CrossMark

ORIGINAL INVESTIGATION

Investigation of genes important in neurodevelopment disorders in adult human brain

Gilles Maussion[1,2] · Alpha B. Diallo[1,2] · Carolina O. Gigek[2,3] · Elizabeth S. Chen[2,3] ·
Liam Crapper[1,2] · Jean-Francois Théroux[1,2] · Gary G. Chen[1,2] · Cristina Vasuta[1,2] ·
Carl Ernst[1,2,4,5]

Abstract Several neurodevelopmental disorders (NDDs) are caused by mutations in genes expressed in fetal brain, but little is known about these same genes in adult human brain. Here, we test the hypothesis that genes associated with NDDs continue to have a role in adult human brain to explore the idea that NDD symptoms may be partially a result of their adult function rather than just their neurodevelopmental function. To demonstrate adult brain function, we performed expression analyses and ChIPseq in human neural stem cell(NSC) lines at different developmental stages and adult human brain, targeting two genes associated with NDDs, *SATB2* and *EHMT1,* and the WNT signaling gene TCF7L2, which has not been associated with NDDs. Analysis of DNA interaction sites in neural stem cells reveals high (40–50 %) overlap between proliferating and differentiating cells for each gene in temporal space. Studies in adult brain demonstrate that consensus sites are similar to NSCs but occur at different genomic locations. We also performed expression analyses using BrainSpan data for NDD-associated genes *SATB2*, *EHMT1*, *FMR1*, *MECP2*, *MBD5*, *CTNND2*, *RAI1*, *CHD8*, *GRIN2A*, *GRIN2B*, *TCF4*, *SCN2A*, and *DYRK1A* and find high expression of these genes in adult brain, at least comparable to developing human brain, confirming that genes associated with NDDs likely have a role in adult tissue. Adult function of genes associated with NDDs might be important in clinical disease presentation and may be suitable targets for therapeutic intervention.

Abbreviations
NSC   Neural stem cells
FBC   Fetal brain cells
iPSC-NPC Induced pluripotent stem cell-neural progenitor cells

G. Maussion and A. B. Diallo contributed equally to this work.

✉ Carl Ernst
  carl.ernst@mcgill.ca

1 Department of Psychiatry, McGill University, Montreal, QC H4H 1R3, Canada

2 McGill Group for Suicide Studies, Douglas Hospital, Montreal, QC H4H 1R3, Canada

3 Disciplina de Genética, Departamento de Morfologia e Genética, Escola Paulista de Medicina/Universidade Federal de São Paulo, São Paulo, SP, Brazil

4 Department of Human Genetics, McGill University, Montreal, QC H4H 1R3, Canada

5 Douglas Hospital Research Institute, 6875 LaSalle Boulevard, Frank Common Building, Room 2101.2, Verdun, QC H4H 1R3, Canada

Background

Neurodevelopmental disorders (NDDs) are a heterogeneous spectrum of features that present as any disorder of early childhood that affects the brain and include seizure disorders, intellectual disability, autism spectrum disorders, and many others. Current classification of NDDs is based on presenting symptoms and clinical judgement, though major advances in genetics have identified mutations in several genes in individuals with NDDs, at once complicating and clarifying diagnoses. Genetic data suggest that both variable expressivity (individuals with mutations in the same gene but presenting different clinical symptoms) and locus heterogeneity (individuals with similar clinical

symptoms but with genetic lesions at different locations in the genome) are common, blurring diagnostic boundaries.

Because NDDs are diagnosed during early life periods, an intense focus has been placed on understanding the role of genetic variation known to cause disease in fetal or early brain development in animal and cell models. While this is unquestionably a reasonable study avenue, it does not occlude the idea that these same genes may have a role in adult brain. The function of genes implicated in neurodevelopmental disorders have thus not undergone the kind of investigation that one might normally expect for genes with such a high relevance to human health. Most importantly, it may be that patient phenotypes are not solely due to potential genetic deficits during brain development, but rather because of the ongoing function of these same genes in adult brain. This could have implications for the development of treatment for NDDs because it implies a component of the disease may be due in part to dysfunction of the adult activities of these genes. Undoubtedly, any genetically driven treatment regime must mount the hurdle of treatment delivery, but the angle proposed here may provide some hope that NDD phenotypes are not simply fixed after neural circuitry formation. The hope with the current study is to explore the adult function of some genes that are well known to cause a neurodevelopmental disease phenotype and to begin to assess their spatiotemporal expression and DNA association patterns. Our goal is to eventually provide treatment for some NDDs and we believe that studying the effects of genes associated with NDDs will help to dissect the critical periods of their activity that are causing disease. For some genes, it may be that a function in the adolescent period is critically important to solidifying a disease progression begun during early brain development. If so, targeting these specific functions may become particular therapeutic avenues. Given that there is almost no treatment available beyond symptom management for any NDD, we believe the current investigation to be an important first step.

Several neurodevelopmental disorders are caused by mutations in DNA association factors implicated in brain development (Krumm et al. 2014), two of which will be investigated here because we have experience working with them and wanted to test our hypothesis in genes with which we have worked previously. *SATB2* is a nuclear matrix association protein thought to associate with AT-rich regions of the genome, where mutations cause craniofacial anomalies, severe speech delay, feeding difficulties, and autism spectrum disorders (Docker et al. 2014; Talkowski et al. 2012). *SATB2* is also a molecular marker of cortical neurons (Alcamo et al. 2008; Britanova et al. 2008). *EHMT1* is a histone methyltransferase, capable of adding methyl groups to amino acid residues such as lysine 9 on Histone 3 to affect gene transcription. Mutations in *EHMT1*

cause severe intellectual disability and behavioral disorders (Kleefstra et al. 2012). Throughout this study, we also included the transcription factor *TCF7L2*, a high-mobility group-box protein (Grosschedl et al. 1994; Roose and Clevers 1999). Diseases associated with variation in *TCF7L2* include type 2 diabetes (Florez et al. 2006) though it has not been implicated in NDDs despite its role in WNT signaling, a pathway strongly suspected of being involved in NDDs (Kalkman 2012). We included *TCF7L2* because it is a well-studied transcription factor, is thought to be expressed ubiquitously in all tissues across developmental stages, has excellent ChIP antibodies available, and known DNA consensus binding regions. In this sense, *TCF7L2* serves as a form of technical control across experiments.

Methods

This work was reviewed and approved by the ethical review board of the Douglas Hospital Research Institute of McGill University.

Neural stem cell (NSC) lines

Two neural stem cell lines are used in the current study; the FBC line is derived from human fetal brain (ReNCellVM, Millipore SCC008), while the iPSC-NPC (induced pluripotent stem cell-derived neural progenitor cells) line was derived from skin cells of a healthy human subject. Both NSC lines were stored as stocks in liquid nitrogen and can be grown indefinitely. Both lines were grown under identical conditions: cells were grown on poly-L-ornithine/laminin (Sigma, St. Louis, MO, USA) coated 6-well plates and maintained in 70 % DMEM, 2 % B27, 1 % Penicillin/Streptomycin (Life Technologies Foster City, CA, USA), 30 % Ham's F12 (Mediatech Herndon, VA, USA), with 20 ng/mL bFGF (R&D Systems, Minneapolis, MN, USA), 20 ng/mL EGF, and 5 μg/mL heparin (Sigma). To differentiate cells, we removed growth factors and let cells differentiate for 30 days. Media were changed every 3 days, whether cells were proliferating or differentiating.

Whole-cell recordings

Primary cell cultures were prepared as described above then transferred to a solution for whole-cell patch clamp recordings. The extracellular HEPES-based saline contained, in mM: 140 NaCl, 5.4 KCl, 2 CaCl$_2$, 1 MgCl$_2$, 15 Hepes, and 10 glucose (pH 7.3–7.4; 295–305 mOsmol). Cells were identified based on their morphology using an inverted microscope (Nikon Eclipse, E600FN), and recordings were performed at room temperature. Whole-cell patch clamp recordings were obtained using borosilicate

pipettes (3–6 MΩ), filled with intracellular solution that contained, in mM: 154 potassium gluconate, 2 EGTA, 1 MgCl2, 10 phosphocreatine, 10 HEPES, 2 Mg-ATP, (pH 7.2–7.3; 275–285 mOsmol). The resistance of the pipettes was determined using Ohm's law (V = IR), by injecting a small current in the circuit, to drop the voltage 5 mV from holding. This was done when the pipette was in the bath position before patching. Where indicated, TTX (200 nM) and TEA (50 μM) were added to the saline solution. Data were acquired using a Multiclamp 700B amplifier (Axon Instruments, Foster City, CA). Currents were filtered at 2 kHz and digitized at 20 kHz. Responses were analyzed offline using Clampfit (Molecular Devices) and GraphPad (Prism). Data were expressed as mean ± s.e.m.

ChIP quality control experiments and neural stem cell (NSC) characterization

Chromatin immunoprecipitation is dependent on excellent antibodies that accurately interact with proteins of interest (Landt et al. 2012). While all three proteins assessed in this study have several antibodies available, including ChIP-grade antibodies, we performed in-house assays to ensure quality (Supplemental Fig. 1A–C).

We selected two neural stem cell lines with which to assess proliferating and differentiating cell stages. We purposefully chose two cell lines from different donors and derived through different methods [FBCs: extracted from fetal brain; iPSC-NPC: derived from human skin, reprogrammed with Yamanaka factors (Takahashi et al. 2007) and differentiated into NSCs] with the conservative rationale that fundamental DNA association binding properties should be independent of cell source, genetic background, and cell fate. Though neural stem cell lines are well known to form heterogeneous cell types (Aradi et al. 2004), we reasoned that identifying targets common across diverse cell types would provide extremely high confidence calls. Supplemental Fig. 1D–E shows relative expression of well-known markers across cell lines, fetal brain, and adult frontal cortex. Markers for proliferating cells include *OCT4* (negative), *NANOG* (negative), *NESTIN* (positive), *MUSASHI* (positive) and *PAX6* (positive). We used a battery of markers including *GFAP, OLIG2, VGAT*, and *TH* in differentiating cells. Finally, we characterized the cell lines electrophysiologically (Supplemental Fig. 1F–L).

Adult human brain tissue

All frozen human brain tissue (n = 6) was supplied by the Bell-Douglas Hospital Brain Bank (https://www.douglas.qc.ca/section/brain-bank-131) and was derived from inferior orbital–frontal cortex (Brodmann Area 10) from psychiatrically screened (Dumais et al. 2005) subjects who

died in a motor vehicle accident (n = 5) or a heart attack (n = 1). The mean age was 36 (range 19–61 years), mean brain pH was 6.45 (range 6.3–6.75), and a mean postmortem interval was less than 24 h for all cases. All subjects were screened for 50 different toxicological substances, including most street drugs and common medications. All subjects were negative for all toxicological screens.

RNA Quantitative PCR

RNA from cells was extracted using RNeasy MinElute Cleanup Kit (Qiagen) and commercially available RNA from eight different tissues (Ambion Total RNA: liver, kidney, spinal cord, adult frontal lobe, fetal brain, lung and testis) was used for cross-tissue analyses. Induced pluripotent stem cells and fibroblasts were from control subjects and provided by the Harvard Stem Cell Institute and Coriell, respectively. cDNA was synthetized using M-MLV Reverse Transcriptase (Invitrogen) and Taqman probes were used for all genes assessed. Real-time PCR reactions were run in triplicate using the ABI 7900HT Fast Real-Time PCR System and data was collected using the Sequence Detection System (SDS) software (Life Technologies). We checked the expression of four housekeeping genes (TBP, NONO, GAPDH and ACTB, Supplemental Fig. 2). We performed an assessment of stability across all cell lines and tissues. All RNA qPCR are normalized to GAPDH and ACTB. Cell numbers and RNA concentrations were identical across experiments.

Chromatin immunoprecipitation (ChIP)

For ChIP, 10^7 cells were dissociated by trypsin, crosslinked with 1 % formaldehyde (Tousimis NC9611804) for 10 min, pelleted and re-suspended in lysis buffer with protease inhibitors. Samples were sonicated with Labsonic M Ultrasonic Homogenizer (Labsonic Sartorius Stedim, EUA) for 7 × 30 s cycles at 30 % power. Chromatin was sheared at 15 1-s pulses followed by 2-min rest periods at 50 % power using the same ultrasonic homogenizer (200–1000 bp). Brain tissue underwent a similar procedure though 100 mg of tissue from frontal cortex was used, and where ChIP was performed in a single pooled reaction from all six brain samples. We followed the exact procedure laid out here (Matevossian and Akbarian 2008) for cell separation. One hundred μg/μL of DNA from either cell lines or pooled brain samples was incubated with antibodies (5 μg/μL) overnight at 4 °C overnight in dilution buffer. Several antibodies were tested in brain tissue prior to ChIP (SATB2: Epitomics:2819-1, Abcam: Ab51502; TCF7L2: Millipore 17-10109; EHMT1: Abcam: ab41969). We used genes reported to be bound by each factor as positive control tests prior to ChIPseq (SATB2 target: *CTIP2A4*

(Chen et al. 2008); TCF7L2 target: *AXIN2* (Valenta et al. 2003); EHMT1 target: *MAGEB16* (Link et al. 2009). IP washes, elution and crosslink reversal were performed with the Magna ChIP A—Chromatin Immunoprecipitation Kit (Millipore), according to manufacturer's instructions. IP and input DNA were purified by phenol chloroform extraction, precipitated in ethanol, and re-suspended in sterile water. We used the Illumina ChIP-Seq DNA sample prep kit (IP-102-1001) for all experiments and followed instructions from the manufacturer. All ChIP experiments were performed independently for each individual.

ChIP sequencing and analysis

Libraries were sequenced in indexed pools of 4–6 using a HiSeq 2000 (Illumina) at the McGill University and Genome Quebec Innovation Center. In all cases, we used 50 bp single-end sequencing and sequenced one library of IgG control to help establish background parameters. We used Bowtie 2, an ultrafast and memory-efficient tool for aligning sequencing reads to long reference sequences, with the very sensitive parameter (allowing 0 mismatches in a seed alignment) and MACS (Zhang et al. 2008) for peak calling with default parameters, except for the *p* value cutoff which we set to 10^{-6}, instead of 10^{-5}. We fixed the duplicate number at five, meaning duplicates were included in all analyses but where only five duplicates were allowed per read. Any peaks detected in the IgG control were systematically removed from all SATB2, TCF7L2, and EHMT1 peaks called. We visually inspected peaks to assess background level and duplicates in the Integrated Genomics Viewer. To associate ChIP peaks with genes, we considered any regions 5 Kb upstream of a gene and any region 2 Kb downstream of gene. To calculate overlapping peaks across independent experiments, we used the BEDTools command intersect tool to screen for overlaps between two sets of genomic features. We considered peaks with one or more base pairs to be overlapping. For peak coordinates, we used the peak outputs defined by MACS under default conditions.

DNA Quantitative PCR (ChIP validation)

To validate ChIPseq data, we reperformed ChiP experiments using freshly grown cell lines and performed a new IP reaction. We selected two genomic loci per gene based on peaks across all cell lines. The primers used to validate ChIP-sequencing results were designed using Primer Quest online tool (http://www.idtdna.com/primerquest/home/index). The primers are listed in the Supplemental Table 1. The PCRs were performed using a 7900HT real-time PCR machine (Applied Biosystems). Each well contains 10 μL

of 2× SybrGreen PCR master Mix, 10 pmol of forward and reverse primers, 1 μL of DNA and H_2O qsp 20 μL. Absolute quantifications were performed for every region of interest and control regions using calibration curves made of serial dilutions (between 3 pg and 50 ng of DNA from blood). As control regions, we have randomly chosen two regions extracted from our ChIP-sequencing results and which were predicted not to be associated with TCF7L2 SATB2 or EHMT1. Absolute quantifications of regions of interest were normalized with those obtained for these two negative control regions.

Results

SATB2, *TCF7L2*, and *EHMT1* are expressed in proliferating and differentiating neural stem cells and adult human brain

We selected several tissues from human to assess how expression levels compared to non-neuronal tissues—several NDDs have clinical features not only in brain, implying that the genes have important functions elsewhere in the body. Across multiple tissues (Fig. 1a–c), we found that expression of *SATB2* was strongest in fetal brain, but was present in cortical (frontal lobe) and subcortical (hippocampus) brain structures and testes. Expression in other tissues assessed by qPCR is negligible with exception of the testes. For *EHMT1*, we found high expression in fetal brain, higher than either adult cortical or subcortical structures, with levels of expression in adult brain comparable to all other tissues assesses. *EHMT1* expression is notably strong in testes. For *TCF7L2*, a gene not associated with NDDs, we found a slight increase in expression in the adult frontal lobe compared to fetal brain tissue, though *TCF7L2* is expressed across all tissues assessed.

Clinical phenotypes associated with *SATB2* or *EHMT1* mutations have implicated several organs beyond brain, notably palate, bone and brachial arch derivatives in *SATB2* mutation cases (Zarate et al. 2015), and cardiac anomalies in *EHMT1* mutation cases (Kleefstra et al. 2006; Tachibana et al. 2005). To determine if expression of these genes is higher in organs known to be associated with disease in adult tissues, we took advantage of RNA sequencing data from GTEx (Mele et al. 2015) (Supplementary Fig. 3). For *EHMT1*, we observed no major differences (>fivefold) in expression between tissues. Similarly, SATB2 showed consistent expression across tissues, though there was notable increase in expression in the spleen compared to all other tissues. Finally, *TCF7L2* shows high variation in expression across tissues, where it is for example, highly expressed in blood vessels but low in blood cells. Variation *TCF7L2*
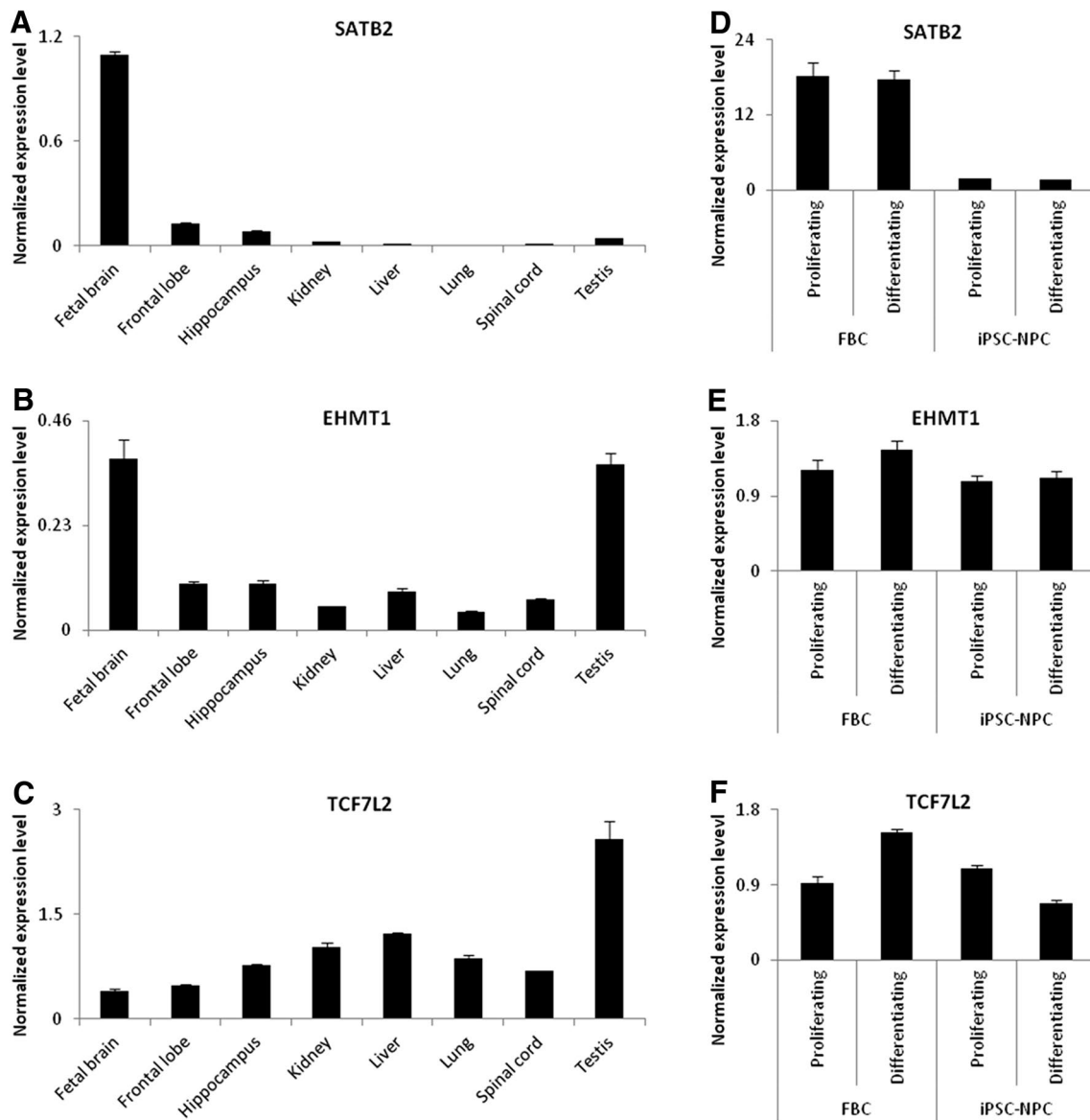
Fig. 1 mRNA expression of SATB2, TCF7L2, and EHMT1 in human tissues and neural stem cell lines

is associated with diabetes (Grant et al. 2006), a result of dysfunctional *TCF7L2* signaling in the pancreas (Mitchell et al. 2015). Notably, *TCF7L2* expression in pancreas is one of the lowest areas of expression compared to other assessed tissues in GTEx (Supplementary Fig. 3C). This finding and those from the *EHMT1* and *SATB2* suggest that relative expression level should not be used to guide where disease is most likely to occur. In other words, the mere fact that a gene is expressed in a given tissue means mutations in that gene could cause disease in that organ, even if expression is higher in other tissues.

We assessed the expression of each gene in FBC and iPSC-NPC cell lines at two functionally defined neurodevelopmental time points (Fig. 1d–f). These time points differ in their exposure to bFGF and EGF—proliferating cells are maintained in culture with these factors, while differentiating cells are post-mitotic and have persisted for 30 days in the absence of both factors. We could detect expression of each gene at all time points, but there was a notable difference in *SATB2* expression in both proliferating and differentiating stages in the iPSC-NPC line, suggesting large differences in the degree of expression of *SATB2* in differentially derived cell lines. With this exception, we observed relatively consistent expression of all genes, including *TCF7L2*, in both proliferating or differentiating stages.

Expression of *SATB2, EHMT1*, and *TCF7L2* across multiple brain regions and developmental time windows

Results from our expression analyses suggest that both SATB2 and EHMT1 are expressed in adult brain, but that there is lower expression than in fetal brain. To explore the expression of these genes in adult brain, we took advantage of data developed by the BrainSpan consortium (Kang et al. 2011), to better understand regional effects of expression patterns in adult brain (Fig. 2). For *SATB2*, we observed a strong cortical expression pattern, with essentially zero expression in cerebellum and sub-cortical structures (Fig. 2a) in adult brain. This suggests that *SATB2*, a gene implicated in cortical neuron development

(Szemes et al. 2006), continues to be expressed in cortical cells in adult brain across all regions assessed. EHMT1 had relatively consistent expression across all adult brain structures, though with a notable increase in cerebellum (Fig. 2a).

We wanted to understand how expression of *SATB2* and *EHMT1* changed over human neurodevelopmental time windows in cortical and subcortical structures (Fig. 2b, c). In dorsal–lateral prefrontal cortex, for both *SATB2* and *EHMT1,* we found higher expression in post-conception weeks 8–36, with expression of both decreasing but stabilizing from the early postnatal years until age 40 years. We observed a similar trend in a subcortical structure (hippocampus), including for *SATB2*, with expression higher earlier in development. We continue to
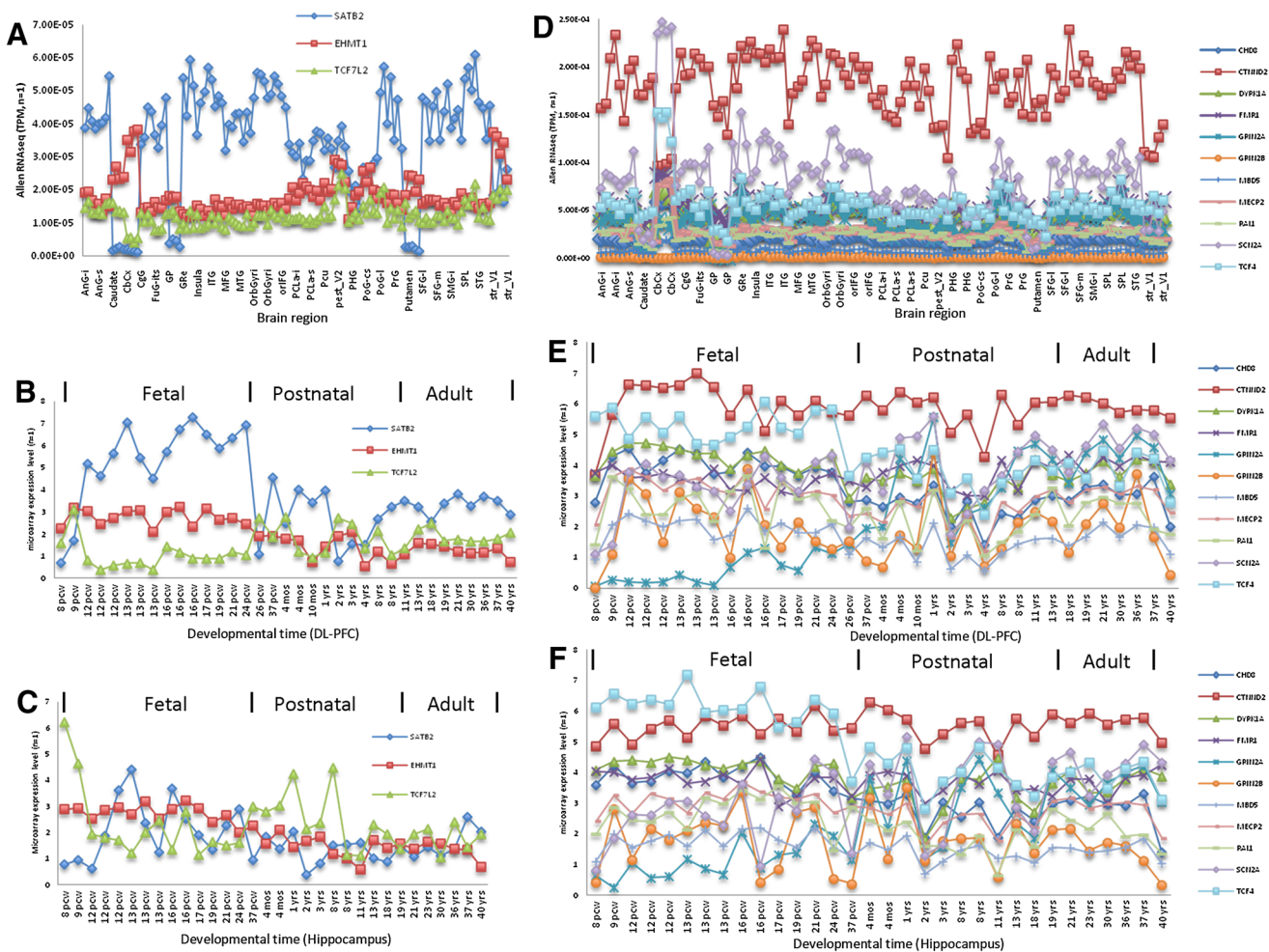

Fig. 2 Allen Brain Atlas data showing gene expression across brain tissues and neurodevelopmental time. **a** RNAseq data showing relative expression values across several adult brain structures. **b** Microarray expression patterns of *SATB2, EHMT1*, and *TCF7L2* in DL-PFC (a cortical structure) from post-conception week (PCW) 8 to age 40 years. **c** The expression patterns of *SATB2, EHMT1*, and *TCF7L2* in hippocampus (a subcortical structure) from post-conception week

(PCW) 8 to age 40 years. **d** RNAseq data showing relative expression values across several adult brain structures for genes associated with NDDs. **e** Microarray expression patterns of genes implicated in NDDs in DL-PFC (a cortical structure) from post-conception week (PCW) 8 to age 40 years. **f** The expression patterns of genes implicated in NDDs in hippocampus (a subcortical structure) from post-conception week (PCW) 8 to age 40 years

include *TCF7L2* in these analyses and found that *TCF7L2* is relatively stable across brain regions and through different developmental time windows in either cortical or subcortical structures.

Expression of genes implicated in NDDs show expression levels in adult brain similar to fetal brain

The purpose of this study is to demonstrate that genes associated with NDDs might also have important function in adult human brain. While we chose to study two genes in depth, we reasoned that including information on well-known genes implicated in NDDs could support this idea. To this end, we plotted Allen Brain Atlas data for several genes (Fig. 2d–f), similar to what we did for *SATB2, EHMT1*, and *TCF7L2*. First, we plotted region-specific information in the adult brain and found that most genes showed consistent expression across all brain regions. The exception to this was for *SCN2A* and *TCF4*, which showed high levels of expression in the cerebellum, while *CTNND2* showed low levels of expression, at least compared to the their levels of expression in other brain regions. *SCN2A* also showed decreased expression in the striatum (caudate nucleus and putamen). Next, we assessed the consistency of expression over developmental time windows, again using one cortical structure (Fig. 2e) and one subcortical structure (Fig. 2f). Our first observation is that most of these genes are stable across developmental time windows whether in cortex or hippocampus. We did observe, however, decreased expression of *TCF4* after birth in both structures, and increased expression of *GRIN2A* over time—which is a well-known effect for this NMDA receptor subunit (Zhong et al. 1995). Using expression level as a proxy for functional action, these data suggest that most genes important in neurodevelopment disorders are expressed in adult brain at levels similar to fetal brain.

ChIPSeq in proliferating and differentiating cells

To assess the degree of spatiotemporal differences in binding across proliferating and differentiating neural stem cells, we performed a large-scale ChIPseq experiment using antibodies targeting SATB2, EHMT1, and TCF7L2 proteins. These experiments were performed in two NSCs at two developmental stages (Table 1 shows all sequencing data related to these experiments). We wanted to understand if two defined developmental timepoints in the same cell line showed different or similar binding patterns. While we do not consider cell line developmental stages to approximate human brain developmental time periods, they do represent distinct phases of development. If genes continue to be expressed at both time points (Fig. 1), one might expect them to target similar regions of the genome. For SATB2, we found that 976 peaks were common (i.e., with at least 1 basepair overlapping raw ChIP peaks) across all cell line experiments, corresponding to 17–34 % of all peaks across cell lines (Fig. 3a; Table 1). When we looked across cell line (FBC or iPSC-NPC) or cell state (proliferating or differentiating), we also found significant overlap. Testing within the same NSC line at different neurodevelopmental stages, we found a 43 % overlap of peaks in iPSC-NPC and a 48 % overlap in FBC peaks, suggesting that just under half of SATB2 targets are common in proliferating and differentiating NSCs. Assessing differences across NSCs at the same developmental stage, we found that 41 % of peaks in proliferating cells and 41 % of peaks in differentiating cells are common, despite the different origins, genetic background, and cell fate of the different NSC lines. *SATB2* expression patterns were drastically lower in the iPSC-NSC line (Fig. 1d), though there appeared to be no difference in the total number of ChIP peaks between these two cell lines. This suggests that DNA binding properties are not necessarily a direct function of expression level. These results imply

Table 1 Sequencing reads used to calculate peaks for each independent analysis for each of three different antibodies

| Targeted protein	Mapping	FBC pro	FBC dif	iPSC-NPC pro	iPSC-NPC dif	Brain
SATB2	Total reads	52,730,285	27,284,216	49,309,774	33,707,859	33,606,420
	Unique aligned	5,386,800	1,561,965	4,376,548	2,254,914	829,697
	Peaks	4655	2830	5447	3407	3467
TCF7L2	Total reads	50,213,400	30,628,062	48,197,171	27,816,136	20,590,701
	Unique aligned	4,638,199	2,066,939	4,704,725	3,161,808	2,998,465
	Peaks	4840	1721	2699	5718	2595
EHMT1	Total reads	54,149,894	23,058,763	67,129,321	42,328,222	30,014,331
	Unique aligned	11,193,218	5,240,846	4,944,462	3,978,119	15,659,260
	Peaks	43,995	744	2918	5572	193,375
EHMT1 (replicate)	Total reads	18,742,992	36,617,759			
	Unique aligned	7,888,963	14,959,428			
	Peaks	90,608	894			
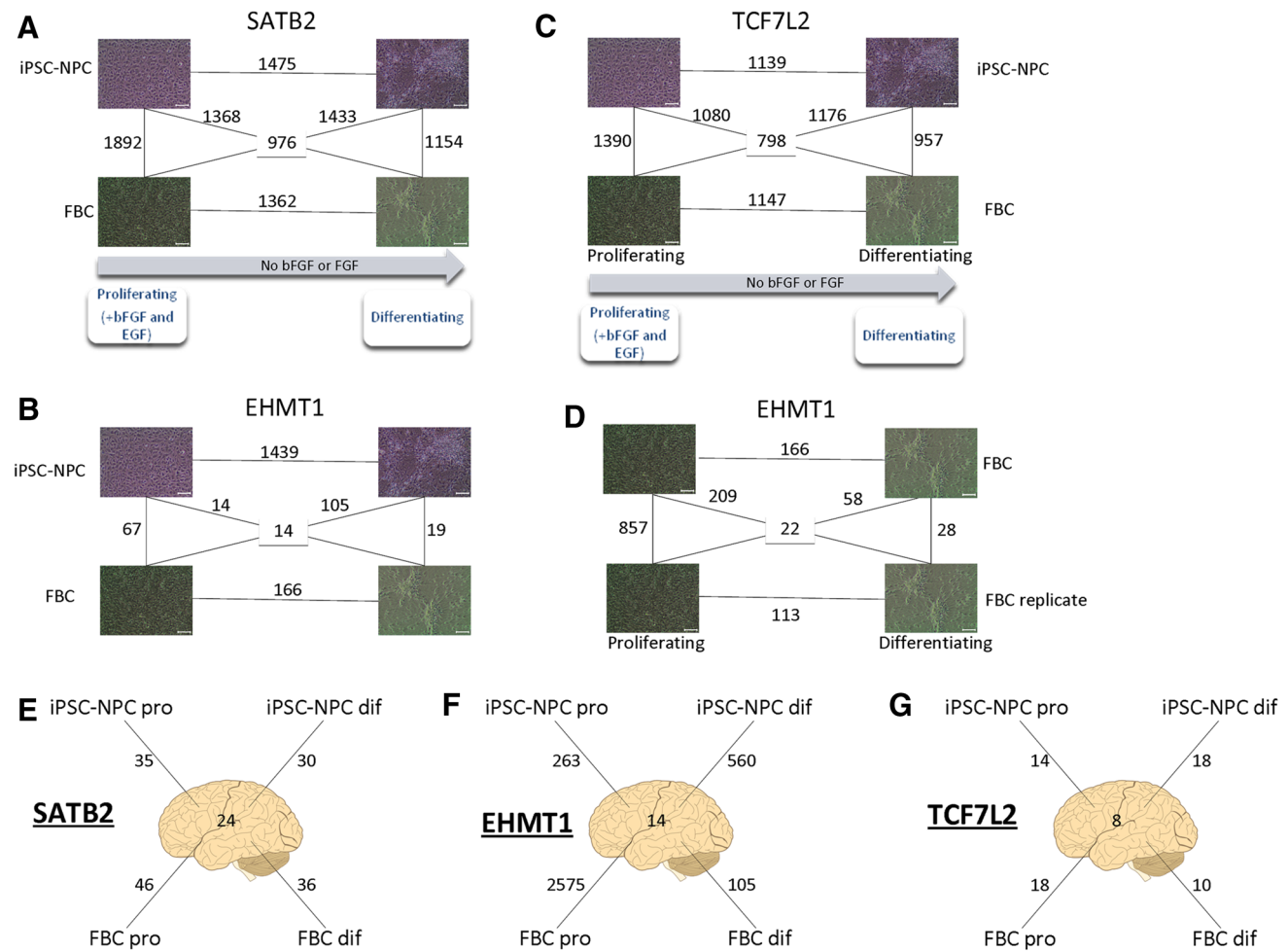
Fig. 3 DNA association overlap regions in neural stem cell lines and adult human brain. **a** A *square* design setup to show the number of overlapping peaks across developmental stages (proliferating and differentiating) and neural stem cell lines (iPSC-NPC and FBC) in the anti-SATB2 ChIPseq experiment. Images represent what NSCs look like at each stage and *scale bar* is 100 μM—these images are the same for all analyses. A *number* depicted on a *line* connecting two analyses represents the number of common peaks (>1 bp overlapping) that were identified while the *middle box* is the number of peaks common to all analyses (see Table 1 for the total peaks identified per analysis). **b** Graphic showing the number of common peaks between each analysis for the anti-EHMT1 ChIPseq experiment. **c** Graphic showing the number of common peaks between each analysis for the anti-TCF7L2 ChIPseq experiment. **d** A replicate experiment performed only in FBCs in a new batch of proliferating and differentiating cells. **e** This graphic shows the number of overlapping peaks in the anti-SATB2 ChIP experiments identified between each NSC line at each stage and adult human brain. The *number* in the *middle* represents the total number of common peaks across all five ChIP experiments (adult brain, 2 NSC lines at two developmental stages). *pro* proliferating and *dif* differentiating. **f** The common ChIP peaks in the anti-EHMT1 ChIP experiment. **g** The common ChIP peaks in the anti-TCF7L2 ChIP experiment

a high degree of heterogeneity across different cell lines (~60 % of binding regions are different) and a somewhat high degree of heterogeneity across developmental time periods in the same cell line (~45 % of targets are identical). Still, it is not clear what even the expected level of overlap should be for either of these analyses, so this provides at least some guidance for future experiments on the spatiotemporal function of DNA association proteins.

Analysis of EHMT1 ChIP results was very different than SATB2 binding patterns (Fig. 3b). First, we observed peculiar peak frequencies between FBC developmental stages, where there were many peaks in the proliferating state and very few in the differentiating state (Table 1), a result not explained by differential expression of EHMT1 at each stage (Fig. 1e). Twenty-two percent of peaks identified in differentiating FBCs were observed in proliferating FBCs, but only 0.3 % of peaks seen in proliferating FBCs were seen in differentiating FBCs. This was not the case in iPSC-NPCs across developmental stages, where we observed 49 % of peaks overlapping (Table 1), similar to results from SATB2 possibly suggesting peculiarities with FBCs. Across NSCs but in the same developmental stage, we observed

<2.5 % overlap at either stage. We were concerned that an undetected technical error occurred in the FBC EHMT1 ChIPseq experiments, possibly explaining the wide variation in peak number between proliferating and differentiating FBCs. We therefore reperformed ChIPseq using a newly thawed batch of FBCs in both proliferating and differentiating states (FBC replicate). For this experiment, we found an identical relationship as that first observed: there were many peaks in FBC proliferating cells, and very few in FBC differentiating cells, and with low overlap of peaks between FBC proliferating cells in initial experiment and replicate experiment (Fig. 3d). This suggested to us that EHMT1 may be very active in FBC proliferating cells, possibly suggesting that EHMT1 associates with the genome in a dynamic way, highly dependent on cell state.

For TCF7L2, we found an almost identical pattern to what was identified for SATB2, that is, a significant degree of overlap of TCF7L2 peaks across NSCs and at developmental stages, where 798 peaks were common across all four independent ChIPseq experiments (Fig. 3c). Specifically, 67 % of peaks were common between FBC proliferating and differentiating cells, while 42 % were common in iPSC-NPCs in proliferating and differentiating states. Across NSCs but in the same developmental stage, we found a 51 % overlap in proliferating cells, and a 55 % overlap in differentiating cells (Table 1).

ChIPSeq in adult human frontal cortex

A major purpose of this study is to determine if genes important in NDDs have a role in adult brain, a question that may lay a foundation for eventual therapeutic intervention because adult function of mutant genes might be treatable, whereas developmental function mostly cannot (due to timing of detection of genetic anomaly, accessibility in utero, and wiring of the nervous system), and one which might suggest that some clinical features of what are considered NDDs are in fact caused by postnatal function of genes of interest such as *EHMT1* and *SATB2*. To provide a proxy for 'adult function', we reasoned that identifying ChIP peaks in adult human brain tissue is a reasonable first step to determine if these genes have a role in adult human brain; we acknowledge that the current experiments are not functional; however, we do think that demonstrating expression in adult brain and identifying DNA association regions are reasonable output measures to begin to assess this question.

Similar to cell lines, we were able to detect several ChIP peaks for each gene (Table 1) in adult frontal cortex (ChIP was performed in tissue from six independent control individuals). For SATB2 and TCF7L2 proteins, we found ChIP peak frequencies similar to cell lines (Table 1), but a much higher number for EHMT1 peaks in brain than in cell lines. We first asked what the overlap was between ChIP peaks

common in cell lines compared to adult frontal cortex. There were 3467 SATB2 genomic targets identified in adult human brain, of which only 24 were common with high confidence (peaks present in all four independent ChIPseq experiments) NSC targets, and we show the overlap with each cell line (Fig. 3e). There was exceptionally low overlap levels between each NSC experiment and brain, suggesting that SATB2 binding regions are more common between cell lines than between any cell line with adult brain. While we assume that ~60 % of SATB2 binding regions can be attributed to differences in cell source (Fig. 3) in NSCs, the profoundly low level of overlap between adult brain and NSC peaks suggests that SATB2 may have very different function in adult brain than it does in proliferating or differentiating NSCs. We include the caveat that brain tissue from frontal cortex is a mixed population of cells, very different from NSCs, and that proteins were extracted fresh from cells whereas postmortem brain has been stored at −80 °C for some time (see "Methods"). That said, our ChIP QC (Supplemental Fig. 1A–C) was performed in adult brain tissue, so we can be confident that ChIP is technically sound in adult brain.

For EHMT1 in human brain, we observed 193,375 peaks, a number three- to fourfold higher than what we observed in cell lines (Table 1). Given the predicted ubiquity and context specificity of EHMT1 DNA binding, we observed in FBCs even on replicate experiments, it may be that this reflects different action of EHMT1 in the 6 different individuals used in this pooled brain analysis. In this case, it suggests that the action of EHMT1 might be highly context, state, and individual dependent—a finding perhaps not surprising for a protein that operates at the genomic level and needs to respond dynamically to cell stressors and other external cues.

TCF7L2, our 'positive control' for technical procedures throughout because it is such a well-studied protein, showed a similar result to SATB2—there was exceptionally low overlap between all cell lines and adult brain.

To demonstrate peak calls in these data, we provide two ChIP peaks per gene, with one brain-specific peak and one cell line-specific peak (Fig. 4) and show the absence of any sequencing reads in the comparison tissue (i.e., cell line compared to brain). In Fig. 4, we show IGV images with raw sequencing reads delineating the called peak. For SATB2 protein for example, there are read pile-ups in an intron of the *E2F3* gene in all cell lines, with no reads from brain ChIPseq data. Conversely, for *CPT1C*, there is a pile-up of reads in DNA extracted from adult frontal cortex, but no reads in this same region from any cell line. The *CPT1C* gene is carnitine palmitoyltransferase 1C, a gene implicated in feeding behaviors in mouse knockout studies (Gao et al. 2009; Wolfgang et al. 2006), where feeding difficulties are a phenotype of individuals with *SATB2* deficiency (Rosenfeld et al. 2009).
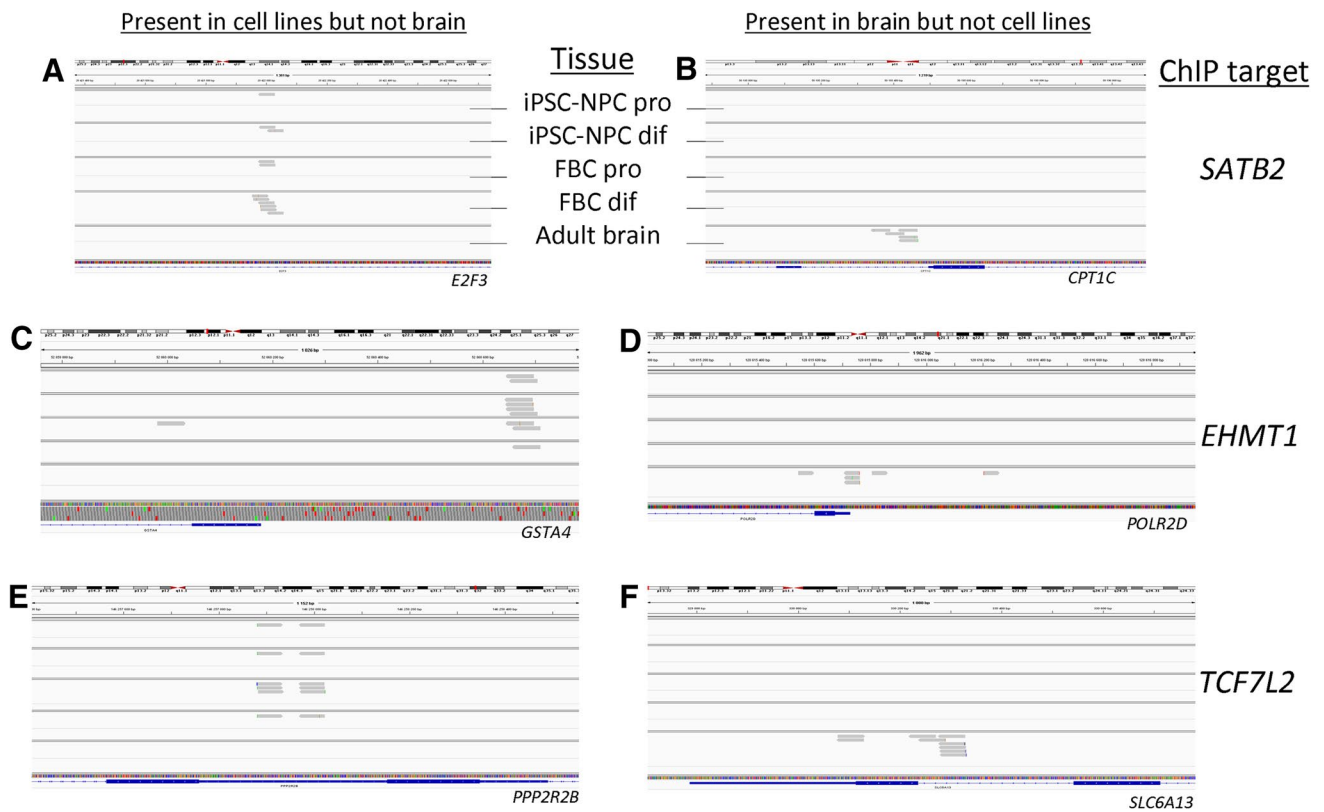
Fig. 4 ChIPSeq data demonstrating differential binding of by neurodevelopmental stage. These Integrated Genomics Viewer image shows the position of this genomic location on the chromosome by a *vertical red bar*, as well as the genetic coordinates for SATB2 ChIPseq experiment. There are five windows within this image and each one represents sequencing reads (*thick gray bars*) from independent ChIPseq experiments from both NSC lines at two neurodevelopmental stages and adult human brain. The *blue line* at the bottom represents the gene, which is also named in italics. **a**, **b** SATB2; **c**, **d** EHMT1; **e**, **f** TCF7L2 (color figure online)

In our 'positive control' ChIPseq targeting *TCF7L2* protein, we observed a similar pattern: there is binding of TCF7L2 in *PPP2R2B* apparently in dimer in all cell lines but not one read in brain; conversely, the neurotransmitter transporter *SLC6A13* shows read pile up in brain but not a single read from any cell line. These raw data examples argue for a much more dynamic genomic association for at least these proteins. Temporal and spatial patterns of DNA association may be extremely dynamic for all transcription factors, dependent on developmental stage or cell context. DNA association maps of individual proteins such as ENCODE may be specific to the context of the experiments in which they were conducted, where developmental stage is one of many possible contexts. Future mapping studies might consider developmental DNA association patterns of proteins.

Replication of ChIP with qPCR confirmation of select genomic targets

While we are confident in our ChIP calling parameters and the fact that we have excellent overlap of 2/3 genes across

4 independent cell types, we validate a selection of ChIP peaks using DNA squired from freshly grown cells and a new ChIP experiment using the same antibodies. To perform these experiments, we selected peaks called across all cell lines for each of SATB2, EHTM1, and SATB2 ChIPseq and performed DNA-based targeted qPCR. We used two negative control regions not found in the ChIPseq data to normalize qPCR (Fig. 5). For each gene, we selected targets that showed high read coverage across all 4 cell lines, and that were the most statistically significant compared to background levels as determined by MACS.

For EHMT1, we attempted to validate ChIP data using genomic loci in *GFAP*, a marker of astrocytes and radial glial cells, and *MALAT1* across all 4 cell lines (Fig. 5c, d). Here, we could observe increased DNA in 3 of 4 cell lines for GFAP and 1 of 4 cell lines for MALAT1, with no signal above baseline in the other cell lines. This suggests to us that EHMT1 may bind very dynamically to DNA, as suggested by our ChIPseq data that did not replicate across experiments in an identical cell line, yet the validation in some cell lines of these targets suggests that EHMT1 is
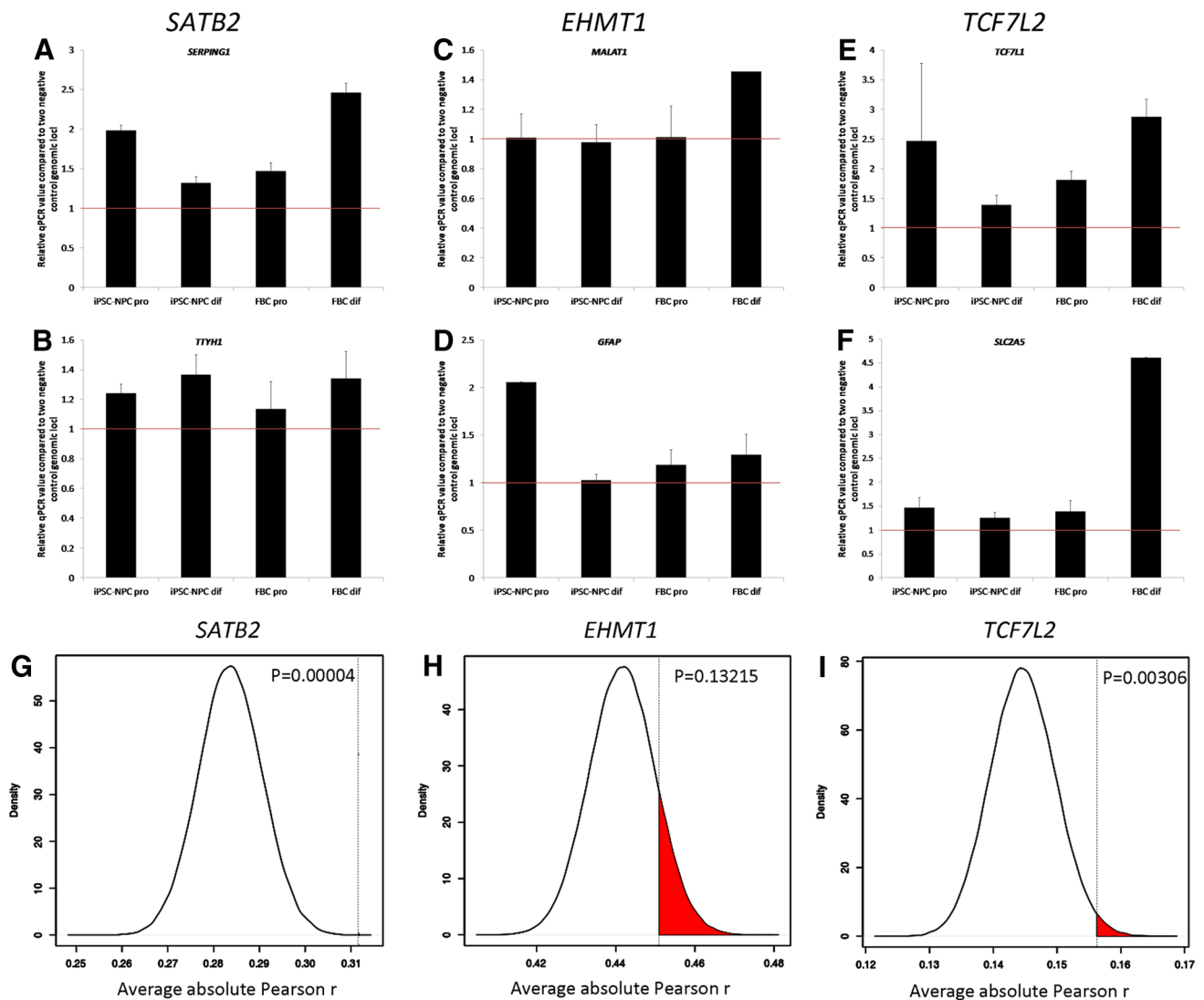
Fig. 5 Validation of ChIP peaks in a new ChIP experiment for each gene and association of ChIP peaks with brain gene expression. **a–f** Replication of ChIP peaks in two targeted genomic loci associated with genes. *Bar graphs* are specific to each cell line and the height above the *red line* represents the enrichment above negative control loci. **g–i** ChIPseq peaks are enriched for genes that correlate with transcription factor expression in BrainSpan. Each chip target was correlated to all the genes in the genome using the BrainSpan developmental gene expression database for the VFC, IPC, HIP, DTH, and MD. Genes that were bound by SATB2 and TCF7L2 were significantly more correlated than expected by chance (color figure online)

not just random noise. More work will need to be done to understand the dynamics of EHMT1 binding to DNA.

For TCF7L2, we could validate both of the selected targets in all 4 cell lines (Fig. 5e, f). Here, we selected genomic loci in *TCF7L1* and the glucose transported *SLC2A5*, in line with TCF7L2's action in diabetes and in a feedback loop with other family members in the WNT signaling pathway. In the case of TCF7L2, a recent study performed ChIPseq targeting this protein in several cell lines but not brain cells (Frietze et al. 2012). To compare how ChIPseq peaks in our study compare to ChIPseq peaks

in different non-neuronal cell lines, we intersected the consensus peak file from TCF7L2 peaks (common across all four ChIPseq experiments) with each cell lines from the Frietze et al. study. First, we asked how many peaks in the TCF7L2 file were seen in any of 6 lines used by Frietze et al., and found that 62 % of peaks we detected were observed in at least one cell line. Next, we used the Szymkiewicz–Simpson coefficient to assess the similarity between each dataset (where a coefficient of 1 represents perfect overlap. HCT116: 0.084; HEK293: 0.028; HeLa: 0.075; HepG2:0.034; MCF7: 0.024; Panc1: 0.064).

Correlation of expression of SATB2 and TCF7L2 genes with their binding targets in BrainSpan data

To further show that ChIP targets are valid, we asked whether there was a correlation of gene expression in brain between each of SATB2, EHMT2, and SATB2 and their target genes. To do this, we used data from BrainSpan at all timepoints from that study and using 5 brain regions. We generated a Pearson R value for each data point using absolute value (so as not to assume negative or positive regulation) and then asked what the probability of observing a correlation was between the DNA association gene (i.e., SATB2, EHMT1, and TCF7L2) and their ChIPseq-identified targets. The R values for all other genes was used in a permutation test to assess whether these was an enrichment of high R values in the target genes. In Fig. 5g–i, we show this graphically, where we find a strongly significant effect for SATB2 and TCF7L2, but not for EHMT1.

DNA consensus motifs in cell lines and adult brain

Proteins that bind or associate with DNA are often associated with particular consensus motifs. Given that SATB2 and TCF7L2 proteins bind DNA associated with distinct loci in adult brain and NSCs (EHMT1 association is unclear), we next wanted to test whether these proteins had distinct DNA consensus motifs at different stages, or if consensus motifs were the same across developmental stages. Thus, there are two hypotheses being explored to explain the lack of overlap between brain tissue and NSCs: (1) these proteins have multiple targets with the same consensus motif in the human genome but only access some of them some of the time (depending on developmental stage for example), or (2) that the proteins associate with DNA at different consensus sites—a molecular explanation for this hypothesis may be that these proteins might associate with different protein complexes in adult brain compared to NSCs, guiding them to genomic regions with different consensus motifs.

To test these diverging hypotheses, we used the high confidence peaks from NSCs, and all peaks from brain. We include EHMT1 data from iPSC-NSCs (where there was good replication between proliferating and differentiating cells) and adult brain, but consider the EHMT1 analysis lower confidence due to the wide variation in peak calls between samples. We used the HOMER algorithm to determine de novo consensus binding motifs that were enriched across each group. Figure 6 shows the five most significant de novo consensus motifs for each analysis as predicted by the HOMER algorithm. For SATB2, the two most significant consensus motifs were identical between the peaks called from all four cell lines and adult brain, specifically the CAATATGG motif. This suggests that the SATB2 consensus binding motif is likely the same in adult human brain as in NSCs, but that

these may be at different genomic locations, since there is almost no overlap between brain and NSCs (Fig. 3e). Notably, this site also includes a central core 'AATAT' domain, in line with SATB2's function in associating with AT-rich domains implied by its name ('special AT-rich sequence-binding protein'). For TCF7L2 protein, we found that the GATCGGGTGT motif was the most significant enriched across common peaks from the NSC experiment, and that 'GATC' occurred in 4/5 of these top five motifs. This is notably because a recent paper (Frietze et al. 2012) identified CTTTGATC as the major 8-mer motif for TCF7L2 across multiple non-neuronal cell lines. This suggests that 'GATC' might form part of a core binding motif of TCF7L2 and that flanking regions of a core motif might provide cell type or developmental stage specificity. In adult brain, we identified TGATCC in the fourth ranked motif suggesting that TCF7L2 likely binds to this GATC core in brain as well as NSCs.

We observed some similarities between brain and iPSC-NSCs from the EHMT1 protein as well, despite the variation in peaks form this ChIP and its function as a DNA association protein rather than a DNA binding protein. The motif 'CCCA' occurred in two of the five highest ranked (most significant) motifs from brain (Rank 1 and 4), and occurred in the third most significant rank from NSCs. Three of five most significant NSCs motifs included 'CAAA', while 'AAAC' was present in the fifth ranked brain motif and the fifth ranked NSC motif. Together these results support the first hypothesis that these three protein associate with similar DNA consensus motifs in NSCs and adult brain, but at different genomic loci.

Enrichment of ChIP peaks at genomic regulatory elements

Using data available on the UCSC genome browser, we next asked whether ChIP peaks for each of SATB2, EHMT1, or TCF7L2 were associated with transcription start sites, binding sites of other transcription factors, or specific histone marks associated with gene expression. Figure 7 shows density plots of genomic features found within 10 Kb up or downstream of all consensus ChIP peaks for each gene. The plot represents the likelihood of observing a particular genomic feature in or within 10 Kb of a peak, where the total area under the curve corresponds to all possible features. Across each gene, we observe an increased likelihood of observing transcription factor binding sites directly in the peak rather than upstream or downstream.

Potential function of SATB2, EHMT1, and TCF7L2 reflects neurodevelopment in NSCs and neuron function in adult brain

If DNA association regions are different for these proteins in NSCs compared to adult brain, what might this

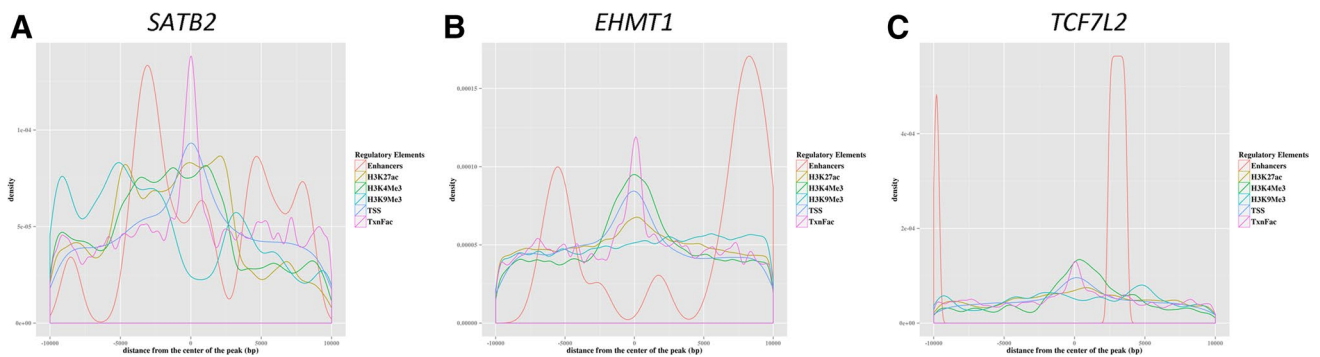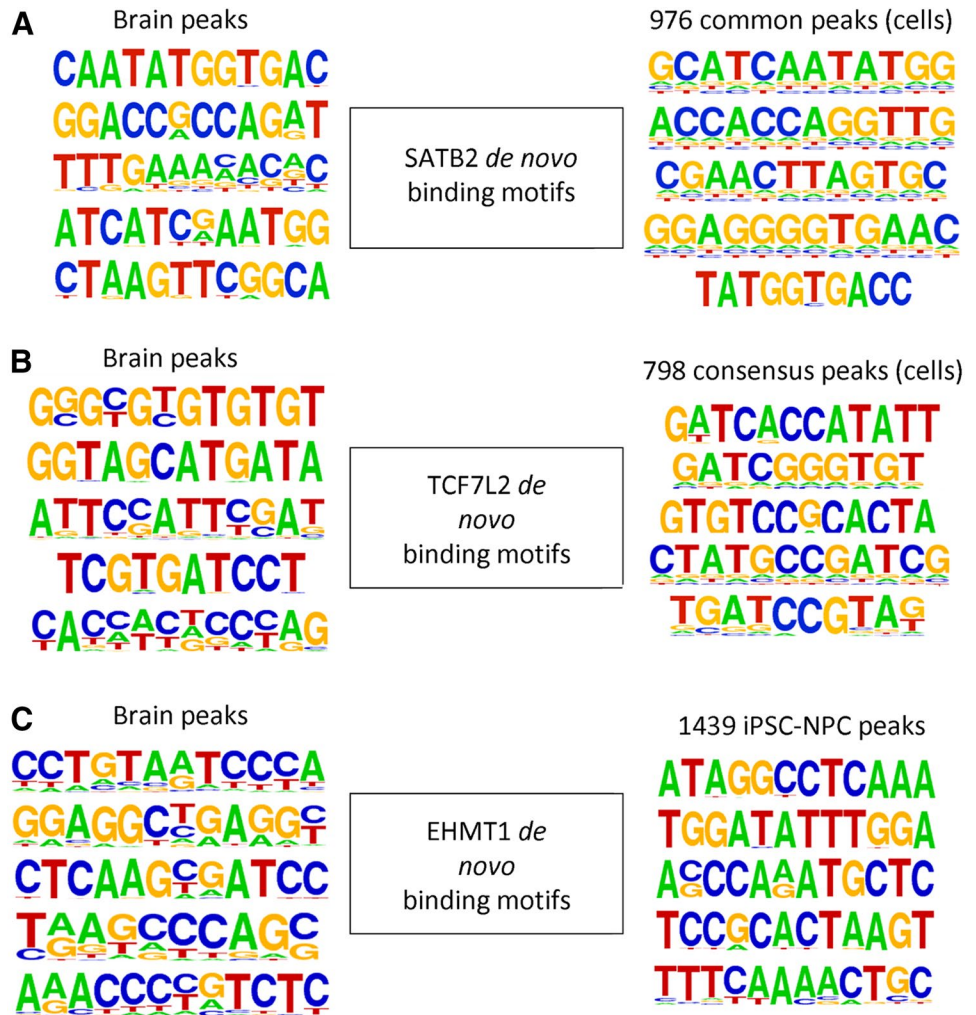**Fig. 6** De novo consensus sites most significant in ChIP peaks for each of SATB2. EHMT1, and TCF7L2

Fig. 7 The relationship between ChIP peaks and transcription start sites. Plots represent all peaks from consensus reads and the likelihood of observing a given feature at any distance from the peak, where the area beneath the curve corresponds to density

differential binding be doing? To understand this, we mapped ChIP peaks to genomic features and performed a Gene Ontology (GO) analysis of those peaks that mapped near genes (5 Kb upstream and 2 Kb downstream). For SATB2 (Fig. 8), we found that brain peaks were mostly

concentrated in genes (>57 %), while common peaks across NSCs showed a similar pattern, but with a higher percentage of peaks found in genes and promoter/5′UTR regions (>65 %). Performing a GO analysis using genes that map to these peaks, we found that GO terms specific
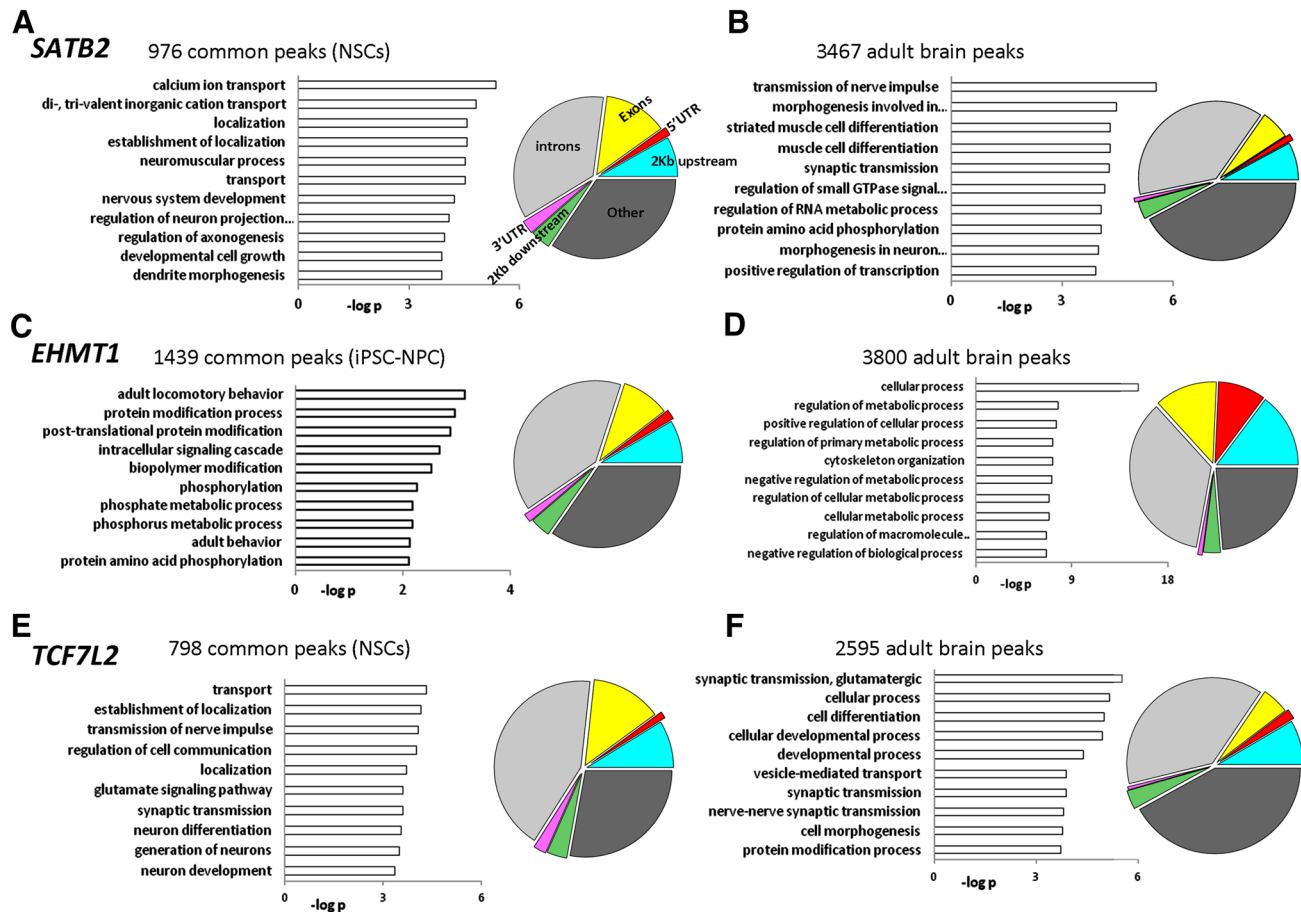
Fig. 8 TF binding loci reflect neurodevelopment in NSCs and neuron function in adult brain. **a** Genomic distribution for those peaks from the SATB2 ChIP experiment common to all four independent NSC ChIPseq experiments and a Gene Ontology analysis showing the most significant GO terms for peaks found near or within genes. **b** Genomic distribution for those peaks from the SATB2 ChIP experiment common to the adult brain ChIPseq experiment and a Gene Ontology analysis showing the most significant GO terms for peaks found near or within genes. **c** Genomic distribution for those peaks from the EHMT1 ChIP experiment common to all four independent NSC ChIPseq experiments and a Gene Ontology analysis showing the most significant GO terms for peaks found near or within genes. **d**

Genomic distribution for those peaks from the EHMT1 ChIP experiment common to the adult brain ChIPseq experiment and a Gene Ontology analysis showing the most significant GO terms for peaks found near or within genes. **e** Genomic distribution for those peaks from the TCF7L2 ChIP experiment common to all four independent NSC ChIPseq experiments and a Gene Ontology analysis showing the most significant GO terms for peaks found near or within genes. **f** Genomic distribution for those peaks from the TCF7L2 ChIP experiment common to the adult brain ChIPseq experiment and a Gene Ontology analysis showing the most significant GO terms for peaks found near or within genes

to NSC consensus peaks were related to neurodevelopment and peaks in human brain were associated with nerve cell transmission. For EHMT1, we used the all-common peaks from the iPSC-NSC experiments (i.e., proliferating and differentiating cells, 1439 common peaks). In these cell lines, we identified GO terms associated with protein modification and metabolic processes, while in adult human brain EHMT1 peaks were also associated with genes involved in metabolic processes. Performing a GO analysis for TCF7L2 peaks, we find a similar pattern to SATB2; that is, TCF7L2 preferentially associates with genes, but that a higher percentage of peaks in NSCs associate with genes and 5′UTR/promoter regions (>72 %) than in adult brain

(>58 %). For the TCF7L2 gene ontology analysis, we find more neurodevelopment terms in NSCs, and more terms related to 'neuron' or 'synapse' in brain, suggesting that TCF7L2 has a role in neurodevelopment in NSCs and in neuron function in adult brain. We also observed the presence of 'synaptic transmission, glutamatergic' as the most significant GO term in adult brain for TCF7L2 peaks, and 'glutamate signaling pathway' as the fifth most significant GO term in NSCs. While GO categories can be overly general, the specificity of these glutamate receptor categories in both NSCs and adult brain is notable. In particular, it suggests a role for TCF7L2 in the regulation of glutamate receptor subunits in NSCs and adult brain, which to

our knowledge has not been reported to be a function of TCF7L2. Together these gene ontology data support a model where DNA associating factors target genes important in neurodevelopment in NSCs and in synapse maintenance or neuronal function in adult brain.

Discussion

The purpose of the current study was to investigate expression and DNA association of genes important in neurodevelopment in adult human brain. The motivation for this study was the lack of therapeutic targets for neurodevelopmental disorders—a lack of hope has percolated the NDD research community where these disorders are seen as untreatable because of the late identification of genetic mutations in the clinic and the 'fixing' of neuronal projections in neurodevelopment (i.e., disease is wired in and cannot be altered after the fact). We highlight here that genes associated with neurodevelopmental disorders continue to be expressed across defined neural stem cell stages and in adult brain and, at least for *SATB2* and *EHMT1*, and many genes well known to be associated with NDDs continue to have a role in adult, post-mitotic neurons. We suggest that clinical manifestation of NDDs may be a result of not just the dysfunctional activity of specific genes in development but also the dysfunction of the same genes in the adult brain. Targeting the adult function of genes might be a reasonable approach to treating NDDs.

This study sought to highlight that genes implicated in NDD may have a role adult brain. To do this, we chose two genes we have worked with previously, *SATB2* and *EHMT1*, and included a gene not previously associated with NDDs but which is implicated in WNT signaling for our experiments. We surely could have chosen different or more genes for these experiments; however, we argue that our approach was reasonable and achieved our goal of highlighting adult brain function. We think that expression levels are a proxy for function, but more studies are needed on the explicit role of each of these genes in adult brain, particularly if the aim of treatment is to be pursued.

There are several technical details of these experiments that warrant discussion. We used two very different neural stem cell lines and found a high degree of overlap (>40 %) of DNA binding regions across cell lines and, and a high degree of overlap (43–67 %) across NSC developmental stages within each cell line. This was true only for SATB2 and TCF7L2 proteins, both of which directly associate with DNA, and likely suggests that this same pattern may be true for other transcription factors expressed at both neurodevelopmental stages. For EHMT1—a histone methyltransferase that associates with DNA but does not directly bind DNA—we saw little overlap across any neurodevelopmental stage.

DNA binding regions for all proteins in adult brain are almost completely different from NSC lines, though the same DNA consensus sites were associated with each protein. DNA association may be a function of both the presence of an appropriate DNA consensus motif and at a specific developmental stage. The dynamics of protein action are not well understood, though it has become clear from projects like ENCODE (Zhang et al. 2012) that the same protein can associate with different genomic regions in different cell types (e.g., liver vs kidney), and our work further highlights this within the nervous system at different developmental stages.

EHMT1 DNA binding patterns appear to be cell type and cell stage specific. EHMT1 does not directly bind DNA, but has an important function in human neurodevelopmental given the severe disorder caused by EHMT1 haploinsufficiency (Talkowski et al. 2012). How then does EHMT1 specify its genomic targets? We found almost no overlap of EHMT1 across proliferating and differentiating NSCs in the FBC line, mostly because there were so many peaks observed at the proliferating time point, and our replication experiment suggested that this was not a technical artifact. In the iPSC-NPC line, we saw a high degree of overlap (>40 %) of ChIP peaks in proliferating and differentiating cells, possibly suggesting that EHMT1 does indeed target specific genomic locations when cells differentiate down a linear cell program, since only the iPSC-NPC line was able to form electrophysiologically defined neurons. Perhaps the heterogeneity of FBCs is so high, and EHMT1 function so ubiquitous, that no overlap could be observed in the FBCs.

We used an arbitrary (but functionally defined) definition of cell stage and highly heterogeneous samples. It may be that the patterns that we observed are specific to the limited windows we defined, so a future analysis might look at more comprehensive temporal windows; for example at day-60 and day-90 post-differentiation stages, rather than just the day-30 stage used here. We used two very different neural stem cell lines and brain tissue from frontal cortex. While we see this as a strength (given the high degree of overlap across NSCs at the same developmental stage), it may be that differential binding patterns by developmental stage are different than that observed here if one were to use the same tissue across all developmental stages; however, this is technically challenging, if not impossible, should one want to examine for example, the earliest neurodevelopmental stage and the late adult stage.

We used DNA binding patterns as a proxy for potential function of each DNA association protein, thus we cannot say precisely what association with DNA might mean. This decision though was re-enforced by the results from the GO analysis, whereby the actual genes where binding sites were observed were remarkably close to what one

might expect from each DNA association protein. We did consider performing DNA association protein knockdown experiments to correlate with ChIP binding peaks; however, recent studies suggest that the relationship between DNA association and effects on transcription are not necessarily direct (Cheng et al. 2012; Jaeger and Manu 2012), further limiting interpretation of these results. Future work with respect to these genes and disease might determine which genomic targets lead to a direct effect on expression of nearby genes and at what time period.

Currently, neurodevelopmental disorders are considered disorders of genomic architecture, whereby clinical features of disease are caused by errors that occur in utero or in perinatal brain development periods (Geschwind and Levitt 2007; Tebbenkamp et al. 2014). Our work suggests a broader view of some genes implicated in autism or other neurodevelopmental disorders is warranted; these genes undeniably have important and disease related functions in early developmental periods; however, they appear also to have a role in adult human brain, at least when using expression and DNA association as an output measure. We suggest that diseases associated with mutations in some of these genes might result from altered action in the adult period and not from action in early neurodevelopment. This contributes to the idea that treatment of disorders traditionally considered neurodevelopmental (e.g., autism) might have therapeutic value (Castren et al. 2012; Ehninger et al. 2008), where targeting the adult function of mutated genes that cause disease might improve clinical outcomes.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

References

Alcamo EA, Chirivella L, Dautzenberg M, Dobreva G, Farinas I, Grosschedl R, McConnell SK (2008) Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. Neuron 57:364–377

Aradi I, Santhakumar V, Soltesz I (2004) Impact of heterogeneous perisomatic IPSC populations on pyramidal cell firing rates. J Neurophysiol 91:2849–2858

Britanova O, de Juan Romero C, Cheung A, Kwan KY, Schwark M, Gyorgy A, Vogel T, Akopov S, Mitkovski M, Agoston D, Sestan N, Molnar Z, Tarabykin V (2008) Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex. Neuron 57:378–392

Castren E, Elgersma Y, Maffei L, Hagerman R (2012) Treatment of neurodevelopmental disorders in adulthood. J Neurosci 32:14074–14079

Chen B, Wang SS, Hattox AM, Rayburn H, Nelson SB, McConnell SK (2008) The Fezf2-Ctip2 genetic pathway regulates the fate choice of subcortical projection neurons in the developing cerebral cortex. Proc Natl Acad Sci USA 105:11382–11387

Cheng C, Alexander R, Min R, Leng J, Yip KY, Rozowsky J, Yan KK, Dong X, Djebali S, Ruan Y, Davis CA, Carninci P, Lassman T, Gingeras TR, Guigo R, Birney E, Weng Z, Snyder M, Gerstein M (2012) Understanding transcriptional regulation by integrative analysis of transcription factor binding data. Genome Res 22:1658–1667

Docker D, Schubach M, Menzel M, Munz M, Spaich C, Biskup S, Bartholdi D (2014) Further delineation of the SATB2 phenotype. Eur J Hum Genet 22:1034–1039

Dumais A, Lesage AD, Boyer R, Lalovic A, Chawky N, Menard-Buteau C, Kim C, Turecki G (2005) Psychiatric risk factors for motor vehicle fatalities in young men. Can J Psychiatry 50:838–844

Ehninger D, Li W, Fox K, Stryker MP, Silva AJ (2008) Reversing neurodevelopmental disorders in adults. Neuron 60:950–960

Florez JC, Jablonski KA, Bayley N, Pollin TI, de Bakker PI, Shuldiner AR, Knowler WC, Nathan DM, Altshuler D, Diabetes Prevention Program Research G (2006) TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. N Engl J Med 355:241–250

Frietze S, Wang R, Yao L, Tak YG, Ye Z, Gaddis M, Witt H, Farnham PJ, Jin VX (2012) Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. Genome Biol 13:R52

Gao XF, Chen W, Kong XP, Xu AM, Wang ZG, Sweeney G, Wu D (2009) Enhanced susceptibility of Cpt1c knockout mice to glucose intolerance induced by a high-fat diet involves elevated hepatic gluconeogenesis and decreased skeletal muscle glucose uptake. Diabetologia 52:912–920

Geschwind DH, Levitt P (2007) Autism spectrum disorders: developmental disconnection syndromes. Curr Opin Neurobiol 17:103–111

Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 38:320–323

Grosschedl R, Giese K, Pagel J (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. Trends Genet 10:94–100

Jaeger J, Manu Reinitz J (2012) Drosophila blastoderm patterning. Curr Opin Genet Dev 22:533–541

Kalkman HO (2012) A review of the evidence for the canonical Wnt pathway in autism spectrum disorders. Mol Autism 3:10

Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AM, Pletikos M, Meyer KA, Sedmak G, Guennel T, Shin Y, Johnson MB, Krsnik Z, Mayer S, Fertuzinhos S, Umlauf S, Lisgo SN, Vortmeyer A, Weinberger DR, Mane S, Hyde TM, Huttner A, Reimers M, Kleinman JE, Sestan N (2011) Spatio-temporal transcriptome of the human brain. Nature 478:483–489

Kleefstra T, Brunner HG, Amiel J, Oudakker AR, Nillesen WM, Magee A, Genevieve D, Cormier-Daire V, van Esch H, Fryns JP, Hamel BC, Sistermans EA, de Vries BB, van Bokhoven H (2006) Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. Am J Hum Genet 79:370–377

Kleefstra T, Kramer JM, Neveling K, Willemsen MH, Koemans TS, Vissers LE, Wissink-Lindhout W, Fenckova M, van den Akker WM, Kasri NN, Nillesen WM, Prescott T, Clark RD, Devriendt K, van Reeuwijk J, de Brouwer AP, Gilissen C, Zhou H, Brunner HG, Veltman JA, Schenck A, van Bokhoven H (2012) Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. Am J Hum Genet 91:73–82

Krumm N, O'Roak BJ, Shendure J, Eichler EE (2014) A de novo convergence of autism genetics and molecular neuroscience. Trends Neurosci 37:95–105

Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, Chen Y, DeSalvo G, Epstein C, Fisher-Aylor KI, Euskirchen G, Gerstein M, Gertz J, Hartemink AJ, Hoffman MM, Iyer VR, Jung YL, Karmakar S, Kellis M, Kharchenko PV, Li Q, Liu T, Liu XS, Ma L, Milosavljevic A, Myers RM, Park PJ, Pazin MJ, Perry MD, Raha D, Reddy TE, Rozowsky J, Shoresh N, Sidow A, Slattery M, Stamatoyannopoulos JA, Tolstorukov MY, White KP, Xi S, Farnham PJ, Lieb JD, Wold BJ, Snyder M (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res 22:1813–1831

Link PA, Gangisetty O, James SR, Woloszynska-Read A, Tachibana M, Shinkai Y, Karpf AR (2009) Distinct roles for histone methyltransferases G9a and GLP in cancer germ-line antigen gene regulation in human cancer cells and murine embryonic stem cells. Mol Cancer Res 7:851–862

Matevossian A, Akbarian S (2008) A chromatin assay for human brain tissue. J Vis Exp. doi:10.3791/717

Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ, Johnson R, Segre AV, Djebali S, Niarchou A, Consortium GT, Wright FA, Lappalainen T, Calvo M, Getz G, Dermitzakis ET, Ardlie KG, Guigo R (2015) Human genomics. The human transcriptome across tissues and individuals. Science 348:660–665

Mitchell RK, Mondragon A, Chen L, McGinty JA, French PM, Ferrer J, Thorens B, Hodson DJ, Rutter GA, Da Silva Xavier G (2015) Selective disruption of Tcf7l2 in the pancreatic beta cell impairs secretory function and lowers beta cell mass. Hum Mol Genet 24:1390–1399

Roose J, Clevers H (1999) TCF transcription factors: molecular switches in carcinogenesis. Biochim Biophys Acta 1424:M23–M37

Rosenfeld JA, Ballif BC, Lucas A, Spence EJ, Powell C, Aylsworth AS, Torchia BA, Shaffer LG (2009) Small deletions of SATB2 cause some of the clinical features of the 2q33.1 microdeletion syndrome. PLoS One 4:e6568

Szemes M, Gyorgy A, Paweletz C, Dobi A, Agoston DV (2006) Isolation and characterization of SATB2, a novel AT-rich DNA binding protein expressed in development- and cell-specific manner in the rat brain. Neurochem Res 31:237–246

Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev 19:815–826

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872

Talkowski ME, Rosenfeld JA, Blumenthal I, Pillalamarri V, Chiang C, Heilbut A, Ernst C, Hanscom C, Rossin E, Lindgren AM, Pereira S, Ruderfer D, Kirby A, Ripke S, Harris DJ, Lee JH, Ha K, Kim HG, Solomon BD, Gropman AL, Lucente D, Sims K, Ohsumi TK, Borowsky ML, Loranger S, Quade B, Lage K, Miles J, Wu BL, Shen Y, Neale B, Shaffer LG, Daly MJ, Morton CC, Gusella JF (2012) Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. Cell 149:525–537

Tebbenkamp AT, Willsey AJ, State MW, Sestan N (2014) The developmental transcriptome of the human brain: implications for neurodevelopmental disorders. Curr Opin Neurol 27:149–156

Valenta T, Lukas J, Korinek V (2003) HMG box transcription factor TCF-4′s interaction with CtBP1 controls the expression of the Wnt target Axin2/Conductin in human embryonic kidney cells. Nucleic Acids Res 31:2369–2380

Wolfgang MJ, Kurama T, Dai Y, Suwa A, Asaumi M, Matsumoto S, Cha SH, Shimokawa T, Lane MD (2006) The brain-specific carnitine palmitoyltransferase-1c regulates energy homeostasis. Proc Natl Acad Sci USA 103:7282–7287

Zarate YA, Perry H, Ben-Omran T, Sellars EA, Stein Q, Almureikhi M, Simmons K, Klein O, Fish J, Feingold M, Douglas J, Kruer MC, Si Y, Mao R, McKnight D, Gibellini F, Retterer K, Slavotinek A (2015) Further supporting evidence for the SATB2-associated syndrome found through whole exome sequencing. Am J Med Genet A 167:1026–1032

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

Zhang J, Zhang H, Liu J, Tu X, Zang Y, Zhu J, Chen J, Dong L, Zhang J (2012) miR-30 inhibits TGF-beta1-induced epithelial-to-mesenchymal transition in hepatocyte by targeting Snail1. Biochem Biophys Res Commun 417:1100–1105

Zhong J, Carrozza DP, Williams K, Pritchett DB, Molinoff PB (1995) Expression of mRNAs encoding subunits of the NMDA receptor in developing rat brain. J Neurochem 64:531–539