Transcriptional suppression from KMT2D loss disrupts cell cycle and hypoxic responses in neurodevelopmental models of Kabuki syndrome

Giovanni A. Carosso[1,2], Leandros Boukas[1,2,3], Jonathan J. Augustin[2,4,6], Ha Nam Nguyen[5], Briana L. Winer[2], Gabrielle H. Cannon[2], Johanna D. Robertson[2], Li Zhang[2], Kasper D. Hansen[2,3], Loyal A. Goff[2,6], Hans T. Bjornsson[2,7,8,9]*

[1]Predoctoral Training Program in Human Genetics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[2]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[3]Department of Biostatistics, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[4]Predoctoral Training Program in Biochemistry, Cellular, and Molecular Biology, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[5]Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[6]Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[7]Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[8]Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik 101, Iceland.
[9]Landspitali University Hospital, Reykjavik, 101, Iceland.
*Correspondence: hbjorns1@jhmi.edu

Abstract

Chromatin modifiers act to coordinate gene expression changes critical to neuronal differentiation from neural progenitor cells (NPCs). Histone methyltransferase KMT2D promotes transcriptional activation and its loss causes the intellectual disability disorder Kabuki syndrome 1 (KS1) in ways that are not well understood. We conducted parallel studies of proliferation, differentiation, and transcription and chromatin profiling in KMT2D-deficient human and mouse models of brain development to define KMT2D-dependent functions in neurodevelopmental contexts, including adult-born hippocampal NPCs *in vivo*. We report cell-autonomous proliferation, cell cycle progression, and survival deficits accompanied by transcriptional perturbance of oxygen-sensing, central carbon metabolism, and mRNA/protein metabolism pathways, with functional experiments demonstrating disruption of KMT2D-dependent hypoxia responses and proteasome functions. Importantly, signatures of precocious neuronal differentiation accompanied defects in NPC maintenance. Together, our findings support a model in which KMT2D loss causes transcriptional suppression of metabolic programs for NPC maintenance, resulting in exhaustion of precursors required for adult hippocampal neurogenesis.

Main

Trithorax group proteins promote chromatin accessibility by exerting antagonistic functions against Polycomb group transcriptional suppressors to favor activated gene expression states[1]. Fine-tuning of cell type transitions during neuronal development depends critically on this duality, as evidenced

40 by severe neurodevelopmental defects caused by variants in chromatin-modifying genes[2]. Loss-of-
41 function variants in genes encoding two such enzymes, lysine-specific methyltransferase 2D (KMT2D)
42 and X-linked lysine-specific demethylase 6A (KDM6A) cause the intellectual disability (ID) disorder
43 Kabuki syndrome (KS1 and KS2, respectively)[3,4]. Up to 74%[5] of cases result from mutations in *KMT2D*
44 (KS1), encoding a major histone H3 lysine 4 (H3K4) methyltransferase which catalyzes chromatin-
45 opening modifications that promote gene transcription at context-specific targets. Developmental
46 requirements of KMT2D in cardiac precursors[6], B cells[7,8], muscle and adipose[9], and epithelial tissues[10]
47 have been linked, respectively, to *KMT2D*-associated cardiac, immunologic, and oncogenic contexts[11],
48 yet studies into the effects of KMT2D deficiency in neurodevelopment have been limited.

49 We previously described a mouse model, *Kmt2d*$^{+/\beta geo}$, demonstrating characteristic features seen
50 in KS1 individuals including postnatal growth restriction, craniofacial abnormalities, and visuospatial
51 memory impairments associating with a paucity of adult-born neuroblasts in the hippocampal dentate
52 gyrus (DG)[12]. Moreover, decreased grey matter volume in the human DG was recently described in
53 human KS1 patients, further implicating this region in KS1-associated ID phenotypes[13]. The continual
54 birth and integration of new neurons makes adult neurogenesis the most potent form of lifelong plasticity
55 in the mammalian brain[14]. During late stages of embryonic development, a subset of multipotent neural
56 progenitor cells (NPCs) persists in the DG, remaining subject to an array of intrinsic and extrinsic factors
57 affecting NPC maintenance, i.e. self-renewal, proliferation, and differentiation, throughout adult life, and
58 mounting evidence tightly links metabolic rewiring[15] and hypoxic states[16] to NPC maintenance.
59 Understanding the role of KMT2D in these processes will be critical to development of postnatal
60 therapeutic interventions in KS1-associated ID. Furthermore, robust models of neurodevelopment with
61 consistent KS1 phenotypes are needed to appreciate the target gene- and cell type-specific impact of
62 KMT2D loss on brain development. Such studies have the potential to identify primary molecular
63 consequences of KMT2D loss and facilitate targeted disease-modifying therapeutic approaches.

64 Here, we report that KMT2D loss disrupts proliferation, cell cycle progression, and cell survival in
65 a manner consistent across mouse and human KS1 neurodevelopment models. Using single-cell
66 transcriptome analysis in patient iPSC-derived models, we identify preferential perturbance in early,
67 cycling NPCs and signatures of precocious maturation in differentiating NPCs. We then compare
68 transcriptional and KMT2D binding profiles to prioritize KS1-associated target genes in mouse
69 hippocampal cells specifically lacking KMT2D catalytic activity, identifying a perturbance of pathways
70 involved in oxygen-sensing, central carbon metabolism, and mRNA/protein degradation, which prompted
71 functional experiments demonstrating KMT2D-dependent cellular responses to hypoxia and proteasome
72 function. Finally, we explore relevancy of these findings *in vivo*, identifying depletion of hippocampal

NPCs accompanied by an *in vivo* cell cycle defect and overlapping transcriptional signatures in *Kmt2d*-deficient adult mice.

Results

KS1 patient-derived cells exhibit defects in proliferation, cell cycle progression, and survival

We reprogrammed skin biopsy fibroblasts to generate induced pluripotent stem cells (iPSCs) from a female with molecularly confirmed KS1 bearing a heterozygous *KMT2D* mutation (c.7903C>T:p.R2635*), and characteristic KS1 features including growth retardation, facial features, and congenital heart disease. Neuropsychological testing in a previous longitudinal study[17] of this patient had revealed a visuospatial memory impairment similar to that described in a retrospective study of KS1 patients[12]. KS1 iPSCs (KS1-1) bearing normal 46, XX karyotype (**Supplementary Fig. 1a**) and characteristic iPSC morphology (**Fig. 1a**) were selected for comparison against two independently generated, previously described iPSC lines (C1-2, C3-1) from unrelated healthy controls[18]. KS1 iPSCs were predicted to be *KMT2D* haploinsufficient due to nonsense-mediated decay (NMD), and RT-qPCR confirmed reduced *KMT2D* mRNA with equal expression of exons upstream and downstream of the mutation site (**Supplementary Fig. 1b, c**). Flow cytometric analyses following pulse of cell division marker 5-ethynyl-2′-deoxyuridine (EdU) demonstrated lower proliferation rates (~25%) in KS1 iPSCs compared to controls (**Fig. 1b**), accompanied by a shift in cell cycle occupancy (**Fig. 1c, Supplementary Fig. 1d**) toward overrepresentation of G_2/M phase cells (24% higher in KS1) and underrepresentation of G_0/G_1 phase cells (23% lower in KS1).

We next generated NESTIN-expressing NPCs through parallel differentiation of KS1 and control iPSCs using an established small molecule inhibition protocol[19]. RT-qPCR confirmed *KMT2D* deficiency in KS1 NPCs (**Supplementary Fig. 1e**), and both KS1 and control NPCs displayed normal NPC morphology (**Fig. 1d, Supplementary Fig. 1f**). Immunostaining against histone marks in KS1 NPCs detected significantly reduced fluorescence of open chromatin mark H3K4me2 (~65%) and increased fluorescence of closed chromatin marks H3K27me3 (~99%) and H3K9me3 (~180%) compared to controls (**Supplementary Fig. 1g-i**). Compared to iPSC data, KS1 NPCs relative to control showed a markedly greater reduction in dividing cells as detected by EdU incorporation (~47%, **Fig. 1e**) as well as fewer mitotic divisions over a 72-hour period as detected by CellTrace Violet (**Supplementary Fig. 1j**). Unlike in iPSCs, cell cycle abnormalities were not observed in KS1 NPCs (**Fig. 1f, Supplementary Fig. 1k**) suggesting this phenotype may be cell-type dependent or lost during induced differentiation. However, flow cytometric analysis revealed higher proportions of cell death in KS1 cells compared to controls among both iPSCs (~130%) and NPCs (~115%) (**Fig. 1g, Supplementary Fig. 1l**).

106 Finally, to determine whether the cell cycle G_2/M bias in KS1 iPSCs occurs in additional KS1
107 patients and primary cells, we performed cell cycle analysis in primary fibroblasts from two additional
108 molecularly confirmed KS1 patients (KS1-2, KS1-3) in addition to KS1-1 and two additional healthy
109 controls. KS1 and control fibroblasts were treated with nocodazole to synchronize in G_2/M phase, and
110 flow cytometric DNA content analysis at 0- and 3-hours post-release demonstrated exit from G_2/M phase
111 among control fibroblasts, but this was markedly decreased in fibroblasts of KS1 patients (**Fig. 1h-i**).
112 Thus, delayed exit from G_2/M was consistent in primary cells from three independent KS1 patients.
113
114 **Deletion of the KMT2D SET methyltransferase domain in isogenic hippocampal cells**
115 **recapitulates KS1-associated defects in proliferation, cell cycle, and survival**
116 We selected the mouse hippocampal HT22 cell line for detailed analysis of KMT2D catalytic
117 function in neuronal context[20]. gDNA sequence encoding the SET methyltransferase domain was deleted
118 using a CRISPR-Cas9 strategy combining an upstream small guide RNA (sgRNA-1, in exon 52) with
119 either sgRNA-2 (exon 54) or sgRNA-3 (intron 54) to create deletions of 570 or 660 bp, respectively, as
120 biological replicates (**Supplementary Fig. 2a**). Targeted cells were clonally expanded for selection of
121 heterozygous (*Kmt2d*$^{+/\Delta}$) and homozygous (*Kmt2d*$^{\Delta/\Delta}$) mutated cell lines for comparison against the
122 parental wild-type line (*Kmt2d*$^{+/+}$). Editing efficiency was verified by PCR and Sanger sequencing of
123 genomic DNA (**Supplementary Fig. 2b-d**). For effects on expression, *Kmt2d* mRNA was analyzed by
124 RT-qPCR comparing exons upstream (exon 15) and within (exon 54) the deletion site (**Supplementary**
125 **Fig. 2e**), demonstrating that mRNA message of the upstream exon was unaffected, while the targeted
126 exon had ~50% decreased message in *Kmt2d*$^{+/\Delta}$ cells and 100% decreased message in *Kmt2d*$^{\Delta/\Delta}$ cells,
127 indicating deletion but not degradation of *Kmt2d* transcript. Furthermore, immunofluorescence using
128 KMT2D antibody with antigenicity upstream of the deletion confirmed presence of KMT2D protein in
129 *Kmt2d*$^{+/+}$ cells as well as *Kmt2d*$^{\Delta/\Delta}$ cells (**Fig. 2a**).
130 72 hours after equal-density plating, cell counts revealed proliferation rates ~39% lower in
131 *Kmt2d*$^{\Delta/\Delta}$ cells compared to wild-type (**Fig. 2b**), further reflected by CellTrace Violet assay measuring
132 less dilution of fluorescent tracer, i.e. fewer cell divisions, in *Kmt2d*$^{\Delta/\Delta}$ daughter cells compared to wild-
133 type over 72 hours (**Fig. 2c**, left). Importantly, genotype had no effect on parental cell uptake of the
134 fluorescent tracer (**Fig. 2c**, right). Cell cycle occupancy at steady-state was analyzed by flow cytometry
135 against marker of proliferation Ki-67 (KI67), in combination with DNA label, revealing enrichment of G_2/M
136 cells and paucity of G_0/G_1 phase cells among *Kmt2d*$^{\Delta/\Delta}$ cells compared to wild-type (**Fig. 2d,**
137 **Supplementary Fig. 2f)**. To better characterize temporal dynamics of cell cycle progression, we
138 synchronized cells in G_2/M phase using nocodazole, and analyzed DNA content at various timepoints

139 after release (**Fig. 2e**). By 3 hours post-release, wild-type cells showed significantly greater exit from
140 G_2/M phase compared to *Kmt2d*$^{\Delta/\Delta}$ cells, with concomitant increase of G_0/G_1 phase daughter cells lasting
141 up to 18 hours (**Fig. 2e', e''**). Finally, detailed analysis of cell death was performed using flow cytometric
142 detection of fluorescence coupled to caspase-3/7-mediated substrate cleavage to label early apoptotic
143 cells. Compared to wild-type, the proportion of apoptotic cells was several-fold larger in both *Kmt2d*$^{+/\Delta}$
144 cells (~287%) and *Kmt2d*$^{\Delta/\Delta}$ cells (~478%) (**Fig. 2f**).
145 Findings of proliferative defects, G_2/M cell cycle delay, and increased apoptosis in *Kmt2d*-deficient
146 hippocampal cells upon specific disruption of the SET catalytic domain make this an attractive model for
147 detailed transcriptional and chromatin profiling of KMT2D function in neurodevelopmental contexts.
148
149 **Transcriptional suppression of metabolic genes in cycling cells, and precocious pro-neuronal**
150 **gene expression in differentiating cells among KS1 patient-derived NPCs**
151 To interrogate transcriptional effects of *KMT2D* loss in the context of NPC differentiation
152 trajectory, we performed single-cell RNA sequencing (scRNA-seq) in iPSCs and NPCs from the KS1
153 patient and controls (**Supplementary Fig. 3a**). We first inspected cell clusters for expression of known
154 iPSC or NPC markers to further verify cell identities, confirming that iPSCs displayed characteristic
155 expression of pluripotency markers including POU domain, class 5, transcription factor 1 (*POU5F1*)
156 (**Supplementary Fig. 3b**), while NPCs expressed known markers of early neural lineage including paired
157 box protein Pax-6 (*PAX6*) **(Supplementary Fig. 3c**). A subset of NPCs further expressed neuronal
158 maturation markers including tubulin beta 3 class-III (*TUBB3/TUJ1*) (**Supplementary Fig. 3d**).
159 Differential expression analysis in iPSCs and NPCs identified genes downregulated (down
160 hereafter) or upregulated (up hereafter) in KS1 patient relative to healthy controls (**Supplementary**
161 **Figure 3e**). KS1 iPSCs displayed strong transcriptional suppression among 421 significantly differentially
162 expressed genes (DEGs), with 372 genes (~88%) down and only 49 genes (~12%) up (**Supplementary**
163 **Table 1**). In contrast, NPCs showed less directional bias with 346 significant DEGs identified, among
164 which 147 genes (~42%) were down and 199 genes (~58%) were up (**Supplementary Table 2**).
165 Intersection of DEG lists showed that 40 genes were shared down in KS1 iPSCs and NPCs, 10 genes
166 were shared up, and surprisingly, 66 genes were found to be down in KS1 iPSCs yet up in KS1 NPCs
167 (**Supplementary Figure 3f, Supplementary Table 3**). Notably, among genes down we observed factors
168 important to stem cell maintenance including transcription factor SOX-2 (*SOX2*), proliferation-associated
169 protein 2G4 (*PA2G4*), protein lin-28 homolog A (*LIN28A*), G1/S-specific cyclin D1 (*CCND1*), and
170 lactadherin (*MFGE8*)[21]. 9 HOX genes, known KMT2D targets, were also down as well as
171 neurodevelopmental disorder-linked genes RNA-binding motif protein 8A (*RBM8A*)[22], small nuclear

172 ribonucleoprotein-associated protein N (*SNRPN*), and necdin (*NDN*). We performed Overrepresentation
173 Analysis (ORA) in iPSC and NPC DEG sets to interrogate enrichment of molecular pathways specifically
174 among down or up genes (**Fig. 3a**). The strongest gene network enrichments were found to reflect those
175 DEGs down in KS1 iPSCs, yet up in KS1 NPCs, i.e. pathways most strongly downregulated in KS1 iPSCs
176 were the same pathways upregulated in KS1 NPCs, compared to respective controls. These networks
177 included NMD, translation initiation, and protein targeting. In contrast, genes down in KS1 NPCs showed
178 enrichment for chaperone-mediated protein folding, G_2/M cell cycle progression, and endoplasmic
179 reticulum/phagosome pathways. Among shared down genes in both KS1 iPSCs and NPCs, the strongest
180 enrichments were in target genes of the transcription factors upstream stimulatory factor 1 (USF1),
181 regulating glucose homeostasis and hypoxia response, Myc proto-oncogene protein (MYC), and signal
182 transducer and activator of transcription 1 (STAT1).

183 To better resolve transcriptional signatures of neural development, we next focused specifically
184 on NPCs from KS1 patient and controls. We used Uniform Manifold Approximation and Projection
185 (UMAP) to visualize single cells in a manner that displays high-dimensionality data while preserving both
186 local and global relationships[23]. Cells of both control NPC lines were tightly clustered, indicating similar
187 expression profiles, in contrast to a distinct separation of KS1 patient cells which appears to gradually
188 lessen in a subset (top) of cells that more closely resemble controls (**Fig. 3b**). By partitioning cells based
189 on expression of cell cycle phase-specific genes (**Fig. 3c**) or maturation stage-specific genes (**Fig. 3d**),
190 we defined an NPC differentiation trajectory consisting of early or cycling NPCs, transitioning NPCs, and
191 differentiating NPCs. We determined that cycling cells comprise the vast majority of NPCs analyzed and
192 exhibit the greatest KS1-associated expression differences (**Fig. 3d**, purple cells), while expression
193 profiles of transitioning and differentiating NPCs show gradual convergence of gene expression (**Fig. 3d**,
194 grey and yellow cells).

195 We next asked whether differences in cell cycle phase composition might account for the strong
196 KS1-associated differential expression in this population. We analyzed cells partitioned by cycle phase
197 to assign individual NPCs into subsets of G_1/G_0 (~55-61% of total cells), S (~20-21%), and G_2/M cells
198 (~19-24%). As expected, S and G_2/M cells are seen less among differentiating NPCs than cycling, yet
199 proportions did not appreciably differ by genotype, in agreement with flow cytometry data in the same
200 cells (**Supplementary Fig. 3g**, **Fig. 1f**). Thus, cycle phase composition is unlikely to drive expression
201 differences in this model. We next interrogated differential expression specifically in NPC subsets of
202 G_1/G_0, S, and G_2/M cells, and observed similar proportions of KS1 up and down genes in each phase
203 subset (55-66% and 33-35%, respectively), with 102 up and 85 down in G_1/G_0 phase cells, 28 up and 15
204 down in S phase cells, and 28 up and 14 down in G_2/M phase cells (**Supplementary Table 4**). Individual

205 DEGs were nearly identical in S and G_2/M phase subsets, and overlapped with DEGs from the G_1/G_0
206 phase subset. Thus, transcriptional effects in KS1 NPCs are due to gene expression differences in G_1/G_0
207 phase cells, and not from differences in cell cycle phase composition.

208 Apart from increased rates of cell death (**Fig. 1g**), another factor that could contribute to a paucity
209 of proliferative NPCs (**Fig. 1e**) would be precocious cell maturation resulting in depletion of cycling
210 precursors. To explore this in detail with scRNA-seq data, we profiled neural differentiation markers
211 (**Supplementary Fig. 3h**) to reveal a continuum of increasingly maturing NPCs ranging from immature
212 cells (*PAX6*$^{+}$) to the most differentiated cells (*MAP2*$^{+}$) (**Fig. 3e**). We further restricted analysis to
213 transitioning and differentiating NPCs, defining a trajectory that enabled parsing of cells into binned
214 deciles of increasing maturation (**Supplementary Fig. 3i**). Quantification of cell densities revealed strong
215 bias of KS1 NPCs in the most differentiated bins relative to controls (**Fig. 3f**). These transcriptional
216 signatures of precocious maturation were corroborated at protein level by flow cytometric analysis, finding
217 KS1 NPCs had increased MAP2 fluorescence and reciprocally decreased PAX6 fluorescence relative to
218 control (**Supplementary Fig. 3j, j'**). Finally, we analyzed DEGs specifically within cycling, transitioning,
219 or differentiating NPC subsets to determine if particular gene networks drive transcriptional differences
220 in a stage-specific manner (**Supplementary Table 4**). The most prevalent networks overall, NMD,
221 translation initiation, and protein targeting, were strongly enriched regardless of maturation stage (**Fig.
222 3g**). In contrast, DEGs in transitioning NPCs, and to a lesser extent cycling NPCs, show enrichment of
223 genes comprising the Notch signaling pathway including delta-like protein 3 (*DLL3*), protein jagged-1
224 (*JAG1*), transcription factor HES-5 (*HES5*), and *CCND1*. Cycling NPCs, but not transitioning or
225 differentiating NPCs, were enriched for DEGs in anaerobic glycolysis pathways including enolase 1,
226 (*ENO1*), pyruvate kinase PKM (*PKM*), L-lactate dehydrogenase L-chain (*LDHA*), fructose-biphosphate
227 aldolase A (*ALDOA*), triosephosphate isomerase (*TPI1*), and polyadenylate-binding protein 1 (*PABPC1*).

228 Together, these results link potent transcriptional effects on cellular metabolism to cell-
229 autonomous proliferation defects in KS1 patient-derived stem cell models, raising the question of whether
230 these features generalize to other independent KMT2D-deficient neuronal models.

231

232 **Globally suppressed transcription upon loss of the KMT2D SET methyltransferase domain, with**
233 **enrichment of KMT2D-bound genes in glycolytic and oxygen-sensing pathways**

234 We performed high-coverage RNA-seq in three biological *Kmt2d*$^{\Delta/\Delta}$ replicates for comparison
235 against the parental *Kmt2d*$^{+/+}$ line, each in technical triplicate, followed by differential expression analysis.
236 This revealed 575 significant DEGs at a False Discovery Rate (FDR) of 0.05 in *Kmt2d*$^{\Delta/\Delta}$ cells compared
237 to wild-type (**Fig. 4a, Supplementary Table 5**), with Principal Component Analysis (PCA) showing clear

238 separation of mutant libraries from wild-type (**Supplementary Fig. 4a-c**). A large majority of DEGs,
239 ~76%, were down in *Kmt2d*$^{\Delta/\Delta}$ cells (436 genes), revealing strong global suppression of gene expression.
240 The gene with highest downward fold change was the epigenetic modifier methylcytosine dioxygenase
241 (*Tet2),* which catalyzes conversion of 5-methylcytosine to 5-hydroxymethylcytosine during active DNA
242 demethylation, and known KMT2D target genes such as krueppel-like factor 10 (*Klf10*)[12] were also
243 observed. ORA for gene networks enriched among *Kmt2d*$^{\Delta/\Delta}$ down DEGs yielded glycolysis, HIF1
244 signaling, Ras signaling, autophagy, BMP signaling, ERK1/ERK2 cascade, and others, while *Kmt2d*$^{\Delta/\Delta}$
245 up DEGs were enriched in fewer networks including ribosome biogenesis, poly(A) RNA binding, and Myc
246 target genes (**Fig. 4b**). Interestingly, against a defined set of transcripts known to co-localize with
247 translation machinery at synaptic partitions in neurons[24], *Kmt2d*$^{\Delta/\Delta}$ down, but not up, DEGs were highly
248 enriched (**Supplementary Fig. 4d**).

249 Since KMT2D-mediated transcriptional promotion occurs in a target gene-specific manner, we
250 reasoned that among the 575 observed *Kmt2d*$^{\Delta/\Delta}$ DEGs, a subset of down genes found to also bind
251 KMT2D itself in wild-type cells would more likely represent primary transcriptional consequences of
252 KMT2D loss, whereas other DEGs may reflect secondary effects of broader cellular perturbations. We
253 performed chromatin immunoprecipitation and high-throughput sequencing (ChIP-seq) using KMT2D
254 antibody in wild-type and *Kmt2d*$^{\Delta/\Delta}$ HT22 cells, first analyzing normal KMT2D binding in wild-type cells to
255 yield 3,756 significant peaks (**Supplementary Fig. 4e, Supplementary Table 6**). Of these, 1,235 peaks
256 (~33%) occur within 5 kb of gene promoters (**Supplementary Table 7**). These promoter-proximal KMT2D
257 peaks (KMT2D-bound genes) were most enriched in mRNA 3'UTR binding, Rho GTPase signaling,
258 nuclear factor kappa B subunit 1 (NFKB) signaling, translation elongation, oxidative stress-induced
259 senescence, HIF1 signaling, and others (**Fig. 4c**). Finally, we intersected a list of KMT2D-bound,
260 *Kmt2d*$^{\Delta/\Delta}$ DEGs to reveal just 74 such genes (**Table 1**), of which a large majority (~85% or 63 genes)
261 were as expected downregulated (**Fig. 4d**), including insulin-like growth factor 1 (*Igf1*), fos-like antigen 2
262 (*Fosl2*), and at least 20 previously described KMT2D target genes in other tissues[7,25]. These 63 KMT2D-
263 bound, *Kmt2d*$^{\Delta/\Delta}$ down DEGs were most significantly enriched for face morphogenesis, glycolytic
264 process, protein kinase B signaling, hypoxia response, and cell proliferation genes, and 29 of these 63
265 genes are known to be HIF-regulated[26] (**Fig. 4e, Table 1**). KMT2D ChIP peaks at these genes showed
266 clustering of KMT2D peaks on promoters (**Fig. 4f**), often overlapping CpG islands, of genes such as
267 angiopoietin-like 4 (*Angptl4*), Cbp/p300-interacting transactivator 2 (*Cited2*), transducer of Erbb2,1
268 (*Tob1*), DNA damage-inducible transcript (*Ddit4*), and FOS-like antigen 2 (*Fosl2*), but also clustering at
269 alternative transcription start sites as was observed with retinoic acid receptor alpha (*Rara*).

270 Noting that a disproportionately high fraction of KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ down DEGs control
271 hypoxia response, and as the critical HIF1A stability regulator Von Hippel-Landau tumor suppressor (*Vhl*)
272 was also significantly down in *Kmt2d*$^{\Delta/\Delta}$ cells, we further interrogated oxygen-sensing transcription in both
273 *Kmt2d*$^{\Delta/\Delta}$ cells as well as KS1 patient-derived iPSCs and NPCs. Surprisingly, downregulated genes in
274 each model were significantly enriched for HIF1A direct targets and genes containing the hypoxia
275 response element (HRE) RCGTG motif, in addition to known hypoxia response genes (**Supplementary**
276 **Fig. 4f**).
277
278 ***Kmt2d*-dependent HIF1A regulation and transactivation, and cell survival in response to hypoxia**
279 Given consistent dysregulation of oxygen-sensing pathways across multiple KS1 models, we
280 hypothesized that *Kmt2d* deficient HT22 cells may show differential response from wild-type cells under
281 hypoxic conditions. To explore this, we subjected *Kmt2d*$^{+/+}$ cells, *Kmt2d*$^{+/\Delta}$ cells, and *Kmt2d*$^{\Delta/\Delta}$ cells to
282 normoxic (20% O_2) or hypoxic (1% O_2) conditions for 72 hours. RT-qPCR analysis of canonical HIF1A
283 transactivation targets *Vegfa, Bnip3, Ddit3, Cdkn1a,* and *Tgfb3* in wild-type cells revealed induction upon
284 hypoxia, whereas *Kmt2d*$^{+/\Delta}$ and *Kmt2d*$^{\Delta/\Delta}$ cells failed to induce these genes to the same levels (**Fig. 5a**).
285 Stabilized, or activated HIF1A undergoes nuclear translocation, so we then quantified the proportion of
286 cellular HIF1A fluorescence colocalizing in the nucleus under normoxic and hypoxic conditions (**Fig. 5b,**
287 **Supplementary Fig. 5a**). Paradoxically, even under normoxic conditions *Kmt2d*$^{\Delta/\Delta}$ cells exhibited ~1.98-
288 fold greater proportion (~17.5%) of nuclear HIF1A than wild-type (~8.8%), and further increased upon
289 hypoxic exposure (**Fig. 5c**). In *Kmt2d*$^{+/\Delta}$ cells, nuclear HIF1A under normoxia (~7.3%) was similar to wild-
290 type, but upon hypoxic exposure sharply increased to ~24.8%. Wild-type cells showed only modestly
291 increased nuclear HIF1A in this hypoxia paradigm, suggesting a lower threshold for hypoxic HIF1A
292 activation in *Kmt2d*$^{+/\Delta}$ cells in addition to constitutive HIF1A activation in *Kmt2d*$^{\Delta/\Delta}$ cells. Under normal
293 oxygen conditions, we confirmed over-representation in G_2/M phase of the cell cycle in *Kmt2d*$^{\Delta/\Delta}$ cells,
294 and hypoxia increased G_2/M phase occupancy in all genotypes (**Fig. 5d**). Flow cytometric quantification
295 of caspase-3/7 activity confirmed higher apoptotic cell proportions under normoxia in *Kmt2d*$^{+/\Delta}$ and
296 *Kmt2d*$^{\Delta/\Delta}$ cells compared to wild-type, and apoptotic cell numbers sharply decreased in all genotypes
297 under hypoxia (**Fig. 5e, e'**). In contrast, necrotic cell proportions detected by SYTOX fluorescence were
298 similar in all genotypes under normoxia, but under hypoxia decreased only in wild-type cells and *Kmt2d*$^{+/\Delta}$
299 cells, while marginally increasing in *Kmt2d*$^{\Delta/\Delta}$ cells (**Fig. 5e, 5e''**).
300 Comparing transcriptional data from KS1 patient cells and *Kmt2d*$^{\Delta/\Delta}$ mouse hippocampal cells
301 relative to respective healthy controls, we also noted strong enrichments of DEGs controlling mRNA and
302 protein metabolism including protein/nucleic acid deglycase (*PARK7*), serine-threonine protein kinase

303 (*PINK1*), X-box binding protein 1 (*XBP1*), ubiquitin carboxyl-terminal hydrolase 13 (*USP13*), RNA
304 demethylase ALKBH5 (*ALKBH5*), insulin-like growth factor 2 mRNA binding protein 1 (*IGF2BP1*), and
305 U6 small nuclear RNA (adenine-(43)-N(6))-methyltransferase (*METTL16*). Strikingly, while ~76% of all
306 DEGs in *Kmt2d*$^{\Delta/\Delta}$ cells were downregulated, evaluation of epitranscriptome modifiers controlling mRNA
307 metabolic rates[27–29] revealed that 22 were upregulated while only 2 were downregulated. Furthermore,
308 mutation of a large protein such as KMT2D with over 5,500 amino acids, without proper degradation,
309 could overwhelm proteasome function as has been described in the context of DNMT1 mutation[30]. Upon
310 evaluation of transcripts known to undergo regulation by NMD[31], we indeed observed significant
311 enrichment among our DEGs of NMD-regulated transcripts, as well as genes directly targeted by UPF1
312 for NMD degradation (**Supplementary Fig. 5b, c**). At the protein level, detection using Proteostat reagent
313 revealed constitutive build-up of misfolded protein aggregates in *Kmt2d*$^{\Delta/\Delta}$ cells relative to wild-type, as
314 well as in wild-type cells treated with proteasome inhibitor MG-132 as positive control, and flow cytometric
315 analysis confirmed increased Proteostat intensity in both the *Kmt2d*$^{\Delta/\Delta}$ cells and treated positive controls
316 (**Supplementary Fig. 5d**). Interestingly, aggregate levels were reduced after hypoxic exposure in both
317 genotypes (**Supplementary Fig. 5d'**). Finally, we explored whether this rescue also occurred in KS1
318 patient fibroblasts. While flow cytometric analysis demonstrated high inter-patient variation of Proteostat
319 fluorescence at baseline, hypoxic exposure consistently reduced Proteostat fluorescence in KS1
320 fibroblasts but not in controls, suggestive of genotype-dependent effects (**Supplementary Fig. 5e**).
321

322 ***Kmt2d*-deficient mice exhibit depletion of adult-born neural precursors, with cell cycle defects**
323 **and transcriptional suppression of cellular metabolic pathways**

324 We next asked whether defects from *in vitro* KS1 models affecting proliferation, cell cycle, and
325 survival are detectable *in vivo* in a characterized KS1 mouse model (*Kmt2d*$^{+/\beta geo}$) with heterozygous
326 *Kmt2d* SET domain loss leading to a truncation. *Kmt2d*$^{+/\beta geo}$ mice previously were found to exhibit
327 visuospatial memory impairments associating with reductions of DCX^{+} or EdU^{+} cells in the DG[12,32], but
328 neurogenic lineage progression has not been described in these mice.

329 We used immunostaining against stage-specific markers (**Fig. 6a, a'**) to compare abundance of
330 individual DG cell precursor stages (**Supplementary Fig. 6a**) of adult mice 8 weeks old, comparing
331 *Kmt2d*$^{+/\beta geo}$ mice to sex- and age-matched *Kmt2d*$^{+/+}$ littermates. We observe significantly fewer neural
332 progenitors in *Kmt2d*$^{+/\beta geo}$ mice compared to *Kmt2d*$^{+/+}$ mice at all stages analyzed (**Fig. 6b**), indicating
333 that heterozygous *Kmt2d* loss restricts normal steady states of NPCs in the hippocampus. Importantly,
334 quiescent radial glia-like cells (qRGLs, Nestin^{+}/MCM2^{-}) were ~39% less numerous in *Kmt2d*$^{+/\beta geo}$ mice
335 compared to *Kmt2d*$^{+/+}$ mice, indicating a baseline depletion of the adult neural stem cell pool. Activated

336 RGLs (aRGLs, Nestin$^+$/MCM2$^+$) were ~43% less numerous compared to *Kmt2d*$^{+/+}$ littermates, and
337 intermediate progenitor cells (IPCs) (Nestin$^-$/MCM2$^+$), were ~26% fewer. We confirm our prior
338 observations[12,32] of reduced representation of neuroblasts (NBs, DCX$^+$) which were 28% decreased in
339 *Kmt2d*$^{+/\beta geo}$ mice compared to *Kmt2d*$^{+/+}$ littermates. From these data, we then calculated a lineage
340 progression index to approximate the expansion potential of each successive neurogenic cell type.
341 Although *Kmt2d*$^{+/\beta geo}$ mice showed fewer total cells of each type at steady-state, the lineage progression
342 index at each cell-type transition did not differ significantly in mutants, in fact appearing modestly higher
343 at the aRGL-IPC transition in *Kmt2d*$^{+/\beta geo}$ mice, suggesting that cell-type transition defects are not
344 observable at any particular stage analyzed (**Fig. 6c**).

345 By stratifying our analysis anatomically along the septotemporal axis of the DG, we observe that
346 aRGL reductions in *Kmt2d*$^{+/\beta geo}$ mice were more pronounced in the septal DG than the temporal region
347 (**Fig. 6d**), congruous with *in vivo* spatial memory defects previously observed[12] which have been localized
348 to this specific DG region[33]. Because DCX$^+$ NBs characteristically migrate radially during differentiation,
349 we compared radial distances of DCX$^+$ cell bodies from the SGZ plane and observed increased distances
350 in *Kmt2d*$^{+/\beta geo}$ mice (**Supplementary Fig. 6b**). Finally, despite reductions of NPC populations in
351 *Kmt2d*$^{+/\beta geo}$ mice, we observed no numeric differences among mature neurons in the DG (NEUN$^+$,
352 **Supplementary Fig. 6c**), nor were any gross abnormalities uncovered with post-mortem MRI volumetric
353 analysis (**Supplementary Fig. 6d, Supplementary Table 8**).

354 For RNA-seq analysis and cell cycle studies in *Kmt2d*$^{+/\beta geo}$ mice, we employed a saturating EdU
355 pulse paradigm to label dividing cells in 8-week old mice, and sampled micro-dissected DG within 1 cell
356 cycle (16 hours) in order to capture the full complement of adult-born, dividing NPCs (**Supplementary**
357 **Fig. 6e**), then purifying EdU$^+$ nuclei by fluorescence-activated cell sorting (FACS) (**Fig. 6e,**
358 **Supplementary Fig. 6f**). Analysis of DNA content in the EdU$^+$ population revealed enrichment in G_2/M
359 phase and paucity of G_0/G_1 phase (**Fig. 6f**). We next performed transcriptional profiling by RNA-seq in
360 EdU$^+$ DG nuclei, yielding 827 DEGs (**Supplementary Table 9**). Among 416 down genes in *Kmt2d*$^{+/\beta geo}$
361 nuclei, the most significant enrichments were for misfolded protein binding, TCA cycle, proteasome
362 complex, oxygen response, poly(A) RNA-binding genes (**Fig. 6h**), including *Pink1*, *Xbp1*, *Usp13*,
363 phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*Pik3ca*), E3 ubiquitin protein
364 ligase SMURF2 (*Smurf2*), gamma-amino butyric acid receptor-associated protein (*Gabarap*), and insulin-
365 like growth factor-binding protein 2 (*Igfbp2*). Among 411 up genes in these nuclei were significant
366 enrichments for translation, p53 binding, and HIF1 targets including *Pabpc1* and hypoxia-inducible factor
367 1A (*Hif1a*) itself. Notably, the upstream pro-apoptosis gene caspase-8 (*Casp8*) was also significantly
368 upregulated in these *Kmt2d*$^{+/\beta geo}$ proliferative nuclei.

369 Given the observed transcriptional suppression in poly(A) RNA-binding genes, we reasoned that
370 improper 3'UTR-mediated mRNA metabolism could lead to accumulation of transcripts influencing NPC
371 maturation. Indeed, despite little overall bias in up- or downregulation among *Kmt2d*$^{+/\beta geo}$ proliferative
372 nuclei, interrogation of canonical positive regulators of neuronal differentiation revealed a marked
373 predominance of pro-neural transcripts upregulated, having only 3 genes down but 14 genes up,
374 including copine-1 (*Cpne1*), focal adhesion kinase 1 (*Ptk2*), Ras-related protein Rab-11A (*Rab11a*),
375 retinoblastoma-associated protein 1 (*Rb1*), and neuronal migration protein doublecortin (*Dcx*) itself.
376 Interestingly, in KS1 patient NPCs, *DCX* in fact was the single most significantly upregulated gene, along
377 with others upregulated such as nuclear receptor subfamily 2, group F, member 1 (*NR2F1*), pro-neural
378 transcription factor HES-1 (*HES1*), and GABAergic interneuron-promoting ladybird homeobox 1 (*LBX1*).
379 *Kmt2d*$^{\Delta/\Delta}$ hippocampal cells compared to wild-type had significantly higher transcript levels of brain-
380 derived neurotrophic factor (*Bdnf*) and neuron-specific microtubule element (*Tubb3/Tuj1*).
381 Together, *in vivo* studies of adult neurogenesis in *Kmt2d*$^{+/\beta geo}$ mice reveal reduced representation
382 of adult DG NPCs from the earliest stages, and cell cycle defects accompanied by metabolic
383 transcriptional features conserved both in *Kmt2d*$^{\Delta/\Delta}$ hippocampal cells and KS1 patient-derived stem
384 cells. Thus, neurodevelopmental disruptions from KMT2D deficiency observed cell-autonomously *in vitro*
385 are recapitulated *in vivo* in a heterogeneous adult NPC niche in KS1 mice.
386
387 **Discussion**
388 Developmental disabilities are estimated to affect nearly 2% of school-age children[34].
389 Improvements in genetic diagnosis have revealed a major emerging cause of ID, the Mendelian disorders
390 of epigenetic machinery, caused by genetic lesions in genes encoding chromatin-modifying enzymes that
391 lead to transcriptional perturbance through disruption of local chromatin states[2]. In KS1, mutations in the
392 histone methyltransferase KMT2D have phenotypic effects in multiple tissues, but neurodevelopmental
393 aspects are currently poorly understood. KS1 diagnoses are typically made during infancy, but the
394 inherent reversibility of chromatin modifications raises the possibility that a detailed understanding of
395 KMT2D activity in neuronal cells could identify molecular targets for postnatal interventions in KS1
396 patients.
397 Here, we report that KMT2D-deficient human and mouse NPC models, *in vitro* and *in vivo*,
398 demonstrate disruption of normal cell cycle progression. Previously, disrupted cell cycle progression had
399 been described upon KMT2D loss in cardiac precursor and B cell precursor models[6,8]. However, these
400 studies observed an increase in cells in G_1/G_0, and S phases with a concurrent decrease in G_2/M
401 compared to wildtype cells, while our findings in NPC models suggest overrepresentation of G_2/M in KS1

402 versus controls. Collectively, these data suggest that disruption of normal progression through the cell
403 cycle may be a unifying feature in KS1, given that characteristic phenotypes include congenital heart
404 disease, immune dysregulation, and intellectual disability[11]. These affected tissues differ in embryonic
405 origin, suggesting that genes important to KS1 phenotypes support basic cellular homeostatic functions
406 regarding housekeeping, energy production, or cell cycle progression, rather than genes with purely
407 tissue-specific function. Notably, Singh and colleagues[35] recently described dynamic changes in H3K4
408 methylation at developmental gene promoters during cell cycle progression, supporting a cycle-
409 dependent instructive role for chromatin modifiers in developmental lineage priming.

410 It is possible that differential representations of cell cycle phase in KS1 cells compared to control
411 could confound some of our studies, as cellular composition may artificially create or hide gene
412 expression changes. However, this would only be relevant for genes that show significant changes in
413 individual phases of the cell cycle, and a recent estimate suggested that only about 850 genes show such
414 characteristics[36]. Moreover, we sought to eliminate cycle-related confounders by analyzing phase-
415 specific gene expression in KS1 patient NPCs, but future studies might focus on purified cells from
416 individual cell cycle phases. KS1 NPCs, in contrast to other models, did not demonstrate a cell cycle
417 phenotype or strong bias toward downregulated gene expression, and while this may reflect strong
418 transcriptional effects induced during the *in vitro* differentiation protocol, other reasons cannot be
419 excluded.

420 We describe concordance of two additional cellular phenotypes in KMT2D-deficient NPCs:
421 proliferation defects and transcriptional states indicating suppression of central carbon metabolism and
422 oxygen sensing. We observe these phenotypes both and *in vivo* and *in vitro*, generating human and
423 mouse cellular KS1 models of neural disruption which may be useful for future screening of potential
424 therapies and testing mechanistic hypotheses, and further highlighting the utility of human-derived
425 models of neurodevelopment[37]. We find these phenotypes to be cell-autonomous in KMT2D-deficient
426 cells, suggesting that gene expression studies in these cells could yield disease-relevant KMT2D target
427 genes and pathways. Furthermore, these models demonstrate comparable cellular and transcriptional
428 phenotypes that result both from haploinsufficiency of *KMT2D* (patient iPSCs, NPCs), and variants that
429 more specifically target the region encoding *Kmt2d*'s methyltransferase domain (HT22 hippocampal cells,
430 *Kmt2d*$^{+/\beta geo}$ mice), indicating that loss, in either dosage or catalytic function, of KMT2D is pathogenic
431 independent of mutation type.

432 Our findings suggest that adult neurogenesis defects, which we previously described in
433 *Kmt2d*$^{+/\beta geo}$ mice compared to wild-type littermates, originate from a diminution of the RGL stem cell pool
434 itself rather than from stalling in particular cell type transitions. No gross changes in brain region volumes
435 or NEUN^{+} mature neuron abundance in the GCL were observed. The adult DG stem cell niche coalesces

436 perinatally, suggesting a developmental component to the observed *in vivo Kmt2d*$^{+/\beta geo}$ phenotype,
437 although further experiments are needed. While adult-born NPCs comprise only a small fraction of the
438 largely post-mitotic adult hippocampus, they uniquely display transient functions of heightened
439 connectivity with outsized influence on existing hippocampal circuitry[38], and further studies on the effects
440 of KMT2D loss on these learning-associated circuits will be required. Notably, the gene most potently
441 suppressed in *Kmt2d*$^{\Delta/\Delta}$ hippocampal cells was *Tet2*, which itself was recently shown to rescue age-
442 related declines in adult neurogenesis and cognition when overexpressed[39].

443 Genome-wide chromatin profiling revealed that cells lacking functional KMT2D methyltransferase
444 activity have a subset of differentially-expressed genes which are themselves KMT2D-bound,
445 mechanistically linking KMT2D dysfunction to the transcriptional suppression observed in each of our
446 neural KS1 models. We describe systematic suppression of hypoxia response pathways, and
447 demonstrate KMT2D-dependent differences in hypoxia response including induction of necrotic cell
448 death and aberrant HIF1A activation. Findings of increased HIF1A activation, yet dampened HIF1A-
449 responsive transcription programs in KS1 cells, are compatible with chronic HIF1A activity as a
450 compensatory change in response to the mutation. Future studies should interrogate differences in local
451 chromatin environment at critical hypoxia response genes in KS1 models, as well as any compensatory
452 changes emerging in response to chronic activation of HIF1A. Interestingly, at least 33 chromatin
453 modifiers are known to impact HIF1A-mediated transactivation, 11 of these associating with
454 developmental disease or cancers, yet KMT2D had not yet been implicated[40]. Two features of adult-born
455 hippocampal NPCs make hypoxia-responsive and metabolic transcription pathways attractive for future
456 studies. First, these cells develop in the SGZ niche, where a locally hypoxic microenvironment enhances
457 their survival[41], making them potentially more vulnerable to compromised oxygen-sensing mechanisms
458 as they exit this microenvironment upon migration. Second, NPC maturation is coupled to a metabolic
459 rewiring from glycolysis in NPCs to oxidative phosphorylation in neurons. Zheng and colleagues[15]
460 recently found this metabolic switch, marked by suppression of glycolytic genes (*Ldha, Hk2*), to be
461 essential for neuronal maturation. We observe suppression of these and related genes across KS1
462 models, accompanied by upregulation of pro-neural genes. Thus, future studies should determine
463 whether rescue of KMT2D-dependent glycolysis programs is sufficient to prevent precocious
464 differentiation and niche depletion in KS1 NPC models, or whether analogous metabolic defects underlie
465 previously described premature activation of terminal differentiation genes in *KMT2D*-deficient
466 keratinocytes[10].

467 In summary, our findings suggest that KMT2D deficiency disrupts neural development by
468 negatively impacting NPC maintenance programs of cell cycle progression, proliferation, and survival.

469 Chromatin profiling in hippocampal cells identified KMT2D-regulated gene programs transcriptionally
470 suppressed across KS1 model systems, identifying genes encoding critical regulators of oxygen-sensing
471 and central carbon metabolism, accompanied by signatures of altered mRNA/protein metabolism and
472 precocious neuronal differentiation. We functionally demonstrate KMT2D-dependent responses to
473 hypoxia, accompanied by *in vivo* demonstration of stem cell maintenance defects echoing cell cycle and
474 proliferation defects observed both in KMT2D-deficient hippocampal cell culture and KS1 patient-derived
475 neurodevelopment models. Together, our findings support a model in which KMT2D loss transcriptionally
476 suppresses hypoxic response programs for NPC maintenance, causing exhaustion of precursors
477 required for adult hippocampal neurogenesis.
478
479 **Methods**
480 **Animals**
481 The *Kmt2d*$^{+/\beta geo}$ allele (Mll2Gt(RRt024)Byg) was generated by Bay Genomics (University of California)
482 through the random insertion of a gene trap vector. *Kmt2d*$^{+/\beta geo}$ mice were fully backcrossed to C57Bl/6
483 background over 10+ generations, and are born roughly in expected Mendelian ratios when bred to wild-
484 type. Animals were housed in a 14-hour light/10-hour dark cycle with free access to food and water. All
485 experiments were performed using mouse protocols approved by the Animal Care and Use Committee
486 of Johns Hopkins University School of Medicine and are in accordance with National Institutes of Health
487 (NIH) guidelines for mouse care and handling. All experiments compare mutant mice against age- and
488 sex-matched wild-type littermates.
489
490 **Patient-derived iPSCs, NPCs, and fibroblasts**
491 Media and reagents are listed (**Table S10**). Informed consent was obtained according to institutional IRB
492 and ISCRO protocols approved by Johns Hopkins University School of Medicine before skin biopsy and
493 fibroblast culture from molecularly confirmed KS1 patients (KS1-1, KS1-2, KS1-3) of the JHU Epigenetics
494 and Chromatin Clinic. Patient KS1-1 was consented for stem cell derivation and cells were reprogrammed
495 using non-integrating Sendai virus vectors (CytoTune-iPS 2.0, ThermoFisher Scientific). 5 days post-
496 induction, cells were transferred to mouse embryonic fibroblast (MEF) feeder plates in iPSC media and
497 monitored for colony formation. 21 days post-induction, colonies were manually selected for optimal iPSC
498 morphology and quality for propagation. Karyotype analysis by G-banding confirmed 46, XX normal
499 female karyotype in KS1-1. Generation and characterization of healthy control lines (C3-1 and C1-2) has
500 been previously described[18]. Feeder MEFs were derived from E13.5 CF-1 mouse embryos and mitotically
501 inactivated by irradiation. iPSCs were enzymatically passaged every 4-8 days using collagenase and

502 mechanical dissociation. NPCs were induced from iPSCs as previously described[19]. Briefly, we
503 synergistically inhibited signaling of glycogen synthase kinase 3 (GSK3), transforming growth factor β
504 (TGF-β), γ-secretase, and Notch signaling pathways using small molecules CHIR99021 (4 μM),
505 SB431542 (3 μM), and Compound E (0.1 μM), in the presence of hLIF (10 ng/ml) and ROCK inhibitor (5
506 μM) for 7 days. NPCs were split with Accutase and propagated in neural induction medium in feeder-free
507 conditions on a Matrigel.
508
509 **CRISPR-Cas9 deletions**
510 Media and reagents are listed (**Table S10**). HT22 mouse hippocampal cells were obtained commercially
511 (MilliporeSigma) and maintained in HT22 media (DMEM with 10% FBS, pen/strep, GlutaMAX). sgRNAs
512 targeting two loci spanning the *Kmt2d* SET domain-encoding region, with cut sites in exon 52 and either
513 exon 54 or intron 54, were designed and selected for on-target activity scores >90% (crispr.mit.edu,
514 **Table S10**), and integrated into Cas9 plasmid (pSpCas9BB-2A-puro v2.0 (PX459), Addgene). Plasmids
515 were delivered to HT22 cells at 20% confluency using Lipofectamine 2000 according to manufacturer
516 protocol. After puromycin selection, heterozygous and homozygous cells were clonally expanded and
517 maintained in culture according to standard protocols. Targeted clone genotypes were verified by PCR
518 (primers listed) and Sanger sequenced at the Johns Hopkins Genetic Resources Core Facility (GRCF).
519 A subset of clones appearing heterozygous by PCR (upper and lower DNA bands of expected size) were
520 found to bear strand invasion in the upper band (**Supplementary Fig. 2.g**) and were removed from
521 analyses (**Fig. 2a-d**). One such line, mutant Clone 1, was thus grouped with homozygous clones 2 and
522 3 for analysis, as both alleles were targeted in these three clones. Subsequent functional assays were
523 performed using Sanger-verified heterozygous and homozygous lines. Cells were tested negative for
524 mycoplasma contamination.
525
526 **RNA-seq in HT22 cells: library preparation**
527 Cells were plated at equal density and sampled at 60% confluency. Total RNA was isolated from three
528 biological replicates of *Kmt2d*$^{\Delta/\Delta}$ clones and *Kmt2d*$^{+/+}$ wild-type parental cells using Direct-Zol RNA
529 MicroPrep (Zymo Research), and libraries were constructed in technical triplicate using NEBNext Poly(A)
530 Magnetic Isolation Module and NEBNext UltraII RNA Library Prep Kit for Illumina (New England BioLabs),
531 with size selection by AMPure XP beads (Beckman Coulter), according to manufacturer protocols. Library
532 quantification and quality checks were done using KAPA Library Quantification Kit for Illumina (KAPA
533 Biosystems), High Sensitivity DNA Kit on BioAnalyzer (Agilent), and Qubit dsDNA HS Assay (Life

534 Technologies). Paired end 50 bp reads were obtained for pooled libraries using Illumina HiSeq 2500 at
535 the JHU GRCF High Throughput Sequencing Center.
536
537 **RNA-seq in HT22 cells: data analysis**
538 Sequencing reads were pseudoaligned to the mouse reference transcriptome (GRCm38) and transcript
539 abundances were subsequently quantified using Salmon[42]. We then used the tximport R package[43] to
540 convert the transcript abundances into normalized gene-level counts, by setting the
541 "countsFromAbundance" parameter equal to "lengthScaledTPM". Next, we used the edgeR[44] and
542 limma[45] R packages to log2 transform these gene-level counts, estimate the mean-variance relationship,
543 and calculate weights for each observation. In order to account for the correlation between technical
544 replicates of the same clone when performing the differential analysis, we fit a mixed linear model, using
545 the function "duplicateCorrelation" from the statmod R package[46] to block on clone. The differential
546 analysis was then performed using the limma R package. Differentially expressed genes were called with
547 0.05 as the cutoff for the False Discovery Rate. When performing the principal component analysis,
548 transcript abundances were first converted into gene-level counts using the tximport R package, with the
549 "countsFromAbundance" parameter equal to "no". Then, we applied a variance stabilizing transformation
550 to these gene-level counts using the "vst" function from the DESeq2 R package[47] with the parameter
551 "blind" set to "TRUE", and subsequently estimated the principal components using the 1000 most variable
552 genes.
553
554 **scRNA-seq: library preparation**
555 NPCs were induced in parallel from each iPSC line (KS1-1, C1-2, C3-1) under identical conditions, and
556 passaged three times before sampling. iPSCs were detached from MEF feeders using collagenase (200
557 units/ml). iPSCs and NPCs were dissociated to single-cell suspension using Accutase. Cell counts and
558 viability were analyzed using Countess II Automated Cell Counter (ThermoFisher Scientific). scRNA-seq
559 libraries were created with Chromium Single Cell 3' Library & Gel Bead Kit v2 (10x Genomics) according
560 to manufacturer protocol. Targeted cell recovery for each sample was 5,000 cells. Sufficient cDNA for
561 library construction was achieved using 20 amplification cycles for iPSC libraries and 16 cycles for NPC
562 libraries. Sample indexing was achieved using 11 PCR cycles for iPSC libraries and 5 cycles for NPC
563 libraries. scRNA-seq libraries were sequenced using Illumina NextSeq 500.
564
565 **scRNA-seq: data analysis**

566 Sequencing output was processed through the Cell Ranger 2.1.0 preprocessing pipeline using default
567 parameters with the exception of --expect-cells=5000 for `cellranger count` and --normalize=none for
568 `cellranger aggr`. Reads were quantified against hg19 using the 10x reference genome and
569 transcriptome builds (refdata-cellranger-GRCh38-1.2.0). The aggregated raw count matrix was then used
570 as input for the Monocle2 single-cell RNAseq framework. Differential gene expression analysis was
571 performed on all NPCs and iPSCs with respect to genotype (KS1 patient vs healthy control) and was
572 performed using the Monocle2[48] likelihood ratio test (0.1% FDR, Monocle2 LRT, Benjamini-Hochberg
573 corrected) with `num_genes_expressed` added as a nuisance parameter to both the full and reduced
574 models. The directionality of the differential gene test was determined by calculating the mean gene
575 expression across all KS1 patient-derived and healthy control cells respectively, evaluating the relative
576 fold change. High-variance genes were selected as those with a positive residual to the estimated
577 dispersion fit and a mean number of reads per cell >=0.0005. Cell cycle stage was determined by profiling
578 cell cycle associated genes across all cells and assigning cell cycle state using the R/Bioconductor
579 package scran[49]. Dimensionality reduction and visualization was performed via UMAP[23] on the
580 log10(counts + 1) of the high variance genes in the NPC dataset. The first 10 principal components were
581 used as input for UMAP using the default parameters of the R/CRAN package umap. Cells were assigned
582 to clusters using Monocle2's implementation of the louvain community detection algorithm. Learned
583 clusters were then aggregated by hand based on marker gene expression into three clusters
584 ("Differentiating", "Transitioning", "Cycling"). Differential gene expression within clusters, and between
585 genotypes was performed as described above. The "Differentiating" cluster was then segregated, and a
586 smooth line was fitted using a linear regression. This line was determined to represent the direction of
587 differentiation by examination of marker genes (Supplemental Figures 3h and 3i). The residuals of this fit
588 were then plotted and deciles were calculated containing equal number of cells along the axis of
589 differentiation. The number of cells in each decile was then counted with respect to genotype.
590

591 **ChIP-seq: library preparation**

592 *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ HT22 cells were sampled at 70% confluency and processed for pull-down with
593 ChIP-grade KMT2D antibody (Millipore Sigma) according to ENCODE guidelines. Sonicated, reverse-
594 crosslinked chromatin served as input control. Briefly, ~300 million cells per cell line were crosslinked in
595 1% formaldehyde, quenched with 0.125 M glycine, and cell lysate supernatants were collected for
596 immediate processing or snap-frozen for storage at -80°C. Nuclei diluted in 1 ml RIPA buffer were
597 sonicated using Bioruptor (Diagenode) for 6 cycles of 5 minutes (60 seconds on/30 seconds off) in ice-
598 cold water bath. Supernatants containing sheared chromatin were pre-cleared with Protein A Dynabeads

599 (ThermoFisher Scientific) and incubated overnight at 4°C with 8 µg KMT2D antibody. ChIP DNA was
600 recovered by Dynabead incubation (overnight at 4°C plus 6 hours at room temperature) before 6
601 sequential salt washes of increasing stringency, then eluted and reverse crosslinked overnight at 65°C.
602 DNA was purified using DNA Clean and Concentrator (Zymo Research) and quantified using High
603 Sensitivity DNA Kit on BioAnalyzer (Agilent), and Qubit dsDNA HS Assay (Life Technologies). DNA
604 libraries were constructed using NEBNext UltraII DNA Library Prep Kit for Illumina (New England
605 BioLabs) and quantified using KAPA Library Quantification Kit for Illumina (KAPA Biosystems). Paired
606 end 75 bp reads were obtained for pooled libraries using Illumina HiSeq 2500 at the JHU GRCF High
607 Throughput Sequencing Center.
608
609 **ChIP-seq: data analysis**
610 Sequencing reads were aligned to the mouse reference genome (mm10) using Bowtie2[50]. Then,
611 duplicate reads were removed with the function MarkDuplicates from Picard
612 (http://broadinstitute.github.io/picard/). Peaks were subsequently called using MACS2[51], with the "keep-
613 dup" parameter equal to "all". After peak calling, we excluded all peaks that overlapped with blacklisted
614 regions provided by ENCODE (https://sites.google.com/site/anshulkundaje/projects/blacklists). To
615 identify genes whose promoters were bound by KMT2D, we defined promoters as 10kb regions centered
616 around the TSS. We obtained TSS coordinates using the "exonsBy" function from the
617 "TxDb.Mmusculus.UCSC.mm10.knownGene" R package (Team BC, Maintainer BP (2018).
618 TxDb.Mmusculus.UCSC.mm10.knownGene: Annotation package for TxDb object(s). R package version
619 3.4.4.), with the "by" parameter equal to "gene".
620
621 **Purification of EdU$^+$ nuclei**
622 Mice were given 150 mg/kg EdU by intraperitoneal injection and sampled after 16 hours. Dentate gyrus
623 (DG) was micro-dissected in ice-cold PBS immediately following sacrifice by halothane inhalation. Total
624 nuclei were purified as described[52] with addition of RNase inhibitor to all buffers (**Table S10**). Briefly,
625 bilateral DG from individual mice were dounce-homogenized in 1 ml lysis buffer and layered above a
626 sucrose gradient for ultracentrifugation at 28,600 RPM for 2 hours at 4°C. Supernatant was aspirated
627 and nuclei were resuspended in Click-iT EdU AlexaFluor-488 Flow Cytometry Kit buffer with addition of
628 RNAse inhibitor, and incubated 30 minutes at room temperature. Samples were passed through 40 µm
629 filter, stained with 1 µg/ml DAPI, and kept on ice before sorting. Nuclear lysates processed identically
630 from non-EdU-injected mice served as negative controls during sorting with Beckman Coulter MoFlo Cell

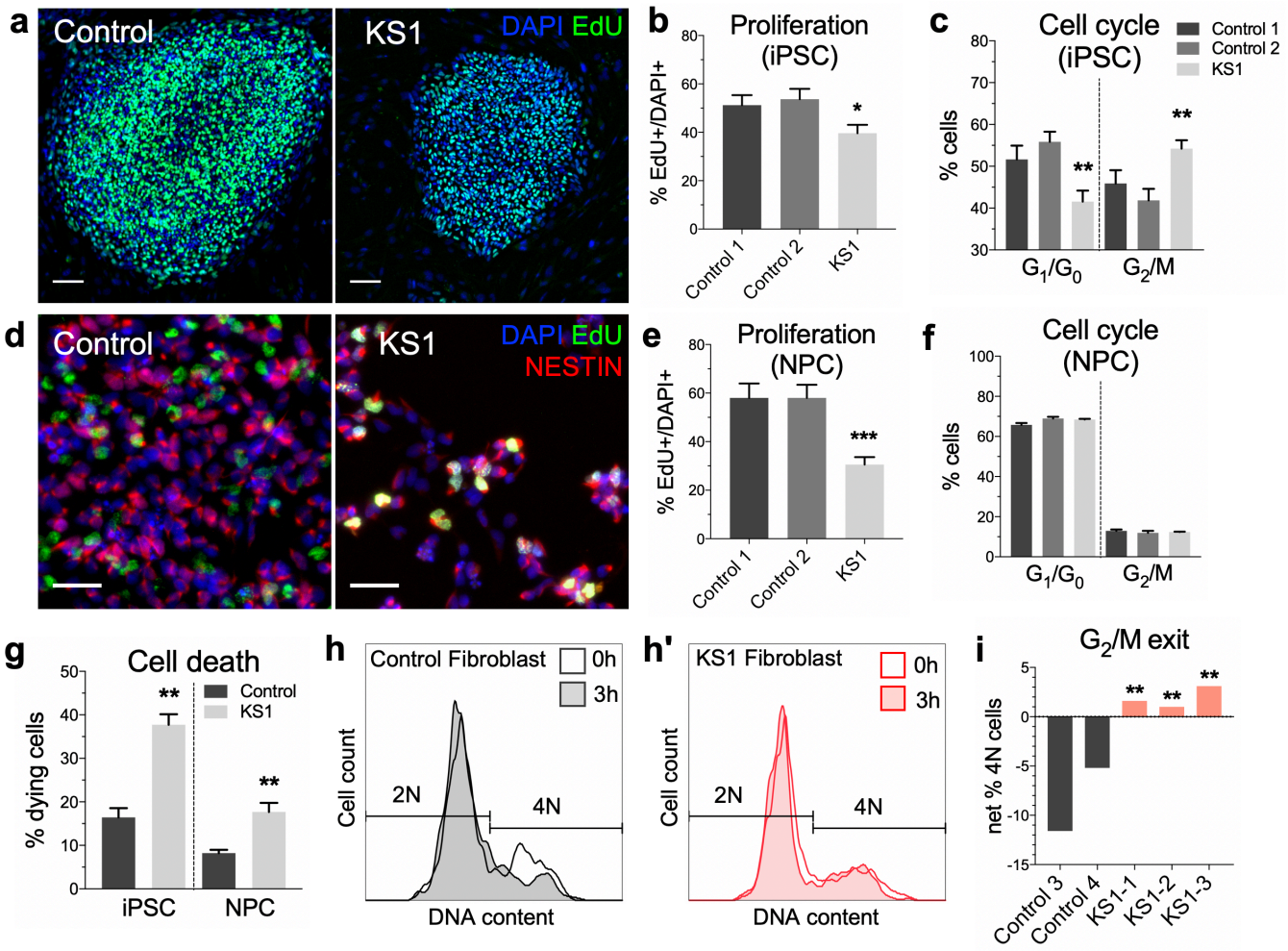
631 Sorter as above. Cell cycle analysis by DNA content was performed with analysis gates discriminating
632 2N and 4N cells by DAPI fluorescence.
633
634 **RNA-seq: EdU$^+$ nuclei**
635 Purified EdU$^+$ nuclei from *Kmt2d*$^{+/\beta geo}$ female mice (n=3) and sex- and age-matched littermates (n=3)
636 were sorted into Smart-Seq 2 lysis buffer (2 μL Smart-Seq2 lysis buffer with RNase inhibitor, 1 μL oligo-
637 dT primer, and 1 μL dNTPs), briefly spun by tabletop microcentrifuge, and snap-frozen on dry ice. Lysates
638 were stored at -80°C until cDNA conversion. Nuclei were processed according to a modified Smart-seq2
639 protocol[53]. Briefly, lysates were thawed to 4°C, heated to 72°C for 5 minutes, and immediately placed on
640 ice. Template-switching first-strand cDNA synthesis was performed using a 5'-biotinylated TSO oligo.
641 cDNAs were amplified using 22 cycles of KAPA HiFi PCR and 5'-biotinylated ISPCR primer. Amplified
642 cDNA was cleaned using 1:1 ratio of Ampure XP beads and approximately 250 pg was input to a one-
643 quarter-sized Nextera XT tagmentation reaction. Tagmented fragments were amplified for 20 enrichment
644 cycles and dual indexes were added to each well to uniquely label each library. Concentrations were
645 assessed with Quant-iT PicoGreen dsDNA Reagent (Invitrogen) and samples were diluted to ~2nM and
646 pooled. Pooled libraries were sequenced on the Illumina HiSeq 2500 platform to a target mean depth of
647 ~8 x 105 bp paired-end fragments per cycle at the Johns Hopkins Genetic Resources Core Facility. For
648 all libraries, paired-end reads were aligned to the mouse reference genome (mm10) using HISAT2[54] with
649 default parameters except: -p 8. Aligned reads from individual samples were quantified against a
650 reference genome (GENCODE vM8)[55]. Quantification was performed using cuffquant with default
651 parameters and the following additional arguments: --no-update-check –p 8. Normalized expression
652 estimates across all samples were obtained using cuffnorm with default parameters[56].
653
654 **RT-qPCR**
655 Total RNA was isolated by RNeasy Mini (Qiagen) and cDNA libraries were constructed with High-
656 Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to manufacturer protocols.
657 Probes were from Taqman (ThermoFisher Scientific, **Table S10**).
658
659 **Immunostaining, confocal imaging, and processing**
660 Coronal brain sections (30 µm) through the entire dentate gyrus (every sixth slice) were maintained in
661 serial order. Brains were paraformaldehyde-fixed by transcardial perfusion and post-fixed in 4% PFA
662 overnight at 4°C before cryoprotection by sequential overnight incubations at 4°C with sucrose
663 concentrations of 10%, 20% and 30% in phosphate buffer. Brains were sectioned by cryostat (Leica),

664 directly mounted to charged slides, and stored at -80°C. Antigen retrieval (DakoCytomation) was
665 performed at 95°C for 20 minutes. After blocking, overnight incubation at 4°C in primary antibodies (**Table**
666 **S10**) was followed by 2-hour room temperature incubation with AlexaFluor-conjugated secondary
667 antibodies at 1:500 dilution. Blocking buffer of 0.1% Triton X-100 in TBS contained 6% serum matched
668 to the host species of secondary antibody. Tiled, z-stacked images were acquired using an inverted Zeiss
669 LSM780 FCS AxioObserver confocal microscope and Zen software (Zeiss) to encompass the entire GCL
670 area of each section using 10X, 20X, or 40X objective. Images were quantified using Imaris (BitPlane).
671 Cell counts for each GCL slice analyzed were corrected by the total GCL area multiplied by z-thickness
672 of the image, and expressed as cells/mm^3. DCX$^+$ neuroblast size and distance from SGZ plane were
673 measured in Fiji (NIH) by experimenters blinded to condition and genotype. iPSCs and NPCs were plated
674 at equal density and fixed with 4% PFA for 15 minutes at room temperature, blocked and permeabilized
675 with 0.25% Triton X-100 and 10% donkey serum in PBS, then stained for 1 hour at room temperature in
676 primary antibodies and 1 hour at room temperature in secondary antibodies. iPSC and NPC images were
677 acquired using EVOS FL Cell Imaging System (ThermoFisher Scientific) and analyzed using Fiji (NIH).
678 HT22 cells were cultured on glass coverslips and stained and imaged as above using Zeiss LSM780.
679
680 **FACS and analysis**
681 Flow cytometric analysis was carried out using FACSverse instrument and FACSsuite software (BD
682 Biosciences), and FACS was performed using Beckman Coulter MoFlo Cell Sorter with proper gate
683 settings and doublet discrimination (**Supplementary Fig. 1m**) at the Johns Hopkins Bloomberg School
684 of Public Health Flow Cytometry Core. Cell samples containing at least 10,000 cells were analyzed from
685 technical triplicate culture wells and analysis was performed in FlowJo v10 (Tree Star Inc). Feeder-free
686 iPSCs and NPCs, HT22 cells, or patient fibroblasts were enzymatically dissociated to single cell
687 suspension. For antibody staining, cells were fixed with 4% PFA for 15 minutes at room temperature,
688 blocked and permeabilized with 0.25% Triton X-100 and 10% donkey serum in PBS, and incubated with
689 primary antibody for 30 minutes at room temperature, followed by 30 minutes in secondary antibody and
690 5 minutes in DAPI (1 µg/ml) for DNA content. Unstained and secondary-only samples served as control.
691 For proliferation analysis, cells were sampled after 30-minute pulse of EdU (10 µM) using Click-iT EdU
692 Flow Cytometry Assay (ThermoFisher Scientific) according to manufacturer protocol. Proteostat reagent
693 (Enzo Life Sciences), CellTrace Violet, and CellEvent caspase-3/7 reagent (ThermoFisher Scientific)
694 were used according to manufacturer protocol. Proteasome inhibition was achieved using 5 µM MG-132
695 for 16 hours (Enzo Life Sciences). Cell cycle synchronization in HT22 cells or patient fibroblasts was

696 achieved by treatment with 250 ng/ml nocodazole (Sigma) for 18 hours before being released into
697 complete media free of nocodazole for indicated time points.
698
699 **Magnetic Resonance Imaging (MRI)**
700 3D T2-weighted MRI (9.4T) was performed in mouse brains following PFA perfusion-fixation by the JHU
701 Department of Radiology Division of MR Research. Atlas-based volumetric analysis was performed in 25
702 brain regions, including hippocampus and olfactory bulb. Structural and ventricular volumes were
703 corrected by total brain volume for analysis (DtiStudio).
704
705 **Acknowledgments and disclosures**
706 HTB is funded through an Early Independence Award from the National Institutes of Health (NIH,
707 DP5OD017877) and through the Louma G. Foundation. Imaging was performed at the Johns Hopkins
708 School of Medicine Microscope Facility with NIH support (S10OD016374). Karyotypes were analyzed
709 by the Johns Hopkins Cytogenomics Laboratory with support from NICHD (1 U54 HD079123-01A1).
710 HTB is a consultant for Millennium Pharmaceuticals. We thank Michael Sherman for assistance with
711 image quantification, Manisha Aggarwal for MRI analysis, Mark Sandusky for schematic designs.
712 Hongjun Song and Kai Ge provided critical reagents and advice. Hal Dietz and Gregg Semenza
713 provided indispensable guidance.
714
715 **Author contributions**
716 GAC and HTB conceived the study; GAC and HTB wrote the manuscript; GAC, HNN, GC, JDR, LZ
717 performed experiments; GAC, LB, JA, KDH and LG analyzed data.

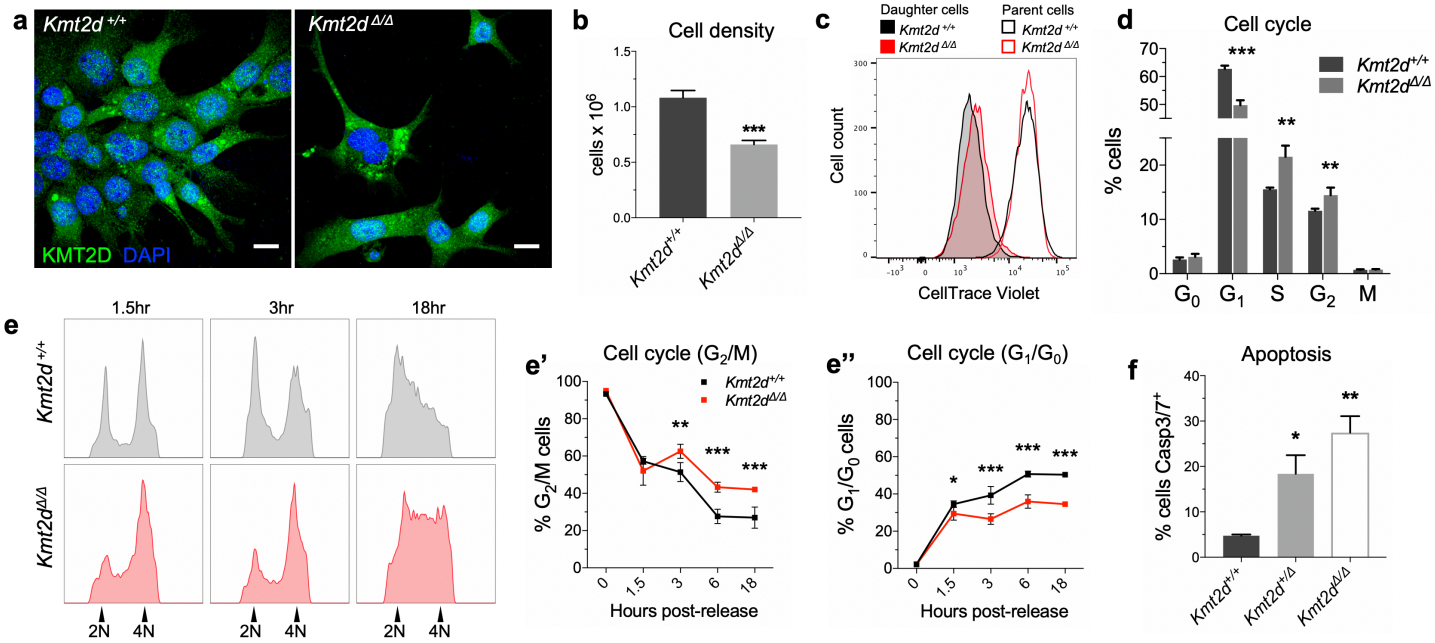
718 **Supplementary Figures**
719 Supplementary Figure 1: iPSC and NPC validations and additional phenotyping
720 Supplementary Figure 2: CRISPR-targeted HT22 cells
721 Supplementary Figure 3: iPSC and NPC single-cell RNA-seq analysis
722 Supplementary Figure 4: HT22 cell RNA-seq analysis
723 Supplementary Figure 5: Additional functional studies of hypoxia response
724 Supplementary Figure 6: Additional phenotyping of *Kmt2d*$^{+/\beta geo}$ mice
725
726 **Supplementary Tables**
727 Supplementary Table 1: Differential gene expression in iPSCs (KS1 versus controls)
728 Supplementary Table 2: Differential gene expression in NPCs (KS1 versus controls)
729 Supplementary Table 3: Intersected DEG sets in iPSCs and NPCs
730 Supplementary Table 4: DEGs in NPCs stratified by cell cycle phase subset
731 Supplementary Table 5: Differential gene expression in HT22 cells (*Kmt2d*$^{+/+}$ versus *Kmt2d*$^{\Delta/\Delta}$)
732 Supplementary Table 6: KMT2D ChIP-seq peaks in HT22 cells *(Kmt2d*$^{+/+}$ cells)
733 Supplementary Table 7: KMT2D-bound genes in HT22 cells *(Kmt2d*$^{+/+}$ cells)
734 Supplementary Table 8: MRI volumetric comparisons in wild-type and *Kmt2d*$^{+/\beta geo}$ mice
735 Supplementary Table 9: Differential gene expression in EdU+ DG nuclei of *Kmt2d*$^{+/\beta geo}$ mice
736 Supplementary Table 10: Reagents

Fig. 1: KS1 patient-derived cells exhibit defects in proliferation, cell cycle progression, and survival



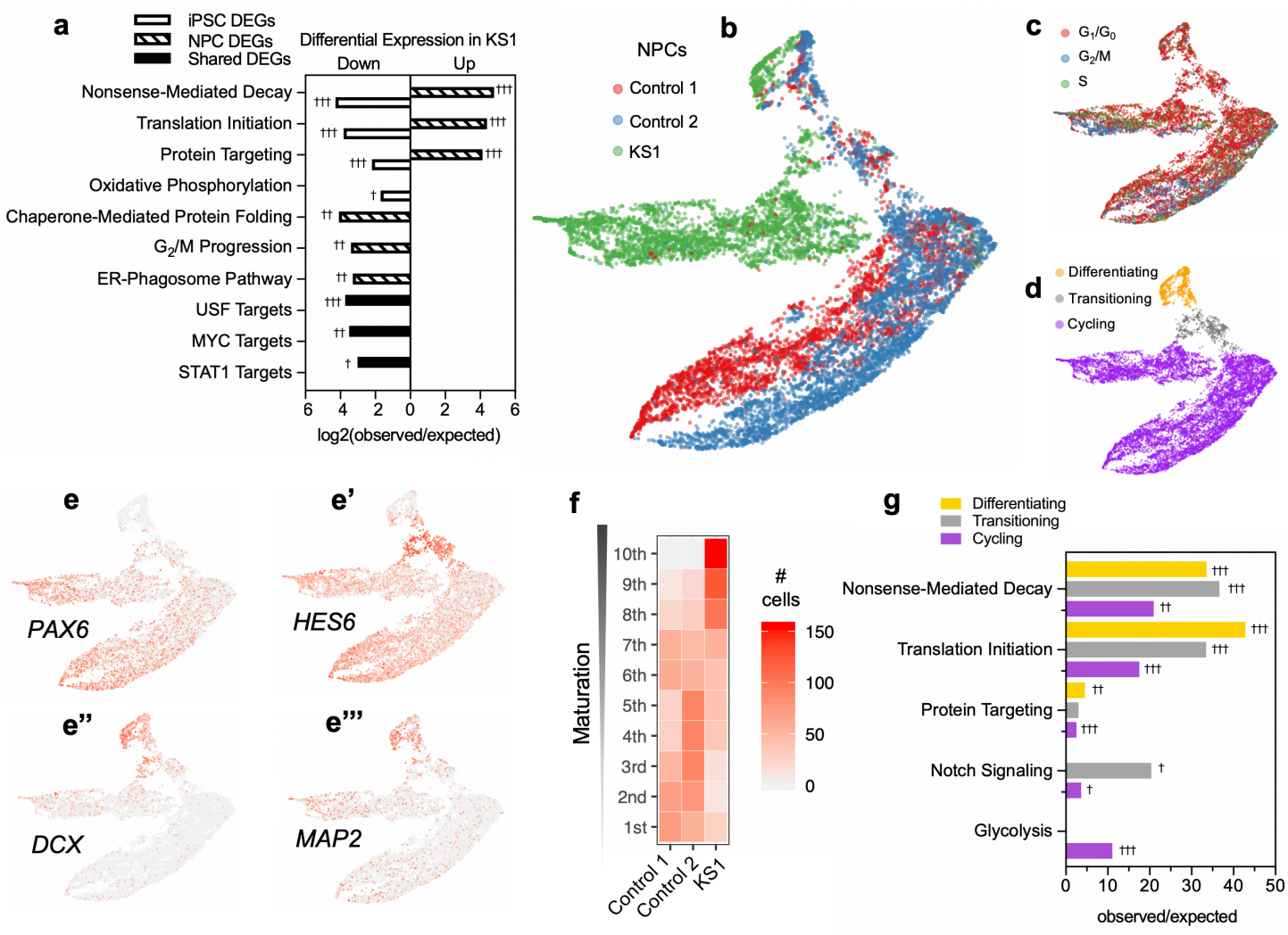
a, **b**, Representative immunostaining of iPSCs from KS1 patient (c.7903C>T:p.R2635*) and healthy controls showing normal colony morphology (**a**). Proliferative cells were pulsed with EdU for 30 minutes and quantified by flow cytometry (**b**). **c**, Cell cycle analysis in iPSCs, discriminating 2N and 4N DNA content (G_1/G_0 and G_2/M, respectively) by flow cytometry using DAPI fluorescence. **d**, Representative immunostaining of NPCs induced from iPSCs of KS1 patient and controls, showing expected morphology, NESTIN expression, and proliferative cells (EdU^+). Proliferative cells were pulsed with EdU for 30 minutes and quantified by flow cytometry (**e**). **f**, Cell cycle analysis in NPCs, discriminating 2N and 4N DNA content (G_1/G_0 and G_2/M, respectively) by flow cytometry using DAPI fluorescence. **g,** Quantification of dying cells by flow cytometric scatter profiles in iPSCs and NPCs from KS1 patient and controls. **h-i**, Cell cycle analysis by flow cytometry in fibroblasts from healthy controls and KS1 patients (KS1-1, KS1-2, KS1-3) using DAPI fluorescence (**h-h'**), and quantification at 0 and 3 hours after nocodazole enrichment of G_2/M phase cells (**i**). Bars indicate mean ± SEM. Asterisks indicate significance from control mean, student's t-test (*$p<0.05$, **$p<0.01$, ***$p<0.001$).

Fig. 2: Deletion of the KMT2D SET methyltransferase domain in isogenic hippocampal cells recapitulates KS1-associated defects in proliferation, cell cycle, and survival



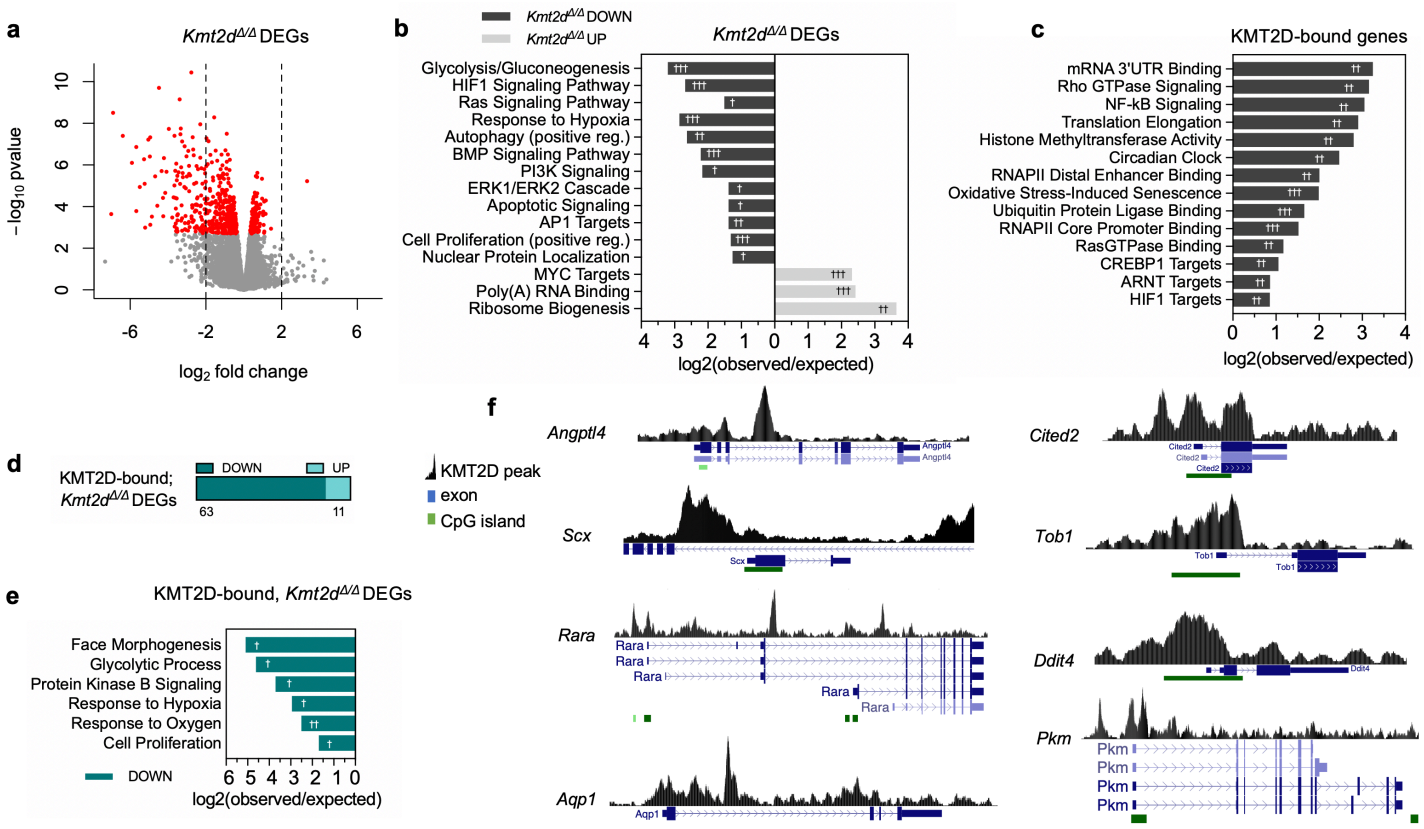
HT22 mouse hippocampal cells targeted by CRISPR-Cas9 to delete the catalytic SET domain without degradation of total *Kmt2d* transcript. **a**, Representative immunostaining confirms KMT2D protein in wild-type cells (*Kmt2d*$^{+/+}$) and homozygous deletion (*Kmt2d*$^{\Delta/\Delta}$) cell lines. **b**, Quantification of cell counts at 72 hours following equal density plating of *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells. **c**, Analysis of cell division with CellTrace Violet dye in *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells over 72 hours. Less dye dilution in daughter cells at 72 hours (i.e. higher mean fluorescence intensity) indicates fewer cell divisions. No difference in fluorescence seen at the 0-hour dye uptake control timepoint. **d**, Flow cytometric quantification of cell cycle composition at steady state in HT22 cells using flow cytometry against Ki67 and DAPI. **e**, Time course analysis of cell cycle by DNA content from 0 to 18 hours after release from G_2/M arrest with nocodazole, in *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells, and quantification of cells in G_1/G_0 (**e**') and G_2/M (**e**''). **f**, Flow cytometric quantification of early apoptotic cells using fluorescence coupled to caspase-3/7-mediated peptide cleavage in *Kmt2d*$^{+/+}$, *Kmt2d*$^{+/\Delta}$, and *Kmt2d*$^{\Delta/\Delta}$ cells. Bars indicate mean ± SEM. Asterisks indicate significance from control mean, student's t-test (*p<0.05, **p<0.01, ***p<0.001).

Fig. 3: Transcriptional suppression of metabolic genes in cycling cells, and precocious pro-neural gene expression in differentiating cells among KS1 patient-derived NPCs



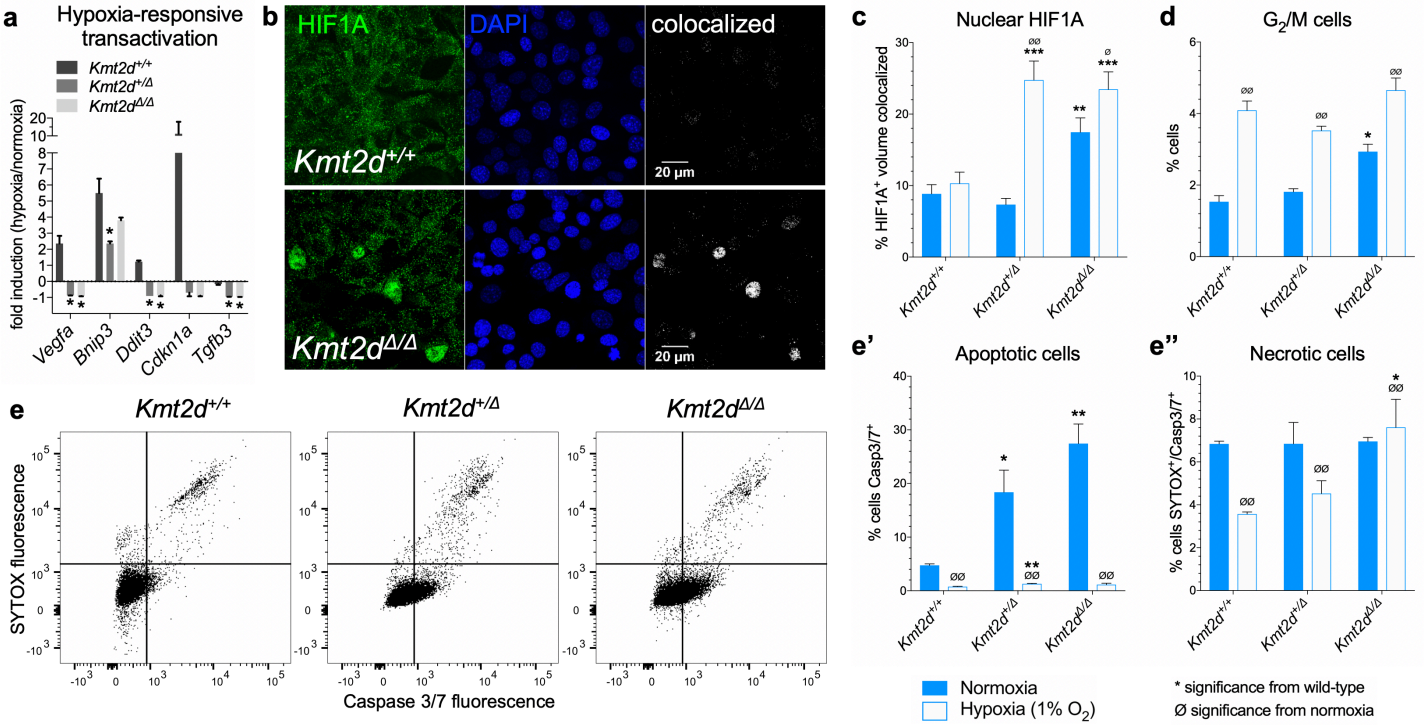
Single-cell RNA-seq profiling in patient-derived cells. **a**, Gene networks most enriched among differentially expressed genes (DEGs) in KS1 patient iPSCs and NPCs relative to respective healthy controls, and DEGs shared in both cell types. **b**, **c**, **d**, Uniform Manifold Approximation Projection analysis of NPCs, clustered to visualize similarity of gene expression between cells. Individual NPCs identified by Patient ID barcodes (**b**) were partitioned into subsets according to cell cycle phase as defined by phase-specific transcription (**c**), and NPC maturation status as defined by stage-specific marker transcription (**d**). **e**, UMAPs annotated by expression intensities of NPC markers to reveal the maturation trajectory. Representative genes (*PAX6, HES6, DCX, MAP2*) demonstrate progression of marker intensities characteristic of neurodevelopment from early NPC (**e**) to differentiating NPC (**e'''**). **f**, Heatmap comparing density of differentiating NPCs in KS1 and controls along the maturation trajectory, using marker expression to define binned deciles from earliest (1st) to most differentiated (10th) decile. **g**, Differential gene expression in distinct NPC subsets (cycling, transitioning, and differentiating), analyzed to identify the most enriched gene networks among DEGs in each subset († FDR<0.05, †† FDR<0.01, ††† FDR<0.001).

Fig. 4: Globally suppressed transcription upon loss of the KMT2D SET methyltransferase domain, with enrichment of KMT2D-bound genes in glycolytic and oxygen-sensing pathways



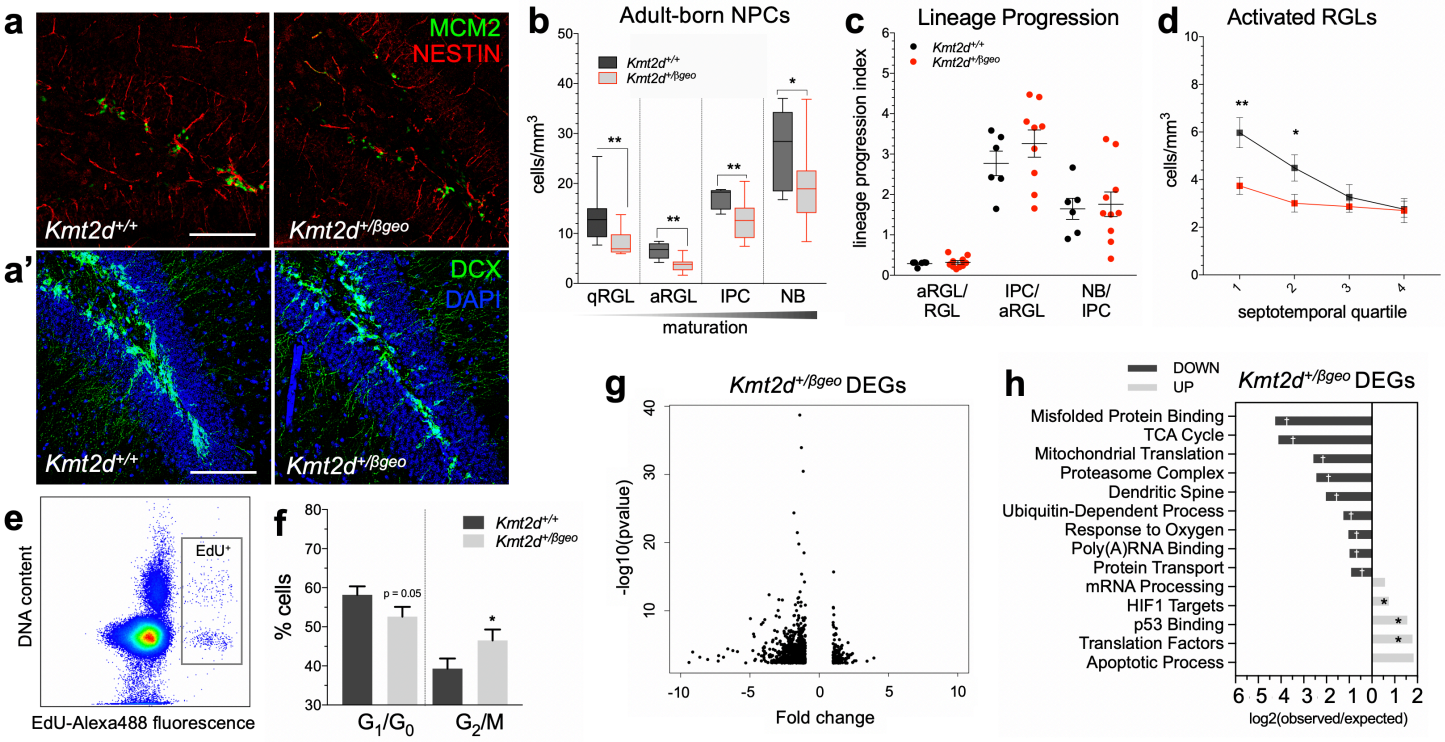
a, **b**, RNA-seq analysis of differential expression in *Kmt2d*$^{\Delta/\Delta}$ HT22 mouse hippocampal cells relative to *Kmt2d*$^{+/+}$ cells. Volcano plot illustrates skewing of DEGs (~76%) toward downregulation in *Kmt2d*$^{\Delta/\Delta}$ cells (**a**) and gene networks showing highest enrichment among DEGs down- or up-regulated in *Kmt2d*$^{\Delta/\Delta}$ cells relative to *Kmt2d*$^{+/+}$ (**b**). **c-f**, ChIP-seq chromatin profiling of global KMT2D binding sites in wild-type HT22 cells. Gene networks showing highest fold change in enrichment among genes proximal to KMT2D peaks (<5 kb to promoter, (**c**)). Intersection of genes observed to be both KMT2D-bound and differentially-expressed in *Kmt2d*$^{\Delta/\Delta}$ cells relative to wild-type, with proportions of genes up- and down-regulated in mutants (**d**) and gene networks showing highest enrichment among KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ DEGs (**e**). Representative KMT2D ChIP-seq read densities at KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ down-regulated genes, including oxygen-sensing pathway genes *Angptl4, Rara, Ddit4, Pkm, Aqp1,* and *Cited2*, annotated with KMT2D binding peaks (black), RefSeq gene annotations (blue), and CpG islands (green) (**f**) († FDR<0.05, †† FDR<0.01, ††† FDR<0.001).

815 **Fig. 5: *Kmt2d*-dependent HIF1A regulation and transactivation, and cell survival in response to**
816 **hypoxia**



817
818
819 Analysis of hypoxia response in *Kmt2d*$^{+/+}$, *Kmt2d*$^{+/\Delta}$, and *Kmt2d*$^{\Delta/\Delta}$ cells, following 72 hours in normoxic
820 (20% O_2) or hypoxic (1% O_2) conditions. **a**, Fold induction of HIF1A target gene mRNA, in hypoxia
821 relative to normoxia, by RT-qPCR analysis (*significance from wild-type induction; negative fold
822 changes plotted as negative reciprocal). **b**, HIF1A nuclear colocalization analysis. Representative Z-
823 stacked confocal images of *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells under normoxic condition are shown (**b**) with
824 quantifications of colocalized HIF1A^{+} and DAPI^{+} voxels (**c**), and %G_2/M phase cells (**d**) in all conditions.
825 **e**, Flow cytometric quantifications of cell death, with CellEvent fluorescence detecting caspase-3/7-
826 mediated peptide cleavage in apoptotic cells (**e'**), and SYTOX reagent for necrotic cells (**e''**). Bars
827 indicate mean ± SEM, student's t-test (*$p<0.05$, **$p<0.01$, ***$p<0.001$).
828

Fig. 6: *Kmt2d*-deficient mice exhibit depletion of adult-born neural precursors, with cell cycle defects and transcriptional suppression of cellular metabolic pathways



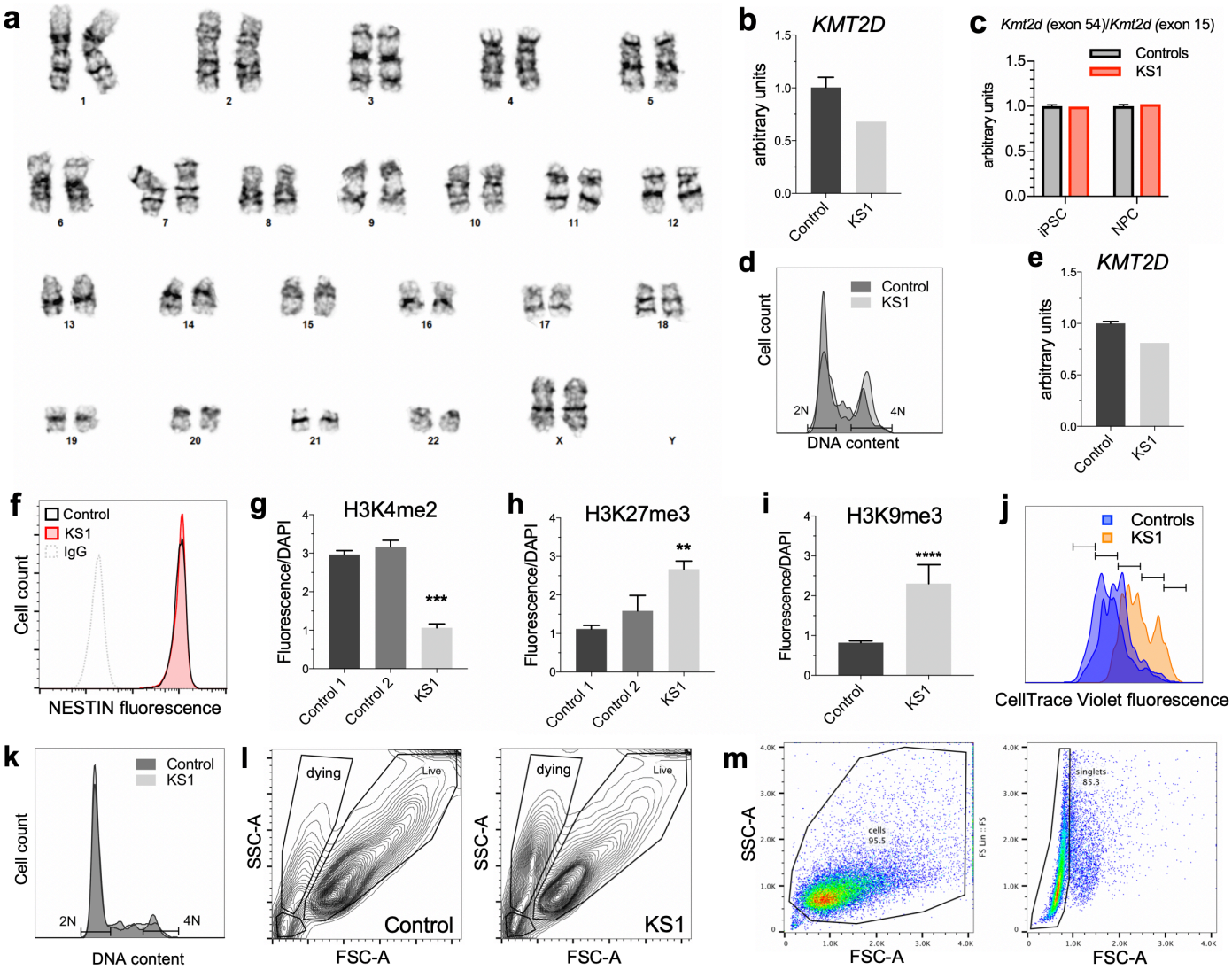
a, **a'**, Representative immunostaining of neurogenesis markers in the dentate gyrus (DG) of adult *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\beta/geo}$ mice. Earliest stages of NESTIN^{+} radial glia-like cells (RGLs) in either quiescent (MCM2^{-}) or activated (MCM2^{+}) states (qRGL and aRGL, respectively), MCM2^{+}/NESTIN^{-} intermediate progenitor cells (IPCs) (**a**), and DCX^{+} neuroblasts (NBs) (**a'**). **b-d**, Quantification of qRGL, aRGL, IPC, and NB cell densities in adult *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\beta/geo}$ mice (**b**). Comparison of lineage progression index, an approximation of expansion potential for each cell type transition (**c**), indicates absence of genotype-associated blockages at any particular cell-type transition analyzed. **d**, Stratified quantification of aRGL density along the septotemporal axis of the DG in *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice, indicating preferential disruption at the septal DG, a brain region linked to visuospatial memory. **e**, **f**, Flow cytometric purification of cycling nuclei from microdissected DG of *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice sampled 16 hours after EdU pulse (**e**). Cell cycle analysis in EdU^{+} nuclei of G_1/G_0 and G_2/M phase using DAPI fluorescence (**f**). **g**, **h,** RNA-seq analysis of differential gene expression in purified EdU^{+} DG nuclei from *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\beta/geo}$ mice. Gene networks most enriched among DEGs down- or up-regulated in *Kmt2d*$^{\beta/geo}$ nuclei, showing transcriptional suppression of cellular metabolic pathways and upregulation of HIF1 target genes. Bars indicate mean ± SEM. Asterisks indicate significance from control mean, student's t-test (*p<0.05, **p<0.01, ***p<0.001), († FDR<0.05, †† FDR<0.01, ††† FDR<0.001).

852 **Table 1**: KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ down DEGs in hippocampal cells. Positive fold changes (0 to 1.1)
853 indicate genes up-regulated and KMT2D-bound. Negative fold changes (-5 to 0) indicate genes down-
854 regulated and KMT2D-bound. *previously described KMT2D target genes[7,25]
855

Fold change	Number of genes	Genes
0 to 1.1	11	*Dazap1, Psmc2, Eif3b, Washc4, Ran, Lrpprc, Imp4*, Sf3a2*, Bend3, Hspd1*, Dusp4*
-1 to 0	37	*Pik3ip1*, Ssc5d, Plekha1, Tinagl1*, Slc29a1, Rras, Nudt18, Slc25a29, F3, Pgpep1, Lrrn4cl, Gpr146, Aldoa*, Dbp, Cpq, Pkm, Socs3, Map3k14*, Rdm1, Rara, Nfkbia*, Klf10*, Tob1, Irf2bp2, Fosl2, Ddit4*, Trim47, Btbd1, Plin2, Mxra8, Spg20, Cited2*, S100a11, Il34, Bnip3l*, Nfil3*, Shb*
-2 to -1	16	*Ccl9, Crip2, Msln, Cp*, Scx, Klhl30*, Vdr*, Tnfsf13, Fcgr4, Emilin1*, Oaf, Aqp1, C1qtnf1, Baalc, Macrod1, Fblim1*
-3 to -2	8	*Ly6c1, Ly6a, Lctl, Plin4, Angptl4*, Tns4, Pkp3, Igf1**
-5 to -3	2	*Gjb4*, Mfap4*

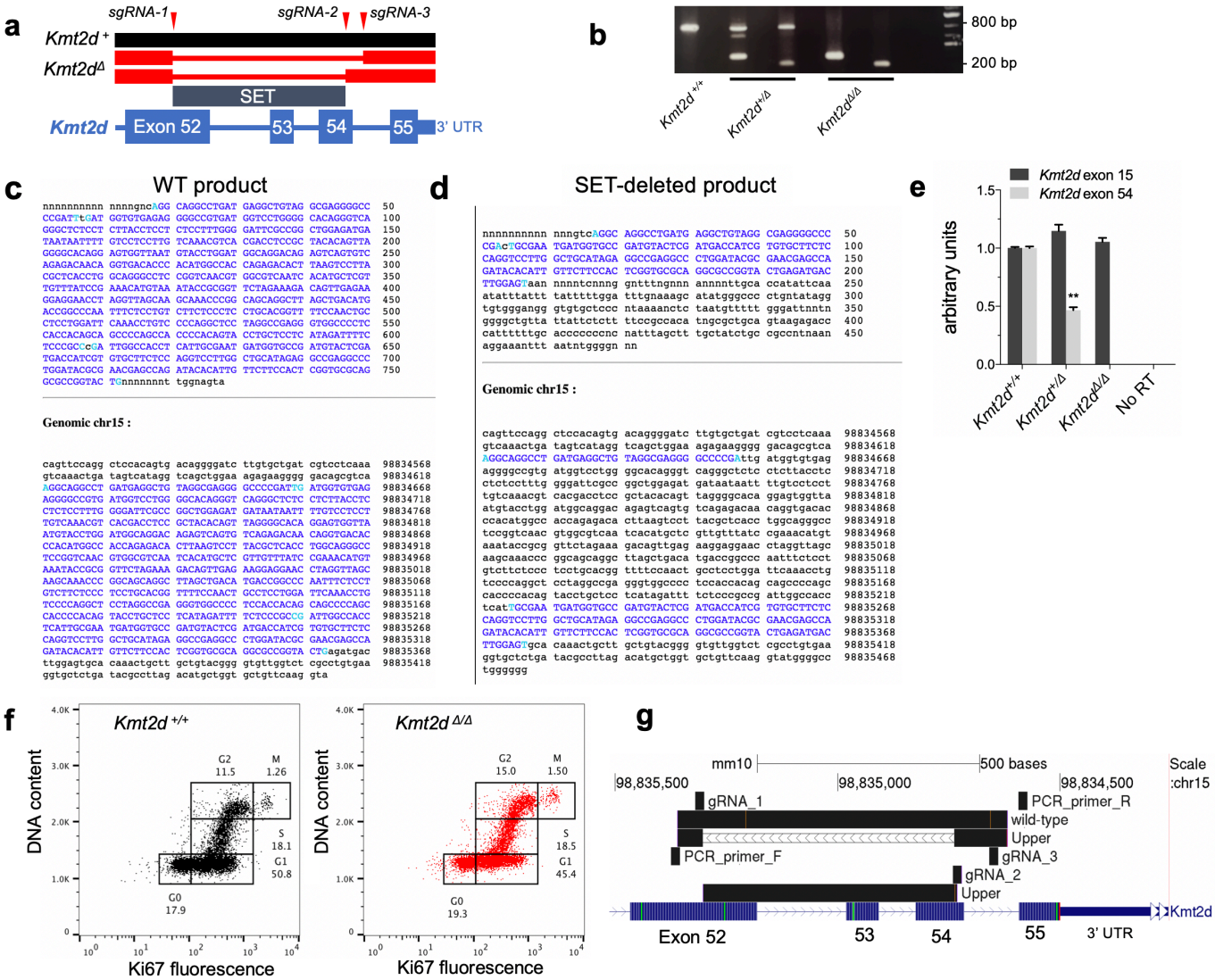
856
857

Supplementary Fig.1: iPSC and NPC line validations and additional phenotyping



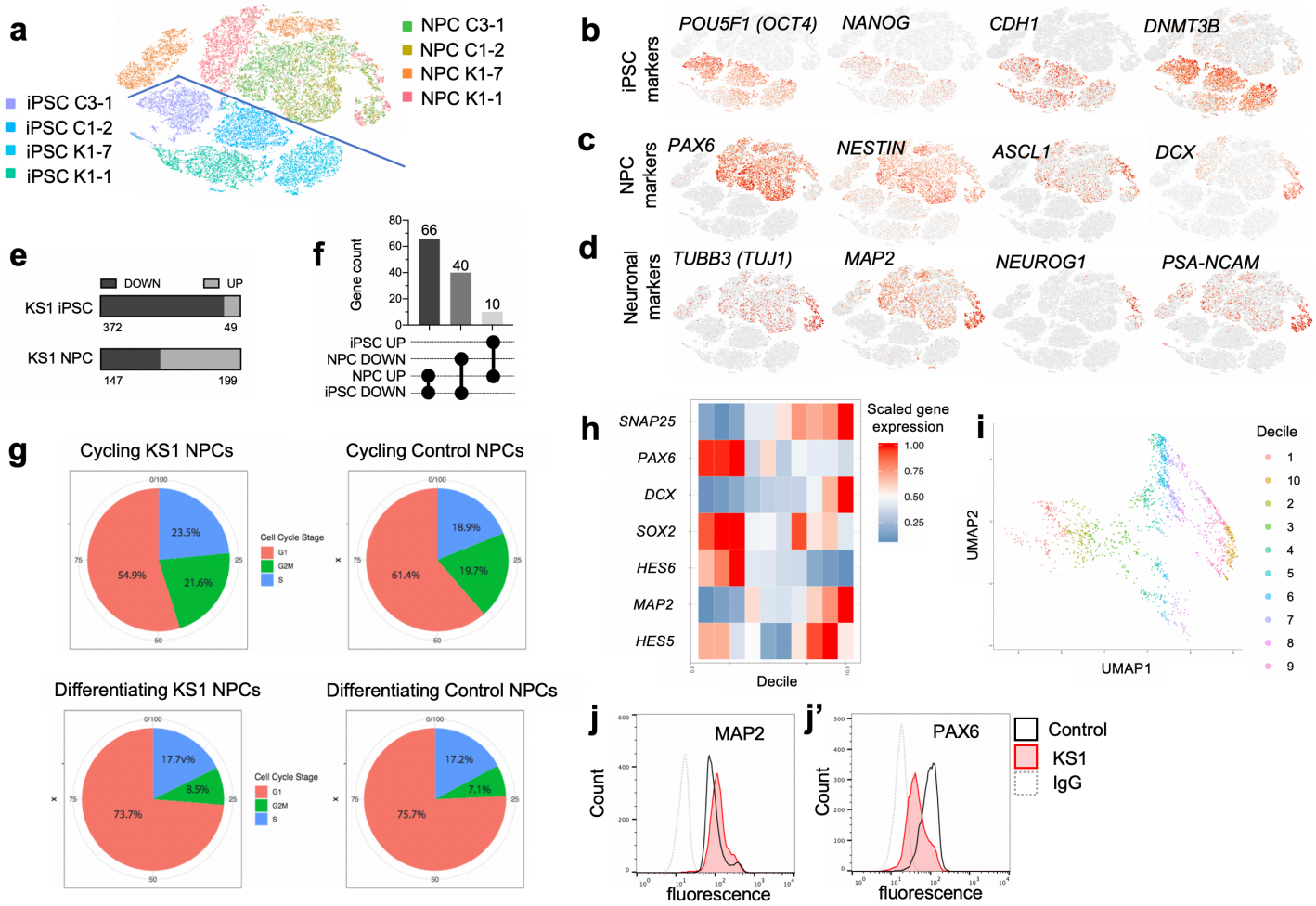
a, 46, XX normal female karyotype in KS1-1 iPSCs **b**, RT-qPCR analysis of *KMT2D* (exon 15) expression in KS1 iPSCs compared to two healthy control iPSC lines (C1-2 and C3-1). **c**, RT-qPCR demonstrating equivalent exonic ratios of KMT2D exon 15 and exon 54. **d**, Flow cytometric analysis of DNA content by DAPI fluorescence in iPSCs. **e**, RT-qPCR analysis of *KMT2D* (exon 15) expression in NPCs derived from the KS1 and control iPSC lines. **f**, Flow cytometric analysis of NESTIN fluorescence intensity in KS1 and control NPCs. **g-i,** Immunofluorescence analysis of open chromatin mark H3K4me2 (**g**) and closed chromatin marks H3K27me3 (**h**) and H3K9me3 (**i**) in KS1 and control NPCs. **j**, CellTrace Violet generational tracking showing fewer divisions (i.e. higher dye intensity) in patient-derived NPCs over 72 hours. **k**, Flow cytometric analysis of DNA content by DAPI fluorescence in NPCs. **l**, Sample flow cytometric gating for detection of scatter profiles indicative of cell death-associated cellular condensation. **m**, representative gating of viable cells and doublet discrimination in immunofluorescence-based flow cytometric analyses of iPSCs and NPCs.

Supplementary Fig. 2: CRISPR-targeted HT22 cells



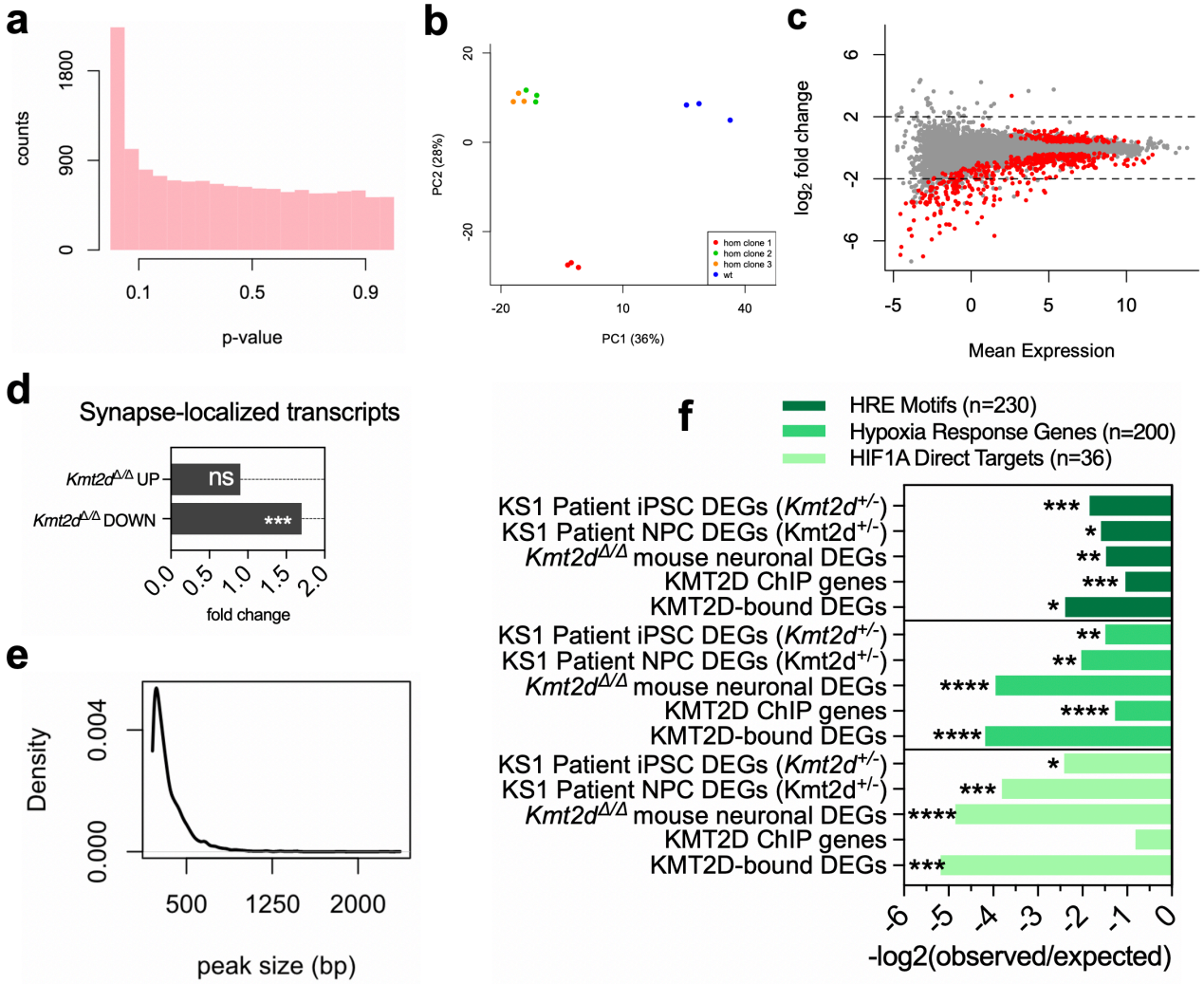
a, Schematic depiction of CRISPR-Cas9 strategy to delete a region spanning the *Kmt2d* SET domain in the HT22 mouse hippocampal cell line. upstream sgRNA-1 (exon 52) was combined with either sgRNA-2 (exon 54) or sgRNA-3 (intron 54) creating deletions of 570 or 660 bp, respectively. **b**, PCR with probes flanking sgRNA cut sites confirms deletion in experimental cell lines (*Kmt2d* $^{+/\Delta}$ and *Kmt2d*$^{\Delta/\Delta}$) compared to wild-type (*Kmt2d*$^{+/+}$). **c**, **d**, Sanger sequences of wild-type (*Kmt2d*$^{+}$) and targeted (*Kmt2d*$^{\Delta}$) alleles in HT22 cells, mapped to *Kmt2d* locus (mm10) on mouse chromosome 15 using UCSC Genome Browser. **e**, RT-qPCR analysis of *Kmt2d* mRNA in HT22 cell lines. Comparison of exons upstream and within the deletion site (exons 15 and 54) to demonstrate deletion, but not degradation of the *Kmt2d* transcript. **f**, Cell cycle gating in flow cytometric analysis using Ki67 and DAPI to discriminate individual stages (G_0, G_1, S, G_2, M) in *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells. **g,** Cells bearing strand invasion in targeted alleles ("Upper" PCR band tracks mapping to plus and minus strand) were excluded from functional analyses.

Supplementary Fig. 3: iPSC and NPC single-cell RNA-seq analysis



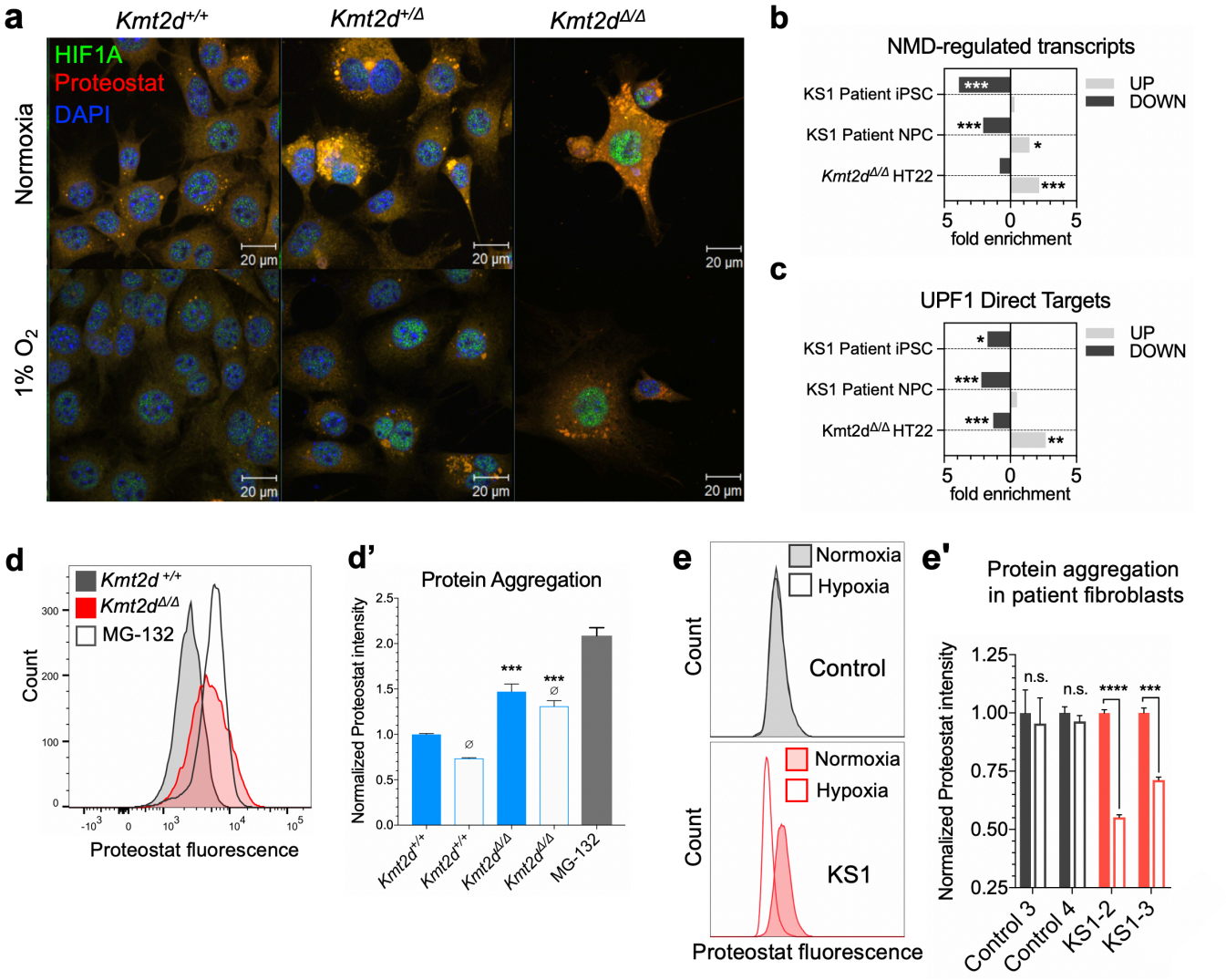
a, t-stochastic neighbor embedding (tSNE) representation of iPSC and NPC libraries sequenced on 10XGenomics platform. Cell clusters colored by cell type and Patient ID. iPSCs and NPCs derived from patient K1-7 were excluded from downstream analysis due to abnormal karyotype. **b**, **c**, **d**, Representative tSNE of iPSC, NPC, and neuronal markers demonstrating expected cell identities and revealing a gradient of cell maturation. **e**, **f**, Proportions of DEGs down- or up-regulated in KS1 patient iPSCs or NPCs compared to respective healthy controls and DEG lists intersected for overlaps among down-regulated and up-regulated genes (**f**). **g**, Cell cycle phase proportions in cycling versus differentiating NPCs from KS1 patient and controls. **h, i,** transcriptional marker-based partitioning of NPCs into bins (deciles) of increasing maturation. **j**, Flow cytometric validation with selected markers in NPCs from KS1 patient and controls. Fluorescence intensities of early NPC marker *PAX6* (**j**) and mature NPC marker *MAP2* (**j'**) are shown.

Supplementary Fig. 4: HT22 cell RNA-seq and ChIP-seq analysis



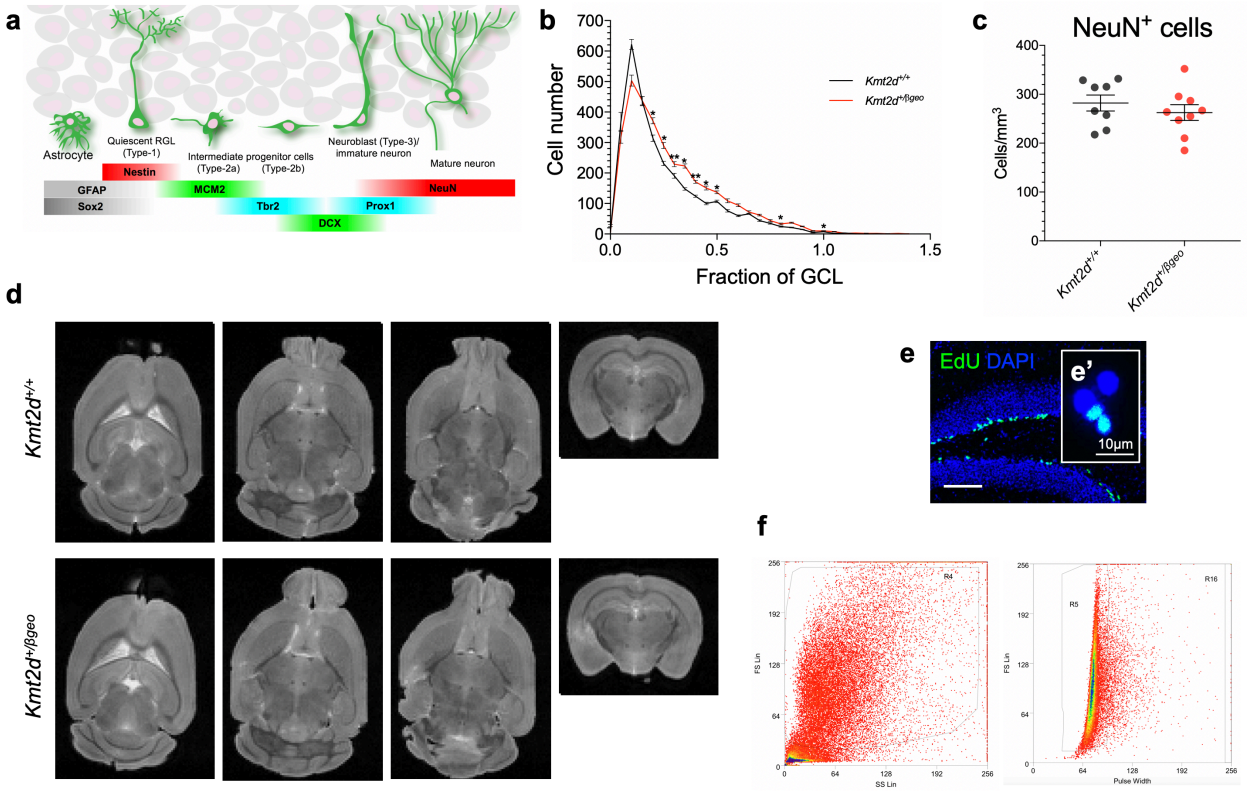
a, P-value distribution in *Kmt2d*$^{\Delta/\Delta}$ DEGs relative to wild-type determined by RNA-seq. **b**, PCA visualizing expression differences in wild-type and *Kmt2d*$^{\Delta/\Delta}$ HT22 cells, illustrating separation of Clone 1 which bears strand invasion. **c**, MA plot illustrating relationships between average expression and log2 fold change for DEGs. **d**, Significant enrichment of *Kmt2d*$^{\Delta/\Delta}$ DEGs among published synapsed-localized transcripts[24]. **e,** KMT2D ChIP-seq analysis comparing KMT2D peak density and peak size (bp) in wild-type HT22 cells. **f**, Significant enrichments of Hypoxia Response genes, HIF1A Direct Target genes, and genes containing the Hypoxia Response Element (HRE) RCGTG motif among observed DEGs in KS1 Patient iPSCs, KS1 Patient NPCs, *Kmt2d*$^{\Delta/\Delta}$ HT22 cells, as well as genes KMT2D-bound in *Kmt2d*$^{+/+}$ HT22 cells, and KMT2D-bound, down-regulated genes in *Kmt2d*$^{\Delta/\Delta}$ HT22 cells). Fisher's Exact Test (*p<0.05, **p<0.01, ***p<0.001).

Supplementary Fig. 5: Additional functional studies of hypoxia response in HT22 cells



a, Representative confocal images of *Kmt2d*$^{+/\Delta}$, *Kmt2d*$^{\Delta/\Delta}$, and wild-type HT22 cells in normoxic and hypoxic (72 hours with 1% O_2) conditions. Immunofluorescence against HIF1A for nuclear localization studies and Proteostat for analysis of misfolded protein aggregates. **b**, **c**, Significant enrichment of NMD-regulated transcripts[31] (**b**) and direct UPF1 targets (**c**) among observed DEGs in patient-derived cells and HT22 cells. Fisher's Exact Test (*p<0.05, **p<0.01, ***p<0.001). **d-e** Flow cytometric analysis of Proteostat fluorescence under normoxic or hypoxic conditions, with quantification, in HT22 cells (**d, d'**) and patient fibroblasts (**e-e'**). Bars indicate mean ± SEM. Asterisks indicate significance from wild-type mean for each condition, student's t-test (*p<0.05, **p<0.01, ***p<0.001). Empty set symbol indicates significance from normoxic condition for each genotype, student's t-test (ø<0.05, øø<0.01).

Supplementary Fig. 6: Additional phenotyping of *Kmt2d*$^{+/\beta geo}$ mice



a, Schematic depicting marker expression during sequential stages of adult neurogenesis in the dentate gyrus (DG). **b**, Quantification of DCX+ cell body distance from SGZ plane in 8-week-old mice. **c**, Quantification of NEUN+ cells in the granule cell layer (GCL) of 8-week-old mice. **d**, Sample images of T2-weighted MRI (9.4T) in PFA-fixed brains of female mice 4 months old. **e, e'**, Confocal images illustrate specificity of EdU uptake in the SGZ niche of the adult DG (**e**) and purified nuclei (**e'**). **f,** Sample FACS gating for viable nuclei and doublet discrimination during purification of cycling EdU^{+} nuclei purified from *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice at 16 hours post-EdU pulse for RNA-seq and cell cycle analysis.

951 Bibliography
952

953 1. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome regulation by polycomb
954 and trithorax: 70 years and counting. *Cell* **171,** 34–57 (2017).
955 2. Bjornsson, H. T. The Mendelian disorders of the epigenetic machinery. *Genome Res.* **25,** 1473–
956 1481 (2015).
957 3. Ng, S. B. *et al.* Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat.*
958 *Genet.* **42,** 790–793 (2010).
959 4. Miyake, N. *et al.* KDM 6 A Point Mutations Cause K abuki Syndrome. *Human mutation* **34,** 108–
960 110 (2013).
961 5. Hannibal, M. C. *et al.* Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome. *Am.*
962 *J. Med. Genet. A* **155A,** 1511–1516 (2011).
963 6. Ang, S.-Y. *et al.* KMT2D regulates specific programs in heart development via histone H3 lysine 4
964 di-methylation. *Development* **143,** 810–821 (2016).
965 7. Ortega-Molina, A. *et al.* The histone lysine methyltransferase KMT2D sustains a gene expression
966 program that represses B cell lymphoma development. *Nat. Med.* **21,** 1199–1208 (2015).
967 8. Zhang, J. *et al.* Disruption of KMT2D perturbs germinal center B cell development and promotes
968 lymphomagenesis. *Nat. Med.* **21,** 1190–1198 (2015).
969 9. Lee, J.-E. *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation
970 during cell differentiation. *Elife* **2,** e01503 (2013).
971 10. Lin-Shiao, E. *et al.* KMT2D regulates p63 target enhancers to coordinate epithelial homeostasis.
972 *Genes Dev.* **32,** 181–193 (2018).
973 11. Froimchuk, E., Jang, Y. & Ge, K. Histone H3 lysine 4 methyltransferase KMT2D. *Gene* **627,** 337–
974 342 (2017).
975 12. Bjornsson, H. T. *et al.* Histone deacetylase inhibition rescues structural and functional brain deficits
976 in a mouse model of Kabuki syndrome. *Sci Transl Med* **6,** 256ra135 (2014).
977 13. Boisgontier, J. *et al.* Anatomical and functional abnormalities on MRI in kabuki syndrome.
978 *NeuroImage: Clinical* (2018). doi:10.1016/j.nicl.2018.11.020
979 14. Gonçalves, J. T., Schafer, S. T. & Gage, F. H. Adult neurogenesis in the hippocampus: from stem
980 cells to behavior. *Cell* **167,** 897–914 (2016).
981 15. Zheng, X. *et al.* Metabolic reprogramming during neuronal differentiation from aerobic glycolysis
982 to neuronal oxidative phosphorylation. *Elife* **5,** (2016).
983 16. Mohyeldin, A., Garzón-Muvdi, T. & Quiñones-Hinojosa, A. Oxygen in stem cell biology: a critical
984 component of the stem cell niche. *Cell Stem Cell* **7,** 150–161 (2010).
985 17. Sanz, J. H., Lipkin, P., Rosenbaum, K. & Mahone, E. M. Developmental profile and trajectory of
986 neuropsychological skills in a child with Kabuki syndrome: implications for assessment of
987 syndromes associated with intellectual disability. *Clin Neuropsychol* **24,** 1181–1192 (2010).
988 18. Wen, Z. *et al.* Synaptic dysregulation in a human iPS cell model of mental disorders. *Nature* **515,**
989 414–418 (2014).
990 19. Li, W. *et al.* Rapid induction and long-term self-renewal of primitive neural precursors from human
991 embryonic stem cells by small molecule inhibitors. *Proc. Natl. Acad. Sci. USA* **108,** 8299–8304
992 (2011).
993 20. Morimoto, B. H. & Koshland, D. E. Induction and expression of long- and short-term
994 neurosecretory potentiation in a neural cell line. *Neuron* **5,** 875–880 (1990).
995 21. Zhou, Y. *et al.* Autocrine mfge8 signaling prevents developmental exhaustion of the adult neural
996 stem cell pool. *Cell Stem Cell* **23,** 444–452.e4 (2018).
997 22. Mao, H. *et al.* Rbm8a haploinsufficiency disrupts embryonic cortical development resulting in

998 microcephaly. *J. Neurosci.* **35,** 7003–7018 (2015).
999 23. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and
1000 projection. *JOSS* **3,** 861 (2018).
1001 24. Merkurjev, D. *et al.* Synaptic N6-methyladenosine (m6A) epitranscriptome reveals functional
1002 partitioning of localized transcripts. *Nat. Neurosci.* **21,** 1004–1014 (2018).
1003 25. Guo, C. *et al.* Global identification of MLL2-targeted loci reveals MLL2's role in diverse signaling
1004 pathways. *Proc. Natl. Acad. Sci. USA* **109,** 17603–17608 (2012).
1005 26. Semenza, G. L. HIF-1: mediator of physiological and pathophysiological responses to hypoxia. *J.*
1006 *Appl. Physiol.* **88,** 1474–1480 (2000).
1007 27. Yoon, K.-J., Vissers, C., Ming, G.-L. & Song, H. Epigenetics and epitranscriptomics in temporal
1008 patterning of cortical neural progenitor competence. *J. Cell Biol.* **217,** 1901–1914 (2018).
1009 28. Davalos, V., Blanco, S. & Esteller, M. Snapshot: messenger RNA modifications. *Cell* **174,** 498–
1010 498.e1 (2018).
1011 29. Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic*
1012 *Acids Res.* **46,** D303–D307 (2018).
1013 30. Baets, J. *et al.* Defects of mutant DNMT1 are linked to a spectrum of neurological disorders. *Brain*
1014 **138,** 845–861 (2015).
1015 31. Hurt, J. A., Robertson, A. D. & Burge, C. B. Global analyses of UPF1 binding and function reveal
1016 expanded scope of nonsense-mediated mRNA decay. *Genome Res.* **23,** 1636–1650 (2013).
1017 32. Benjamin, J. S. *et al.* A ketogenic diet rescues hippocampal memory defects in a mouse model of
1018 Kabuki syndrome. *Proc. Natl. Acad. Sci. USA* **114,** 125–130 (2017).
1019 33. Morris, A. M., Churchwell, J. C., Kesner, R. P. & Gilbert, P. E. Selective lesions of the dentate
1020 gyrus produce disruptions in place learning for adjacent spatial locations. *Neurobiol. Learn. Mem.*
1021 **97,** 326–331 (2012).
1022 34. Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T. & Saxena, S. Prevalence of intellectual
1023 disability: a meta-analysis of population-based studies. *Res. Dev. Disabil.* **32,** 419–436 (2011).
1024 35. Singh, A. M. *et al.* Cell-Cycle Control of Bivalent Epigenetic Domains Regulates the Exit from
1025 Pluripotency. *Stem Cell Rep.* **5,** 323–336 (2015).
1026 36. Whitfield, M. L. *et al.* Identification of genes periodically expressed in the human cell cycle and
1027 their expression in tumors. *Mol. Biol. Cell* **13,** 1977–2000 (2002).
1028 37. Bershteyn, M. *et al.* Human iPSC-Derived Cerebral Organoids Model Cellular Features of
1029 Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell Stem Cell* **20,** 435–449.e4
1030 (2017).
1031 38. Ge, S., Yang, C.-H., Hsu, K.-S., Ming, G.-L. & Song, H. A critical period for enhanced synaptic
1032 plasticity in newly generated neurons of the adult brain. *Neuron* **54,** 559–566 (2007).
1033 39. Gontier, G. *et al.* Tet2 Rescues Age-Related Regenerative Decline and Enhances Cognitive
1034 Function in the Adult Mouse Brain. *Cell Rep.* **22,** 1974–1981 (2018).
1035 40. Luo, W. & Wang, Y. Epigenetic regulators: multifunctional proteins modulating hypoxia-inducible
1036 factor-α protein stability and activity. *Cell Mol. Life Sci.* **75,** 1043–1056 (2018).
1037 41. Chatzi, C., Schnell, E. & Westbrook, G. L. Localized hypoxia within the subgranular zone
1038 determines the early survival of newborn hippocampal granule cells. *Elife* **4,** e08722 (2015).
1039 42. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-
1040 aware quantification of transcript expression. *Nat. Methods* **14,** 417–419 (2017).
1041 43. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level
1042 estimates improve gene-level inferences. [version 2; referees: 2 approved]. *F1000Res.* **4,** 1521
1043 (2015).

1044 44. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential
1045 expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2010).
1046 45. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and
1047 microarray studies. *Nucleic Acids Res.* **43,** e47 (2015).
1048 46. Giner, G. & Smyth, G. K. Statmod: probability calculations for the inverse Gaussian distribution.
1049 *arXiv preprint arXiv:1603.06687* (2016).
1050 47. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-
1051 seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).
1052 48. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by
1053 pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32,** 381–386 (2014).
1054 49. Lun, A. T. L., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of
1055 single-cell RNA-seq data with Bioconductor. [version 2; referees: 3 approved, 2 approved with
1056 reservations]. *F1000Res.* **5,** 2122 (2016).
1057 50. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–
1058 359 (2012).
1059 51. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
1060 52. Rizzardi, L. *et al.* Neuronal brain region-specific DNA methylation and chromatin accessibility are
1061 associated with neuropsychiatric disease heritability. *BioRxiv* (2017). doi:10.1101/120386
1062 53. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat.*
1063 *Methods* **10,** 1096–1098 (2013).
1064 54. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory
1065 requirements. *Nat. Methods* **12,** 357–360 (2015).
1066 55. Mudge, J. M. & Harrow, J. Creating reference gene annotation for the mouse C57BL6/J genome
1067 assembly. *Mamm. Genome* **26,** 366–378 (2015).
1068 56. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with
1069 TopHat and Cufflinks. *Nat. Protoc.* **7,** 562–578 (2012).
1070
1071
1072