HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2021 May 09.

Published in final edited form as:

Cell Rep. 2021 April 27; 35(4): 109042. doi:10.1016/j.celrep.2021.109042.

5-hydroxymethylcytosine is dynamically regulated during forebrain organoid development and aberrantly altered in Alzheimer's disease

Janise N. Kuehner[1,7], **Junyu Chen**[1,7], **Emily C. Bruggeman**[1], **Feng Wang**[1], **Yangping Li**[1], **Chongchong Xu**[2,4,5], **Zachary T. McEachin**[4], **Ziyi Li**[6], **Li Chen**[3], **Chadwick M. Hales**[5], **Zhexing Wen**[2,4,5,*], **Jingjing Yang**[1,*], **Bing Yao**[1,8,*]

[1]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA

[2]Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, USA

[3]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN, USA

[4]Department of Cell Biology, Emory University School of Medicine, Atlanta, GA, USA

[5]Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

[6]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[7]These authors contributed equally

[8]Lead contact

SUMMARY

5-hydroxymethylcytosine (5hmC) undergoes dynamic changes during mammalian brain development, and its dysregulation is associated with Alzheimer's disease (AD). The dynamics of 5hmC during early human brain development and how they contribute to AD pathologies remain largely unexplored. We generate 5hmC and transcriptome profiles encompassing several developmental time points of healthy forebrain organoids and organoids derived from several familial AD patients. Stage-specific differentially hydroxymethylated regions demonstrate an acquisition or depletion of 5hmC modifications across developmental stages. Additionally, genes concomitantly increasing or decreasing in 5hmC and gene expression are enriched in

*Correspondence: zhexing.wen@emory.edu (Z.W.), jingjing.yang@emory.edu (J.Y.), bing.yao@emory.edu (B.Y.).

AUTHOR CONTRIBUTIONS

Conceptualization, J.N.K., J.C., Z.W., J.Y., and B.Y.; formal analysis, J.N.K., J.C., Y.L., L.C., and Z.L.; investigation, E.C.B., F.W., and C.X.; resources, Z.T.M., C.M.H., C.X., and Z.W.; writing – original draft, J.N.K., J.C., E.C.B., Z.W., J.Y., and B.Y.; writing – review & editing, J.N.K., J.C., Y.L., Z.W., J.Y., and B.Y.; visualization, J.N.K. and J.C.; supervision and project administration, Z.W., J.Y., and B.Y.; funding acquisition, B.Y., J.N.K., J.Y., and Z.W.

DECLARATION OF INTERESTS

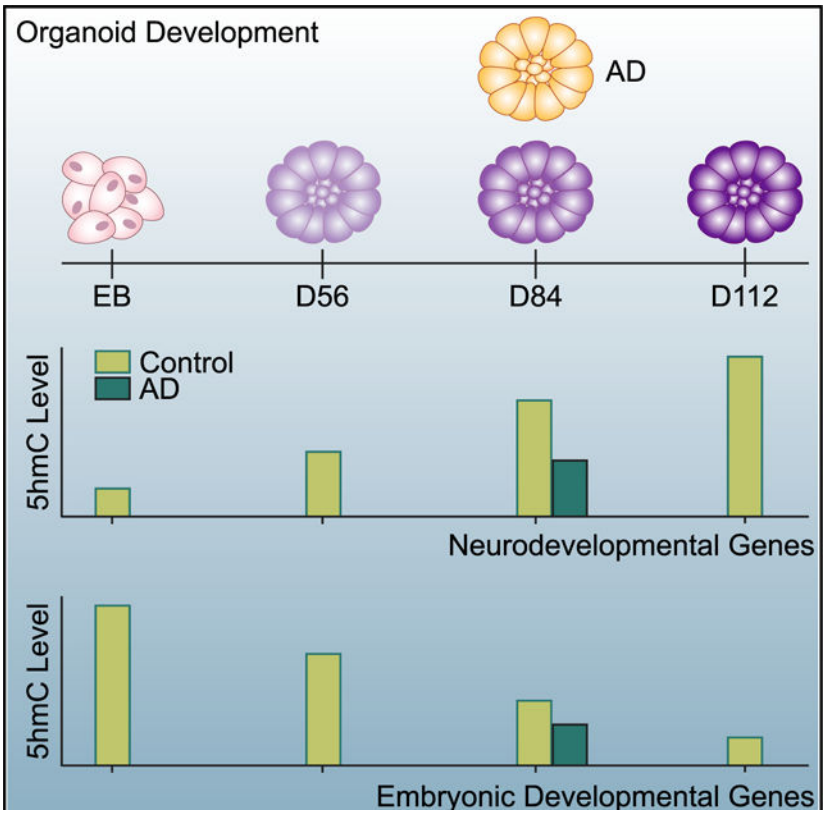
The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109042.

neurobiological or early developmental processes, respectively. Importantly, our AD organoids corroborate cellular and molecular phenotypes previously observed in human AD brains. 5hmC is significantly altered in developmentally programmed 5hmC intragenic regions in defined fetal histone marks and enhancers in AD organoids. These data suggest a highly coordinated molecular system that may be dysregulated in these early developing AD organoids.

Graphical Abstract



In brief

Kuehner et al. use forebrain organoids derived from healthy controls to study the dynamics of 5hmC across early brain development. In addition, organoids derived from several AD patients reveal aberrant 5hmC patterns that could disrupt early neuronal networks and contribute to the onset of AD later in life.

INTRODUCTION

Epigenetics refers to heritable changes in gene expression without altering the DNA sequence through mechanisms such as DNA and histone modifications and non-coding RNAs. Mounting evidence implicates critical roles for DNA modifications, specifically 5-hydroxymethylcytosine (5hmC), in regulating brain development (Kuehner et al., 2019). 5hmC emerged as a key DNA modification in the nervous system due to its significant enrichment in the brain (Kriaucionis and Heintz, 2009) and its ability to regulate neuronal-

specific gene expression during neural progenitor cell differentiation (Li et al., 2017). Due to the sample restraint, very few studies have investigated the 5hmC landscape during human early brain development spanning several developmental stages (Wen et al., 2014a). Previous studies that profiled 5hmC in human fetal brain tissues have lacked comprehensive genome-wide coverage that expands beyond the coding regions of the genome (Spiers et al., 2017). Understanding the continuous dynamics of 5hmC throughout early brain development could reveal how crucial neurodevelopmental milestones are attained, and how failure to achieve these milestones could be detrimental to normal brain development and function, or even contribute to neurological diseases.

Genome-wide sequencing studies have suggested that abnormalities in 5hmC distribution and function could be critical factors contributing to Alzheimer's disease (AD) (Shu et al., 2016; Zhang et al., 2020). AD is the most common neurodegenerative disease worldwide (James et al., 2014) and is characterized by extensive memory loss and cognitive impairments and the accumulation of amyloid-beta (Aβ) plaques, phosphorylated Tau, and neurofibrillary tangles (Hardy and Higgins, 1992). Despite considerable efforts, the molecular mechanisms underlying AD pathogenesis remain elusive, especially before the onset of AD pathology and symptoms. Evidence suggests that long before the cognitive impairments of AD manifest, there are structural and functional brain defects (Busche and Konnerth, 2016; Sun et al., 2009); however, it is unknown whether alterations in DNA modifications have also manifested. Initial attempts to profile 5hmC in AD brains have revealed conflicting results due to differences between species (mouse versus human), lack of comprehensive brain developmental time points, and postmortem delay (Bradley-Whitman and Lovell, 2013; Condliffe et al., 2014). Genome-wide and brain-region-specific 5hmC profiling in late-stage mouse and postmortem AD samples has detected a global reduction in 5hmC as well as differentially hydroxymethylated regions (DhMRs) (Bernstein et al., 2016; Zhang et al., 2020). These initial studies have laid the foreground for further investigation into how aberrations in the 5hmC landscape could contribute to AD pathology.

Transgenic mouse and human postmortem brains have been leading models for studying the basic mechanisms and human-specific features of late-stage AD, respectively. Another human-specific model for AD is the 3-dimensional brain organoid derived from human induced pluripotent stem cells (iPSCs) that recapitulates fetal brain development at the molecular level (Qian et al., 2016). Brain organoids are an attractive model system for studying early development and neurological diseases because they (1) can model disease progression spanning a comprehensive timeline, (2) retain the complexity of a multicellular tissue/organ while being maintained in a cell-culture-like environment (Takebe and Wells, 2019), and (3) can recapitulate human brain development *in vitro* (Itskovitz-Eldor et al., 2000). Therefore, we generated 5hmC and transcriptome profiles encompassing several developmental time points of healthy forebrain organoids and organoids derived from several familial AD (fAD) patients. Our organoids are comparable to early human fetal brain development, spanning the 12- to 24-week post-conception period (Qian et al., 2016; Kelava and Lancaster, 2016). Furthermore, they allow us to obtain a comprehensive picture of 5hmC dynamics during early neural development and how aberrations in 5hmC might contribute to AD.

RESULTS

Genome-wide profiling of 5hmC in forebrain organoids during early brain development

Forebrain organoids were cultured from a healthy iPSC line using miniature SpinΩ bioreactors (Qian et al., 2016) to study the dynamics of 5hmC during early brain development. We used embryoid body (EB) and forebrain organoids that had been cultured for 8 days (EBs), 56 days (D56), 84 days (D84), and 112 days (D112) (Figure 1A) to model the early developing fetal brain. Using a 5hmC-selective chemical labeling method (hMe-Seal or 5hmC sequencing [5hmC-seq]) (Song et al., 2011), we generated genome-wide 5hmC profiles from replicated samples at each of the organoid developmental time points. Our 5hmC and RNA-sequencing (RNA-seq) data showed high Pearson correlations (>0.9) and clustered together by principal-component analysis (PCA), indicating sufficient reproducibility among multiple sample replicates (Figures S1A–S1F). The different profiles of EBs compared to mature organoids occur due to their distinct cellular composition (Qian et al., 2016).

Using the peak-calling tool MACS2, we observed that ~50% more 5hmC peaks were called in EBs compared to the later time points (Figure 1B), which is in agreement with our global 5hmC dot blot and quantification analysis (Figures S1G and S1H). In general, 5hmC is largely distributed in intronic and intergenic regions, specifically repetitive elements (Figures 1C and S1I), and the proportion of 5hmC peaks in coding regions appeared to be relatively constant across all stages. Notably, the identified 5hmC regions do not appear to be the same regions retaining 5hmC peaks during development (Figure S1J). Finally, enrichment analysis revealed that 5hmC peaks were enriched in intragenic regions and depleted in intergenic regions across the human genome (Figure 1D), which is consistent with previous findings (Szulwach et al., 2011).

We next focused on the dynamic 5hmC patterns during organoid neurodevelopment. Average 5hmC read counts were plotted globally (Figures 1E, 1G, 1I, and 1K) and across gene bodies, promoters, and intergenic regions (Figure S1K). We observed that EBs have a distinct 5hmC pattern from the D56, D84, and D112 organoids, which were all comparable (Figures 1E, 1G, 1I, and 1K). For example, 5hmC is more enriched in the gene body and promoter region of ankyrin 1 (*ANK1*), a gene important for cellular proliferation, in EBs compared to the later developmental stages (Figure 1F). Other neurodevelopmental-specific genes such as *DRD2*, *NTRK1*, and *TUBB2B* have stage-specific 5hmC enrichment in D56, D84, and D112, respectively (Figures 1H, 1J, and 1L). Based on the data presented, we have shown that the 5hmC landscape is distinct from the multipotent EB stage to the neural-lineage-committed developing organoid stages.

Dynamics of 5hmC regulation during forebrain organoid development

To investigate the detailed dynamics of 5hmC regulation during forebrain organoid development, we identified differentially hydroxymethylated regions (DhMRs) at each developmental stage. By comparing 5hmC peaks identified in D56 organoids with those from EB samples, we found this transition generated the most DhMRs (Figure 2A). As the organoids became further differentiated, the number of established and disappearing DhMRs

continued to decrease but remained evenly distributed across all chromosomes (Figure S2A). Given that a substantial amount of DhMR fluctuation occurs early in neurodevelopment, we identified stage-specific DhMRs (Tables S1 and S2). We observed a total of 101,907 stage-specific 5hmC-enriched peaks (Figure 2B; Table S1). In the EB-specific 5hmC-enriched peaks, functional analysis revealed that the annotated genes were largely within developmental genes such as *TNF* (Figure S2B). Comparatively, stage-specific 5hmC peaks of the other 3 stages are enriched within genes critical for nervous system development and other neurobiological processes (e.g., *PAX6*, *AKT1*, and *SNCA*) (Figures S2C–S2E). Similar analyses were completed for stage-specific 5hmC-depleted regions (Figure 2C; Table S2). Both the EB- and D56-specific 5hmC-depleted regions showed a reduction of 5hmC in key neurodevelopmental genes such as *NEUROG1* (Figure S2F) and *GFAP* (Figure S2G), suggesting that their expression may be specific to more mature neurodevelopmental stages. As expected, D84- and D112-specific 5hmC-depleted regions were located in developmental genes such as *WNT10A* (Figure S2H) and *FGFR1* (Figure S2I) and involved in repressing multicellular organism developmental processes. Collectively, these findings support 5hmC as a critical epigenomic mark for brain development, especially during the differentiation of early nervous system structures to mature brain structures.

We next investigated DhMRs that showed continual 5hmC accumulation or depletion across the developmental stages, because these regions are more likely to be programmed and important for proper development (Figure 2D). Among the total identified DhMRs, we found that 13,249 (13%) showed accumulation and 19,350 (19%) showed depletion in 5hmC levels (Figure S2J), giving a total of 32,599 DhMRs of interest. Interestingly, DhMRs with continual 5hmC accumulation appeared to gradually gain 5hmC modifications during development (Figure 2E), whereas DhMRs with continual 5hmC depletion showed an instantaneous loss of 5hmC from the EB to other stages (Figure 2F). These data suggest distinct dynamics for 5hmC acquisition versus 5hmC depletion throughout neuronal development.

Intragenic 5hmC is positively associated with gene expression (Pastor et al., 2011), and thus we further explored genes that harbored DhMRs with continual 5hmC accumulation or depletion and showed a continual increase or decrease in gene expression as determined by our RNA-seq data (Figure 2G). We found 314 concomitantly increasing genes and 171 non-concomitantly increasing genes across the developmental stages (Figure S2K). Using Gene Ontology (GO) analysis, we found that the 314 concomitantly increasing genes were enriched in neurodevelopmental processes, supporting the importance of 5hmC accumulation for proper brain development (Figure 2H). These results were also confirmed using the Genomic Regions Enrichment of Annotations Tool (Figure S2N). *SOX11*, a critical transcription factor in embryo and brain development (Bergsland et al., 2006), was identified among this group of genes (Figure 2J). The 171 non-concomitantly increasing genes did not reveal any significant biological processes, potentially suggesting an indirect correlation with 5hmC and less relevance to neurodevelopmental processes. Analysis of the continuously decreasing DhMRs revealed 601 concomitantly decreasing genes and 361 non-concomitantly decreasing genes across developmental stages (Figures S2L and S2M). GO analyses showed that the 601 concomitantly decreasing genes were enriched in general developmental processes (Figures 2I and S2O). The embryonic growth factor *FGF8* was

identified among this group of genes and displayed enrichment that was restricted to the EB stage (Figure 2K). Notably, the non-concomitantly decreasing genes were also enriched in neurobiological processes, but were more specifically involved in neuronal synapse processes (Figure S2M). Overall, these data illustrate the synergism between 5hmC and gene expression during organoid development, where the continual regulation of 5hmC appears to strongly affect gene expression and foster proper neurodevelopment.

Recent studies have linked the presence of 5hmC at enhancer regions as a possible mechanism by which 5hmC promotes active gene expression (Wen et al., 2014a). We plotted all the 5hmC read counts from the developmental stages across all the fetal enhancer regions (Gao and Qian, 2020) to investigate their correlation. As the organoids aged, higher 5hmC read counts were observed in fetal enhancer regions (Figure 2L), further supporting a key role of 5hmC in gene regulation. Next, we wanted to consider the histone profile of enhancer regions that harbored either continuously increasing or decreasing 5hmC DhMRs to characterize the 5hmC-histone crosstalk at these enhancer regions. Interestingly, enhancer regions overlapped with DhMRs showing 5hmC accumulation (Figure 2M; 4,222) were poised (enriched with H3K4me1), and may later become active because the accumulation of 5hmC is known to facilitate an accessible chromatin environment (Mahé et al., 2017). On the other hand, enhancer regions overlapped with DhMRs with continual 5hmC depletion (Figure 2N; 751) were active (enriched with H3K27ac) and could become inactivated later during development as 5hmC levels continued decreasing. Collectively, these findings suggest that the level of 5hmC at enhancer regions could affect their ability to distally regulate their target genes.

AD forebrain organoids recapitulate hallmark AD pathologies

We generated forebrain organoids from four fAD iPSC lines (AD-01: *PSEN1* Y155H; AD-02: *PSEN1* M139V; AD-03: *PSEN1* intron 4 deletion; AD-04: *APP* V717I) and three healthy controls (C-03, C-09, C-21). To confirm that our AD organoids recapitulated hallmark pathologies observed in the brain of AD patients, we performed immunofluorescence staining of phosphorylated Tau proteins and Ab aggregates on organoids cultured for 84 days (Figures 3A and 3B). Consistent with previous findings (Raja et al., 2016), we found that both phosphorylated Tau proteins and Ab aggregates were significantly increased in all AD organoid lines compared to controls. Immunoblotting and quantification revealed a 3-fold increase of phosphorylated Tau in AD organoids (Figure 3C). Accumulation of Aβ−40 and Aβ−42 peptides is associated with AD pathogenesis, and our AD organoids showed significant enrichment of both peptides individually as measured by an ELISA (Figure 3D).

To further investigate the genome-wide 5hmC alterations in AD human organoids, three organoid lines carrying fAD risk mutations (AD-01: *PSEN1* Y155H; AD-03: *PSEN1* intron 4 deletion; AD-04: *APP* V717I) and three control (C-03, C-09, C-21) lines were harvested at 84 days for 5hmC-seq profiling (Figure 1A). The 5hmC-seq and RNA-seq data were of high quality, showing high Pearson correlations (>0.8) among replicates (Figure S3). Significantly, both our computational and experimental data revealed a global reduction of 5hmC in AD organoids compared to their controls (Figures 3E–3G), which is consistent

with previous studies using AD models (Raja et al., 2016). Interestingly, despite the observed global 5hmC reduction, the overall distribution of the proportion of 5hmC-enriched peaks at intra- and intergenic regions remained comparable, except at enhancer regions where the proportion of 5hmC-enriched peaks increased from 20.5% in controls to 22.4% in AD organoids (Figure 3H). Collectively, our AD organoids recapitulate human AD pathologies and known epigenomic signatures, validating them as an appropriate model system to study AD.

5hmC is globally altered in AD organoids

Using the same peak-calling and DhMR identification approach described in Figure 2A, overlapping 5hmC regions in 3 independent AD and control lines were first identified. We found a total of 67,466 common peaks across control organoids and 59,632 common peaks in the AD organoids. These overlapping 5hmC regions were then used to identify AD-specific DhMRs, where 9,428 AD-specific enriched DhMRs and 16,362 AD-specific depleted DhMRs were found (Figures 4A, 4B, S4A, and S4B). Given that 5hmC has been linked to gene expression (Pastor et al., 2011), we analyzed the differential gene expression patterns between control and AD organoids using our RNA-seq data (Figure 4C). Numerous neurodevelopmental genes and AD risk genes were identified among these differentially expressed genes (Figure S4B). To ensure these findings were not due to changes in cell composition, we performed cellular deconvolution analysis using MuSiC (Wang et al., 2019). No substantial changes in the estimated cellular proportions were observed (Figure S4C). These findings suggest that in organoids derived from fAD patients, a reduction in 5hmC levels could consequently initiate subtle alterations in the early neuronal gene expression profile.

To further investigate how a reduction of 5hmC in AD organoids could reshape the gene expression profile, we annotated all of the AD-specific enriched or depleted DhMRs (Figures 4A and 4B) to their respective genes. Of the AD-enriched DhMRs, we identified 676 genes that showed an increase in gene expression and enrichments exclusively in neurodevelopmental pathways (Figures 4D and S4D). *GRIN3A*, which encodes an NMDA receptor subunit of a glutamate-gated ion channel (Figure 4F), was found among this group of genes. On the other hand, 463 genes showing decreased expression were enriched in basic developmental processes, although most of the terms were not found to be significant (Figures S4D and S4E). Investigation of the AD-depleted DhMRs revealed 1,172 genes with simultaneous decreasing gene expression (Figure 4E) and 1,379 genes showing an increase in expression (Figures S4G and S4F). The genes in both these groups were largely enriched in basic developmental processes, such as the centromere gene *CENPO* (Figure 4G). Our data suggest that genes displaying increases of 5hmC and expression in AD organoids are enriched in neurodevelopmental processes. On the other hand, irrespective of the impact on gene expression, the regions where 5hmC was lost were predominately occurring in genes that regulate basic developmental processes. Interestingly, we found 17 AD risk genes previously identified from genome-wide association studies (GWASs) that contained at least one DhMR in our AD organoids (Figure S4H). Taken together, these findings are reflective of the sophisticated nature of neurodevelopment and support that AD-specific 5hmC

changes that are occurring early in development may cause subtle disruptions in the neuronal network that could contribute to the onset of AD later in life.

To determine how the loss of 5hmC in AD organoids could impact organoid development, we overlapped our AD-specific DhMRs with our continuously increasing (Figure 4H) and decreasing regions (Figure 4J) previously identified in Figure 2D. The combined AD-specific DhMRs that overlapped with continuously increasing DhMRs (n = 1,083) were also enriched in neurobiological processes (Figure 4I). Similarly, the combined AD-specific DhMRs that overlapped with continuously decreasing DhMRs (n = 588) were enriched in early developmental processes (Figure 4K). Collectively, these data indicate that the dysregulation of 5hmC modifications found in fAD organoids could affect structural brain development as early as fetal development.

Given that crosstalk between different epigenomic mechanisms can impact gene expression, we sought to understand how various histone marks may be affected by aberrant 5hmC levels in AD organoids. The Jaccard index (Jaccard, 1912) was used to quantify the overlap between our AD-specific DhMRs and published fetal brain histone marks (Kundaje et al., 2015; Yan et al., 2016; Figure 4L). Interestingly, the Jaccard index between AD-depleted DhMRs and the active histone mark H3K4me3 showed the highest enrichment, suggesting that 5hmC depletion is more likely to orchestrate the presence of H3K4me3 to co-regulate gene expression. AD-enriched DhMRs appear to moderately overlap with the active enhancer marks H3K27ac and H3K4me1, which may indicate 5hmC alterations in AD organoids could also affect the identity and activity of these enhancer regions. We next performed a similar analysis on the AD-specific DhMRs that was described for Figures 2M and 2N. The common regions between fetal brain enhancers and AD-specific DhMRs appear to be more enriched with the active histone marks H3K27ac and H3K4me3, again suggesting that change of 5hmC in AD organoids could influence active enhancer activities (Figure 4M). These observations indicate that altered 5hmC in our AD organoids could have multifaceted epigenomic roles such as directly modulating transcription, influencing histone marks, and determining enhancer activities and identities.

Several recent studies have revealed strong 5hmC alteration in human postmortem brains of late-onset AD (Table S3). However, whether these alterations have already occurred in early brain development remains unexplored. To that end, we compared the DhMRs identified from our AD organoid model with those from five published postmortem AD brains (Gasparoni et al., 2018; Lardenoije et al., 2019; Li et al., 2019; Qin et al., 2020; Zhao et al., 2017) and found very little DhMR overlap between our early developing AD organoids and those from postmortem AD brain samples. On the other hand, a recent paper profiled DNA cytosine modifications in early- and late-onset AD using cultured, patient-derived iPSCs differentiated into 2D cortical neurons (Fetahu et al., 2019). We found 37–760 times more overlapping DhMRs between our AD-specific DhMRs and the 2D neurons derived from another fAD patient line (*PSEN1* mutation L286V) (Table S3). Many genes associated with these common DhMRs were also associated with key neurodevelopmental processes. Overall, these findings provide conserved 5hmC alterations in two early AD models that could be further explored for mechanistic relevance in AD pathology.

DISCUSSION

5hmC acquisition and depletion in coding and non-coding regions during neurodevelopment

Our analyses support different dynamics for 5hmC acquisition versus 5hmC depletion throughout neuronal development. 5hmC modifications that showed a continuous enrichment throughout development did so in a more gradual manner. The steady accumulation of 5hmC could be important to promote neuronal maturation and thus brain development. On the contrary, 5hmC that showed a continuous depletion trend throughout development declined more rapidly from the pluripotent EB stage to the neuronally differentiated D56 stage, which may be required to commence the transitions between developmental stages. In fact, when we examined the subset of concomitantly increasing genes (Figure 2H), they were strongly represented in genes associated with general neurodevelopmental processes, whereas concomitantly decreasing genes (Figure 2I) were largely enriched in general developmental and proliferative processes. Interestingly, non-concomitantly decreasing genes were strongly represented in highly specified neuronal processes, such as synapse development and function (Figure S2O). From this, one could infer that 5hmC acquisition may have a role in overall neuronal architecture and morphology, whereas 5hmC depletion may be more involved in “fine-tuning” neuronal functions and transitioning away from undifferentiated stages and more toward neuronally committed lineages.

In the context of the central nervous system, our understanding of the relationship between 5hmC and enhancer regions is still premature. Crosstalk of 5hmC with other epigenetic modifications, such as histone marks, can modulate the chromatin architecture and ultimately regulate gene expression (Mahé et al., 2017). We found that 5hmC was more strongly associated with enhancer regions in the developing organoids than it was in the undifferentiated EBs. When we specifically looked at DhMRs that continued to increase or decrease in 5hmC, these regions associated with poised or active enhancer regions, respectively, implying a role for 5hmC in promoting cell differentiation through its relationship with enhancers to upregulate cell-type-specific differentiating genes (Sérandour et al., 2012). Collectively, both pieces of data support the underappreciated importance of 5hmC regulation in non-coding and enhancer regions to induce neuronal-specific gene regulation during early fetal brain development.

Forebrain organoid model of AD and the impact of 5hmC global alterations

In forebrain organoids derived from fAD patients carrying various *PSEN1* mutations or an *APP* mutation, we demonstrated at the cellular level that our AD organoids recapitulate hallmarks of human AD pathology, despite being reminiscent of fetal brain stages between 12 and 24 weeks post-conception (Qian et al., 2016). PSC-derived models better recapitulate the structure and function of fetal tissues compared to their adult tissues (Takebe and Wells, 2019), making them the earliest human model to study early neurodevelopment. These findings validate the use of AD forebrain organoids as a promising AD model.

In our study, we found a significant global reduction of 5hmC in the AD organoids compared to controls. Although we identified nearly 2 times fewer AD-enriched DhMRs compared to AD-depleted DhMRs, ectopic enrichment of 5hmC associated with increased gene expression specifically affected neurodevelopmental processes. In support of this, we identified numerous critical neurodevelopmental genes and AD risk genes that were primarily upregulated in AD organoids. Regions that became ectopically depleted for 5hmC were involved in regulating basic developmental processes irrespective of how their expression patterns changed. Noticeably, our RNA-seq data revealed that 56.7% of the dysregulated genes were inappropriately upregulated, despite the significant global 5hmC reduction we observed in the AD organoids. One possibility could be the increased proportion of 5hmC being distributed to enhancer regions that was observed only in AD organoids. Interestingly, we also found that enhancer regions overlapped with our AD-specific DhMRs were enriched for both active and poised enhancer histone marks. This speaks to a mechanism whereby these subtle 5hmC alterations during early brain development might not result in structural damage but could affect the delicate neuronal networks making the AD-predisposed brain more vulnerable to AD pathogenesis. These findings collectively support that aberrant 5hmC dynamics disrupt the timing of neurodevelopment in the fetal brain carrying fAD risk mutations.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resource/reagents should be directed to and will be fulfilled by the Lead Contact, Bing Yao (bing.yao@emory.edu)

Materials availability—This study did not generate new unique reagents.

Data and code availability—The accession number for the 5hmC-seq and RNA-seq data generated in this paper is GSE151818.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The control iPSC lines were generously provided by Dr. Gary Bassell's laboratory from Emory University. The fAD fibroblasts were generously provided by Dr. Selina Wray, from UCL Queen Square Institute of Neurology and then generated into fAD iPSCs by Dr. Chadwick Hales, from Emory University. Fibroblasts and iPS cells were collected and stored under Emory IRB approved protocol IRB00064365. iPSCs were generated using the Invitrogen CytoTune kits with Sendai virus and Yamanaka factors. Samples were maintained in standard growth conditions or cryopreserved in liquid nitrogen for longer term storage.

Human forebrain-specific organoid cultures—All fAD patient-derived iPSC lines (AD-01: *PSEN1* Y155H (male); AD-02: *PSEN1* M139V (female); AD-03: *PSEN1* intron4 deletion (female); AD-04: *APP* V717I (male)) and the three healthy control (C-03 (male), C-09 (male), C-21 (female)) iPSC lines (provided by Drs. Chadwick Hales' and Gary Bassell's laboratory at Emory University) were cultured on irradiated MEFs in human iPSC medium consisting of D-MEM/F12 (GIBCO #11330–032), 20% Knockout Serum

Replacement (KSR, GIBCO #10828028), 1X Glutamax (GIBCO #35050061), 1X MEM Non-essential Amino Acids (NEAA, GIBCO #11140050), 100 μM β- Mercaptoenthanol (GIBCO #21985023), and 10 ng/ml human basic FGF (bFGF, PeproTech #100–18B) as described (Wen et al., 2014b). Forebrain-specific organoids were generated as previously described (Qian et al., 2016). Briefly, human iPSC colonies were detached from the feeder layer with 1 mg/ml collagenase treatment (Termo Fisher Scientific #17104019) for 1 hour and suspended in embryonic body (EB) medium, consisting of FGF-2-free iPSC medium supplemented with 2 μM Dorsomorphin (Tocris #3093) and 2 μM A-83 (Tocris #692) in non-treated polystyrene plates for 4 days with a daily medium change. On days 5–6, half of the medium was replaced with induction medium consisting of DMEM/F12, 1X N2 Supplement (GIBCO #17502048), 10 μg/ml Heparin (Sigma), 1X Penicillin/Streptomycin (10,000 U/mL, GIBCO #15140122), 1X Non-essential Amino Acids, 1X Glutamax, 4 ng/ml WNT-3A (R&D Systems), 1 μM CHIR99021 (Tocris #4423), and 1 μM SB-431542 (Tocris #1614). On day 7, organoids were embedded in Matrigel (Corning #354277) and continued to grow in induction medium for 6 more days. On day 14, embedded organoids were mechanically dissociated from Matrigel by pipetting up and down onto the plate with a 5ml pipette tip. Typically, 10 – 20 organoids were transferred to each well of a 12-well spinning bioreactor (SpinΩ) containing differentiation medium, consisting of DMEM/F12, 1X N2 and B27 Supplements (Thermo Fisher Scientific #17504044), 1X Penicillin/Streptomycin, 100 μM β- Mercaptoenthanol (Invitrogen), 1X MEM NEAA, 2.5 μg/ml Insulin (Sigma #I0516). At day 71, differentiation medium was exchanged with maturation medium, consisting of Neurobasal (GIBCO #21103049), 1X B27 Supplement, 1X Penicillin/Streptomycin, 1X β-Mercaptoenthanol, 0.2 mM Ascorbic Acid (Sigma #1043003), 20 ng/ml BDNF (Peprotech #450–02), 20 ng/ml GDNF (Peprotech #450–10), 1 ng/ml TFGβ (Peprotech), and 0.5 mM cAMP (StemCell Technologies #73884). All media were changed every other day.

METHOD DETAILS

Organoid immunocytochemistry—Forebrain organoids were processed for immunocytochemistry as previously described (Qian et al., 2016). Briefly, whole organoids were fixed in 4% Paraformaldehyde (Polysciences #18814–10) in Phosphate Buffered Saline (BPS) for 30–60 min at room temperature. Organoids were washed 3 times with PBS and then incubated in 30% sucrose (Sigma #S5016) solution overnight. Organoids were embedded in tissue freezing medium (General Data #TFM-5) and sectioned with a cryostat (Leica). For immunostaining, freezing medium was washed with PBS before permeabilization with 0.2% Triton-X in PBS for 1 hr. Tissues were then blocked with blocking medium consisting of 10% donkey serum (Millipore #S30) in PBS with 0.1% Tween-20 (PBST) for 30 min. Primary antibodies diluted in blocking solution were applied to the sections overnight at 4°C. After washing with PBST, secondary antibodies diluted in blocking solution were applied to the sections for 1hr at room temperature. Finally, sections were washed with PBST and stained with DAPI. All images were captured by Nikon Eclipse Ti-E microscope. Quantitative analyses were conducted on randomly picked cortical structures in a blind fashion using ImageJ software (Schneider et al., 2012).

DNA and RNA isolation—Embryoid bodies and organoids were collected after 56, 84 and 112 days of culture and were immediately frozen on dry ice and stored at −80°C. Tissue

was first homogenized in a lysis buffer (10mM Tris pH 8.0, 5mM EDTA, 200mM NaCl, 0.2% SDS) with 25 μL proteinase K (20mg/ml) using a hand-held pestle homogenizer then was incubated at 55°C overnight. After the overnight digestion, the lysates were brought to room temperature and incubated with 5 μL of RNase A solution (20mg/ml) for at least 2 hours at room temperature. DNA was extracted by adding equal volume of buffered phenol:chloroform:isoamyl alcohol (25:24:1 ratio) and centrifuged in Phase-Lock tubes at 15,000 RPM at room temperature. Supernatant was transferred to clean tubes. An equal volume of isopropanol was then added to the supernatant and mixed well at room temperature to precipitate the DNA. The DNA was then centrifuged at 10,000 g for 10 minutes at room temperature, and then washed in 70% ethanol. After all ethanol was removed, the DNA pellet was eluted in nuclease-free water and incubated at 55°C for 1 hour before storing at −20°C. The DNA was quantified by Nanodrop and Qubit, and quality confirmed by a gel.

A separate aliquot of tissue was used for RNA isolation. Tissue was homogenized in TRIzol using a hand-held pestle homogenizer and incubated in TRIzol for at least 5 minutes. Chloroform (1:5 ratio) was added, the tubes shaken, and incubated at room temperature for 15 minutes. Samples were then centrifuged at 12,000 g for 15 minutes at 4°C. The top aqueous layer was transferred to a clean tube, and the RNA was precipitated in 3M NaAc pH 5.2 (10:1 ratio), 4 μL of glycogen (5mg/ml), 100% isopropanol (1:1 ratio) overnight at −80°C. The next day, the samples were centrifuged at 15,000 RPM for 20 minutes at 4°C. The resulting RNA pellet was washed once in 75% ethanol, centrifuged at 7,500 g for 10 minutes at 4°C. The washed RNA pellet was dissolved in nuclease-free water. RNA was quantified by Nanodrop and quality confirmed by a gel.

Dot blot—Genomic DNA was blotted onto a Hybond nylon membrane (Amersham, GE Healthcare) using a Bio-Rad Dot Blot apparatus (#1706545, Bio-Rad) and washed three times with 6X saline-sodium citrate buffer with 15 minutes of vacuum. The DNA was hybridized to the membrane at 85°C for 30 minutes. Immunoblotting was performed by first blocking the membrane in 5% milk/0.2% tris-buffered saline with Tween-20 for 30 minutes then incubating overnight in primary antibody in 5% milk (5hmC antibody 1:2000, Active Motif #39769) at 4°C with rotation. Secondary antibody incubation was anti-rabbit HRP-linked IgG (Cell Signaling #7074S) 1:5000 in 5% milk for 1 hour. Signal was detected with ECL substrate (Denville Scientific HyGLO #E2400) and imaged with autoradiography film and a Konica Minolta film processor (SRX-101A). Films were scanned into digital form, then pixel densitometry quantification was performed using ImageJ software (Schneider et al., 2012).

Western blot analysis—Human iPSC-derived forebrain organoids were lysed in RIPA buffer (150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS; 50 mM Tris, pH 8.0) containing Complete Protease Inhibitor Cocktail (Roche). Samples were left on ice for 30 min and sonicated briefly. The insoluble fraction was removed by centrifugation at 15,000 rpm for 15 min at 4°C. Protein concentration was determined by BCA protein assay kit (Bio-Rad). 2X SDS sample buffer (Bio-Rad) containing 5% β-mercaptoethanol (Sigma) was added to equal amounts of protein. Proteins were then separated by 4%–15% SDS-

PAGE (Bio-Rad) and transferred to PVDF (0.2 μm) or nitrocellulose membrane (0.45 μm). 5% dried milk in TBST (Tris buffered saline with 0.1% Tween 20) was incubated for blocking, and membranes were applied with specific antibodies. After washing with TBST and incubation with horseradish peroxidase-conjugated anti-rabbit or anti-mouse IgG (Santa Cruz Biotechnology), the antigen-antibody was detected by chemiluminescence (ECL; Pierce) and X-ray film (GE Healthcare).

ELISA analysis—Aβ concentration was measured from forebrain organoid supernatants using commercial ELISA kit for Aβ (1–40) and Aβ (1–42) (Thermo Fisher Scientific #KHB3481 and #KHB3544 respectively) following the manufacturer's instruction. Briefly, media samples were incubated (4 hours) in primary antibodies against the COOH-terminus of the 1–40 or 1–42 Aβ sequence in pre-coated 96 well plates (pre-coated with monoclonal antibody specific to human Aβ 1–40 or 1–42) followed by aspiration and four washes (in washing buffer). And then the samples were subsequently incubated with HRP-conjugated secondary antibodies followed by aspiration, four washes, and addition of HRP substrate (3,3′,5,5′-tetramethylbenzidine). The reaction was stopped using 1 N sulfuric acid and absorption was measured at 450 nm in a Synergy HT microplate reader (BioTek). Absolute values were calculated from a standard curve and plotted as either picogram/ml (pg/ml) or Aβ42/40 ratio per sample.

5hmC capture—Five μg of genomic DNA was sonicated to 300–400 base pair (bp) using a Covaris focused ultrasonicator. 5hmC capture was performed according to the method described in Song et al. (2011). First, a glucosyltransfer reaction was performed using 2.5μl of T4 phage β-glucosyltransferase enzyme (10,000U/ml; New England BioLabs #M0357L) and 100μM UDP-6-N3-glucose (Jena Biosciences #CLK-076) and incubated at 37°C for 2 hours. After purification with AMPure XP beads (Beckman Coulter A63881), the glucosylated 5hmC-containing DNA fragments were biotinylated with 150μM disulfide biotin linker (Click Chemistry Tools A112–5) at 37°C for 2 hours. After purification with AMPure XP beads, the biotinylated 5hmc-containing DNA fragments were pulled down using Dynabeads MyOne Streptavidin C1 beads (ThermoFisher Scientific #65002) and were washed three times with Binding/Washing buffer (20mM Tris pH 7.5, 1mM EDTA, 2M NaCl, 0.01% Tween-20). The 5hmC-containing DNA fragments were then removed from the beads with fresh 100mM dithiothreitol for 2 hours with rotation at room temperature. After final purification with AMPureXP beads, the 5hmC-enriched DNA fragments were eluted in nuclease-free water and quantified by Qubit.

Library preparation and high-throughput sequencing—Enriched DNA from 5hmC capture were subjected to library construction using the NEBNext Ultra II DNA Library Prep kit for Illumina (New England BioLabs) according to manufacturer's protocol. Briefly, 1–7ng of 5hmC-enriched DNA or un-enriched genomic DNA was utilized for each library construction. An Agilent 2100 Bioanalyzer was used to confirm purity and fragmentation size of the final libraries. RNA-seq libraries were generated using TruSeq RNA Sample Prep V2 kit (New England BioLabs) with 1 μg of RNA to first obtain poly-A-enriched RNA and synthesize cDNA, and then the same library construction protocol was followed using 5ng

of cDNA. Libraries were sequenced pair-end (150bp) on an Illumina HiSeq platform by Admera Health, LLC.

QUANTIFICATION AND STATISTICAL ANALYSIS

All details of the data processing, statistical analyses, and Bioinformatics software used in this paper are provided in Method details. All statistical analyses procedures are described in the corresponding figure legends. One-Way ANOVA was used when making comparisons between more than two groups in Figure S1H. Two group difference tests in Figures 3C, 3D, and 3G were performed using unpaired t test. All statistical tests were done using GraphPad 9.02 (Prism). Significance levels were set to $\alpha = 0.05$ with *p value < 0.05, ** p value < 0.01, *** p value < 0.001.

RNA-seq cellular deconvolution—Cellular deconvolution was conducted on the organoid RNA-seq data with MuSiC (Wang et al., 2019), a method that utilizes cell-type-specific single-cell RNA sequencing (scRNA-seq) data as a reference panel to quantify cell type compositions of samples with bulk RNA-seq data. We obtained the scRNA-seq data of postmortem human brain from the ROSMAP single cell study (Mathys et al., 2019) as reference in the deconvolution. Following the software recommendations, we provided the raw read counts of both bulk RNA sequencing and the single cell RNA sequencing data as inputs. We applied MuSiC with all default settings and obtained cell type proportions in one step without clustering them into higher level groups. The output proportions "Est.prop.weighted" were used as the final proportion estimations.

5hmC-seq and RNA-seq data processing—Raw 150 bp paired-end ChIP–seq reads were mapped to the hg19 reference genome using bowtie2 (v2.2.6) (Langmead et al., 2009) with the parameters "–no-discordant–no-mixed" to prevent discordant alignments and alignments for the individual mates. Flags "-F 4" and "-q 10" were used in samtools (v 1.9) (Li et al., 2009) to exclude unmapped reads and reads with low Mapping Quality (MAPQ) values less than 10 for each replicate in all stages. Technical replicates with multiple sequencings were combined using the "merge" function in samtools. The genome was segmented into 1-kb consecutive bins and the reads were normalized with respect to the sample with the smallest total numbers of read counts. Bins with less than 10 counts summed over all samples were removed and the correlation of the normalized reads intensity was calculated between biological replicates on remaining bins across all stages. Principal component analysis (PCA) was performed to visualize replicate clusters using the binned normalized read counts. All biological replicates with Pearson correlation > 0.8 per time point were included in our analysis. The 5hmC levels between control and AD organoids was determined by segmenting the genome into 1-kb bins, where only those bins exceeding 100 read counts were considered as high confident 5hmC regions when determining significance. The MACS2 algorithm (v 2.1.0.) (Zhang et al., 2008) was used to call 5hmC enriched peaks for each organoid sample by comparing to the merged input organoid sample. By default, MACS2 normalizes all enriched peaks. Biological replicate peak files were combined using the "intersect" function in bedtools (v 2.27.0) (Quinlan and Hall, 2010) and only 5hmC peaks that were identified in all replicates were considered for our analyses. Based on a previous publication (Wang et al., 2018), the default parameters for the

bedtools intersect function report a 1bp minimum overlap between regions. HOMER (Hypergeometric Optimization of Motif EnRichment) (v4.9) (Heinz et al., 2010) was used to annotate identified peak regions to their corresponding nearby genes, with an additional flag "-annStats" to also annotate genomic features to these peak regions. 5hmC enrichment was estimated by calculating the ratio of observed verse expected probability for 5hmC peaks annotated to the specific genomic feature. The observed probability was the length of the 5hmc peaks that covers the related genomic regions versus the length of the total 5hmC peaks, and the expected probability was the length of the total regions of the specific genomic feature divided by the whole genome length. Further, EB peaks were overlapped with peak regions in other stages by using the "intersect" function in bedtools (v2.27.0).

Raw RNA-seq reads were aligned to the human genome build hg19 using TopHat2 (v1.3.3) (Trapnell et al., 2012). Cufflinks (v2.2.1) (Trapnell et al., 2012) was used to generate fragments per kilobase million (FPKM) of all RefSeq genes and averaged per condition for downstream analyses. RNA-seq read count correlations were determined among all organoid samples and PCA was used to cluster organoid samples with FPKM data. All replicates per stage showed high correlation (Pearson's correlation > 0.85). Using these aligned RNA-seq data, differential gene expression analysis was conducted using Cuffdiff (v2.2.1) (Trapnell et al., 2012) with respect to control and AD organoids.

Identification of differentially hydroxymethylated regions (DhMRs)— Developmental stage specific DhMRs were determined by established or disappeared 5hmC peaks in the current stage compared to the previous stage as previously described (Wang et al., 2018). These peaks were identified by using the "windowbed" function in bedtools (v2.29.2) by extended 300bp on either side of the 5hmC peak regions. Peak regions with no surrounding peaks within 300bp in the previous stage were defined as established DhMRs and peak regions in the previous stage with no surrounding peaks within 300bp in current stage were defined as disappeared DhMRs in the current stage. Developmental stage-specific DhMRs were identified by grouping the regions based on the stage with either the highest or lowest normalized read count. To identify AD organoid-specific DhMRs, 5hmC peaks from three replicate peak files for each AD organoid line were first overlapped using the "intersect" function in bedtools (v 2.27.0) to obtain common 5hmC peak regions per line. Then, using the same "windowbed" function described above, DhMRs per AD line were identified by comparing to healthy control (C-03). Three sets of DhMRs from each line were finally overlapped, again using the "windowbed" function, to obtain shared AD-specific DhMRs to avoid person to person variations.

Bioinformatics analysis—To identify stage-specific regions, 5hmC read counts were enumerated in all DhMRs per replicate and normalized to the sample with smallest total reads in the whole genome. The averages of the read counts of all replicates per stage were taken and compared across stages. DhMRs with inconsistent established or disappeared patterns were removed, for example, D56 established DhMRs with more averaged reads in EB stage than in D56 stage were excluded. Stage specific upregulated DhMRs were defined as DhMRs with the highest read counts in the current stage and downregulated DhMRs were defined as DhMRs with the lowest read counts in the current stage. Heatmaps and metaplots

were generated by Ngsplot tool (Shen et al., 2014) to validate the enrichment patterns of 5hmc within these DhMRs. To enhance the visibility of the heatmaps, the read counts were normalized by subtracting the median count per region and divided by the median absolute deviation per region. The corresponding genes of these stage specific DhMRs were annotated by HOMER (v4.9) using default settings, and later used for Gene ontology (GO) analysis (Ashburner et al., 2000; The Gene Ontology Consortium, 2019) or Genomic Regions Enrichment of Annotations Tool (GREAT) analysis (McLean et al., 2010). To identify DhMRs that showed a continual increase or decrease in 5hmC read counts across developmental stages, averaged normalized 5hmC read counts were counted and grouped based on the stage with either the highest or lowest number or read counts. These DhMRs were then annotated to their corresponding genes by HOMER (v4.9), and those located in intergenic regions were removed from further analysis. Only genes also showing a continual increase or decrease in expression were used for gene ontology. Genes showing a continuous increase in 5hmC and a continuous increase in gene expression will be referred to as concomitantly increasing genes, whereas genes continuously increasing in 5hmC but decreasing in gene expression will be referred to as non-concomitantly increasing genes. Genes showing a continuous decrease in both 5hmC and gene expression will be referred to as concomitantly decreasing genes, whereas genes continuously decreasing in 5hmC, but increasing in gene expression will be referred to as non-concomitantly decreasing genes. GO analysis were performed to identify functional patterns enriched with these genes. To investigate possible links between 5hmC and enhancer regions, published enhancer regions from human fetal brains and astrocytes available on the online database EnhancerAtlas (http://www.enhanceratlas.org/downloadv2.php) (Gao and Qian, 2020) were overlapped with DhMRs that showed either a continual increase or decrease in 5hmC. EnhancerAtlas was also used to predict and annotate genes that may be regulated by these enhancers. Furthermore, raw sequence data from human fetal brain histone modifications (H3K4me3, H3K36me3 and H3K9ac, H3K9me3 and H3K27me3, H3K4me1 and H3K27ac) from the Epigenome Roadmap Project (http://www.roadmapepigenomics.org/) (Kundaje et al., 2015) were also obtained, mapped and overlapped with the above enhancer regions, where the average normalized read counts of these histone marks per stage were plotted. These histone marks were also used to assess the Jaccard similarity coefficient (Jaccard, 1912) between 5hmC peaks and histone peaks at enhancer regions that overlapped with AD-specific DhMRs. The Jaccard index quantifies the proportion of these overlapping regions. To compare AD organoid DhMRs to published human postmortem AD brains or patient derived iPSCs differentiated into 2D cortical neurons, DhMRs were identified using the same analysis for identifying the organoid DhMRs described above. Overlapping DhMRs were annotated to their corresponding genes and used for gene ontology.

Gene ontology (GO) analysis—Functional annotation analysis was conducted in the GO Consortium classification system (http://geneontology.org) (Ashburner et al., 2000; The Gene Ontology Consortium, 2019). We then clustered GO terms to a representative functional term and plotted the most significant (−log10(FDR)) to show their statistical significance. These results were confirmed using GREAT (McLean et al., 2010) which assigns biological processes directly from our identified DhMRs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the following funding sources: NIH grants R01AG062577, R01MH117122, and R01AG064786 (to B.Y.), GM008492 training grant (to J.N.K.), R35GM138313 (to J.Y.), and R01AG065611, R01NS107505, R01MH121102, P50AG025688, and R21MH123711, Department of Defense (DoD) grant W81XWH1910353, and an Edward Mallinckrodt, Jr. Foundation grant (to Z.W.).

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al.; The Gene Ontology Consortium (2000). Gene Ontology: tool for the unification of biology. Nat. Genet 25, 25–29. [PubMed: 10802651]

Bergsland M, Werme M, Malewicz M, Perlmann T, and Muhr J (2006). The establishment of neuronal properties is controlled by Sox4 and Sox11. Genes Dev 20, 3475–3486. [PubMed: 17182872]

Bernstein AI, Lin Y, Street RC, Lin L, Dai Q, Yu L, Bao H, Gearing M, Lah JJ, Nelson PT, et al. (2016). 5-hydroxymethylation-associated epigenetic modifiers of Alzheimer's disease modulate Tau-induced neurotoxicity. Hum. Mol. Genet 25, 2437–2450. [PubMed: 27060332]

Bradley-Whitman MA, and Lovell MA (2013). Epigenetic changes in the progression of Alzheimer's disease. Mech. Ageing Dev 134, 486–495. [PubMed: 24012631]

Busche MA, and Konnerth A (2016). Impairments of neural circuit function in Alzheimer's disease. Philos. Trans. R. Soc. Lond. B Biol. Sci 371, 20150429. [PubMed: 27377723]

Condliffe D, Wong A, Troakes C, Proitsi P, Patel Y, Chouliaras L, Fernandes C, Cooper J, Lovestone S, Schalkwyk L, et al. (2014). Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain. Neurobiol. Aging 35, 1850–1854. [PubMed: 24679604]

Fetahu IS, Ma D, Rabidou K, Argueta C, Smith M, Liu H, Wu F, and Shi YG (2019). Epigenetic signatures of methylated DNA cytosine in Alzheimer's disease. Sci. Adv 5, eaaw2880. [PubMed: 31489368]

Gao T, and Qian J (2020). EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. Nucleic Acids Res. 48 (D1), D58–D64. [PubMed: 31740966]

Gasparoni G, Bultmann S, Lutsik P, Kraus TFJ, Sordon S, Vlcek J, Dietinger V, Steinmaurer M, Haider M, Mulholland CB, et al. (2018). DNA methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex. Epigenetics Chromatin 11, 41. [PubMed: 30045751]

The Gene Ontology Consortium (2019). The Gene Ontology resource: 20 years and still GOing strong. Nucleic Acids Res 47 (D1), D330–D338. [PubMed: 30395331]

Hardy JA, and Higgins GA (1992). Alzheimer's disease: the amyloid cascade hypothesis. Science 256, 184–185. [PubMed: 1566067]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589. [PubMed: 20513432]

Itskovitz-Eldor J, Schuldiner M, Karsenti D, Eden A, Yanuka O, Amit M, Soreq H, and Benvenisty N (2000). Differentiation of human embryonic stem cells into embryoid bodies comprising the three embryonic germ layers. Mol. Med 6, 88–95. [PubMed: 10859025]

Jaccard P (1912). The distribution of the flora in the alpine zone. New Phytol. 11, 37–50.

James BD, Leurgans SE, Hebert LE, Scherr PA, Yaffe K, and Bennett DA (2014). Contribution of Alzheimer disease to mortality in the United States. Neurology 82, 1045–1050. [PubMed: 24598707]

Kelava I, and Lancaster MA (2016). Stem cell models of human brain development. Cell Stem Cell 18, 736–748. [PubMed: 27257762]

Kriaucionis S, and Heintz N (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324, 929–930. [PubMed: 19372393]

Kuehner JN, Bruggeman EC, Wen Z, and Yao B (2019). Epigenetic regulations in neuropsychiatric disorders. Front. Genet 10, 268. [PubMed: 31019524]

Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, et al.; Roadmap Epigenomics Consortium (2015). Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330. [PubMed: 25693563]

Langmead B, Trapnell C, Pop M, and Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. [PubMed: 19261174]

Lardenoije R, Roubroeks JAY, Pishva E, Leber M, Wagner H, Iatrou A, Smith AR, Smith RG, Eijssen LMT, Kleineidam L, et al. (2019). Alzheimer's disease-associated (hydroxy)methylomic changes in the brain and blood. Clin. Epigenetics 11, 164. [PubMed: 31775875]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, and Durbin R; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. [PubMed: 19505943]

Li X, Yao B, Chen L, Kang Y, Li Y, Cheng Y, Li L, Lin L, Wang Z, Wang M, et al. (2017). Ten-eleven translocation 2 interacts with forkhead box O3 and regulates adult neurogenesis. Nat. Commun 8, 15903. [PubMed: 28660881]

Li P, Marshall L, Oh G, Jakubowski JL, Groot D, He Y, Wang T, Petronis A, and Labrie V (2019). Epigenetic dysregulation of enhancers in neurons is associated with Alzheimer's disease pathology and cognitive symptoms. Nat. Commun 10, 2246. [PubMed: 31113950]

Mahé EA, Madigou T, Sérandour AA, Bizot M, Avner S, Chalmel F, Palierne G, Métivier R, and Salbert G (2017). Cytosine modifications modulate the chromatin architecture of transcriptional enhancers. Genome Res 27, 947–958. [PubMed: 28396520]

Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, Menon M, He L, Abdurrob F, Jiang X, et al. (2019). Single-cell transcriptomic analysis of Alzheimer's disease. Nature 570, 332–337. [PubMed: 31042697]

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, and Bejerano G (2010). GREAT improves functional interpretation of *cis*-regulatory regions. Nat. Biotechnol 28, 495–501. [PubMed: 20436461]

Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P, et al. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473, 394–397. [PubMed: 21552279]

Qian X, Nguyen HN, Song MM, Hadiono C, Ogden SC, Hammack C, Yao B, Hamersky GR, Jacob F, Zhong C, et al. (2016). Brain-region-specific organoids using mini-bioreactors for modeling ZIKV exposure. Cell 165, 1238–1254. [PubMed: 27118425]

Qin L, Xu Q, Li Z, Chen L, Li Y, Yang N, Liu Z, Guo J, Shen L, Allen EG, et al. (2020). Ethnicity-specific and overlapping alterations of brain hydroxymethylome in Alzheimer's disease. Hum. Mol. Genet 29, 149–158. [PubMed: 31814020]

Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. [PubMed: 20110278]

Raja WK, Mungenast AE, Lin YT, Ko T, Abdurrob F, Seo J, and Tsai LH (2016). Self-organizing 3D human neural tissue derived from induced pluripotent stem cells recapitulate Alzheimer's disease phenotypes. PLoS ONE 11, e0161969. [PubMed: 27622770]

Schneider CA, Rasband WS, and Eliceiri KW (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675. [PubMed: 22930834]

Sérandour AA, Avner S, Oger F, Bizot M, Percevault F, Lucchetti-Miganeh C, Palierne G, Gheeraert C, Barloy-Hubler F, Péron CL, et al. (2012). Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. Nucleic Acids Res 40, 8255–8265. [PubMed: 22730288]

Shen L, Shao N, Liu X, and Nestler E (2014). ngs.plot: quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics 15, 284. [PubMed: 24735413]

Shu L, Sun W, Li L, Xu Z, Lin L, Xie P, Shen H, Huang L, Xu Q, Jin P, and Li X (2016). Genome-wide alteration of 5-hydroxymenthylcytosine in a mouse model of Alzheimer's disease. BMC Genomics 17, 381. [PubMed: 27207465]

Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen CH, Zhang W, Jian X, et al. (2011). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat. Biotechnol 29, 68–72. [PubMed: 21151123]

Spiers H, Hannon E, Schalkwyk LC, Bray NJ, and Mill J (2017). 5-hydroxymethylcytosine is highly dynamic across human fetal brain development. BMC Genomics 18, 738. [PubMed: 28923016]

Sun B, Halabisky B, Zhou Y, Palop JJ, Yu G, Mucke L, and Gan L (2009). Imbalance between GABAergic and glutamatergic transmission impairs adult neurogenesis in an animal model of Alzheimer's disease. Cell Stem Cell 5, 624–633. [PubMed: 19951690]

Szulwach KE, Li X, Li Y, Song CX, Wu H, Dai Q, Irier H, Upadhyay AK, Gearing M, Levey AI, et al. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat. Neurosci 14, 1607–1616. [PubMed: 22037496]

Takebe T, and Wells JM (2019). Organoids by design. Science 364, 956–959. [PubMed: 31171692]

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, and Pachter L (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc 7, 562–578. [PubMed: 22383036]

Wang C, Liu X, Gao Y, Yang L, Li C, Liu W, Chen C, Kou X, Zhao Y, Chen J, et al. (2018). Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. Nat. Cell Biol 20, 620–631. [PubMed: 29686265]

Wang X, Park J, Susztak K, Zhang NR, and Li M (2019). Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. Nat. Commun 10, 380. [PubMed: 30670690]

Wen L, Li X, Yan L, Tan Y, Li R, Zhao Y, Wang Y, Xie J, Zhang Y, Song C, et al. (2014a). Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. Genome Biol. 15, R49. [PubMed: 24594098]

Wen Z, Nguyen HN, Guo Z, Lalli MA, Wang X, Su Y, Kim NS, Yoon KJ, Shin J, Zhang C, et al. (2014b). Synaptic dysregulation in a human iPS cell model of mental disorders. Nature 515, 414–418. [PubMed: 25132547]

Yan L, Guo H, Hu B, Li R, Yong J, Zhao Y, Zhi X, Fan X, Guo F, Wang X, et al. (2016). Epigenomic landscape of human fetal brain, heart, and liver. J. Biol. Chem 291, 4386–4398. [PubMed: 26719341]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. 9, R137. [PubMed: 18798982]

Zhang Y, Zhang Z, Li L, Xu K, Ma Z, Chow HM, Herrup K, and Li J (2020). Selective loss of 5hmC promotes neurodegeneration in the mouse model of Alzheimer's disease. FASEB J. 34, 16364–16382. [PubMed: 33058355]

Zhao J, Zhu Y, Yang J, Li L, Wu H, De Jager PL, Jin P, and Bennett DA (2017). A genome-wide profiling of brain DNA hydroxymethylation in Alzheimer's disease. Alzheimers Dement. 13, 674–688. [PubMed: 28089213]

Highlights

- 5hmC is dynamically regulated during forebrain organoid development
- AD organoids recapitulate cellular and molecular pathologies seen in patient brains
- Aberrant 5hmC in AD organoids could subtly disrupt early neuronal networks
- 5hmC crosstalk with histone marks could affect enhancer activity in AD organoids

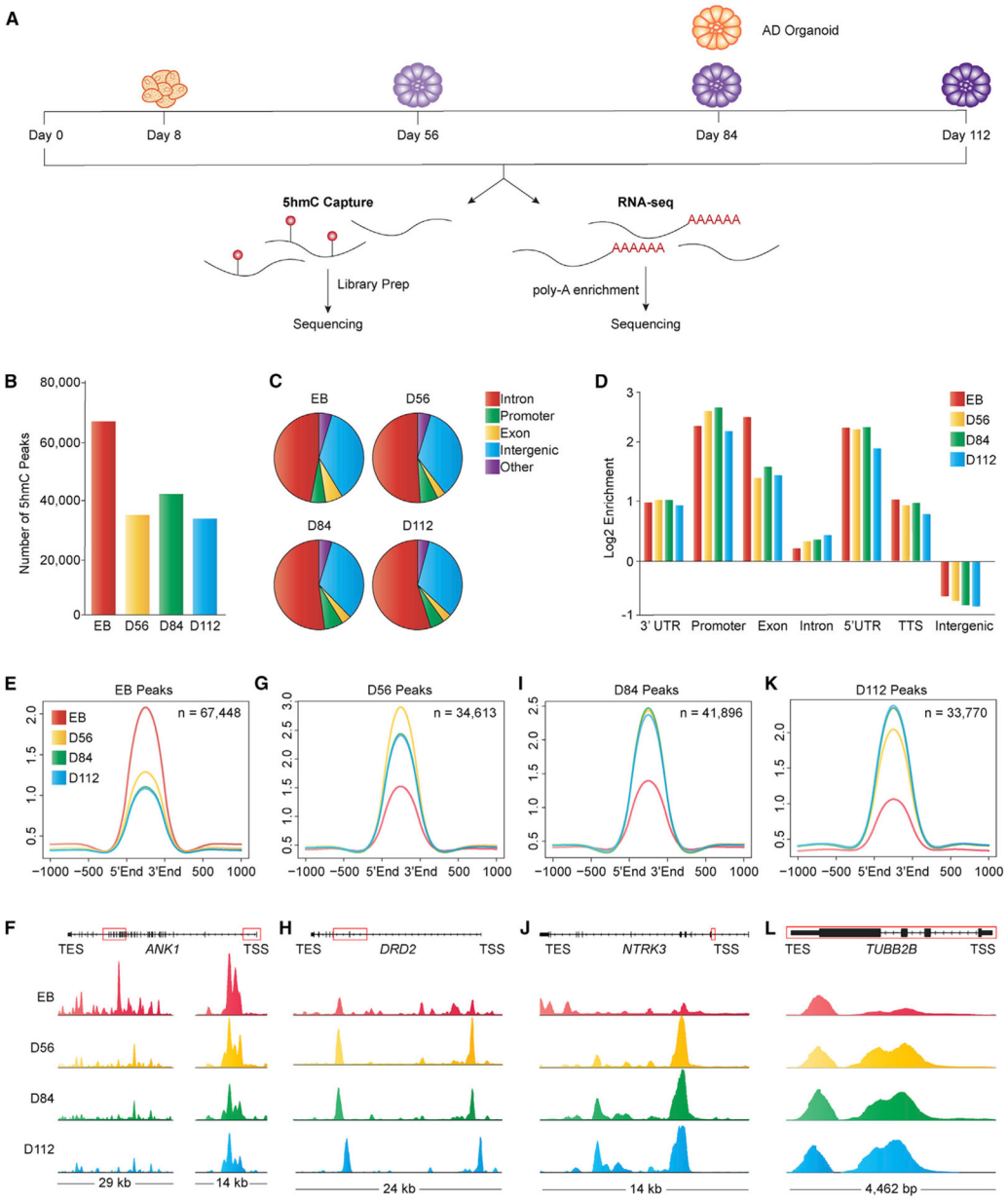


Figure 1. Genome-wide profiling of 5hmC in forebrain organoids during development

(A) Schematic of the collection time points of forebrain organoids derived from controls and patients with Alzheimer's disease (AD) for genome-wide 5hmC and RNA sequencing: day 8 embryoid bodies (EBs), day 56 (D56), day 84 (D84), day 112 (D112), and AD organoid at day 84.

(B) Number of 5hmC peaks identified across developmental stages.

(C) Distribution of 5hmC peaks across genomic features in the human genome.

(D) Enrichment of 5hmC peaks at 3′ and 5′ untranslated regions (3′ UTR and 5′ UTR), promoters, exons, introns, transcription termination sites (TTSs), and intergenic regions.

(E, G, I, and K) Average normalized 5hmC read counts across 5hmC peaks for EBs (E), D56 (G), D84 (I), and D112 (K).

(F, H, J, and L) Normalized 5hmC counts at peak regions identified in *ANK1* (F), *DRD2* (H), *NTRK3* (J), and *TUBB2B* (L) in forebrain organoids across developmental stages. Red boxes indicate where on the gene the displayed peak region(s) originated.

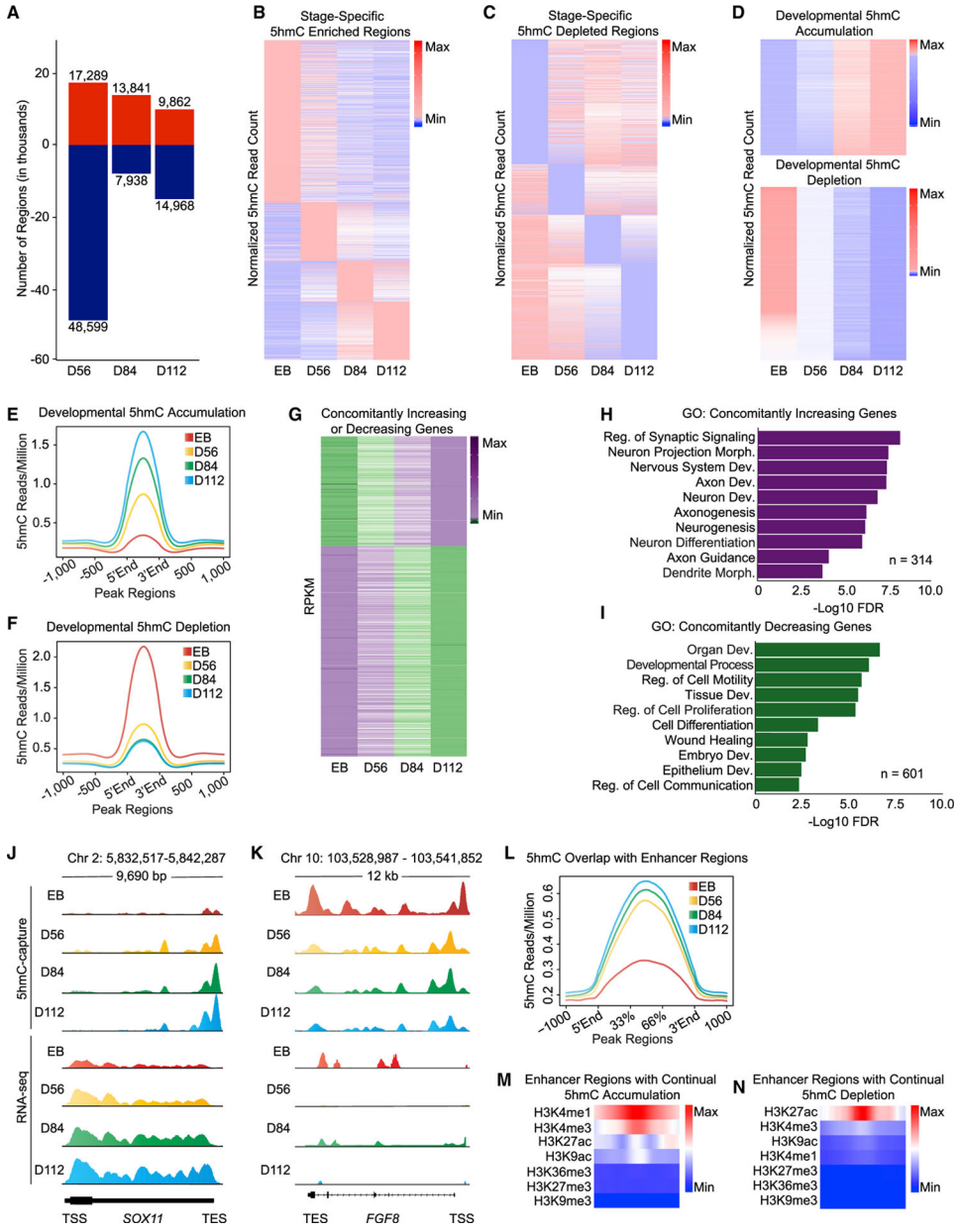


Figure 2. Dynamics of 5hmC regulation during forebrain organoid development

(A) Number of established and disappeared 5hmC peaks at D56, D84, and D112.

(B–D) Heatmaps of developmental-stage-specific DhMRs, where the color scale represents normalized 5hmC read counts.

(B) DhMRs that were enriched in developmental stages.

(C) DhMRs that were depleted in developmental stages.

(D) DhMRs with continual 5hmC accumulation (top) and continual 5hmC depletion (bottom) during organoid development.

(E and F) Average normalized 5hmC read counts per stage with continual 5hmC accumulation (E) and continual 5hmC depletion (F).

(G) Heatmap of RPKM (reads per kilobase per million) for genes that concomitantly increase in 5hmC and gene expression (top) or concomitantly decrease in 5hmC and gene expression (bottom).

(H and I) Gene Ontology (GO) analysis of the genes that concomitantly increase in 5hmC and gene expression (H) and concomitantly decrease in 5hmC and gene expression (I). Reg., regulation; Dev., development; Morph., morphogenesis.

(J and K) Normalized 5hmC read count and transcriptome across the concomitantly increasing *SOX11* gene (J) or the concomitantly decreasing *FGF8* gene (K).

(L) Average normalized 5hmC read counts per stage across enhancer regions.

(M and N) Enrichment of histone modifications at enhancer regions from fetal brains overlapped with DhMRs that continually accumulated (M) or lost (N) 5hmC.

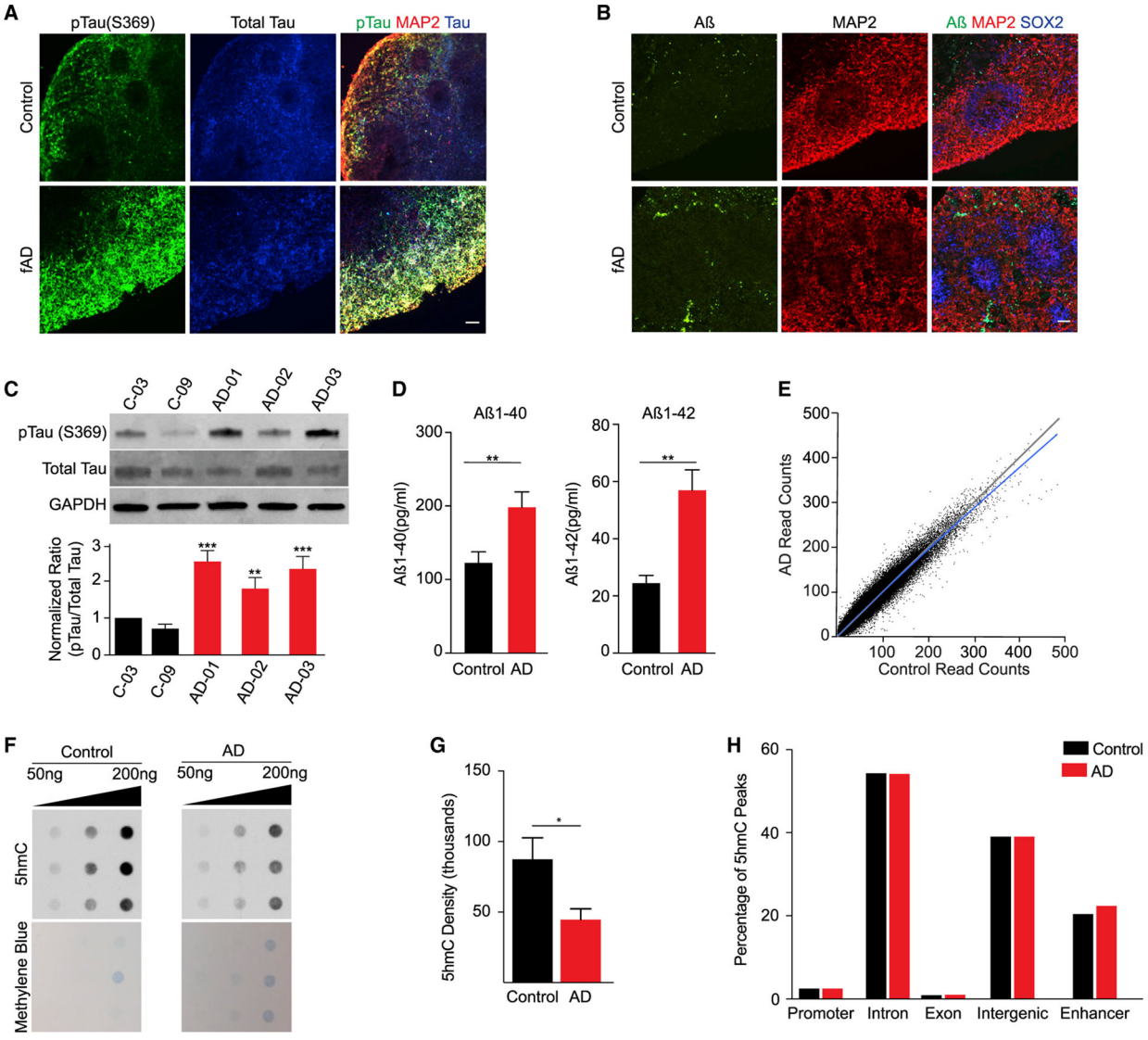


Figure 3. AD organoids recapitulate hallmark pathologies of human AD brains

(A) Representative phosphorylated Tau immunostaining of fAD organoids at day 84 and controls; scale bar = 50 μm.

(B) Representative amyloid-beta (Aβ) immunostaining of fAD organoids at day 84 and controls; scale bar = 50 μm.

(C) Immunoblot of phosphorylated and total Tau protein derived from independent control organoid lines (n = 2 biological replicates done in triplicate) and independent fAD patient organoid lines (n = 3 biological replicates done in triplicate) (*p < 0.05, **p < 0.01, ***p < 0.001, unpaired t test; data are represented as mean ± SEM).

(D) ELISA quantification of Aβ1–40 and Aβ1–42 peptide levels in control and fAD organoids (**p < 0.01, unpaired t test, n = 3 biological replicates; data are represented as mean ± SEM).

(E) Average normalized 5hmC read counts across the whole genome show that 5hmC is significantly depleted ($p = 2.513 \times 10^{-7}$ by unpaired t test) in AD versus control organoids.

(F) 5hmC dot blot showing whole-organoid 5hmC enrichment in controls versus AD organoids at day 84 (top). Methylene blue staining confirms equal amounts of DNA were loaded per sample (bottom).
(G) Quantification of 5hmC dot blot in controls and AD organoids ($p < 0.05$, unpaired t test, $n = 3$ biological replicates; data are represented as mean ± SEM).
(H) Proportions of 5hmC peaks across genomic features in control and AD organoids.

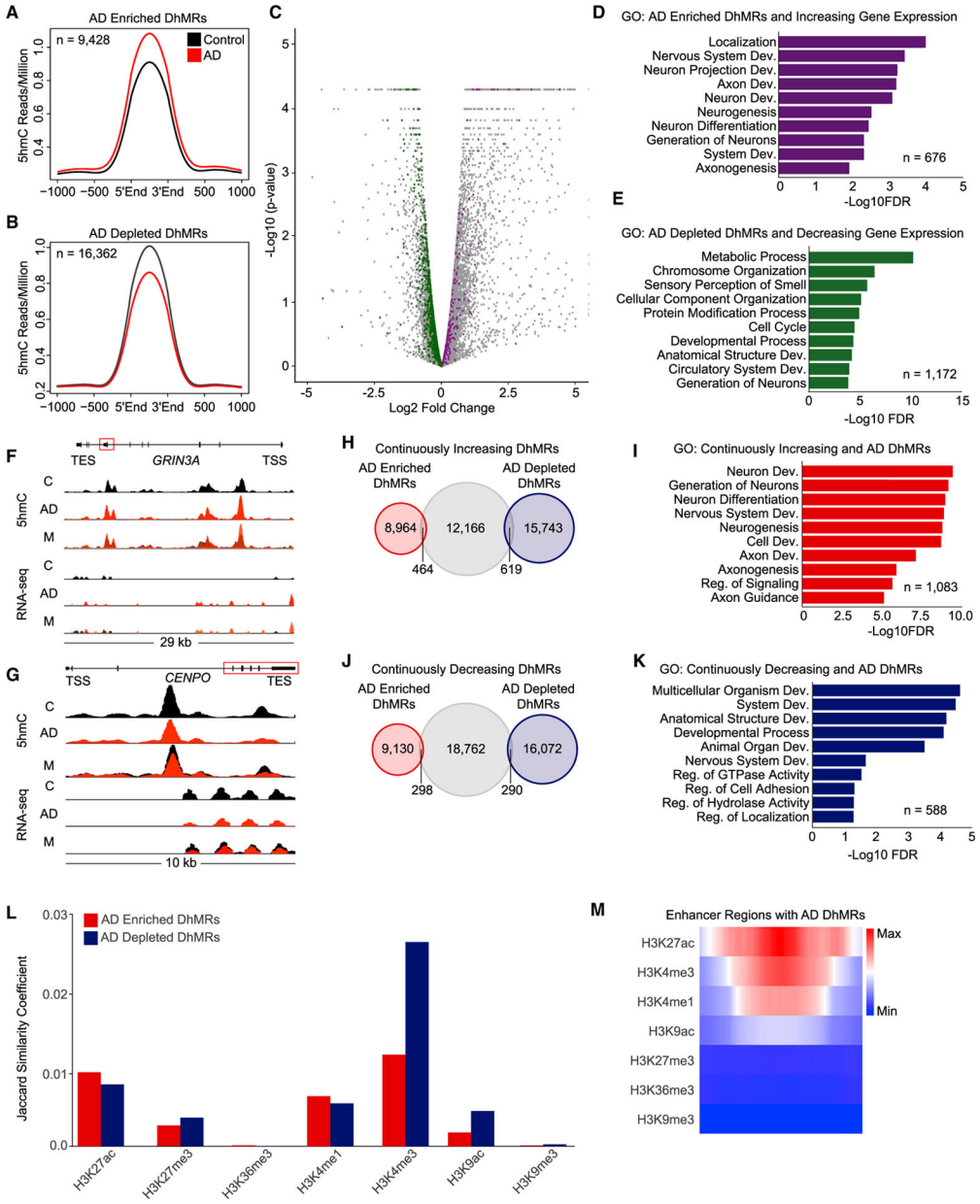


Figure 4. Aberrant alteration of 5hmC in AD organoids

(A and B) Average normalized 5hmC read counts at AD-enriched and AD-depleted DhMRs.

(C) Differentially expressed genes (n = 7,976 downregulated genes and n = 10,458 upregulated genes in AD organoids). Green dots: AD-depleted DhMRs with decreasing gene expression. Purple dots: AD-enriched DhMRs with increasing gene expression.

(D and E) GO analysis of genes annotated to AD-enriched DhMRs with increasing gene expression (D) and AD-depleted DhMRs with decreasing gene expression (E) in AD organoids.

(F and G) Normalized 5hmC read counts and transcriptome of *GRIN3A*, which depicts AD-enriched DhMRs with increasing gene expression (F), or *CENPO*, which depicts AD-

depleted DhMRs with decreasing gene expression (G). Red boxes indicate where on the gene the displayed peak region(s) originated from. C, control; M, merge.
(H–K) Venn diagrams and corresponding GO analysis results with respect to overlapped AD-enriched and -depleted DhMRs and continual 5hmC accumulation (H and I) or continual 5hmC depletion (J and K) during development.
(L) AD-depleted DhMRs are most similar to H3K4me3 regions, whereas AD-enriched DhMRs are most similar to H3K27ac and H3K4me1.
(M) Enrichment of fetal brain histone modifications in the overlapped regions between fetal brain enhancer regions and AD-specific DhMRs (n = 3,688 enhancers).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
5-hydroxymethylcytosine(5-hmC); Rabbit polyclonal	Active Motif	Cat#39769; RRID:AB10013602
Phospho-Tau (Ser396) Polyclonal	Thermo Fisher Scientific	Cat#44–752G; RRID:AB_2533745
Tau Monoclonal (T46)	Thermo Fisher Scientific	Cat#13–6400; RRID:AB_2533025
Polyclonal MAP2	Novus	Cat#NB300–213; RRID:AB_2138178
β-amyloid (D54D2) XP; Rabbit monoclonal	Cell Signaling	Cat#8243S; RRID:AB_2797642
Human SOX2 Affinity Purified Polyclonal	R and D Systems	Cat#AF2018; RRID:AB_355110
Chemicals, peptides, and recombinant proteins		
T4 phage β-glucosyltransferase	New England Biolabs	Cat#M0357L
UDP-6-N3-glucose	Jena Biosciences	Cat#CLK-076
AMPure XP beads	Beckman Coulter	A63881
disulfide biotin linker	Click Chemistry Tools	A112–5
Dynabeads MyOne Streptavidin C1 beads	Thermo Fisher Scientific	Cat#65002
Critical commercial assays		
NEBNext Ultra II DNA Library Prep kit for Illumina	New England Biolabs	Cat#E7645L
TruSeq RNA Sample Prep V2 kit	New England Biolabs	RS-122–2001
Deposited data		
Raw and analyzed sequencing data	This paper	GSE151818
Raw images	This paper	https://data.mendeley.com/datasets/bxp7jh6yw5/1
PSEN1 mutated neural progenitor cell line	Fetahu et al., 2019	Table S3
AD postmortem brains	Qin et al., 2020	Table S3
AD postmortem brains	Zhao et al., 2017	Table S3
AD postmortem brains	Lardenoije et al., 2019	Table S3
AD postmortem brains	Gasparoni et al., 2018	Table S3
AD postmortem brains	Li et al., 2019	GSE110732
Fetal brain histone modifications	Yan et al., 2016	Table S3
Epigenome RoadMap Project: fetal brain histone modifications	Kundaje et al., 2015	http://www.roadmapepigenomics.org/
Enhancer Atlas 2.0	Gao and Qian, 2020	http://www.enhanceratlas.org/
Experimental models: Cell lines		
Controls (CTRL): human induced pluripotent stem cells (hiPCS)	Provided by Dr. Gary Bassell (Emory University)	N/A
Familial Alzheimer Diseased (FAD): human induced pluripotent stem cells (hiPCS)	Provided by Dr. Chadwick Hales (Emory University)	N/A
Alzheimer Diseased (AD) human fibroblasts	Provided by Dr. Selina Wray (UCL Queen Square Institute of Neurology)	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
National Center for Microscopy and Imaging Research: ImageJ Mosaic Plug-ins	Schneider et al., 2012	https://ncmir.ucsd.edu/downloads/montaging_plugins.shtm RRID:SCR_001935
Bowtie v.2.2.6	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/index.shtml RRID:SCR_005476
MACS2 v.2.1.0	Zhang et al., 2008	https://github.com/taoliu/MACS RRID:SCR_013291
ngs.plot v.2.61	Shen et al., 2014	https://code.google.com/p/ngsplot/ RRID:SCR_011795
HOMER v.4.9	Heinz et al., 2010	http://homer.ucsd.edu/homer/mer/ RRID:SCR_010881
TopHat2 v.1.3.3	Trapnell et al., 2012	https://ccb.jhu.edu/software/tophat/index.shtml RRID:SCR_013035
Cuffdiff v.2.2.1	Trapnell et al., 2012	http://cufflinks.cbcb.umd.edu/ RRID:SCR_001647
BEDTools	Quinlan and Hall, 2010	https://github.com/arq5x/bedtools2 RRID:SCR_006646
SAMTOOLS v.1.9	Li et al., 2009	http://samtools.sourceforge.net/ RRID:SCR_002105
GO Consortium	Ashburner et al., 2000	http://geneontology.org
MuSiC	Wang et al., 2019	https://github.com/xuranw/MuSiC
Genomic Regions Enrichment of Annotations Tool (GREAT)	McLean et al., 2010	http://great.stanford.edu/public/html/