Precocious neuronal differentiation and disrupted oxygen responses in Kabuki syndrome

Giovanni A. Carosso[1,2], Leandros Boukas[1,2,3], Jonathan J. Augustin[2,4,5], Ha Nam Nguyen[6], Briana L. Winer[2], Gabrielle H. Cannon[2], Johanna D. Robertson[2], Li Zhang[2], Kasper D. Hansen[2,3], Loyal A. Goff[2,5], 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, United States.
[2]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, United States.
[3]Department of Biostatistics, Johns Hopkins University School of Medicine, Baltimore, United States.
[4]Predoctoral Training Program in Biochemistry, Cellular, and Molecular Biology, Johns Hopkins University School of Medicine, Baltimore, United States.
[5]Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, United States.
[6]Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, United States.
[7]Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, United States.
[8]Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland.
[9]Landspitali University Hospital, Reykjavik, Iceland.
*Correspondence: hbjorns1@jhmi.edu

Abstract

Chromatin modifiers act to coordinate gene expression changes critical to neuronal differentiation from neural progenitor cells (NPCs). *KMT2D* encodes a histone methyltransferase that promotes transcriptional activation, and is frequently mutated in cancers and in the majority (>70%) of patients diagnosed with the congenital, multisystem intellectual disability (ID) disorder Kabuki syndrome 1 (KS1). While critical roles for KMT2D are established in various non-neuronal tissues, the effects of KMT2D loss in brain cell development have not been described. We conducted parallel studies of proliferation, differentiation, transcription, and chromatin profiling in KMT2D-deficient human and mouse models to define KMT2D-dependent functions in neurodevelopmental contexts, including adult-born hippocampal neural stem cells (NSCs) and NPCs *in vivo*. We report cell-autonomous defects in proliferation, cell cycle, and cell survival, accompanied by NSC depletion and precocious differentiation of NPCs *in vitro and in vivo*. Transcriptional suppression in KMT2D-deficient cells indicated perturbation of hypoxia-responsive cellular metabolism pathways, and functional experiments confirmed abnormalities of neuronal hypoxia response in cells with inactivated KMT2D. Together, our findings support a model in which loss of KMT2D function suppresses expression of oxygen-responsive gene programs in developing neural progenitors, resulting in precocious neuronal differentiation and exhaustion of the adult-born hippocampal progenitor pool.

Introduction

Trithorax group proteins promote chromatin accessibility by exerting antagonistic functions against Polycomb group transcriptional suppressors to activate gene expression (1). Fine-tuning of cell type transitions during neuronal development depends critically on this duality, as evidenced by severe neurodevelopmental defects caused by variants in numerous chromatin-modifying genes (2). Loss-of-function variants in genes encoding two such enzymes, lysine-specific methyltransferase 2D (KMT2D) and X-linked lysine-specific demethylase 6A (KDM6A/UTX) cause the intellectual disability (ID) disorder Kabuki syndrome (KS1 and KS2, respectively) (3, 4). Up to 74% (5) of KS cases result from mutations in *KMT2D* (KS1), encoding a major histone H3 lysine 4 (H3K4) methyltransferase which catalyzes the placement of chromatin-opening modifications promoting gene expression at context-specific targets. Developmental requirements of KMT2D in cardiac precursors (6), B cells (7, 8), muscle and adipose (9), and epithelial tissues (10) have been linked, respectively, to *KMT2D*-associated cardiac, immunologic, and oncogenic contexts (11), yet the effects of KMT2D deficiency in neurodevelopment are not yet understood.

We previously described a mouse model of Kabuki syndrome, *Kmt2d*$^{+/\beta geo}$, demonstrating characteristic features of KS1 patients including postnatal growth retardation, craniofacial abnormalities, and visuospatial memory impairments, associating with decreased adult hippocampal neurogenesis in the dentate gyrus (DG) (12). Decreased grey matter volume in the human DG was subsequently observed in KS1 patients, further implicating this region in KS1-associated ID phenotypes (13). The continual birth and integration of new neurons makes adult neurogenesis the most potent form of lifelong plasticity in the mammalian brain (14), and has now been established in humans (15). During late embryonic stages, a subset of multipotent progenitors persists in the DG (16), becoming subject to an array of intrinsic and extrinsic factors affecting NSC maintenance, i.e. self-renewal, proliferation, and differentiation into NPCs and neurons, throughout adult life. Mounting evidence tightly links metabolic rewiring (17) and hypoxic states in the DG (18) to regulation of NSC maintenance and NPC differentiation, and understanding the role of KMT2D in these processes will be critical to development of postnatal interventions in KS1-associated ID. Furthermore, robust cellular models of neurodevelopmental disruption in KS1 are needed to appreciate the target gene- and cell type-specific impact of KMT2D loss on brain function. Such studies have the potential to identify primary molecular consequences of KMT2D loss and facilitate targeted disease-modifying therapeutic approaches.

Here, we first report that KMT2D loss disrupts proliferation, cell cycle, and cell survival in a manner consistent across mouse and human KS1 neurodevelopment models. We then compare

72 transcriptional and KMT2D binding profiles to prioritize KS1-associated target genes in mouse
73 hippocampal cells specifically lacking KMT2D catalytic activity, identifying disruption of HIF1-
74 regulated, oxygen-responsive cellular metabolism pathways, prompting functional experiments which
75 demonstrated KMT2D-dependent neuronal responses to hypoxia. To assess translational potential of
76 our results in humans, we generated KS1 patient-derived iPSC and NPC models, confirming
77 proliferation, cell cycle, and cell survival defects, before using single-cell transcriptome analysis to
78 probe gene expression perturbances along the NPC development trajectory. Surprisingly, we
79 observed hallmarks of precocious neuronal maturation in differentiating KS1 NPCs. Finally, we explore
80 relevancy of these findings *in vivo*, identifying depletion of adult-born hippocampal NSCs,
81 accompanied by cell cycle defect, global transcriptional suppression, and precocious NPC
82 differentiation in *Kmt2d*-deficient adult mice.
83
84

Results

Genetic ablation of the *Kmt2d* SET methyltransferase domain disrupts cell proliferation and cell cycle in a cell-autonomous manner

We selected the HT22 mouse hippocampal neuronal cell line (19) for detailed analysis of KMT2D function in neuronal context. We generated distinct *Kmt2d*-inactivating alleles (biological replicates on an isogenic background) and analyzed each one in both heterozygosity and homozygosity, enabling gene dosage-dependent studies. Specifically, gDNA sequence encoding the SET methyltransferase domain was deleted by CRISPR-Cas9 with an upstream small guide RNA ($sgRNA^{up}$) in exon 52 and either $sgRNA^{1}$ (exon 54) or $sgRNA^{2}$ (intron 54), resulting in deletions of 565 ($Kmt2d^{\Delta 1}$) or 654 bp ($Kmt2d^{\Delta 2}$), respectively (**Supplementary Figure 1A**). Mutations were characterized to base-pair resolution by Sanger DNA sequencing and *in silico* translation to predict protein products (**Supplementary Figure 1A'**). Targeted cells were identified by PCR (**Supplementary Figure 1B**) and clonally expanded to establish heterozygous ($Kmt2d^{+/\Delta}$) and homozygous ($Kmt2d^{\Delta/\Delta}$) cell lines for comparison against the parental wild-type line ($Kmt2d^{+/+}$). *Kmt2d* mRNA was analyzed by RT-qPCR demonstrating ~50% decreased message of the targeted region in $Kmt2d^{+/\Delta}$ cells and no message in $Kmt2d^{\Delta/\Delta}$ cells (**Supplementary Figure 1C**), and confirmed specificity of the deletion, leaving intact mRNA message of exons upstream of the deletion site. Immunofluorescence using a KMT2D antibody, with antigenicity upstream of our mutations, confirmed presence of KMT2D protein in all genotypes, demonstrating distinctly nuclear distribution in $Kmt2d^{+/+}$ cells but more diffuse distribution in $Kmt2d^{+/\Delta}$ and $Kmt2d^{\Delta/\Delta}$ cells, as well as uniformly nuclear expression of neuronal nuclei marker RBFOX3/NeuN (**Figure 1A**).

Proliferation analysis, performed 72 hours after equal-density plating, revealed cell densities ~52% lower in $Kmt2d^{+/\Delta}$ cells and ~39% lower in $Kmt2d^{\Delta/\Delta}$ cells, compared to wild-type (**Figure 1B**). Proliferation defects were supported independently by CellTrace Violet generational tracking, detecting less dilution of a fluorescent tracer, i.e. fewer cell divisions, in $Kmt2d^{\Delta/\Delta}$ daughter cells compared to wild-type over 72 hours (**Figure 1C**, left, **Supplementary Figure 1D**). As control, genotype had no effect on parental cell uptake of the fluorescent tracer (**Figure 1C**, right). Flow cytometric analysis of cell cycle occupancy at steady-state, using marker of proliferation Ki-67 (KI67) and a DNA label, revealed that $Kmt2d^{+/\Delta}$ cells and $Kmt2d^{\Delta/\Delta}$ cells were enriched for S and G_2 phase, compared to wild-type (**Figure 1D, Supplementary Figure 1E-E'**). To characterize temporal dynamics of cell cycle progression, we synchronized cells in G_2/M phase using nocodazole, and analyzed DNA content at timepoints after release (**Figure 1E**). We observed wild-type cells exiting G_2/M phase at

117 significantly higher rates than *Kmt2d*$^{\Delta/\Delta}$ cells, beginning at 3 hours and up to 18 hours after release
118 (**Figure 1E'**). Finally, cell death was profiled by flow cytometric detection of caspase-3/7 substrate
119 cleavage and SYTOX uptake, distinguishing early apoptotic and necrotic cells. Apoptotic, but not
120 necrotic, cell proportions were greater in both *Kmt2d*$^{+/\Delta}$ cells (~287%) and *Kmt2d*$^{\Delta/\Delta}$ cells (~478%)
121 compared to wild-type (**Figure 1F**).
122 Findings of proliferative defects, G_2/M cell cycle delay, and increased apoptosis in *Kmt2d*-
123 inactivated hippocampal cells, upon selective targeting of the SET catalytic domain, support a cell-
124 autonomous role for KMT2D activity and make this an attractive model for detailed transcriptional and
125 chromatin profiling of KMT2D function in a neuronal context.
126

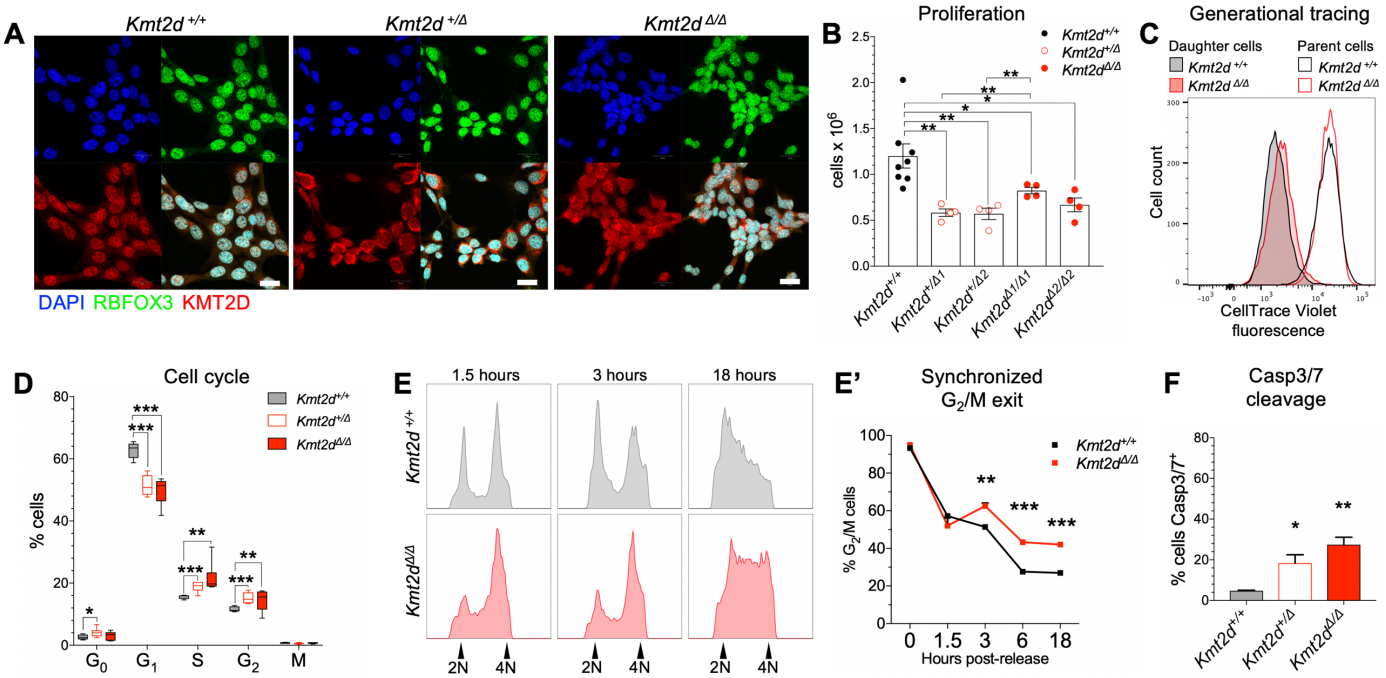


Figure 1. Genetic ablation of the *Kmt2d* SET methyltransferase domain disrupts cell proliferation and cell cycle in a cell-autonomous manner. (**A**) Representative immunostaining against KMT2D and neuronal nuclei marker RBFOX3 in HT22 mouse hippocampal cells, including wild-type (*Kmt2d*$^{+/+}$) and KMT2D-inactivated cells (*Kmt2d*$^{+/\Delta}$ and *Kmt2d*$^{\Delta/\Delta}$). (**B**) Decreased proliferation in *Kmt2d*-inactivated cells quantified 72 hours after equal density plating (7 wells per cell line). (**C**) Generational tracking by CellTrace Violet dye in *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells over 72 hours. Higher mean fluorescence intensity indicates less dye dilution, i.e. fewer cell divisions, in *Kmt2d*$^{\Delta/\Delta}$ daughter cells (left) at 72 hours. Negative control, parental cells confirm equal, genotype-independent uptake of CellTrace (right) at 0 hours. (**D**) Flow cytometric quantification of cell cycle at steady-state in HT22 cells (5 wells per cell line) by gating on Kl67 and DAPI fluorescence (DNA content). (**E**) *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells synchronized in G_2/M and released for time course analysis of G_2/M exit, by DNA content, up to 18 hours after release, and (**E'**) quantification of cells in G_2/M (technical triplicates per time point). (**F**) Flow cytometric quantification of early apoptotic cells using fluorescence coupled to caspase-3/7 peptide cleavage in *Kmt2d*$^{+/+}$, *Kmt2d*$^{+/\Delta}$, and *Kmt2d*$^{\Delta/\Delta}$ cells (3-6 wells per cell line). Bars indicate mean±SEM. Asterisks indicate significance from control mean, Student's t-test (*$p<0.05$, **$p<0.01$, ***$p<0.001$).

146 **Suppressed transcription of KMT2D- and HIF-regulated hypoxia response genes upon loss of**
147 **the KMT2D SET methyltransferase domain**

148 We performed high-coverage RNA-seq comparing three *Kmt2d*$^{\Delta/\Delta}$ clones against the parental
149 *Kmt2d*$^{+/+}$ line, each in technical triplicate, followed by differential expression analysis. Libraries robustly
150 clustered by genotype, with clear separation of *Kmt2d*$^{\Delta/\Delta}$ cells from *Kmt2d*$^{+/+}$ by Principal Component
151 Analysis (PCA) of gene expression, yielding 575 significant differentially-expressed genes (DEGs) at a
152 False Discovery Rate (FDR) of 0.05 in *Kmt2d*$^{\Delta/\Delta}$ cells compared to *Kmt2d*$^{+/+}$ (**Figure 2A-A',**
153 (**Supplementary Figure 2A-B**, **Supplementary Table 1**). ~76% of significant DEGs (436 genes) were
154 downregulated (down hereafter) in *Kmt2d*$^{\Delta/\Delta}$ cells, consistent with strong global suppression of gene
155 expression upon *Kmt2d* inactivation, including known KMT2D target genes such as Krueppel-like
156 factor 10 (*Klf10*) (12). Overrepresentation analysis (ORA) determined several significantly enriched
157 gene networks among *Kmt2d*$^{\Delta/\Delta}$ down DEGs, including glycolysis, hypoxia-inducible factor 1 (HIF1)
158 signaling, RAS signaling, autophagy, and others, while *Kmt2d*$^{\Delta/\Delta}$ upregulated DEGs (up hereafter) were
159 enriched in fewer networks (**Figure 2B**).

160 KMT2D-mediated transcriptional activation occurs in a target gene-specific manner, so we
161 reasoned that among the 575 observed *Kmt2d*$^{\Delta/\Delta}$ DEGs, a subset of genes found to also bind KMT2D
162 itself, in wild-type cells, would more likely represent direct transcriptional consequences of *Kmt2d*
163 inactivation, whereas non-bound DEGs could reflect indirect effects. We performed chromatin
164 immunoprecipitation followed by high-throughput sequencing (ChIP-seq) with a previously validated
165 ChIP-grade KMT2D antibody (9) in *Kmt2d*$^{+/+}$ HT22 cells. This identified 3,756 KMT2D binding peaks
166 significantly enriched over input (**Supplementary Table 2**), of which ~10% occur inside promoters and
167 ~33% (1,235 peaks) occur within 5 kb of a gene transcription start site (TSS±5kb, **Supplementary**
168 **Figure 2C-E**). To account for promoter and enhancer interactions (9, 10, 20), we reasoned that
169 TSS±5kb peaks more likely constitute a cis-regulatory function of KMT2D binding on proximal genes,
170 and we will refer to these as KMT2D-bound genes. The 1,463 KMT2D-bound genes in neuronal cells
171 (**Supplementary Table 3**) were significantly enriched in mRNA 3'UTR binding, rho GTPase signaling,
172 circadian clock, translation elongation, oxidative stress-induced senescence, HIF1 signaling, and other
173 pathways (**Supplementary Figure 2F**).

174 We then intersected KMT2D-bound genes with *Kmt2d*$^{\Delta/\Delta}$ DEGs to reveal 74 putative direct
175 target genes (**Table 1**), of which ~85% (63 genes) were as expected downregulated (**Figure 2C**),
176 including insulin-like growth factor 1 (*Igf1*), fos-like antigen 2 (*Fosl2*). At least 20 observed KMT2D-
177 bound, *Kmt2d*$^{\Delta/\Delta}$ DEGs were previously described as KMT2D targets in other tissues (7, 21),

178 suggesting some KMT2D target conservation across cell types. KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ down-DEGs
179 were most significantly enriched for face morphogenesis, glycolytic process, protein kinase B
180 signaling, hypoxia response, and cell proliferation pathways (**Figure 2C'**), and surprisingly, 29 of these
181 63 genes are known to be HIF1-regulated (22). KMT2D-ChIP peaks on HIF1-regulated genes clustered
182 at promoters and enhancers, often overlapping CpG islands in genes such as *Fosl2*, with others
183 clustering at alternative TSSs, as in retinoic acid receptor alpha (*Rara*), or in enhancer-like peaks as in
184 DNA-damage-inducible transcript 4 (*Ddit4*) (**Figure 2D**, **Supplementary Figure 2G-G'**).
185 A significant fraction of KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ DEGs control oxygen-responsive metabolism,
186 warranting interrogation of KMT2D and HIF1A binding site overlaps which could implicate a cis-
187 regulatory role for KMT2D in cellular oxygen responses. We first intersected our significant KMT2D-
188 ChIP peak sequences with regions previously determined HIF1A-bound in embryonic heart (23),
189 yielding 423 regions overlapping with base pair resolution (**Figure 2E**). Like KMT2D, HIF1A showed
190 ~10% of peaks located inside promoters, but among overlapping KMT2D/HIF1A-bound peaks this
191 fraction approached ~40%, supporting cooperative cis-regulatory activity by these factors (**Figure**
192 **2E', Supplementary Figure 2E**). Remarkably, we observed 289 TSS±5kb genes, as defined above, for
193 these overlapped KMT2D/HIF1A-bound peaks, 8 of which are *Kmt2d*$^{\Delta/\Delta}$ DEGs: *Ddit4*, heat shock
194 protein family D member 1 (*Hspd1*), Cbp/p300-interacting transactivator with Glu/Asp-rich C-terminal
195 domain 2 (*Cited2*), DAZ associated protein 1 (*Dazap1*), WASH complex subunit 4 (*Washc4*), transducer
196 of (*Errb2*), 1 (*Tob1*), carboxypeptidase Q (*Cpq*), leucine rich pentatricopeptide repeat containing 4
197 (*Lrpprc*).
198 To determine if KMT2D/HIF1A-regulated genes generalize to additional tissues, and restricting
199 analysis to promoters, we next interrogated independent sets of genes with experimentally validated,
200 hypoxia-induced HIF1A binding at the promoter (24). Of 86 validated genes, 5 were KMT2D-bound,
201 *Kmt2d*$^{\Delta/\Delta}$ down-DEGs in neuronal cells, a 23.3-fold greater frequency than that expected by chance
202 (odds ratio 23.3, p=4.74e-6): *Ddit4*, *Cited2*, aldolase A (*Aldoa*), Bcl2/adenovirus E1B 19-KD protein-
203 interacting protein 3-like (*Bnip3l*), and ceruloplasmin (*Cp*). Of 81 genes validated in three or more
204 distinct tissues, 3 were KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ down-DEGs in neuronal cells (odds ratio=14.03,
205 p=0.002): *Klf10*, *Rara*, and *Ddit4*. Of note, these genes are known to have important roles regulating
206 cellular differentiation in addition to mediating oxygen responses.
207 Given the significant occurrence of *Kmt2d*$^{\Delta/\Delta}$ down-DEGs for oxygen response in neuronal cells
208 at normoxic culture conditions, including shared and direct KMT2D/HIF1A targets, we hypothesized
209 that KMT2D may be required for transcriptional response to changes in cellular oxygen pressure. To
210 test this in a *Kmt2d* dosage-dependent manner, we subjected *Kmt2d*$^{+/+}$ cells, *Kmt2d*$^{+/\Delta}$ cells, and

211 *Kmt2d*$^{\Delta/\Delta}$ cells to normoxic (20% O_2) or hypoxic (1% O_2) conditions, then analyzed hypoxia-induced
212 gene expression. RT-qPCR analysis of the canonical HIF1A targets vascular endothelial growth factor
213 A (*Vegfa*), Bcl2/adenovirus E1B 19-KD protein-interacting protein 3 (*Bnip3*), DNA-damage-inducible
214 transcript 3 (*Ddit3*), and cyclin-dependent kinase inhibitor 1A (*Cdkn1A*) revealed the expected gene
215 induction, up to several-fold, upon hypoxic exposure over baseline in *Kmt2d*$^{+/+}$ cells; in contrast,
216 *Kmt2d*$^{+/\Delta}$ and *Kmt2d*$^{\Delta/\Delta}$ cell lines failed to induce expression of hypoxia-inducible genes to comparable
217 levels (**Figure 2F**).
218 Upon cellular sensing of low oxygen, activated HIF1A becomes stabilized against cytoplasmic
219 degradation and undergoes nuclear translocation, so we next quantified the proportion of cellular
220 HIF1A fluorescence colocalizing in the nucleus, under normoxic and hypoxic conditions
221 (**Supplementary Figure 2H**). Unexpectedly, under normoxic conditions *Kmt2d*$^{\Delta/\Delta}$ cells exhibited ~2-
222 fold greater proportion of nuclear HIF1A (~17.5%) than *Kmt2d*$^{+/+}$ cells (~8.8%), and this difference
223 increased further upon hypoxic exposure (**Supplementary Figure 2H'**). In *Kmt2d*$^{+/\Delta}$ cells, nuclear
224 HIF1A fraction (~7.3%) was comparable to wild-types in normoxia, but in hypoxia sharply increased to
225 ~24.8%. *Kmt2d*$^{+/+}$ cells in the same hypoxia paradigm showed modestly increased nuclear HIF1A,
226 suggesting a lower threshold for hypoxic HIF1A translocation in *Kmt2d*$^{+/\Delta}$ cells and constitutive HIF1A
227 translocation in *Kmt2d*$^{\Delta/\Delta}$ cells. We again observed over-representation in G_2/M phase occupancy in
228 normoxia in *Kmt2d*$^{\Delta/\Delta}$ cells compared to *Kmt2d*$^{+/+}$ cells, and observed a mild increase of G_2/M phase
229 occupancy in hypoxia for all genotypes (**Supplementary Figure 2H''**). Flow cytometric quantification
230 of caspase-3/7 activity demonstrated sharp decreases in apoptotic cell fractions in all genotypes
231 under hypoxia (**Supplementary Figure 2I-I'**). In contrast, necrotic cell fractions detected by SYTOX
232 fluorescence were similar across genotypes in normoxia, but under hypoxia decreased only in
233 *Kmt2d*$^{+/+}$ cells and *Kmt2d*$^{+/\Delta}$ cells, while marginally increasing in *Kmt2d*$^{\Delta/\Delta}$ cells (**Supplementary Figure**
234 **2I''**).
235 Taken together, data from *Kmt2d*-inactivated mouse neuronal cells suggest that KMT2D plays
236 an important role in positive regulation of cellular proliferation and oxygen-responsive, HIF1-inducible
237 gene expression, including direct targets important to cellular differentiation. The cell-autonomous
238 nature of these *Kmt2d*-dependent phenotypes raises the question of whether such effects apply to
239 human cells, and furthermore what, if any, consequences occur in the context of neuronal
240 differentiation.
241
242

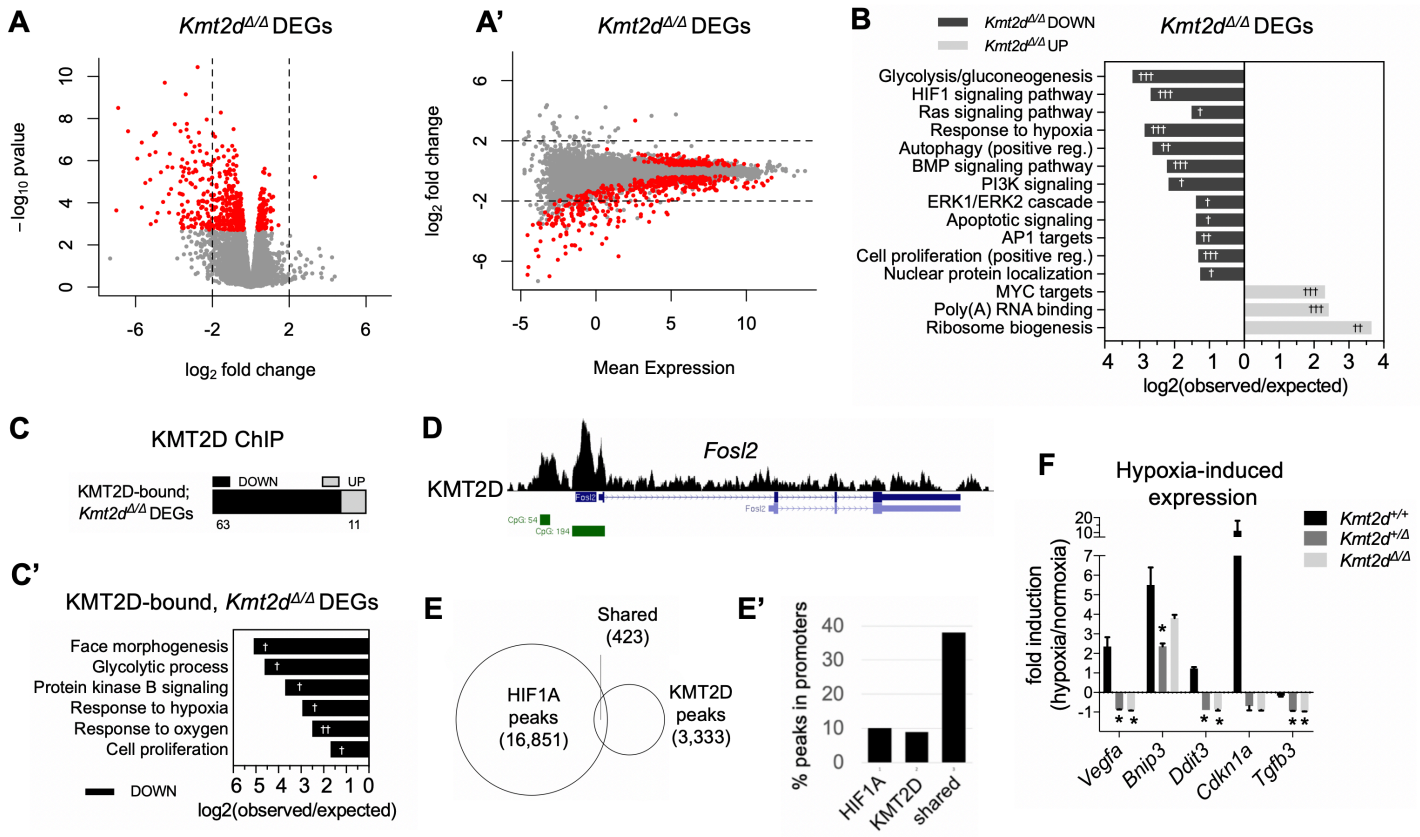


Figure 2. Suppressed transcription of KMT2D- and HIF-regulated hypoxia response genes upon loss of the *Kmt2d* SET methyltransferase domain in neuronal cells. (**A-A'**) Expression analysis by RNA-seq in HT22 mouse neuronal cells reveals 575 significant differentially expressed genes (DEGs) in *Kmt2d*$^{\Delta/\Delta}$ clones (3 biological replicates) relative to *Kmt2d*$^{+/+}$ cells, each in technical triplicate. Fold changes in expression indicate most significant *Kmt2d*$^{\Delta/\Delta}$ DEGs (~76%, red dots) are downregulated in *Kmt2d*$^{\Delta/\Delta}$ cells, plotted against (**A**) p-value and (**A'**) mean expression. (**B**) Gene networks significantly enriched among down- or up-regulated *Kmt2d*$^{\Delta/\Delta}$ DEGs. (**C**) *Kmt2d*$^{\Delta/\Delta}$ DEGs which are also KMT2D-bound, as determined by ChIP-seq chromatin profiling in *Kmt2d*$^{+/+}$ HT22 cells, and (**C'**) gene networks significantly enriched among KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ DEGs. (**D**) Representative ChIP-seq track of a KMT2D-bound, *Kmt2d*$^{\Delta/\Delta}$ DEG depicting KMT2D binding peaks (black), RefSeq gene annotations (blue), and CpG islands (green). (**E**) Overlapping loci of observed KMT2D-ChIP peaks in HT22 cells and HIF1A-ChIP peaks in embryonic heart (23). (**E'**) Overlapping KMT2D/HIF1A peak regions, compared to individually bound regions, are enriched at gene promoters. (**F**) RT-qPCR analysis of hypoxia-induced gene expression in *Kmt2d*$^{+/+}$, *Kmt2d*$^{+/\Delta}$, and *Kmt2d*$^{\Delta/\Delta}$ cells, following 72 hours in normoxia (20% O_2) or hypoxia (1% O_2), with fold induction of HIF1A target gene mRNA (2 biological replicates per genotype, each in technical triplicate; negative fold changes plotted as negative reciprocal). Fisher's Exact Test (†FDR<0.05, ††FDR<0.01, †††FDR<0.001); *significance from wild-type induction, Student's t-test (*p<0.05, **p<0.01, ***p<0.001).

264 **KS1 patient-derived cells recapitulate KMT2D-associated defects in proliferation and cell cycle**

265 We reprogrammed skin biopsy fibroblasts to generate induced pluripotent stem cells (iPSCs)
266 from a female with molecularly confirmed KS1 bearing a heterozygous nonsense *KMT2D* mutation
267 (c.7903C>T:p.R2635*), and characteristic KS1 features including growth retardation, classical facial
268 features, and congenital heart disease. Neuropsychological testing in a previous longitudinal study (25)
269 of this patient had revealed a visuospatial memory impairment similar to that described in
270 retrospective and prospective studies of KS1 patients (12, 26). KS1 iPSCs (KS1-1) bearing normal 46,
271 XX karyotype (**Supplementary Figure 3A**) and characteristic iPSC morphology (**Figure 3A**) were
272 selected for comparison against two independently generated, previously described iPSC lines (C1-2,
273 C3-1) from unrelated healthy controls (27). KS1 iPSCs were predicted to be *KMT2D* haploinsufficient
274 due to nonsense-mediated decay (NMD), and RT-qPCR confirmed decreased *KMT2D* mRNA
275 compared to controls, with equivalently decreased expression of exons upstream and downstream of
276 the mutation site as expected for a nonsense mutation (**Supplementary Figure 3B-C**). Flow
277 cytometric analyses following pulse of cell division marker 5-ethynyl-2′-deoxyuridine (EdU)
278 demonstrated lower proliferation rates (~25%) in KS1 iPSCs compared to controls (**Figure 3B**),
279 accompanied by a shift in cell cycle occupancy (**Figure 3C, Supplementary Figure 3D**) toward
280 overrepresentation of G_2/M phase cells (24% higher in KS1) and underrepresentation of G_0/G_1 phase
281 cells (23% lower in KS1).

282 We next generated nestin (NES)-expressing NPCs through parallel differentiation of KS1 and
283 control iPSCs using an established small molecule inhibition protocol (28). RT-qPCR confirmed
284 decreased *KMT2D* also in KS1 NPCs (**Supplementary Figure 3E**) and both KS1 and control NPCs
285 displayed normal NPC morphology (**Figure 3D, Supplementary Figure 3F**). Immunostaining against
286 histone marks in KS1 NPCs detected significantly reduced fluorescence of open chromatin mark
287 H3K4me2 (~65%) and increased fluorescence of closed chromatin marks H3K27me3 (~99%) and
288 H3K9me3 (~180%) compared to controls (**Supplementary Figure 3G-I**). Compared to iPSC data, KS1
289 NPCs relative to control showed a markedly greater reduction in dividing cells as detected by EdU
290 incorporation (~47%, **Figure 3E**) as well as fewer mitotic divisions over a 72-hour period as detected
291 by CellTrace Violet (**Supplementary Figure 3J**). Unlike in iPSCs, cell cycle abnormalities were not
292 observed in KS1 NPCs (**Figure 3F, Supplementary Figure 3K**) suggesting this phenotype may be
293 cell-type dependent or lost during induced differentiation. However, flow cytometric analysis revealed
294 higher proportions of cell death in KS1 cells compared to controls among both iPSCs (~130%) and
295 NPCs (~115%) (**Figure 3G, Supplementary Figure 3L**).

296 Finally, to determine whether the cell cycle G_2/M bias in KS1 iPSCs occurs in additional KS1
297 patients and unmanipulated primary cells, we performed cell cycle analysis in primary fibroblasts from
298 two additional molecularly confirmed KS1 patients (KS1-2, KS1-3) in addition to KS1-1 and two
299 additional healthy controls. KS1 and control fibroblasts were treated with nocodazole to synchronize
300 in G_2/M phase, and flow cytometric DNA content analysis at 0- and 3-hours post-release
301 demonstrated exit from G_2/M phase among control fibroblasts, but this was markedly decreased in
302 fibroblasts of KS1 patients (**Figure 3H-H'**). Thus, delayed exit from G_2/M was consistent in primary
303 cells from three independent KS1 patients.
304
305

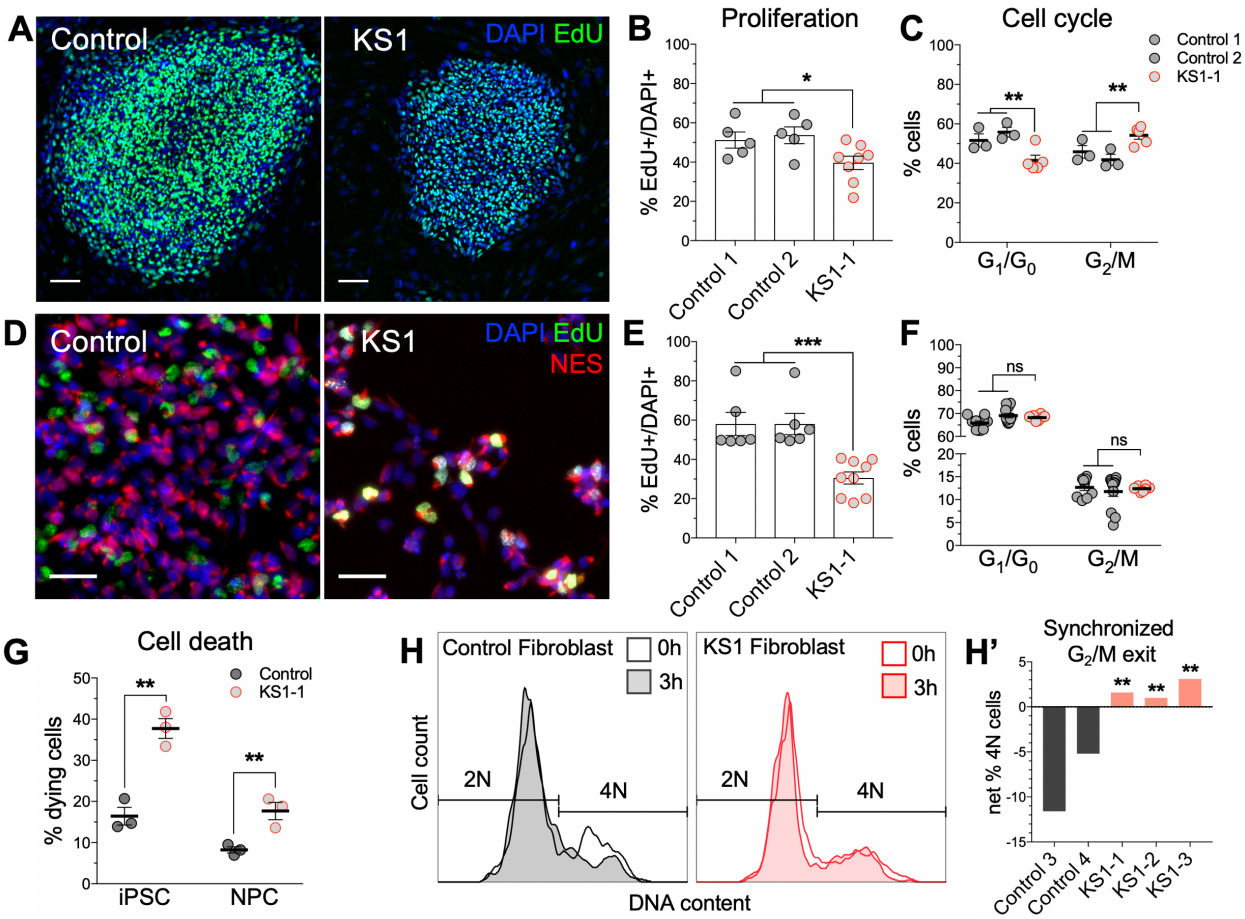


Figure 3. KS1 patient-derived cells recapitulate KMT2D-associated defects in proliferation and cell cycle. (**A**) Representative immunostaining of iPSCs from a confirmed *KMT2D*$^{+/-}$ KS1 patient (c.7903C>T:p.R2635*) and healthy controls, showing normal colony morphology. (**B**) Proliferative cells (5-8 wells per cell line) were pulsed with EdU for 30 minutes and quantified by flow cytometry. (**C**) Cell cycle analysis in iPSCs (3-5 wells per cell line), 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 NPC morphology, nestin (NES) staining, and proliferative cells (EdU$^+$). (**E**) Proliferative cells (6-9 wells per cell line) were pulsed with EdU for 30 minutes and quantified by flow cytometry. (**F**) Cell cycle analysis in NPCs (8-12 wells per cell line), 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 (3 wells per cell line) from KS1 patient and controls. (**H-H'**) Synchronized G_2/M exit analysis by flow cytometry in fibroblasts from additional KS1 patients (KS1-1, KS1-2, KS1-3) and healthy controls (Controls 3 and 4), in triplicate per cell line. Cells were enriched for G_2/M phase using nocodazole and analyzed by DAPI fluorescence to and quantify (**H'**) G_2/M phase cell fractions at 0 and 3 hours after release. Bars indicate mean ± SEM. Asterisks indicate significance from control mean. Student's t-test (*p<0.05, **p<0.01, ***p<0.001).

325 **Transcriptional suppression of metabolic genes in cycling cells, and hallmarks of precocious**
326 **neuronal differentiation in KS1 patient-derived NPCs**

327 To interrogate transcriptional consequences of *KMT2D* loss in the context of neuronal
328 differentiation, we performed single-cell RNA sequencing (scRNA-seq) in iPSCs and NPCs from the
329 KS1 patient and controls (**Supplementary Figure 4A**). We first inspected cell clusters for expression
330 of known iPSC or NPC markers to further verify cell identities, confirming that iPSCs displayed
331 characteristic expression of pluripotency markers including POU domain, class 5, transcription factor
332 1 (*POU5F1*) (**Supplementary Figure 4B**), while NPCs expressed known markers of early neural lineage
333 including paired box protein Pax-6 (*PAX6*) (**Supplementary Figure 4C**). A subset of NPCs further
334 expressed neuronal maturation markers including microtubule-associated protein 2 (*MAP2*)
335 (**Supplementary Figure 4D**).

336 First, differential expression analysis in iPSCs and NPCs identified genes downregulated or
337 upregulated in KS1 patient relative to healthy controls (**Supplementary Figure 4E**). KS1 iPSCs
338 displayed strong transcriptional suppression among 421 significantly differentially expressed genes
339 (DEGs), with 372 genes (~88%) down and only 49 genes (~12%) up (**Supplementary Table 4**). In
340 contrast, NPCs, following small molecule induction, showed less directional bias with 346 significant
341 DEGs identified, among which 147 genes (~42%) were down and 199 genes (~58%) were up
342 (**Supplementary Table 5**). We reasoned that despite differences in global expression trends between
343 iPSCs and NPCs, genes shared down in both KS1 iPSCs and NPCs may be more likely to represent
344 robust targets of KMT2D. Intersection of DEG lists showed that 40 genes were shared down in KS1
345 iPSCs and NPCs and 10 genes were shared up (**Supplementary Figure 4F-G, Supplementary Table**
346 **6**). Shared down genes included the glycolysis genes aldehyde dehydrogenase 7 family member A1
347 (*ALDH7A1*), enolase 1 (*ENO1*), and triosephosphate isomerase 1 (*TPI1*), as well as factors important to
348 stem cell maintenance including proliferation-associated protein 2G4 (*PA2G4*) and protein lin-28
349 homolog A (*LIN28A*). Interestingly, some genes such as *TPI1* and *PA2G4* had also been observed
350 down in *Kmt2d*-inactivated mouse neuronal cells (**Supplementary Table 1**). Similar to *Kmt2d*-deficient
351 neuronal cells, downregulated genes in both patient-derived models were significantly enriched for
352 HIF1A direct targets, genes containing the hypoxia-responsive element (HRE) 5'-RCGTG-3' motif, and
353 known hypoxia response genes (**Supplementary Figure 4H**).

354 We next focused specifically on NPCs from KS1 patient and controls to better understand
355 transcriptional abnormalities during neuronal differentiation. We used Uniform Manifold Approximation
356 and Projection (UMAP) to visualize single cells in a manner that displays high-dimensionality data

357 while preserving both local and global relationships (29). Cells of both control NPC lines were tightly
358 clustered, indicating similar expression profiles, in contrast to a distinct separation of KS1 patient cells
359 which appears to gradually lessen in a subset (top) of cells that more closely resemble controls
360 (**Figure 4A**). We then used marker expression to partition single-cell libraries into developmentally
361 informative subsets as follows. First, as control, we verified that differences in cell cycle phase
362 composition do not account for KS1-associated differential gene expression in NPCs (**Supplemental**
363 **Figure 5A-A'', Supplementary Table 7**). Next, we partitioned cells by expression of maturation stage-
364 specific genes to defined an NPC differentiation trajectory consisting of early or "cycling" NPCs,
365 "transitioning" NPCs, and "differentiating" NPCs (**Figure 4B**). We observed that cycling cells comprise
366 the vast majority of NPCs analyzed and exhibit the greatest KS1-associated expression differences
367 (**Figure 4B**, purple cells), while expression profiles of transitioning and differentiating NPCs show
368 gradual convergence of gene expression (**Figure 4B**, grey and yellow cells, respectively). We analyzed
369 DEGs exclusively within cycling, transitioning, and differentiating NPC subsets to determine if
370 particular gene networks drive transcriptional differences in a stage-specific manner (**Figure 4C,**
371 **Supplementary Table 7**). The most prevalent networks overall, NMD, translation initiation, and protein
372 targeting, were strongly enriched regardless of maturation stage. In contrast, DEGs in transitioning
373 NPCs, and to a lesser extent cycling NPCs, show enrichment of genes comprising the Notch signaling
374 pathway including delta-like protein 3 (*DLL3*), protein jagged-1 (*JAG1*), transcription factor HES-5
375 (*HES5*), and cyclin D1 (*CCND1*). Cycling NPCs had DEGs enriched in glycolysis pathways including
376 *ENO1*, *TPI1*, polyadenylate-binding protein 1 (*PABPC1*), L-lactate dehydrogenase L-chain (*LDHA*),
377 *ALDOA,* and pyruvate kinase PKM (*PKM*), the latter two genes having also been observed as KMT2D-
378 bound, *Kmt2d*$^{\Delta/\Delta}$ down genes in mouse neuronal cells.

379 Apart from increased rates of KS1 cell death (**Figure 3G**), another factor potentially contributing
380 to the observed decrease of proliferative KS1 NPCs (**Figure 3E**) would be a change in cellular
381 differentiation, such as precocious cell maturation, resulting in depletion of cycling precursors. To
382 explore this in detail with scRNA-seq data, we examined expression of markers ranging from
383 immature cells (*PAX6*$^{+}$) to the most differentiated cells (*MAP2*$^{+}$) (**Figure 4D-D'**). We further restricted
384 analysis exclusively to the transitioning and differentiating, i.e. "maturing" NPC subset
385 (**Supplementary Figure 5A'**), defining a trajectory that enabled parsing of cells into binned deciles of
386 increasing maturation (**Supplementary Figure 5B-E''**). Quantification of cell densities revealed strong
387 bias of KS1 NPCs in the most matured bins relative to controls (**Figure 4E**), i.e. greater representation
388 of mature NPCs from KS1 than controls. These transcriptional signatures of precocious maturation

389 were corroborated at protein level by flow cytometric analysis, finding KS1 NPCs had increased MAP2
390 fluorescence and reciprocally decreased PAX6 fluorescence relative to control (**Figure 4F-F'**).
391 Together, these results link transcriptional suppression of cellular metabolic pathways to cell-
392 autonomous proliferation defects in *KMT2D*-deficient KS1 patient-derived stem cell models, revealing
393 striking similarities to phenotypes and gene expression profiles in *Kmt2d*-deficient mouse neuronal
394 cells. Furthermore, scRNA-seq data suggest that precocious differentiation may contribute to KS1-
395 associated neurodevelopmental defects, however a stronger demonstration would be to observe
396 these phenotypes *in vivo*, in an animal model established to phenocopy ID-related features of KS1
397 patients such as visuospatial memory impairments.
398
399

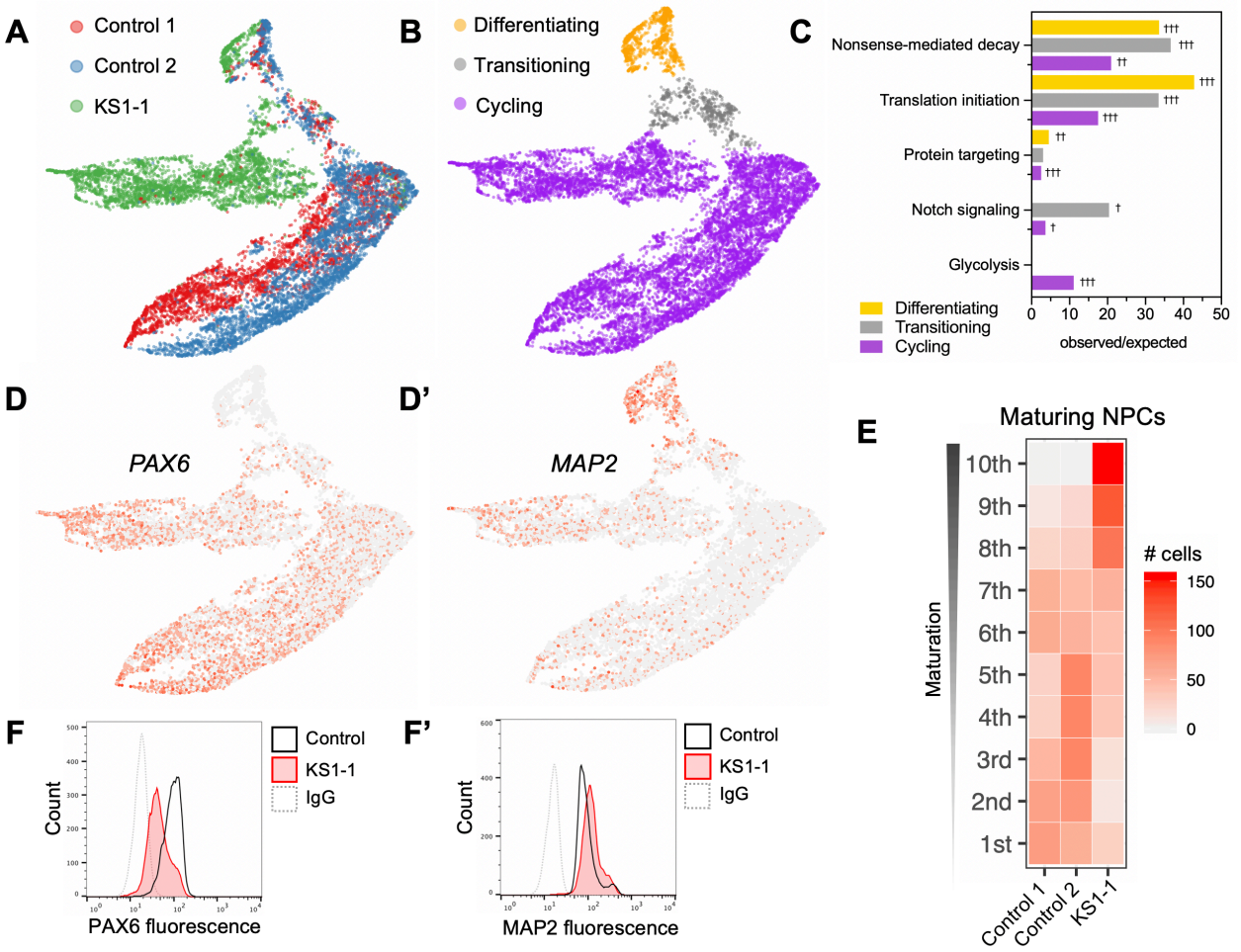


Figure 4. Transcriptional suppression of metabolic genes in cycling cells, and hallmarks of precocious neuronal differentiation in KS1 patient-derived NPCs. (**A**) Single-cell RNA-seq profiling in patient and healthy control iPSC-derived NPCs (~5,000 cells per patient), with Uniform Manifold Approximation Projection (UMAP) to visualize gene expression differences between cells. (**B**) NPCs partitioned by NPC maturation stage as defined by stage-specific marker expression, and (**C**) enriched gene networks, analyzed exclusively among DEGs for each NPC subset (cycling, transitioning, and differentiating). (**D-D'**) UMAPs annotated by relative expression intensities of NPC markers revealing the maturation trajectory from (**D**) early NPC (*PAX6*$^{+}$) to (**D'**) differentiating NPC (*MAP2*$^{+}$). (**E**) Heatmap comparing density of maturing NPCs along the maturation trajectory, defined by binned marker expression from earliest (1st) to most differentiated (10th) deciles, with KS1 cells disproportionately occupying the most mature bins. (**F-F'**) Flow cytometric validation of protein marker expression in NPCs from KS1 patient and controls, plotting fluorescence intensities of (**F**) early NPC marker PAX6 and (**F'**) mature NPC marker MAP2. (Fisher's Exact Test, †FDR<0.05, ††FDR<0.01, †††FDR<0.001).

416 ***In vivo* NSC depletion and precocious NPC differentiation in a *Kmt2d*$^{+/\beta geo}$ mouse model of**
417 **Kabuki syndrome**

418 Finally, we asked whether proliferative defects, transcriptional suppression, and precocious
419 differentiation phenotypes observed *in vitro* using mouse neuronal cells and KS1 patient-derived stem
420 cell models generalize to *in vivo* studies in an established KS1 mouse model. *Kmt2d*$^{+/\beta geo}$ mice, bearing
421 a heterozygous *Kmt2d* SET domain mutation, were previously found to exhibit visuospatial memory
422 impairments and reductions of doublecortin (DCX^{+}) and pulsed EdU^{+} cells in the DG (12, 30), but
423 detailed analysis of neurogenic lineage progression and cycling NPC phenotypes in *Kmt2d*$^{+/\beta geo}$ mice
424 has not been performed.

425 For cell cycle and RNA-seq analysis in *Kmt2d*$^{+/\beta geo}$ mice, we employed an EdU pulse paradigm
426 to label dividing cells in 8-week old mice, and sampled micro-dissected DG within 1 cell cycle (16
427 hours) in order to capture the full complement of adult-born, dividing NPCs (**Figure 5A-A'**), then
428 purifying EdU^{+} nuclei by fluorescence-activated cell sorting (FACS) (**Figure 5B**, **Supplementary**
429 **Figure 6A**). Analysis of DNA content in the EdU^{+} population revealed enrichment in G_2/M phase and
430 paucity of G_0/G_1 phase (**Figure 5C**), as observed in KMT2D-deficient *in vitro* models. We next
431 performed transcriptional profiling by RNA-seq in purified EdU^{+} DG nuclei, yielding 827 DEGs
432 (**Supplementary Figure 6B**, **Supplementary Table 8**). Among 416 down genes in *Kmt2d*$^{+/\beta geo}$ nuclei,
433 the most significant enrichments were for misfolded protein binding, TCA cycle, proteasome complex,
434 oxygen response, and poly(A) RNA-binding genes (**Supplementary Figure 6C**). Among genes
435 upregulated we observed the upstream pro-apoptosis gene caspase-8 (*Casp8*) in these *Kmt2d*$^{+/\beta geo}$
436 proliferative nuclei. Given the observed downregulation of poly(A) RNA-binding genes, we considered
437 the possibility that improper 3'UTR-mediated mRNA metabolism could lead to accumulation of
438 transcripts influencing NPC maturation. Indeed, despite little overall bias in up- or downregulation
439 among *Kmt2d*$^{+/\beta geo}$ proliferative nuclei, interrogation of canonical positive regulators of neuronal
440 differentiation revealed a marked predominance of pro-neural transcripts upregulated, having only 3
441 genes down but 14 genes up, including copine-1 (*Cpne1*), focal adhesion kinase 1 (*Ptk2*), ras-related
442 protein RAB11A (*Rab11A*), and retinoblastoma-associated protein 1 (*Rb1*). Interestingly, KS1 patient
443 NPCs had also shown upregulated pro-neural genes such as nuclear receptor subfamily 2, group F,
444 member 1 (*NR2F1*), pro-neural transcription factor HES-1 (*HES1*), and the GABAergic interneuron-
445 promoting ladybird homeobox 1 (*LBX1*), and *Kmt2d*$^{\Delta/\Delta}$ mouse neuronal cells *in vitro* compared to wild-
446 type had significantly higher mRNA levels of brain-derived neurotrophic factor (*Bdnf*) and neuron-
447 specific microtubule element (*Tubb3*/*Tuj1*). Such pro-neural gene expression observed across KS1
448 models raises the possibility that NPC differentiation states could be altered in *Kmt2d*$^{+/\beta geo}$ mice.

449 To examine NPC lineage progression *in vivo*, we compared stage-specific cell abundances
450 both at steady-state and after pulse-label birth dating of adult-born NPCs. We used immunostaining
451 markers (**Supplementary Figure 6D**) to compare abundance of individual DG cell precursor stages of
452 adult mice 8 weeks old, comparing *Kmt2d*$^{+/\beta geo}$ mice to sex- and age-matched *Kmt2d*$^{+/+}$ littermates
453 (**Figure 5D-D'**).
454 At steady-state, overall we observed significantly fewer neural progenitors in *Kmt2d*$^{+/\beta geo}$ mice
455 compared to *Kmt2d*$^{+/+}$ mice at all stages analyzed, indicating that heterozygous *Kmt2d* loss impacts
456 multiple NPC stages in the adult DG (**Figure 5D''**). Importantly, quiescent radial glia-like (qRGL,
457 NES^{+}/MCM2^{-}) NSCs were ~39% less numerous in *Kmt2d*$^{+/\beta geo}$ mice compared to *Kmt2d*$^{+/+}$ mice,
458 indicating a baseline depletion of the adult hippocampal NSC pool. Activated RGL (aRGL,
459 NES^{+}/MCM2^{+}) NSCs were ~43% less numerous compared to *Kmt2d*$^{+/+}$ littermates, and intermediate
460 progenitor cells (IPCs) (NES^{-}/MCM2^{+}), were ~26% fewer. We confirmed our prior observations (12, 30)
461 of reduced representation at the neuroblast, or immature neuron, stage (DCX^{+}) of NPCs, finding a 28%
462 decrease in *Kmt2d*$^{+/\beta geo}$ mice compared to *Kmt2d*$^{+/+}$ littermates. By stratifying our analysis anatomically
463 along the septotemporal axis of the DG, we observe that aRGL NSC reductions in *Kmt2d*$^{+/\beta geo}$ mice
464 were more pronounced in the septal DG than the temporal region (**Supplementary Figure 6E-E'**),
465 congruous with *in vivo* spatial memory defects previously observed (12) which have been localized to
466 this specific DG region (31). Because DCX^{+} NPCs characteristically migrate radially during
467 differentiation, we compared radial distances of DCX^{+} cell bodies from the SGZ plane and observed
468 increased distances in *Kmt2d*$^{+/\beta geo}$ mice compared to *Kmt2d*$^{+/+}$ littermates (**Supplementary Figure 6F**).
469 Finally, despite reductions of NPC populations in *Kmt2d*$^{+/\beta geo}$ mice, we observed no numeric
470 differences among mature neurons in the DG (RBFOX3/NeuN^{+}, **Supplementary Figure 6G**), nor were
471 any gross abnormalities uncovered with post-mortem MRI volumetric analysis (**Supplementary Figure
472 6H, Supplementary Table 9**).
473 From these data, we then calculated a lineage progression index to approximate the expansion
474 potential of each successive neurogenic cell type. Although *Kmt2d*$^{+/\beta geo}$ mice showed fewer total cells
475 of each type at steady-state, the lineage progression index at each cell-type transition did not differ
476 significantly in *Kmt2d*$^{+/\beta geo}$ mice, in fact appearing modestly higher in early-stage transitions in
477 *Kmt2d*$^{+/\beta geo}$ mice, suggesting that particular cell-type transition impairments are not responsible for the
478 adult neurogenesis defect in *Kmt2d*$^{+/\beta geo}$ mice (**Supplementary Figure 6I**). We did note a substantially
479 higher coefficient of variation in *Kmt2d*$^{+/\beta geo}$ mice at the qRGL-to-aRGL progression stage, suggesting

480 impaired coordination of NSC activation or increased stochasticity in mitotic entry of *Kmt2d*$^{+/\beta geo}$ NSCs
481 (**Supplemental Figure I'**).
482 Pulse-labeling combined with marker-based imaging enables precise measurement of birth
483 dates, i.e. mitotic division, of specific cell types. To resolve temporal dynamics of NPC differentiation
484 in *Kmt2d*$^{+/\beta geo}$ and wild-type mice, we pulsed adult mice with EdU for a period of 2 weeks, during
485 which time a subset of labeled DG cells is expected to reach the late NPC or immature neuron stage,
486 characterized by the radial extension of a DCX^{+} neuronal process from the cell body toward the
487 hippocampal molecular layer. In contrast, a different subset of labeled cells bearing a NES^{+} qRGL
488 NSC-like process represents the population of cells that have not yet progressed beyond a stem-like
489 state. Thus, by quantifying EdU-labeled cells exhibiting either a DCX^{+} neuronal process (EdU^{+}/DCX^{+}) or
490 a NES^{+} qRGL NSC-like process (EdU^{+}/NES^{+}) (**Figure 5E, Supplementary Figure 7A**), one can
491 compare relative differentiation states achieved by these cells, and a higher proportion of EdU^{+}/DCX^{+}
492 cells would indicate precocious differentiation. Indeed, while steady-state cell numbers in this
493 experiment again showed overall decreases in both NES^{+} qRGL NSCs and DCX^{+} NPCs in *Kmt2d*$^{+/\beta geo}$
494 mice compared to wild-type (**Supplementary Figure 7A'**), among pulsed cells the *Kmt2d*$^{+/\beta geo}$ mice
495 exhibited a higher fraction of EdU^{+}/DCX^{+} immature neurons than *Kmt2d*$^{+/+}$ littermates (**Figure 5E'**). In
496 other words, *Kmt2d*$^{+/\beta geo}$ mouse DG cells born in the preceding 2 weeks occupied a more advanced
497 differentiation state than wild-type DG cells born in the same window.
498 Together, studies of adult neurogenesis in *Kmt2d*$^{+/\beta geo}$ mice suggest that neurodevelopmentally
499 relevant *in vivo* consequences of KMT2D loss accompany cell cycle defects and transcriptional
500 profiles similar to those expected from findings in our mouse and human *in vitro* models. We found
501 reduced representation of adult DG NPCs beginning with the earliest NSC stages, suggesting baseline
502 depletion of the adult-born NSC pool in the KS1 mouse model of ID. Furthermore, the demonstration
503 of precocious differentiation *in vivo* supports our hypothesis that transcriptional perturbations due to
504 KMT2D loss may lead to premature neuronal differentiation. Thus, changes in proliferation, cell cycle,
505 gene expression, and neural differentiation due to KMT2D deficiency that we observed cell-
506 autonomously *in vitro,* are recapitulated *in vivo* in the heterogeneous adult DG NPC niche of KS1 mice,
507 providing a putative cellular mechanism of ID etiology in KS1.
508
509
510

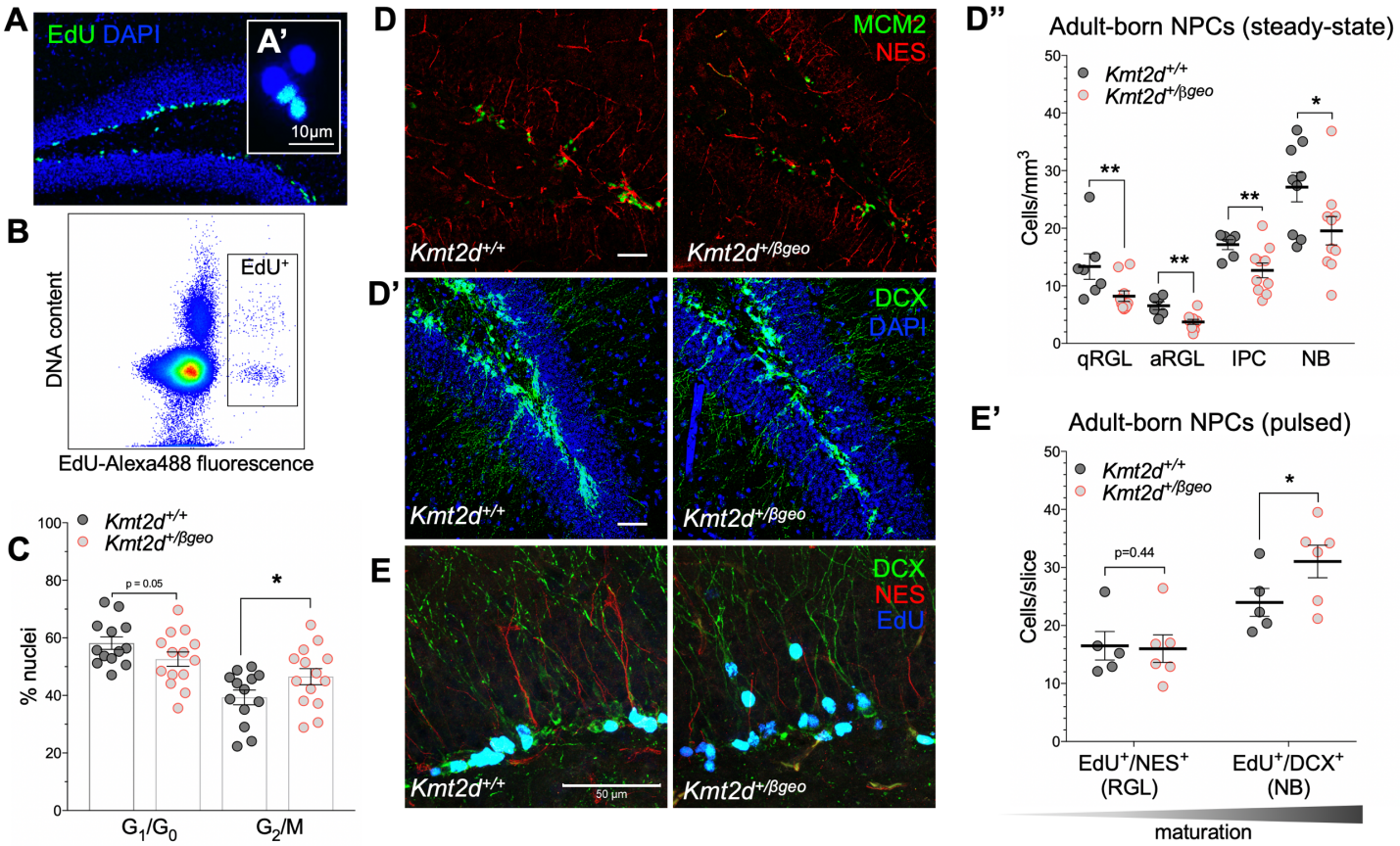


Figure 5. *In vivo* NSC depletion and precocious NPC differentiation in a mouse model of Kabuki syndrome. (**A**) Immunostaining of cycling, EdU-labeled DG cells and isolated nuclei from microdissected DG (**A'**) with flow cytometric purification (**B**) of labeled nuclei. (**C**) Cell cycle analysis of EdU$^+$ nuclei from *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice sampled 16 hours after EdU pulse, using DAPI fluorescence (13-14 mice per genotype). (**D-D''**) Representative immunostaining of steady-state neurogenesis markers in the dentate gyrus (DG) of adult *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice (6-10 mice per genotype, 7-10 slices per mouse). Earliest stages of nestin (NES)$^+$ radial glia-like (RGL) NSCs in either quiescent (MCM2$^-$) or activated (MCM2$^+$) states (qRGL and aRGL, respectively), MCM2$^+$/NES$^-$ intermediate progenitor cells (IPCs) (**D**), and DCX$^+$ NPCs (**D'**). (**D''**) Quantification of qRGL, aRGL, IPC, and DCX$^+$ NPCs cell densities in adult *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\beta/geo}$ mice. (**E**) Representative immunostaining from *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice (5-6 mice per genotype, 10 slices per mouse) of EdU pulse-labeled cells extending a NES$^+$ process (NSCs) or DCX$^+$ process (NPCs), with (**E'**) quantification of double-labeled cells. Bars indicate mean ± SEM. Asterisks indicate significance from control mean, Student's t-test (*p<0.05, **p<0.01, ***p<0.001).

528 **Discussion**

529 The Mendelian disorders of epigenetic machinery (MDEM) are a major emerging cause of ID
530 (2). In these congenital syndromes, mutations in genes encoding chromatin-modifying enzymes
531 disrupt local chromatin states and gene expression, resulting in multisystem anomalies. Though
532 individually rare, the MDEM provide valuable insights to the functional consequences of altered
533 epigenetic states. The ID disorder KS1 is caused by mutations in the histone methyltransferase
534 *KMT2D*, but mechanistic links to neurodevelopmental and cognitive consequences in patients have
535 not been established. KS1 diagnoses are typically made after birth, but the inherent reversibility of
536 chromatin modifications raises the possibility that a detailed understanding of KMT2D activity in
537 neuronal cells could identify specific molecular targets for postnatal interventions in KS1 patients.

538 Here, we report that KMT2D-deficient human and mouse neurodevelopment models, *in vitro*
539 and *in vivo*, demonstrate similar patterns of transcriptional suppression, proliferative defects, and
540 precocious cellular differentiation. These phenotypes were cell-autonomous *in vitro*, suggesting that 1)
541 chromatin and gene expression studies in neurogenic cell types could yield disease-relevant KMT2D
542 targets and 2) these cellular models provide platforms for screening of novel therapeutic strategies or
543 targeted manipulations. In this study, we perform extensive transcription and KMT2D profiling in these
544 models and observe systematic suppression of hypoxia response pathways, particularly among
545 HIF1A-regulated genes that are also directly KMT2D-bound in neuronal cells. Physically overlapping
546 KMT2D- and HIF1A-bound genomic loci were observed even across different tissues, namely on gene
547 promoters in neuronal and cardiac cells, raising the possibility of shared etiologies in embryonically
548 distinct KS1-affected organ systems. Indeed, KMT2D-deficient neuronal cells, in contrast to isogenic
549 wild-type cells, were unable to mount a characteristic hypoxia-inducible gene activation response
550 when exposed to low-oxygen conditions, demonstrating oxygen sensing defects in KS1, and future
551 studies could interrogate KS1 neuronal differentiation phenotypes upon targeted manipulations of
552 oxygen sensing loci.

553 The implication of hypoxia-responsive expression defects in KS1 suggests clinical relevance of
554 two recent neurodevelopmental findings. First, the adult hippocampal NPC niche harbors locally
555 hypoxic, but dynamic microenvironments, and the hypoxic state positively influences NPC survival
556 (32). Thus, compromised oxygen-sensing is expected to render NPCs particularly vulnerable to
557 changes in oxygen pressure experienced by differentiating, migrating NPCs. Second, NPC maturation
558 is coupled to a metabolic rewiring from glycolysis, in early NPCs, to oxidative phosphorylation in
559 immature neurons. Zheng and colleagues (17) recently found this metabolic switch, marked by
560 suppression of glycolytic genes (*Ldha, Hk2*), to be essential for neuronal maturation. We observed

561 suppression of these and related genes in KS1 models, accompanied by upregulation of pro-neuronal
562 differentiation genes, as well as precocious differentiation of pulsed adult DG NPCs and concomitant
563 depletion of the NSC pool *in vivo* in the KS1 mouse model.

564 Analogous findings of premature activation of terminal differentiation genes, reduced
565 proliferation, and precocious maturation in KMT2D-depleted keratinocytes were recently linked to
566 disorganized epidermal stratification (10). Furthermore, in KMT2D-deficient cardiomyocytes, loss of
567 H3K4me2 at KMT2D-bound hypoxia response genes associated with cell cycle and proliferative
568 defects in heart development (6). In contrast, conditional KMT2D deletion in B cells conferred
569 proliferative advantage and impaired cell maturation in a stage-dependent manner, despite significant
570 up-regulation of differentiation genes (8, 9). Thus, while KMT2D's role in enhancer-mediated, cell type-
571 specific gene expression during differentiation is well-established (11), phenotypic manifestations
572 appear cell type- and stage-dependent. We now extend phenotypes of transcriptional perturbance,
573 hypoxia response, cell cycle, proliferation, and premature differentiation to KMT2D-deficient neuronal
574 contexts. Phenotypic concordance across tissues of disparate embryonic origin suggests that KMT2D
575 targets important to KS1 phenotypes support basic cellular homeostatic functions related to
576 housekeeping, energy production, and cell cycle progression, rather than genes with purely brain-
577 specific function. Furthermore, we report concordant phenotypes both from nonsense *KMT2D*
578 mutations (patient iPSCs and NPCs), and mutations limited to the KMT2D methyltransferase domain
579 (HT22 neuronal cells, *Kmt2d*$^{+/\beta geo}$ mice), indicating that loss of either gene dosage or catalytic function
580 of KMT2D can be pathogenic.

581 Present results indicate that the adult hippocampal neurogenesis defect, which we previously
582 found to associate with visuospatial memory defects in *Kmt2d*$^{+/\beta geo}$ mice, originates from a depletion of
583 the DG NSC pool itself rather than from stalling at any particular cell type transition. In fact, despite
584 having fewer NPCs overall, we observed the *Kmt2d*$^{+/\beta geo}$ NPC population to occupy a more advanced
585 maturation state than that of wild-type littermates. While adult-born NPCs comprise a small fraction of
586 the adult hippocampus, they wield an outsized influence on DG networks which, strikingly, extends
587 even to the contralateral hippocampus in promoting visuospatial memory ability (33). This bilateral,
588 network-mediated performance improvement was abolished upon transient silencing of young, but not
589 older, adult-born DG NPCs, suggesting that NPC maturation rates and cognitive effects are coupled
590 (33). Such studies increase the likelihood that precocious differentiation, in the context of developing
591 neurons, could measurably impact visuospatial memory. Furthermore, multispecies comparisons
592 demonstrate that measured decreases in neurogenesis rates are consistent with accelerated neuronal
593 maturation rates across the lifespan (34).

594 The apparent paradox of increased HIF1A activation, despite blunted hypoxia-responsive
595 expression in *Kmt2d*$^{+/\Delta}$ and *Kmt2d*$^{\Delta/\Delta}$ neuronal cells raises two possibilities. First, chronic HIF activity
596 may reflect cellular compensatory efforts in the absence of negative feedback from hypoxia gene
597 induction. In this case, heterochromatic environments at HIF-binding genes could prevent induction.
598 Alternatively, cellular oxygen sensing could be coupled to chromatin states and gene expression in a
599 HIF-independent manner. A pair of independent studies recently discovered direct oxygen sensing by
600 KDM6A/UTX, the H3K27 demethylase lost in KS2 patients, as well as the H3K4/H3K36 demethylase
601 5A (KDM5A), directly producing HIF-independent changes in chromatin states, cell differentiation, and
602 cell fate (35, 36). Specifically, independent findings of upregulated histone methylation at H3K4,
603 H3K27, H3K9, and H3K36 in hypoxia, linked directly to cell maturation in widespread cell types,
604 suggest that KS-associated targets are unusually vulnerable to oxygen pressure changes, and support
605 mechanistic links between our neuronal hypoxia response and differentiation findings. Remarkably,
606 the latter study found that genes showing hypoxia-upregulated H3K4me3 peaks were significantly
607 enriched for HIF targets, and these loci were almost exclusively in promoters, precisely where we
608 observed KMT2D peaks to overlap HIF1A peaks. Strikingly, loss of KDM5A, whose activity opposes
609 that of KMT2D at H3K4 sites, resulted in elevated hypoxia-responsive expression, i.e. an effect
610 inversely proportional to the present KS1-associated transcriptional suppression of hypoxia-response
611 genes such as *Klf10* and *Bnip3l.* A number of histone demethylases, and at least 33 chromatin
612 modifiers in total, are known to impact hypoxia response genes, with 11 of these associating with
613 developmental disease or cancers, yet KMT2D and other histone methyltransferases have not yet
614 been implicated (37).
615 In summary, our findings suggest that KMT2D deficiency disrupts neurogenesis by negatively
616 impacting stem cell maintenance functions including cell cycle progression, proliferation, and survival,
617 with concomitant NSC depletion and precocious neuronal differentiation *in vivo*. Chromatin and
618 expression profiling identified KMT2D- and HIF-regulated gene programs suppressed across KS1
619 model systems, which mechanistically link hypoxia pathways with phenotypes observed by us and
620 others. We functionally demonstrate KMT2D-dependent neuronal hypoxia responses and *in vivo*
621 neurodevelopmental disruptions predicted from our cellular work. Together, our findings support a
622 model in which KMT2D loss transcriptionally suppresses oxygen sensing programs critical to early
623 NPC maintenance, resulting in precocious differentiation and exhaustion of precursors required for
624 adult hippocampal neurogenesis.
625

626 **Methods**

627 **Animals**

628 The *Kmt2d*$^{+/\beta geo}$ allele (Mll2Gt(RRt024)Byg) was generated by Bay Genomics (University of California)
629 through the random insertion of a gene trap vector. *Kmt2d*$^{+/\beta geo}$ mice were fully backcrossed to
630 C57Bl/6 background over more than 10 generations in the Bjornsson lab, and are born roughly in
631 expected Mendelian ratios when bred to wild-type. Animals were housed in a 14-hour light/10-hour
632 dark cycle with free access to food and water. All experiments compare mutant mice against age- and
633 sex-matched wild-type littermates.

634

635 **Patient-derived iPSCs, NPCs, and fibroblasts**

636 Media and reagents are listed (**Supplementary Table 10**). Skin biopsy fibroblasts were cultured from
637 molecularly confirmed KS1 patients (KS1-1, KS1-3, KS1-4) of the JHU Epigenetics and Chromatin
638 Clinic. Patient KS1-1 was consented for stem cell derivation and cells were reprogrammed using non-
639 integrating Sendai virus vectors (CytoTune-iPS 2.0, ThermoFisher Scientific, Waltham, MA). 5 days
640 post-induction, cells were transferred to mouse embryonic fibroblast (MEF) feeder plates in iPSC
641 media and monitored for colony formation. 21 days post-induction, colonies were manually selected
642 for optimal iPSC morphology and quality for propagation. Karyotype analysis by G-banding confirmed
643 46, XX normal female karyotype in KS1-1. Generation and characterization of healthy control lines (C3-
644 1 and C1-2) has been previously described (27). Feeder MEFs were derived from E13.5 CF-1 mouse
645 embryos and mitotically inactivated by irradiation. iPSCs were enzymatically passaged every 4-8 days
646 using collagenase and mechanical dissociation. NPCs were induced from iPSCs as previously
647 described (28). Briefly, we synergistically inhibited signaling of glycogen synthase kinase 3 (GSK3),
648 transforming growth factor β (TGF-β), γ-secretase, and Notch signaling pathways using small
649 molecules CHIR99021 (4 μM), SB431542 (3 μM), and Compound E (0.1 μM), in the presence of hLIF
650 (10 ng/ml) and ROCK inhibitor (5 μM) for 7 days. NPCs were split with Accutase and propagated in
651 neural induction medium in feeder-free conditions on a Matrigel.

652

653 **CRISPR-Cas9 deletions**

654 Media and reagents are listed (**Supplementary Table 10**). HT22 mouse hippocampal cells were
655 obtained commercially (MilliporeSigma, Burlington, MA) and maintained in HT22 media (DMEM with
656 10% FBS, pen/strep, GlutaMAX). sgRNAs targeting two loci spanning the *Kmt2d* SET domain-
657 encoding region, with cut sites in exon 52 and either exon 54 or intron 54, were designed and selected
658 for on-target activity scores >90% (crispr.mit.edu, **Supplementary Table 10**), and integrated into

659 Cas9 plasmid (pSpCas9BB-2A-puro v2.0 (PX459), Addgene, Watertown, MA). Plasmids were
660 delivered to HT22 cells at 20% confluency using Lipofectamine 2000 according to manufacturer
661 protocol. After puromycin selection, heterozygous and homozygous cells were clonally expanded and
662 maintained in culture according to standard protocols. Targeted clone genotypes were verified by
663 PCR (primers listed) and Sanger sequenced at the Johns Hopkins Genetic Resources Core Facility
664 (GRCF). A subset of clones appearing heterozygous by PCR (upper and lower DNA bands of expected
665 size) were found to bear strand invasion in the upper band (**Supplementary Figure 1B**) and were
666 removed from analyses. One such line, mutant Clone 1, was thus grouped with homozygous clones 2
667 and 3 for analysis, as both alleles were targeted in all three clones. Subsequent cellular assays were
668 performed using Sanger-verified heterozygous and homozygous lines (Figures 1 and 2).
669
670 **RNA-seq in HT22 cells: library preparation**
671 Cells were plated at equal density and sampled at 60% confluency. Total RNA was isolated from three
672 biological replicates of *Kmt2d*$^{\Delta/\Delta}$ clones and *Kmt2d*$^{+/+}$ wild-type parental cells using Direct-Zol RNA
673 MicroPrep (Zymo Research, Irvine, CA), and libraries were constructed in technical triplicate using
674 NEBNext Poly(A) Magnetic Isolation Module and NEBNext UltraII RNA Library Prep Kit for Illumina
675 (New England BioLabs, Ipswich, MA), with size selection by AMPure XP beads (Beckman Coulter,
676 Brea, CA), according to manufacturer protocols. Library quantification and quality checks were done
677 using KAPA Library Quantification Kit for Illumina (KAPA Biosystems, Wilmington, MA), High Sensitivity
678 DNA Kit on BioAnalyzer (Agilent, Santa Clara, CA), and Qubit dsDNA HS Assay (Life Technologies,
679 Carlsbad, CA). Paired end 50 bp reads were obtained for pooled libraries using Illumina HiSeq 2500 at
680 the JHU GRCF High Throughput Sequencing Center.
681
682 **RNA-seq in HT22 cells: data analysis**
683 Sequencing reads were pseudoaligned to the mouse reference transcriptome (GRCm38) and
684 transcript abundances were subsequently quantified using Salmon (38). We then used the tximport R
685 package (39) to convert the transcript abundances into normalized gene-level counts, by setting the
686 "countsFromAbundance" parameter equal to "lengthScaledTPM". Next, we used the edgeR (40) and
687 limma (41) R packages to log2 transform these gene-level counts, estimate the mean-variance
688 relationship, and calculate weights for each observation. In order to account for the correlation
689 between technical replicates of the same clone when performing the differential analysis, we fit a
690 mixed linear model, using the function "duplicateCorrelation" from the statmod R package (42) to
691 block on clone. The differential analysis was then performed using the limma R package. Differentially

692 expressed genes were called with 0.05 as the cutoff for the False Discovery Rate (FDR). When
693 performing the principal component analysis, transcript abundances were first converted into gene-
694 level counts using the tximport R package, with the “countsFromAbundance” parameter equal to “no”.
695 Then, we applied a variance stabilizing transformation to these gene-level counts using the “vst”
696 function from the DESeq2 R package (43) with the parameter “blind” set to “TRUE”, and subsequently
697 estimated the principal components using the 1000 most variable genes.
698
699 **scRNA-seq: library preparation**
700 NPCs were induced in parallel from each iPSC line (KS1-1, C1-2, C3-1) under identical conditions, and
701 passaged three times before sampling. iPSCs were detached from MEF feeders using collagenase
702 (200 units/ml). iPSCs and NPCs were dissociated to single-cell suspension using Accutase. Cell
703 counts and viability were analyzed using Countess II Automated Cell Counter (ThermoFisher
704 Scientific). scRNA-seq libraries were created with Chromium Single Cell 3’ Library & Gel Bead Kit v2
705 (10x Genomics) according to manufacturer protocol. Targeted cell recovery for each sample was
706 5,000 cells. Sufficient cDNA for library construction was achieved using 20 amplification cycles for
707 iPSC libraries and 16 cycles for NPC libraries. Sample indexing was achieved using 11 PCR cycles for
708 iPSC libraries and 5 cycles for NPC libraries. scRNA-seq libraries were sequenced using Illumina
709 NextSeq 500.
710
711 **scRNA-seq: data analysis**
712 Sequencing output was processed through the Cell Ranger 2.1.0 preprocessing pipeline using default
713 parameters with the exception of --expect-cells=5000 for `cellranger count` and --normalize=none for
714 `cellranger aggr`. Reads were quantified against hg19 using the 10x reference genome and
715 transcriptome builds (refdata-cellranger-GRCh38-1.2.0). The aggregated raw count matrix was then
716 used as input for the Monocle2 single-cell RNAseq framework. Differential gene expression analysis
717 was performed on all NPCs and iPSCs with respect to genotype (KS1 patient vs healthy control) and
718 was performed using the Monocle2 (44) likelihood ratio test (0.1% FDR, Monocle2 LRT, Benjamini-
719 Hochberg corrected) with `num_genes_expressed` added as a nuisance parameter to both the full and
720 reduced models. The directionality of the differential gene test was determined by calculating the
721 mean gene expression across all KS1 patient-derived and healthy control cells respectively, evaluating
722 the relative fold change. High-variance genes were selected as those with a positive residual to the
723 estimated dispersion fit and a mean number of reads per cell >=0.0005. Cell cycle stage was
724 determined by profiling cell cycle associated genes across all cells and assigning cell cycle state using

725 the R/Bioconductor package scran (45). Dimensionality reduction and visualization was performed via
726 UMAP (29) on the log10(counts + 1) of the high variance genes in the NPC dataset. The first 10
727 principal components were used as input for UMAP using the default parameters of the R/CRAN
728 package umap. Cells were assigned to clusters using Monocle2's implementation of the louvain
729 community detection algorithm. Learned clusters were then aggregated by hand based on marker
730 gene expression into three clusters ("Differentiating", "Transitioning", "Cycling"). Differential gene
731 expression within clusters, and between genotypes was performed as described above. The
732 "Differentiating" cluster was then segregated, and a smooth line was fitted using a linear regression.
733 This line was determined to represent the direction of differentiation by examination of marker genes
734 (**Supplemental Figure 5B-E''**). The residuals of this fit were then plotted and deciles were calculated
735 containing equal number of cells along the axis of differentiation. The number of cells in each decile
736 was then counted with respect to genotype.
737
738 **ChIP-seq: library preparation**
739 *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ HT22 cells were sampled at 70% confluency and processed for pull-down with
740 ChIP-grade KMT2D antibody (Millipore Sigma) according to ENCODE guidelines. Sonicated, reverse-
741 crosslinked chromatin served as input control. Briefly, ~300 million cells per cell line were crosslinked
742 in 1% formaldehyde, quenched with 0.125 M glycine, and cell lysate supernatants were collected for
743 immediate processing or snap-frozen for storage at -80°C. Nuclei diluted in 1 ml RIPA buffer were
744 sonicated using Bioruptor (Diagenode, Denville, NJ) for 6 cycles of 5 minutes (60 seconds on/30
745 seconds off) in ice-cold water bath. Supernatants containing sheared chromatin were pre-cleared with
746 Protein A Dynabeads (ThermoFisher Scientific) and incubated overnight at 4°C with 8 µg KMT2D
747 antibody. ChIP DNA was recovered by Dynabead incubation (overnight at 4°C plus 6 hours at room
748 temperature) before 6 sequential salt washes of increasing stringency, then eluted and reverse
749 crosslinked overnight at 65°C. DNA was purified using DNA Clean and Concentrator (Zymo Research)
750 and quantified using High Sensitivity DNA Kit on BioAnalyzer (Agilent), and Qubit dsDNA HS Assay
751 (Life Technologies). DNA libraries were constructed using NEBNext UltraII DNA Library Prep Kit for
752 Illumina (New England BioLabs) and quantified using KAPA Library Quantification Kit for Illumina
753 (KAPA Biosystems). Paired end 75 bp reads were obtained for pooled libraries using Illumina HiSeq
754 2500 at the JHU GRCF High Throughput Sequencing Center.
755
756 **ChIP-seq: data analysis**

757 Sequencing reads were aligned to the mouse reference genome (mm10) using Bowtie2 (46). Then,
758 duplicate reads were removed with the function MarkDuplicates from Picard
759 (http://broadinstitute.github.io/picard/). Peaks were subsequently called using MACS2 (47), with the
760 "keep-dup" parameter equal to "all". After peak calling, we excluded all peaks that overlapped with
761 blacklisted regions provided by ENCODE
762 (https://sites.google.com/site/anshulkundaje/projects/blacklists). To identify genes whose promoters
763 were bound by KMT2D, we defined promoters as 10kb regions centered around the TSS. We obtained
764 TSS coordinates using the "exonsBy" function from the "TxDb.Mmusculus.UCSC.mm10.knownGene"
765 R package (Team BC, Maintainer BP (2018). TxDb.Mmusculus.UCSC.mm10.knownGene: Annotation
766 package for TxDb object(s). R package version 3.4.4.), with the "by" parameter equal to "gene".
767
768 **Purification of EdU$^+$ nuclei**
769 Mice were given 150 mg/kg EdU by intraperitoneal injection and sampled after 16 hours. Dentate
770 gyrus (DG) was micro-dissected in ice-cold PBS immediately following sacrifice by halothane
771 inhalation. Total nuclei were purified as described (48) with addition of RNase inhibitor to all buffers
772 (**Supplementary Table 10**). Briefly, bilateral DG from individual mice were dounce-homogenized in 1
773 ml lysis buffer and layered above a sucrose gradient for ultracentrifugation at 28,600 RPM for 2 hours
774 at 4°C. Supernatant was aspirated and nuclei were resuspended in Click-iT EdU AlexaFluor-488 Flow
775 Cytometry Kit buffer with addition of RNAse inhibitor, and incubated 30 minutes at room temperature.
776 Samples were passed through 40 µm filter, stained with 1 µg/ml DAPI, and kept on ice before sorting.
777 Nuclear lysates processed identically from non-EdU-injected mice served as negative controls during
778 sorting with Beckman Coulter MoFlo Cell Sorter as above. Cell cycle analysis by DNA content was
779 performed with analysis gates discriminating 2N and 4N cells by DAPI fluorescence.
780
781 **RNA-seq: EdU$^+$ nuclei**
782 Purified EdU$^+$ nuclei from *Kmt2d*$^{+/\beta geo}$ female mice (500 nuclei pooled from 3 mice) and sex- and age-
783 matched littermates (500 nuclei pooled from 3 mice) were sorted into Smart-Seq 2 lysis buffer (2 µL
784 Smart-Seq2 lysis buffer with RNase inhibitor, 1 µL oligo-dT primer, and 1 µL dNTPs), briefly spun by
785 tabletop microcentrifuge, and snap-frozen on dry ice. Lysates were stored at -80°C until cDNA
786 conversion. Nuclei were processed according to a modified Smart-seq2 protocol (49). Briefly, lysates
787 were thawed to 4°C, heated to 72°C for 5 minutes, and immediately placed on ice. Template-
788 switching first-strand cDNA synthesis was performed using a 5'-biotinylated TSO oligo. cDNAs were
789 amplified using 20 cycles of KAPA HiFi PCR and 5'-biotinylated ISPCR primer. Amplified cDNA was

790 cleaned using 1:1 ratio of Ampure XP beads and approximately 250 pg was input to a one-quarter-
791 sized Nextera XT tagmentation reaction. Tagmented fragments were amplified for 12 enrichment
792 cycles and dual indexes were added to each well to uniquely label each library. Concentrations were
793 assessed with Quant-iT PicoGreen dsDNA Reagent (Invitrogen) and samples were diluted to ~2nM
794 and pooled. Pooled libraries were sequenced on the Illumina HiSeq 2500 platform to a target mean
795 depth of ~8 x 105 bp paired-end fragments per cycle at the Johns Hopkins Genetic Resources Core
796 Facility. For all libraries, paired-end reads were aligned to the mouse reference genome (mm10) using
797 HISAT2 (50) with default parameters except: -p 8. Aligned reads from individual samples were
798 quantified against a reference genome (GENCODE vM8) (51). Quantification was performed using
799 cuffquant with default parameters and the following additional arguments: --no-update-check –p 8.
800 Normalized expression estimates across all samples were obtained using cuffnorm with default
801 parameters (52).
802
803 **RT-qPCR**
804 Total RNA was isolated by RNeasy Mini (Qiagen, Venlo, Netherlands) and cDNA libraries were
805 constructed with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA)
806 according to manufacturer protocols. Experiments were performed in technical triplicate and no
807 explicit power analysis was used to determine sample numbers. Probes were from Taqman
808 (ThermoFisher Scientific, **Supplementary Table 10**).
809
810 **Immunostaining, confocal imaging, and processing**
811 Coronal brain sections (30 μm) through the entire dentate gyrus (every sixth slice) were maintained in
812 serial order. Brains were paraformaldehyde-fixed by transcardial perfusion and post-fixed in 4% PFA
813 overnight at 4°C before cryoprotection by sequential overnight incubations at 4°C with sucrose
814 concentrations of 10%, 20% and 30% in phosphate buffer. Brains were sectioned by cryostat (Leica,
815 Wetzlar, Germany), directly mounted to charged slides, and stored at -80°C. Antigen retrieval
816 (DakoCytomation) was performed at 95°C for 20 minutes. After blocking, overnight incubation at 4°C
817 in primary antibodies (**Supplementary Table 10**) was followed by 2-hour room temperature incubation
818 with AlexaFluor-conjugated secondary antibodies at 1:500 dilution. Blocking buffer of 0.1% Triton X-
819 100 in TBS contained 6% serum matched to the host species of secondary antibody. Tiled, z-stacked
820 images were acquired using an inverted Zeiss LSM780 FCS AxioObserver confocal microscope and
821 Zen software (Zeiss) to encompass the entire GCL area of each section using 10X, 20X, or 40X
822 objective. Images were quantified using Imaris (BitPlane, Zurich, Switzerland) by experimenters

823 blinded to genotype. Cell counts for each GCL slice analyzed were corrected by the total GCL area
824 multiplied by z-thickness of the image, and expressed as cells/mm^3. For pulse-label experiments, mice
825 were injected with 150 mg/kg EdU in saline every 48 hours from P41-P54 and sampled on P55 to be
826 processed for staining as before. DCX^+ neuroblast distance from SGZ plane was measured in Fiji (NIH,
827 Bethesda, MD). No explicit power analysis was used to determine sample numbers. iPSCs and NPCs
828 were plated at equal density and fixed with 4% PFA for 15 minutes at room temperature, blocked and
829 permeabilized with 0.25% Triton X-100 and 10% donkey serum in PBS, then stained for 1 hour at
830 room temperature in primary antibodies and 1 hour at room temperature in secondary antibodies.
831 iPSC and NPC images were acquired using EVOS FL Cell Imaging System (ThermoFisher Scientific)
832 and analyzed using Fiji. HT22 cells were cultured on glass coverslips and stained and imaged as
833 above using Zeiss LSM780.
834
835 **FACS and analysis**
836 Flow cytometric analysis was carried out using FACSverse instrument and FACSsuite software (BD
837 Biosciences), and FACS was performed using Beckman Coulter MoFlo Cell Sorter with proper gate
838 settings and doublet discrimination (**Supplemental Figure 3M, Supplemental Figure 6A**) at the Johns
839 Hopkins Bloomberg School of Public Health Flow Cytometry Core. Cell samples containing at least
840 10,000 cells were analyzed from technical triplicate culture wells and analysis was performed in
841 FlowJo v10 (Tree Star Inc, Ashland, OR). No explicit power analysis was used to determine sample
842 numbers. Feeder-free iPSCs and NPCs, HT22 cells, or patient fibroblasts were enzymatically
843 dissociated to single cell suspension. For antibody staining, cells were fixed with 4% PFA for 15
844 minutes at room temperature, blocked and permeabilized with 0.25% Triton X-100 and 10% donkey
845 serum in PBS, and incubated with primary antibody for 30 minutes at room temperature, followed by
846 30 minutes in secondary antibody and 5 minutes in DAPI (1 μg/ml) for DNA content. Unstained and
847 secondary-only samples served as control. For proliferation analysis, cells were sampled after 30-
848 minute pulse of EdU (10 μM) using Click-iT EdU Flow Cytometry Assay (ThermoFisher Scientific)
849 according to manufacturer protocol. Proteostat reagent (Enzo Life Sciences, Ann Arbor, MI), CellTrace
850 Violet, and CellEvent caspase-3/7 reagent (ThermoFisher Scientific) were used according to
851 manufacturer protocol. Proteasome inhibition was achieved using 5 μM MG-132 for 16 hours (Enzo
852 Life Sciences). Cell cycle synchronization in HT22 cells or patient fibroblasts was achieved by
853 treatment with 250 ng/ml nocodazole (Sigma, St. Louis, MO) for 18 hours before being released into
854 complete media free of nocodazole for indicated time points.
855

856 **Magnetic Resonance Imaging (MRI)**

857 3D T2-weighted MRI (9.4T) was performed in mouse brains following PFA perfusion-fixation by the
858 JHU Department of Radiology Division of MR Research. Atlas-based volumetric analysis was
859 performed in 25 brain regions, including hippocampus and olfactory bulb. Structural and ventricular
860 volumes were corrected by total brain volume for analysis (DtiStudio). No explicit power analysis was
861 used to determine sample numbers.

862

863 **Statistics**

864 Statistical analyses (excluding high-throughput data) were calculated using GraphPad Prism (version
865 7.0b), and p-values less than or equal to 0.05 were considered significant (*$p<0.05$, **$p<0.01$,
866 ***$p<0.001$). Gene set enrichments were calculated by Overrepresentation analysis (ORA) with
867 WebGestalt (53), or with Fisher's Exact Test function using R version 3.5.2. Statistics for high-
868 throughput experiments are described in Methods.

869

870 **Study Approval**

871 All experiments were performed using mouse protocols approved by the Animal Care and Use
872 Committee of Johns Hopkins University School of Medicine and are in accordance with National
873 Institutes of Health (NIH) guidelines for mouse care and handling. Informed consent for KS1 patient
874 cell collection and iPSC derivation was obtained according to institutional IRB and ISCRO protocols
875 approved by Johns Hopkins University School of Medicine.

876

877 **Acknowledgments**

878 Imaging was performed at the Johns Hopkins School of Medicine Microscope Facility with NIH
879 support (S10OD016374). Karyotypes were analyzed by the Johns Hopkins Cytogenomics Laboratory
880 with support from NICHD (1 U54 HD079123-01A1). We thank Michael Sherman for assistance with
881 image quantification, Manisha Aggarwal for MRI analysis, Mark Sandusky for schematic designs.
882 Hongjun Song and Kai Ge provided critical reagents and advice. Hal Dietz and Gregg Semenza
883 provided indispensable guidance.

884

885 **Conflict of interest statement**

886 HTB is a consultant for Millennium Pharmaceuticals, otherwise authors declare no conflict of interest exists.

887

888 **Funding**

HTB is funded through an Early Independence Award from the National Institutes of Health (NIH, DP5OD017877) and through a grant from the Louma G. Foundation.

Author contributions

GAC and HTB conceived the study; GAC and HTB wrote the manuscript; GAC, HNN, GC, JDR, LZ performed experiments; GAC, LB, JA, KDH and LG analyzed data.

Data availability

All high-throughput data is publicly available. For HT22 cell RNA-seq and ChIP-seq, and DG nuclei RNA-seq, see GEO Accession #GSE126167. For scRNA-seq data, see GEO Accession #GSE126027. Scripts for scRNA-seq analysis are available at https://github.com/Jaugust7/Kabuki-Syndrome-scRNA-analysis.

901 **References**

902 1. Schuettengruber B, Bourbon H-M, Di Croce L, Cavalli G. Genome regulation by polycomb and
903 trithorax: 70 years and counting.. *Cell* 2017;171(1):34–57.
904 2. Bjornsson HT. The Mendelian disorders of the epigenetic machinery.. *Genome Res.*
905 2015;25(10):1473–1481.
906 3. Ng SB et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome.. *Nat.*
907 *Genet.* 2010;42(9):790–793.
908 4. Miyake N et al. KDM 6 A Point Mutations Cause K abuki Syndrome.. *Human mutation*
909 2013;34(1):108–110.
910 5. Hannibal MC et al. Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome.. *Am. J.*
911 *Med. Genet. A* 2011;155A(7):1511–1516.
912 6. Ang S-Y et al. KMT2D regulates specific programs in heart development via histone H3 lysine 4 di-
913 methylation.. *Development* 2016;143(5):810–821.
914 7. Ortega-Molina A et al. The histone lysine methyltransferase KMT2D sustains a gene expression
915 program that represses B cell lymphoma development.. *Nat. Med.* 2015;21(10):1199–1208.
916 8. Zhang J et al. Disruption of KMT2D perturbs germinal center B cell development and promotes
917 lymphomagenesis.. *Nat. Med.* 2015;21(10):1190–1198.
918 9. Lee J-E et al. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during
919 cell differentiation.. *Elife* 2013;2:e01503.
920 10. Lin-Shiao E et al. KMT2D regulates p63 target enhancers to coordinate epithelial homeostasis..
921 *Genes Dev.* 2018;32(2):181–193.
922 11. Froimchuk E, Jang Y, Ge K. Histone H3 lysine 4 methyltransferase KMT2D.. *Gene* 2017;627:337–
923 342.
924 12. Bjornsson HT et al. Histone deacetylase inhibition rescues structural and functional brain deficits in
925 a mouse model of Kabuki syndrome.. *Sci Transl Med* 2014;6(256):256ra135.
926 13. Boisgontier J et al. Anatomical and functional abnormalities on MRI in kabuki syndrome..
927 *Neuroimage Clin.* [published online ahead of print: November 19, 2018]; doi:10.1016/j.nicl.2018.11.020
928 14. Gonçalves JT, Schafer ST, Gage FH. Adult neurogenesis in the hippocampus: from stem cells to
929 behavior.. *Cell* 2016;167(4):897–914.
930 15. Moreno-Jiménez EP et al. Adult hippocampal neurogenesis is abundant in neurologically healthy
931 subjects and drops sharply in patients with Alzheimer's disease.. *Nat. Med.* [published online ahead of
932 print: March 25, 2019]; doi:10.1038/s41591-019-0375-9
933 16. Berg DA et al. A common embryonic origin of stem cells drives developmental and adult
934 neurogenesis.. *Cell* [published online ahead of print: March 13, 2019]; doi:10.1016/j.cell.2019.02.010
935 17. Zheng X et al. Metabolic reprogramming during neuronal differentiation from aerobic glycolysis to
936 neuronal oxidative phosphorylation.. *Elife* 2016;5. doi:10.7554/eLife.13374
937 18. Mohyeldin A, Garzón-Muvdi T, Quiñones-Hinojosa A. Oxygen in stem cell biology: a critical
938 component of the stem cell niche.. *Cell Stem Cell* 2010;7(2):150–161.
939 19. Morimoto BH, Koshland DE. Induction and expression of long- and short-term neurosecretory
940 potentiation in a neural cell line.. *Neuron* 1990;5(6):875–880.
941 20. Wang C et al. Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition..
942 *Proc. Natl. Acad. Sci. USA* 2016;113(42):11871–11876.
943 21. Guo C et al. Global identification of MLL2-targeted loci reveals MLL2's role in diverse signaling
944 pathways.. *Proc. Natl. Acad. Sci. USA* 2012;109(43):17603–17608.
945 22. Semenza GL. HIF-1: mediator of physiological and pathophysiological responses to hypoxia.. *J.*
946 *Appl. Physiol.* 2000;88(4):1474–1480.

947 23. Guimarães-Camboa N et al. Hif1α represses cell stress pathways to allow proliferation of hypoxic
948 fetal cardiomyocytes.. *Dev. Cell* 2015;33(5):507–521.
949 24. Benita Y et al. An integrative genomics approach identifies Hypoxia Inducible Factor-1 (HIF-1)-
950 target genes that form the core response to hypoxia.. *Nucleic Acids Res.* 2009;37(14):4587–4602.
951 25. Sanz JH, Lipkin P, Rosenbaum K, Mahone EM. Developmental profile and trajectory of
952 neuropsychological skills in a child with Kabuki syndrome: implications for assessment of syndromes
953 associated with intellectual disability.. *Clin Neuropsychol* 2010;24(7):1181–1192.
954 26. Harris J, Mahone EM, Bjornsson HT. Molecularly confirmed Kabuki (Niikawa-Kuroki) syndrome
955 patients demonstrate a specific cognitive profile with extensive visuospatial abnormalities.. *J. Intellect.*
956 *Disabil. Res.* [published online ahead of print: February 14, 2019]; doi:10.1111/jir.12596
957 27. Wen Z et al. Synaptic dysregulation in a human iPS cell model of mental disorders.. *Nature*
958 2014;515(7527):414–418.
959 28. Li W et al. Rapid induction and long-term self-renewal of primitive neural precursors from human
960 embryonic stem cells by small molecule inhibitors.. *Proc. Natl. Acad. Sci. USA* 2011;108(20):8299–
961 8304.
962 29. McInnes L, Healy J, Saul N, Großberger L. UMAP: uniform manifold approximation and projection.
963 *JOSS* 2018;3(29):861.
964 30. Benjamin JS et al. A ketogenic diet rescues hippocampal memory defects in a mouse model of
965 Kabuki syndrome.. *Proc. Natl. Acad. Sci. USA* 2017;114(1):125–130.
966 31. Morris AM, Churchwell JC, Kesner RP, Gilbert PE. Selective lesions of the dentate gyrus produce
967 disruptions in place learning for adjacent spatial locations.. *Neurobiol. Learn. Mem.* 2012;97(3):326–
968 331.
969 32. Chatzi C, Schnell E, Westbrook GL. Localized hypoxia within the subgranular zone determines the
970 early survival of newborn hippocampal granule cells.. *Elife* 2015;4:e08722.
971 33. Zhuo J-M et al. Young adult born neurons enhance hippocampal dependent performance via
972 influences on bilateral networks.. *Elife* 2016;5. doi:10.7554/eLife.22429
973 34. Snyder JS. Recalibrating the relevance of adult neurogenesis.. *Trends Neurosci.* 2019;42(3):164–
974 178.
975 35. Chakraborty AA et al. Histone demethylase KDM6A directly senses oxygen to control chromatin
976 and cell fate.. *Science (80-.).* 2019;363(6432):1217–1222.
977 36. Batie M et al. Hypoxia induces rapid changes to histone methylation and reprograms chromatin..
978 *Science (80-.).* 2019;363(6432):1222–1226.
979 37. Luo W, Wang Y. Epigenetic regulators: multifunctional proteins modulating hypoxia-inducible
980 factor-α protein stability and activity.. *Cell Mol. Life Sci.* 2018;75(6):1043–1056.
981 38. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware
982 quantification of transcript expression.. *Nat. Methods* 2017;14(4):417–419.
983 39. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates
984 improve gene-level inferences. [version 2; peer review: 2 approved]. *F1000Res.* 2015;4:1521.
985 40. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression
986 analysis of digital gene expression data.. *Bioinformatics* 2010;26(1):139–140.
987 41. Ritchie ME et al. limma powers differential expression analyses for RNA-sequencing and
988 microarray studies.. *Nucleic Acids Res.* 2015;43(7):e47.
989 42. Giner G, Smyth GK. Statmod: probability calculations for the inverse Gaussian distribution.. *arXiv*
990 *preprint arXiv:1603.06687* 2016;
991 43. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq
992 data with DESeq2.. *Genome Biol.* 2014;15(12):550.

993 44. Trapnell C et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal
994 ordering of single cells.. *Nat. Biotechnol.* 2014;32(4):381–386.
995 45. Lun ATL, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell
996 RNA-seq data with Bioconductor. [version 2; peer review: 3 approved, 2 approved with reservations].
997 *F1000Res.* 2016;5:2122.
998 46. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2.. *Nat. Methods*
999 2012;9(4):357–359.
1000 47. Zhang Y et al. Model-based analysis of ChIP-Seq (MACS).. *Genome Biol.* 2008;9(9):R137.
1001 48. Rizzardi L et al. Neuronal brain region-specific DNA methylation and chromatin accessibility are
1002 associated with neuropsychiatric disease heritability. *BioRxiv* [published online ahead of print: March
1003 24, 2017]; doi:10.1101/120386
1004 49. Picelli S et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells.. *Nat.*
1005 *Methods* 2013;10(11):1096–1098.
1006 50. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements..
1007 *Nat. Methods* 2015;12(4):357–360.
1008 51. Mudge JM, Harrow J. Creating reference gene annotation for the mouse C57BL6/J genome
1009 assembly.. *Mamm. Genome* 2015;26(9–10):366–378.
1010 52. Trapnell C et al. Differential gene and transcript expression analysis of RNA-seq experiments with
1011 TopHat and Cufflinks.. *Nat. Protoc.* 2012;7(3):562–578.
1012 53. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive, powerful,
1013 flexible and interactive gene set enrichment analysis toolkit.. *Nucleic Acids Res.* 2017;45(W1):W130–
1014 W137.
1015
1016
1017

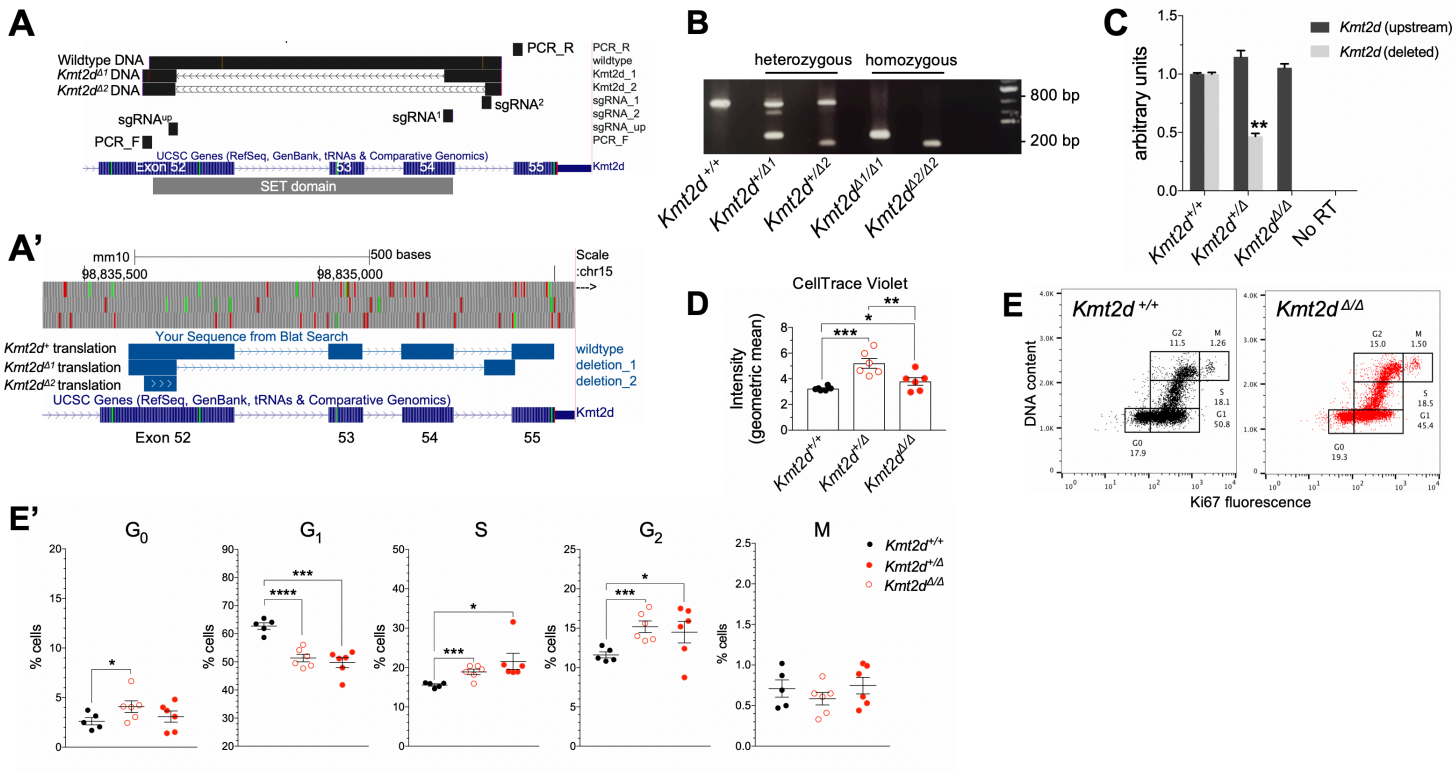
1018 **Figures and Table**
1019 Figure 1. Genetic ablation of the *Kmt2d* SET methyltransferase domain disrupts cell proliferation and
1020 cell cycle in a cell-autonomous manner.
1021 Figure 2. Suppressed transcription of KMT2D- and HIF-regulated hypoxia response genes upon loss
1022 of the *Kmt2d* SET methyltransferase domain in neuronal cells.
1023 Figure 3. KS1 patient-derived cells recapitulate KMT2D-associated defects in proliferation and cell
1024 cycle.
1025 Figure 4. Transcriptional suppression of metabolic genes in cycling cells, and hallmarks of precocious
1026 neuronal differentiation in KS1 patient-derived NPCs.
1027 Figure 5. *In vivo* NSC depletion and precocious NPC differentiation in a mouse model of Kabuki
1028 syndrome.
1029 Table 1: KMT2D-bound genes that are differentially expressed in *Kmt2d*$^{\Delta/\Delta}$ neuronal cells.
1030
1031
1032 **Supplemental Figures**
1033 Supplementary Figure 1: CRISPR-targeted HT22 cells
1034 Supplementary Figure 2: HT22 cell RNA-seq and ChIP-seq analysis
1035 Supplementary Figure 3: iPSC and NPC validations and phenotyping
1036 Supplementary Figure 4: iPSC and NPC single-cell RNA-seq analysis
1037 Supplementary Figure 5. Stratified scRNA-seq analysis of NPCs
1038 Supplementary Figure 6: Adult neurogenesis phenotypes in *Kmt2d*$^{+/\beta geo}$ mice
1039 Supplementary Figure 7: Pulse-labeling to birth-date adult-born NPCs *in vivo*
1040
1041
1042 **Supplemental Tables**
1043 Supplementary Table 1: Differentially expressed genes in HT22 cells (*Kmt2d*$^{+/+}$ versus *Kmt2d*$^{\Delta/\Delta}$)
1044 Supplementary Table 2: KMT2D ChIP-seq peaks in HT22 cells (*Kmt2d*$^{+/+}$ cells)
1045 Supplementary Table 3: KMT2D-bound genes in HT22 cells (*Kmt2d*$^{+/+}$ cells)
1046 Supplementary Table 4: Differentially expressed genes in iPSCs (KS1 versus controls)
1047 Supplementary Table 5: Differentially expressed genes in NPCs (KS1 versus controls)
1048 Supplementary Table 6: Intersected gene sets of iPSCs and NPCs
1049 Supplementary Table 7: Differentially expressed genes in NPCs (stratified by subsets)
1050 Supplementary Table 8: Differentially expressed genes in EdU^{+} DG nuclei of *Kmt2d*$^{+/\beta geo}$ mice
1051 Supplementary Table 9: MRI volumetric comparisons in wild-type and *Kmt2d*$^{+/\beta geo}$ mice
1052 Supplementary Table 10: Reagents

1053

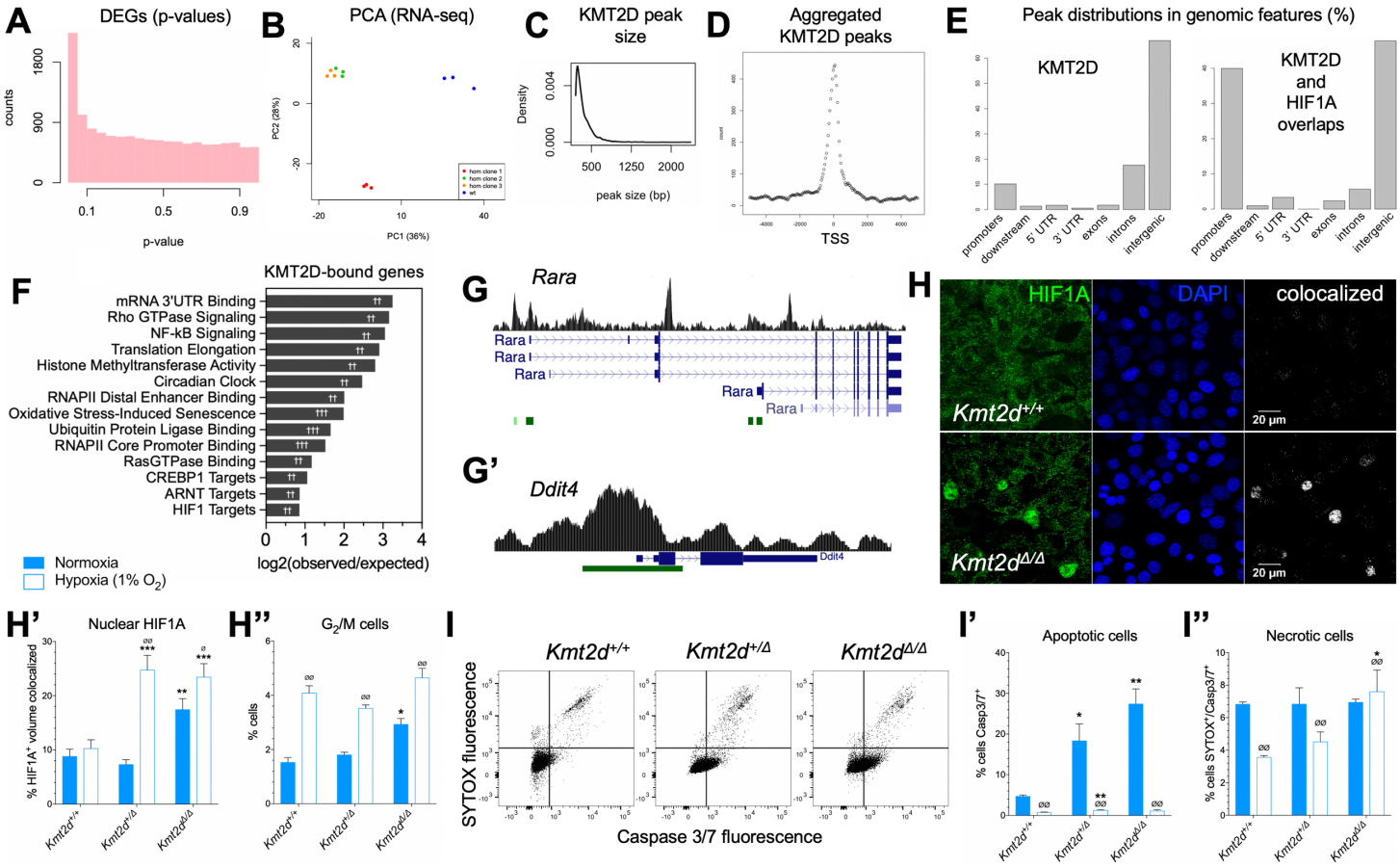
Fold change	Number of genes	Genes
0 to 1.1	11	***Dazap1***, *Psmc2, Eif3b,* ***Washc4****, Ran,* ***Lrpprc****, Imp4*, Sf3a2*, Bend3,* ***Hspd1*****, Dusp4*
-1 to 0	37	*Pik3ip1*, Ssc5d, Plekha1, Tinagl1*, Slc29a1, Rras, Nudt18, Slc25a29, F3, Pgpep1, Lrrn4cl, Gpr146,* ***Aldoa*****, Dbp,* ***Cpq****, 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*

1054

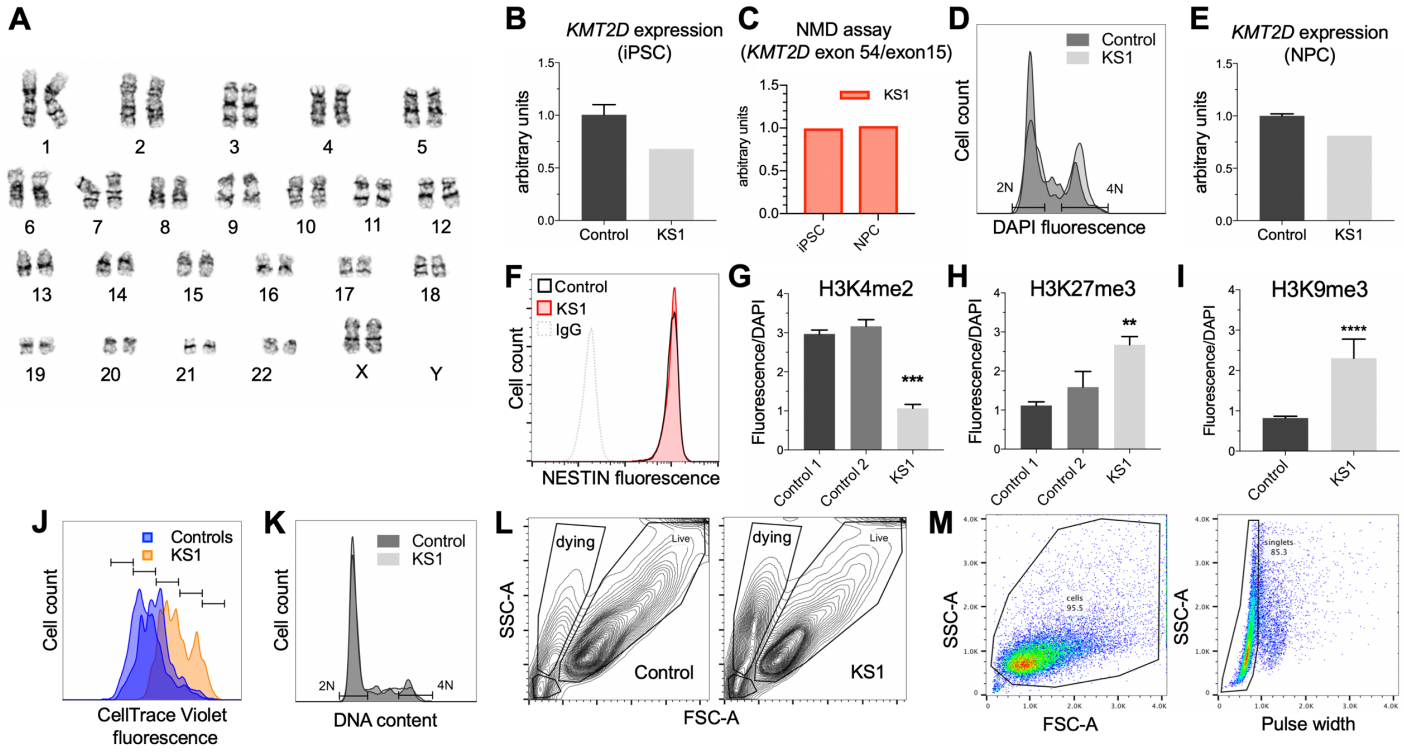
1055 **Table 1**: **KMT2D-bound genes that are differentially expressed in *Kmt2d*$^{\Delta/\Delta}$ neuronal cells.**
1056 Positive fold changes (0 to 1.1) indicate genes up-regulated and KMT2D-bound. Negative fold
1057 changes (-5 to 0) indicate genes down-regulated and KMT2D-bound. *previously described KMT2D
1058 target genes (7, 21), **bold** genes having experimentally validated HIF1A binding in the promoter (24).
1059



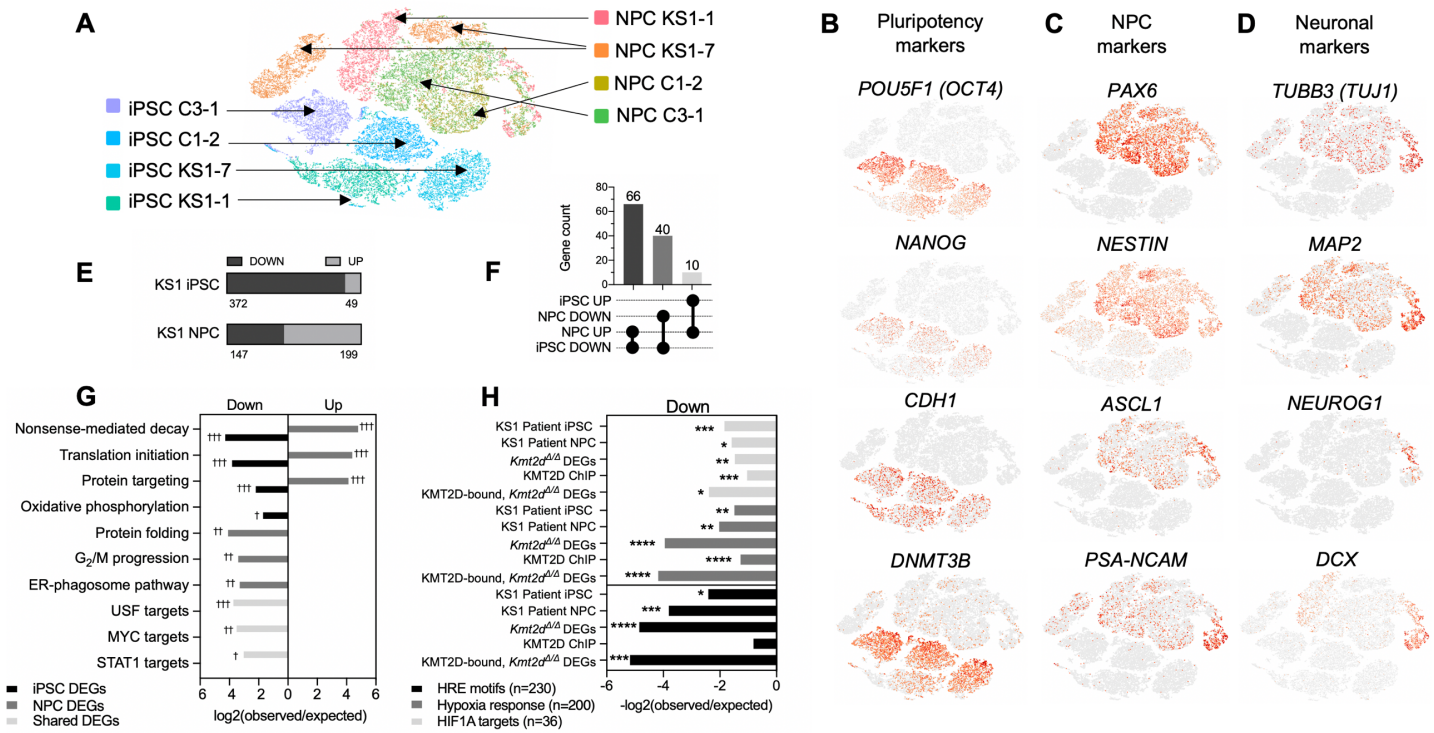
Supplementary Figure 1. CRISPR-targeted HT22 cells. (**A**) Sanger-sequenced DNA of wild-type (*Kmt2d*$^{+}$) and targeted (*Kmt2d*$^{\Delta}$) alleles in HT22 cells, mapped with sgRNAs and PCR primers, to *Kmt2d* locus (mm10) on chromosome 15. (**A'**) Mapping of Sanger-sequenced DNA after *in silico* translation to predict amino acid sequences illustrates premature termination codons (PTC) created in *Kmt2d*$^{\Delta 1}$ and *Kmt2d*$^{\Delta 2}$ alleles. (**B**) PCR with probes flanking sgRNA cut sites identifies experimental cell lines (*Kmt2d* $^{+/\Delta}$ and *Kmt2d*$^{\Delta/\Delta}$) compared to wild-type (*Kmt2d*$^{+/+}$). (**C**) RT-qPCR analysis of mRNA using probes spanning upstream exons (15-16) or exons within the deletion site (53-54). (**D**) CellTrace fluorescence intensities after 72 hours of proliferation analysis in HT22 cells. Increased intensity indicates less dye dilution, i.e. fewer cell divisions. (**E**) Cell cycle gating by 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, and (**E'**) quantification of each cycle phase.



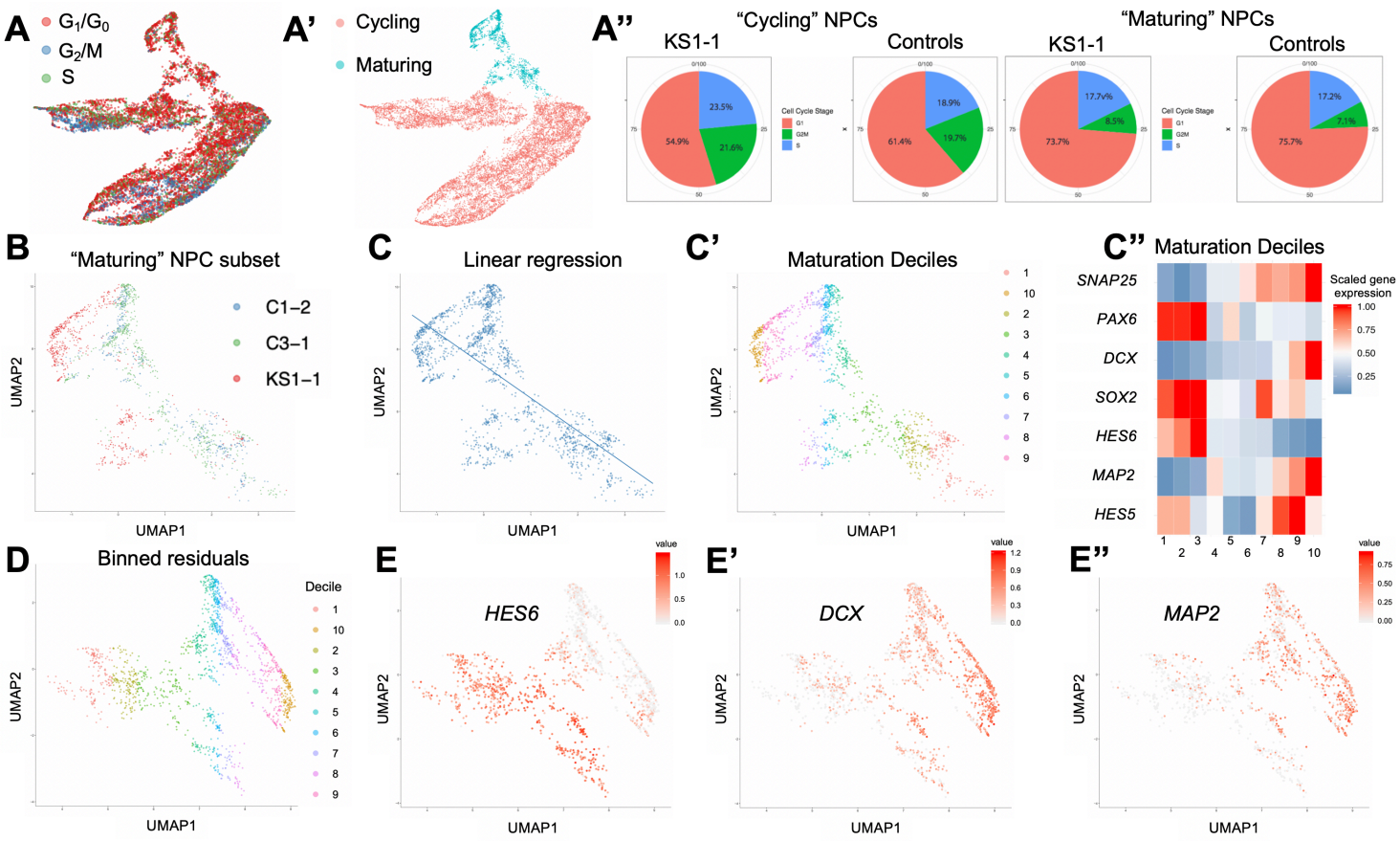
Supplementary Figure 2. HT22 cell RNA-seq and ChIP-seq analysis. (**A**) P-value distribution in *Kmt2d*$^{\Delta/\Delta}$ DEGs relative to wild-type indicate a well-calibrated test. (**B**) PCA visualizing clear expression differences in wild-type and *Kmt2d*$^{\Delta/\Delta}$ HT22 cells, illustrating separation of Clone 1 which bears strand invasion. (**C**) KMT2D ChIP-seq analysis comparing KMT2D peak density and peak size (bp) in wild-type HT22 cells. (**D**) Validation of expected KMT2D peak distributions about gene TSSs and (**E**) genomic features. (**F**) Gene networks showing highest fold change in enrichment among genes proximal to KMT2D peaks (<5 kb to promoter). (**G**) KMT2D peaks clustered at alternate TSSs of *Rara* gene and (**G'**) enhancer-like peaks at *Ddit4* gene. (**H**) HIF1A nuclear colocalization analysis (200-350 cells per condition). Representative Z-stacked confocal images of *Kmt2d*$^{+/+}$ and *Kmt2d*$^{\Delta/\Delta}$ cells under normoxic condition are shown with quantifications of colocalized HIF1A^{+} and DAPI^{+} voxels (**H'**), and %G_2/M phase cells measured by flow cytometry (12 wells per cell line) (**H''**) in all conditions. (**I**) Flow cytometric quantifications of cell death, with CellEvent fluorescence detecting caspase-3/7-mediated peptide cleavage in apoptotic cells (**I'**), and SYTOX reagent for necrotic cells (**I''**) (3-6 wells per cell line). Bars indicate mean ± SEM, Student's t-test (*p<0.05, **p<0.01, ***p<0.001). Fisher's Exact Test († FDR<0.05, †† FDR<0.01, ††† FDR<0.001).



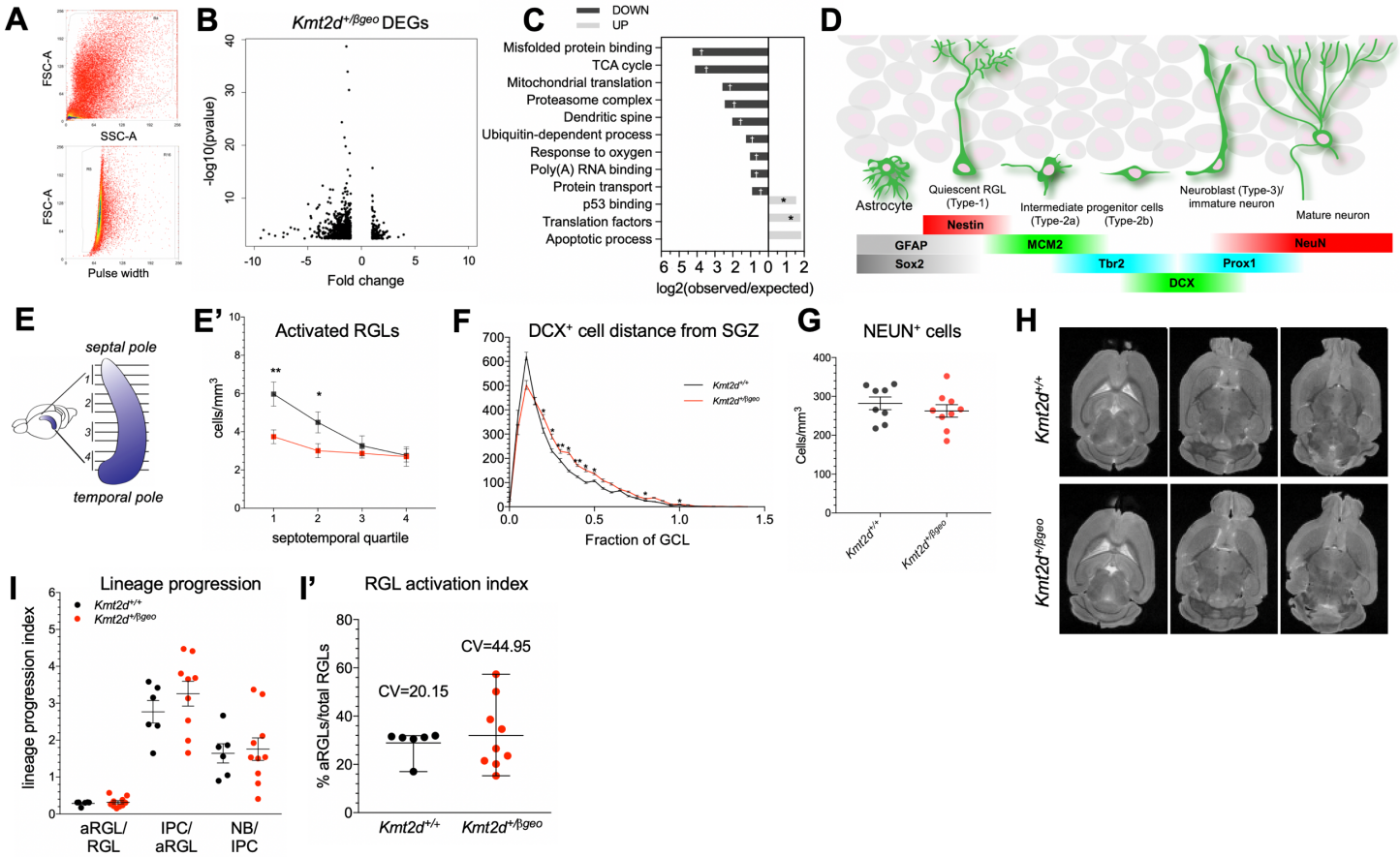
Supplementary Figure 3. iPSC and NPC line validations and additional phenotyping. (**A**) 46, XX normal female karyotype in KS1-1 iPSCs. (**B**) RT-qPCR analysis of *KMT2D* (exon 15) expression in KS1 iPSCs compared to two healthy control iPSC lines (C1-2 and C3-1), measured in technical triplicate. (**C**) RT-qPCR demonstrating equivalent exonic ratios of *KMT2D* exon 15 to exon 54, measured in technical triplicate, consistent with NMD of the entire transcript. (**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, measured in technical triplicate. (**F**) Flow cytometric analysis of NESTIN fluorescence intensity in KS1 and control NPCs. (**G-I**) Immunofluorescence imaging analysis of open chromatin mark H3K4me2 (**G**) and closed chromatin marks H3K27me3 (**H**) and H3K9me3 (**I**) in KS1 and control NPCs (9 images per line). (**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. Bars indicate mean ± SEM, Student's t-test (*p<0.05, **p<0.01, ***p<0.001).



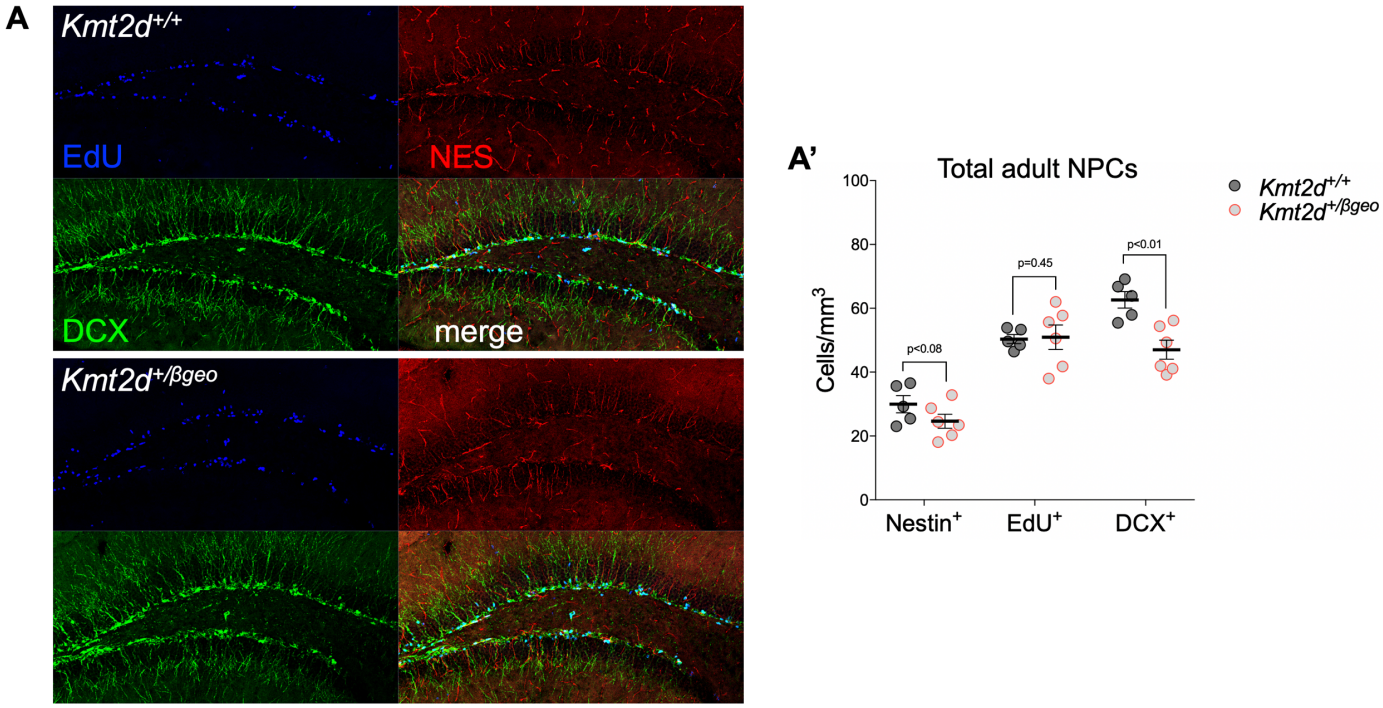
Supplementary Figure 4. iPSC and NPC single-cell RNA-seq analysis. (**A**) t-stochastic neighbor embedding (tSNE) representation of iPSC and NPC libraries sequenced on 10XGenomics platform. Cell clusters colored by cell type and Patient ID. iPSCs and NPCs derived from patient K1-7 were excluded from downstream analysis due to abnormal karyotype. (**B-D**) Representative tSNE of iPSC, NPC, and neuronal markers demonstrating expected cell identities and revealing a gradient of cell maturation. (**E**) Proportions of DEGs down- or up-regulated in KS1 patient iPSCs or NPCs compared to respective healthy controls, (**F**) DEG lists intersected for overlaps among down-regulated and up-regulated genes, and (**G**) 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. (**H**) 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 KMT2D ChIP-seq genes in wild-type 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, † FDR<0.05, †† FDR<0.01, ††† FDR<0.001).



Supplementary Figure 5. Stratified scRNA-seq analysis of NPCs. (**A-A''**) Uniform Manifold Approximation Projection (UMAP) of single-cell NPC libraries partitioned by (**A**) cell cycle marker expression into subsets of G_1/G_0, S, and G_2/M cells, used for cycle phase-stratified differential expression analysis (**Supplementary Table 7**) to rule out confounding differences in cell cycle phase composition on NPC transcriptome comparisons. (**A'**) Subset of "Cycling" versus non-cycling, "Maturing" NPCs, which includes "Transitioning" and "Differentiating" cells as defined (**Figure 4B**), and (**A''**) UMAP-based cell cycle occupancies consistent with experimental FACS data (**Figure 3F**). (**B-C'**) UMAPs of the segregated population of Differentiating NPCs displaying (**B**) library patient ID's, (**C**) smooth linear regression fitted to define the maturation trajectory and (**C'**) binned deciles of progressively maturing cells along the regression. (**C''**) Relative expression of selected NPC markers defining directionality of the maturation trajectory. (**D**) Binned residuals used to calculate deciles containing equal number of cells along the axis of differentiation. (**E-E''**) Representative NPC marker expression plotted over binned residuals.



Supplementary Figure 6. Phenotyping of *Kmt2d*$^{+/\beta geo}$ mice. (**A**) 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. (**B**) RNA-seq analysis of differential gene expression in purified EdU^{+} DG nuclei from *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice. (**C**) Gene networks most enriched among DEGs down- or up-regulated in *Kmt2d*$^{+/\beta geo}$ nuclei, showing transcriptional suppression of cellular metabolic pathways. (**D**) Schematic depicting marker expression during sequential stages of adult neurogenesis in the dentate gyrus (DG). (**E**) Serial ordering of perfusion-fixed brain slices enables anatomically-sequenced analysis of neurogenesis, for (**E'**) 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. (**F**) Quantification of DCX^{+} cell body distance from SGZ plane in 8-week-old mice (9-10 mice per genotype, at least 100 cells per mouse). (**G**) Quantification of NEUN^{+} cells in the granule cell layer (GCL) of 8-week-old mice (8-9 mice per genotype). (**H**) Sample images of T2-weighted MRI (9.4T) in PFA-fixed brains of female mice 4 months old. (**I-I'**) Comparison of lineage progression index, an approximation of expansion potential for each cell type transition, indicates absence of genotype-associated blockages at any particular cell-type transition analyzed, and (**I'**) increased Coefficient of Variance (CV) in RGL activation rates in *Kmt2d*$^{+/\beta geo}$ mice. Bars indicate mean ± SEM, Student's t-test (*p<0.05, **p<0.01, ***p<0.001). Fisher's Exact Test († FDR<0.05, †† FDR<0.01, ††† FDR<0.001).



Supplementary Figure 7. Pulse-labeling to birth-date adult-born NPCs *in vivo*. (**A**) Representative immunostaining from *Kmt2d*$^{+/+}$ and *Kmt2d*$^{+/\beta geo}$ mice (5-6 mice per genotype, 10 slices per mouse) of EdU pulse-labeled cells extending a NES^{+} process (NSCs) or DCX^{+} process (NPCs), showing the entire DG area used for quantification. (**A'**) Steady-state quantification of NSCs, NPCs, and EdU-labeled cells for immunostaining quantification in pulse-labeling differentiation experiment, confirming steady-state reduction of adult neurogenesis in *Kmt2d*$^{+/\beta geo}$ mice, despite their increase of pulsed immature neurons (**Figure 5E-E'**). Bars indicate mean ± SEM, Student's t-test (*p<0.05, **p<0.01, ***p<0.001).