bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Histone variant H2BE enhances chromatin accessibility in neurons to promote synaptic gene expression and long-term memory

Emily R. Feierman[1,3,4], Sean Louzon[2,3,4], Nicholas A. Prescott[6,7], Tracy Biaco[6,7], Qingzeng Gao[6], Qi Qiu[3], Kyuhyun Choi[5], Katherine C. Palozola[3,4], Anna J. Voss[1,3,4], Shreya Mehta[3,4], Camille Quaye[3,4], Katherine Lynch[3,4], Marc Fuccillo[5], Hao Wu[3,4], Yael David[6,7], Erica Korb[3,4]*

[1]Neuroscience Graduate Group, [2]Cell and Molecular Biology Graduate Group, [3]Department of Genetics, [4]Epigenetics Institute, [5]Department of Neuroscience, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, [6]Chemical Biology Program, Memorial Sloan Kettering Cancer Center, [7]Tri-institutional PhD Program in Chemical Biology, New York, NY, *corresponding author ekorb@pennmedicine.upenn.edu

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

Regulation of histones occurs through multiple mechanisms including exchange with histone variants. Unlike canonical histones, variants are replication-independent and therefore accumulate in post-mitotic cells such as neurons. While recent findings link variants to neurological and neuropsychiatric disorders, few are well studied in the context of the brain. H2BE is the single H2B variant found outside germline tissues, yet its expression and effects on chromatin remained unclear. We applied new tools including novel antibodies, biochemical assays, and sequencing approaches to reveal broad H2BE expression in the brain and its role in regulating chromatin structure, neuronal transcription, and mouse behavior. H2BE is enriched at promoters and enhances chromatin accessibility. We further identify a single amino acid driving these accessibility changes. Lastly, we show that H2BE is critical for synaptic gene expression and long-term memory. Together, these data reveal a novel mechanism linking histone variants to chromatin regulation and neuronal function underlying memory.

INTRODUCTION

Dynamic regulation of transcription is an integral part of a neuron's ability to form functional synapses and store information. Chromatin, the structure of DNA wrapped around histone proteins, allows cells in the brain to translate diverse environmental signals into a transcriptional profile that supports long-lasting memories and adaptive behavioral responses[1–3]. As a result, disruption of mechanisms that regulate chromatin and transcription has widespread effects and can lead to the onset of several neurological disorders, including autism spectrum disorder and related neurodevelopmental disorders[4–9] and neurodegeneration[10–13]. However, the molecular mechanisms through which histone proteins mediate complex neuronal functions such as memory formation are not fully understood.

Histone proteins are the central components of chromatin, affecting DNA compaction and thus, transcription. Two copies of each of the canonical histone proteins—H2A, H2B, H3, and H4—form an octamer around which

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

~146 base pairs of DNA are wound[14]. Chromatin is regulated by the addition of post-translational modifications onto histones and by expression of histone variants. Histone variants, which are encoded by separate genes, can substitute for the canonical forms, and are involved in regulation of many cellular processes and gene expression[15]. While canonical histone proteins are transcribed and translated during S phase of the cell cycle and deposited concurrently with DNA replication[16], histone variants are unique in that they can be synthesized throughout the cell cycle and continue to be generated and deposited into chromatin throughout the lifespan of terminally-differentiated cells[17,18]. This feature makes histone variants particularly critical to post-mitotic cells such as neurons. To date, several histone variants have been implicated in neuronal function and regulation of neuronal transcription[10,19–26]. Changes in the expression of histone variants have been observed in animals following exposure to diverse environmental factors including drugs of abuse[27] and social stress[23].

While variants of H2A and H3 are well characterized, few studies have investigated variants of histone H2B in animals. A recent study identified seven mammalian H2B variants[28]. Of those, only one known variant, H2BE, is expressed outside of germline tissues. The histone variant H2BE was discovered in the mouse main olfactory epithelium, where it affects neuronal longevity and olfactory neuron function[29]. This study concluded that H2BE was only expressed in olfactory epithelial neurons based on transcript expression data and that this remarkable mechanism was a unique feature of the atypical cell turnover that occurs within this sensory system. However, due to a lack of a specific antibody, analysis of H2BE protein expression could not be performed and limited sequencing approaches available at the time precluded analysis of the mechanisms through which H2BE regulates chromatin. While emerging evidence demonstrates the importance of histone variants in the brain and in neurological disorders[10,19–26], to the best of our knowledge no subsequent studies have been published in the intervening decade. Thus, whether H2BE is expressed outside of the olfactory system and how this unusual protein affects chromatin and cellular functions beyond olfaction has remained a mystery.

Here, we sought to define the function of H2BE in regulating chromatin and examine whether it functions beyond the olfactory epithelium. We generated a highly specific antibody for H2BE, allowing for comprehensive

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

analysis of H2BE throughout the brain. In addition, we combine biochemistry, mouse models, and animal behavior with combinatorial analysis of various next-generation sequencing datasets to characterize genomic localization of H2BE as well as its effects on chromatin structure, neuronal transcription, and cognition. We show that H2BE is expressed throughout the brain. In neurons, H2BE is expressed at gene promoters, conferring an open chromatin conformation. Further, we localize the major effects of H2BE on chromatin to a single amino acid. Loss of H2BE results in widespread compaction of chromatin and transcriptional disruption leading to synaptic dysfunction and impaired long-term memory. In summary, this work reveals a novel mechanism by which a histone variant contributes to the complex chromatin environment in the brain, and links single amino acid changes in histone proteins to regulation of neuronal transcription and cognition.

RESULTS

H2BE is expressed at promoters in cortical neurons

The histone variant H2BE differs from its canonical counterpart, H2B, by only five amino acids. The five amino acid changes in H2BE are distributed throughout the length of the protein, with one in each tail and three that lie within the globular domain (Fig 1a). A major challenge to studying this variant has been the lack of tools to directly and specifically detect endogenous H2BE protein. To overcome this, through a partnership with MilliporeSigma, we developed and tested a highly specific antibody against the first ten amino acids at the N-terminal of H2BE. This antibody has no cross-reactivity to canonical H2B or H3 tail peptides (Supp Fig 1a). To further confirm the specificity of the antibody, we obtained the previously reported H2BE knockout (KO) mouse[29], in which the coding sequence of *H2bc21* was replaced with a membrane-targeted mCherry reporter sequence (Supp Fig 1b). We confirmed the loss of H2BE transcript using RNA-sequencing of WT and KO primary cortical neurons (Supp Fig 1c). As previously reported[29], KO mice are viable and heterozygous crosses generate expected Mendelian ratios (Supp Fig 1d).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We extracted histones from adult mouse brain tissue and detected H2BE in the cortex, hippocampus, and cerebellum of WT mice (Fig 1b) indicating much more widespread expression than previously reported. Importantly, we observed no signal in any KO tissue demonstrating antibody specificity in tissue. We also measured H2BE in whole protein lysates from WT and KO cortical tissue and primary cultured neurons derived from E16.5 embryonic cortices and similarly observed H2BE expression in WT tissue and cultured neurons with no detectable signal in KOs (Supp Fig 1e-f). We further confirmed specificity using immunohistochemistry of WT and KO sections of adult mice (Supp Fig 1g) and found that H2BE co-localizes with the neuronal marker NeuN (Fig 1c), confirming the expression of H2BE in neurons *in vivo*. H2BE is also present in several tissues throughout the body but enriched in the brain relative to other tissues (Supp Fig 1h). We also found high levels of H2BE expression in the main olfactory epithelium consistent with previous findings originally identifying H2BE transcript as highly enriched in this tissue[29]. These data provide the first characterization of H2BE protein expression, and importantly, demonstrate that it is expressed outside the olfactory system with detectable levels throughout the mouse brain. This expands the relevance of H2BE beyond what was previously believed to be a variant exclusively expressed in sensory neurons of the olfactory epithelium. Further, while H2A and H3 variants are now well characterized in multiple cell types, this analysis of H2BE expression provides the first evidence of an H2B variant that is widely expressed in multiple tissues.

We next assessed the abundance of H2BE relative to the total pool of H2B in mouse cortex. H2BE was nearly absent in the cortex at birth, while its abundance increased over time and reached 0.24% of all H2B species by adulthood (Fig 1d; Supp Fig 1i). This low percentage is notable given that other histone proteins such as H3.3 and H2A.Z accumulate in post-mitotic cells such as neurons and become the dominant species[19,22]. However, histone variants can also have specific genomic enrichment patterns[9,19,22,30–33]. We therefore speculated that if H2BE is concentrated in specific genomic locations, this would allow it to exert significant effects on chromatin while only accounting for a small percentage of the total H2B protein pool. To test this, we performed CUT&Tag with sequencing[34] on primary cortical neurons from WT and KO E16.5 embryos following 12 days in culture. We

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

first validated our antibody for CUT&Tag by comparing enrichment of H2BE signal over background signal in KO cells (Supp Fig 1j) and by confirming a high concordance of peak sites between replicates (Supp Fig 1k). We found that H2BE expression is highly enriched in core promoter regions throughout the genome (Fig 1e-f; Supp Fig 1l) with most identified peaks near a transcription start site (TSS) (Fig 1g). Gene ontology analysis shows that H2BE is enriched specifically at genes related to important neuronal functions, including dendrite morphogenesis, synaptic vesicle endocytosis, and others (Supp Fig 1m). Further, H2BE enrichment is positively correlated with gene expression in WT neurons (Fig 1h-k), suggesting a relationship between H2BE enrichment and expression of neighboring genes. Together, these data indicate that H2BE is expressed at low levels relative to other H2B proteins in the mouse cortex but is highly enriched at core promoters, a regulatory region critical for the control of gene transcription.

H2BE promotes open chromatin via interactions at the histone-DNA interface

H2BE has 3 unique amino acids that lie within the globular domain. Of these, asparagine at position 75 (N75) and serine at position 97 (S97) are embedded within the nucleosome core, while the isoleucine at site 39 (I39) lies within the histone-DNA interface[14] (Fig 1a). Due to the location of these amino acids within the nucleosome core (N75 and S97), we tested whether H2BE affects nucleosome stability. We generated nucleosomes containing either two copies of canonical H2B or variant H2BE with canonical H2A, H3, and H4. Importantly, there were no changes in thermal stability between H2B- and H2BE-containing nucleosomes. Both species yielded two melting peaks at ~72 °C (Tm1) and ~85 °C (Tm2) corresponding to H2A/H2B dimer and H3/H4 tetramer dissociation, respectively (Fig 2a). This indicates that H2BE-specific amino acids do not affect nucleosome stability.

We next tested the ability of H2BE to compact DNA due to the H2BE-specific amino acid located at the DNA interface (I39). Nucleosome arrays were generated by combining H2B- or H2BE-containing octamers with 12x601 DNA and compaction was measured using atomic force microscopy[35]. We found that canonical H2B-containing arrays occupy a smaller surface area compared to H2BE-containing arrays (Fig 2b-c), without a

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

change in volume of the nucleosomes (Supp Fig 2a), demonstrating that the incorporation of H2BE confers a more open chromatin conformation relative to canonical H2B. To confirm these findings, we used Mg^{2+} precipitation assays[36] and found that H2BE arrays display a lower propensity to oligomerize in the presence of increasing concentrations of Mg^{2+} (Supp Fig 2b). Together, these assays revealed that H2BE nucleosome arrays are qualitatively more open compared to H2B arrays without affecting nucleosome stability. Further, these data indicate that the effect of H2BE on chromatin compaction is intrinsic to the amino acid sequence of the protein as the only components present in these *in vitro* assays include unmodified histones and DNA without other protein complexes or the addition of histone post translational modifications. Thus, the effect of H2BE in decreasing chromatin compaction occurs irrespective of posttranslational modifications and cellular signaling mechanisms and is an innate property of H2BE.

These biochemical assays were performed using synthetic nucleosomes containing 2 copies of either canonical H2B or H2BE. However, in a cellular context, it is possible that H2BE coexists with H2B in a nucleosome. We therefore assessed the effect of H2BE on chromatin accessibility in a physiological context using WT and KO primary cortical neurons paired with the assay for transposase-accessible chromatin with sequencing (ATAC-seq)[37]. We first performed an analysis of chromatin accessibility at accessible genomic regions including transcription start sites (TSSs) and enhancers in an unbiased fashion regardless of whether a peak was detected in WT neurons. We found a decrease in accessibility in KO neurons at TSSs (Supp Fig 2c-e). Of all TSSs with differential accessibility in KO, 99.98% (8,559 sites) were downregulated, while only 0.02% (2 sites) were upregulated. We observed a similar decrease in accessibility at enhancer regions (Supp Fig 2f-h), with no sites showing significantly increased accessibility. We also examined accessibility genome-wide, focusing on peaks detected in WT neurons regardless of genomic region, and again found that loss of H2BE results in marked decreased chromatin accessibility (Fig 2d-g; Supp Fig 2i-j).

To determine the relationship between H2BE enrichment and chromatin accessibility, we measured ATAC-seq signal at all H2BE peaks identified by CUT&Tag. H2BE peaks showed decreased accessibility in KO neurons (Fig

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2h-k). Of all H2BE peaks, 53.2% had decreased accessibility in KO, and zero had increased accessibility (Fig 2k). Further, we found that sites with the greatest enrichment of H2BE in WT neurons also showed the greatest magnitude change in ATAC signal between WT and KO, suggesting a direct correlation between H2BE occupancy and chromatin compaction upon H2BE loss (Fig 2l-n). We then subset H2BE peaks by genomic region and quantified the number of peaks with decreased accessibility in KO. We found that 77.5% H2BE-bound transcription start sites had decreased accessibility, while more modest changes were observed at H2BE peaks in the gene body (40.3%) and in intergenic H2BE peak regions (28.0%) (Supp Fig 2k-l). These data demonstrate a direct role of H2BE in modulating chromatin accessibility in neurons, particularly at transcription start sites.

We speculated that the amino acid difference found at I39 drives the effects on chromatin accessibility based on its location at the nucleosome-DNA interface (Fig 3a). To test the necessity and sufficiency of this single amino acid for the effect of H2BE on accessibility, we generated mutant nucleosomes in which the H2BE-specific isoleucine at site 39 is mutated to valine (H2BE-I39V), the amino acid found at that site on canonical H2B, and the converse where the valine at site 39 on canonical H2B was converted to isoleucine (H2B-V39I) (Fig 3b, Supp Fig 3a). We then measured thermal stability of the mutant nucleosomes and detected no differences between H2B, H2BE, H2BE-I39V, and H2B-V39I nucleosomes (Fig 3c). We then constructed chromatin arrays from these nucleosomes (Supp Fig 3b) and measured compaction with a Mg^{2+} precipitation assay. We found that arrays containing H2BE-I39V are less soluble than WT H2BE arrays in the presence of Mg^{2+}, indicating that H2BE-I39V are more compact than H2BE arrays (Fig 3d). Further, the V39I mutation on canonical H2B was sufficient to make H2B-V39I nucleosome arrays more soluble than arrays containing WT H2B (Fig 3e). Together, these data demonstrate that the I39 on H2BE is both necessary and sufficient for H2BE to promote open chromatin. Finally, we generated viral vectors containing WT H2BE or mutant H2BE-I39V and used these viruses to re-express H2BE in KO neurons (Supp Fig 3c). We then performed ATAC-seq to measure global changes in accessibility. As we previously showed, WT neurons have more open chromatin than KO neurons. We confirmed that these results were not simply due to viral infection or overexpression of a histone using a virus expressing a GFP reporter only

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

or canonical H2B, respectively. We also found that re-expressing WT H2BE was sufficient to open chromatin around the TSS. In contrast, neurons expressing I39V mutants were indistinguishable from KO neurons (Fig 3f-g). Taken together, these data illustrate that this amino acid is necessary for the effect of H2BE on chromatin compaction.

H2BE promotes synaptic gene expression

Given the profound changes detected in chromatin accessibility in KO neurons, we next performed RNA-sequencing (RNA-seq) on WT and KO primary cortical neurons to determine how loss of H2BE affects gene expression. We saw broad changes in gene expression in KO neurons, with 1,790 downregulated genes and 1,764 upregulated genes (Fig 4a). We note that *H2bc21* is called as upregulated in KO in our differential expression analysis because of signal mapping to the 3'UTR located downstream of the H2BE coding sequence (Supp Fig 1d). In the KO, the H2BE coding sequence is replaced by mCherry[29] allowing for continued expression of the UTR and has therefore been removed from further analysis. Downregulated genes are enriched for functions related to neuronal morphology and axon growth while upregulated genes are enriched for functions related to cellular metabolism and mitochondrial function (Fig 4b). Because many of the downregulated genes were related to synapse function, we also analyzed differentially expressed genes with SynGO[39] which confirmed disruption of critical synaptic functions particularly those related to regulation of synaptic transmission and signaling (Supp Fig 4a).

To better understand the properties of differentially expressed genes, we looked at basal expression levels in WT of genes that were downregulated, upregulated, or unaffected in KO. We found that genes that were downregulated in KO had higher expression in WT cells compared to both upregulated genes and non-differentially expressed genes (Fig 4c). These data suggest that highly expressed genes are more susceptible to downregulation upon loss of H2BE. Next, we compared RNA-seq data to ATAC-seq data to determine the relationship between changes in accessibility and gene expression. This analysis revealed that ~60% of downregulated genes show decreased accessibility at the transcription start site and/or proximal enhancers for

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that gene (Fig 4d). Further, genes downregulated in KO showed the greatest decrease in accessibility by ATAC-seq (Fig 4e). Taken together, these data suggest that downregulated genes are the most highly expressed in WT and show the greatest changes in accessibility. Downregulated genes that did not overlap with ATAC peaks, as well as upregulated genes may be due to secondary or compensatory mechanisms or other yet unknown functions of H2BE. Finally, overlap of CUT&Tag, ATAC, and RNA sequencing datasets identified genes with H2BE enrichment in WT neurons, changes in chromatin compaction in KO, and decreased transcription in KO, including *Slc8a2 and Tuba1a* (Fig 4f; Supp Fig 4b-e).

H2BE affects neuronal gene expression in the brain

We next wanted to understand the function of H2BE in the brain. Given that H2BE is expressed at negligible levels early in life, we predicted H2BE loss would not have substantial effects on cell identity. Rather we predicted it may affect gene expression within cell types as a consequence of the cumulative effects of H2BE loss into adulthood. We therefore performed single-nucleus Drop-seq[40] on cortical tissue from adult WT and KO mice. We identified 20 unique cell clusters, including 10 excitatory neuron clusters, 5 inhibitory neurons, one cluster of claustrum neurons, and 4 non-neuronal subtypes (Fig 5a). We did not detect notable changes in cluster identity across genotypes (Fig 5b, Supp Fig 5a), providing additional evidence that the effect of H2BE accumulates after development. Neuronal clusters were annotated according to the cortical layer they occupy, or—if unidentifiable by cortical layer—according to the gene most differentially expressed in that cluster relative to all other excitatory or inhibitory neuronal clusters. We identified differentially expressed genes in almost all neuronal clusters (Fig 5c; Supp Fig 5b), and also in non-neuronal cells, including astrocytes, oligodendrocytes, and microglia (Supp Fig 5b).

Of all excitatory neurons, the two clusters from cortical layer 2/3 had the most differential gene expression. Looking at the excitatory cluster with the greatest number of downregulated genes, from layer 2/3 pyramidal neurons (Ex_L2/3_2), we observed substantially more downregulated genes (114) than upregulated (21) which was true of most cell types (Fig 5d; Supp Fig 5c). These downregulated genes are involved in synaptic functions

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

such as cell-cell adhesion (*Nlgn1, Nrxn1*), pre- and postsynaptic membrane assembly (*Il1rapl1, Magi2*), and trans-synaptic signaling (*Grm7, Tenm2*) (Fig 5e). We next overlapped the down- and upregulated genes from the top 5 excitatory clusters, identifying 14 downregulated genes and 7 upregulated genes in common (Fig 5f-g; Supp Fig 5c). The inhibitory cluster with the most differentially expressed genes, inhibitory cluster with Phactr1 expression (Inh_Phactr1), included 284 down- and 18 upregulated genes (Fig 5h). Gene ontology analysis of these genes reveals similar annotations to those from layer 2/3 pyramidal neurons (Ex_L2/3_2), though with lower enrichment scores (Fig 5i). We overlapped the differential gene lists from the top 3 inhibitory clusters and identified 12 downregulated and 2 upregulated genes in common between these clusters (Fig 5j-k; Supp Fig 5d). Notably, there are 6 downregulated genes and 2 upregulated genes found in all top excitatory and inhibitory clusters shown here, indicating a pronounced role for H2BE in the expression of these genes across all neuronal cell types. In summary, H2BE is essential for the expression of synaptic genes in the adult mouse brain with similar disruptions to genes that regulate synaptic function observed across multiple cell types.

H2BE regulates synaptic strength and is required for long-term memory

These intersecting changes in synaptic gene expression raise the possibility that knocking out H2BE will lead to weakened or otherwise impaired synapses. To assess synaptic function, we performed field recordings of pharmacologically isolated excitatory postsynaptic potentials (fEPSP) from the stratum radiatum of the CA1 region in acute hippocampal slices. We electrically stimulated Schaffer collateral fibers with increasing input voltage, observing that the slope of the resulting fEPSP increases in both WT and KO. However, the response is significantly lower in KO, suggesting that synapses are weaker in KO in response to the same intensity stimulus (Fig 6a). To rule out that these differences in synaptic strength resulted from altered recruitment of Schaffer collaterals, we also measured fiber volley amplitude in response to the same stimuli and saw no difference between WT and KO (Fig 6b-c). Together, these data suggest that excitatory glutamatergic synaptic transmission is weakened in KO neurons supporting our finding that H2BE promotes the expression of synaptic genes.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Alterations in synaptic function can cause deficits in cognition. Therefore, we performed a battery of behavioral assays to analyze the role of H2BE in cognition and learning using an H2BE-KO mouse. First, we assessed olfactory function due to the high level of H2BE expression in the mouse olfactory tissues[29]. Previous work found that H2BE KO caused a deficit in olfactory discrimination in a learning paradigm. However, based on findings that H2BE is expressed throughout the brain (Fig. 1b, Supp Fig1f-h), it is unclear if these deficits are due to olfactory dysfunction or indicative of broader learning and memory deficits. We therefore first performed a test of olfactory habituation and dishabituation[41] through repeated exposure to the same odorant followed by exposure to novel odorants. Mice are expected to show less interest in odors after repeated presentations and increased interest when new odors are introduced. We found that KO mice have intact habituation and dishabituation to a variety of non-social and social olfactory cues (Fig 6d). There is a significant effect of sex on time spent interacting with the olfactory cues, with females showing greater interaction time with odor cues. However, within sex there was no effect of genotype on interaction with any of the cues (Supp Fig 6a). This behavior serves as an important control behavior and suggests intact olfactory processing. Regardless, in all tests (except for sociability measures) we removed all odor cues to remove potential confounds in the case that subtle deficits were not detected. Notably, these findings fit with published data demonstrating that KOs have intact odor evoked electrical responses in the olfactory epithelium[29] but raise the possibility that deficits may be due to other cognitive disruptions. In an open field arena, WT and KO mice traveled the same distance when allowed to explore freely, though there was an overall effect of sex, with male mice traveling a shorter distance during the trial compared to female mice (Fig 6e; Supp Fig 6b). There was no effect of H2BE KO on anxiety-like behavior as assessed by time spent in the center zone of an open field arena (Supp Fig 6c). These behavioral tests indicate KO mice have intact olfaction and no detectable mobility deficits or changes in anxiety and exploratory behaviors.

Due to the well-documented role of synaptic genes in social behaviors[42–44], we tested the sociability of WT and KO mice. Mice were placed in a three-chamber apparatus containing a rock or mouse in each of the outer

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

chambers and a neutral center chamber. Mice were allowed to explore freely, and we measured the amount of time spent interacting with the mouse and rock. A discrimination index was calculated for each mouse by finding the difference between interaction time with the mouse and rock as a proportion of total time spent interacting with the two. Both WT and KO mice show a preference for the mouse (as measured by a positive discrimination index) (Fig 6f). However, subtle changes were detected in average interaction duration with the mouse observed (Supp Fig 6d-f) suggesting largely intact sociability with minor changes in social behavior.

We next tested KO mice and control littermates in a T-maze to test working memory. In the T-maze, mice were placed at the entrance to a T-shaped maze and allowed to explore freely. Mice with intact working memory and normal exploratory behavior typically alternate between the three arms before returning to the first arm visited. WT and KO mice exhibited similar rates of 3-arm alternations with no effect of sex (Fig 6g), suggesting that exploratory behavior and working memory remain intact in KO mice.

Finally, we tested long-term memory in mice, specifically selecting tests that rely on proper synapse regulation[45–47]. In a novel object recognition test, mice are exposed to two identical objects in an arena and allowed to explore freely. On the following day, one of the objects is replaced with a novel object, and mice are returned to the arena. Due to a natural preference for novelty, mice that recognize the familiar object during testing spend more time exploring the novel object than the familiar object. We observed that mice lacking H2BE were not able to distinguish between a novel object and a familiar object presented 24 hours earlier with no effect of sex (Fig. 6h), suggesting that long-term memory is disrupted in KO mice. Importantly, there were no differences between groups in total time spent with both objects during learning or testing (Supp Fig 6g-h), indicating that deficits were not due to differences in object exploration.

We also assessed mice in two complementary tests of long-term memory—cued and contextual fear conditioning. In these assays, mice are exposed to two foot-shocks paired with a tone cue. For recall, mice are placed back in the chamber and freezing behavior is quantified as a readout of fear memory. In contextual

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

conditioning trials, mice are returned to the chamber or "context" 24 hours or 14 days post-training and freezing behavior is measured. For cued conditioning trials, mice are placed in an altered context and re-exposed to the tone cue alone (no foot-shock) 48 hours or 15 days post-training and freezing is measured during the tone presentations. In both cued and contextual recall, there was no effect of sex. However, KO mice exhibit less freezing behavior 14 (or 15) days post-training compared to their control littermates (Fig 6i-j). There was no difference in freezing behavior prior to the initial tone-shock pair ("Pre-shock") or in response to the first tone presentation ("Tone 1"), illustrating that there was nothing inherently fearful about the context or tone cue alone. Interestingly, in both contextual and cued conditioning, knockout mice showed a trend toward decreased freezing at 24 hour (contextual) or 48 hour (cued) post-training, however this difference did not reach significance. Together, these results demonstrate a critical role for H2BE in long-term memory, while olfaction, motor behavior, exploration, social behavior, and working memory are largely spared.

In summary, these findings support an overarching model in which H2BE promotes accessibility at critical and highly expressed neuronal genes to support synaptic responses and robust memory. Thus, when H2BE is lost, chromatin adopts a more closed state and synaptic gene expression is reduced, leading to synaptic dysfunction and impaired long-term memory (Fig 7).

DISCUSSION

Here, we show widespread expression of H2BE throughout the brain, with specific enrichment at gene promoters. The expression of H2BE results in greater chromatin accessibility which is reliant on a single amino acid that lies at the interface between the histone octamer and nucleosomal DNA. Further, we show that loss of H2BE results in broad changes in gene expression in both primary neuronal cultures and in neurons derived from the adult brain. Lastly, we show that H2BE is critical for long-term memory. In summary, we provide evidence of a novel example of how histone variants in the brain are critical to controlling chromatin structure, transcription, and memory.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The five amino acids that differ between H2BE and canonical H2B are distributed across the length of the protein rather than clustered together. This provides a unique opportunity to link a single amino acid site to the distinct properties of H2BE and sets H2BE apart from other better understood histone variants[48,49]. We found that one of these amino acids, I39, is critical for the effect of H2BE on chromatin. This places H2BE as a novel example of cases in which a single amino acid on a histone protein is linked to large-scale changes in chromatin regulation. This has recently been shown in two other variants, H2A.Z and H3.3, where a single amino acid is linked to neurological function and the onset of neurodevelopmental disorders[9,26]. Additionally, emerging research on oncohistones has revealed how single amino acid mutations in histone variants cause genome-wide changes that underlie various cancers[50–53]. Our work adds to this emerging theme in chromatin biology and provides a novel example of a variant-specific amino acid that mediates robust changes in chromatin structure.

This study expands upon previous work linking H2BE expression levels in olfactory tissues to olfactory neuron longevity[29]. In accordance with this study, we see substantial enrichment of H2BE in the mouse main olfactory epithelium (Supp Fig 1i). However, the development of an H2BE-specific antibody, H2BE nucleosomes, and chromatin arrays, along with the expansion of next-generation sequencing technologies has allowed us to define H2BE expression throughout the brain and its function in chromatin regulation. We found no overt deficits in basic olfactory function, consistent with previous results from electro-olfactograms showing no change in electrical responses to odor presentations in the olfactory epithelium of H2BE KO mice[29]. We therefore expanded on the behavioral characterization of H2BE KO mice, showing that long-term memory is impaired in KO mice, while other important behaviors such as motor function, exploratory behavior, social behavior, and working memory remain intact. Notably, none of our findings disagree with previous work but use newly available tools to provide an in-depth analysis of H2BE expression and function.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

With the finding that H2BE is expressed throughout the brain, several questions remain outstanding regarding the mechanism of H2BE expression and its function. We've shown that H2BE expression is enriched at promoters, yet the mechanism by which this enrichment occurs remains unknown. One possibility is that this expression pattern is mediated by chaperone proteins that specifically deposit H2BE at promoters. Alternatively, greater histone turnover at regions of open chromatin may result in the accumulation of H2BE at those sites which then serves to maintain accessible chromatin states. Further, while we observe large-scale changes in transcription upon knocking out H2BE, only a subset of downregulated genes (~60%) can be explained by a loss of accessibility. Further study is required to explain the remaining 40% of downregulated genes and the genes that are upregulated upon loss of H2BE. We speculate that some of these differential gene expression patterns may be explained by secondary or compensatory mechanisms, or by functions of H2BE that have not yet been discovered. We also found that H2BE is expressed in other cell types other than neurons in the brain. H2BE KO mice have no overt health issues or disruptions in basic motor and cognitive functions, suggesting that the memory deficits that we observed are not due to major health issues. However, it is plausible that H2BE has distinct roles in other cell types. Further, the role of H2BE in other cell types in the brain or other systems could emerge in other biological contexts, including disease, aging, or other perturbations.

Lastly, there is a growing body of clinical data linking histone variants to neurological disorders. To date H2BE has not been linked to neurological or neurodevelopmental disorders. However, a recent study of human histone genes suggests that H2BE is intolerant to mutations, suggesting that mutations in H2BE may be pathogenic[54]. We anticipate that with increased use of whole exome sequencing in patients with unexplained neurodevelopmental disorders, it is likely that future work will reveal mutations within H2BE, similar to disorders recently linked to other histone variant proteins[9,10]. In summary, this work reveals how histone variant H2BE contributes to the complex chromatin environment in the brain. This research furthers our understanding of mechanisms by which neurons control transcription and ultimately govern behavior. Further, this work

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

provides insight into novel mechanisms underlying memory and the increasing number of neurodevelopmental disorders linked to disruptions in chromatin regulation.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements: We thank Drs. Catherine Dulac and Stephen Santoro for sharing reagents and mouse lines. Mass spectrometry was performed with the CHOP-Penn Proteomics Core. Behavior procedures were performed with the Neurobehavior Testing Core at University of Pennsylvania. We acknowledge the use of the MSK Molecular Cytology Core Facility, supported by the NCI Cancer Center Support Grant (CCSG, P30 CA08748), Cycle for Survival, and the Marie-Josée and Henry R. Kravis Center for Molecular Oncology. E.H. is supported by NIH grant F31MH126576. S.L, K.P., and E.K., were supported by NIH grants 1DP2MH129985 and R00MH111836 and by the Klingenstein-Simons Fellowship from the Esther A. & Joseph Klingenstein Fund and the Simons Foundation, the Alfred P. Sloan Foundation Research Fellowship (FG-2020-13529), the Brain and Behavior Research Foundation NARSAD Young Investigator Award, and pilot funding from the Epigenetics Institute at the University of Pennsylvania. N.A.P. is supported by NIH grant F99CA264420.

Competing interest statement: The authors have no conflicts of interest.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

384 **REFERENCES**

385

386 1. Borrelli, E., Nestler, E. J., Allis, C. D. & Sassone-Corsi, P. Decoding the Epigenetic Language of Neuronal
387 Plasticity. *Neuron* **60**, 961–974 (2008).

388 2. Gallegos, D. A., Chan, U., Chen, L.-F. & West, A. E. Chromatin Regulation of Neuronal Maturation and
389 Plasticity. *Trends Neurosci* **41**, 311–324 (2018).

390 3. Zovkic, I. B. Epigenetics and memory: an expanded role for chromatin dynamics. *Current Opinion in*
391 *Neurobiology* **67**, 58–65 (2021).

392 4. Iossifov, I. *et al.* The contribution of de novo coding mutations to autism spectrum disorder. *Nature*
393 **515**, 216–221 (2014).

394 5. Parikshak, N. N. *et al.* Integrative Functional Genomic Analyses Implicate Specific Molecular Pathways
395 and Circuits in Autism. *Cell* **155**, 1008–1021 (2013).

396 6. O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo
397 mutations. *Nature* **485**, 246–250 (2012).

398 7. Iossifov, I. *et al.* De Novo Gene Disruptions in Children on the Autistic Spectrum. *Neuron* **74**, 285–299
399 (2012).

400 8. Thudium, S., Palozola, K., L'Her, É. & Korb, E. Identification of a transcriptional signature found in
401 multiple models of ASD and related disorders. *Genome Res.* **32**, 1642–1654 (2022).

402 9. Greenberg, R. S., Long, H. K., Swigut, T. & Wysocka, J. Single Amino Acid Change Underlies Distinct
403 Roles of H2A.Z Subtypes in Human Syndrome. *Cell* **178**, 1421-1436.e24 (2019).

404 10. Bryant, L. *et al.* Histone H3.3 beyond cancer: Germline mutations in *Histone 3 Family 3A and 3B* cause a
405 previously unidentified neurodegenerative disorder in 46 patients. *Sci. Adv.* **6**, eabc9207 (2020).

406 11. Berson, A., Nativio, R., Berger, S. L. & Bonini, N. M. Epigenetic Regulation in Neurodegenerative
407 Diseases. *Trends Neurosci* **41**, 587–598 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

12. Nativio, R. *et al.* An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease. *Nat Genet* **52**, 1024–1035 (2020).

13. Nativio, R. *et al.* Dysregulation of the epigenetic landscape of normal aging in Alzheimer's disease. *Nat Neurosci* **21**, 497–505 (2018).

14. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).

15. Martire, S. & Banaszynski, L. A. The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biol* **21**, 522–541 (2020).

16. Marzluff, W. F., Wagner, E. J. & Duronio, R. J. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet* **9**, 843–854 (2008).

17. Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. The human and mouse replication-dependent histone genes. *Genomics* **80**, 487–498 (2002).

18. Talbert, P. B. & Henikoff, S. Histone variants — ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11**, 264–275 (2010).

19. Maze, I. *et al.* Critical Role of Histone Turnover in Neuronal Transcription and Plasticity. *Neuron* **87**, 77–94 (2015).

20. Zovkic, I. B., Paulukaitis, B. S., Day, J. J., Etikala, D. M. & Sweatt, J. D. Histone H2A.Z subunit exchange controls consolidation of recent and remote memory. *Nature* **515**, 582–586 (2014).

21. Yang, Y. *et al.* Chromatin remodeling inactivates activity genes and regulates neural coding. *Science* **353**, 300–305 (2016).

22. Stefanelli, G. *et al.* Learning and Age-Related Changes in Genome-wide H2A.Z Binding in the Mouse Hippocampus. *Cell Rep* **22**, 1124–1131 (2018).

23. Lepack, A. E. *et al.* Aberrant H3.3 dynamics in NAc promote vulnerability to depressive-like behavior. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12562–12567 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

24. Narkaj, K. *et al.* Blocking H2A.Z Incorporation via Tip60 Inhibition Promotes Systems Consolidation of Fear Memory in Mice. *eNeuro* **5**, (2018).

25. Dunn, C. J. *et al.* Histone Hypervariants H2A.Z.1 and H2A.Z.2 Play Independent and Context-Specific Roles in Neuronal Activity-Induced Transcription of *Arc/Arg3.1* and Other Immediate Early Genes. *eNeuro* **4**, ENEURO.0040-17.2017 (2017).

26. Khazaei, S. *et al.* Single substitution in H3.3G34 alters DNMT3A recruitment to cause progressive neurodegeneration. *Cell* **186**, 1162-1178.e20 (2023).

27. Wimmer, M. E. *et al.* H3.3 Barcoding of Nucleus Accumbens Transcriptional Activity Identifies Novel Molecular Cascades Associated with Cocaine Self-administration in Mice. *J. Neurosci.* **39**, 5247–5254 (2019).

28. Raman, P. *et al.* Novel Classes and Evolutionary Turnover of Histone H2B Variants in the Mammalian Germline. *Mol Biol Evol* **39**, msac019 (2022).

29. Santoro, S. W. & Dulac, C. The activity-dependent histone variant H2BE modulates the life span of olfactory neurons. *eLife* **1**, e00070 (2012).

30. Gamble, M. J., Frizzell, K. M., Yang, C., Krishnakumar, R. & Kraus, W. L. The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes Dev* **24**, 21–32 (2010).

31. Sun, Z. *et al.* Transcription-associated histone pruning demarcates macroH2A chromatin domains. *Nat Struct Mol Biol* **25**, 958–970 (2018).

32. Seo, J. *et al.* Genome-wide profiles of H2AX and γ-H2AX differentiate endogenous and exogenous DNA damage hotspots in human cells. *Nucleic Acids Res* **40**, 5965–5974 (2012).

33. Palmer, D. K., O'Day, K., Trong, H. L., Charbonneau, H. & Margolis, R. L. Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 3734–3738 (1991).

34. Kaya-Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**, 1930 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

35. Krzemien, K. M., Beckers, M., Quack, S. & Michaelis, J. Atomic force microscopy of chromatin arrays reveal non-monotonic salt dependence of array compaction in solution. *PLoS ONE* **12**, e0173459 (2017).

36. Schwarz, P. M., Felthauser, A., Fletcher, T. M. & Hansen, J. C. Reversible Oligonucleosome Self-Association: Dependence on Divalent Cations and Core Histone Tail Domains. *Biochemistry* **35**, 4009–4015 (1996).

37. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology* **109**, (2015).

38. Arimura, Y. *et al.* Cancer-associated mutations of histones H2B, H3.1 and H2A.Z.1 affect the structure and stability of the nucleosome. *Nucleic Acids Res* **46**, 10007–10018 (2018).

39. Koopmans, F. *et al.* SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* **103**, 217-234.e4 (2019).

40. Hu, P. *et al.* Dissecting Cell-Type Composition and Activity-Dependent Transcriptional State in Mammalian Brains by Massively Parallel Single-Nucleus RNA-Seq. *Molecular Cell* **68**, 1006-1015.e7 (2017).

41. Yang, M. & Crawley, J. N. Simple Behavioral Assessment of Mouse Olfaction. *CP Neuroscience* **48**, (2009).

42. Spratt, P. W. E. *et al.* The Autism-Associated Gene Scn2a Contributes to Dendritic Excitability and Synaptic Function in the Prefrontal Cortex. *Neuron* **103**, 673-685.e5 (2019).

43. Yan, Y. *et al.* ASH1L haploinsufficiency results in autistic-like phenotypes in mice and links Eph receptor gene to autism spectrum disorder. *Neuron* **110**, 1156-1172.e9 (2022).

44. O'Connor, E. C., Bariselli, S. & Bellone, C. Synaptic basis of social dysfunction: a focus on postsynaptic proteins linking group- I m G lu R s with AMPAR s and NMDAR s. *Eur J of Neuroscience* **39**, 1114–1129 (2014).

45. Clarke, J. R., Cammarota, M., Gruart, A., Izquierdo, I. & Delgado-García, J. M. Plastic modifications induced by object recognition memory processing. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2652–2657 (2010).

46. Choi, D. I. *et al.* Synaptic correlates of associative fear memory in the lateral amygdala. *Neuron* **109**, 2717-2726.e3 (2021).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

47. Sheynikhovich, D., Otani, S., Bai, J. & Arleo, A. Long-term memory, synaptic plasticity and dopamine in rodent medial prefrontal cortex: Role in executive functions. *Front. Behav. Neurosci.* **16**, 1068271 (2023).

48. Hake, S. B. & Allis, C. D. Histone H3 variants and their potential role in indexing mammalian genomes: the 'H3 barcode hypothesis'. *Proc Natl Acad Sci U S A* **103**, 6428–6435 (2006).

49. Giaimo, B. D., Ferrante, F., Herchenröther, A., Hake, S. B. & Borggrefe, T. The histone variant H2A.Z in gene regulation. *Epigenetics & Chromatin* **12**, 37 (2019).

50. Nacev, B. A. *et al.* The expanding landscape of 'oncohistone' mutations in human cancers. *Nature* **567**, 473–478 (2019).

51. Bagert, J. D. *et al.* Oncohistone mutations enhance chromatin remodeling and alter cell fates. *Nat Chem Biol* **17**, 403–411 (2021).

52. Lewis, P. W. *et al.* Inhibition of PRC2 Activity by a Gain-of-Function H3 Mutation Found in Pediatric Glioblastoma. *Science* **340**, 857–861 (2013).

53. Stafford, J. M. *et al.* Multiple modes of PRC2 inhibition elicit global chromatin alterations in H3K27M pediatric glioma. *Sci. Adv.* **4**, eaau5935 (2018).

54. Lubin, E., Bryant, L., Aicher, J., Li, D. & Bhoj, E. Analysis of histone variant constraint and tissue expression suggests five potential novel human disease genes: H2AFY2, H2AFZ, H2AFY, H2AFV, H1F0. *Hum Genet* **141**, 1409–1421 (2022).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE LEGENDS

Figure 1. H2BE is expressed at promoters in neurons. (a) H2B and H2BE amino acid sequence. Variant amino acids are highlighted in red. (b) Representative western blot for H2BE and H2B in brain tissue from 3–5-month-old WT and KO mice (ctx = cortex; hpc = hippocampus; crb = cerebellum). H2B serves as loading control. (c) Representative immunofluorescent images of cortical tissue from 3-month-old mice stained with H2BE, neuronal marker NeuN, and nuclear marker DAPI. Scale bar = 20µm. (d) LC/MS-MS quantification of H2BE in cortical tissue across age (n = 6-10 biological replicates per age with 3-5 from both male and female combined; data represents mean +/- SEM; one-way ANOVA with Dunnett's multiple comparisons test; ***p<.001, ****p<.0001). No difference was found between male and female mice. (e) Metaplot of H2BE CUT&Tag profiling in WT primary cortical neurons. Plot shows read counts per million mapped reads (RPM) between the transcription start site (TSS) and transcription end site (TES) +/- 2kb (n = 3 biological replicates per genotype). (f) Genomic distribution of H2BE enrichment sites relative to the entire mouse genome (Chi-square test; *p<.05, ***p<.001). (g) Distribution of H2BE enrichment sites (top) and the entire genome binned by 500bp (bottom) relative to the nearest TSS. (h) Metaplot comparison of H2BE CUT&Tag at genes according to expression level. "Not expressed" was defined as genes with average normalized read counts <3 by RNA-sequencing (n = 8,150 genes). Remaining genes were binned into three equally sized groups by average normalized read counts (n = 6,028 genes per group). KO shows average across all genes regardless of expression. Plot shows read counts per million mapped reads (RPM) around the transcription start site (TSS) +/- 2kb. (n = 3 biological replicates per genotype). (i) H2BE CUT&Tag peak scores assigned by MACS3 by expression level as defined above (one-way ANOVA with post-hoc pairwise t-tests with Bonferroni correction; ****p<.0001, n.s. = not significant). (j-k) RNA-sequencing (top) and CUT&Tag (bottom) gene tracks for representative mid-expressed (*Ankrd27*) and low-expressed (*E2f7*) genes.

Figure 2. H2BE promotes open chromatin. (a) Thermal stability assay of H2B- and H2BE-containing nucleosomes yields two melting peaks (Tm) corresponding to dimer and tetramer dissociation, respectively (n=3 biological

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

525 replicates per group; Welch's t-tests with Holm-Šídák correction). (b) Representative micrographs of H2B- or
526 H2BE-containing nucleosome arrays. Scale bar = 100µm. (c) Volume/surface area quantification of compaction
527 from atomic force microscopy of H2B or H2BE nucleosome arrays (n = 60 H2B arrays, 40 H2BE arrays; Welch's
528 t-test; ****p<.0001). Surface area = Total # of pixels occupied by array. Volume = Total pixel intensity of array.
529 (d) Heat map and (e) metaplot comparison of ATAC-seq average signal from WT and KO primary cortical neurons
530 at all WT peaks. Plot shows read counts per million mapped reads (RPM) at all measured peaks +/- 1kb (n = 3
531 WT biological replicates, 4 KO biological replicates). (f) Normalized ATAC-seq read counts at all ATAC peak sites
532 detected in WT neurons (unpaired t-test; ****p<.0001). (g) Volcano plot showing differential chromatin
533 accessibility between WT and KO at WT ATAC peaks. Blue = peaks that decrease in KO; red = peaks that increase
534 in KO; gray = peaks below significance cut-off of FDR<.05. (h) Heat map and (i) metaplot comparison of ATAC-
535 seq average signal from WT and KO primary cortical neurons at all H2BE binding sites identified by CUT&Tag.
536 Plot shows read counts per million mapped reads (RPM) at all measured peaks +/- 1kb (n = 3 WT biological
537 replicates, 4 KO biological replicates). (j) Normalized ATAC-seq read counts at H2BE binding sites (unpaired t-
538 test; ****p<.0001) (k) Volcano plot showing differential chromatin accessibility at H2BE binding sites. Blue =
539 peaks that decrease in KO; red = peaks that increase in KO; gray = peaks below significance cut-off of FDR<.05.
540 (l) Correlation between H2BE reads in WT neurons and changes in ATAC-seq signal between WT and KO
541 (Spearman correlation; $\rho = 0.74$; $p < 2.2x10^{-16}$). (m-n) ATAC-seq gene tracks for representative mid-expressed
542 (*Ankrd27*) and low-expressed (*E2f7*) genes.

543 **Figure 3. Single amino acid at the histone-DNA interface drives H2BE effect on chromatin.** (a) Nucleosome
544 surrounded by DNA shows that the H2BE-specific amino acid at site 39 lies at the histone-DNA binding interface.
545 Cyan=H2B; red=V39; white=H2A/H3/H4. (b) Schematic of canonical H2B, WT H2BE, and mutant H2BE
546 sequences. (c) Thermal stability assay of H2BE and H2BE-I39V nucleosomes (n = 3 biological replicates per group;
547 2-way ANOVA). (d) Mg^{2+} precipitation assay of chromatin arrays containing WT H2BE or H2BE-I39V (n=4 per
548 group) (e) Mg^{2+} precipitation assay of chromatin arrays containing canonical H2B or H2B-V39I (n=8 per group)

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(f) ATAC-seq average signal at all genes (n = 3-8 biological replicates per condition). Plot shows read counts per million mapped reads (RPM) around the transcription start site (TSS) +/- 2kb. (g) ATAC-seq gene tracks for *Ankrd27*.

Figure 4. H2BE promotes synaptic gene expression. (a) Volcano plot showing differentially expressed genes (FDR<.05, absolute fold change > 1.25) between WT and KO primary cortical neurons (n = 4 biological replicates per genotype [2 male, 2 female]). Blue = downregulated; red = upregulated. (b) Gene ontology enrichment analysis of down- (left) and upregulated (right) genes. (c) Normalized RNA-seq read counts in WT neurons for genes downregulated, upregulated or unchanged in KO (one-way ANOVA with post-hoc pairwise t-tests with Bonferroni correction; ****p<.0001, n.s. = not significant; DEG = differentially expressed genes). (d) Pie chart showing which downregulated genes have a decreased ATAC-seq peak in KO at a proximal enhancer and/or promoter for that gene in WT. (e) ATAC-seq fold change (KO/WT) for downregulated, upregulated, and unchanged genes in KO (one-way ANOVA with post-hoc pairwise t-tests with Bonferroni correction; ****p<.0001, n.s. = not significant; DEG = differentially expressed genes). (f) CUT&Tag, ATAC-seq, and RNA-seq genome browser tracks at the *Slc8a2* locus in WT and KO cortical neurons. See Supplemental Figure 4 for additional example tracks.

Figure 5. H2BE is required for appropriate synaptic function in the brain. (a) UMAP of single-nucleus transcriptomic profiles from adult (2-4 months) mouse cortices (n = 3 biological replicates per genotype). Colors represent cluster identity. Ex = excitatory neurons; Inh = inhibitory neurons. (b) UMAP clustering with dots representing each nucleus colored by genotype. Relative distribution of WT and KO cells found by cluster shown in Supplemental Figure 5. (c) Number of differentially expressed genes within each cluster for all excitatory and inhibitory neuronal clusters. (d) Volcano plot showing differential expression from WT and KO in a representative cluster of layer 2/3 cortical excitatory neurons (Ex_L2/3_2). Blue = downregulated; red = upregulated. (e) Gene ontology enrichment analysis of downregulated genes in cluster Ex_L2/3_2. Upregulated shown in Supplemental Figure 5. (f-g) Overlap of differentially expressed down- (f) and upregulated (g) genes

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

from the 5 excitatory neuron clusters with the greatest number of differentially expressed genes. (h) Volcano plot showing differential expression from WT and KO cortical inhibitory neurons in a representative cluster, Inh_Phactr1. Blue = downregulated; red = upregulated. (i) Gene ontology enrichment analysis of downregulated genes in cluster Inh_Phactr1. (j-k) Overlap of differentially expressed down- (j) and upregulated (k) genes from the 3 inhibitory neuron clusters with the greatest number of differently expressed genes.

Figure 6. H2BE is required for long-term memory. (a) Input-output curves from field extracellular postsynaptic potentials and (b) afferent fiber volleys following stimulation of hippocampal Schaffer collaterals. The line represents the inferred responses through linear regression (WT r^2=0.6748, KO r^2=0.5006), while each dot corresponds to the averaged slope from each experiment based on stimulus intensity. (n = 11/13 from 5/4 animals [2-3 males and females per genotype]; two-way ANOVA [stimulus intensity x genotype] with Šídák's multiple comparisons test; *p<.05, **p<.01, ***p<.001). (c) Representative traces from WT and KO. (d) Interaction time with a scented cotton swab during presentations of water, two non-social odors (almond, vanilla), and social odors (same sex, opposite sex) for three consecutive trials per scent. Line and shaded area represent mean +/- SEM (n = 20 WT mice [8 male, 12 female], 24 KO mice [10 male, 14 female]; two-way ANOVA [genotype x sex] for each odorant. No differences were detected between genotypes. Significant differences were detected between sex (graphed separately in Supplemental Figure 6a). (e) Distance traveled in an open field arena during a 5-minute trial (n = 19 WT mice [9 male, 10 female], 19 KO mice [11 male, 8 female]; two-way ANOVA [genotype x sex]). Sexes graphed separately in Supplemental Figure 6b. (f) Discrimination index for interaction time between a mouse and a rock during a 3-chamber sociability SEM (n = 20 WT mice [8 male, 12 female], 24 KO mice [10 male, 14 female]; two-way ANOVA [genotype x sex]). DI = ([mouse – rock] / [mouse + rock]). No effects of sex were observed. (g) Percent correct trials of spontaneous alternation in a T-maze trial (n = 19 WT mice [9 male, 10 female], 19 KO mice [11 male, 8 female]; two-way ANOVA [genotype x sex]). No effects of sex were observed. (h) Discrimination index during a novel object recognition task (n = 19 WT mice [9 male, 10 female], 19 KO mice [11 male, 8 female]; two-way ANOVA [genotype x sex]; *p<.05)]). DI = ([novel – familiar]

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

/ [novel + familiar]). No effects of sex were observed. (i-j) Percent freezing during cued (i) and contextual (j) fear conditioning (n = 20 WT mice [8 male, 12 female], 24 KO mice [10 male, 14 female]; two-way ANOVA [genotype x sex] for each phase of testing; **p<.01).

Figure 7. Model of H2BE function. In a WT mouse, H2BE is expressed at promoters throughout the genome, where its expression causes chromatin to open and higher levels of synaptic gene expressions. This supports healthy synapses and proper neuronal function. In contrast, upon loss of H2BE chromatin remains in a closed state, there is loss of synaptic gene expression, synaptic dysfunction, and impaired long-term memory.

Supplemental Figure 1. (a) Dot blot analysis of H2BE antibody against N terminal peptides for H2BE, H2B, and H3. (b) Schematic of the H2BE-KO/Gap43-mCherry KI allele as generated by Santoro & Dulac (2012). (c) RNA-sequencing heat map of WT and KO primary cortical neurons at the *H2bc21* locus shows no reads in the H2BE coding sequence in KO neurons. (d) Pie chart of progeny from $H2BE^{+/-}$ x $H2BE^{+/-}$ crosses show expected Mendelian ratios (n = 126 mice from 26 litters). (e) Representative western blot for H2BE and H2B in cortical tissue. WT and KO demonstrates anti-H2BE antibody specificity. H2B serves as loading control. (f) Representative western blot for H2BE and H2B in cortical tissue from WT and KO primary cortical neurons demonstrates anti-H2BE antibody specificity. H2B serves as loading control. (g) Representative immunofluorescent images of cortical tissue from mice stained with H2BE and the nuclear marker DAPI demonstrates anti-H2BE antibody specificity. Scale bar = 30µm. (h) Representative western blot for H2BE and H2B in various tissues from WT and KO mice. (i) Representative MS/MS spectra of the H2B and H2BE peptides. (j) Metaplot comparison of CUT&Tag sequencing average signal from WT and KO primary cortical neurons demonstrates anti-H2BE antibody specificity. Plot shows read counts per million mapped reads (RPM) around the transcription start site (TSS) +/- 2kb (n = 1 biological replicate per genotype). (k) Overlap analysis of H2BE peaks detected by CUT&Tag from 3 biological replicates. (l) Genomic distribution of H2BE peaks detected by CUT&Tag. (m) Gene ontology enrichment analysis of genes where H2BE is enriched.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Figure 2. (a) Volume of H2B or H2BE nucleosome arrays. (n = 60 H2B arrays, 40 H2BE arrays; Welch's t-test). Volume = total pixel intensity of each array. (b) Mg^{2+} precipitation assay for chromatin arrays containing canonical H2B or variant H2BE (n = 4 per group). (c) Metaplot comparison of ATAC-seq average signal from WT and KO primary cortical neurons at all transcription start sites. Plot shows read counts per million mapped reads (RPM) around the transcription start site (TSS) +/- 2kb (n = 3 WT biological replicates, 4 KO biological replicates). (d) Volcano plot showing differential chromatin accessibility between WT and KO at transcription start sites. Blue = peaks that decrease in KO; red = peaks that increase in KO; gray = peaks below significance cut-off of FDR<.05. (e) Normalized ATAC-seq read counts at transcription start sites with significantly different accessibility in WT (left) or KO (right). (unpaired t-test; ****p<.0001). (f) Metaplot comparison of ATAC-seq average signal from WT and KO primary cortical neurons at all enhancers identified in brain tissue. Plot shows read counts per million mapped reads (RPM) at all measured enhancers +/- 1kb (n = 3 WT biological replicates, 4 KO biological replicates). (g) Volcano plot showing differential chromatin accessibility between WT and KO at enhancers. Blue = peaks that decrease in KO; red = peaks that increase in KO; gray = peaks below significance cut-off of FDR<.05. (h) Normalized ATAC-seq read counts at transcription start sites with significantly different accessibility in WT (left) or KO (right). (unpaired t-test; ****p<.0001). (i) Summary plot (top) of peaks with decreased (purple) or increased (blue) accessibility in KO with heatmap (bottom) showing each TSS. The majority of TSS regions show decreased accessibility in KO. (j) Normalized ATAC-seq read counts at all WT ATAC-seq peaks with significantly different accessibility in WT (left) or KO (right). (unpaired t-test; ****p<.0001). (k) Percent of all detected ATAC-seq peaks within H2BE binding sites at promoters, within the gene body, or intergenic that are significantly decreased in KO. (l) Summary plot (top) and heatmap (bottom) of CUT&Tag reads and ATAC-seq reads at all H2BE binding sites, partitioned by genomic region.

Supplemental Figure 3. (a) Mononucleosome electrophoretic mobility shift assay using native polyacrylamide gel electrophoresis to validate nucleosome assembly, stained using ethidium bromide to visualize DNA. (b) Chromatin array electrophoretic mobility shift assay using native agarose-polyacrylamide gel electrophoresis

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(APAGE) to validate chromatin array assembly, stained with ethidium bromide. (c) Western blot for H2BE and H2B in protein lysates from WT and KO neurons, and KO neurons infected with viruses expressing WT H2BE, H2BE-I39V, canonical H2B, or a GFP control.

Supplemental Figure 4. (a) Synaptic gene ontology enrichment analysis by SynGO of genes where H2BE is enriched. (b) Overlap analysis of genes with H2BE enrichment, decreased accessibility in KO, and decreased transcription in KO. (c) CUT&Tag, ATAC-seq, and RNA-seq genome browser tracks at the *Tuba1a* locus in WT and KO cortical neurons. (d-e) RNA-seq tracks from WT and KO primary cortical neurons at (d) *Ankrd27* and (e) *E2f7* previously shown in Figs 1 and 2.

Supplemental Figure 5. (a) Analysis of the difference between the proportion of WT and KO nuclei in each cluster. Data represents the fold change in WT vs KO nuclei in the cluster and a confidence interval for the magnitude difference. Clusters with fold difference >1.8 and FDR<.05 highlighted in pink. Ex_Angpt1 cluster contains only small number of nuclei with 84 nuclei in WT and 45 in KO. (b) Number of differentially expressed genes within glial clusters and claustral neurons from single-nucleus RNA-seq. (c-d) CUT&Tag, ATAC-seq, and RNA-seq genome browser tracks at the *Tuba1a* locus in WT and KO cortical neurons. Note that these gene ontology analyses produced the same results because of a high concordance between neuronal clusters and by nature of the small gene lists produced by snRNA-seq.

Supplemental Figure 6. (a) Olfactory habituation/dishabituation data separated by sex. Line and shaded area represent mean +/- SEM (n = 20 WT mice [8 male, 12 female], 24 KO mice [10 male, 14 female]). (b) Distance traveled in an open field arena during a 5-minute trial separated by sex (n = 19 WT mice [9 male, 10 female], 19 KO mice [11 male, 8 female]; two-way ANOVA [genotype x sex]; main effect of sex; ***p<.001). (c) Time spent in center zone of an open field arena during a 5-minute trial sex (n = 19 WT mice [9 male, 10 female], 19 KO mice [11 male, 8 female]; two-way ANOVA [genotype x sex]). No effects of sex were observed. (d) Total interaction time with the cylinders containing the mouse or rock. (n = 20 WT mice [8 male, 12 female], 24 KO

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

667 mice [10 male, 14 female]; two-way ANOVA [genotype x sex]). No effects of sex were observed. (e-f) Mean
668 interaction time per visit to (e) mouse or (f) rock cylinder (n = 20 WT mice [8 male, 12 female], 24 KO mice [10
669 male, 14 female]; two-way ANOVA [genotype x sex]; *p<0.05). No effects of sex were observed. (g-h) Total
670 interaction time with both objects during (g) training and (h) testing (n = 19 WT mice [9 male, 10 female], 19 KO
671 mice [11 male, 8 female]; two-way ANOVA [genotype x sex]). No effects of sex were observed.

672

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

673 **METHODS**

674 **Mice**

675 An H2BE-KO mouse was generated as described previously[1]. In brief, the endogenous H2BE CDS was replaced
676 with a membrane-targeted mCherry reporter sequence in C57Bl/6 mice. All mice were housed in a 12-hour light-
677 dark cycle and fed a standard diet. All experiments were conducted in accordance with and approval of the
678 IACUC.

679 **Histone extraction**

680 Snap-frozen brain tissues were homogenized in nuclear isolation buffer (NIB; 15mM Tris-HCl pH 7.5, 60mM KCl,
681 1mM $CaCl_2$, 15mM NaCl, 5mM $MgCl_2$, 250mM sucrose supplemented by protease inhibitor [Roche
682 04693124001], phosphatase inhibitor [Roche 04906837001], 1mM DTT, 1mM PMSF) + 0.3% NP-40 using a pre-
683 chilled dounce and pestle. Sample were then incubated on ice for 5 minutes, and centrifuged for 5 minutes at
684 1000g at 4°C. The pellet was washed 1x in NIB without NP-40 and centrifuged for 5 minutes at 1000g at 4°C. The
685 pellet was then resuspended in 1mL cold 0.4N H_2SO_4 and rotated overnight at 4C. Following the overnight
686 incubation, samples were pelleted for 5 minutes at 10,000g at 4°C and the supernatant was transferred to a
687 new tube. Trichloroacetic acid was added to 25% by volume, and the cells were left on ice at 4°C overnight. Cells
688 were again pelleted for 5 minutes at 10,000g at 4°C, and the supernatant was discarded. Samples were
689 centrifuged for 5 minutes at 16,000g at 4°C and supernatant was discarded. The pellet was washed 2X with ice-
690 cold acetone. After the second wash, samples were air-dried. The pellet was resuspended in molecular biology-
691 grade H_2O, sonicated in a Biorupter for 10 min, and then incubated at 50°C in a thermomixer shaking at
692 1000rpm. Samples were centrifuged for 10 minutes at 16,000g at 4°C and the supernatant containing the
693 histone fraction was collected.

694 **Western blotting**

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Protein lysates or histone samples were mixed with 5X Loading Buffer (5% SDS, 0.3M Tris pH 6.8, 1.1mM Bromophenol blue, 37.5% glycerol), boiled for 10 minutes, and cooled on ice. Protein was resolved by 16% Tris-glycine SDS-PAGE, followed by transfer to a 0.45-µm PVDF membrane (Sigma-Aldrich IPVH00010) for immunoblotting. Membranes were blocked for 1 hour at RT in 5% milk in 0.1% TBST and probed with primary antibody overnight at 4C. Membranes were incubated with secondary antibody for 1 hour at RT. Antibodies are shown in Supplemental Table.

Immunohistochemistry

Brain sectioning

Mice were anesthetized with isoflurane and transcardially perfused with 4% paraformaldehyde. Brains were collected, rinsed with ice-cold PBS, placed in 15% sucrose solution at 4°C and allow to settle to the bottom. Brains were then transferred to a 30% sucrose and again allowed to settle to the bottom. Brains were then cryopreserved in O.C.T. Brains were coronally sectioned at 10µm using a cryostat and mounted on Superfrost Plus slides.

Staining

Sections were washed 3X at RT in ash buffer (0.3% Triton X-100 in 1X PBS), then blocked at RT for 1h in Blocking buffer (5% normal goat serum, 0.3% Triton X-100, 1% BSA in 1X PBS). Sections were washed 3X, then incubated with primary antibody (diluted in Antibody Dilution buffer [0.3% Triton X-100, 1% BSA inX PBS]) in a humid chamber overnight at 4°C. Sections were washed 3X, then incubated with secondary antibody (diluted in Antibody Dilution buffer) in a humid chamber for 2h at RT. After three additional washes, sections were incubated with 1ug/uL DAPI (Thermo Scientific 62247) for 10 min at RT. Coverslips were mounted onto slides using Prolong Gold mounting medium. Slides were imaged on a Leica DM IL microscope with a 40X objective.

Quantitative mass spectrometry

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In-Solution Digestion

Samples were solubilized and digested per the S-Trap Micro (Protifi) manufacturer's protocol[2]. Briefly, samples were solubilized in 50 µL of extraction buffer containing 5% sodium dodecyl sulfate (SDS, Affymetrix), 50mM TEAB (pH 8.5, Sigma), and protease inhibitor cocktail (Roche cOmplete, EDTA free), reduced in 5mM TCEP (Thermo), alkylated in 20mM iodoacetamide (Sigma), then acidified with phosphoric acid (Aldrich) to a final concentration of 1.2%. Samples were diluted with 90% methanol (Fisher) in 100 mM TEAB, then loaded onto an S-trap column and washed three times with 50/50 chloroform/methanol (Fisher) followed by three washes of 90% methanol in 100 mM TEAB. A 1:10 ratio (enzyme: protein) of Trypsin (Promega) and LysC (Wako) suspended in 20µL 50mM TEAB was added, and samples were digested for 1.5 hours at 47 °C in a humidity chamber. After incubation, peptides were eluted with an additional 40 µL of 50 mM TEAB, followed by 40 µL of 0.1% trifluoroacetic acid (TFA) (Pierce) in water, and finally 40 µL of 50/50 acetonitrile:water (Fisher) in 0.1% TFA. Eluates were combined and desalted directly using Phoenix peptide cleanup kit (PreOmics) per manufactures protocol, dried by vacuum centrifugation and reconstituted in 0.1% TFA containing iRT peptides (Biognosys, Schlieren, Switzerland).

Target Assay Development

Parallel Reaction Monitoring (PRM) assay[3,4] was developed using target peptides identified from data dependent acquisition (DDA) analysis of recombinant murine histone H2BE- and a tryptic digest of enriched murine histones. A peptide unique to H2BE (EIQTSVR) and a peptide common to other H2B variants (EIQTAVR) were selected. Heavy isotope-labelled peptides were synthesized, quantified by amino acid analysis, aliquoted, and lyophilized by JPT Peptide Technologies, Berlin. To test and optimize the PRM method, 8ng of heavy isotope peptides were initially spiked into 250ng of tryptic E.coli digest and injected on column. After PRM method optimization, heavy labelled peptides were spiked into 2ug Histone tryptic so that 5ng was injected on column for each peptide.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Mass Spectrometry Data Acquisition

Samples were analyzed on a Q-Exactive HF mass spectrometer (Thermofisher Scientific San Jose, CA) coupled with an Ultimate 3000 nano UPLC system and an EasySpray source. Peptides were loaded onto an Acclaim PepMap 100 75um x 2cm trap column (Thermo) at 5uL/min and separated by reverse phase (RP)-HPLC on a nanocapillary column, 75 µm id × 50cm 2um PepMap RSLC C18 column (Thermo). Mobile phase A consisted of 0.1% formic acid and mobile phase B of 0.1% formic acid/acetonitrile. Peptides were eluted into the mass spectrometer at 300 nL/min with each RP-LC run comprising a 90-minute gradient from 3% B to 45% B.

The mass spectrometer was set to repetitively scan m/z from 300 to 1400 (R = 240,000) followed by data-dependent MS/MS scans on the twenty most abundant ions, minimum AGC 1e4, dynamic exclusion with a repeat count of 1, repeat duration of 30s, and resolution of 15000. The AGC target value was 3e6 and 1e5, for full and MSn scans, respectively. MSn injection time was 160 ms. Rejection of unassigned and 1+,6-8 charge states was set.

The parallel reaction monitoring (PRM) method[3,4] combined two scan events. For the full scan we used 150–2000 m/z mass selection, resolution 120,000 at m/z 200, automatic gain control (AGC) target value of 3e6, and maximum injection time of 200ms. The targeted MS/MS was run at an Orbitrap resolution of 30,000 at m/z 200, an AGC target value of 5e6, and maximum fill time of 200ms. and a target isolation window of 1.2 m/z. Fragmentation was performed with normalized collision energy (NCE) of 27 eV. Table 1 shows targeted inclusion list with scheduled retention times for heavy and light versions of EIQTAVR and EIQTSVR., along with unscheduled acquisition of iRT peptides for internal quality control.

Table 1. The PRM scheduled list for heavy and light peptides

Mass [m/z]	Formula [M]	Species [z]	Polarity	Start[min]	End[min]	N(CE)	MSX	ID	Comment

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

408.73233		2	Positive	25	35	27			EIQTAVR(light)
413.73647		2	Positive	25	35	27			EIQTAVR(heavy)
416.72979		2	Positive	25	35	27			EIQTSVR(light)
421.73393		2	Positive	25	35	27			EIQTSVR(heavy)

System Suitability and Quality Control

The suitability of Q Exactive HF instrument was monitored using QuiC software (Biognosys, Schlieren, Switzerland) and Skyline[5] for the analysis of the spiked-in iRT peptides in each sample. Meanwhile, as a measure for quality control, we injected standard *E. coli* protein digest prior to and after injecting sample set and collected the data in the Data Dependent Acquisition (DDA) mode. The collected data were analyzed in MaxQuant[6] and the output was subsequently visualized using the PTXQC[7] package to track the quality of the instrumentation.

MS data processing and analysis

The raw files for DDA analysis were processed with MaxQuant version 2.0.1.0 using its default workflow. The reference *E. coli* proteome from UniProt was concatenated with murine histones and common protein contaminants and used for the raw data search.

Skyline 21.2.0.568 with its default settings was used to process the PRM raw data. The heavy-isotope label for the C-terminal Arg was set as static modification. All peak integrations were reviewed manually and the sum of all transitions was calculated for light and heavy peptides. For each sample, light/heavy peak area ratio was calculated and multiplied by the known heavy peptide amount per injection to estimate the light peptide abundance value.

Primary neuronal culture

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Cortices were dissected from E16.5 C57BL/6J embryos and cultured in supplemented neurobasal medium (Neurobasal [Gibco 21103-049], B27 [Gibco 17504044], GlutaMAX [Gibco 35050- 061], Pen-Strep [Gibco 15140-122]) in TC-treated 12- or 6well plates coated with 0.05 mg/mL Poly-D-lysine (Sigma-Aldrich A-003-E). At 3-4 DIV, neurons were treated with 0.5 µM AraC. For all experiments using cultured cortical neurons, neurons were collected at 12 DIV.

CUT&Tag-sequencing

Library preparation & sequencing

Input samples were 400K primary cortical neurons per biological replicate. CUT&Tag was performed according to published protocols[8]. Prior to sequencing, library size distribution was confirmed by capillary electrophoresis using an Agilent 4200 TapeStation with high sensitivity D1000 reagents (5067-5585), and libraries were quantified by qPCR using a KAPA Library Quantification Kit (Roche 07960140001). Libraries were sequenced on an Illumina NextSeq550 instrument (42-bp read length, paired end).

Data processing and analysis

Reads were mapped to *Mus musculus* genome build mm10 with Bowtie 2 (v2.4.5)[9]. Six million reads from each biological replicate were subset and each condition was then merged across biological replicates (SAMtools[10] v1.15). Heatmaps were generated using deepTools[11] (v3.5.1). Metaplots were generated using ngs.plot[12] (v2.63) against the mouse genome. Peaks were called using MACS3[13] (v3.0.0b1) and annotated using Homer[14] (v4.10) . For downstream analysis, we used a Peak Score cutoff of 25 and removed peaks that were assigned to 'ChrUn' (unknown chromosome) by Homer. The R package GenomicDistributions[15] (v1.6.0) was used to analyze the genomic distribution of peaks. IGV tools[16] (2.12.3) was used to generate genome browser views.

To compare CUT&Tag signal to gene expression, normalized read counts from RNA-sequencing of WT primary neuronal cultures (see RNA-sequencing methods section) were used to generate gene lists by expression level.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Genes with base mean < 3 were defined as 'not expressed'. The remaining genes comprise the 'all expressed' categorization, and these genes were further divided into 3-quantiles (by base mean) to define 'low expression', 'mid expression' and 'high expression'.

Gene ontology

For gene ontology analysis, gene names were assigned to peak coordinates using Homer. Peaks that were annotated as 'intergenic', 'non-coding', and 'NA' by Homer were not included in GO analysis. PANTHER[17,18] (v18.0) was used to perform an overrepresentation test against the biological process complete ontology using default parameters. The *Mus musculus* genome was used as a background gene list. For conciseness and visualization, parent terms were excluded and only the most specific GO terms were plotted.

Dot Blot

0.2-µm PVDF paper was emersed in methanol and placed on TBS-soaked filter paper. Peptides were dotted onto the PVDF paper at serial dilutions and allowed to dry for 4 hours. The membrane was then wet in methanol again and equilibrated in TBS before blocking for 1 hour in 3% non-fat dry milk dissolved in TBS. The membrane was then incubated in antibody solution for 1 hour at room temperature and washed 3 times for five minutes in TBS-T (TBS with 0.1% Tween-20) before incubation with secondary antibody for 1 hour. Following secondary, the membrane was again washed 3 times for five minutes before imaging.

Protein extraction

Tissues were homogenized in 500uL RIPA buffer (10% sucrose, 1% SDS, 5mM HEPES pH 7.9, 10mM sodium butyrate in MilliQ water supplemented by protease inhibitor [Roche 04693124001], phosphatase inhibitor [Roche 04906837001], 1mM DTT, 1mM PMSF) using a pre-chilled dounce and pestle. Samples were then titrated 5x with a 26.3 gauge needle and centrifuged for 15 min at max speed at 4°C. Supernatant containing protein lysates were transferred to a new tube.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Chromatin Biochemistry Methods

DNA preparation for chromatin assembly

Mononucleosome-sized 601 DNA was prepared by PCR amplification of a DNA template containing one copy of the Widom 601 DNA sequence[19]. PCR products were then pooled and purified with a QIAquick PCR Purification Kit (Qiagen), using water to elute from the final columns. Eluents were pooled, frozen, and lyophilized before being resuspended in buffer TE (10 mM Tris-HCl pH 7.6, 0.1 mM EDTA), quantified by NanoDrop OneC, and adjusted to a final concentration of approximately 1-1.5 g/L.

DNA templates used for chromatin fibers were prepared essentially as described before[20]. *E. coli* DH5α cells were transformed with a pWM530 vector bearing 12 repeats of the Widom 601 nucleosome positioning sequence with a 30 bp linker and used to inoculate 6 L of luria broth under ampicillin selection and grown at 37 °C for 18-24 hrs. Cultures were harvested and DNA purified using a Plasmid Giga Kit (Qiagen). Purified DNA was resuspended in buffer TE and adjusted to a concentration of ~1 g/L. Plasmid was then digested with EcoRI and EcoRV (NEB) overnight to generate linear chromatin templates (i.e., dslDNA or 601^{18}-L. Complete digestion was determined by agarose gel electrophoresis in 1X TAE buffer. Next, digestion reactions were fractionated by PEG 6000-induced precipitation as described previously to separate the ~3.2 kb desired chromatin substrates from the 200-300 bp plasmid backbone fragments[21]. Precipitated DNA was next resuspended in buffer TE, purified by phenol/chloroform extraction, and precipitated with absolute ethanol as described[22]. DNA was finally resuspended in buffer TE, quantified by NanoDrop OneC, and adjusted to a concentration of 1-1.5 g/L. Buffer DNA was likewise prepared by digestion of a similar construct containing 8 repeats of the 155 bp mouse mammary tumor virus (MMTV) weak nucleosome positioning sequence with EcoRV, PEG precipitation, and phenol/chloroform extraction.

Histone expression and purification

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Recombinant human histones were expressed and purified as previously described[19]. Briefly, H2A, H2B, H2BE, H3.1, H4, and histone mutants H2B V39I and H2BE I39V were expressed in *E. coli* BL21(DE3) or C41(DE3)) cells at 37 °C until reaching an OD_{600} of 0.6-0.8, before induction with 500 µM IPTG for 4 hours at 37 °C. Bacteria were then harvested by centrifugation at 6000 x g for 25 minutes.

Cell pellets were resuspended in lysis buffer (200 mM NaCl, 20 mM Tris-HCl pH 7.6, 1 mM EDTA, 1 mM β-mercaptoethanol) and then lysed via sonication. Lysate was cleared by centrifugation at 30,000 x g for 20 minutes and the remaining insoluble pellet was resolubilized in extraction buffer (6 M guanidine HCl, 1 mM DTT, 1x PBS, pH 7) and nutated at 4 °C overnight. The extraction was cleared by centrifugation at 30,000 x g for 40 minutes and filtered using a 0.45-micron syringe filter (Fisher Scientific). Filtered extraction was diluted 1:1 with HPLC buffer A (0.1% trifluoroacetic acid (TFA) in water) before purification with reverse-phase HPLC on a 30-70% buffer B (0.1% TFA in 90% acetonitrile and 10% water). Absorbance at 280 nm was used to observe desired peaks and fractions were collected using an automated fraction collector. Purity of fractions was determined using a quadrupole LC-MS. Fractions deemed pure were combined, lyophilized and stored at -80 °C.

Histone octamer assembly

To assemble histone octamers, monomeric core histones were resuspended in unfolding buffer (20 mM Tris-HCl pH 7.6, 6 M guanidine hydrochloride, 1 mM DTT) and quantified by measuring A280 using a NanoDrop. Core histones were next combined at the following stoichiometries to generate a slight excess of H2A/H2B dimer to assist with subsequent purification: 1.05:1.05:1:1 H2A:H2B:H3:H4. Total protein concentration of the mixture was adjusted to 1 g/L, and samples were dialyzed against octamer refolding buffer (10 mM Tris-HCl pH 7.6, 2 M NaCl, 1 mM EDTA, 1 mM DTT) at 4 °C using 3.5K MWCO Slide-A-Lyzer dialysis cassettes. A total of three rounds of dialysis against refolding buffer were performed, with the first exchange going overnight and the subsequent two lasting at least 6 hrs.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Following the final dialysis, samples were harvested from dialysis cassettes, cleared by centrifugation, and concentrated using 30K MWCO Amicon centrifugal filter units. Finally, samples were centrifuged at 17,0000 xg for a minimum of 10 minutes at 4 °C before being injected onto an AKTA 25L FPLC instrument and resolved over a SuperDex 200 10/300 Increase column, using octamer refolding buffer for the liquid phase. Fractions were analyzed by SDS-PAGE on 12 % acrylamide gels, and those containing octamers were pooled and concentrated using Amicon 30K MWCO centrifugal filter unites. Octamers were quantified by A280, adjusted to 50 % (v/v) glycerol, and stored at -20 °C for future use.

Chromatin reconstitution

Nucleosome core particles (NCPs), linear nucleosome arrays, and circular nucleosome arrays were all prepared by the same method of salt gradient dialysis with all steps at 4 °C. Substrate DNA and recombinant histone octamers were combined in approximately equimolar quantities (with respect to the expected nucleosome load of the DNA, i.e., 1 for mononucleosomal DNA, or 12 for chromatin arrays) with final buffer conditions identical to octamer refolding buffer – 10 mM Tris-HCl pH 7.6, 2 M NaCl, 0.1 mM EDTA, 1 mM DTT. For nucleosome array assemblies MMTV DNA was also added, but at a lower stoichiometry to act as a buffer that facilitates array assembly without competing for octamer occupancy (0.2:1 MMTV:NCP). Optimal assembly stoichiometries were determined empirically for all templates, with Octamer:DNA stoichiometries being approximately 1.2:1 for mononucleosomes and 1.6:1 for 601 nucleosome arrays.

Assembly reactions were combined and mixed by gentle pipetting, centrifuged for 5 min at 17,000 xg at 4 °C, and then added to Slide-A-Lyzer Mini dialysis buttons pre-moistened in array initial buffer (10 mM Tris-HCl pH7.6, 1.4 M NaCl, 0.1 mM EDTA, 1 mM DTT) and dialyzed against 200 mL of the same buffer for 1 hr. Next, a peristaltic pump was used to transfer 350 mL of dilution buffer (10 mM Tris-HCl pH 7.6, 10 mM NaCl, 0.1 mM EDTA, 1 mM DTT) to the samples in initial buffer with a flow rate of 0.5-1.0 mL/min. After all dilution buffer was transferred, arrays were left to dialyze for at least 1 hr or up to overnight. Samples were next moved to a fresh

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

350 mL of dilution buffer and dialyzed for 6 hrs. Lastly, samples were moved 300 mL of fresh dilution buffer and dialyzed for 1-2 hrs before harvesting.

Mononucleosome samples were simply pipetted out of dialysis buttons, centrifuged for 5 min at 17,000 xg, and supernatant (in case any precipitation was present) was quantified, subject to EMSA to validate assembly efficiency, and stored at 4 °C for up to 2 months. Chromatin array samples were similarly harvested to yield a mixture of assembled chromatin fibers, MMTV mononucleosomes, and free MMTV DNA. To separate the chromatin fibers, an equal volume of precipitation buffer (10 mM Tris-HCl pH 7.6, 10 mM NaCl, 10 mM $MgCl_2$) was added to samples followed by a 20-minute incubation on ice. Next, samples were centrifuged for 10 minutes at 17,000 xg. Supernatant was gently pipetted off samples to avoid disturbing the barely visible pellets, to which a desired volume of array dilution buffer was next added. Pellets were left on ice for 10 minutes undisturbed to allow for the gradual resuspension of chromatin pellets before being quantified. Finally, EMSA was used to validate assemblies and harvests. Chromatin arrays were stored at 4 °C for up to one week before use.

Electrophoretic mobility shift assays (EMSAs)

Mononucleosome EMSAs were performed using polyacrylamide gel electrophoresis in 5 % acrylamide, 0.5 X TBE gels. A solution of 1 M sucrose was used as a loading buffer (1:3 dilution) for nucleosome samples. Gels were run for 30-40 minutes at a constant voltage of 130 V, stained with SYBR Gold dye diluted 1:10,000 in 0.5 X TBE buffer for 5-10 minutes, and DNA migration was visualized on an Amersham AI600 imager (GE/Cytiva) using the UV 312 nm channel. If needed, protein migration was visualized after DNA imaging by staining gels with Imperial Protein Stain and visualized on the AI600 instrument in the colorimetric channel.

Chromatin array EMSAs were performed similarly, except with a different gel formulation. Agarose-polyacrylamide gel electrophoresis (APAGE) gels were cast using the Mini-PROTEAN Tetra Handcast system (Bio-Rad). Before combining gel-casting reagents, the gel-facing sides of the 1.5 mm spacer plates, short plates, and 10-well gel combs were lubricated with a thin layer of 50 % (v/v) glycerol. Next, ultrapure water, 50X TAE

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(sufficient for a final concentration of 0.5 X), and powdered agarose (sufficient for a final concentration of 1 % w/v) were heated until dissolved. This solution was rapidly combined with 40 % acrylamide (37.5:1 mono:bis) solution (Bio-Rad) (sufficient for a final concentration of 2 %), APS (sufficient for 0.125 %), and TEMED (sufficient for 0.04 %), and added into the assembled gel casting cassettes. Gels were allowed to cool and polymerize for at least 1 hr. Once cool, gels were pre-run at 4 °C in 0.5 X TAE buffer for 3 hrs at 100 V. Next, chromatin samples could be loaded onto gels using sucrose loading solution and the gels run for approximately 1 hr at 120 V at room temperature before visualization as mononucleosome EMSAs.

Atomic force microscopy

DNA-protein complexes were imaged using an Asylum MFP 3D Bio AFM (Oxford Instruments, Goleta CA) with an Olympus AC240TS probe in tapping mode at room temperature. The samples were prepared a suitable concentration (0.5-1.0 ng/µL), then 40 uL of prepared samples were slowly deposited to a freshly cleaved AP-mica for 5 minutes and rinsed with 1 mL ultrapure deionized water twice before being gently dried with UHP argon gas.

AFM images were collected at a speed of 0.5-1 Hz at 512 × 512-pixel resolution, with an image size of 2 µm. For analysis, raw images were exported into 8-bit grayscale Tiff images using the Asylum Research's Igor Pro software and imported into FIJI/ImageJ (NIH) for detection of single particles and quantification of volume, surface area, and volume/surface area ratio using as has been done previously for studies of chromatin compaction via AFM[23]. In order to assess single chromatin particles, rather than potential clusters of multiple fibers or residual MMTV mononucleosomes, only particles with volumes measuring between 2,000 and 15,000 nm^3 were included in analyses.

Magnesium-dependent self-association assay

Chromatin compaction was tested by magnesium-drive self-association as described previously[24]. Briefly, a magnesium solution (100 mM $MgCl_2$, 10 mM Tris-HCl pH 7.6, 10 mM NaCl) was titrated into the sample solution

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to raise the magnesium concentration in 0.5 mM increments. After each addition, samples were allowed to sit on ice for 10 min and then centrifuged for an additional 10 min at 17,000 xg at 4 °C. The concentration of soluble DNA from nucleosome arrays was measured by NanoDrop.

Nucleosome differential scanning fluorimetry (DSF) assays

Nucleosome stability was measured using a Protein Thermal Shift kit (Applied Biosystems) with the following modifications. The assay was conducted at 10 µL volumes with 10X SYPRO Orange dye (Invitrogen) and nucleosome dilution buffer (10 mM Tris-HCl pH 7.6, 100 mM NaCl, 0.1 mM EDTA, and 1 mM DTT) in 384-well plates. Final nucleosome concentrations used were approximately 60 ng/µL of DNA. Fluorescence melt curve data was acquired using a QuantStudio 5 Real-Time PCR System (Applied Biosystems) with the following method: initial ramp rate of 1.6 °C/s to 25 °C with a 5-minute hold time at 25 °C, followed by a second ramp rate of 0.05 °C/s to 99.9 °C with a 2-minute hold time at 99.9 °C. Melting temperatures were calculated using the Protein Thermal Shift software (Applied Biosystems).

ATAC-sequencing

Library preparation & sequencing

Input samples were 400K primary cortical neurons per biological replicate. Neurons were collected by scraping in cold lysis buffer (10 mM Tris-Cl (pH 7.5), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% (vol/vol) NP-40, 0.1% (vol/vol) Tween-20 and 0.01% (vol/vol) digitonin) and washed in wash buffer (10 mM Tris-Cl (pH 7.5), 10 mM NaCl, 3 mM $MgCl_2$ and 0.1% (vol/vol) Tween-20). Transposition was performed with Tagment DNA TDE1 (Illumina, 15027865). Transposition reactions were cleaned with AMPure XP beads (Beckman, A63880), and libraries were generated by PCR with NEBNext High-Fidelity 2× PCR Master Mix (NEB, M0541). Prior to sequencing, library size distribution was confirmed by capillary electrophoresis using an Agilent 4200 TapeStation with high sensitivity D1000 reagents (5067-5585), and libraries were quantified by qPCR using a KAPA Library Quantification Kit

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Roche 07960140001). Libraries were sequenced on an Illumina NextSeq550 instrument (42-bp read length, paired end).

Data processing and analysis

Reads were mapped to *Mus musculus* genome build mm10 with Bowtie 2 (v2.4.5)[9]. Fifty million reads were subsetted from each biological replicate and each condition was then merged across biological replicates (SAMtools[10] v1.15). Heatmaps were generated using deepTools[11] (v3.5.1). Metaplots were generated using ngs.plot[12] (v2.63) against the mouse genome. Peaks were called using MACS3[13] (v3.0.0b1) and annotated using Homer[14] (v4.10) . For downstream analysis, we used a Peak Score cutoff of 25 and removed peaks that were assigned to 'ChrUn' (unknown chromosome) by Homer. Differential ATAC peaks were called using DiffBind[25] (v3.4.11). The following genomic regions were used as input data for DiffBind analysis: WT ATAC peak coordinates from MACS3 (Fig 2d-g), H2BE binding site coordinates from MACS3 peak calling of CUT&Tag data (Fig 2h-k), TSS coordinates +/- 500bp from mm10, enhancer regions from brain tissue[26]. IGV tools[16] (2.12.3) was used to generate genome browser views.

To compare ATAC signal within different genomic regions, we subset H2BE binding sites based on their Homer annotations into promoter ('promoter-TSS' in Homer), gene body ('exon' and 'intron' in Homer), and intergenic ('intergenic' in Homer).

Viral infection

Viral constructs

The GFP control plasmid was obtained from Addgene, pLenti-CMV-MCS-GFP-SV-puro (Addgene plasmid 73582).

The H2BE and H2B constructs were generated as previously described[1] and were provided by the Dulac lab. The H2B and H2BE coding sequences were then moved to the pLenti backbone through Gibson assembly. Primers were designed through NEB Builder to separately amplify H2B or H2BE and the pLenti backbone. Each primer

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

contained a non-complementary region on its 5'-end that corresponded to approximately 10 bases on the other template. PCR amplification was performed with Q5 High-Fidelity DNA Polymerase (NEB M0491S). Following DPN1 digestion, fragments were ligated together using NEBuilder HiFi DNA Assembly Master Mix (NEB E2621S) and transformed into NEB 5-alpha Competent *E. coli* cells (E2621S). Plasmid sequence was verified through Plasmidsaurus long read sequencing.

The H2BE-I39V construct was generated by site-directed mutagenesis of the H2BE backbone using Pfu Turbo HotStart DNA polymerase (Agilent, 600322-51), and primers were created using the DNA-based primer design feature of the online PrimerX tool. Constructs were verified by Sanger sequencing.

Lentiviral production

HEK293T cells were cultured in high-glucose DMEM growth medium (Corning 10-013-CV), 10% FBS (Sigma-Aldrich F2442-500ML), and 1% Pen-Strep (Gibco 15140-122). Calcium phosphate transfection was performed with Pax2 and VSVG packaging plasmids. Viral media was removed 2 h after transfection and collected at 48 and 72 h later. Viral media was passed through a 0.45-µm filter and precipitated for 48 hours with PEG-it solution (40% PEG-8000 [Sigma-Aldrich P2139-1KG], 1.2 M NaCl [Fisher Chemical S271-1]). Viral particles were pelleted and resuspended in 200µL PBS.

Neuronal infection

At 10 DIV, neurons were transduced overnight with lentivirus containing the constructs described above. Virus was removed the following day, and neurons were cultured for one additional day.

RNA-sequencing

Library preparation & sequencing

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Input samples were 400 primary cortical neurons from 4 WT biological replicates and 4 KO biological replicates. RNA was isolated using Zymo Quick-RNA Miniprep Plus Kit (R1057). Prior to library preparation, RNA integrity was confirmed using an Agilent 4200 TapeStation with high sensitivity RNA reagents (5067-5579). Sequencing libraries were prepared using the TruSeq Stranded mRNA kit (Illumina 20020595). Prior to sequencing, library size distribution was confirmed by capillary electrophoresis using an Agilent 4200 TapeStation with high sensitivity D1000 reagents (5067-5585), and libraries were quantified by qPCR using a KAPA Library Quantification Kit (Roche 07960140001). Libraries were sequenced on an Illumina NextSeq1000 instrument (66-bp read length, paired end).

Data processing and analysis

Reads were mapped to *Mus musculus* genome build mm10 with Star (v2.7.9a). The R packages DESeq2[27] (v1.38.3) and limma (v3.54.2) via edgeR[28] (v3.40.2) were used to perform differential gene expression analysis. We defined genes as differentially expressed where FDR < 0.05 and absolute fold change 1.25. Volcano plots were generated using VolcaNoseR[29]. IGV tools[16] (2.12.3) was used to generate genome browser views.

Gene ontology

PANTHER[17,18] (v18.0) was used to perform an overrepresentation test against the biological process complete ontology using default parameters. SynGO[30] was used for synaptic gene ontologies and overrepresentation tests of differentially expressed genes. All expressed genes (defined as base mean 3) was used as a background gene list. For conciseness and visualization, parent terms were excluded and only the most specific GO terms were plotted.

Single-nucleus Drop-sequencing (sNucDrop-seq)

Nuclei isolation

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Snap-frozen brain tissues were homogenized in 1mL Buffer A (0.25M sucrose, 50mM Tris-HCl pH7.4, 25mM KCl, 5mM $MgCl_2$ supplemented by EDTA-free protease inhibitor [Roche 4693159001]) using a pre-chilled dounce and pestle. Homogenate was then transferred to a pre-chilled 15mL conical tube and mixed with 6mL Buffer B (2.3M sucrose, 50mM Tris-HCl pH7.4, 25mM KCl, 5mM $MgCl_2$). An additional 2mL Buffer A was used to rinse leftover homogenate from the dounce and combined with the sample. The homogenate was gently transferred to a pre-chilled 15mL ultracentrifuge tube containing 2mL Buffer C (1.8M sucrose, 50mM Tris-HCl pH7.4, 25mM KCl, 5mM $MgCl_2$). Nuclei were pelleted at 100,000 x g for 1.5hr at 4C using a SWI41 rotor. The supernatant was discarded and 1.5mL Buffer D (0.01% BSA in 1X PBS with 0.5U/uL RNase inhibitor [Lucigen 30281-2]) was gently added to the nuclei pellet and incubated on ice 20min. The nuclei pellet were resuspended and the suspension was transferred to a 1.5mL lo-bind tube.

Library preparation and sequencing

The single-nucleus suspensions were individually diluted to a concentration of 100 nuclei/mL in DPBS containing 0.01% BSA. Approximately 1.5 mL of this single-nucleus suspension was loaded for each sNucDrop-seq run. The single-nucleus suspension was then co-encapsulated with barcoded beads (ChemGenes) using an Aquapel-coated PDMS microfluidic device (mFluidix) connected to syringe pumps (KD Scientific) via polyethylene tubing with an inner diameter of 0.38mm (Scientific Commodities) (Macosko et al., 2015). Barcoded beads were resuspended in lysis buffer (200 mM Tris-HCl pH8.0, 20 mM EDTA, 6% Ficoll PM-400 (GE Healthcare/Fisher Scientific), 0.2% Sarkosyl (Sigma-Aldrich), and 50 mM DTT (Fermentas; freshly made on the day of run) at a concentration of 120 beads/mL. The flow rates for nuclei and beads were set to 4,000 mL/hr, while QX200 droplet generation oil (Bio-rad) was run at 15,000 mL/hr. A typical run lasts 20 min. Droplet breakage with Perfluoro-1-octanol (Sigma-Aldrich), reverse transcription and exonuclease I treatment were performed, as previously described[31], with minor modifications. For up to 120,000 beads, 200 µL of reverse transcription (RT) mix (1x Maxima RT buffer (ThermoFisher), 4% Ficoll PM-400, 1 mM dNTPs (Clontech), 1 U/mL RNase inhibitor, 2.5 mM Template Switch Oligo (TSO: AAGCAGTGGTATCAACGCAGAGTGAATrGrGrG), and 10 U/ mL Maxima H

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Minus Reverse Transcriptase (ThermoFisher)) were added. The RT reaction was incubated at room temperature for 30min, followed by incubation at 42C for 120 min. To determine an optimal number of PCR cycles for amplification of cDNA, an aliquot of 6,000 beads was amplified by PCR in a volume of 50 µL (25 µL of 2x KAPA HiFi hotstart readymix (KAPA biosystems), 0.4 µL of 100 mM TSO-PCR primer (AAGCAGTGGTATCAACGCAGAGT, 24.6 µL of nuclease-free water) with the following thermal cycling parameter (95C for 3 min; 4 cycles of 98C for 20 sec, 65C for 45 sec, 72C for 3 min; 9 cycles of 98C for 20 sec, 67C for 45 sec, 72C for 3 min; 72C for 5 min, hold at 4C). After two rounds of purification with 0.6x SPRISelect beads (Beckman Coulter), amplified cDNA was eluted with 10 µL of water. 10% of amplified cDNA was used to perform real-time PCR analysis (1 µL of purified cDNA, 0.2 µL of 25 mM TSO-PCR primer, 5 µL of 2x KAPA FAST qPCR readymix, and 3.8 µL of water) to determine the additional number of PCR cycles needed for optimal cDNA amplification (Applied Biosystems QuantStudio 7 Flex). We then prepared PCR reactions per total number of barcoded beads collected for each sNucDrop-seq run, using 6,000 beads per 50- µL PCR reaction, and ran the aforementioned program to amplify the cDNA for 4 + 10 to 12 cycles. We then tagmented cDNA using the Nextera XT DNA sample preparation kit (Illumina, FC-131-1096), starting with 550 pg of cDNA pooled in equal amounts, from all PCR reactions for a given run. Following cDNA tagmentation, we further amplified the tagmented cDNA libraries with 12 enrichment PCR cycles using the Illumina Nextera XT i7 primers along with the P5-TSO hybrid primer (AATGATACGGCGACCACCGAGATCTACACGCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGT*A*C)[32]. After quality control analysis by Qubit 3.0 (Invitrogen) and a Bioanalyzer (Agilent), libraries were sequenced on an Illumina NextSeq 500 instrument using the 75-cycle High Output v2 Kit (Illumina). We loaded the library at 2.0 pM and provided Custom Read1 Primer (GCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC) at 0.3 mM in position 7 of the reagent cartridge. The sequencing configuration was 20 bp (Read1), 8 bp (Index1), and 60 bp (Read2).

Preprocessing of sNucDrop-seq data

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Paired-end sequencing reads of sNucDrop-seq were processed using publicly available the Drop-seq Tools v1.12 software[32] with some modifications. Briefly, each mRNA read (read2) was tagged with the cell barcode (bases 1 to 12 of read 1) and unique molecular identifier (UMI, bases 13 to 20 of read 1), trimmed of sequencing adaptors and poly-A sequences, and aligned using STAR v2.5.2a to the mouse reference genome assembly (mm10, Gencode release vM13). Because a substantial proportion (~50%) of reads derived from nuclear transcriptomes of mouse cortices were mapped to the intronic regions, the intronic reads were retained for downstream analysis. A custom Perl script was used to retrieve both exonic and intronic reads mapped to predicted strands of annotated genes[31]. Uniquely mapped reads were grouped by cell barcodes. Cell barcodes were corrected for possible bead synthesis errors, using the DetectBeadSynthesisErrors program from the Drop-seq Tools v1.12 software. To generate digital expression matrix, a list of UMIs in each gene (as rows), within each cell (as columns), was assembled, and UMIs within ED = 1 were merged together. The total number of unique UMI sequences was counted, and this number was reported as the number of transcripts of that gene for a given nucleus.

Nuclei clustering and marker gene identification

Raw digital expression matrices were combined and loaded into the R package Seurat. For normalization, UMI counts for all nuclei were scaled by library size (total UMI counts), multiplied by 10,000 and transformed to log scale. Nuclei with a relatively high percentage of UMIs mapped to mitochondrial genes (0.2) were discarded. Moreover, nuclei with <=500 UMI or =5000 UMI, discarded, as were nuclei with <=250 genes. Only genes found to be expressing in 10 cells were retained. The Seurat object was then normalized and transformed using the Seurat functions NormalizeData and SCTransform for each genotype. The Seurat functions SelectIntegrationFeatures (nfeatures = 3000), PrepSCTIntegration, FindIntegrationAnchors and IntegrateData were used to integrate the datasets based on the top 3000 more variable features. Prior to clustering, we performed principal component analysis using the RunPCA function and selected 30 principal components for UMAP non-linear dimensional reduction. Based on UMAP, twenty clusters were identified using the Seurat

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

functions Find Neighbors (dim = 30) and FindClusters (resolution = 0.5). To identify marker genes for each cluster, differential expression analysis was performed using the Seurat function FindAllMarkers. Differentially expressed genes that were expressed at least in 25% cells within the cluster and with a fold change more than 0.25 (log scale) were considered to be marker genes. Marker gene analysis led to the identification of 15 cortical neuron clusters (10 excitatory, 5 inhibitory), 1 subcortical neuron cluster, and 4 non-neuronal clusters.

Differential gene expression analysis

Differential gene expression analysis between WT an KO groups was performed using the Seurat function FindMarkers (min.pct = .00001, logfc.threshold = 0) with a Wilcoxon Rank Sum test. Genes with an adjusted p-value <.05 were considered differentially expressed between WT and KO.

Gene ontology

PANTHER[17,18] (v18.0) was used to perform an overrepresentation test against the biological process complete ontology using default parameters. All genes detected by FindMarkers was used as a background gene list. For conciseness and visualization, parent terms were excluded and only the most specific GO terms were plotted.

Electrophysiology

Mice were deeply anesthetized and trans-cardially perfused with ice-cold aCSF containing (in mM): 124 NaCl, 2.5 KCl, 1.2 HaH2PO4, 24 NaHCO3, 5 HEPES, 13 Glucose, 1.3 MgSO4, 2.5 CaCl2. After perfusion, the brain was quickly removed, submerged and coronally sectioned on a vibratome (VT1200s, Leica) at 400 µm thickness in ice-cold aCSF. Slices were transferred to NMDG based recovery solution at 32ºC of the following composition (in mM): 92 NMDG, 2.5 KCl, 1.2 NaH2PO4, 30 NaHCO3, 20 HEPES, 25 Glucose, 5 Sodium ascorbate, 2 Thiourea, 3 Sodium pyruvate, 10 MgSO4, 0.5 CaCl2. After 12-15 minutes recovery, slices were transferred to room temperature aCSF chamber (20-22ºC) and left for at least 1 hour before recording. Following recovery, slices

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were placed in a recording chamber, fully submerged at a flow rate of 1.4~1.6 mL/min and maintained at 29-30°C in oxygenated (95% O2, 5% CO2) aCSF, with 100µM Picrotoxin included.

In extracellular recordings, recording pipettes were fabricated by pulling borosilicate glass (World Precision Instruments, TW150-3). These pipettes exhibited a tip resistance ranging from 4 to 5 MΩ when filled with an aCSF. Stimulus pipettes were made by pulling theta glasses (TG150-4, Warner Instrument) and adjusting the tip size to 30~50 um. For measurements of evoked post-synaptic response in Schaffer-collateral pathway, Hippocampal layers were identified using IR-DIC optics (BX51, Olympus), with visual guidance. The stimulus electrode was positioned on the surface of the striatum radiatum (s.r.) from the CA3 direction. A recording electrode was placed on the postsynaptic site in CA1, opposite the stimulus electrode.

Brief electrical pulses (0.2ms) were delivered on each sweep to measure post-synaptic responses every 10 seconds. Once responses stabilized, the intensity of the stimulus was sequentially adjusted for input-output (I-O) measurement. Voltage (V) ranges were from 3V to 30V in increments of 3V. We identified the field excitatory post-synaptic potential (fEPSP) based on the temporal separation of the fiber volley, and then averaged measured the fEPSP slope of the 10-90% range from baseline to peak. Ten responses at the same input were averaged.

Recordings were performed using a MultiClamp 700B (Molecular Devices) and Igor7 (WaveMetrics; recording artist addon, developed by Richard C Gerkin, github : https://github.com/rgerkin/recording-artist), filtered at 2.8 kHz and digitized at 10 kHz. Axon terminals were stimulated with brief (0.2 ms) pulses using isoflex isolator (Voltage control). Data were analyzed using Igor7.

Behavioral assays

Behavioral cohorts

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Male and female WT and KO mice were tested in the behavioral tests described below. Mice were 3-4 months old at the onset of behavioral testing. One cohort of mice was used for olfaction, social choice, and fear conditioning testing, in that order. A separate cohort of mice was used for NOR and T-maze testing. During testing, the experimenter was blinded to genotype of the mice.

Olfactory habituation/dishabituation test

Mice were tested for olfactory habituation & dishabituation using a published protocol[33]. In brief, mice received sequential presentations of cotton swabs scented with different odors in the following sequence: water, almond, vanilla, same sex conspecific, opposite sex conspecific. Each odor was presented in three consecutive trials for 2 min, with an intertrial interval of 1 min. Time spent interacting with each scented swab was manually analyzed by three scorers who were blinded to sex and genotype.

NOR/open field

Novel object recognition and open field were performed as described in Korb 2015[34] and Korb 2017[35]. In brief, mice were placed in an empty arena for 10 minutes for a habituation period that also served as an open field assessment. One day later mice were habituated for an additional 2 min and briefly removed from the arena while two identical objects (either a faucet, a plastic pyramid, a small fish figurine, or stacked Legos) were placed in the box and mice were given 10 min to explore. On the following day, mice were returned to the box with one object they had previously seen and one new object in place of the original object and were allowed to explore for 10 min. All sessions were recorded using EthoVision software. Time spent interacting with each object was manually analyzed. Discrimination index was calculated as (time with novel object − time with familiar object)/(time with novel object + time with familiar object).

3-chamber social choice assay

The social choice test was carried out in a three-chambered apparatus, consisting of a center chamber and two outer chambers. Before the start of the test and in a counter-balanced manner, one end chamber was

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

designated the social chamber, into which a stimulus mouse would be introduced, and the other end chamber was designed the nonsocial chamber. Two identical, clear Plexiglas cylinders with multiple holes to allow for air exchange were placed in each end chamber. In the habituation phase of the test, the experimental mouse freely explores the three chambers with empty cue cylinders in place for 10 min. Immediately following habituation, an age- and sex-matched stimulus mouse was placed in the cylinder in the social chamber while a rock was simultaneously placed into the other cylinder in the nonsocial chamber. The experimental mouse was tracked during the 10 min habituation and 10 min social choice phases. All testing was recorded, and videos were analyzed using ANY-maze software.

T-maze

For T-maze testing mice were placed at the end of the center arm of a T shaped raised arena enclosed with plexiglass. Mice were allowed to freely explore for 7 minutes. Entries into each arm were defined as the full body of the mouse entering (not necessarily including the tail). A success was defined as a consecutive entry into each of the 3 arms without returning to the arm that the mouse had been in immediately prior. Successful triads over the total number of possible triads based on the total entries were calculated as correct triad/(total entries - 2).

Contextual and cued fear conditioning

Mice were handled for 2 minutes each on 3 consecutive days immediately prior to the onset of testing. On training day, mice were placed in individual chambers for 2 min followed by a loud tone lasting 30 s that co-terminated with a 2-s, 1.25-mA foot shock. One minute later mice received another tone-shock pairing and were then left undisturbed for an additional 1 min in the chamber before being returned to their home cage. Freezing behavior, defined as no movement except for respiration, was determined before and after the tone-shock pairings and scored by MedAssociates VideoFreeze software. To test for context-dependent learning, we placed mice back into the same testing boxes 24 hr later for a total of 5 min without any tone or shock, and again

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

179 measured the total time spent freezing. Following an additional 24 hr, we tested for cue-dependent fear
180 memory by placing the mice into a novel chamber consisting of altered flooring, wall-panel inserts, and vanilla
181 scent. After 2 min in the chamber, the cue tone was played for a total of 3 min, and the total time spent freezing
182 during the presentation of this cue tone was recorded. Long-term contextual and cued fear memory were again
183 tested with the same protocol at 14 d (contextual) or 15d (cued) post-training.

184

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

1. Santoro, S. W. & Dulac, C. The activity-dependent histone variant H2BE modulates the life span of olfactory neurons. *eLife* **1**, e00070 (2012).

2. Zougman, A., Selby, P. J. & Banks, R. E. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **14**, 1006–1000 (2014).

3. Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S. & Coon, J. J. Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. *Molecular & Cellular Proteomics* **11**, 1475–1488 (2012).

4. Gallien, S. *et al.* Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol Cell Proteomics* **11**, 1709–1723 (2012).

5. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968 (2010).

6. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301–2319 (2016).

7. Bielow, C., Mastrobuoni, G. & Kempa, S. Proteomics Quality Control: Quality Control Software for MaxQuant Results. *J Proteome Res* **15**, 777–787 (2016).

8. S Kaya-Okur, H. *Bench top CUT&Tag v3*. https://www.protocols.io/view/bench-top-cut-amp-tag-bcuhiwt6 (2020) doi:10.17504/protocols.io.bcuhiwt6.

9. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).

10. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).

11. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160–W165 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

13. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

14. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576–589 (2010).

15. Kupkova, K. *et al.* GenomicDistributions: fast analysis of genomic intervals with Bioconductor. *BMC Genomics* **23**, 299 (2022).

16. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24–26 (2011).

17. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat Genet* **25**, 25–29 (2000).

18. Gene Ontology Consortium *et al.* The Gene Ontology knowledgebase in 2023. *Genetics* **224**, iyad031 (2023).

19. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *Journal of Molecular Biology* **276**, 19–42 (1998).

20. Osunsade, A. *et al.* A Robust Method for the Purification and Characterization of Recombinant Human Histone H1 Variants. *Biochemistry* **58**, 171–176 (2019).

21. Lis, J. T. Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods Enzymol* **65**, 347–353 (1980).

22. Dyer, P. N. *et al.* Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA. in *Methods in Enzymology* vol. 375 23–44 (Elsevier, 2003).

23. Krzemien, K. M., Beckers, M., Quack, S. & Michaelis, J. Atomic force microscopy of chromatin arrays reveal non-monotonic salt dependence of array compaction in solution. *PLoS ONE* **12**, e0173459 (2017).

24. Schwarz, P. M., Felthauser, A., Fletcher, T. M. & Hansen, J. C. Reversible Oligonucleosome Self-Association: Dependence on Divalent Cations and Core Histone Tail Domains. *Biochemistry* **35**, 4009–4015 (1996).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

25. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).

26. Wang, Y. *et al.* SEdb 2.0: a comprehensive super-enhancer database of human and mouse. *Nucleic Acids Res* **51**, D280–D290 (2023).

27. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

28. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

29. Goedhart, J. & Luijsterburg, M. S. VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci Rep* **10**, 20560 (2020).

30. Koopmans, F. *et al.* SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* **103**, 217-234.e4 (2019).

31. Hu, P. *et al.* Dissecting Cell-Type Composition and Activity-Dependent Transcriptional State in Mammalian Brains by Massively Parallel Single-Nucleus RNA-Seq. *Molecular Cell* **68**, 1006-1015.e7 (2017).

32. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

33. Yang, M. & Crawley, J. N. Simple Behavioral Assessment of Mouse Olfaction. *CP Neuroscience* **48**, (2009).

34. Korb, E., Herre, M., Zucker-Scharff, I., Darnell, R. B. & Allis, C. D. BET protein Brd4 activates transcription in neurons and BET inhibitor Jq1 blocks memory in mice. *Nat Neurosci* **18**, 1464–1473 (2015).

35. Korb, E. *et al.* Excess Translation of Epigenetic Regulators Contributes to Fragile X Syndrome and Is Alleviated by Brd4 Inhibition. *Cell* **170**, 1209-1223.e20 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.29.575103; this version posted January 30, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

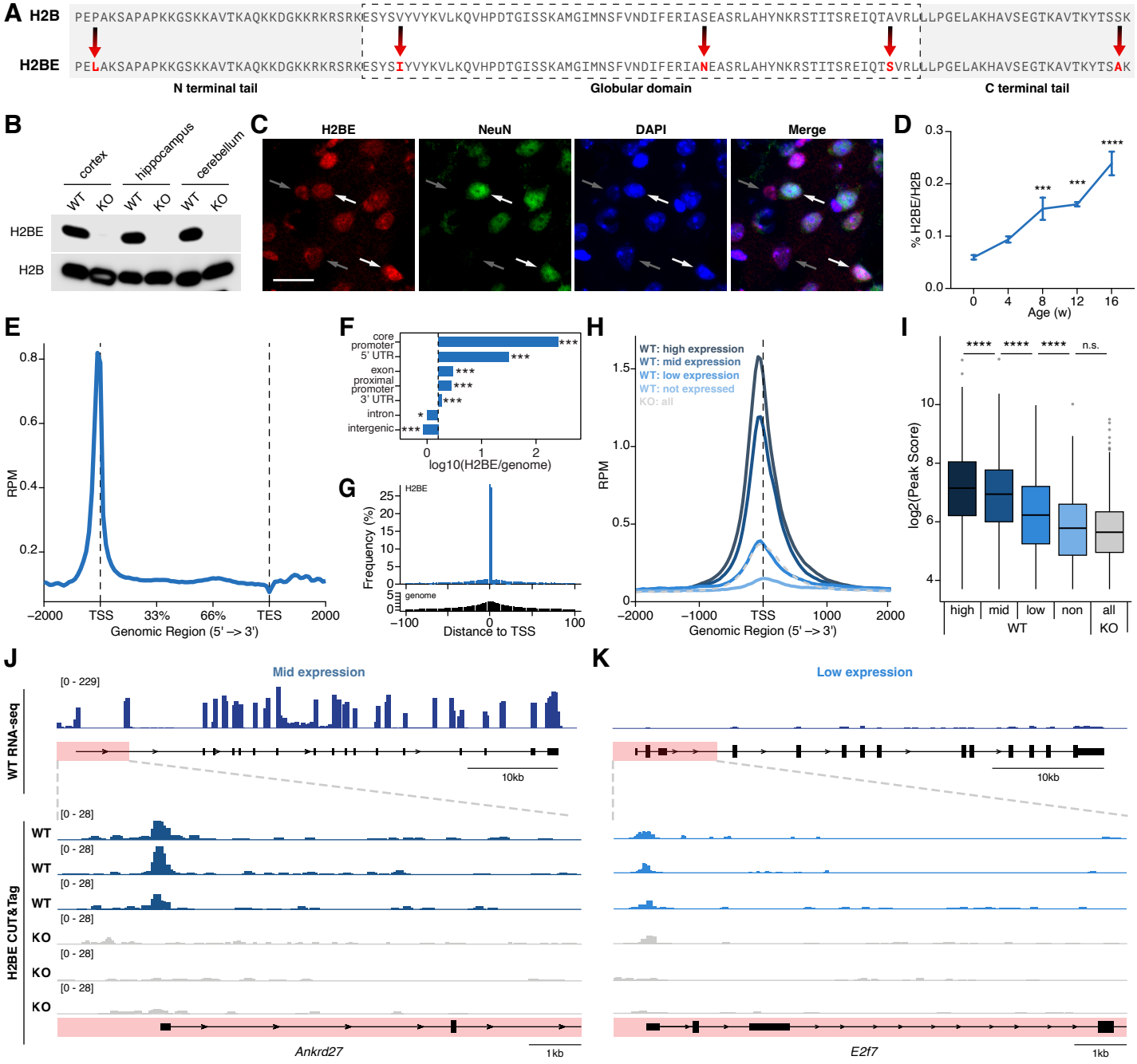


Figure 2

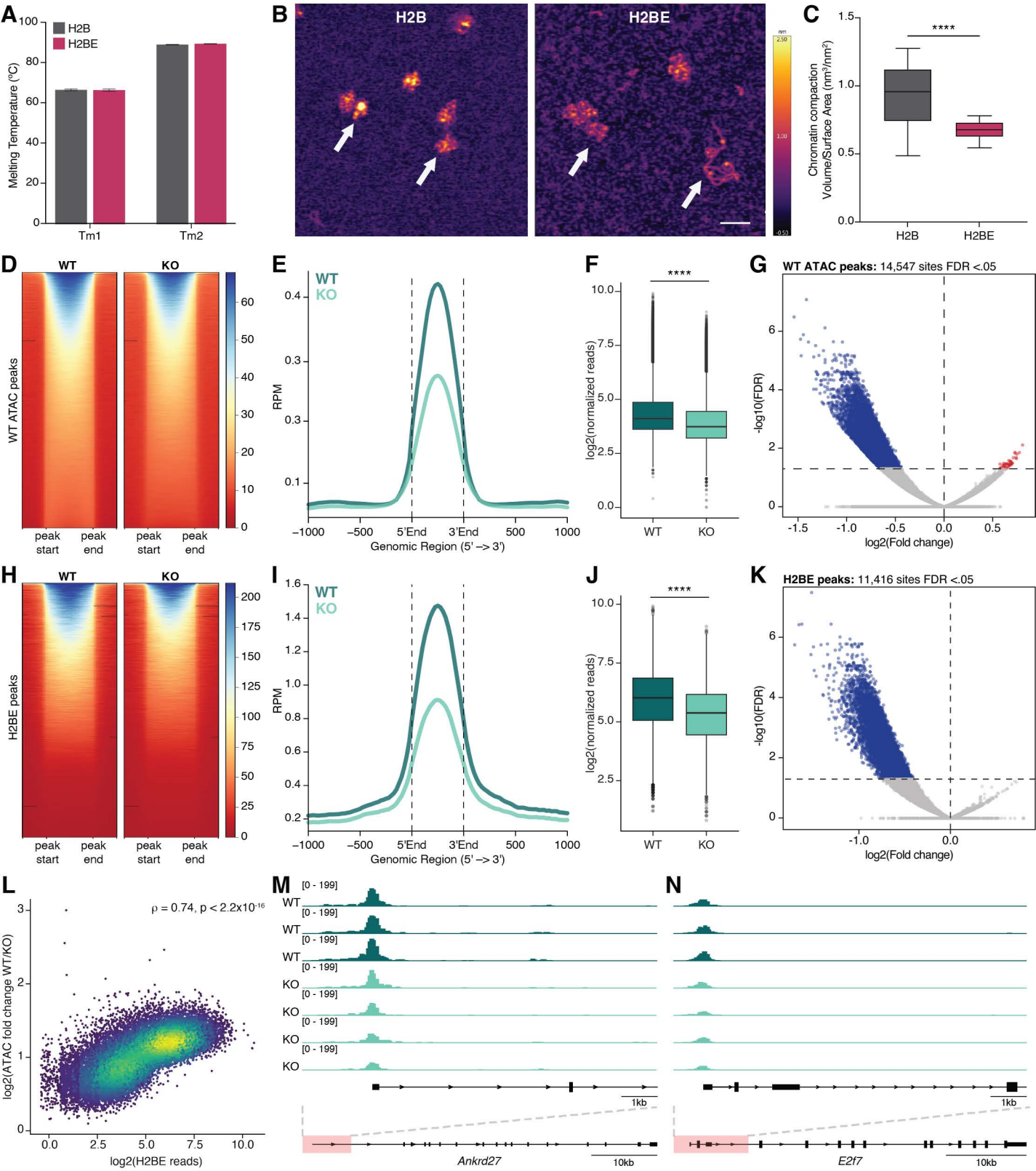
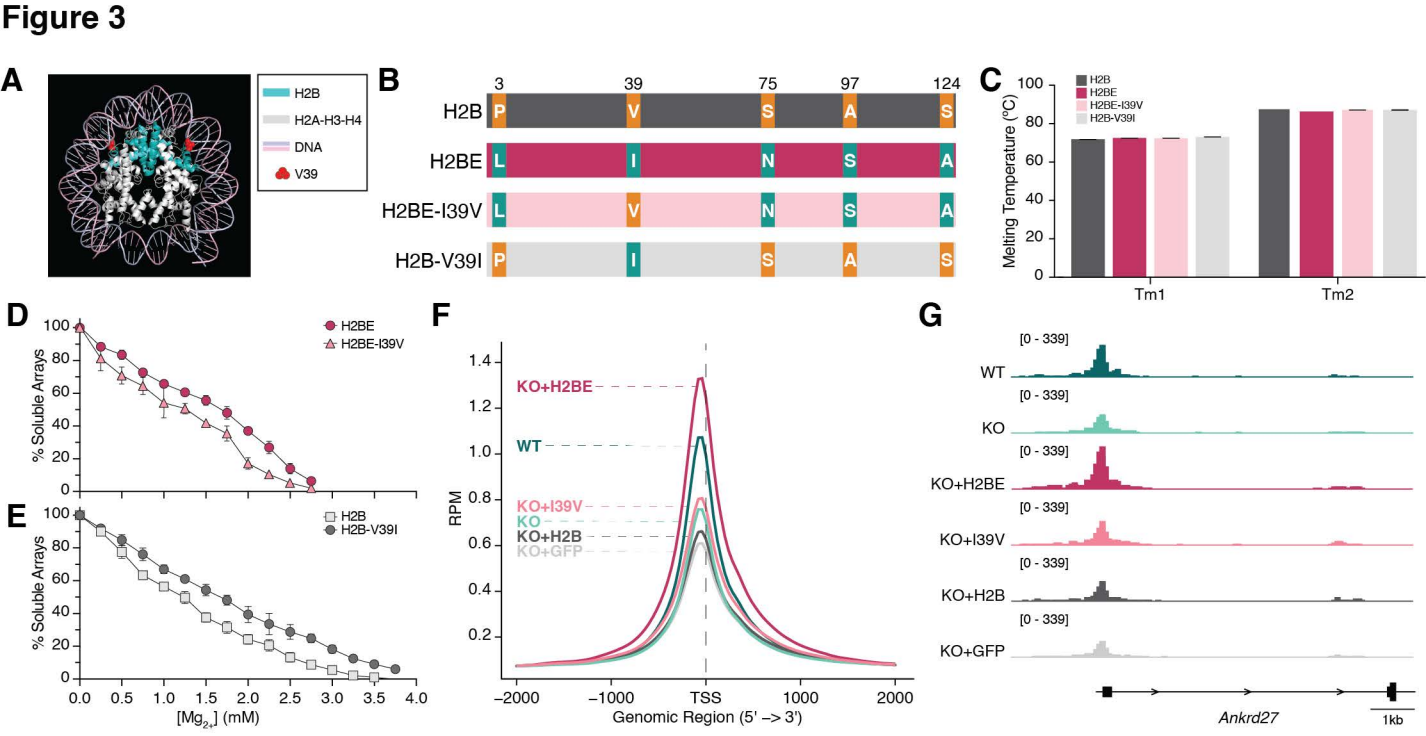

A
H2B
H2BE
Melting Temperature (°C)
Tm1
Tm2
B
H2B
H2BE
C
****
Chromatin compaction
Volume/Surface Area (nm³/nm²)
H2B
H2BE
D
WT
KO
WT ATAC peaks
peak start
peak end
E
WT
KO
RPM
Genomic Region (5' –> 3')
F
****
log2(normalized reads)
WT
KO
G
WT ATAC peaks: 14,547 sites FDR <.05
-log10(FDR)
log2(Fold change)
H
WT
KO
H2BE peaks
I
WT
KO
RPM
Genomic Region (5' –> 3')
J
****
log2(normalized reads)
K
H2BE peaks: 11,416 sites FDR <.05
-log10(FDR)
log2(Fold change)
L
ρ = 0.74, p < 2.2x10⁻¹⁶
log2(ATAC fold change WT/KO)
log2(H2BE reads)
M
WT
KO
Ankrd27
1kb
10kb
N
E2f7
1kb
10kb

Figure 3


A
H2B
H2A-H3-H4
DNA
V39
B
3
39
75
97
124
H2B
H2BE
H2BE-I39V
H2B-V39I
C
H2B
H2BE
H2BE-I39V
H2B-V39I
Melting Temperature (°C)
Tm1
Tm2
D
H2BE
H2BE-I39V
% Soluble Arrays
E
H2B
H2B-V39I
% Soluble Arrays
[Mg2+] (mM)
F
KO+H2BE
WT
KO+I39V
KO
KO+H2B
KO+GFP
RPM
TSS
Genomic Region (5' -> 3')
G
[0 - 339]
WT
KO
KO+H2BE
KO+I39V
KO+H2B
KO+GFP
Ankrd27
1kb

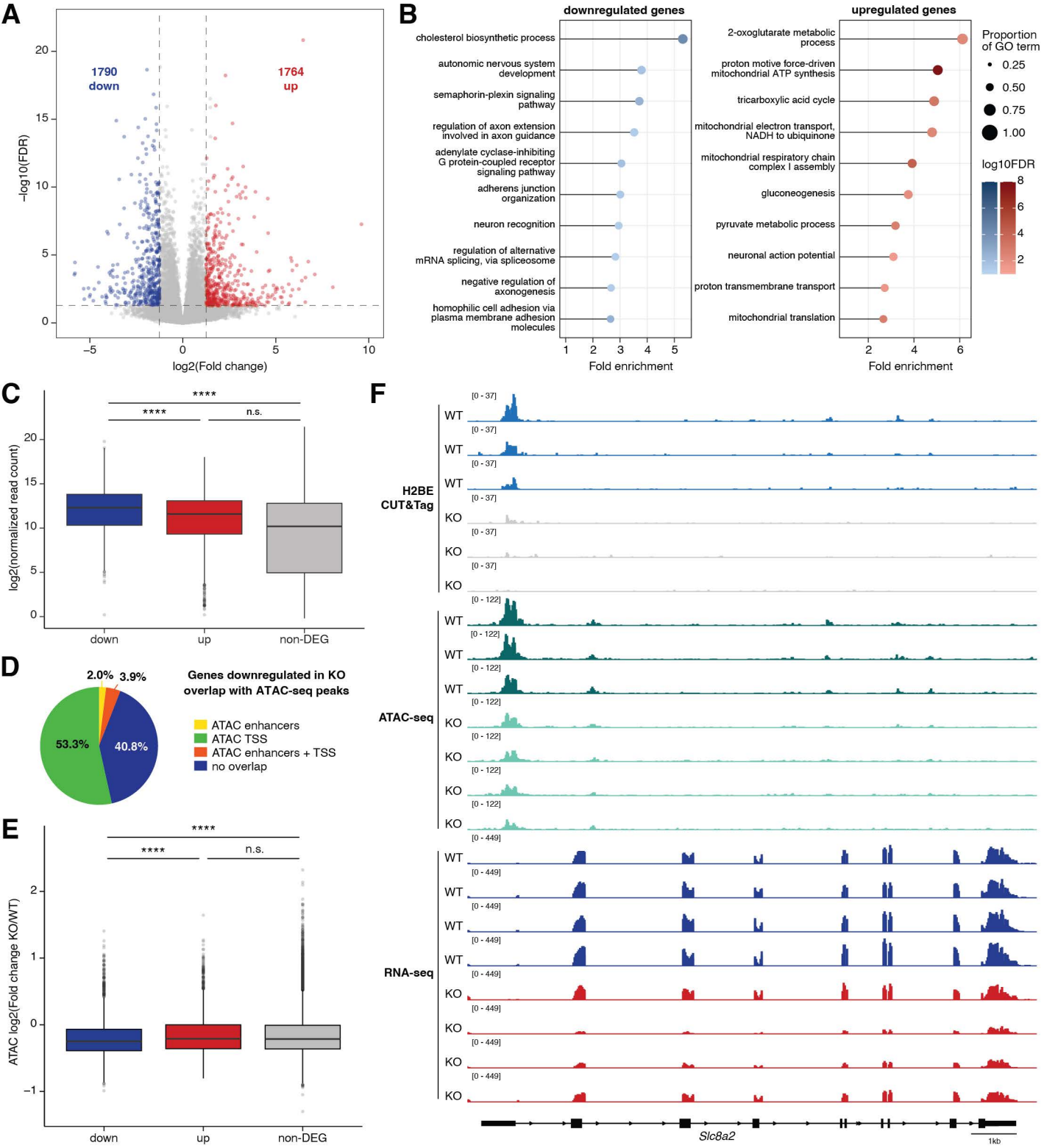

A
1790
down
1764
up
−log10(FDR)
log2(Fold change)
B
downregulated genes
cholesterol biosynthetic process
autonomic nervous system development
semaphorin-plexin signaling pathway
regulation of axon extension involved in axon guidance
adenylate cyclase-inhibiting G protein-coupled receptor signaling pathway
adherens junction organization
neuron recognition
regulation of alternative mRNA splicing, via spliceosome
negative regulation of axonogenesis
homophilic cell adhesion via plasma membrane adhesion molecules
Fold enrichment
upregulated genes
2-oxoglutarate metabolic process
proton motive force-driven mitochondrial ATP synthesis
tricarboxylic acid cycle
mitochondrial electron transport, NADH to ubiquinone
mitochondrial respiratory chain complex I assembly
gluconeogenesis
pyruvate metabolic process
neuronal action potential
proton transmembrane transport
mitochondrial translation
Fold enrichment
Proportion of GO term
0.25
0.50
0.75
1.00
log10FDR
C
log2(normalized read count)
down
up
non-DEG
****
n.s.
D
2.0%
3.9%
53.3%
40.8%
Genes downregulated in KO overlap with ATAC-seq peaks
ATAC enhancers
ATAC TSS
ATAC enhancers + TSS
no overlap
E
ATAC log2(Fold change KO/WT)
down
up
non-DEG
****
n.s.
F
H2BE CUT&Tag
ATAC-seq
RNA-seq
WT
KO
Slc8a2
1kb

Figure 5

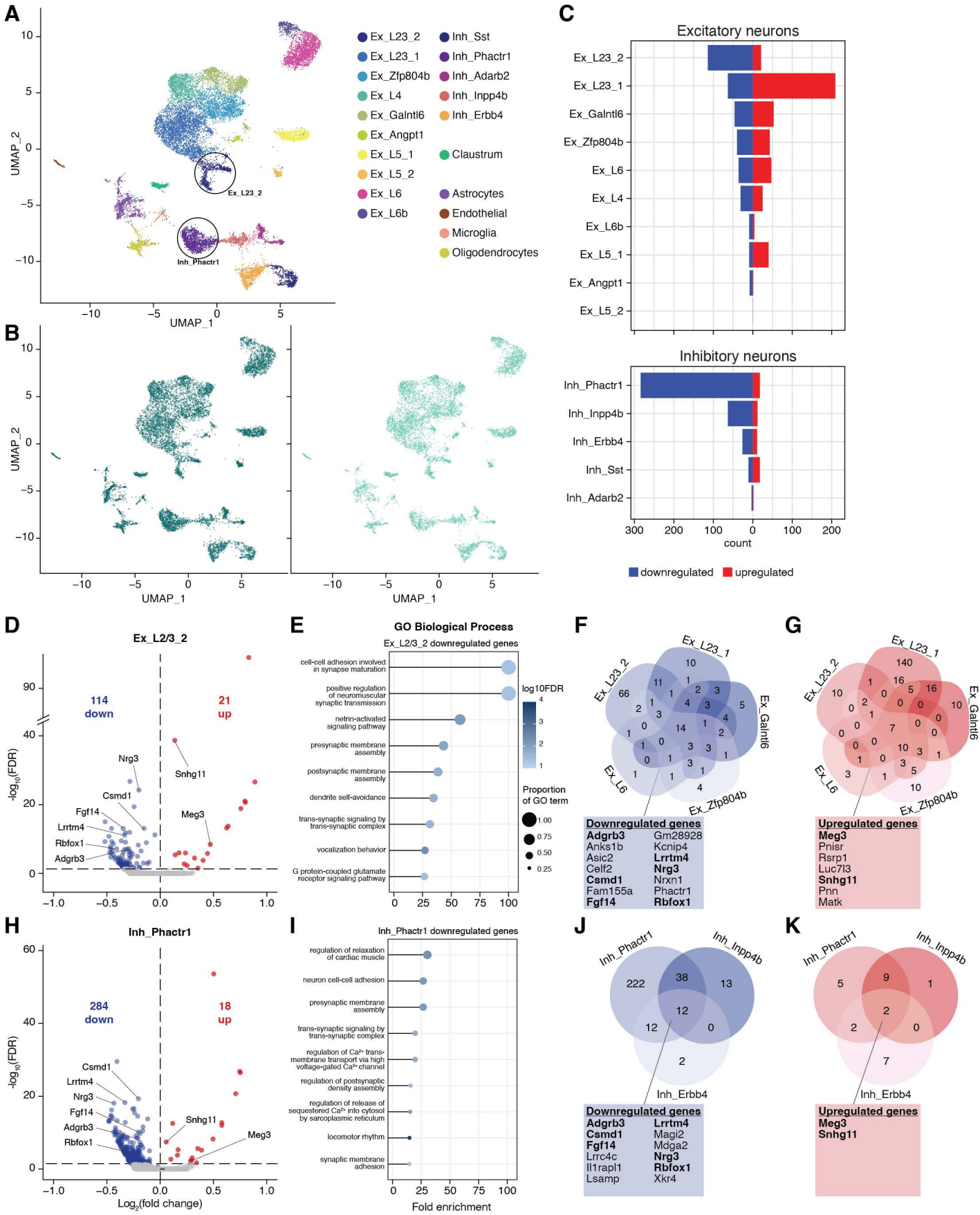

A
Ex_L23_2
Ex_L23_1
Ex_Zfp804b
Ex_L4
Ex_Galntl6
Ex_Angpt1
Ex_L5_1
Ex_L5_2
Ex_L6
Ex_L6b
Inh_Sst
Inh_Phactr1
Inh_Adarb2
Inh_Inpp4b
Inh_Erbb4
Claustrum
Astrocytes
Endothelial
Microglia
Oligodendrocytes
UMAP_1
UMAP_2
B
C
Excitatory neurons
Inhibitory neurons
count
downregulated
upregulated
D
Ex_L2/3_2
114 down
21 up
Nrg3
Snhg11
Csmd1
Fgf14
Meg3
Lrrtm4
Rbfox1
Adgrb3
-log10(FDR)
E
GO Biological Process
Ex_L2/3_2 downregulated genes
cell-cell adhesion involved in synapse maturation
positive regulation of neuromuscular synaptic transmission
netrin-activated signaling pathway
presynaptic membrane assembly
postsynaptic membrane assembly
dendrite self-avoidance
trans-synaptic signaling by trans-synaptic complex
vocalization behavior
G protein-coupled glutamate receptor signaling pathway
log10FDR
Proportion of GO term
F
Ex_L23_1
Ex_L23_2
Ex_Galntl6
Ex_L6
Ex_Zfp804b
Downregulated genes
Adgrb3 Gm28928
Anks1b Kcnip4
Asic2 Lrrtm4
Celf2 Nrg3
Csmd1 Nrxn1
Fam155a Phactr1
Fgf14 Rbfox1
G
Upregulated genes
Meg3
Pnisr
Rsrp1
Luc7l3
Snhg11
Pnn
Matk
H
Inh_Phactr1
284 down
18 up
Csmd1
Lrrtm4
Nrg3
Fgf14
Adgrb3
Rbfox1
Snhg11
Meg3
Log2(fold change)
I
Inh_Phactr1 downregulated genes
regulation of relaxation of cardiac muscle
neuron cell-cell adhesion
presynaptic membrane assembly
trans-synaptic signaling by trans-synaptic complex
regulation of Ca2+ transmembrane transport via high voltage-gated Ca2+ channel
regulation of postsynaptic density assembly
regulation of release of sequestered Ca2+ into cytosol by sarcoplasmic reticulum
locomotor rhythm
synaptic membrane adhesion
Fold enrichment
J
Inh_Phactr1
Inh_Inpp4b
Inh_Erbb4
Downregulated genes
Adgrb3 Lrrtm4
Csmd1 Magi2
Fgf14 Mdga2
Lrrc4c Nrg3
Il1rapl1 Rbfox1
Lsamp Xkr4
K
Upregulated genes
Meg3
Snhg11

Figure 6

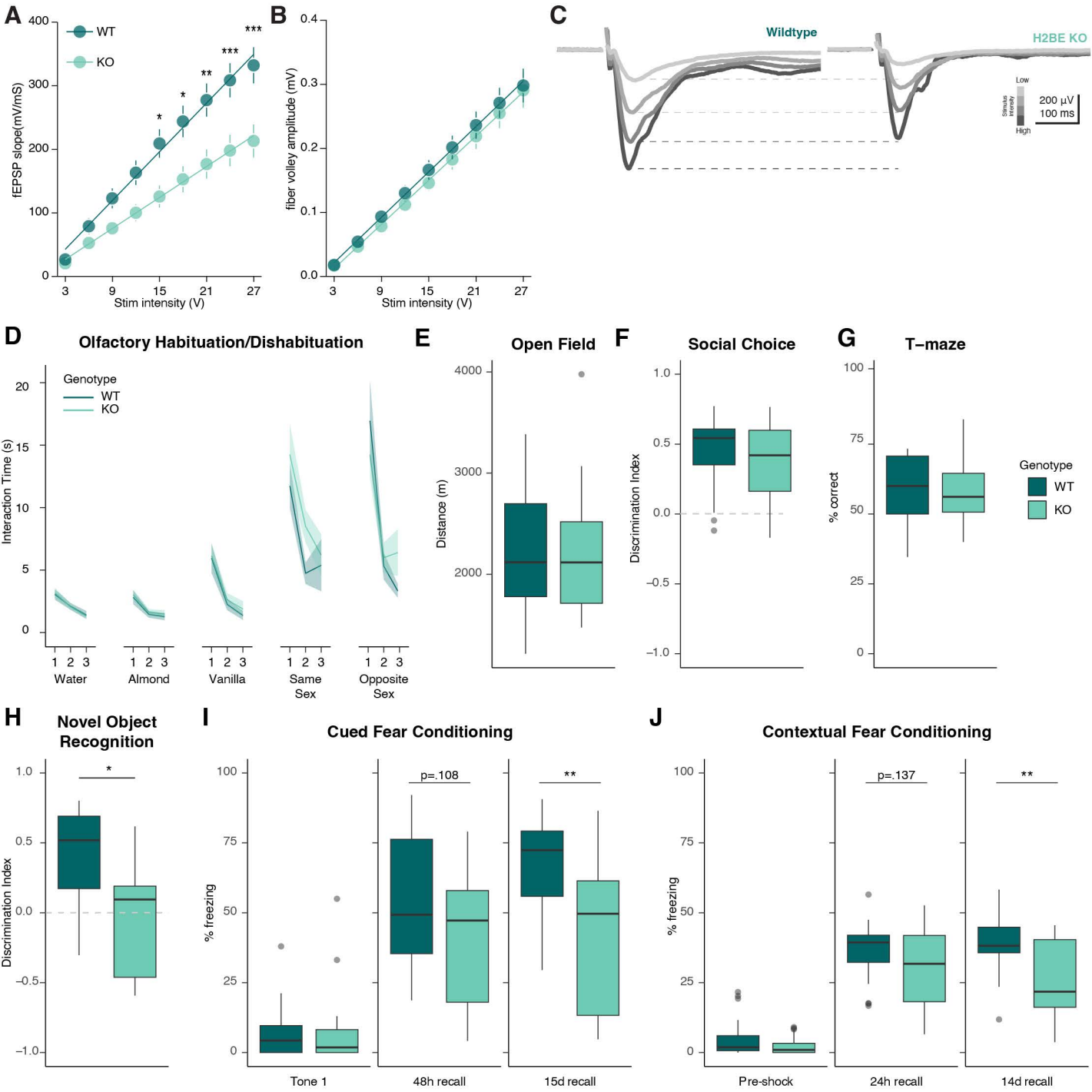
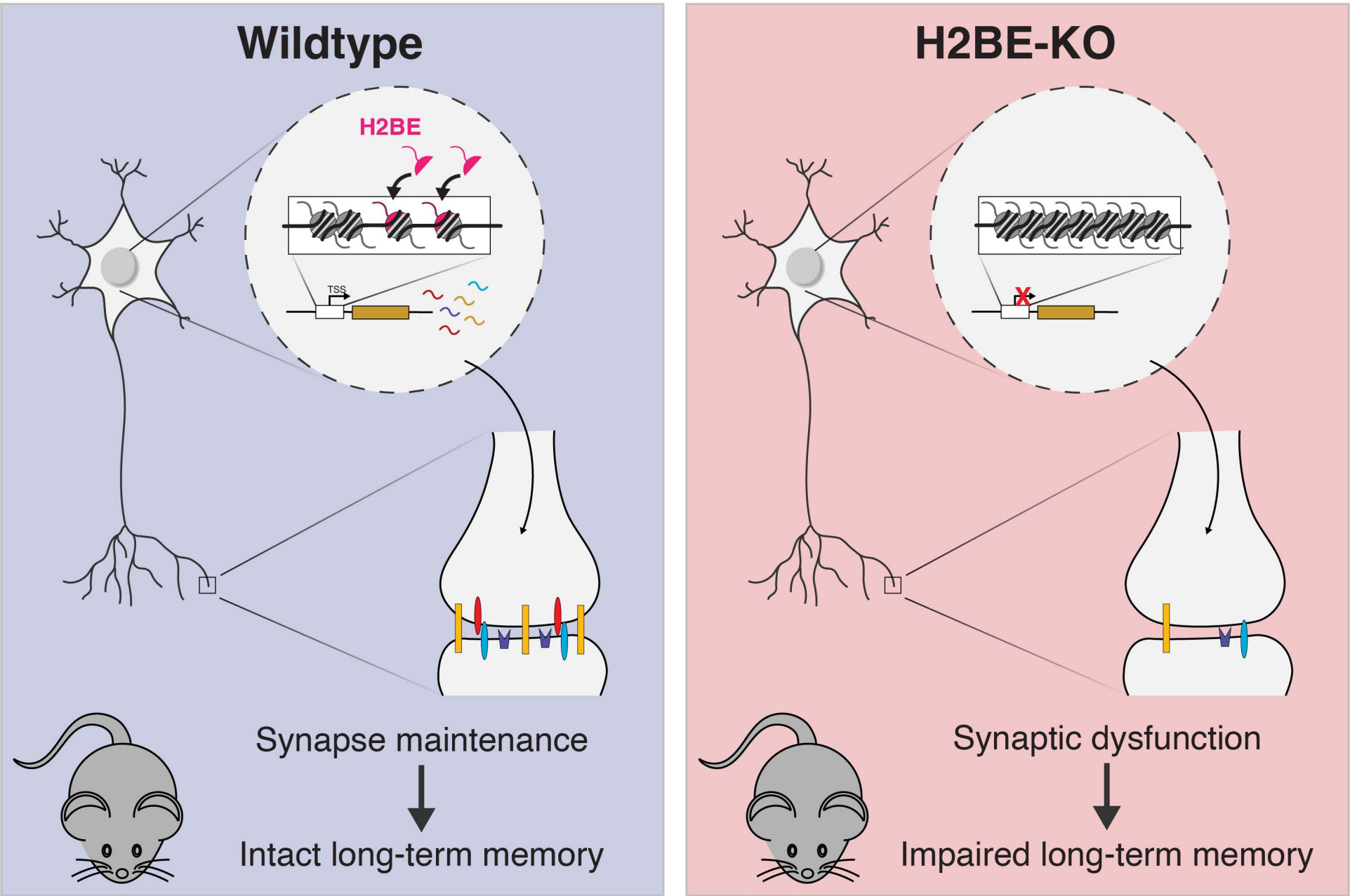
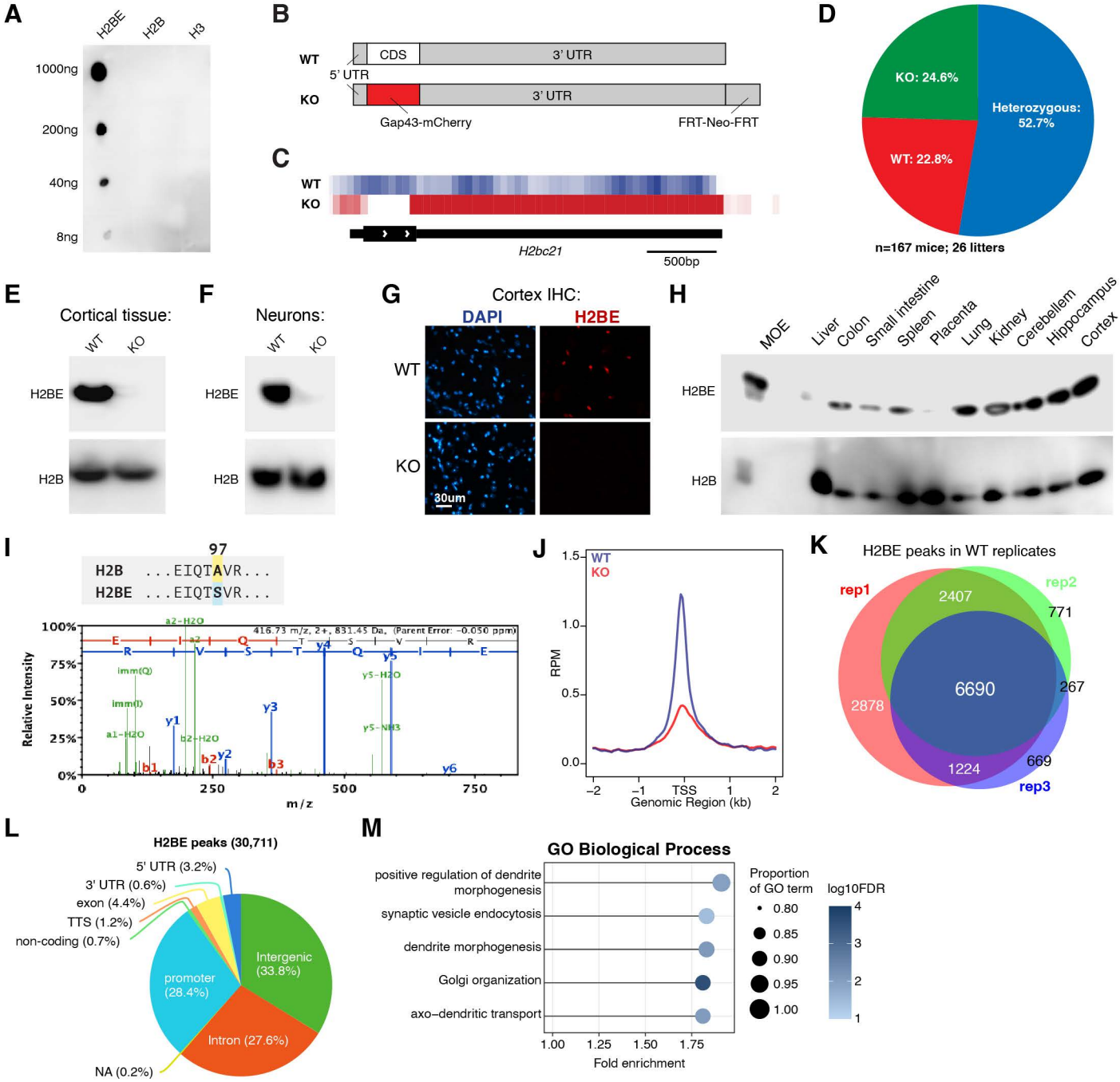

A
WT
KO
fEPSP slope(mV/mS)
Stim intensity (V)
B
fiber volley amplitude (mV)
Stim intensity (V)
C
Wildtype
H2BE KO
Low
Stimulus Intensity
High
200 µV
100 ms
D
Olfactory Habituation/Dishabituation
Genotype
WT
KO
Interaction Time (s)
Water
Almond
Vanilla
Same Sex
Opposite Sex
E
Open Field
Distance (m)
F
Social Choice
Discrimination Index
G
T–maze
% correct
Genotype
WT
KO
H
Novel Object Recognition
Discrimination Index
I
Cued Fear Conditioning
% freezing
p=.108
Tone 1
48h recall
15d recall
J
Contextual Fear Conditioning
% freezing
p=.137
Pre-shock
24h recall
14d recall

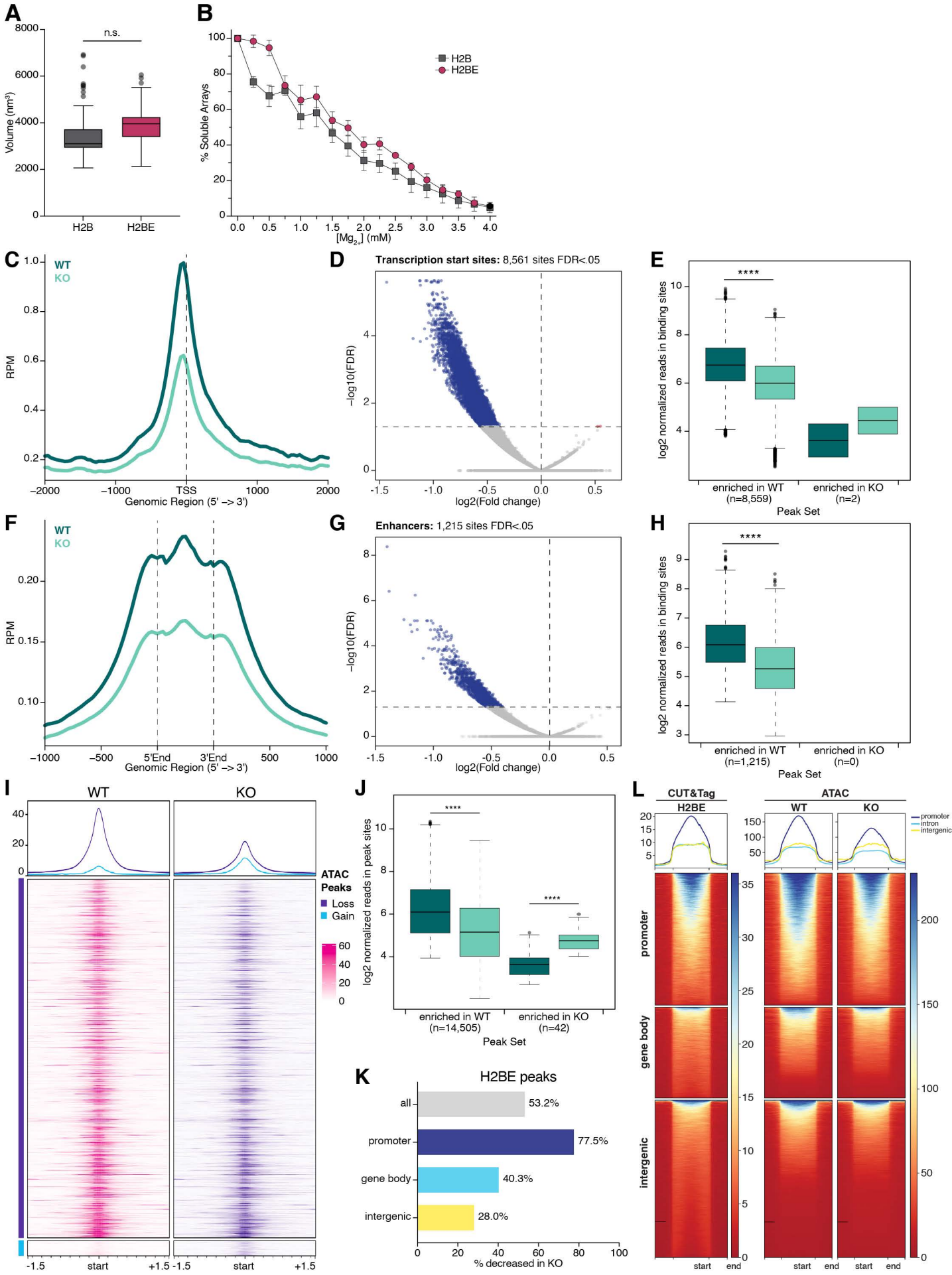
Figure 7


Wildtype
H2BE
TSS
Synapse maintenance
Intact long-term memory
H2BE-KO
Synaptic dysfunction
Impaired long-term memory

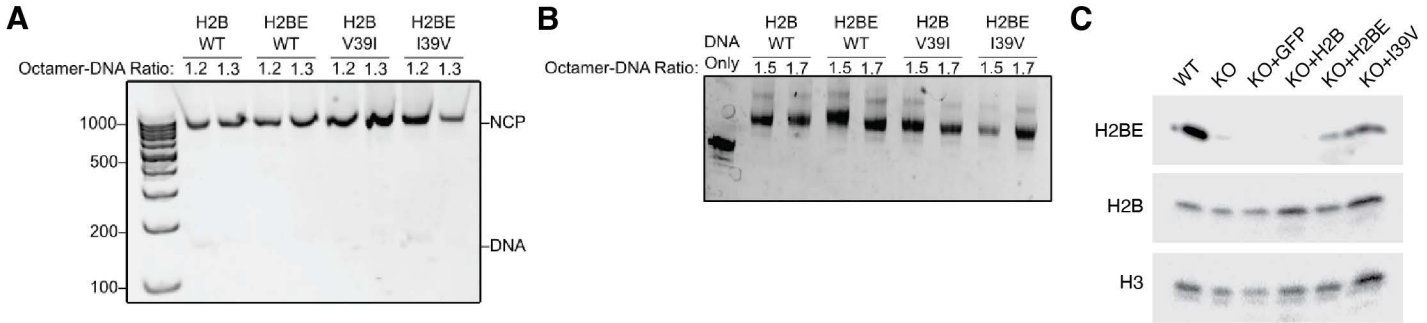
Supplemental Figure 1



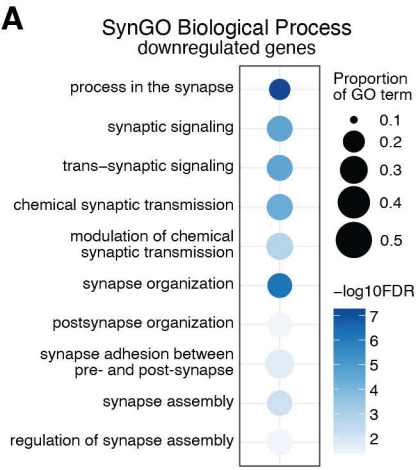
Supplemental Figure 2

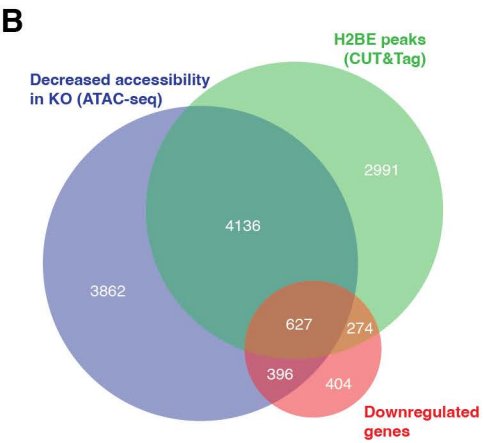


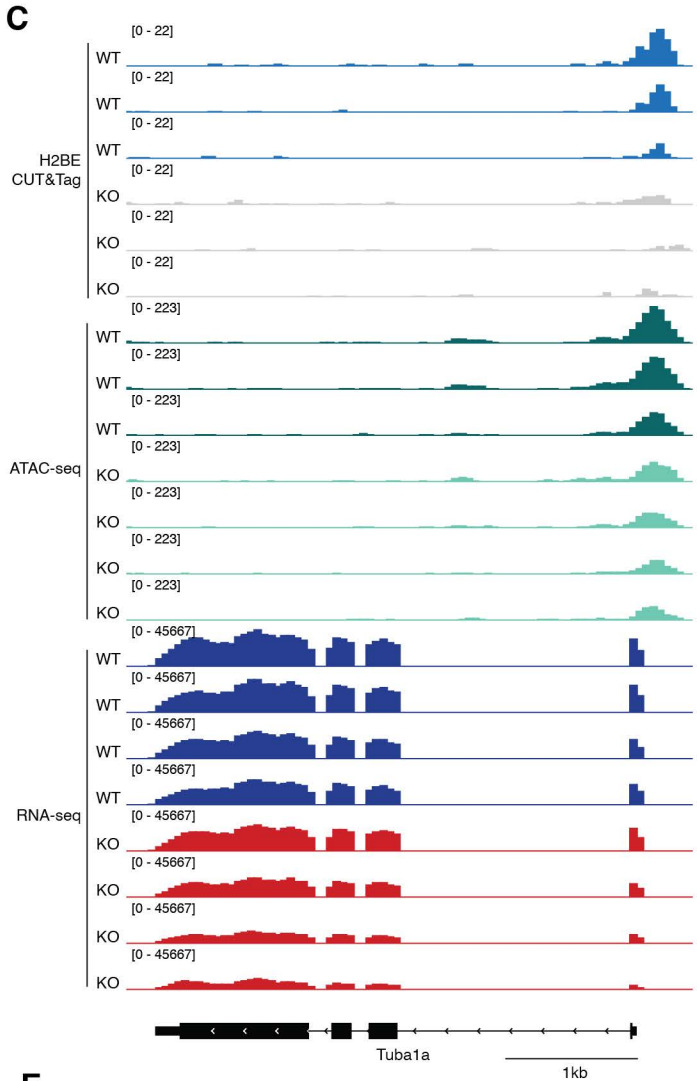
Supplemental Figure 3

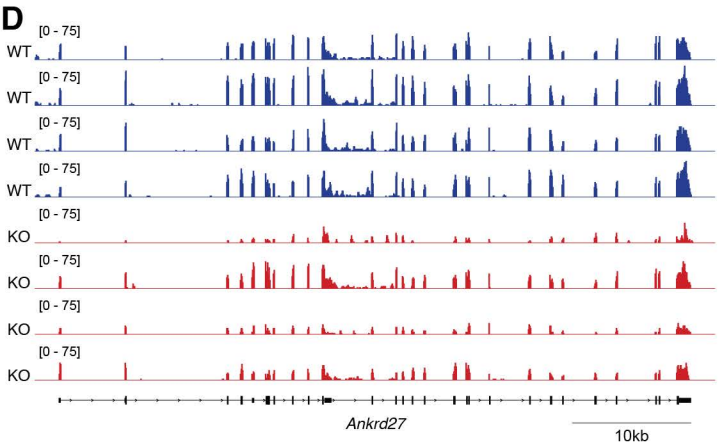


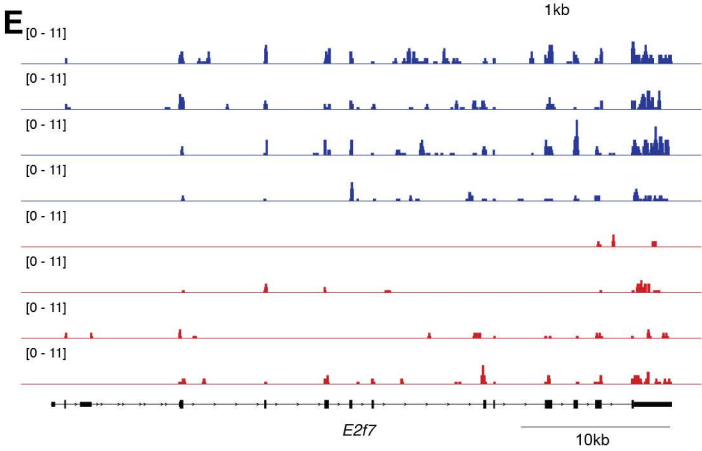
Supplemental Figure 4



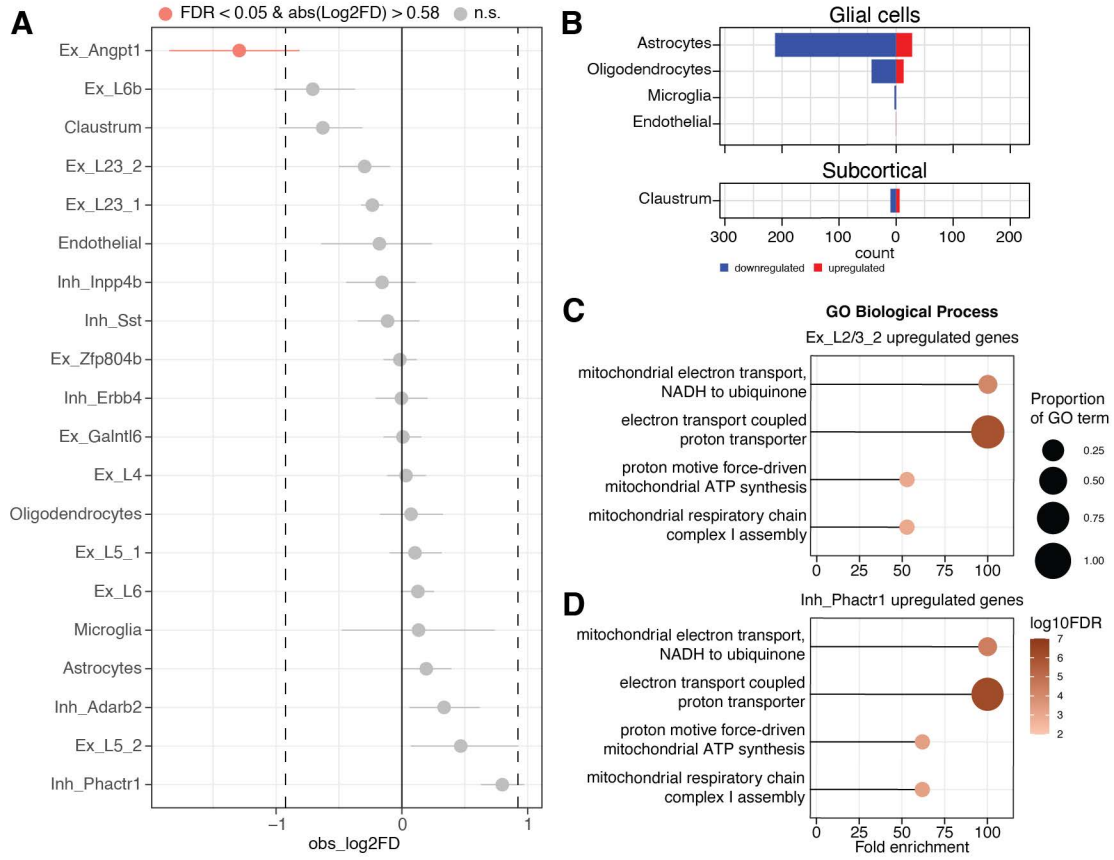
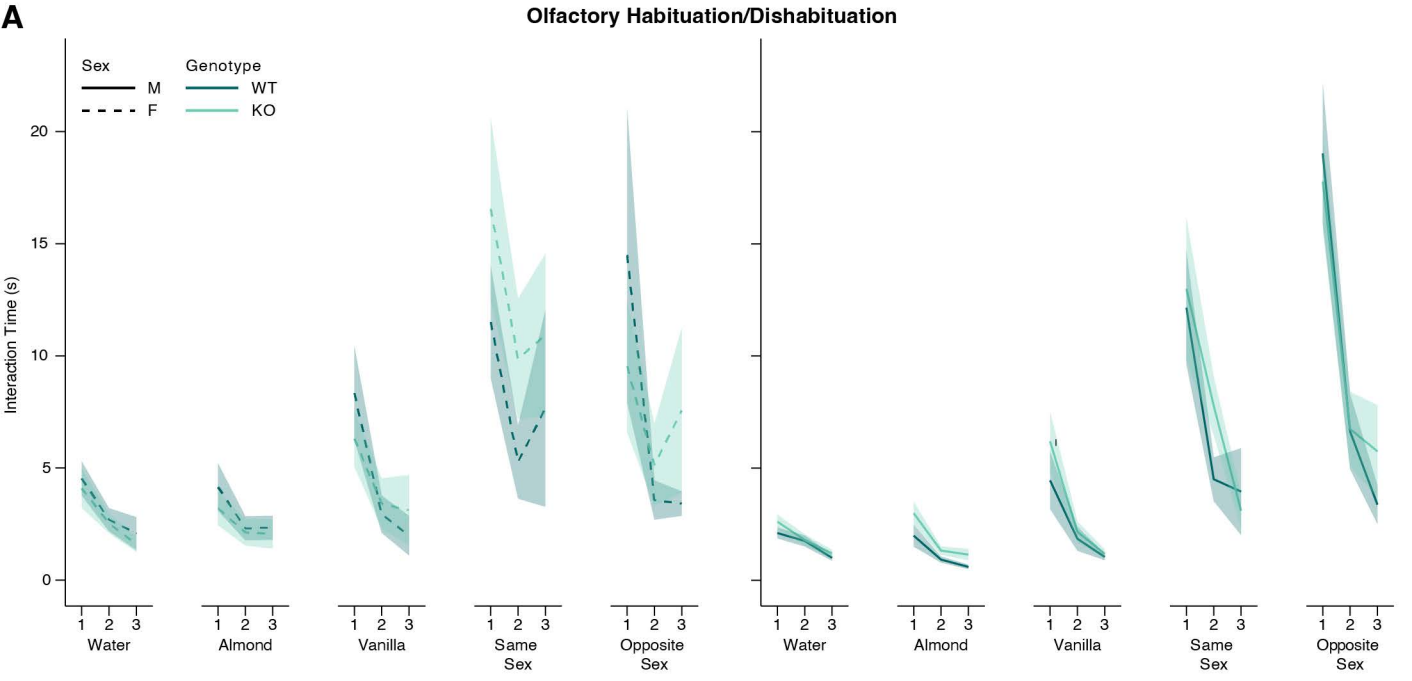
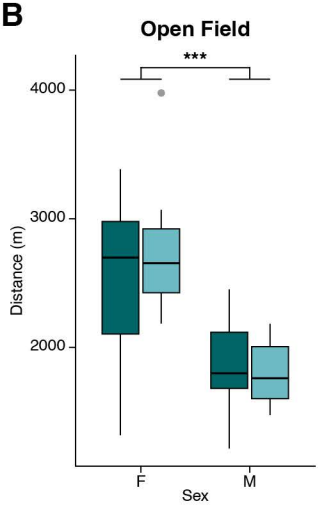
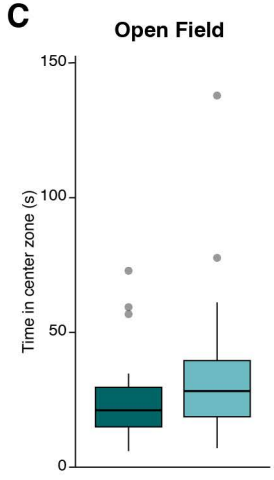
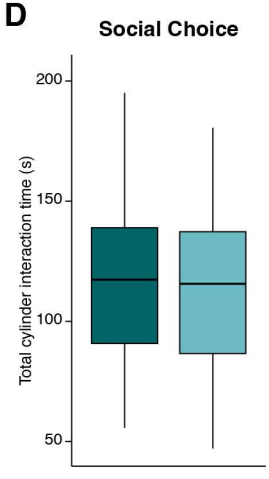
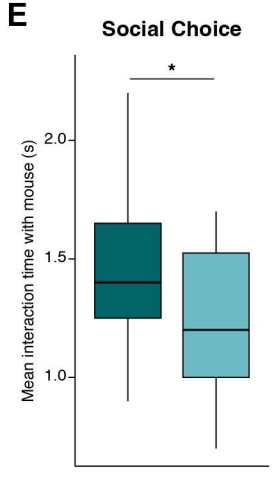
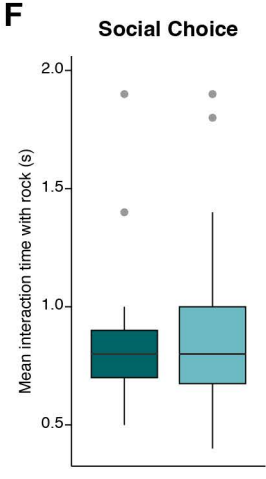
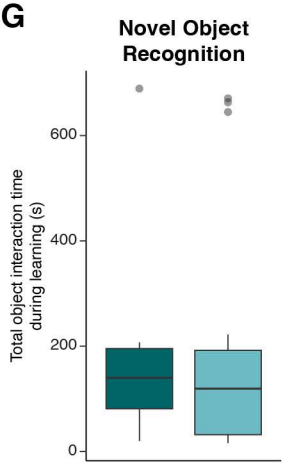
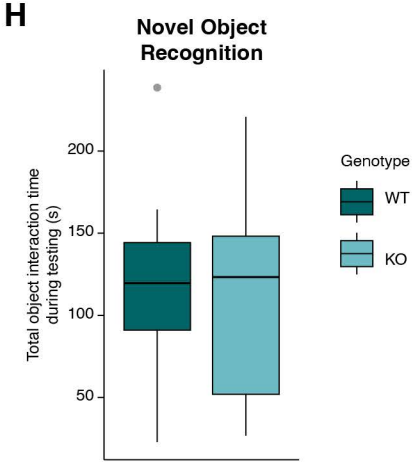








Supplemental Figure 5



Supplemental Figure 6