et al.* IGFBP7 binds to the IGF-1 receptor and blocks its activation by insulin-like growth factors. *Sci. Signal.* **5**, ra92 (2012).

139. Burger, A. M. *et al.* Down-regulation of T1A12/mac25, a novel insulin-like growth factor binding protein related gene, is associated with disease progression in breast carcinomas. *Oncogene* **16**, 2459–2467 (1998).

140. Sprenger, C. C., Damon, S. E., Hwa, V., Rosenfeld, R. G. & Plymate, S. R. Insulin-like growth factor binding protein-related protein 1 (IGFBP-rP1) is a potential tumor suppressor protein for prostate cancer. *Cancer Res.* **59**, 2370–2375 (1999).

141. Ruan, W. *et al.* IGFBP7 plays a potential tumor suppressor role in colorectal carcinogenesis. *Cancer Biol. Ther.* **6**, 354–359 (2007).

142. Scharf, J.-G. & Braulke, T. The role of the IGF axis in hepatocarcinogenesis. *Horm. Metab. Res. Horm. Stoffwechselforschung Horm. Metab.* **35**, 685–693 (2003).

143. Verhagen, H. J. *et al.* IGFBP7 induces apoptosis of acute myeloid leukemia cells and synergizes with chemotherapy in suppression of leukemia cell survival. *Cell Death Dis.* **5**, e1300 (2014).

144. Bolomsky, A. *et al.* Insulin like growth factor binding protein 7 (IGFBP7) expression is linked to poor prognosis but may protect from bone disease in multiple myeloma. *J. Hematol. Oncol.J Hematol Oncol* **8**, 10 (2015).

145. Wajapeyee, N., Serra, R. W., Zhu, X., Mahalingam, M. & Green, M. R. Oncogenic BRAF induces senescence and apoptosis through pathways mediated by the secreted protein IGFBP7. *Cell* **132**, 363–374 (2008).

146. Duron, E. *et al.* Insulin-like growth factor-I and insulin-like growth factor binding protein-3 in Alzheimer's disease. *J. Clin. Endocrinol. Metab.* **97**, 4673–4681 (2012).

147. Anand, R., Gill, K. D. & Mahdi, A. A. Therapeutics of Alzheimer's disease: Past, present and future. *Neuropharmacology* **76 Pt A**, 27–50 (2014).

148. Braak, H., Rüb, U. & Del Tredici, K. Cognitive decline correlates with neuropathological stage in Parkinson's disease. *J. Neurol. Sci.* **248**, 255–258 (2006).

149. Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).

150. Blennow, K., de Leon, M. J. & Zetterberg, H. Alzheimer's disease. *Lancet Lond. Engl.* **368**, 387–403 (2006).

151. Koekkoek, P. S., Kappelle, L. J., van den Berg, E., Rutten, G. E. H. M. & Biessels, G. J. Cognitive function in patients with diabetes mellitus: guidance for daily care. *Lancet Neurol.* **14**, 329–340 (2015).

152. Fernandez, A. M. & Torres-Alemán, I. The many faces of insulin-like peptide signalling in the brain. *Nat. Rev. Neurosci.* **13**, 225–239 (2012).

153. Glasper, E. R., Llorens-Martin, M. V., Leuner, B., Gould, E. & Trejo, J. L. Blockade of insulin-like growth factor-I has complex effects on structural plasticity in the hippocampus. *Hippocampus* **20**, 706–712 (2010).

154. Seibenhener, M. L. & Wooten, M. C. Use of the Open Field Maze to measure locomotor and anxiety-like behavior in mice. *J. Vis. Exp. JoVE* e52434 (2015) doi:10.3791/52434.

155. Illouz, T., Madar, R., Louzoun, Y., Griffioen, K. J. & Okun, E. Unraveling cognitive traits using the Morris water maze unbiased strategy classification (MUST-C) algorithm. *Brain. Behav. Immun.* **52**, 132–144 (2016).

156. Deacon, R. M. J. Assessing nest building in mice. *Nat. Protoc.* **1**, 1117–1119 (2006).

157. Dunham, N. W. & Miya, T. S. A note on a simple apparatus for detecting neurological deficit in rats and mice. *J. Am. Pharm. Assoc. Am. Pharm. Assoc.* **46**, 208–209 (1957).

158. Mena, A. *et al.* Reduced Prepulse Inhibition as a Biomarker of Schizophrenia. *Front. Behav. Neurosci.* **10**, 202 (2016).

159. Cryan, J. F., Mombereau, C. & Vassout, A. The tail suspension test as a model for assessing antidepressant activity: review of pharmacological and genetic studies in mice. *Neurosci. Biobehav. Rev.* **29**, 571–625 (2005).

160. Halder, R. *et al.* DNA methylation changes in plasticity genes accompany the formation and maintenance of memory. *Nat. Neurosci.* **19**, 102–110 (2016).

161. Morris, R. Developments of a water-maze procedure for studying spatial learning in the rat. *J. Neurosci. Methods* **11**, 47–60 (1984).

162. Yee, B. K. & Singer, P. A conceptual and practical guide to the behavioural evaluation of animal models of the symptomatology and therapy of schizophrenia. *Cell Tissue Res.* **354**, 221–246 (2013).

163. Castagné, V., Moser, P., Roux, S. & Porsolt, R. D. Rodent models of depression: forced swim and tail suspension behavioral despair tests in rats and mice. *Curr. Protoc. Neurosci.* **Chapter 8**, Unit 8.10A (2011).

164. Klug, J. R. *et al.* Genetic inhibition of CaMKII in dorsal striatal medium spiny neurons reduces functional excitatory synapses and enhances intrinsic excitability. *PloS One* **7**, e45323 (2012).

165. Dang, M. T. *et al.* Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 15254–15259 (2006).

166. Crawley, J. N. Designing mouse behavioral tasks relevant to autistic-like behaviors. *Ment. Retard. Dev. Disabil. Res. Rev.* **10**, 248–258 (2004).

167. Shoji, H. & Miyakawa, T. Relationships between the acoustic startle response and prepulse inhibition in C57BL/6J mice: a large-scale meta-analytic study. *Mol. Brain* **11**, 42 (2018).

168. Baranek, G. T. Efficacy of sensory and motor interventions for children with autism. *J. Autism Dev. Disord.* **32**, 397–422 (2002).

169. Mielnik, C. A. *et al.* Consequences of NMDA receptor deficiency can be rescued in the adult brain. *Mol. Psychiatry* (2020) doi:10.1038/s41380-020-00859-4.

170. Quesnel-Vallières, M. *et al.* Misregulation of an Activity-Dependent Splicing Network as a Common Mechanism Underlying Autism Spectrum Disorders. *Mol. Cell* **64**, 1023–1034 (2016).

171. Choudhury, R., Singh, S., Arumugam, S., Roguev, A. & Stewart, A. F. The Set1 complex is dimeric and acts with Jhd2 demethylation to convey symmetrical H3K4 trimethylation. *Genes Dev.* **33**, 550–564 (2019).

172. Girdhar, K. *et al.* Cell-specific histone modification maps in the human frontal lobe link schizophrenia risk to the neuronal epigenome. *Nat. Neurosci.* **21**, 1126–1136 (2018).

173. Benayoun, B. A. *et al.* H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **158**, 673–688 (2014).

174. Jakovcevski, M. *et al.* Neuronal Kmt2a/Mll1 histone methyltransferase is essential for prefrontal synaptic plasticity and working memory. *J. Neurosci. Off. J. Soc. Neurosci.* **35**, 5097–5108 (2015).

175. Shen, E. Y. *et al.* 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. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **41**, 3103–3113 (2016).

176. Mukai, J. *et al.* Recapitulation and Reversal of Schizophrenia-Related Phenotypes in Setd1a-Deficient Mice. *Neuron* **104**, 471-487.e12 (2019).

177. Van der Jeugd, A. *et al.* Progressive age-related cognitive decline in tau mice. *J. Alzheimers Dis. JAD* **37**, 777–788 (2013).

178. Serneels, L. *et al.* gamma-Secretase heterogeneity in the Aph1 subunit: relevance for Alzheimer's disease. *Science* **324**, 639–642 (2009).

179. Radde, R. *et al.* Abeta42-driven cerebral amyloidosis in transgenic mice reveals early and robust pathology. *EMBO Rep.* **7**, 940–946 (2006).

180. Latham, N. & Mason, G. From house mouse to mouse house: the behavioural biology of free-living Mus musculus and its implications in the laboratory. *Appl. Anim. Behav. Sci.* **86**, 261–289 (2004).

181. Jirkof, P. Burrowing and nest building behavior as indicators of well-being in mice. *J. Neurosci. Methods* **234**, 139–146 (2014).

182. Udagawa, T. *et al.* Genetic and acute CPEB1 depletion ameliorate fragile X pathophysiology. *Nat. Med.* **19**, 1473–1477 (2013).

183. Na, E. S., Nelson, E. D., Kavalali, E. T. & Monteggia, L. M. The impact of MeCP2 loss- or gain-of-function on synaptic plasticity. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **38**, 212–219 (2013).

184. Ross, P. D. *et al.* Exclusive expression of MeCP2 in the nervous system distinguishes between brain and peripheral Rett syndrome-like phenotypes. *Hum. Mol. Genet.* **25**, 4389–4404 (2016).

185. Silverman, J. L., Yang, M., Lord, C. & Crawley, J. N. Behavioural phenotyping assays for mouse models of autism. *Nat. Rev. Neurosci.* **11**, 490–502 (2010).

186. El-Kordi, A. *et al.* Development of an autism severity score for mice using Nlgn4 null mutants as a construct-valid model of heritable monogenic autism. *Behav. Brain Res.* **251**, 41–49 (2013).

187. Speed, H. E. *et al.* Autism-Associated Insertion Mutation (InsG) of Shank3 Exon 21 Causes Impaired Synaptic Transmission and Behavioral Deficits. *J. Neurosci. Off. J. Soc. Neurosci.* **35**, 9648–9665 (2015).

188. Peñagarikano, O. *et al.* Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell* **147**, 235–246 (2011).

189. Pedersen, C. S., Sørensen, D. B., Parachikova, A. I. & Plath, N. PCP-induced deficits in murine nest building activity: employment of an ethological rodent behavior to mimic negative-like symptoms of schizophrenia. *Behav. Brain Res.* **273**, 63–72 (2014).

190. Belforte, J. E. *et al.* Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. *Nat. Neurosci.* **13**, 76–83 (2010).

191. Demontis, D. *et al.* Association of GRIN1 and GRIN2A-D with schizophrenia and genetic interaction with maternal herpes simplex virus-2 infection affecting disease risk. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **156B**, 913–922 (2011).

192. Liu, Y.-P. *et al.* Association between polymorphisms in the GRIN1 gene 5' regulatory region and schizophrenia in a northern Han Chinese population and haplotype effects on protein expression in vitro. *BMC Med. Genet.* **20**, 26 (2019).

193. Torres-Lista, V. & Giménez-Llort, L. Impairment of nesting behaviour in 3xTg-AD mice. *Behav. Brain Res.* **247**, 153–157 (2013).

194. Lippi, S. L. P., Smith, M. L. & Flinn, J. M. A Novel hAPP/htau Mouse Model of Alzheimer's Disease: Inclusion of APP With Tau Exacerbates Behavioral Deficits and Zinc Administration Heightens Tangle Pathology. *Front. Aging Neurosci.* **10**, 382 (2018).

195. Wesson, D. W. & Wilson, D. A. Age and gene overexpression interact to abolish nesting behavior in Tg2576 amyloid precursor protein (APP) mice. *Behav. Brain Res.* **216**, 408–413 (2011).

196. Filali, M. & Lalonde, R. Age-related cognitive decline and nesting behavior in an APPswe/PS1 bigenic model of Alzheimer's disease. *Brain Res.* **1292**, 93–99 (2009).

197. Bendesky, A. *et al.* The genetic basis of parental care evolution in monogamous mice. *Nature* **544**, 434–439 (2017).

198. Kohl, J. *et al.* Functional circuit architecture underlying parental behaviour. *Nature* **556**, 326–331 (2018).

199. Lin, L., Chen, G., Kuang, H., Wang, D. & Tsien, J. Z. Neural encoding of the concept of nest in the mouse brain. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 6066–6071 (2007).

200. Amigó, J. *et al.* The absence of 5-HT4 receptors modulates depression- and anxiety-like responses and influences the response of fluoxetine in olfactory bulbectomised mice: Adaptive changes in hippocampal neuroplasticity markers and 5-HT1A autoreceptor. *Neuropharmacology* **111**, 47–58 (2016).

201. Heller, H. C. *et al.* Nest building is impaired in the Ts65Dn mouse model of Down syndrome and rescued by blocking 5HT2a receptors. *Neurobiol. Learn. Mem.* **116**, 162–171 (2014).

202. Moy, S. S. *et al.* Social approach in genetically engineered mouse lines relevant to autism. *Genes Brain Behav.* **8**, 129–142 (2009).

203. Yang, S. Y., Yoo, H. J., Cho, I. H., Park, M. & Kim, S. A. Association with tryptophan hydroxylase 2 gene polymorphisms and autism spectrum disorders in Korean families. *Neurosci. Res.* **73**, 333–336 (2012).

204. Egawa, J. *et al.* A detailed association analysis between the tryptophan hydroxylase 2 (TPH2) gene and autism spectrum disorders in a Japanese population. *Psychiatry Res.* **196**, 320–322 (2012).

205. Angoa-Pérez, M. *et al.* Brain serotonin determines maternal behavior and offspring survival. *Genes Brain Behav.* **13**, 579–591 (2014).

206. Svirsky, N., Levy, S. & Avitsur, R. Prenatal exposure to selective serotonin reuptake inhibitors (SSRI) increases aggression and modulates maternal behavior in offspring mice. *Dev. Psychobiol.* **58**, 71–82 (2016).

207. Li, X., Morrow, D. & Witkin, J. M. Decreases in nestlet shredding of mice by serotonin uptake inhibitors: comparison with marble burying. *Life Sci.* **78**, 1933–1939 (2006).

208. Schneider, C. W. & Chenoweth, M. B. Effects of hallucinogenic and other drugs on the nest-building behaviour of mice. *Nature* **225**, 1262–1263 (1970).

209. Szczypka, M. S. *et al.* Dopamine production in the caudate putamen restores feeding in dopamine-deficient mice. *Neuron* **30**, 819–828 (2001).

210. Tsutsui-Kimura, I. *et al.* Dysfunction of ventrolateral striatal dopamine receptor type 2-expressing medium spiny neurons impairs instrumental motivation. *Nat. Commun.* **8**, 14304 (2017).

211. Quintana, D. S. *et al.* Oxytocin pathway gene networks in the human brain. *Nat. Commun.* **10**, 668 (2019).

212. Baskerville, T. A. & Douglas, A. J. Dopamine and oxytocin interactions underlying behaviors: potential contributions to behavioral disorders. *CNS Neurosci. Ther.* **16**, e92-123 (2010).

213. Knobloch, H. S. & Grinevich, V. Evolution of oxytocin pathways in the brain of vertebrates. *Front. Behav. Neurosci.* **8**, 31 (2014).

214. Husarova, V. M. *et al.* Plasma Oxytocin in Children with Autism and Its Correlations with Behavioral Parameters in Children and Parents. *Psychiatry Investig.* **13**, 174–183 (2016).

215. Lee, H.-J., Caldwell, H. K., Macbeth, A. H., Tolu, S. G. & Young, W. S. A conditional knockout mouse line of the oxytocin receptor. *Endocrinology* **149**, 3256–3263 (2008).

216. de Oliveira Pereira Ribeiro, L. *et al.* Evidence for Association Between OXTR Gene and ASD Clinical Phenotypes. *J. Mol. Neurosci. MN* **65**, 213–221 (2018).

217. Lopatina, O., Inzhutova, A., Salmina, A. B. & Higashida, H. The roles of oxytocin and CD38 in social or parental behaviors. *Front. Neurosci.* **6**, 182 (2012).

218. Chatterjee, O. *et al.* An overview of the oxytocin-oxytocin receptor signaling network. *J. Cell Commun. Signal.* **10**, 355–360 (2016).

219. Satoh, Y. *et al.* ERK2 contributes to the control of social behaviors in mice. *J. Neurosci. Off. J. Soc. Neurosci.* **31**, 11953–11967 (2011).

220. Wersinger, S. R., Ginns, E. I., O'Carroll, A.-M., Lolait, S. J. & Young, W. S. Vasopressin V1b receptor knockout reduces aggressive behavior in male mice. *Mol. Psychiatry* **7**, 975–984 (2002).

221. Gjoneska, E. *et al.* Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* **518**, 365–369 (2015).

222. Shulha, H. P. *et al.* Epigenetic signatures of autism: trimethylated H3K4 landscapes in prefrontal neurons. *Arch. Gen. Psychiatry* **69**, 314–324 (2012).

223. Wang, Z. *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019–1031 (2009).

224. Soares, L. M. *et al.* Determinants of Histone H3K4 Methylation Patterns. *Mol. Cell* **68**, 773-785.e6 (2017).

225. Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7285–7290 (2015).

226. Habib, N. *et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).

227. Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).

228. Zeisel, A. *et al.* Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999-1014.e22 (2018).

229. Stewart, A. M., Braubach, O., Spitsbergen, J., Gerlai, R. & Kalueff, A. V. Zebrafish models for translational neuroscience research: from tank to bedside. *Trends Neurosci.* **37**, 264–278 (2014).

230. Antinucci, P. & Hindges, R. A crystal-clear zebrafish for in vivo imaging. *Sci. Rep.* **6**, 29490 (2016).

231. Eichenbaum, H. The role of the hippocampus in navigation is memory. *J. Neurophysiol.* **117**, 1785–1796 (2017).

232. Rendall, A. R., Truong, D. T. & Fitch, R. H. Learning delays in a mouse model of Autism Spectrum Disorder. *Behav. Brain Res.* **303**, 201–207 (2016).

233. Um, S. M. *et al.* NGL-2 Deletion Leads to Autistic-like Behaviors Responsive to NMDAR Modulation. *Cell Rep.* **23**, 3839–3851 (2018).

234. Berkowitz, L. E., Harvey, R. E., Drake, E., Thompson, S. M. & Clark, B. J. Progressive impairment of directional and spatially precise trajectories by TgF344-Alzheimer's disease rats in the Morris Water Task. *Sci. Rep.* **8**, 16153 (2018).

235. Manczak, M., Kandimalla, R., Yin, X. & Reddy, P. H. Hippocampal mutant APP and amyloid beta-induced cognitive decline, dendritic spine loss, defective autophagy, mitophagy and mitochondrial abnormalities in a mouse model of Alzheimer's disease. *Hum. Mol. Genet.* **27**, 1332–1342 (2018).

236. Radyushkin, K. *et al.* Complexin2 null mutation requires a 'second hit' for induction of phenotypic changes relevant to schizophrenia. *Genes Brain Behav.* **9**, 592–602 (2010).

237. Volkert, M. R., Elliott, N. A. & Housman, D. E. Functional genomics reveals a family of eukaryotic oxidation protection genes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 14530–14535 (2000).

238. Oliver, P. L. *et al.* Oxr1 is essential for protection against oxidative stress-induced neurodegeneration. *PLoS Genet.* **7**, e1002338 (2011).

239. Liu, K. X. *et al.* Neuron-specific antioxidant OXR1 extends survival of a mouse model of amyotrophic lateral sclerosis. *Brain J. Neurol.* **138**, 1167–1181 (2015).

240. Jiang, Y. *et al.* Serum secreted miR-137-containing exosomes affects oxidative stress of neurons by regulating OXR1 in Parkinson's disease. *Brain Res.* **1722**, 146331 (2019).

241. Wang, J. *et al.* Loss of Oxidation Resistance 1, OXR1, Is Associated with an Autosomal-Recessive Neurological Disease with Cerebellar Atrophy and Lysosomal Dysfunction. *Am. J. Hum. Genet.* **105**, 1237–1253 (2019).

242. Suh, J., Rivest, A. J., Nakashiba, T., Tominaga, T. & Tonegawa, S. Entorhinal cortex layer III input to the hippocampus is crucial for temporal association memory. *Science* **334**, 1415–1420 (2011).

243. Massaad, C. A. & Klann, E. Reactive oxygen species in the regulation of synaptic plasticity and memory. *Antioxid. Redox Signal.* **14**, 2013–2054 (2011).

244. Popa-Wagner, A., Mitran, S., Sivanesan, S., Chang, E. & Buga, A.-M. ROS and brain diseases: the good, the bad, and the ugly. *Oxid. Med. Cell. Longev.* **2013**, 963520 (2013).

245. El-Kordi, A. *et al.* A single gene defect causing claustrophobia. *Transl. Psychiatry* **3**, e254 (2013).

9. Abbreviations

µg	microgram
µl	microliter
bp	base pairs
g	gravitation unit
h	hour
kDa	kilodalton
L	liter
mM	millimolar
ng	nanogramm
µM	micromolar
°C	degree Celsius
v/v	volume per volume
w/v	weight per volume
%	percent
5HT	Serotonin
ACC	Anterior cingulate cortex
AD	Alzheimer's disease
Arc	Activity-regulated cytoskeleton associated
ASD	Autistic spectrum disorders
BSA	Bovine serum albumin
BrdU	5-Bromo-2'-deoxyuridine
CA	Cornu Ammonis
cAMP	Cyclic adenosine monophosphate
CamKII	Ca^{2+}/calmodulin-dependent protein kinase II
cDNA	complementary DNA
ChIP	Chromatin Immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation sequencing
cKO	Conditional knock-out
CNS	Central nervous systme
CREB	cAMP responsive element-binding protein
CTD	Carboxy terminal domain
DG	Dentate gyrus
DNA	Desoxyribonucleic acid
DTT	Dithiothreitol
EGTA	Ethylene glycol-bis-tetraacetic acid
EDTA	Ethylenediaminetetraacetic acid
ESC	Embryonic stem cells
EPM	Elevated plus maze

FACS	Fluorescent-activated cell sorting
FC	Fold change
FC	Fear conditioning
FDR	False discovery rate
GO	Gene Ontology
HDAC	Histone deacetylase
HDACi	Histone deacetylase inhibitor
HPRT1	Hypoxantine guanine phosphoribosyltransferase 1
IEG	Immediate Early Genes
KMT	Lysine methyltransferase
KDM	Lysine demethylase
LOF	Loss of function
LPA	Linear poly acrylamide
LTP	Long-term potentiation
NGS	Next generation sequencing
NOR	Novel object recognition
NMDA	N-methyl-D-aspartic acid
mRNA	messenger RNA
MAPK	Mitogen-activated protein kinase
MWM	Morris water maze
OF	Open field
padj	p-value adjusted
PBS	Phosphate-buffered saline
PFC	Prefrontoral cortex
PI	Protease inhibitor
PKA	Protein kinase A
PPI	Prepulse inhibition
PTM	Post-translational modifications
PT	Probe test
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	Ribonucleic acid sequencing
RNAPII	RNA polymerase II
RRM	RNA recognition motif
RT	Room temperature
SAHA	Suberoylanilide hydroxamic acid
SDS	Sodium dodecyl sulphate
SEM	Standard error of mean
SET	Su(var)3-9, Enhancer(zeste) and Trithorax
SNT	Novel object recognition
qRT-PCR	Quantitative real-time PCR

TBS	Tris-buffered saline
TF	Transcription factor
TST	Tail suspension test
TSS	Transcription start site
WDR5	WD repeat containing protein 5
WT	Wild-type

10. List of figures

11.   List of tables

12.   Acknowledgments

13.   Curriculum Vitae

<u>Personal</u>

24.01.1993 born to Svetlana Michurina and Oleg Michurin in Moscow, Russia
Married, daughter Matilda (2018) and daughter Zoë (expected early 2021)

<u>Education</u>

06/2017 – 11/2020	Ph.D. Student at Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Göttingen, Germany. Member of International Max Planck Research School in Genome Science
10/2015 – 05/2017	M.Sc. in Molecular Medicine, Georg-August University, Göttingen, Germany
09/2011 – 06/2015	B.Sc. in Biology, Lomonosov Moscow State University, Russia
2004 – 2011	Secondary education, Gymnasium 1517, Moscow, Russia

<u>Related conferences</u>

07/2018	11th FENS Forum of Neuroscience, Berlin, Germany (Poster presentation)
06/2018	Participation in the 68th Lindau Nobel Laureate Meeting as a young scientist
07/2017 – 08/2017	Developmental Neurobiology course, Okinawa Institute of Science and Technology (OIST), Okinawa, Japan (Short talk)
05/2017	EMBO Conference: Chromatin and Epigenetics, EMBL, Heidelberg, Germany

<u>Scholarships</u>

06/2018	Selected as "Young Scientist" participant for the 68th Lindau Nobel Laureate Meeting, Lindau, Germany
06/2017	Selected for "Developmental Neurobiology Course" by Okinawa Institute of Science and Technology (OIST), Okinawa, Japan

<u>Publications</u>

SETD1B enables hippocampal neurons the preeminent expression of learning and memory genes.
<u>Alexandra Michurina</u>, Sadman Sakib, Cemil Kerimogl, Dennis Manfred Krüger, Lalit Kaurani, Rezaul Islam, Julia Cha, Tonatiuh Pena Centeno, Andrea Kranz, Francis Adrian Stewart, Andre Fischer *(Under revision in "EMBO journal")*

H3 lysine 9 acetylation regulates neocortex expansion and folding by control of basal progenitor genesis.

Kerimoglu C., Pham L., Tonchev A., Narayan R., Pirouz M., Sokpor G., Nguyen H., Xie Y., Rosenbusch J., <u>Michurina A.,</u> Sadman S., Seong R.H., Capece V., Kovachev E., Angelova M., Minkov R., Stoyakova A., Staiger J., Fischer A., Tuoc T..
(Under review in "Science Advances")

Steroid structure base of agonistic and antagonistic action of pregna-D'-pentaranes in the transcriptional regulation of nuclear progesterone receptor B in yeast.
<u>Michurina A.O.,</u> Polikarpova A.V., Levina I.S., Kulikova L.E., Zavarzin I.V., Guseva A.A., Morozov I.A., Rubtsov P.M., Smirnova O.V., Shchelkunova T.A. *("Biochemistry" (Mosc.))*