



Published in final edited form as:

Cell. 2018 November 15; 175(5): 1228–1243.e20. doi:10.1016/j.cell.2018.10.006.

N⁶-methyladenine DNA Modification in Glioblastoma

Qi Xie^{1,*}, Tao P. Wu^{2,3,*}, Ryan C. Gimple^{1,4,*}, Zheng Li², Briana C. Prager^{1,4,5}, Qiulian Wu¹, Yang Yu⁶, Pengcheng Wang⁶, Yinsheng Wang⁶, David U. Gorkin⁷, Cheng Zhang⁷, Alexis V. Dowiak⁷, Kaixuan Lin², Chun Zeng⁸, Yinghui Sui⁸, Leo J. Y. Kim^{1,4}, Tyler E. Miller⁴, Li Jiang¹, Christine Lee-Poturalski⁹, Zhi Huang¹⁰, Xiaoguang Fang¹⁰, Kui Zhai¹⁰, Stephen C. Mack¹¹, Maike Sander⁸, Shideng Bao¹⁰, Amber E. Kerstetter-Fogle^{12,13}, Andrew E. Sloan^{12,13}, Andrew Z. Xiao^{2,#}, Jeremy N. Rich^{1,14,15,#}

¹Department of Medicine, Division of Regenerative Medicine, University of California, San Diego, La Jolla CA 92037, USA

²Department of Genetics and Yale Stem Cell Center, Yale School of Medicine, New Haven CT 06520, USA

³Present address: Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston TX 77030, USA

⁴Department of Pathology, Case Western Reserve University, Cleveland, OH 44120, USA

⁵Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland OH 44195, USA

⁶Department of Chemistry, University of California, Riverside CA 92521, USA

⁷Center for Epigenomics, Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla CA 92037, USA

⁸Departments of Pediatrics and Cellular and Molecular Medicine, Pediatric Diabetes Research Center, University of California, San Diego, La Jolla CA 92093, USA

⁹Department of Pharmacology, Case Western Reserve University, Cleveland OH 44195, USA

¹⁰Department of Stem Cell Biology and Regenerative Medicine, Lerner Research Institute, Cleveland Clinic, Cleveland OH 44195, USA

¹¹Department of Pediatrics, Division of Hematology and Oncology, Baylor College of Medicine, Texas Children's Hospital, Houston TX, 77030, USA

#.Correspondence to: Jeremy N. Rich (drjeremyrich@gmail.com) and Andrew Z. Xiao (andrew.xiao@yale.edu).

*These authors contributed equally to this work.

AUTHOR CONTRIBUTIONS

Q.X., T.P.W., J.N.R., A.Z.X. designed experiments. Q.X., T.P.W., R.C.G., J.N.R., A.Z.X. analyzed data and prepared the manuscript. Q.X., T.P.W., R.C.G., Z.L., B.C.P., Q.W., L.J.Y.K., K.L. performed experiments. Y.Y., P.W., Y.W. performed mass spectrometry experiments and analysis. D.U.J., C.Z., A.V.D. performed ATAC-seq experiments and analysis. T.E.M., Z.H., S.B. provided scientific input and edited the manuscript.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DECLARATION OF INTERESTS

The authors declare no financial or other competing interests.

¹²Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland OH, 44106, USA

¹³Department of Neurological Surgery, University Hospitals-Cleveland Medical Center, Cleveland OH, 44106, USA

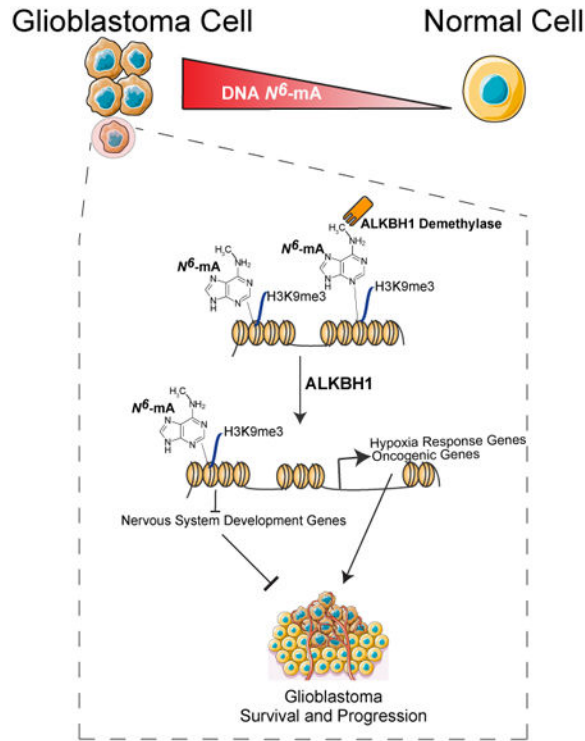
¹⁴Department of Neurosciences, University of California, San Diego, School of Medicine, La Jolla CA 92037, USA

¹⁵Lead Contact

SUMMARY

Genetic drivers of cancer can be dysregulated through epigenetic modifications of DNA. While the critical role of DNA 5-methylcytosine (5mC) in the regulation of transcription is recognized, the functions of other non-canonical DNA modifications remain obscure. Here, we report the identification of novel N(6)-methyladenine (N^6 -mA) DNA modifications in human tissues and implicate this epigenetic mark in human disease, specifically the highly malignant brain cancer, glioblastoma. Glioblastoma markedly upregulated N^6 -mA levels, which co-localized with heterochromatic histone modifications, predominantly H3K9me3. N^6 -mA levels were dynamically regulated by the DNA demethylase, ALKBH1, depletion of which led to transcriptional silencing of oncogenic pathways through decreasing chromatin accessibility. Targeting the N^6 -mA regulator, ALKBH1, in patient-derived human glioblastoma models inhibited tumor cell proliferation and extended survival of tumor-bearing mice, supporting this novel DNA modification as a potential therapeutic target for glioblastoma. Collectively, our results uncover a novel epigenetic node in cancer through the DNA modification, N^6 -mA.

Graphical Abstract



In brief

N6-methyladenine DNA modifications are enriched in human glioblastoma and decreasing this modification can inhibit cancer growth by limiting chromatin accessibility at oncogenic loci

INTRODUCTION

Aberrant epigenetic landscapes promote tumor initiation and progression (Elsasser et al., 2011; Loven et al., 2013; Northcott et al., 2014). To date, the focus of cancer research has been on the global and local aberrations of 5-methylcytosine (5mC). Genome-wide 5mC hypomethylation frequently occurs in cancer genomes, leading to widespread genomic instability and de-repression of repetitive elements (Cadieux et al., 2006; Feinberg and Vogelstein, 1983; Widschwendter et al., 2004). However, local hypermethylation of 5mC events are often found in close proximity to important tumor suppressor genes in a number of cancers (Esteller, 2002). Low grade gliomas and secondary glioblastomas commonly harbor mutations in isocitrate dehydrogenase 1 or 2 (IDH1/2), leading to DNA hypermethylation and manifesting as the glioma-CpG island methylator phenotype (G-CIMP) (Noushmehr et al., 2010; Turcan et al., 2012). Thus, DNA methylation is a critical epigenetic mark that regulates many important developmental processes and plays a fundamental role in cancer.

While numerous epigenetic modifications have been described in addition to DNA methylation, it is still unknown whether tumor cells can adopt novel epigenetic mechanisms that are rarely utilized in normal adult tissues. In addition to the canonical 5mC, other DNA methylation events, including N(6)-methyladenine (N⁶-mA), have been identified in bacteria

(Vanyushin et al., 1968) and a limited number of eukaryotes, such as mosquitoes, plants, *C. elegans* and *Drosophila* (Fu et al., 2015; Greer et al., 2015; Rogers and Rogers, 1995; Yao et al., 2017; Zhang et al., 2015). We recently demonstrated that N^6 -mA DNA methylation occurs in mammals, where it acts mainly as a repressive mark to silence transcription of LINE transposons in mouse embryonic stem cells (ESCs) (Wu et al., 2016) and this modification has also been reported in human tissue (Xiao et al., 2018; Zhu et al., 2018). However, the roles of N^6 -mA in human cancer and in genome biology are largely unknown.

To interrogate N^6 -mA function, we utilized glioblastoma as a model human cancer. Glioblastoma is the most prevalent and aggressive primary intrinsic brain tumor and is characterized by widespread epigenetic dysregulation, including DNA hypermethylation of 5mC and alterations in chromatin remodelling enzymes, such as the polycomb complex proteins, EZH2 and BMI1 (Abdouh et al., 2009; Jin et al., 2017; Suva et al., 2009). Thus, we studied the N^6 -mA DNA modification in human patient-derived glioblastoma tumor models and functionally characterized its genomic localization and role in the regulation of glioblastoma.

RESULTS

Identification of N^6 -mA DNA modifications in human glioblastoma

To investigate the roles of the N^6 -mA DNA modification in human glioblastoma, we extracted genomic DNA from functionally validated human patient-derived glioblastoma stem cell (GSC) models (387, D456, GSC23, and 1919) and primary human tumor specimens (3028 and CW2386), then performed dot blot analysis using an N^6 -mA specific antibody that has been previously validated (Greer et al., 2015; Wu et al., 2016; Zhang et al., 2015). GSCs and primary human tumor specimens displayed strikingly elevated N^6 -mA levels relative to normal human astrocytes (Figure 1A). N^6 -mA levels were independently quantified by mass spectrometry (MS) using an established and highly sensitive (LOQ: 1.6 fmol) mass spectrometry (UHPLC-MS/MS) approach with stable isotope-labeled [$^{15}\text{N}_5$] N^6 -mA as an internal standard for sample enrichment and quantification (Figures S1A and S1B), which demonstrated that N^6 -mA levels in GSCs were elevated more than hundredfold compared to normal human astrocytes (Figure 1B). N^6 -mA DNA levels were also measured by immunofluorescence staining. To eliminate signal from RNA, samples were treated with RNaseA prior to incubation with primary antibodies. N^6 -mA DNA levels detected by immunofluorescence validated dot blot and mass spectrometry results, with much higher N^6 -mA levels in the patient-derived GSCs than in normal human astrocytes (Figures 1C and 1D). Effects of cell culture conditions were ruled out using DNA N^6 -mA immunohistochemistry on a tissue microarray (TMA) with normal brain and glioblastoma tissues, confirming increased DNA N^6 -mA levels in glioblastoma (Figures 1E and 1F). Potential inter-individual heterogeneity in N^6 -mA levels was addressed using immunofluorescence staining of matched glioblastoma and surrounding normal brain tissue from the same patient following surgical resection. Consistently, glioblastoma tumor tissue contained increased levels of N^6 -mA (Figure S1C). N^6 -mA levels were elevated in other central nervous system cancers, including diffuse intrinsic pontine glioma (DIPG), meningioma, and medulloblastoma, compared to normal human astrocytes (Figure S1D).

Of note, a recent study reported high levels of N^6 -mA in normal human adult tissues (at several hundred PPM), with downregulation in gastric and stomach cancers (Xiao et al., 2018). However, our current study, together with other previous studies (Zhu et al., 2018) did not find high levels of N^6 -mA in normal adult tissues or mammalian cells. Collectively, our results support elevation of a novel DNA modification, N^6 -mA, in glioblastoma, with likely involvement in other cancers.

Genomic profiling and pattern of N^6 -mA in glioblastoma

N^6 -mA genomic localization was analyzed using N^6 -mA DIP-seq (DNA immunoprecipitation with anti- N^6 -mA antibodies followed by next-generation sequencing) in two human patient-derived GSC models (387 and GSC23) and one primary human tumor specimen (CW2386) to identify N^6 -mA-enriched genomic regions. The number of N^6 -mA peaks measured by DIP-sequencing in glioblastoma ranged from 7,282 to 17,263 per model (Figure 2A). N^6 -mA peaks were most common in intergenic regions, consistent with our previous findings in murine ESCs (Wu et al., 2016) (Figure 2B). N^6 -mA peaks were present on each chromosome with mild enrichment on chromosomes 7, 19, and 21 (Figure 2C). Gene Ontology (GO) enrichment analysis of the closest gene to each of the shared N^6 -mA peaks revealed an enrichment in neurogenesis and neuronal development pathways (Figures 2D and 2E). GO analysis restricted to peaks within gene bodies revealed similar results (Figures S2A and S2B). Although adenine nucleotides are distributed throughout the genome, selective distribution of N^6 -mA could reflect sequence-specific recognition of molecular regulators. To investigate this, we identified DNA motifs enriched within N^6 -mA regions in the common shared peaks. One of the top ranking motifs (GGAAT) closely resembles human major satellite repeats that are enriched at constitutive heterochromatic regions (Figure S2C). Consistent with findings that DNA N^6 -mA functions as a repressive mark in murine ESCs and brains (Wu et al., 2016; Yao et al., 2017), these data suggest that N^6 -mA-mediated repression of neurodevelopmental pathways may serve a role in heterochromatin formation that contributes to glioblastoma tumorigenesis.

N^6 -mA is enriched in heterochromatin

To explore the molecular mechanism of N^6 -mA as a repressive epigenetic modification, we performed chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) for two heterochromatin marks [histone 3 lysine 9 trimethyl (H3K9me3) and histone 3 lysine 27 trimethyl (H3K27me3)] and a euchromatin mark [histone 3 lysine 4 trimethyl (H3K4me3)] in a patient-derived GSC model. Over 80% of N^6 -mA peaks intersected with regions bearing heterochromatic histone modifications (H3K9me3 and H3K27me3), while few peaks (3.5%) overlapped with euchromatic histone modifications (H3K4me3) (Figures 3A-C). Most commonly, N^6 -mA marks occurred in regions bound by both heterochromatin marks. N^6 -mA marks in regions bound by only a single heterochromatin mark were increased near peaks bound by H3K9me3 alone (26.4%), while peaks bound by H3K27me3 alone (16.2%) were less frequently enriched (Figures 3A and 3B). Aggregation analysis showed that N^6 -mA peaks common to all three glioma models (Figure 2A) were directly superimposed with those of the heterochromatic histone marks (Figure 3C). All N^6 -mA peaks in the 387 patient-derived GSC model, some of which were unique in this model, were strongly correlated with these heterochromatic marks (Figure 3D). Genomic overlap

between the N^6 -mA DNA modification and the canonical 5mC DNA methylation mark was interrogated through whole genome bisulfite sequencing analysis of the 387 patient-derived GSCs and mapping of the intersection with N^6 -mA. As expected, 5mC marks were depleted in promoter regions, but slightly increased in N^6 -mA peaks compared to randomly selected genomic regions (Figures S2D-G). Thus, our data suggest that N^6 -mA strongly co-localizes with heterochromatic histone modifications and with the largely repressive canonical DNA methylation mark, 5mC.

As the N^6 -mA DNA modification acts as a repressive mark that is highly associated with the H3K9me3 heterochromatin histone modification, the high levels of N^6 -mA may function to support GSC survival by repressing key tumor suppressor genes. To investigate this possibility, the Tumor Suppressor Gene (TSG) Database (Zhao et al., 2016) was leveraged to identify potential tumor suppressor genes that contain both N^6 -mA and H3K9me3 enrichment. Several tumor suppressor genes, including CDKN3, RASSF2 and AKAP6, appear to be repressed by N^6 -mA and H3K9me3 (Figure S3).

ALKBH1 is a DNA N^6 -mA demethylase that actively modulates gene expression in human glioblastoma

DNA methylation is dynamically regulated by a number of enzymes. We next investigated the enzymes controlling DNA N^6 -mA modification in glioblastoma. Our previous study in murine ESCs identified Alkylated DNA Repair Protein AlkB Homolog (ALKBH1) as a demethylase for DNA N^6 -mA (Wu et al., 2016), so we focused on its human homolog as a potential regulatory node. *In vitro* DNA demethylation assays demonstrated strong N^6 -mA demethylase catalytic activity of the human ALKBH1 protein toward N^6 -mA in oligonucleotides using dot blotting (Figure 4A) or UHPLC-MS (Figure 4B). To investigate its functions in glioblastoma, ALKBH1 expression was targeted by two independent, non-overlapping small hairpin RNA (shRNA) lentiviral constructs (designated shALKBH1.1008 and shALKBH1.1551) compared to a non-targeting control sequence shRNA insert (shCONT), which increased DNA N^6 -mA levels in two patient-derived GSCs (387 and D456; Figure 4C). As a recent study indicated that ALKBH1 might have activities against tRNA (Liu et al., 2016), we further investigated and found that ALKBH1 does not function as a tRNA-1mA demethylase in patient-derived GSCs (data not shown).

We next aimed to identify N^6 -mA methyltransferases in GSCs. N-6 adenine-specific DNA methyltransferase 1 (N6AMT1) was recently suggested as a DNA N^6 -mA methyltransferase in human cells (Xiao et al., 2018). In light of these findings, we interrogated the function of N6AMT1 in patient-derived GSCs. In contrast to this previous report, genomic knockout of N6AMT1 in GSCs did not alter the levels of DNA N^6 -mA, assessed by dot blot (Figures S4A and S4B). Further, purified N6AMT1 recombinant protein incubated with DNA oligonucleotides *in vitro* did not alter DNA N^6 -mA levels, suggesting that N6AMT1 does not serve as an N^6 -mA methyltransferase, at least in glioblastoma (Figures S4C and S4D). Our findings are consistent with a previous study that also did not support N6AMT1 as an N^6 -mA methyltransferase (Liu et al., 2010). Thus, we focused our attention on ALKBH1, the only definitive modulator of N^6 -mA in GSCs, to understand the role of N^6 -mA regulation in glioblastoma.

Our data suggest ALKBH1 may control specific genetic pathways in glioblastomas. To investigate this, we performed RNA sequencing (RNA-seq) to interrogate the effects of ALKBH1-knockdown on the whole transcriptome. 321 genes were downregulated by twofold or greater, while only 37 genes were upregulated by more than twofold upon ALKBH1 depletion (Figure 4D), suggesting that an increase of N^6 -mA DNA levels primarily represses gene expression in human glioblastoma. To further elucidate the genomic sites of ALKBH1 demethylase activity, N^6 -mA DIP-seq was performed on a patient derived GSCs (387) after transduction with either one of two independent ALKBH1 shRNAs or shCONT. 68% of downregulated genes after ALKBH1 knockdown were associated with genomic regions with increased N^6 -mA levels (Figure 4E), strongly suggesting that the transcriptional downregulation is a direct consequence of elevated N^6 -mA levels due to ALKBH1 depletion. To understand the role of N^6 -mA in regulating chromatin accessibility, ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) was performed following knockdown of ALKBH1. Consistent with our finding of co-localization of N^6 -mA with heterochromatic regions, ALKBH1 knockdown led to a greater number of genomic loci with decreased chromatin accessibility (1,389 sites) than sites with increased accessibility (243 sites) (Figures 4F-I). ATAC-seq peaks, representing sites of open chromatin, negatively correlated with N^6 -mA peaks, indicating that N^6 -mA is depleted at sites of high chromatin accessibility (Figures 4J). Overlaid ATAC-seq and RNA-seq data showed that more genes bearing a significantly decreased accessibility site within 2 Kb of their transcriptional start sites tended to transcriptionally repressed (37 genes) rather than upregulated (19 genes) (Figure 4K). To further investigate the interaction between ALKBH1 and N^6 -mA modifications, biotin-labelled oligonucleotides with or without N^6 -mA modifications were used to perform ALKBH1 pulldowns using whole cell lysates from patient-derived GSCs. N^6 -mA modified oligonucleotides preferentially bound ALKBH1 relative to unlabelled oligonucleotides, suggesting that ALKBH1 binds regions bearing N^6 -mA modifications (Figure S5A). To assess the preference of ALKBH1 to bind N^6 -mA genome-wide, ALKBH1 ChIP-seq identified ALKBH1 binding sites were mapped to N^6 -mA peaks. ALKBH1 ChIP-seq peaks were preferentially enriched within N^6 -mA marked regions compared to the genome background (Figures S5B and S5C). Thus, ALKBH1 functions as a transcriptional activator by binding to regions of N^6 -mA enrichment and removing repressive N^6 -mA marks at selected genomic loci.

ALKBH1 depletion facilitates heterochromatin formation in human glioblastoma through N^6 -mA DNA modification

We next analyzed co-localization of N^6 -mA peaks and genomic sites directly regulated by ALKBH1 (i.e. N^6 -mA peaks gained after ALKBH1 knockdown). Consistent with the strong association of N^6 -mA and heterochromatin marks at a genome-wide level, N^6 -mA sites directly increased upon ALKBH1 knockdown were even more strongly correlated with heterochromatin histone modifications, particularly H3K9me3 (Figures 5A and 5B). Genes marked by N^6 -mA peaks overlapped with those marked by H3K9me3 (Figures 5C, 5D, and S5D). For example, two genes downregulated upon ALKBH1 depletion, RFTN2 and LOX, were marked by increased N^6 -mA levels that co-localized with H3K9me3 demarcated heterochromatin regions (Figures S5G and S5H).

To assess the effects of ALKBH1-mediated alterations of N^6 -mA on the global H3K9me3-marked heterochromatin landscape, we performed H3K9me3 ChIP-seq following ALKBH1 knockdown in a patient derived GSCs, revealing an increased number of H3K9me3 peaks following ALKBH1 knockdown (9,999 peaks with increased H3K9me3 signal vs. 2,995 peaks with decreased H3K9me3 signal) (Figures 5E and 5F), which was confirmed in two additional patient-derived GSCs (Figures S5E and S5F). The genes *SDK1* and *BMT2* gained N^6 -mA peaks co-localizing with gained H3K9me3 signal following ALKBH1 knockdown (Figure 5G). Collectively, these results suggest that N^6 -mA plays an important role in facilitating heterochromatin formation in human glioblastoma, particularly acting through the H3K9me3 histone modification.

ALKBH1 regulates hypoxia-response genes

To interrogate the gene pathways regulated by ALKBH1-sensitive N^6 -mA sites, we performed GO analysis on genes downregulated upon ALKBH1 depletion, revealing enrichment of key oncogenic pathways previously implicated in glioblastoma pathogenesis, including hypoxia responses, nervous system development, cellular and tissue development, and plasma membrane structure pathways (Figures S6A and S6B). Of note, hypoxia response genes were modulated by ALKBH1, even when cells were grown under room air conditions. As hypoxia-related gene signatures were the top enrichment following ALKBH1 knockdown, the comparative transcriptional effects of ALKBH1 knockdown were assessed under hypoxic conditions (1% oxygen) by RNA-seq. Similarly to the normoxic conditions, genes in the hypoxia pathway were downregulated following ALKBH1 knockdown when cells were grown in hypoxia (Figures S6C and S6D). Upregulated genes upon ALKBH1 knockdown in hypoxia included DNA damage and p53 pathway genes (Figures S6E and S6F). To support the clinical significance of these findings, an ALKBH1-regulated gene signature highly correlated with a hypoxia signature in The Cancer Genome Atlas (TCGA) glioblastoma dataset (Figure S6G). To explore the correlation between hypoxia and ALKBH1-dependent N^6 -mA demethylation sites, RNA-seq of GSCs under hypoxia identified over 50% of genes upregulated in hypoxia were suppressed by ALKBH1 knockdown (Figure S6H). N^6 -mA peaks that decreased following hypoxia tended to be located closer to upregulated genes in hypoxia (e.g. *MIAT*) than randomly selected peaks (genome background) (Figures S6I-K). Thus, ALKBH1 regulates key processes previously implicated in GSC biology through alterations in the N^6 -mA landscape. H3K9me3 has been previously linked to repression of hypoxia-induced genes during development and tumorigenesis (Luo et al., 2012), including in glioblastoma (Intlekofer et al., 2015). Collectively, our results demonstrate an unexpected role of N^6 -mA and ALKBH1 in modulating hypoxia-induced genes in human glioblastoma.

N^6 -mA is a potential therapeutic target in glioblastoma

To address the functional significance of N^6 -mA in tumor biology, we investigated ALKBH1 as a potential therapeutic target. ALKBH1 protein levels were modestly, but consistently, elevated in GSCs compared to their matched differentiated non-stem tumor cells by immunoblotting (Figure S7A). Silencing ALKBH1 decreased GSC proliferation (Figure 6A). To control for the potential off-target effects of shRNAs in additional experiments, we depleted ALKBH1 using sgRNAs by CRISPR-Cas9 technology with two independent non-

overlapping single-guide RNAs (sgALKBH1#4 and sgALKBH1#5). Following ALKBH1 depletion with sgRNAs, glioma cells displayed reduced cell growth, consistent with the shRNA results (Figure 6B). Sphere formation is surrogate marker of self-renewal, albeit with caveats. *In vitro* limiting dilution sphere formation assays revealed that ALKBH1 knockdown resulted in a more than tenfold decrease in the frequency of sphere formation (Figure 6C). GSCs are functionally defined by their ability to initiate tumors *in vivo*, prompting us to evaluate the potential anti-tumor effects of ALKBH1 targeting *in vivo*. Two different patient-derived GSCs (387 and D456) transduced with either one of two non-overlapping ALKBH1-targeting shRNAs (shALKBH1) or control non-targeting shRNA (shCONT) were implanted into the brains of immunocompromised mice. Consistent with our *in vitro* results, animals bearing tumors derived from GSCs expressing shRNAs or sgRNAs targeting ALKBH1 displayed increased survival relative to those bearing GSCs expressing shCONT or sgCONT, respectively (Figures 6D and 6E). In light of the role of ALKBH1 in the regulation of hypoxia pathways, we determined the effects of ALKBH1 knockdown in hypoxic conditions. Depleting ALKBH1 mRNA levels by shRNAs in hypoxia revealed reduced cell viability *in vitro*, similarly to in normoxia (Figure S7B). ALKBH1 knockdown in hypoxia impaired sphere formation capacity in a limiting dilution assay, mirroring our findings in normoxia (Figure S7C). Over expression of ALKBH1 diminished N^6 -mA levels in GSCs (Figure S7D). In contrast to the knockdown studies, ALKBH1 over expression did not alter cell viability, self-renewal, or *in vivo* tumor formation (Figures S7E-G). Although N^6 -mA was reduced initially, its levels return to baseline after 3 days, despite continuous over expression of ALKBH1 (data not shown). This implies that compensatory mechanisms, which likely include currently unknown DNA demethylases and methyltransferases, exist to maintain the levels of N^6 -mA within a relatively narrow range that is compatible with cellular survival. In summary, ALKBH1 is essential for maintaining cell viability and stemness properties under both normoxic and hypoxic conditions.

To determine the clinical relevance of these findings, we performed *in silico* studies on the TCGA glioblastoma dataset. While correlation with patient outcome does not universally indicate the importance of any individual gene targets (Kaelin, 2017), ALKBH1 was highly expressed in glioblastomas relative to non-tumor brain tissue and associated with reduced survival and increased glioma grade (Figures 7A-C). An ALKBH1-regulated gene signature, defined by downregulated genes following ALKBH1 knockdown, correlated with tumor grade and shorter patient survival in several datasets (Figures 7D-G). Taken together, our findings demonstrate that ALKBH1 and regulation of N^6 -mA is necessary for glioma growth and may be a therapeutically targetable node.

DISCUSSION

N^6 -mA was originally described as a DNA modification that discriminates the original and newly synthesized DNA strand in bacteria (O'Brown and Greer, 2016). We and others reported that DNA undergoes rare N^6 -mA modification in low and high eukaryotes (Fu et al., 2015; Greer et al., 2015; Wu et al., 2016; Xiao et al., 2018; Yao et al., 2017; Zhang et al., 2015). However, the biological function of N^6 -mA in human cancers remains unclear. Here we elucidated the function of N^6 -mA in human glioblastoma. N^6 -mA levels were elevated in tumor relative to normal brain tissues and in GSCs compared to normal human astrocytes,

suggesting that GSCs co-opt the regulatory potential of N^6 -mA to promote a tumorigenic chromatin landscape. As in murine ESCs, N^6 -mA repressed gene expression. Our ChIP-seq results strongly associated N^6 -mA sites with a common heterochromatin marker H3K9me3 in glioblastoma, which has not been demonstrated previously. H3K9me3 is not involved in transcriptional silencing of N^6 -mA at young LINE-1s in ESCs. These results indicate that glioma cells have hijacked and further adapted N^6 -mA-mediated silencing mechanisms typically employed during early embryogenesis. Previous studies have implicated H3K9me3 in the regulation of DNA methylation and shown that complete removal of DNA methylation affects H3K9me3 levels and disrupts heterochromatin architecture (Du et al., 2015; Saksouk et al., 2014). Our findings corroborate the potential crosstalk between DNA N^6 -mA and H3K9me3 and possible collaboration of these modifications in transcription silencing in human cancers. Future experiments will inform the relationship between N^6 -mA, 5mC, and H3K9me3 to discern the temporal sequence of transcriptional silencing mediated by these marks. The precise roles of N^6 -mA and its interplay with the canonical 5mC DNA methylation mark as well as other histone modifications remain to be uncovered.

We demonstrated that ALKBH1 is a critical N^6 -mA demethylase that acts to regulate a number of gene networks critical to GSC identity, including the hypoxia response pathway. Most epigenetic regulation functions optimally in a specific range with either loss or gain of function leading to disruption of cellular biology. Concordantly, elevation of N^6 -mA levels by targeting its demethylase ALKBH1 inhibited tumor formation. Although it may appear paradoxical, both N^6 -mA levels as well as the N^6 -mA demethylase were upregulated in glioblastoma. This apparent discrepancy can be resolved by understanding that the regulation of N^6 -mA is a highly dynamic process that depends on both a demethylase and a (currently unknown) methyltransferase. We hypothesize that the flux of N^6 -mA is important, not just the absolute levels. We interrogated the role of N^6 -mA in glioblastoma through modulating the only known regulator of N^6 -mA, ALKBH1. Disruption of the dynamic regulation of N^6 -mA through knockdown of ALKBH1 led to decreased proliferation, self-renewal, and tumor formation capacity, suggesting that efficient regulation of N^6 -mA is a GSC dependency.

Although we identified the N^6 -mA demethylase, further studies are required to identify the methyltransferases active in glioblastoma to model N^6 -mA in cancer. We posit that glioblastoma cells have elevated genome-wide N^6 -mA levels, whereas at selected genomic loci, uncontrolled increases will be detrimental to the tumor cells. The active regulation of this modification at local sites by both a demethylase and methyltransferase likely play important roles in the tuning of biological processes. While global N^6 -mA enrichment is localized near genes involved in neurodevelopmental processes, those modifications that are specifically regulated by ALKBH1 tend to be localized near hypoxia response genes. It should be noted that methyl-6-adenosine (m6A) modifications of RNA have been recently reported as well, including in glioblastoma (Cui et al., 2017; Fu et al., 2014; Zhang et al., 2017). The biochemical relationship between N^6 -mA and m6A modifications on DNA and RNA appear to be distinct.

In conclusion, we demonstrated that the N^6 -mA DNA modification exists in human tissues, is dynamically regulated in GSCs by the DNA demethylase ALKBH1, and that maintenance

of this regulatory circuitry is critical to cell survival and proliferation. Our findings reveal a new class of DNA modification in human disease and identify a potential therapeutic vulnerability that can be exploited for cancer therapy. Therefore, our results offer a novel therapeutic and discovery paradigm for deadly cancers.

METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

The authors are willing to distribute all materials, datasets, and protocols used in the manuscript. Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jeremy N. Rich (drjeremyrich@gmail.com).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Glioblastoma stem cell derivation and xenograft maintenance—Glioblastoma tissues were obtained from excess surgical resection samples from patients at the Case Western Reserve University after review by neuropathology with appropriate consent and in accordance with an IRB-approved protocol (090401). All patient studies were conducted in accordance with the Declaration of Helsinki. Glioma stem cell models 387 and 3565 were derived by our laboratory and transferred via a material transfer agreement from Duke University. All GSC models were cultured in Neurobasal media (Invitrogen) supplemented with B27 without vitamin A (Invitrogen), EGF, and bFGF (20 ng/ml each; R&D Systems), sodium pyruvate (Life Technologies, Cat # 11360070), and glutamax (Life Technologies, Cat # 35050061). For normoxia experiments, cells were cultured in cell culture incubators at 37°C with 20% oxygen and 5% carbon dioxide. For hypoxia experiments, cells were cultured in cell culture incubators at 37°C with 1% oxygen and 5% carbon dioxide. To decrease the incidence of cell culture-based artifacts, patient-derived xenografts were produced and propagated as a renewable source of tumor cells for study. Xenografted tumors were dissociated using a papain dissociation system according to the manufacturer's instructions. STR analyses were performed on each tumor model used in this study for authentication. The 3565 GSC model was derived from a glioblastoma from a 32-year old male patient. The T387 GSC model was derived from a GBM from a 76-year old female patient. The D456 GSC model was derived from a glioblastoma biopsy from an 8-year old female patient and was provided as a generous gift from Darell Bigner (Duke University). The GSC23 model was derived from a recurrent glioblastoma biopsy from a 63-year old male patient and was provided as a generous gift by Erik Sulman (MD Anderson Cancer Center). The 1919 GSC model was derived from a glioblastoma from a 53-year old male patient. The 3028 GSC model was derived from a recurrent glioblastoma from a 65-year old female patient. The 2386 tissue was derived from a glioblastoma from a 71-year old female patient. The 2752 tissue was derived from a glioblastoma from a 53-year old male patient. The 2762 tissue was derived from a glioblastoma from a 74-year old male patient.

Glioblastoma Stem Cell Model or Tissue	Patient Age (Years)	Patient Sex
T387	76	Female
3565	32	Male
D456	8	Female
GSC23	63	Male
1919	53	Male
3028	65	Female
2386	71	Female
2752	53	Male
2762	74	Male

In Vivo Tumorigenesis—Intracranial xenografts were created by implanting 10,000 human-derived GSCs into the right cerebral cortex of NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ, The Jackson Laboratory, Bar Harbor, ME, USA) mice at a depth of 3.5 mm. All mouse experiments were performed under an animal protocol approved by the University of California, San Diego Institutional Animal Care and Use Committee. Healthy, wild-type female mice of NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ, The Jackson Laboratory, Bar Harbor, ME, USA) background, 4–6 weeks old, were randomly selected and used in this study for intracranial injection. Mice had not undergone prior treatment or procedures. Mice were maintained in 14 hours light/10 hours dark cycle by animal husbandry staff at the University of California, San Diego, with no more than 5 mice per cage. Housing conditions and animal status were supervised by a veterinarian. Animals were monitored until neurological signs were observed, at which point they were sacrificed. Neurological signs or signs of morbidity included hunched posture, gait changes, lethargy and weigh loss. In parallel survival experiments, mice were observed until the development of neurological signs. Healthy female mice of NSG, 4–6 weeks old, were randomly selected and used in this study for intracranial injection.

Additional experimental cell lines—293FT cells (ThermoFisher Scientific, Cat# R70007) were used to generate lentiviral particles as described in the method details section. 293FT cells were derived from embryonic kidney cells from a female human. Human astrocytes (ThermoFisher Scientific Cat#N7805100) were derived from an 18-week old female fetus and human astrocytes (SciencCell Cat#1800) were derived from a 23-week old human fetus of unknown gender (not reported by the manufacturer). The IOMM-Lee cell line was derived from a meningioma biopsy from a 61 year-old male and was provided as a courtesy of Dr. Randy Jensen (University of Utah). The CH-157MN cell line was derived from a meningioma biopsy from a 41-year old female and was a kind gift from Dr. Yancey Gillespie (University of Alabama-Birmingham). The SU-DIPG-XIII (DIPG 13) cell line was derived from a diffuse intrinsic pontine glioma specimen from a 6 year-old female pretreated with XRT. The SU-DIPG-XVII (DIPG 17) cell line was derived from a diffuse intrinsic pontine glioma specimen from an 8 year old male pretreated with XRT, avastin,

panobinostat, and everolimus. Both DIPG cell lines were a kind gift from Michelle Monje (Stanford University). The ONS-76 cell line was derived from a medulloblastoma specimen from a 2 year-old female. The DAOY cell line was derived from a medulloblastoma specimen from a 4 year-old male. Both medulloblastoma cell lines were a kind gift from Robert Wechsler-Reya (Sanford Burnham Prebys Medical Discovery Institute).

Cell Model	Cell Type	Patient Age (Years)	Patient Sex
293FT Cells	Embryonic Kidney	Embryo	Female
IOMM-Lee	Meningioma	61	Male
CH-157MN	Meningioma	41	Female
SU-DIPG-XIII	DIPG	6	Female
SU-DIPG-XVII	DIPG	8	Male
ONS-76	Medulloblastoma	2	Female
DAOY	Medulloblastoma	4	Male
Human Astrocytes (Thermo Fisher)	Astrocyte	18 weeks	Female
Human Astrocytes (ScienCell)	Astrocyte	23 weeks	Unknown

METHOD DETAILS

Enzymatic hydrolysis of DNA—The enzymatic digestion of DNA was performed as described here. One μg of genomic DNA was first spiked with uniformly ^{15}N -labeled 2'-deoxyadenosine and D_3 -labeled N^6 -methyl-2'-deoxyadenosine. DNA hydrolysis was performed with the addition of nuclease P1 (0.1 U), phosphodiesterase 2 (0.000125 U), erythro-9-(2-hydroxy-3-nonyl)adenine (EHNA) (0.25 nmol) and 4.5 μL solution containing 300 mM sodium acetate (pH 5.6) and 10 mM zinc chloride to the DNA solution. Incubation was carried out at 37°C for 24 hours. A second step digestion was carried out with 0.1 unit of alkaline phosphatase, 0.00025 unit phosphodiesterase I, 6 μL of 0.5 M Tris-HCl buffer (pH 8.9) at 37°C for 2 hours. Subsequently, the digestion mixture was neutralized by the addition of formic acid and the enzymes in the digestion mixture were removed by chloroform extraction.

LC-ESI-MS/MS and MS/MS/MS analysis—The LC-MS/MS and MS/MS/MS experiments were conducted on an LTQ-XL linear ion-trap mass spectrometer (Thermo Fisher Scientific) coupled with an EASY-nLC II system (Thermo Fisher Scientific). HPLC separation was performed by employing a homemade trapping column (150 $\mu\text{m} \times 40$ mm) and an analytical column (75 $\mu\text{m} \times 200$ mm), both packed with Magic C18 AQ (200 \AA , 5 μm , Michrom BioResource, Auburn, CA). The mass spectrometer was set up for acquiring the MS/MS of the $[\text{M} + \text{H}]^+$ ions of 2'-deoxyadenosine (m/z 252.1) and [$^{15}\text{N}_5$]-2'-deoxyadenosine (m/z 257.1), as well as the MS/MS/MS for the further cleavages of the $[\text{M} + \text{H}]^+$ ions for the nucleobase portions of N^6 -methyl-2'-deoxyadenosine (m/z 150.1) and [D_3]- N^6 -methyl-2'-deoxyadenosine (m/z 153.1).

Dot blotting—DNA samples were denatured at 99°C for 10 minutes, cooled down on ice for 3 minutes, neutralized with 10% vol of 6.6 M ammonium acetate. Samples were spotted on the membrane (Amersham Hybond-N+, GE) and air dry for 5 minutes, followed by UV-crosslink (2× auto-crosslink, 1800 UV Stratalinker, STRATAGENE). Membranes were blocked in 5% nonfat dry milk in TBST for 2 hours at room temperature, incubated with *N⁶*-mA antibodies (1:1000, Synaptic Systems, 202-003) for 16 hours at 4°C. After 5 washes, membranes were incubated with HRP linked secondary anti-rabbit IgG antibody (1:5,000, Cell Signaling 7074) for 1 hour at room temperature. Signals were detected with ECL Plus Western Blotting Reagent Pack (GE Healthcare).

Immunofluorescent staining—Cells were fixed in 4% paraformaldehyde and permeabilized in PBS containing 0.3% Triton X-100 for 15 minutes. After permeabilization, samples were treated with 2N HCl for 30 min and subsequently neutralized for 10 minutes with 0.1 M sodium borate buffer pH 8.5. Then, the samples were blocked for 2h in blocking solution (5% donkey serum in PBS, RNaseA was added to a final concentration of 100 µg/ml), followed by incubation with *N⁶*-mA antibodies (1:200, Synaptic Systems, 202-003). Primary antibodies were incubated overnight at 4°C, followed by species appropriate secondary antibodies (1:1000, Alexa 488; Invitrogen Molecular Probes, Eugene, OR) with incubation for 1 hour. Nuclei were stained with DAPI, and slides were then mounted using Fluoromount (Calbiochem, San Diego, CA). Images were taken using a Leica DM4000 Upright microscopy.

Immunohistochemical quantification of glioblastoma tissue microarrays

(TMAs)—DNA *N⁶*-mA levels in glioblastoma were investigated in glioblastoma TMAs (US Biomax, GL805b). Briefly, a TMA of deidentified formalin-fixed, paraffin-embedded (FFPE) glioblastoma specimens was immunostained for *N⁶*-mA antibodies (1:100, Synaptic Systems, 202-003). Secondary antibodies used were EnVision labeled polymer-HRP (horseradish peroxidase) anti-mouse or anti-rabbit as appropriate. Staining was visualized using 3, 3'-diaminobenzidine (DAB) chromogen (Dako, Carpinteria, CA). Each tumor was represented by three separate 2 mm cores on the TMA, with each core embedded in a separate TMA block. Each TMA core was semi-quantified on a relative scale from 0 to 3, with 0 = negative and 3 = strongest.

Vectors and lentiviral transfection—Lentiviral clones to express shRNA directed against ALKBH1 (shALKBH1.1008:TRCN0000235019, shALKBH1.1551:TRCN0000146838), or a control shRNA insert that does not target human and mouse genes (shCONT, SHC002) were obtained from Sigma-Aldrich (St. Louis, MO). The CRISPR design tool from ChopChop (<http://chopchop.cbu.uib.no/index.php>) was used to design the guide RNA (gRNA). Oligonucleotides were purchased from Fisher, and annealed and cloned into LentiCRISPR v2 plasmid, which was a gift from Dr. Feng Zhang (Addgene, plasmid 52961). The target sequence for SgRNAs used were as follows: ALKBH1 SgRNA#4: TCCGCTTCTACCGTCAGAGC, ALKBH1 SgRNA#5: GATCCTGAATTACTACCGCC. N6AMT1 SgRNA#1: GGGCTCGTACACGTCGCTGA, SgRNA#2: AAGCAGAAACGTGTCTCCG, SgRNA#3: GGGCTGGTGGCAGAAATGGT, SgRNA#4: TTGAGGTGGAGTCACTACAT.

Lentiviral particles were generated in 293FT cells in stem cell media with co-transfection with the packaging vectors pCMV-dR8.2 dvpr and pCI-VSVG (Addgene, Cambridge, MA) by Lipofectamine 2000 (Invitrogen).

Proliferation assay—Cell proliferation was measured using Cell-Titer Glo (Promega, Madison, WI). For this assay, 1,000 cells were seeded into each well of a 96-well plate. Cells were plated in 100 μ L of Neurobasal media (Invitrogen) supplemented with B27 without vitamin A (Invitrogen), EGF, and bFGF (20 ng/ml each; R&D Systems), sodium pyruvate, and glutamax. To assess cell viability, 40 μ L of the Cell-Titer Glo (Promega, Madison, WI) reagent was added to each well. Plates were placed on an orbital shaker at 100rpm for 10 minutes to promote cell lysis and covered with aluminum foil to prevent exposure to light. Plates were incubated in the dark for 2 minutes and then plates were read using a luminometer. All data were normalized to day 0 and presented as mean \pm SEM.

Generation of lentiviral particles—293FT cells (ThermoFisher Scientific, Cat# R70007) were used to generate lentiviral particles through co-transfection of the packaging vectors pCMV-dR8.2 dvpr (Addgene, Plasmid # 8455) and pCI-VSVG (Addgene, Plasmid # 1733) with the appropriate overexpression, shRNA, or other plasmid. 293FT cells were seeded at a density of 1.2 million cells in DMEM, high glucose (ThermoFisher Scientific, Cat# 11995073) in 10% Fetal Bovine Serum (ThermoFisher Scientific, Cat# 26140079 with 1% Penicillin-Streptomycin (ThermoFisher Scientific, Cat # 15140122). Cells were incubated for 24 hours prior to transfection. Transfection was performed using LipoD293™ In Vitro DNA Transfection Reagent (SignaGen Laboratories, Cat # SL100668) according to the manufacturer's instructions. Briefly, 5 μ g each of the packaging plasmids and the plasmid of interest were combined into a tube, the transfection reagent was diluted and added followed by a 15 minute incubation. The transfection mixture was then added to the 293FT cells. Media was changed after 16 hours. Virus was collected 48 hours after media change and concentrated using the Lenti-X Concentrator (Clontech Takara Bio USA, Cat # 631232) according to the manufacturers instructions. Viral supernatants were centrifuged at 1,500 x g for 45 minutes and viral pellets were frozen at -80°C for future use.

In vitro limiting dilution assay—For in vitro limiting dilution assays, decreasing numbers of cells per well (20, 10, 5, and 1) were plated in 96-well plates in neurobasal media supplemented with the components listed above. Ten days after plating, the presence and number of neurospheres in each well was quantified. Extreme limiting dilution analysis was performed using software available at <http://bioinf.wehi.edu.au/software/elda>, as previously described.

Glioblastoma Stem Cell Differentiation—For *in vitro* differentiation, glioblastoma stem cells were cultured for one week in DMEM, high glucose (ThermoFisher Scientific, Cat# 11995073) in 10% Fetal Bovine Serum (ThermoFisher Scientific, Cat# 26140079 with 1% Penicillin-Streptomycin (ThermoFisher Scientific, Cat # 15140122) as part of a previously established differentiation protocol. Cellular differentiation was verified by observing depleted expression of SOX2 protein by western blot.

Western Blotting—Cells were lysed in hypotonic buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.5% NP-40; 50 mM NaF with protease inhibitors) on ice for 15 minutes and cleared by centrifugation at 10,000 g at 4°C for 10 minutes. Protein concentration was determined using the Bradford assay (Bio-Rad Laboratories, Hercules, CA). Equal amounts of protein were mixed with reducing Laemmli loading buffer, boiled and electrophoresed on NuPAGE Gels (Invitrogen), then transferred to PVDF membranes (Millipore, Billerica, MA). Blocking was performed for 1 hour with 5% nonfat dry milk in TBST and blotting performed with primary antibodies for 16 hours at 4°C. Antibodies included ALKBH1 (Abeam, ab126596, Cambridge, MA), SOX2 (R&D Systems, AF2018), N6AMT1 (Abcam, ab173804, Cambridge, MA), GAPDH (Cell Signaling Technology, 2118).

Patient database bioinformatics—For survival analyses, TCGA data for survival analysis was accessed through the Gliovis web portal <http://gliovis.bioinfo.cnio.es/> (Bowman et al., 2017).

Intracranial tumor formation and in vivo bioluminescence imaging—GSCs were transduced with lentiviral vectors expressing ALKBH1 or a non-targeting, control (shCONT) shNRA for the knockdown experiments. 36 hours post infection, viable cells were counted and engrafted intracranially into NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ, The Jackson Laboratory, Bar Harbor, ME) mice under a University of California, San Diego Institutional Animal Care and Use Committee (IACUC) approved protocol. In parallel survival experiments, animals were monitored until they developed neurological signs.

N⁶-mA DIP sequencing—Genomic DNA from patient-derived glioblastoma stem cell models was purified with a DNeasy kit (QIAGEN, 69504). For each sample, 5 μ g DNA was sonicated to 200–300 bp with Bioruptor. Then, adaptors were ligated to genomic DNA fragments following the Illumina protocol. The ligated DNA fragments were denatured at 95 degree for 5 minutes. Then, the single-stranded DNA fragments were immunoprecipitated with N⁶-mA antibodies (5 μ g for each reaction, 202-003, Synaptic Systems) overnight at 4°C. N⁶-mA enriched DNA fragments were purified according to the Active Motif hMeDIP protocol. IP DNA and input DNA were PCR amplified with NEBNext indexing primers, and were then subjected to multiplexed library construction and sequencing with Illumina HiSeq sequencing.

ChIP-sequencing—Formaldehyde-fixed cells were lysed and sheared (Branson S220) on wet ice. The sheared chromatin was cleared and incubated overnight at 4°C with 5 μ g H3K9me3 antibody (Abcam, ab8898), H3K27me3 (active motif, 39155), or H3K4me3 (Abcam) respectively. Antibody-chromatin complexes were immunoprecipitated with protein G magnetic Dynal beads (Life Technologies), washed, eluted, reverse crosslinked, and treated with RNase A followed by proteinase K. ChIP DNA was purified using Ampure XP beads (Beckmann Coulter) and then used to prepare sequencing libraries by NEB NEBNext[®] Ultra[™] DNA Library Prep Kit for Illumina for sequencing.

ChIP-sequencing Data Analysis—Raw FASTQ files were trimmed using TrimGalore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and aligned to the hg19 human genome using the Burrows-Wheeler Aligner (BWA) (Zhang et al., 2008). Identical

ChIP-seq sequence reads were collapsed to a single read to avoid PCR duplicates using Picard and SamTools was used to sort and index the BAM files.

N⁶-mA Specific Analysis Parameters: BAM files from the *N⁶-mA* DIP-seq and matched inputs were converted to BED using the Bedtools “bamtoBed” function (Quinlan and Hall, 2010). *N⁶-mA* peaks were called with the SICER algorithm (Xu et al., 2014) using the EPIC wrapper script (<https://github.com/biocore-ntnu/epic>) with input DIP-seq files used as a control and using the default significance cutoff of False Discovery Rate (FDR) < 0.05. Peak intersections were calculated using the Bedtools “intersect” function (Quinlan and Hall, 2010). HOMER was used to perform genome annotation using the “annotatePeaks” function and for motif analysis using the “findMotifsGenome” function (Heinz et al., 2010). CEAS was also used for genome ontology analysis and for determining chromosome distribution of DIP-seq peaks (Ji et al., 2006). For gene set enrichment analysis, peaks were mapped to the closest transcription start site using the HOMER “annotatePeaks” function. Genes were then used as inputs into the online Gene Set Enrichment Analysis web portal (<http://software.broadinstitute.org/gsea/index.jsp>) (Subramanian et al., 2005) (Mootha et al., 2003). Pathway enrichment bubble plots were generated using the Bader Lab Enrichment Map Application (Merico et al., 2010) and Cytoscape (<http://www.cytoscape.org>). For subsequent visualization of *N⁶-mA* signal, bigwig files were generated using the DeepTools “bamCoverage” script using default settings (<http://deeptools.readthedocs.io/>) (Ramirez et al., 2014). Signal tracks were visualized using the Integrative Genomics Viewer, IGV (Robinson et al., 2011; Thorvaldsdottir et al., 2013). For the generation of heatmaps and profile plots, the DeepTools “computeMatrix”, “plotHeatmap”, and “plotProfile” functions were used with specific parameters given in the figure legends of the manuscript.

Histone Modification Specific Analysis Parameters: BAM files from H3K9me3, H3K27me3, and H3K4me3 ChIP-seq experiments and matched inputs were converted to BED using the Bedtools “bamtoBed” function (Quinlan and Hall, 2010). Histone modification peaks were called with the SICER algorithm (Xu et al., 2014) using the EPIC wrapper script (<https://github.com/biocore-ntnu/epic>) with input ChIP-seq files used as a control and using the default significance cutoff of False Discovery Rate (FDR) < 0.05. Intersections between *N⁶-mA* and various histone modification peaks were calculated using the Bedtools “intersect” function (Quinlan and Hall, 2010). To identify regions of gained H3K9me3 enrichments following ALKBH1 knockdown, the SICER algorithm was used as described above with the ALKBH1-knockdown H3K9me3 ChIP-seq file serving as the “treatment” file and the non-targeting shRNA H3K9me3 ChIP-seq file serving as the “control” file using default settings as described above. Reciprocally, to identify regions of lost H3K9me3 enrichments following ALKBH1 knockdown, the SICER algorithm was used as described above with the non-targeting H3K9me3 ChIP-seq file serving as the “treatment” file and the ALKBH1-knockdown H3K9me3 ChIP-seq file serving as the “control” file. Consensus gained and lost peaks were generated by identifying peaks observed in both technical replicates using the bedtools “intersect” function (Quinlan and Hall, 2010) and results are plotted as a volcano plot using the R programming language. For subsequent visualization of histone modification ChIP-seq signal, bigwig files were generated using the DeepTools “bamCoverage” script using default settings

(<http://deeptools.readthedocs.io/>) (Ramirez et al., 2014). Signal tracks were visualized using the Integrative Genomics Viewer, IGV (Robinson et al., 2011; Thorvaldsdottir et al., 2013). For the generation of heatmaps and profile plots, the DeepTools “computeMatrix”, “plotHeatmap”, and “plotProfile” functions were used with specific parameters given in the figure legends of the manuscript.

ALKBH1 ChIP-seq Specific Parameters: Peaks were called using the MACS2 “callpeak” function using default settings using paired-end BAM files from ALKBH1 ChIP-seq experiments using the “-f BAMPE” parameter and the default FDR-corrected q-value of 0.05 with a matched ChIP-seq input file serving as the control. The bedtools “intersect” function was used to calculate overlaps with N^6 -mA peaks, and the bedtools “shuffle” function was used to generate a “genome background” expected overlaps file (Quinlan and Hall, 2010). For subsequent visualization of ALKBH1 ChIP-seq signal, bigwig files were generated using the DeepTools “bamCoverage” script using default settings (<http://deeptools.readthedocs.io/>) (Ramirez et al., 2014). For the generation of heatmaps and profile plots, the DeepTools “computeMatrix”, “plotHeatmap”, and “plotProfile” functions were used with specific parameters given in the figure legends of the manuscript.

RNA-sequencing—Cellular RNA was extracted using the Qiagen miRNeasy Mini Kit (catalog number 217004) according to the manufacturers instructions. Total RNA was prepared for sequencing using the Illumina TruSeq Stranded Total RNA Library Prep Kit. For samples treated in normoxia, FASTQ sequencing files were trimmed and then aligned to the hg19 human genome using the STAR aligner (Dobin et al., 2013). Differential expression analysis was performed using CuffDiff (<http://cole-trapnell-lab.github.io/cufflinks/cuffdiff/>). For samples treated in hypoxia, FASTQ sequencing reads were trimmed using Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and transcript quantification was performed using Salmon in the quasi-mapping mode (Patro et al., 2017). Salmon “quant” files were converted using Tximport (<https://bioconductor.org/packages/release/bioc/html/tximport.html>) and differential expression analysis was performed using DESeq2 (Love et al., 2014). Gene set enrichment analysis was performed by selecting differentially expressed genes (FDR-corrected p-value < 0.05), generating a pre-ranked list using the gene expression fold change as the ranking metric, and inputting a pre-ranked list into the GSEA desktop application (<http://software.broadinstitute.org/gsea/downloads.jsp>) (Subramanian et al., 2005) (Mootha et al., 2003). Pathway enrichment bubble plots were generated using the Bader Lab Enrichment Map Application (Merico et al., 2010) and Cytoscape (<http://www.cytoscape.org>).

ATAC-seq library generation—ATAC-seq was performed on 50,000 nuclei. The samples were permeabilized in cold permabilization buffer (0.2% IGEPAL-CA630 (I8896, Sigma), 1 mM DTT (D9779, Sigma), Protease inhibitor (05056489001, Roche), 5% BSA (A7906, Sigma) in PBS (10010-23, Thermo Fisher Scientific)) for 10 minutes on the rotator in the cold room and centrifuged for 5 min at 500 xg at 4°C. The pellet was resuspended in cold tagmentation buffer (33 mM Tris-acetate (pH = 7.8) (BP-152, Thermo Fisher Scientific), 66 mM K-acetate (P5708, Sigma), 11 mM Mg-acetate (M2545, Sigma), 16 % DMF (DX1730,

EMD Millipore) in Molecular biology water (46000-CM, Corning)) and incubated with Tagmentation enzyme (FC-121-1030; Illumina) at 37 °C for 30 min with shaking 500 rpm. The tagementated DNA was purified using MinElute PCR purification kit (28004, Qiagen). The libraries were amplified using NEBNext High-Fidelity 2X PCR Master Mix (M0541, NEB) with primer extension at 72°C for 5 minutes, denaturation at 98°C for 30s, followed by 8 cycles of denaturation at 98°C for 10 s, annealing at 63°C for 30 seconds and extension at 72°C for 60 seconds. After the purification of amplified libraries using MinElute PCR purification kit (28004, Qiagen), the double size selection was performed using SPRIselect bead (B23317, Beckman Coulter) with 0.55X beads and 1.5X to sample volume. Finally, the libraries were sequenced on HiSeq4000 (Paired-end 50 cycles, Illumina).

ATAC-seq data processing—Adapter-trimmed fastq files were aligned to hg19 by Bowtie2 (Langmead and Salzberg, 2012) using parameters “-X2000 --mm --local”. After filtering by using samtools (Li et al., 2009) with “-q 30 -F 1804 -f 2”, only primary and proper mated reads were left to remove PCR duplicate by using “markduplicate” from Picard tools (<http://broadinstitute.github.io/picard/>). The remaining mapped reads were shift +4 bp and -5 bp for “+” and “-” strand respectively to adjust for TN5 dimer so that the first base of each reads represents the cutting site. Then the peak calling was performed by using MACS2 (Zhang et al., 2008) with “-f bed -p 0.01 -shift -75 -extsize 150”. The output “narrowPeaks” were further filtered out the blacklist regions (Consortium, 2012). The detailed pipeline can be found in https://github.com/epigen-UCSD/atac_seq_pipeline, which is based on Kundaje lab’s pipeline https://github.com/kundajelab/atac_dnase_pipelines with a few small changes documented on the repo.

To calculate gained and lost peaks following ALKBH1 knockdown, ATAC-seq BAM files along with narrowpeak files were analyzed using DiffBind (<https://bioconductor.org/packages/release/bioc/html/DiffBind.html>) (Ross-Innes et al., 2012) and DESeq2 (Love et al., 2014) to identify regions of quantitatively differential chromatin accessibility following ALKBH1 knockdown. DiffBind was used to generate MA-plots and supervised counts heatmaps. For the generation of heatmaps and profile plots, the DeepTools “computeMatrix”, “plotHeatmap”, and “plotProfile” functions were used with specific parameters given in the figure legends of the manuscript.

Bisulfite Sequencing—Genomic DNA was extracted from human patient-derived glioblastoma stem cells and bisulfite conversion was performed. Following paired-end sequencing, FASTQ files separated into 11 paired-end fastq files to facilitate processing. Reads were trimmed using Trim Galore to remove 10bp from the 5’ and 3’ end of both reads. Bowtie2 (Langmead and Salzberg, 2012) was used to map the reads to a pre-prepared hg19 genome in the non-directional and paired-end mode. The De-duplicate function in Bismark (Krueger and Andrews, 2011) was used to remove duplicate reads. All of the mapping results from each pair of small fastq files were merged together using the samtools merge method, with the -n option to sort the reads by names. The function “bismark_methylation_extractor” was applied on the merged mapping results to extract the methylation information. The coverage2cytosine function was used to generate a genome-wide cytosine methylation report. The MakeTagDirectory function in homer was

applied on the cytosine methylation report, with the option of -format bismark, -minCounts 5 -checkGC to calculate percentage of methylation in the CpG locations across the genome. The annotatePeaks.pl script from HOMER (Heinz et al., 2010) was used to obtain the methylation level of selected regions of interest. We calculated the average methylation in the 12,000 bp regions centered on the N^6 -mA peaks as described in the manuscript and TSS. The cytosine methylation report was also used to calculate the methylation level for each individual N^6 -mA peak, 1000 bp regions ahead of the TSS, and also for a bed file with shuffled chip-seq peak across the genome. Three violin plots were generated. Rank-sum test was applied between three of the distributions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis for each plot is listed in the figure legend and/or in corresponding methods above. In brief, all grouped data are presented as mean \pm s.d. All box and whisker plots of expression data are presented as median (middle line of box) \pm the 25 percentile (top and bottom line of box, respectively). P-values presented are calculated by two-sided Mann-Whitney U-test. Kaplan-Meier curves were generated and log-rank (Mantel-Cox) analysis was performed to generate P values using GraphPadPrism software (GraphPad Software). Sample sizes for each experiment are given in corresponding figures and/or methods above. Sizes were chosen based on previous experience with given experiments, or in the case of retrospective analysis, all available samples were included.

DATA AND SOFTWARE AVAILABILITY

All raw and selected processed data files are available on the Gene Expression Omnibus (GEO). All data can be accessed at the SuperSeries accession number: GSE118093 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE118093>). Direct links to the RNA-sequencing data can be found at the SubSeries accession number: GSE117632. Direct links to ATAC-sequencing data can be found at the SubSeries accession number: GSE118092. Direct links to ChIP-sequencing, whole genome bisulfite sequencing, and N^6 -mA DIP-sequencing data can be found at the SubSeries accession number: GSC119081. Additional data will be provided upon request. There are no restrictions on data availability.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank JingTian Zhou (Bioinformatics and Systems Biology, UCSD) for analysis of bisulfite sequencing. We thank Erik Sulman (MD Anderson Cancer Center) for providing the GSC23 glioblastoma cell model. We thank our funding sources: M.S.: DK104202, DK112155, DK078803, DK068471. S.B.: CA184090, NS091080, NS099175. B.C.P: CA217066. R.C.G: CA217065. L.J.Y.K: CA203101. A.Z.X.: R01GM114205-01 and The Ludwig Family Foundation. J.N.R.: CA197718, CA154130, CA169117, CA171652, NS087913, NS089272, NS103434. SCM: CPRIT scholar award, ALSF young investigator award, Rally research grant, BEAR Necessities Pediatric Cancer Foundation Grant, Children's Cancer Research Fund award, and Baylor College of Medicine Junior Faculty Award. AES: A217956-01, CA187780, NS 089272 as well as the Coulter, Ferry Foundation, Kimble and Gerald Kaufman Jr. Foundations and the Peter D. Cristal Chair at University Hospitals. The graphical abstract figure was prepared in part using images from Servier Medical Art by Servier (<https://smart.servier.com/>), which is licenced under a Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>).

REFERENCES

- Abdoh M, Facchino S, Chato W, Balasingam V, Ferreira J, and Bernier G (2009). BMI1 sustains human glioblastoma multiforme stem cell renewal. *J Neurosci* 29, 8884–8896. [PubMed: 19605626]
- Bowman RL, Wang Q, Carro A, Verhaak RG, and Squatrito M (2017). GlioVis data portal for visualization and analysis of brain tumor expression datasets. *Neuro Oncol* 19, 139–141. [PubMed: 28031383]
- Cadioux B, Ching TT, VandenBerg SR, and Costello JF (2006). Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. *Cancer research* 66, 8469–8476. [PubMed: 16951158]
- Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. [PubMed: 22955616]
- Cui Q, Shi H, Ye P, Li L, Qu Q, Sun G, Sun G, Lu Z, Huang Y, Yang CG, et al. (2017). m(6)A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep* 18, 2622–2634. [PubMed: 28297667]
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. [PubMed: 23104886]
- Du J, Johnson LM, Jacobsen SE, and Patel DJ (2015). DNA methylation pathways and their crosstalk with histone methylation. *Nature reviews Molecular cell biology* 16, 519–532. [PubMed: 26296162]
- Elsasser SJ, Allis CD, and Lewis PW (2011). Cancer. New epigenetic drivers of cancers. *Science* 331, 1145–1146. [PubMed: 21385704]
- Esteller M (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* 21, 5427–5440. [PubMed: 12154405]
- Feinberg AP, and Vogelstein B (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 301, 89–92. [PubMed: 6185846]
- Fu Y, Dominissini D, Rechavi G, and He C (2014). Gene expression regulation mediated through reversible m(6)A RNA methylation. *Nat Rev Genet* 15, 293–306. [PubMed: 24662220]
- Fu Y, Luo GZ, Chen K, Deng X, Yu M, Han D, Hao Z, Liu J, Lu X, Dore LC, et al. (2015). N6-methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*. *Cell* 161, 879–892. [PubMed: 25936837]
- Greer EL, Blanco MA, Gu L, Sendinc E, Liu J, Aristizabal-Corrales D, Hsu CH, Aravind L, He C, and Shi Y (2015). DNA Methylation on N6-Adenine in *C. elegans*. *Cell* 161, 868–878. [PubMed: 25936839]
- Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38, 576–589. [PubMed: 20513432]
- Hu Y, and Smyth GK (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods* 347, 70–78. [PubMed: 19567251]
- Intlekofer AM, Dematteo RG, Venneti S, Finley LW, Lu C, Judkins AR, Rustenburg AS, Grinaway PB, Chodera JD, Cross JR, et al. (2015). Hypoxia Induces Production of L-2-Hydroxyglutarate. *Cell metabolism* 22, 304–311. [PubMed: 26212717]
- Ji X, Li W, Song J, Wei L, and Liu XS (2006). CEAS: cis-regulatory element annotation system. *Nucleic Acids Res* 34, W551–554. [PubMed: 16845068]
- Jin X, Kim LJY, Wu Q, Wallace LC, Prager BC, Sanvoranart T, Gimple RC, Wang X, Mack SC, Miller TE, et al. (2017). Targeting glioma stem cells through combined BMI1 and EZH2 inhibition. *Nat Med* 23, 1352–1361. [PubMed: 29035367]
- Kaelin WG Jr. (2017). Common pitfalls in preclinical cancer target validation. *Nature reviews Cancer* 17, 425–440. [PubMed: 28642604]
- Krueger F, and Andrews SR (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572. [PubMed: 21493656]

- Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357–359. [PubMed: 22388286]
- Li H, and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. [PubMed: 19451168]
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. [PubMed: 19505943]
- Liu F, Clark W, Luo G, Wang X, Fu Y, Wei J, Wang X, Hao Z, Dai Q, Zheng G, et al. (2016). ALKBH1-Mediated tRNA Demethylation Regulates Translation. *Cell* 167, 816–828 e816. [PubMed: 27745969]
- Liu P, Nie S, Li B, Yang ZQ, Xu ZM, Fei J, Lin C, Zeng R, and Xu GL (2010). Deficiency in a glutamine-specific methyltransferase for release factor causes mouse embryonic lethality. *Mol Cell Biol* 30, 4245–4253. [PubMed: 20606008]
- Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. [PubMed: 25516281]
- Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, and Young RA (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153, 320–334. [PubMed: 23582323]
- Luo W, Chang R, Zhong J, Pandey A, and Semenza GL (2012). Histone demethylase JMJD2C is a coactivator for hypoxia-inducible factor 1 that is required for breast cancer progression. *Proc Natl Acad Sci U S A* 109, E3367–3376. [PubMed: 23129632]
- Martin M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17, 10–12.
- Merico D, Isserlin R, Stueker O, Emili A, and Bader GD (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5, e13984. [PubMed: 21085593]
- Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34, 267–273. [PubMed: 12808457]
- Northcott PA, Lee C, Zichner T, Stutz AM, Erkek S, Kawauchi D, Shih DJ, Hovestadt V, Zapatka M, Sturm D, et al. (2014). Enhancer hijacking activates GFII family oncogenes in medulloblastoma. *Nature* 511, 428–434. [PubMed: 25043047]
- Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, Pan F, Pelloski CE, Sulman EP, Bhat KP, et al. (2010). Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* 17, 510–522. [PubMed: 20399149]
- O'Brown ZK, and Greer EL (2016). N6-Methyladenine: A Conserved and Dynamic DNA Mark. *Adv Exp Med Biol* 945, 213–246. [PubMed: 27826841]
- Patro R, Duggal G, Love MI, Irizarry RA, and Kingsford C (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14, 417–419. [PubMed: 28263959]
- Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. [PubMed: 20110278]
- Ramirez F, Dundar F, Diehl S, Gruning BA, and Manke T (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 42, W187–191. [PubMed: 24799436]
- Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44, W160–165. [PubMed: 27079975]
- Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, and Mesirov JP (2006). GenePattern 2.0. *Nat Genet* 38, 500–501. [PubMed: 16642009]
- Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, and Mesirov JP (2011). Integrative genomics viewer. *Nature biotechnology* 29, 24–26.
- Rogers JC, and Rogers SW (1995). Comparison of the effects of N6-methyldeoxyadenosine and N5-methyldeoxycytosine on transcription from nuclear gene promoters in barley. *Plant J* 7, 221–233. [PubMed: 7704046]

- Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481, 389–393. [PubMed: 22217937]
- Saksouk N, Barth TK, Ziegler-Birling C, Olova N, Nowak A, Rey E, Mateos-Langerak J, Urbach S, Reik W, Torres-Padilla ME, et al. (2014). Redundant mechanisms to form silent chromatin at pericentromeric regions rely on BEND3 and DNA methylation. *Mol Cell* 56, 580–594. [PubMed: 25457167]
- Sanjana NE, Shalem O, and Zhang F (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11, 783–784. [PubMed: 25075903]
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498–2504. [PubMed: 14597658]
- Soneson C, Love MI, and Robinson MD (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* 4, 1521. [PubMed: 26925227]
- Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, Sabatini DM, Chen IS, Hahn WC, Sharp PA, et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* 9, 493–501. [PubMed: 12649500]
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545–15550. [PubMed: 16199517]
- Suva ML, Riggi N, Janiszewska M, Radovanovic I, Provero P, Stehle JC, Baumer K, Le Bitoux MA, Marino D, Cironi L, et al. (2009). EZH2 Is Essential for Glioblastoma Cancer Stem Cell Maintenance. *Cancer research* 69, 9211–9218. [PubMed: 19934320]
- Thorvaldsdottir H, Robinson JT, and Mesirov JP (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* 14, 178–192. [PubMed: 22517427]
- Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, Campos C, Fabius AW, Lu C, Ward PS, et al. (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* 483, 479–483. [PubMed: 22343889]
- Vanyushin BF, Belozersky AN, Kokurina NA, and Kadirova DX (1968). 5-methylcytosine and 6-methylamino-purine in bacterial DNA. *Nature* 218, 1066–1067. [PubMed: 5656625]
- Widschwendter M, Jiang G, Woods C, Muller HM, Fiegl H, Goebel G, Marth C, Muller-Holzner E, Zeimet AG, Laird PW, et al. (2004). DNA hypomethylation and ovarian cancer biology. *Cancer research* 64, 4472–4480. [PubMed: 15231656]
- Wu TP, Wang T, Seetin MG, Lai Y, Zhu S, Lin K, Liu Y, Byrum SD, Mackintosh SG, Zhong M, et al. (2016). DNA methylation on N(6)-adenine in mammalian embryonic stem cells. *Nature* 532, 329–333. [PubMed: 27027282]
- Xiao CL, Zhu S, He M, Chen, Zhang Q, Chen Y, Yu G, Liu J, Xie SQ, Luo F, et al. (2018). N(6)-Methyladenine DNA Modification in the Human Genome. *Mol Cell*.
- Xu S, Grullon S, Ge K, and Peng W (2014). Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. *Methods Mol Biol* 1150, 97–111. [PubMed: 24743992]
- Yao B, Cheng Y, Wang Z, Li Y, Chen L, Huang L, Zhang W, Chen D, Wu H, Tang B, et al. (2017). DNA N6-methyladenine is dynamically regulated in the mouse brain following environmental stress. *Nature communications* 8, 1122.
- Zhang GQ, Huang H, Liu D, Cheng Y, Liu XL, Zhang WX, Yin RC, Zhang DP, Zhang P, Liu JZ, et al. (2015). N-6-Methyladenine DNA Modification in *Drosophila*. *Cell* 161, 893–906. [PubMed: 25936838]
- Zhang S, Zhao BS, Zhou A, Lin K, Zheng S, Lu Z, Chen Y, Sulman EP, Xie K, Bogler O, et al. (2017). m(6)A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program. *Cancer Cell* 31, 591–606 e596. [PubMed: 28344040]

- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137. [PubMed: 18798982]
- Zhao M, Kim P, Mitra R, Zhao J, and Zhao Z (2016). TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res* 44, D1023–1031. [PubMed: 26590405]
- Zhu S, Beaulaurier J, Deikus G, Wu TP, Strahl M, Hao Z, Luo G, Gregory JA, Chess A, He C, et al. (2018). Mapping and characterizing N6-methyladenine in eukaryotic genomes using single-molecule real-time sequencing. *Genome Res* 28, 1067–1078. [PubMed: 29764913]

HIGHLIGHTS

- N(6)-methyladenine DNA modification is markedly upregulated in glioblastoma.
- N^6 -mA co-localizes with H3K9me3-marked heterochromatin domains.
- The DNA demethylase, ALKBH1, dynamically controls N^6 -mA levels.
- Targeting N^6 -mA regulation kills glioblastoma stem cells.

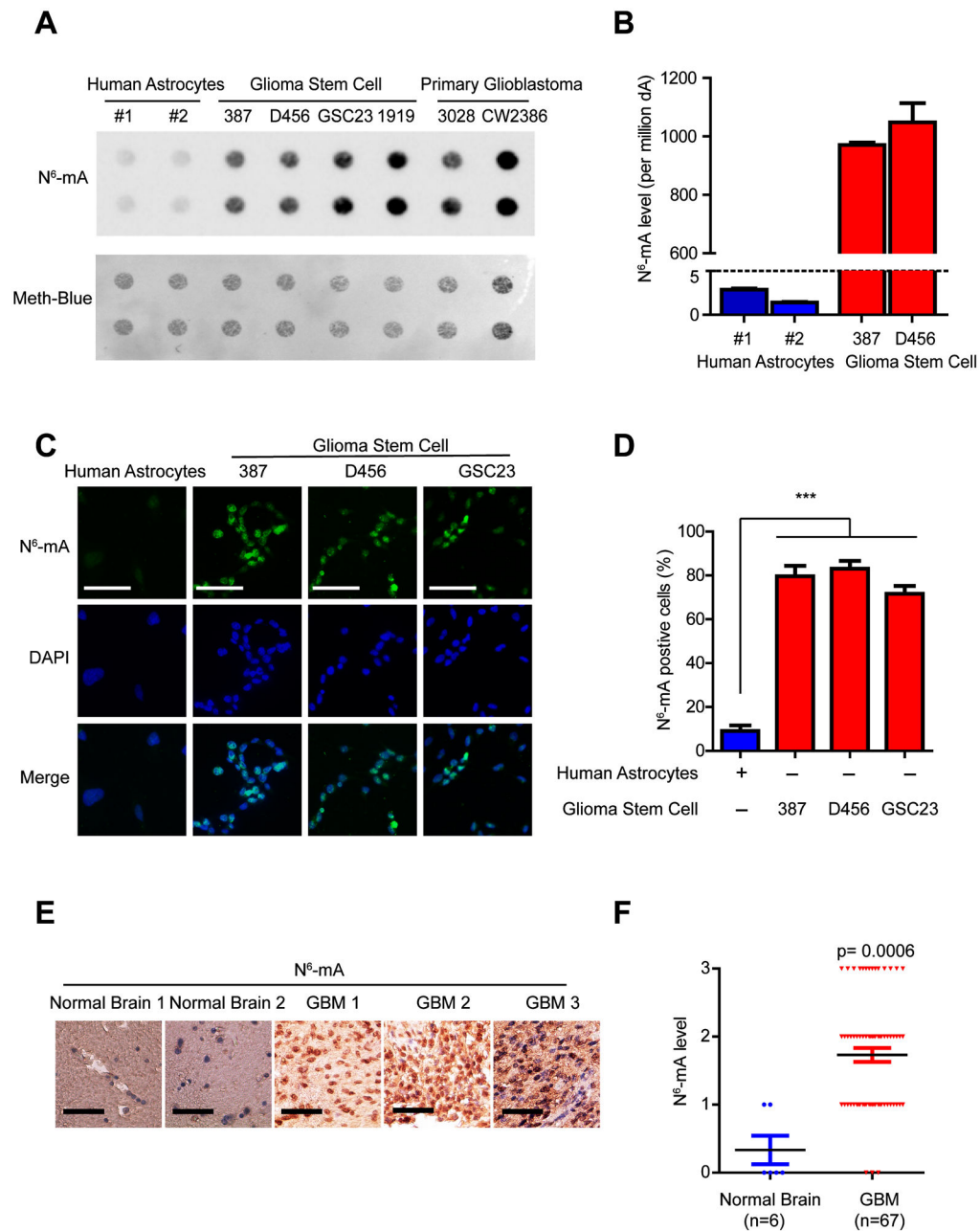


Figure 1. Identification of N(6)-methyladenine (N^6 -mA) DNA modification in human glioblastoma. See also Figure S1.

(A) Levels of the N^6 -mA DNA modification were assessed via DNA dot blot in (1) normal human astrocytes, (2) patient-derived GSC models (387, D456, GSC23, and 1919) and (3) primary human glioblastoma specimens (3028, CW2386) using an N^6 -mA-specific antibody. Methylene blue detected DNA loading.

(B) Mass spectrometry analysis of N^6 -mA in two normal human astrocyte cell lines and two patient-derived GSC models (387 and D456). Data are presented as mean \pm SD. Two replicates were used for each sample. Significance was determined by one-way ANOVA

with Tukey multiple comparison test. $P < 0.0001$ for each human astrocyte vs. GSC comparison.

(C) N^{σ} -mA DNA immunofluorescence in normal human astrocytes and human patient-derived GSC models (387, D456, GSC23). DAPI indicates cell nuclei. Scale bars, 50 μm .

(D) Quantification of percentage of N^{σ} -mA positive cells by immunofluorescence staining in (C). N^{σ} -mA was quantified counting 100 cells from each sample. $N = 3$ slides/cell type. Data are presented as mean \pm SD. Significance assessed by ANOVA. ***, $P < 0.001$.

(E) Immunohistochemistry (IHC) staining of N^{σ} -mA in non-neoplastic brain tissues (6 total) and human primary glioblastoma specimens (67 total) from a tissue microarray. Scale bar, 50 μm .

(F) Quantification of N^{σ} -mA levels from immunohistochemistry staining in (E). N^{σ} -mA levels were scored from low levels (score = 0) to highest levels (score = 3). Data are presented as mean \pm standard deviation. Student's t-test, $P = 0.0006$.

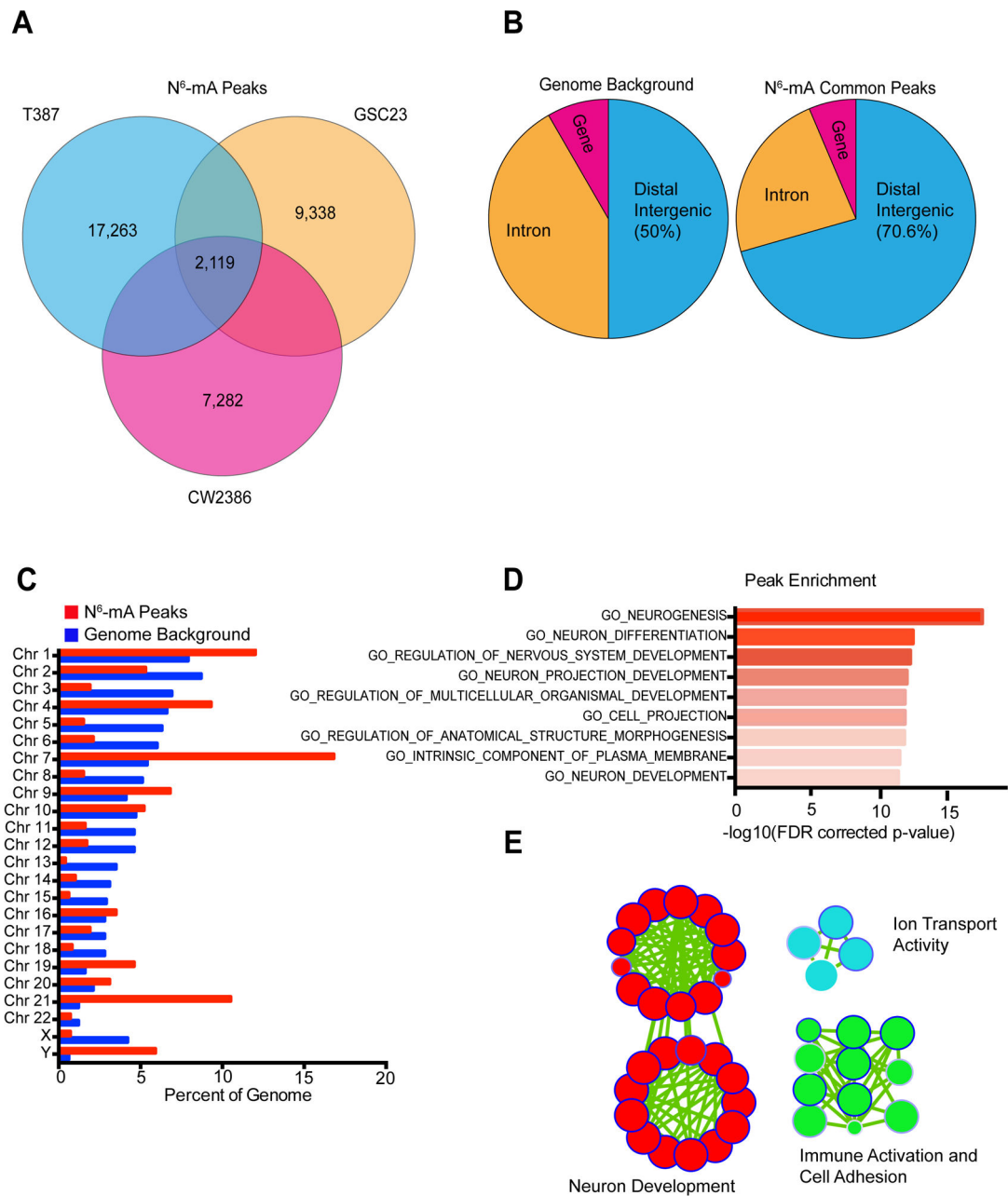


Figure 2. Genomic localization of N⁶-mA enrichment in glioma stem cells. See also Figure S2.

(A) N⁶-mA enrichment was identified using SICER for a human primary glioma sample (CW2386) and two in vitro human patient-derived glioma models (T387 and GSC23). The number of N⁶-mA peaks are shown with the number of shared peaks in the center.

(B) Genome Ontology analysis showing the fraction of common N⁶-mA peaks present in distal intergenic, intronic, or gene regions compared to genome background. (Chi-squared test, $P < 0.0001$)

(C) Genome Ontology analysis showing the fraction of common N⁶-mA peaks on each chromosome compared to genome background. Chromosome 7: $p = 1.2 \times 10^{-208}$;

Chromosome 21: $p = 2.2 \times 10^{-322}$; Chromosome 3: $p = 5.1 \times 10^{-68}$; Chromosome 5: $p = 1.2 \times 10^{-69}$.

(D) Gene Ontology (GO) enrichment analysis of the closest gene to each of the common N^6 -mA peaks. Values are expressed as $-\log_{10}$ (FDR corrected p-value).

(E) Enrichment map demonstrating key pathways identified in the GO enrichment analysis in panel (D). Circles represent individual gene sets. The size of the circle depicts the number of genes in the gene set and the edge color depicts the FDR-corrected p-value of the enrichment, with dark blue representing the most significant gene sets.

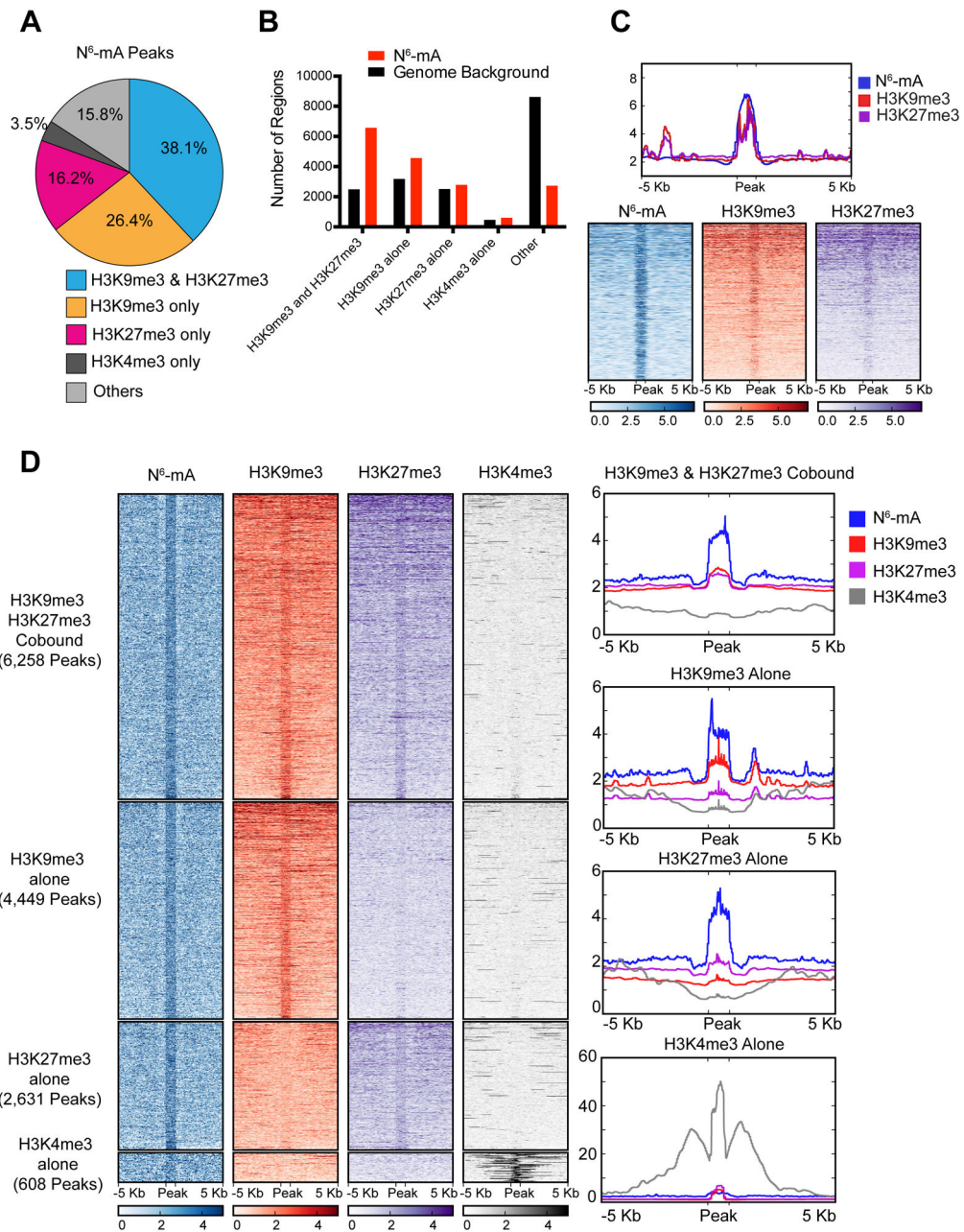


Figure 3. Intersection of N^6 -mA peaks with heterochromatin-associated histone modification domains in glioma stem cells. See also Figure S3.

(A) The fraction of N^6 -mA peaks in a patient derived GSC model (387) that overlap with matched H3K9me3, H3K27me3, and H3K4me3 histone modification domains in the same model.

(B) Number of overlaps between N^6 -mA peaks and histone modification domains in a patient derived GSC model (387) relative to genome background.

(C) Heatmap and Profile plot demonstrating the intersection of N^6 -mA common peaks with H3K9me3 and H3K27me3 signal over scaled window 5kb upstream and downstream of the common N^6 -mA peak.

(D) Heatmap of N^6 -mA peaks in a GSC model (387) divided into those that are cobound with H3K9me3 and H3K27me3, those bound by H3K9me3 or H3K27me3 alone, and those bound by H3K4me3 alone. Signal is shown over a scaled window 5kb upstream and downstream of the N^6 -mA peak and the height of the heatmap is directly proportional to the number of regions present in each segment.

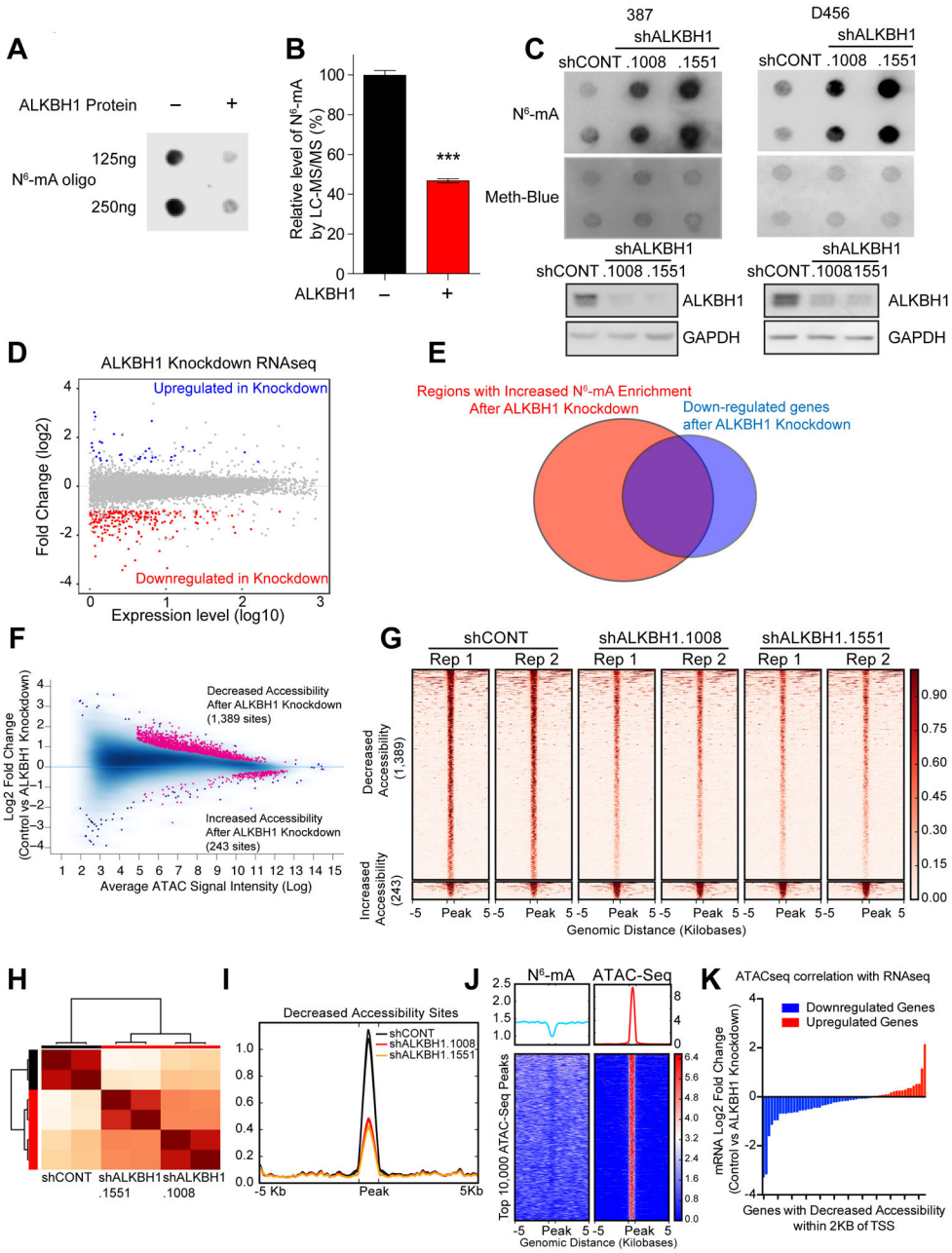


Figure 4. ALKBH1 is a N^6 -mA DNA demethylase in human glioblastoma and contributes to N^6 -mA co-localization with H3K9me3 genome wide. See also Figure S4.

(A) N^6 -mA labelled DNA oligonucleotides were treated in a cell-free in vitro demethylase reaction with recombinant human ALKBH1 proteins. Results are depicted by dot blot after treatment of two quantities of substrate DNA oligonucleotides.

(B) In vitro demethylation reaction was quantified by LC-MS/MS mass spectrometry following addition of ALKBH1 protein to N^6 -mA labelled DNA oligonucleotides. Data are presented as mean \pm standard deviation. (Student's t-test. ***, $P < 0.001$. $N = 3$)

(C) ALKBH1 expression was decreased using two independent shRNAs in human patient-derived GSC models and N^6 -mA levels were assessed using a DNA dot blot. Methylene blue detected DNA loading. ALKBH1 protein level was assessed using western blot.

(D) RNA-seq analysis in a human patient-derived GSC model following ALKBH1 knockdown. Blue dots indicate the most highly upregulated genes following ALKBH1 knockdown (37), while the red dots indicate the most highly downregulated genes following ALKBH1 knockdown (321).

(E) Venn Diagram indicates overlap between (1) genomic regions with gained N^6 -mA after ALKBH1 knockdown and (2) downregulated genes after ALKBH1 knockdown.

(F) Genome wide differentially accessible sites were identified by ATAC-seq. 1,632 sites of differential accessibility (FDR-corrected p-value < 0.05) were identified and are visualized in a MA-plot. 1,389 sites were identified with decreased accessibility after ALKBH1 knockdown while 243 sites displayed increased accessibility. Log2 fold change > 0.5; p < 0.05.

(G) Chromatin accessibility heatmap for the differentially accessible sites (absolute value of Log2 fold change > 1). Two replicates were performed for each treatment group. Signal is shown over a scaled window 5 kb upstream and downstream of the differentially accessible region.

(H) Supervised heatmap showing the correlation between the ATAC-seq counts in non-targeting and ALKBH1-knockdown samples. Two replicates were performed for each sample.

(I) Profile plot showing ATAC-seq signal over sites of decreased chromatin accessibility in ALKBH1-knockdown samples. Scaled signal is shown 5 kilobases upstream and downstream of each ATAC peak.

(J) Heatmap and profile plot showing the N^6 -mA DIP-seq and ATAC-seq signals over the top 10,000 ranked ATAC-seq peaks in the 387 GSC model.

(K) Graph showing the mRNA fold change following ALKBH1 knockdown for genes with sites of decreased accessibility. Among genes with decreased accessibility sites within 2KB of their transcriptional start site, 37 genes were downregulated while 19 genes were upregulated.

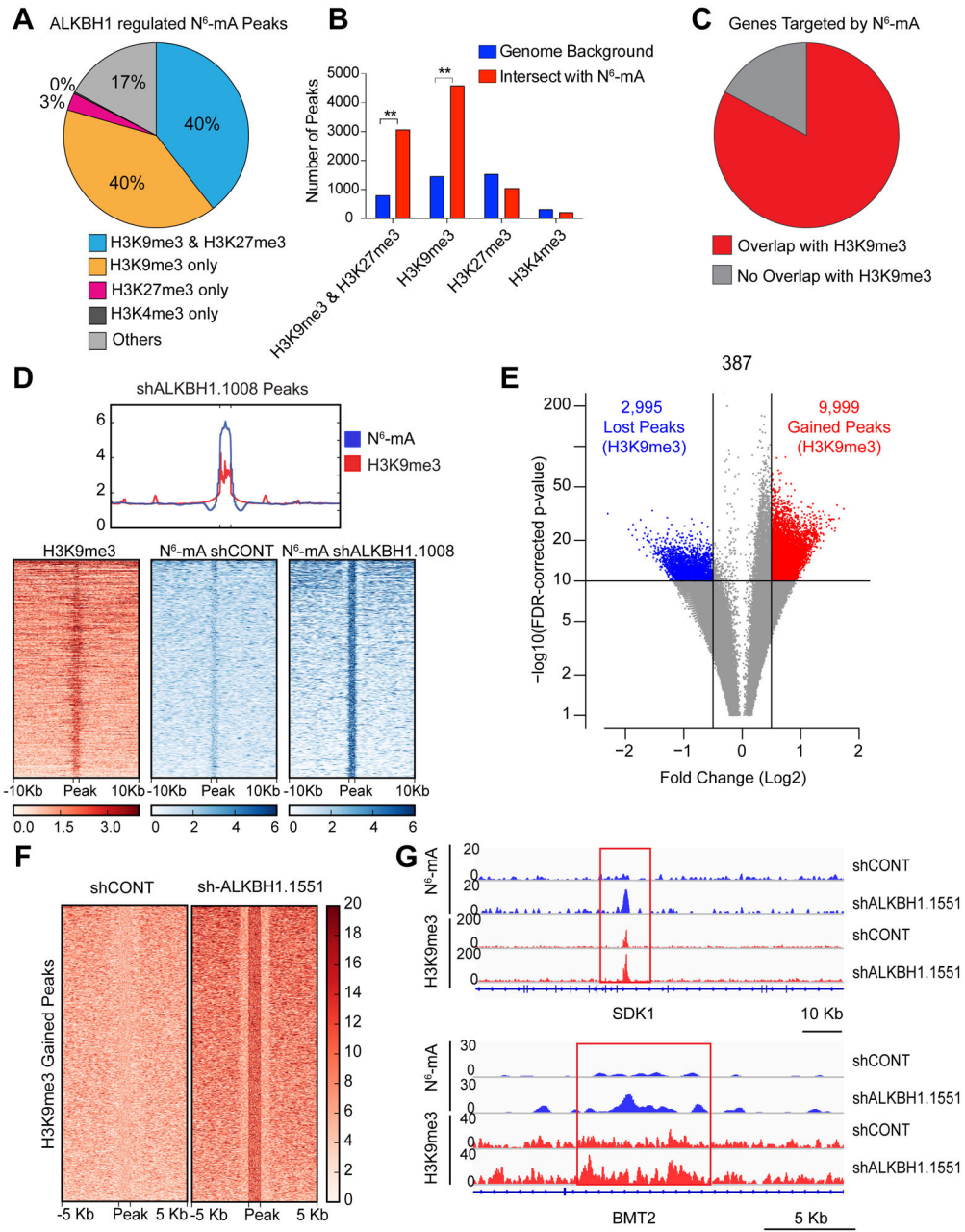


Figure 5. ALKBH1 regulates downstream gene expression through N^6 -mA-dependent heterochromatin formation. See also Figures S5 and S6.

(A) Pie chart depicts the fraction of N^6 -mA peaks gained after ALKBH1 knockdown in a patient derived GSC model (387) that overlap with matched H3K9me3, H3K27me3, and H3K4me3 histone modification domains in the same model.

(B) Overlaps between N^6 -mA peaks gained following ALKBH1 knockdown and histone modification domains in a patient derived GSC model (387) relative to genome background. **, $p < 0.01$.

(C) Pie chart shows the percentage of genes targeted by the N^6 -mA DNA modification that overlap with the H3K9me3 histone modification.

- (D)** Heatmap shows gained N^6 -mA peaks after ALKBH1 knockdown with shRNA (shALKBH1.1008) and co-localization with H3K9me3. Signal is shown over a scaled window 10kb upstream and downstream of the gained N^6 -mA peak.
- (E)** Volcano plot shows the number of H3K9me3 peaks gained (red: 9,999) and lost (blue: 2,995) following ALKBH1 knockdown in the 387 GSC model.
- (F)** Heatmap indicating sites of H3K9me3 enrichment following ALKBH1 knockdown with shALKBH1.1551. Signal is shown over a scaled window 5 kb upstream and downstream of the gained N^6 -mA peak.
- (G)** Examples of genes (SDK1 and BMT2) showing co-localization of gained N^6 -mA DNA modification peaks with gained H3K9me3 peaks following knockdown of ALKBH1 with shALKBH1.1551.

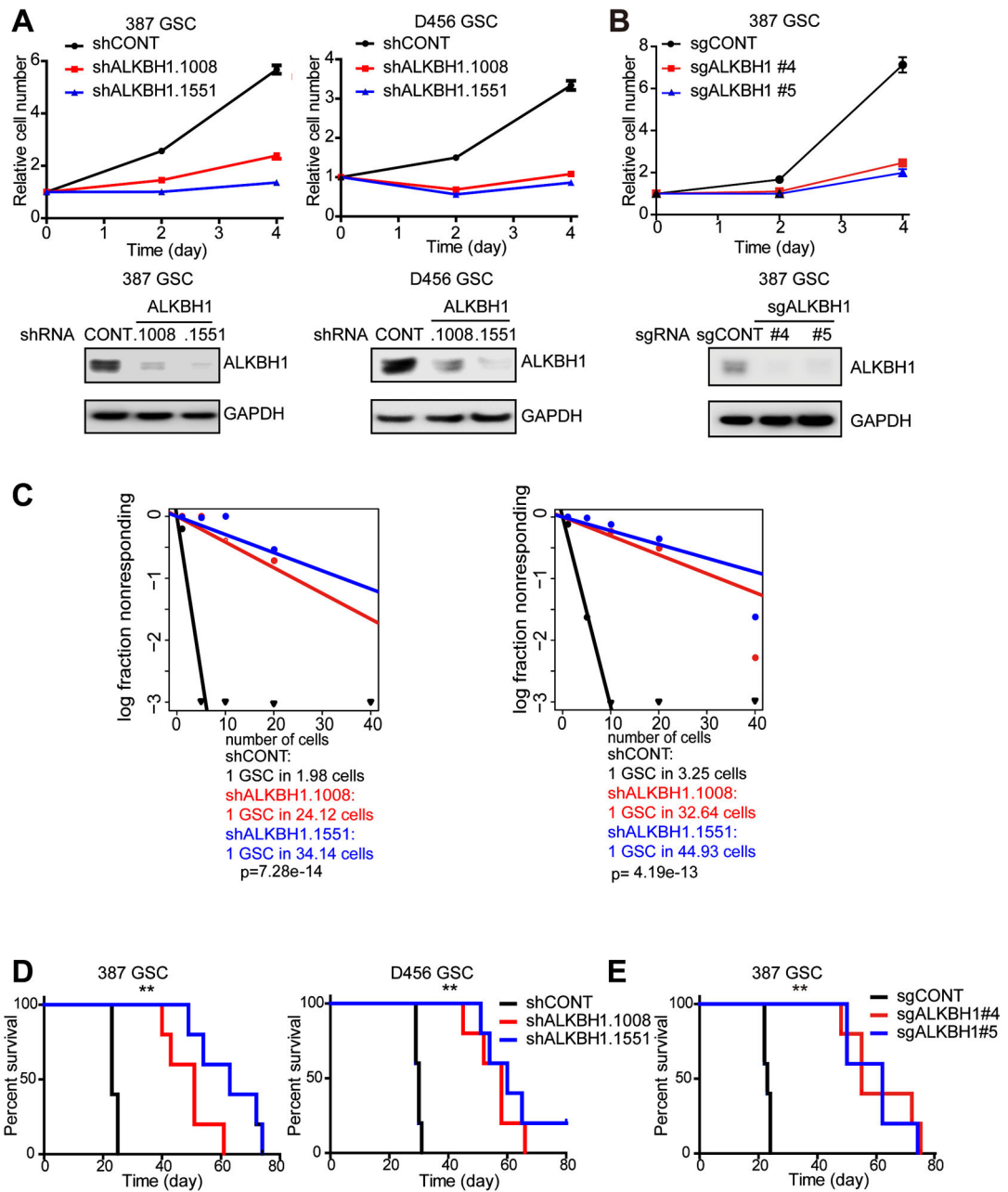


Figure 6. ALKBH1 is essential for glioblastoma stem cell growth, self-renewal, and tumor formation capacity. See also Figure S7.

(A) (Top) Cell proliferation was assessed over a 4 day time course after treatment with a non-targeting control shRNA (shCONT) or two independent non-overlapping shRNAs (shALKBH1.1008 and shALKBH1.1551) in two human patient derived GSC models (387 and D456). Significance was determined by two-way repeated measures ANOVA with Dunnett multiple test correction. $P < 0.0001$. (Bottom) Knockdown efficiency of shRNAs targeting ALKBH1 was assessed by immunoblot.

(B) (Top) Cell proliferation was assessed over a four day time course following treatment with a non-targeting control sgRNA (sgCONT) or two independent non-overlapping

sgRNAs (sgALKBH1#4 and sgALKBH1#5). Significance was determined by two-way repeated measures ANOVA with Dunnett multiple test correction. $P < 0.0001$. (Bottom) Immunoblot showing ALKBH1 protein level following treatment with a non-targeting control sgRNA (sgCONT) and two independent non-overlapping sgRNAs targeting ALKBH1 (sgALKBH1#4 and sgALKBH1#5).

(C) Tumorsphere formation efficiency and self-renewal capacity were measured by extreme in vitro limiting dilution assays (ELDA) in two human patient derived GSC models (387 and D456) after transduction with shCONT or shALKBH1. 387, $p = 7.28e-14$. D456, $p = 4.19e-13$.

(D) Kaplan-Meier curves depict survival of immunocompromised mice bearing intracranial tumors grown from human patient-derived GSC models (387 and D456) following transduction with shCONT or shALKBH1. Significance was determined by log-rank analysis. **, $p < 0.01$. $N = 5$ for each group.

(E) Kaplan-Meier curve depicts survival of immunocompromised mice bearing intracranial tumors grown from human patient-derived GSC models (387) following transduction with single guide RNAs (sgRNA) targeting ALKBH1 or a non-targeting control. Significance was determined by log-rank analysis. **, $p < 0.01$. $N = 5$ for each group.

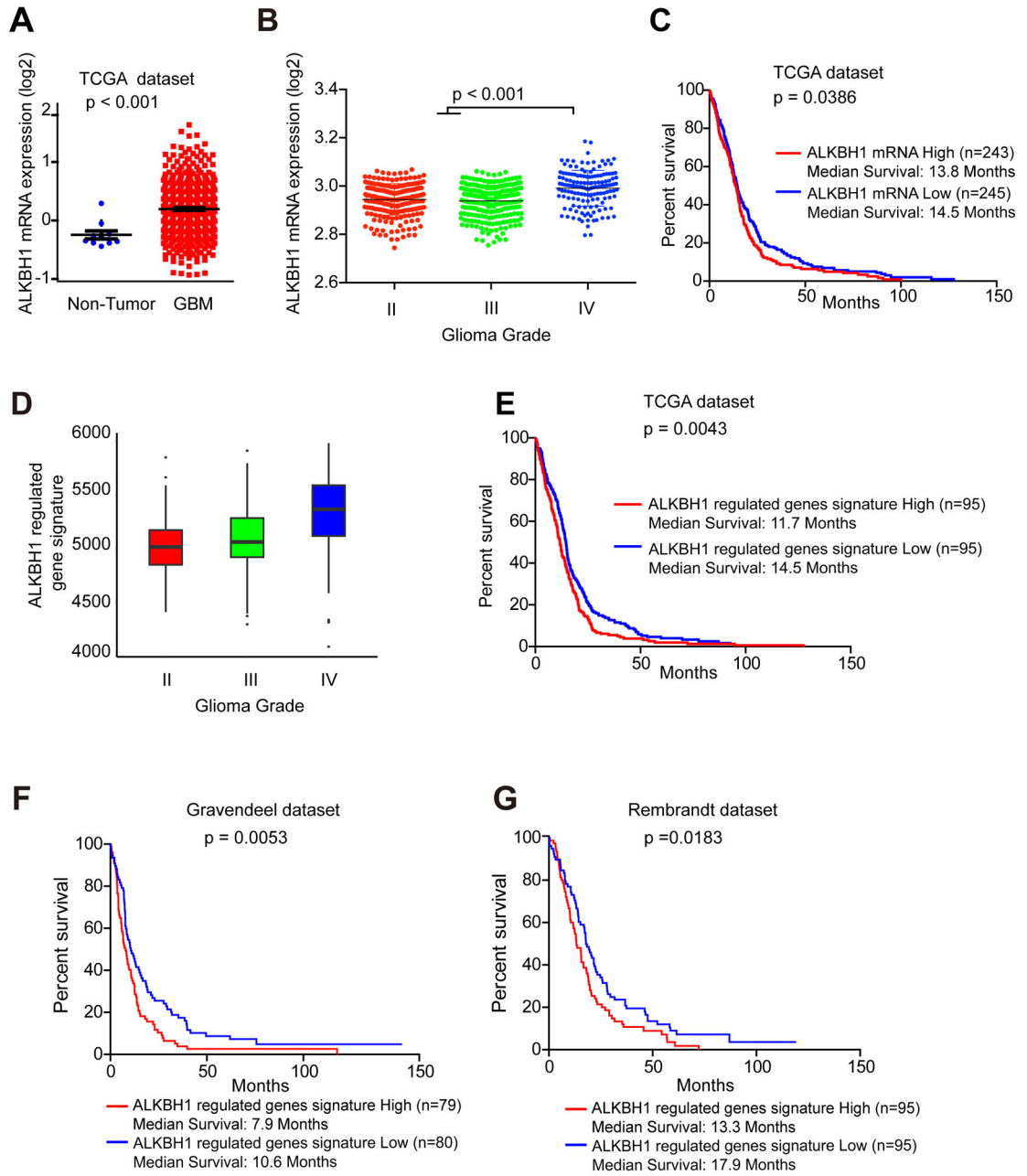


Figure 7. ALKBH1 is associated with poor clinical outcomes in glioblastoma patient datasets. (A) Relative mRNA expression levels of ALKBH1 in non-tumor brain and glioblastoma were determined in The Cancer Genome Atlas dataset. (Student’s t-test, $p < 0.001$) (B) ALKBH1 gene (mRNA) expression was determined for all samples in The Cancer Genome Atlas glioblastoma and low-grade glioma datasets. The mRNA expression is plotted after dividing samples by glioma grade. Ordinary one-way ANOVA with Tukey multiple comparison test was used for statistical analysis, $p < 0.001$ for grade II vs grade IV and grade III vs grade IV comparisons. Grade II: $n = 226$; Grade III: $n = 244$; Grade IV: $n = 150$. (C) Kaplan-Meier curve of patient survival in The Cancer Genome Atlas clinical data set. Patients are stratified by ALKBH1 expression status, with the “ALKBH1 high group”

defined as patients with greater than the median ALKBH1 expression. Significance was determined by log-rank analysis. $P = 0.0386$. Median survival is 13.8 months for the “ALKBH1 mRNA High” group ($n = 243$) and 14.5 months for the “ALKBH1 mRNA Low Group” ($n = 245$).

(D) An ALKBH1 regulated gene expression signature score was calculated for all samples in The Cancer Genome Atlas glioblastoma and low-grade glioma datasets. The signature score is plotted after dividing samples by glioma grade. Significance was determined by one-way ANOVA with tukey multiple test correction. $P < 0.0001$ for Grade II vs. Grade IV; $p < 0.0001$ for Grade III vs. Grade IV; not significant for Grade II vs. Grade III. Grade II: $n = 226$; Grade III: $n = 244$; Grade IV: $n = 150$).

(E) Kaplan-Meier curve of patient survival in The Cancer Genome Atlas clinical data set. Patients are stratified by ALKBH1-regulated gene signature status, with the “High” group defined as patients with greater than the median value of the ALKBH1-regulated gene signature. Significance was determined by log-rank analysis. $P = 0.0043$. Median survival is 11.7 months for the “ALKBH1 signature High” group ($n = 95$) and 14.5 months for the “ALKBH1 signature Low Group” ($n = 95$).

(F) Kaplan-Meier curve of patient survival in the Gravendeel clinical data set. Patients are stratified by ALKBH1-regulated gene signature status, with the “High” group defined as patients with greater than the median value of the ALKBH1-regulated gene signature. (log-rank analysis. $P = 0.0053$)

(G) Kaplan-Meier curve of patient survival in the REMBRANDT clinical data set. Patients are stratified by ALKBH1-regulated gene signature status, with the “High” group defined as patients with greater than the median value of the ALKBH1-regulated gene signature. (log-rank analysis. $P = 0.0183$)