SUPPLEMENTARY DATA

Carella et al., 2019

This document includes the following supplementary information:

- **SUPPLEMENTARY MATERIALS AND METHODS**
- **SUPPLEMENTARY REFERENCES**
- **SUPPLEMENTARY TABLE AND FIGURE LEGENDS**
- **SUPPLEMENTARY TABLES AND FIGURES**

SUPPLEMENTARY MATERIALS AND METHODS

Quantitative RT-PCR

TRIzol® reagent (Invitrogen, 15596018) was used to extract total RNA from both human samples and LN229 cells, according to the manufacturer's instructions. cDNA was synthesized from total RNA (500 ng) using the SuperScript III Reverse Transcriptase Kit (Invitrogen, 18080044), following the manufacturer's recommendations. Quantitative RT-PCR was performed using SYBR Green 2× PCR Master Mix (Applied Biosystems, 4309155) and the corresponding oligonucleotides (listed in **Supplementary Table S2**). qRT-PCR was carried out using the StepOnePlus real-Time PCR System (Applied Biosystems). *GAPDH* was used as house-keeping gene to standardize data, following the ΔΔCt method.

Stable transfections

LN229 cells were stably transfected with protein-expression vectors encoding human TET3 full sequence (GeneCopoeiaTM, EX- I2170-M02), or with empty sequence (GeneCopoeiaTM, EX-EGFP-M11) using Lipofectamine® 2000, according to the manufacturer's recommendations. Selection of stably-transfected clones was performed with 1 mg/mL of G418 disulfate salt (Sigma, A-1720). TET3 expression was verified by qRT-PCR using the specific oligonucleotides listed in **Supplementary Table S2**.

Pyrosequencing assays

DNA methylation and hydroxymethylation patterns of representative CpG sites were analyzed by bisulfite pyrosequencing. Genomic DNA isolation was performed according to standard phenol/chloroform/isoamyl alcohol extraction protocol. Oxidative bisulfite (oxBS) and bisulfite-only (BS) conversion was performed according to the TrueMethyl® Array Kit User Guide (CEGX, Version 2) with some minor modifications. DNA samples (1µg) were cleaned with Agencourt AMPure XP (Beckman Coulter), oxidized with 1 µL of a KRuO4 solution (Alpha Aeser, 375 mM in 0.3 M NaOH), and conversion was performed using Epitect bisulfite kit (Qiagen®). PCR amplification was performed with the specific set of *TET3* oligonucleotides (**Supplementary Table S2**) designed with the specific software PyroMark assay design (version 2.0.01.15) in order to analyze cg11236515, cg03763077 and cg15996154 CpGs. After PCR amplification of oxBS and BS samples, we performed bisulfite pyrosequencing with PyroMark Q24 reagents, equipment and software (Qiagen) and the Vacuum Prep Tool (Biotage, Uppsala, Sweden),

following the respective manufacturers' instructions. 5mC values were obtained from oxBS methylation levels. 5hmC levels were obtained upon subtraction of oxBS methylation values from their corresponding BS treated pair.

DNA methylation arrays

Oxidative bisulfite (oxBS) and bisulfite-only (BS) conversion of empty or stably-transfected TET3 clones DNA was performed according to the protocols previously described for the pyrosequencing assay. DNA samples were then hybridized to the Illumina HumanMethylationEPIC BeadChip array platform following the Illumina Infinium HD methylation protocol at the Spanish Centro Nacional de Genotipado (CEGEN-ISCIII, www.cegen.org). Raw HumanMethylation450 data from brain and glioma samples were obtained from arrayexpress (E-MTAB-6003) or GEO (GSE73895)[1,2]. IDAT files were processed using the R/Bioconductor package *minfi* (version 1.22.1) and probe signals were corrected using the SWAN algorithm [3]. Background correction or control probe normalization was not applied. Probes located in chromosomes X and Y, probes overlapping genetic variants (SNP137Common track from UCSC genome browser), and crossreactive and multimapping probes were discarded for downstream analyses. Estimations of 5mC and 5hmC levels for a given CpG were calculated with the R/bioconductor package *ENmix* (version 1.12.3) using the oxBS.MLE method [4].

Whole genome bisulfite sequencing (WGBS)

Processed WGBS data corresponding to oxBS or BS conversion of normal brain or glioma samples were obtained from Arrayexpress (E-MTAB-5171)[5]. The WGBS data was subsequently filtered to highlight the whole CpG coverage along the *TET3* gene. Estimations of 5mC and 5hmC levels for the corresponding CpGs were calculated with the R/bioconductor package *ENmix* (version 1.12.3) by means of the oxBS.MLE method.

5-hydroxymethylcytosine immunoprecipitation- qPCR assay

Immunoprecipitation of 5hmC was performed using the EpiQuik Hydroxymethylated DNA Immunoprecipitation (hMeDIP) Kit (Epigentek), according to the manufacturer's recommendations. Input, non-specific IgG- and 5 hmC-enriched fractions were obtained from five control brains and five

glioma tumors. All fractions were analyzed by RT-qPCR with oligonucleotides specific to the cg03763077 region (listed in **Supplementary Table S2**). After confirming there were no significant differences between input DNAs, 5hmC relative enrichment was calculated as a Fold Change relative to Input Ct Mean.

Chromatin immunoprecipitation assay

Chromatin immunoprecipitation was performed with control brain and tumor samples as described in the "Chromatin preparation from tissues for chromatin immunoprecipitation (ChIP)" protocol by Abcam, with some minor modifications. Frozen tissue was fixed for 10 min with 1-1.5% formaldehyde solution. Crosslinking was stopped by the addition of glycine to 125 mM. Homogenization was performed on ice using an Eppendorf Micropestle which fitted tightly into the bottom of the Eppendorf tube. Samples were resuspended in Lysis Buffer (50 mM HEPES-KOH pH7.5;140 mM NaCl; 1 mM EDTA pH8; 1% Triton X-100; 0.1% Sodium Deoxycholate; 0.1% SDS; Protease Inhibitors) and sonicated using a Diagenode's Bioruptor® Sonicator to obtain chromatin fragments of 200–500 bp in length.

Chromatin immunoprecipitation of LN229 human glioblastoma cell line treated with SAHA was performed following a standard ChIP protocol. In brief, after treatment, cells were fixed with 1% formaldehyde (Sigma-Aldrich), lysed in sodium dodecyl sulfate (SDS) lysis buffer (1% SDS, 10 mM EDTA, and 50 mM Tris–HCl pH 8.1), and sonicated.

Immunoprecipitations were performed, for both protocols, overnight using antibodies against H4K16ac (Active Motif, 39167) and total histone H3 (Abcam, ab1791) as positive control, or IgG antiserum (Abcam, ab46540) as negative control. Antibody–chromatin complexes were precipitated with Salmon Sperm DNA/Protein A-Agarose beads (Upstate Biotechnologies), then washed and eluted from beads using the corresponding buffer (1% SDS, 100mM NaHCO3). DNA was extracted with phenol–chloroform and was ethanol-precipitated. Immunoprecipitated DNA was analyzed by qRT-PCR. The experiment was performed for TET3 specific regions using SYBR® Green in the StepOnePlus™ Real-Time PCR System (Applied Biosystems) and the corresponding oligonucleotides (listed in **Supplementary Table S2**). All measurements were performed in triplicate and a calibration curve was determined for the primer set.

ChIP-sequencing analyses

BigWig imputed signal tracks incorporating information for 6 histone posttranslational modifications (H3K4me1, H3K4me3, H3K9ac, H3K27ac, H3K36me3 and H3K9me3) corresponding to Brain Inferior Temporal Lobe and Brain Dorsolateral Prefrontal Cortex were obtained from the NIH Roadmap consortium (http://egg2.wustl.edu/roadmap/). BigWig tracks corresponding to the same set of histone marks on two glioblastoma samples were obtained from GSE113816[6]. Normalization of track signals was performed by using the following pipeline: First, BigWig files were converted to Bedgraphs using the UCSC Genome Browser tool *bigWigToBedGraph*. Second, for each Bedgraph file, the total signal was calculated as the sum of the signals obtained in every genomic position detected. Third, every genomic position of the aforementioned Bedgraph files was normalized against its corresponding total signal and the resulting Bedgraphs were converted to Bigwig format using the tool *bedGraphToBigWig* from UCSC genome browser.

High-performance capillary electrophoresis (HPCE) quantification of global histone H4 acetylation.

Histone acid extraction (sulfuric acid plus acetone precipitation) was carried out according to the procedure of Sarg and colleagues [7], with some minor modifications. Global histone H4 acetylation (AcH4) was quantified as previously described [8]. In brief, the histone H4 fraction was purified by reverse-phase high-performance liquid chromatography (HPLC) using a Jupiter C18 column (Phenomenex, Inc.) and eluted with an acetonitrile gradient (20-60%) in 0.3% trifluroacetic acid using an HPLC gradient system (Beckman Coulter). Non-, mono-, di-, tri- and tetra-acetylated histone H4 derivatives were resolved by HPCE. An uncoated fused-silica capillary (Beckman-Coulter; 60.2 cm x 75 mm, effective length 50 cm) was used in a capillary electrophoresis system (P/ACE MDQ, Beckman-Coulter) with 100 mM phosphate buffer (pH 2.0) 0.02% (w/v) HPM-cellulose as running buffer and operating voltages of 12 kV.

Immunohistochemistry

Immunohistochemistry was performed with the EnVision FLEX Mini Kit (DAKO, K8024) and Dako Autostainer system. Paraffin-embedded tissues (2.5-5µm) were deparaffinized, rehydrated and subjected to heat induction epitope retrieval (HIER) at 95ºC and pH 9 for 20 min (DAKO, GV804) in the Pre-Treatment Module, PT-LINK (DAKO). Sections were incubated with TET3 antibody (NovusBiological,

NBP2-13427) diluted in EnVision™ FLEX Antibody Diluent (K8006) (1:100 dilution) for 30 min following the blockage of endogenous peroxidase with EnVision™ FLEX Peroxidase-Blocking Reagent (DM821). Signal was detected using diaminobenzidine chromogen as substrate (DM823) in EnVision™ FLEX Substrate Buffer (DM823) after incubation with Dako EnVision™ FLEX /HRP (DM822). Sections were counterstained with hematoxylin, dehydrated and mounted with permanent medium (Dako mounting medium, CS703).

Western blotting

Proteins were separated on a 12% SDS-PAGE gel and transferred to a polyvinylidene difluoride (PVDF) membrane (Immobilon PSQ, Millipore). The membrane was blocked in phosphate buffered saline buffer with 0.1% Tween 20 (PBS-T) complemented with 10% milk. For the detection of protein specific bands, the following primary antibodies were used at 1:1000 dilution in TBS-T with 5% milk: anti-TET3 (GeneTex, GTX121453), anti-Vimentin (Abcam, ab16700), anti-SOX2 (Merck, ab5603), Anti- ß Tubulin III (Abcam, ab18207) and anti-H3 (Abcam, ab1791). The secondary antibodies used were either anti-rabbit IgG (Amersham Biosciences, NA934) or anti-mouse IgG (Santa Cruz Biotechnology, sc-2005) conjugated to horseradish peroxidase at 1:5000 dilution in TBS-T with 5% milk. Signals were detected using the ECL detection kit (Amersham Biosciences) and Oddysey Fc imaging system, and subsequent bands were quantified with the ImageJ software (version 1.46r).

5-hydroxymethylcytosine Dot blotting

5hmC levels were interrogated on DNA isolated from LN229 cells samples following the recommendations from Diagenode's dot blot protocol. Briefly, samples were resuspended in 0.1 M NaOH and denatured at 99ºC for 5 minutes. After a brief cool down on ice, samples were spun down and neutralized with 0.1 vol of 6.6 M ammonium acetate, and were immediately spotted on nitrocellulose membranes (Amersham Hybond N+). DNA was then UV-crosslinked to the membrane, and the membrane was incubated for 1h with blocking solution (10% milk in TBS-Tween20 0.1%). Next, the membrane was incubated with anti-5hmC rabbit monoclonal antibody (ActiveMotif, ref 39769, 1/500 dilution) for a maximum of 1 hour in blocking solution at room temperature. The membrane was washed 3 times at room temperature with PBS-Tween20 0.1% for a total of 15 minutes and then incubated with

HRP-conjugated secondary antibody (Amersham Biosciences, NA934, 1/5.000 dilution in blocking Solution) for 60 minutes maximum at room temperature. After this, the membrane was washed 3 times with TBS-Tween and the signal was detected with ECL Plus Western Blotting Reagent Pack (GE Healthcare) and Oddysey Fc imaging system. Subsequent signals were quantified with the ImageJ software (version 1.46r).

Immunofluorescence

75,000-80,000 cells of empty and TET3 stably-transfected clones were seeded (for each condition). Cells were fixed with 99.8% Methanol (5 min, room temperature) and incubated for 1 hour at room temperature (under dark conditions) with ß Tubulin III, Vimentin and SOX2 antibody (Anti- ß Tubulin III, Abcam, ab18207; Anti-Vimentin, Abcam, ab16700; Anti-SOX2, Merck, ab5603). Cells were washed with PBS 1x and incubated with Alexa-488-conjugated secondary antibody (Life Technologies, A21441). DNA was stained with DAPI (4', 6-diamidino-2-phenylidole) and images of the immunofluorescence preparations were obtained using a digital camera (AxioCam MRm,Carl Zeiss) connected to a Zeiss AxioObserver Z1 microscope (Carl Zeiss, Germany), using a Plan-Apochromat 63X/1.4 oil lens objective (NA=1.4, working distance=0.19mm).

Colony formation assays

Colony formation assays were initially conducted by plating cells at different densities per well (100, 200, and 400 cells). At day 14, stable colonies were fixed and stained with a mixture of 6.0% Glutaraldehyde (Sigma, 340855) and 0.5% Crystal Violet (Sigma, C3886). After removing the excess dye, number of colonies was counted. Plating efficiency mean represents the ratio of the number of colonies versus the number of cells seeded at different cell dilutions. The assay was performed as described by Franken and colleagues [9].

Neural Differentiation profiling

The differentiation status of TET3 transfected LN229 cell clones was interrogated by using the Neural qPCR profiler with RNA-Quant kit (System Biosciences, RA500A-1). RNA extraction was performed as described in the section on Quantitative RT-PCR assay above. RT and Real-Time qPCR reactions were performed following the manufacturer's recommendations. qRT-PCR was carried out using the

StepOnePlus real-Time PCR System (Applied Biosystems) and *GAPDH* was used as house-keeping gene to standardize data, following the ΔΔCt method.

Region set enrichment analyses

Assignment of differentially hydroxymethylated CpG sites (dhmCpGs) to their corresponding CG Island (CGI) status was performed with the R/Bioconductor package *IlluminaHumanMethylation450anno.ilmn12.hg19* (version 0.6.0). Enrichment of genomic positions was performed with the R/Bioconductor package *TxDb.Hsapiens.UCSC.hg19.knownGene* (version 3.2.2) and *ChIPseeker* (version 1.12.0) [10]. The statistical significance of CGI status and genomic region enrichment were calculated using a two-sided Fisher's tests. Odds Ratios (ORs) were used to indicate the magnitude of the association effect with respect to a particular feature. Only probes common to both the HumanMethylation450 and the HumanMethylationEPIC platform were used for legitimate statistical purposes, along with a background probe set which included all the probes interrogated by the HumanMethylation450 Beadchip.

Chromatin enrichment analyses were conducted with the R/Bioconductor package *LOLA* (version 1.4.0)[11]. Histone mark enrichments (H3K4me1, H3K4me3, H3K27me3, H3K36me3, H3K9me3 and H3K27ac) were calculated using ChIP-seq tracks from 6 Embryonic Stem Cell lines and 7 brain tissue samples obtained from ENCODE and the NIH Roadmap Epigenome consortia [12,13], and further integrated in the *LOLA* extended software environment (data obtained from http://databio.org/regiondb). Chromatin-segmentation data for the same cells/tissue types was obtained from NIH Roadmap's ChromHMM expanded 18-state model (http://egg2.wustl.edu/roadmap/). Enrichments were calculated by performing one-sided Fisher's tests (significance threshold value <0.05) comparing the overlap of dhmCpGs with the dataset of interest.

Gene, KEGG and Reactome ontology analyses

Gene (GO), KEGG (Kyoto Encyclopedia of Genes and Genomes) and Reactome pathway ontology analyses were conducted using the R/Bioconductor package c*lusterProfiler* (version 3.4.2) [14]. Genes with dhmCpGs in brain and TET3 overexpressing LN229 cells were used to interrogate the KEGG

(http://www.kegg.jp/kegg/rest/keggapi.html), GO (R/Bioconductor package *GO.db*, version 3.4.1) and Reactome (R/Bioconductor package *ReactomePA*, version 1.20.2) annotation databases for gene ontology purposes. Total number of genes present in each array platform was used as the background set for legitimate ontology comparisons.

Gene expression arrays

Raw CEL files corresponding to normal brain tissue or TCGA-GBM datasets were obtained from the NCBI Gene Expression Omnibus Database (GEO) under the accession numbers GSE14805 and GSE83130 respectively [15,16]. Raw CEL files from normal brain, purified astrocytes, purified microglia, tumor samples and LN229 cells were obtained from ArrayExpress under the accession numbers E-GEOD-36634, E-GEOD-15824 and E-GEOD-23806[17–19]. Further information regarding clinical/pathological status and accession numbers of these particular samples is provided in **Supplementary Table S5**. Raw CEL files from PBT003 cells transduced with empty control RNA (GSM1970502, GSM1970504, GSM1970506, GSM1970507), TLX shRNA (GSM1970503, GSM1970505) or TET3 shRNA-expressing lentivirus (GSM1970508, GSM1970509, GSM1970510, GSM1970511) were obtained from GEO under the accession number GSE75945 [20]. Microarray analyses were conducted using the R/Bioconductor package *affy* (version 1.54.0). Raw data were further normalized by the robust multi-array averaging method (RMA). Differentially expressed genes in knockdowns or their corresponding control conditions were calculated with the R/Bioconductor package *limma* (version 3.32.2) [21] and fitted using a linear regression model including probe expression signals as response variable, and the variable of interest (empty or knockdown conditions). Statistical significance was calculated by means of an empirical Bayes moderated t-test and multiple testing corrections were applied using the Benjamini-Hochberg method to control for false discovery rate (p-value < 0.003 for TLXKD, or FDR < 0.05 for TET3KD, respectively). In the case of normal and brain cancer samples, datasets were obtained from Auvergne and colleagues [22] (GSE29796) and array data was normalized with the RMA method. A matrix including normalized log2 expression values was used for representation purposes.

RNAseq data

RSEM normalized RNAseq data corresponding to GBM samples were obtained from TCGA consortium [16]

via UCSC Xena Public Data Hub (http://xena.ucsc.edu/)(version 2017-10-13). Fastq files corresponding to 12 purified glia samples and 4 LN229 cells RNA sequencing experiments were obtained from the SRA studies SRP113619, SRP173011 and SRP095447. Further details of these experiments and the samples analyzed in our pipeline are described in **Supplementary Table S5**. Quality control of the sequenced reads was performed with the *FastQC* software (version 0.11.7) and adapter removal was performed with *Trim Galore* (version 0.4.1). Reads were mapped to the reference human genome assembly GRCh37 using *RSEM* (version 1.3.1). Finally, RPKM data relating to 114 brain cortex and 125 cerebellum tissue samples was obtained from the GTEx consortium [23] (www.gtexportal.org). A robust correlation algorithm from the R/CRAN package *robcor* (version 0.1-6, "ssd" method) was used for further statistical comparisons between *TET3* and *RECK* or *WBP1L* expression data.

Survival analyses

Gene expression datasets corresponding to high grade GBM patients were obtained from GEO under the accession number GSE4412 [24]. Kaplan-Meier curves were generated with the PROGgeneV2 platform [25] (http://watson.compbio.iupui.edu/chirayu/proggene/database). Sample stratification criteria towards low or high gene expression groups were selected based on the median expression value of each particular gene across patients. p-values indicated significant differences in event rates between Kaplan-Meier curves and were calculated with the Log-rank test function.

Statistical analyses

Group comparisons were performed by fitting a general linear model to each dataset. Contrasts were defined to test the different hypotheses (comparisons between levels of the main group variable) on each of the models. Repeated measures designs, involving the measurement of a specified quantity at different time points, were modeled using a general linear model, with a main group predictor and a blocking variable describing the time of observation. Time was modeled as an ordinal variable and contrasts were defined based on the main group variable. Correlation between variables was assessed using the Spearman rank correlation coefficient. The resulting p-values were adjusted using the Bonferroni method for controlling the Family-Wise Error Rate (FWER). Goodness-of-fit was assessed for

each of the models using graphical methods and the Shapiro-Wilk test on the residuals. A Wilcoxon rank-sum (group comparisons) or signed-rank (comparisons across time) test was used for each pairwise comparison if the residuals showed strong deviations from normality and there were sufficient observations to apply the test. All statistical analyses were conducted using the R statistical programming language (version 3.4.0).

SUPPLEMENTARY REFERENCES

1. Fernandez AF, Bayón GF, Sierra MI, Urdinguio RG, Toraño EG, García MG, Carella A, López V, Santamarina P, Pérez RF, Belmonte T, Tejedor JR, et al. Loss of 5hmC identifies a new type of aberrant DNA hypermethylation in glioma. *Hum Mol Genet* 2018;27:3046–59.

2. Johnson KC, Houseman EA, King JE, von Herrmann KM, Fadul CE, Christensen BC. 5-Hydroxymethylcytosine localizes to enhancer elements and is associated with survival in glioblastoma patients. *Nat Commun* 2016;7:13177.

3. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol* 2012;13:R44.

4. Xu Z, Taylor JA, Leung Y-K, Ho S-M, Niu L. oxBS-MLE: an efficient method to estimate 5-methylcytosine and 5-hydroxymethylcytosine in paired bisulfite and oxidative bisulfite treated DNA. *Bioinforma Oxf Engl* 2016;32:3667–9.

5. Raiber E-A, Beraldi D, Martínez Cuesta S, McInroy GR, Kingsbury Z, Becq J, James T, Lopes M, Allinson K, Field S, Humphray S, Santarius T, et al. Base resolution maps reveal the importance of 5-hydroxymethylcytosine in a human glioblastoma. *NPJ Genomic Med* 2017;2:6.

6. Chen X, Zhang M, Gan H, Wang H, Lee J-H, Fang D, Kitange GJ, He L, Hu Z, Parney IF, Meyer FB, Giannini C, et al. A novel enhancer regulates MGMT expression and promotes temozolomide resistance in glioblastoma. *Nat Commun* 2018;9:2949.

7. Sarg B, Koutzamani E, Helliger W, Rundquist I, Lindner HH. Postsynthetic trimethylation of histone H4 at lysine 20 in mammalian tissues is associated with aging. *J Biol Chem* 2002;277:39195–201.

8. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Pérez-Rosado A, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* 2005;37:391–400.

9. Franken NAP, Rodermond HM, Stap J, Haveman J, van Bree C. Clonogenic assay of cells in vitro. *Nat Protoc* 2006;1:2315–9.

10. Yu G, Wang L-G, He Q-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinforma Oxf Engl* 2015;31:2382–3.

11. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinforma Oxf Engl* 2016;32:587–9.

12. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57–74.

13. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, et al. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518:317–30.

14. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J Integr Biol* 2012;16:284–7.

15. Hodgson JG, Yeh R-F, Ray A, Wang NJ, Smirnov I, Yu M, Hariono S, Silber J, Feiler HS, Gray JW, Spellman PT, Vandenberg SR, et al. Comparative analyses of gene copy number and mRNA expression in glioblastoma multiforme tumors and xenografts. *Neuro-Oncol* 2009;11:477–87.

16. Brennan CW, Verhaak RGW, McKenna A, Campos B, Noushmehr H, Salama SR, Zheng S, Chakravarty D, Sanborn JZ, Berman SH, Beroukhim R, Bernard B, et al. The somatic genomic landscape of glioblastoma. *Cell* 2013;155:462–77.

17. Sim FJ, Windrem MS, Goldman SA. Fate determination of adult human glial progenitor cells. *Neuron Glia Biol* 2009;5:45–55.

18. Grzmil M, Morin P, Lino MM, Merlo A, Frank S, Wang Y, Moncayo G, Hemmings BA. MAP kinase-interacting kinase 1 regulates SMAD2-dependent TGF-β signaling pathway in human glioblastoma. *Cancer Res* 2011;71:2392–402.

19. Günther HS, Schmidt NO, Phillips HS, Kemming D, Kharbanda S, Soriano R, Modrusan Z, Meissner H, Westphal M, Lamszus K. Glioblastoma-derived stem cell-enriched cultures form distinct subgroups according to molecular and phenotypic criteria. *Oncogene* 2008;27:2897–909.

20. Cui Q, Yang S, Ye P, Tian E, Sun G, Zhou J, Sun G, Liu X, Chen C, Murai K, Zhao C, Azizian KT, et al. Downregulation of TLX induces TET3 expression and inhibits glioblastoma stem cell self-renewal and tumorigenesis. *Nat Commun* 2016;7:10637.

21. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.

22. Auvergne RM, Sim FJ, Wang S, Chandler-Militello D, Burch J, Al Fanek Y, Davis D, Benraiss A, Walter K, Achanta P, Johnson M, Quinones-Hinojosa A, et al. Transcriptional differences between normal and glioma-derived glial progenitor cells identify a core set of dysregulated genes. *Cell Rep* 2013;3:2127–41.

23. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015;348:648–60.

24. Freije WA, Castro-Vargas FE, Fang Z, Horvath S, Cloughesy T, Liau LM, Mischel PS, Nelson SF. Gene expression profiling of gliomas strongly predicts survival. *Cancer Res* 2004;64:6503–10.

25. Goswami CP, Nakshatri H. PROGgene: gene expression based survival analysis web application for multiple cancers. *J Clin Bioinforma* 2013;3:22.

SUPPLEMENTARY TABLE AND FIGURE LEGENDS

Supplementary Table S1 – Clinico/pathological characteristics of normal brain and tumor samples used for DNA methylation analyses and pyrosequencing experiments.

Supplementary Table S2 – Oligonucleotide sequences used for quantitative RT-PCR, pyrosequencing, ChIP real-time PCR, hMeDIP and DNA sequencing experiments.

Supplementary Table S3 – Estimations of 5mC and 5hmC levels corresponding to the CpG probes located along the *TET* genes in the HumanMethylation450K array platform and WGBS data.

Supplementary Table S4 – *TET3* gene expression data corresponding to gene expression arrays (HT_HG-U133A) and RNAseq experiments (IlluminaHiSeq_RNASeqV2) from the TCGA-GBM cohort. Information regarding the CIMP and experimental classification of these samples is also included for grouping purposes.

Supplementary Table S5 – Gene expression data corresponding to gene expression arrays (HG-U133_Plus_2) and RNAseq experiments for TET and HPRT1 genes in normal brain, tumor samples and LN229 cells. Information regarding the clinical variables and experimental classification of these samples is included in additional columns.

Supplementary Table S6 – List of differentially methylated or hydroxymethylated CpG sites in brain versus glioma comparisons. Columns indicate genomic coordinates as well as related information including relation to CpG Island, associated gene name, associated region type and differences in methylation between brain and glioma samples for each significant CpG site (cutoff > 0.15). Raw data obtained from E-MTAB-6003.

Supplementary Table S7 – List of differentially methylated or hydroxymethylated CpG sites in TET3 overexpressing LN229 cells versus control counterparts. Columns indicate genomic coordinates as well as related information including relation to CpG Island, associated gene name, associated region type and

differences in methylation between TET3 transfected clones and control cells for each significant CpG site (cutoff > 0.25). Raw data deposited in E-MTAB-6001.

Supplementary Table S8 – List of differentially expressed genes in human primary glioblastoma stem cell lines upon TET3KD or TLX1KD conditions. Columns indicate the extent of the fold change and a series of related statistics including the associated gene name and the direction of the change for each significant probe. Raw data obtained from GSE75945.

Supplementary Figure S1 – **DNA methylation profiling along *TET* genes in human brain and glioma samples.** (**A** to **F**) Overall 5mC and 5hmC levels observed for *TET* genes in brain and glioma samples obtained from two different HumanMethylation450K datasets. (**G-I**) High content profiling of 5mC and 5hmC obtained from WGBS experiment. Accession numbers of the corresponding datasets are indicated, and all CpG probes identified in these analyses are represented. To facilitate further interpretation purposes, methylation array probes are mapped to their corresponding genomic coordinates.

Supplementary Figure S2 – **Biological validation in an independent cohort of samples confirms that loss of *TET3* 5hmC and gain of 5mC are frequent phenomena in brain tumor.** (**A** to **C**) Box plots represent methylation and hydroxymethylation levels of CpGs cg11236515, cg03763077 and cg15996154 as determined by pyrosequencing experiments. For cg03763077 we analyzed, in duplicate, 12 control brain (six grey matter and six white matter) and 10 glioblastoma samples. For cg11236515 and cg15996154 we validated, in duplicate, 12 brain and 12 glioma samples. (**D** to **E**) Box plots indicate methylation and hydroxymethylation levels of the aforementioned CpG sites in white and grey matter or glioma samples. Wilcoxon rank sum tests were applied and p-values were adjusted by applying the Bonferroni correction. ***: p-value<0.001. (**G**) Additional validation of 5hmC loss in glioma by immunoprecipitation and qRT-PCR at the cg03763077 position in *TET3*. A total of 5 brain samples and 5 gliomas were analyzed (Mann Whitney test, *:p-value < 0.05).

Supplementary Figure S3 – Epigenetic profiling of *TET* genes in healthy brain tissue and glioblastoma multiforme. (**A** to **C**) BigWig tracks reflect the normalized signal of 6 histone marks along the genomic

coordinates of *TET1, TET2,* and *TET3* genes obtained from Brain Inferior Temporal Lobe and Brain Dorsolateral Prefrontal Cortex samples, and two glioblastoma samples generated from patient derived xenografts. For further interpretation purposes, their corresponding genomic coordinates are indicated below the charts. (**D**) BigWig tracks corresponding to HPRT1 gene, which is included as a control.

Supplementary Figure S4 – Computational validation using TCGA-GBM data indicates that TET3 levels are reduced in human glioblastoma. (**A**) *TET3* gene expression data in normal glia tissue and TCGA-GBM samples. Data values were obtained from gene expression array experiments. (**B** and **C**) Deconvolution of the array TCGA-GBM samples based on their CIMP status (**B**) or their gene expression signatures (**C**). (**D**) *TET3* gene expression data in normal brain and TCGA-GBM samples. Data values were obtained from RNAseq experiments. (**E** and **F**) Deconvolution of the previous TCGA-GBM samples based on their CIMP status (**E**) or their gene expression signatures (**F**). **(G** and **H)** Survival curves related to DNA methylation (**G**) and gene expression (**H**) clusters adapted from [16]. p-value refers to differences in event rates between the Kaplan-Meier curves and was calculated with the Log-rank test function. For all samples, data values and group classifications are listed in **Supplementary Table S4**.

Supplementary Figure S5 – **Expression of *TET* genes is compromised in glioma samples.** (**A**) Barplots depicting the normalized expression levels of *TET* genes and HPRT1 in a panel of normal brain samples, including white matter, cortex, purified astrocytes and purified microglia (blue tones), and a set of tumor samples and glioblastoma LN229 cells, comprising Oligodendrogliomas, Astrocytomas, Glioblastoma multiforme and the aforementioned cell line (red/purple tones) obtained from gene expression arrays. (**B**) Line bars indicate the mean expression levels of the aforementioned groups. Each line is colored according to the color code displayed in panel A. (C) Barplots representing normalized expression levels (TPM) of *TET* genes and HPRT1 in a cohort of purified glia samples (blue) or 4 LN229 datasets (purple) obtained from publicly available data. Additional information regarding sample sources, data values and group classifications are provided in **Supplementary Table S5**.

Supplementary Figure S6 – **TET3 protein expression is diminished in glioblastoma samples.** (**A**) Additional images of white and grey matter immunohistochemical staining with TET3 antibody. White

matter presents positive TET3 nuclear staining in contrast with tumoral sections, which essentially display background levels of TET3. Magnification: 20×.

Supplementary Figure S7 – Effectiveness of epigenetic drug treatments in LN229 cells. (**A**) Dot blots indicate the extent of 5hmC signal (left) or the signal of the loaded DNA (right) in control cells, cells treated with Vitamin C (50 ug/ml) for 72 h, or cell clones transfected with TET expression vector. (**B**) Fold enrichment quantification of the dot blot signals upon normalization versus the DNA loading control and over control cells (Student t-test, ***:p-value < 0.001). (C) Chromatogram indicating the increase in total H4ac+ signal in LN229 cells upon SAHA treatment. (**D**) Bar graph depicting global histone H4 acetylation analyses observed upon SAHA treatment of LN229 cells as estimated by HPLC/HPCE in three independent experiments (Mann Whitney test, ***:p-value < 0.001).

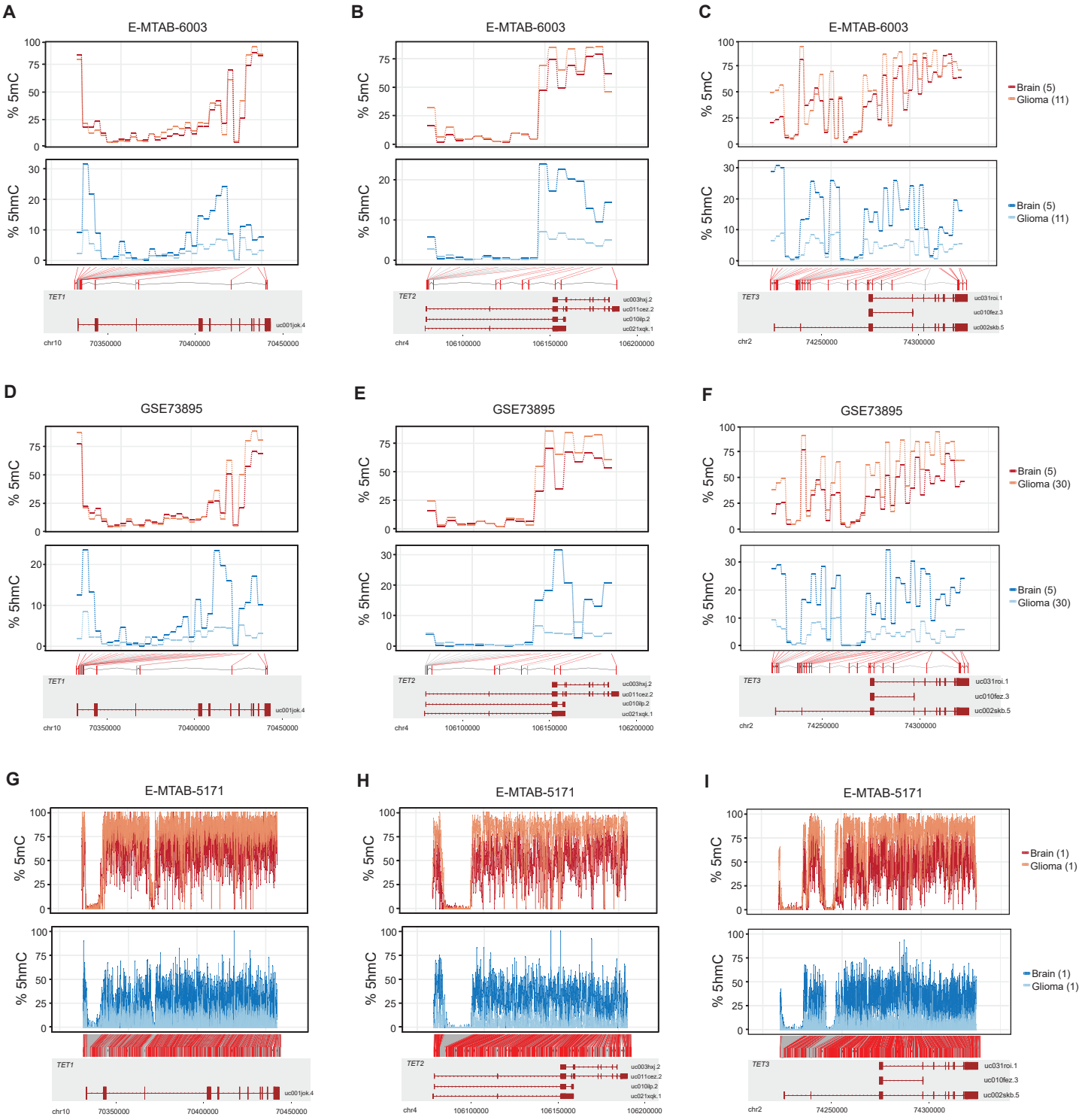
Supplementary Figure S8 – Characterization of the differentiation status of stably-transfected LN229 cells. (**A**) Expression levels of TET3 and β-Tubulin proteins obtained by western blot analyses in TET3 overexpressing LN229 cells or control conditions. (**B**) Fold enrichment quantification of TET3 levels. β-Tubulin was used as loading control. (**C**) Barplot indicates the results of the qPCR neural profiler experiment. Genes are colored according to their relationship with the differentiation categories indicated on the right. Only genes with a significant threshold (+/- 0.5 log2FC) were selected for visualization purposes. (**D**) Expression levels of Vimentin, SOX2 and H3 proteins obtained by western blot analyses in LN229 cells exposed to control or TET3 overexpressing conditions. (**E**) Fold enrichment quantification of protein bands (Vimentin or SOX2 normalized versus H3) for the aforementioned conditions. For all the western blots, quantifications were obtained with the ImageJ software. (Student t-test, ***:p-value < 0.001, *:p-value < 0.05).

Supplementary Figure S9 – Chromatin signatures of 5hmC in brain and TET3 transfected cells. (**A** and **B**) Histone mark enrichment analyses of 5hmC in brain and TET3 transfected cells. Heatmap indicates OR of significant dhmCpGs in brain (**A**) and in TET3 overexpressing LN229 cells (**B**) across six common histone modifications from the NIH Roadmap Epigenome consortium. Brain tissues and ES cells selected for the comparisons are displayed on the right side of the plot. (**C** and **D**) Chromatin state enrichment analyses of 5hmC in brain and TET3 transfected LN229 cells. Chromatin segmentation data for the given

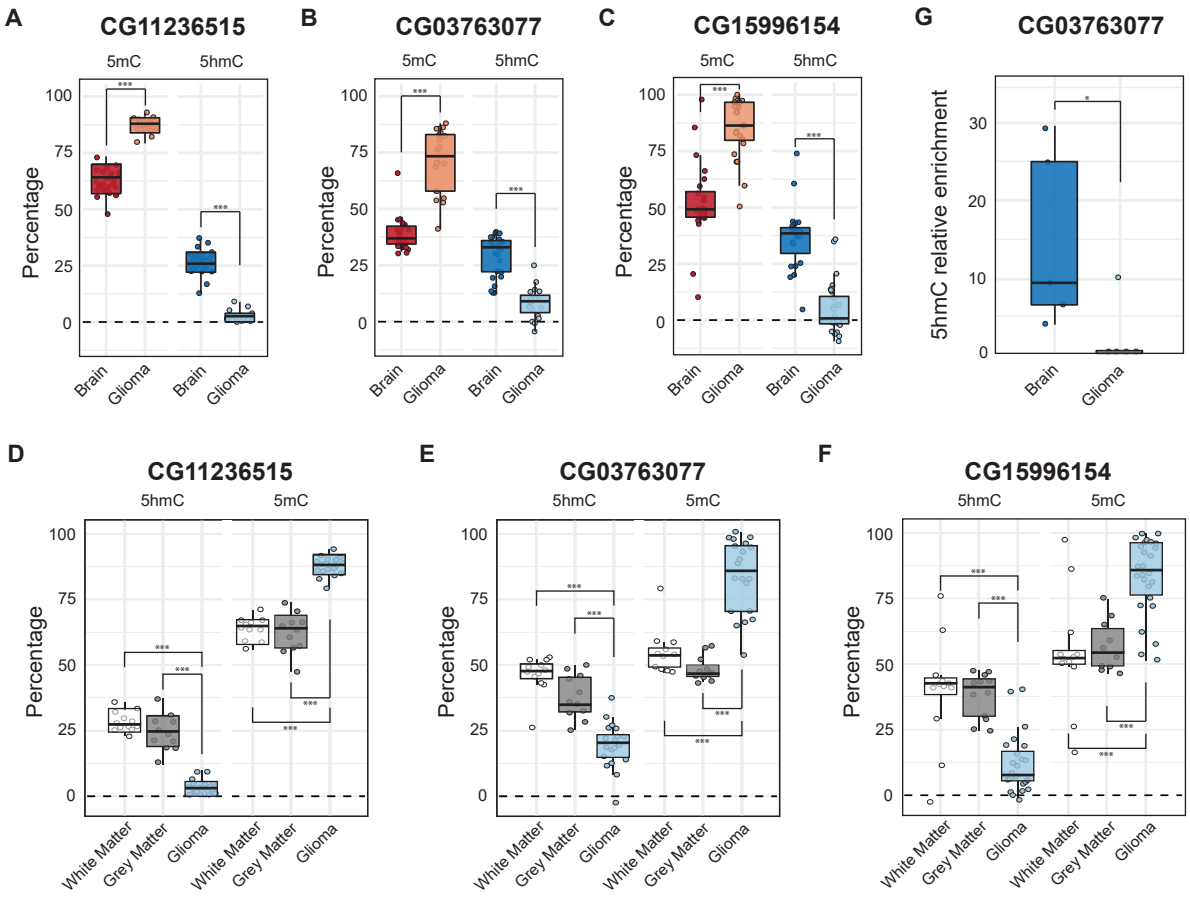
tissues/cell lines was obtained from the NIH Roadmap Epigenome ChromHMM expanded 18-state model. Heatmap depicts OR enrichment of significant dhmCpGs in brain (**C**) or TET3 transfected cells (**D**) across 18 chromatin states for the same brain tissues/ES cells as the previous histone mark analysis. For all graphs, color intensity indicates the degree of the significant OR enrichments.

Supplementary Figure S10 – **Partial restoration of 5hmC levels of *RECK* and *WBP1L* genes upon TET3 overexpression.** (**A** and **B**) Overall hydroxymethylation levels observed for *RECK* and *WBP1L* genes in brain and glioma (top) or in TET3 transfected or control LN229 cells (bottom). All CpG probes located in the previously mentioned candidate genes are represented. To facilitate further interpretation purposes, methylation array probes are mapped to their corresponding genomic coordinates.

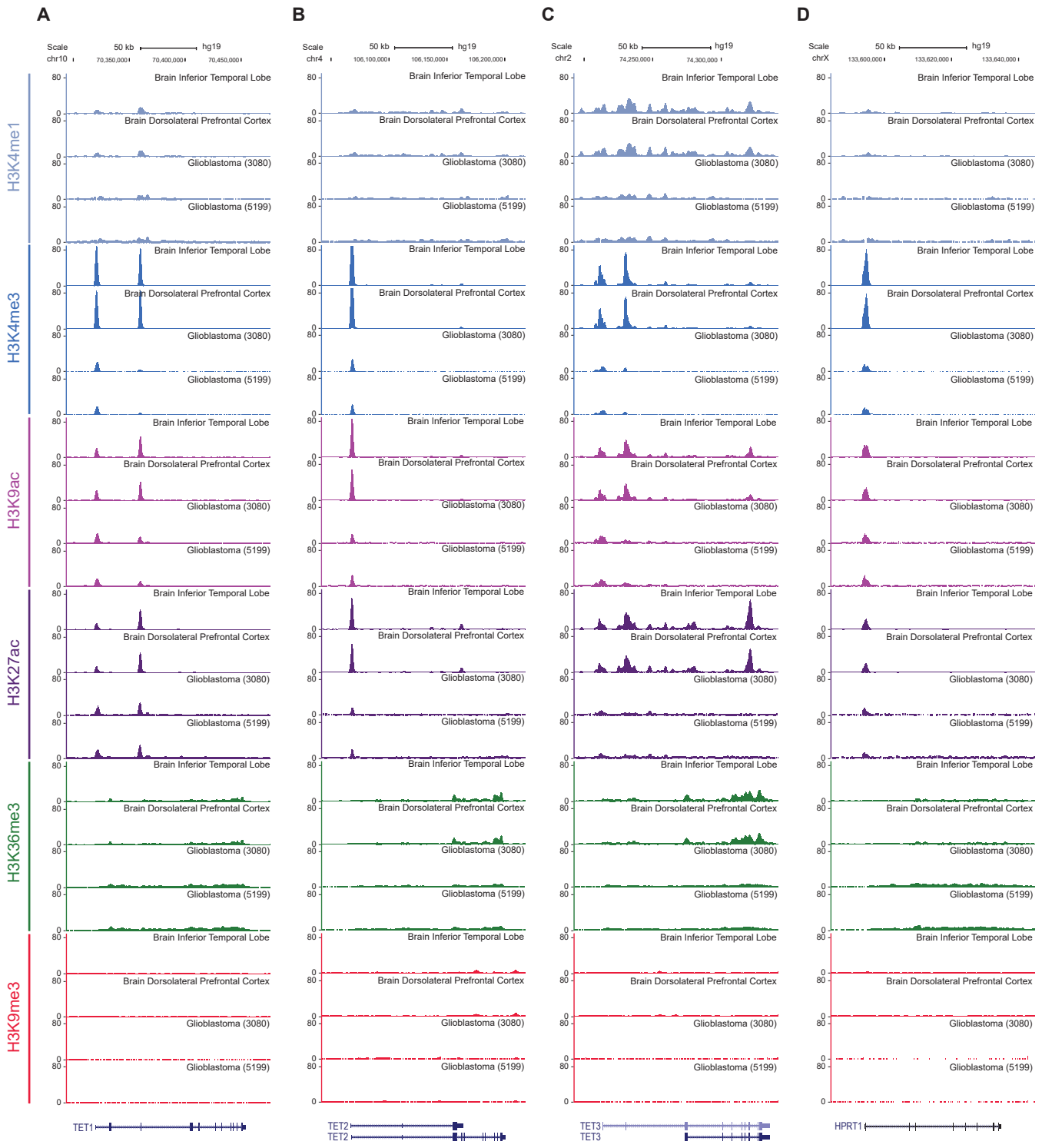
SUPPLEMENTARY FIGURES



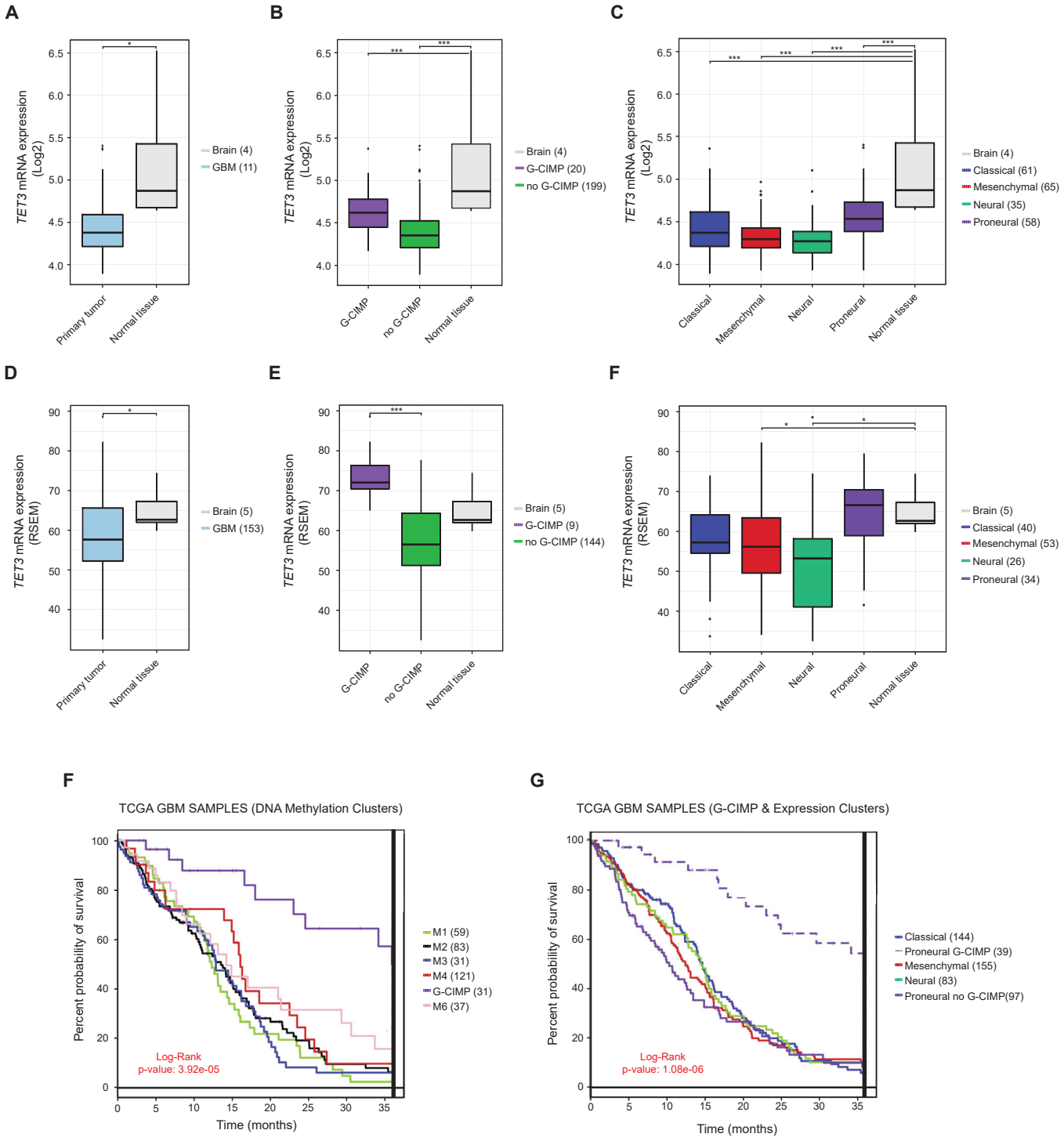
Supplementary Figure S1 - DNA methylation profiling along *TET* genes in human brain and glioma samples.



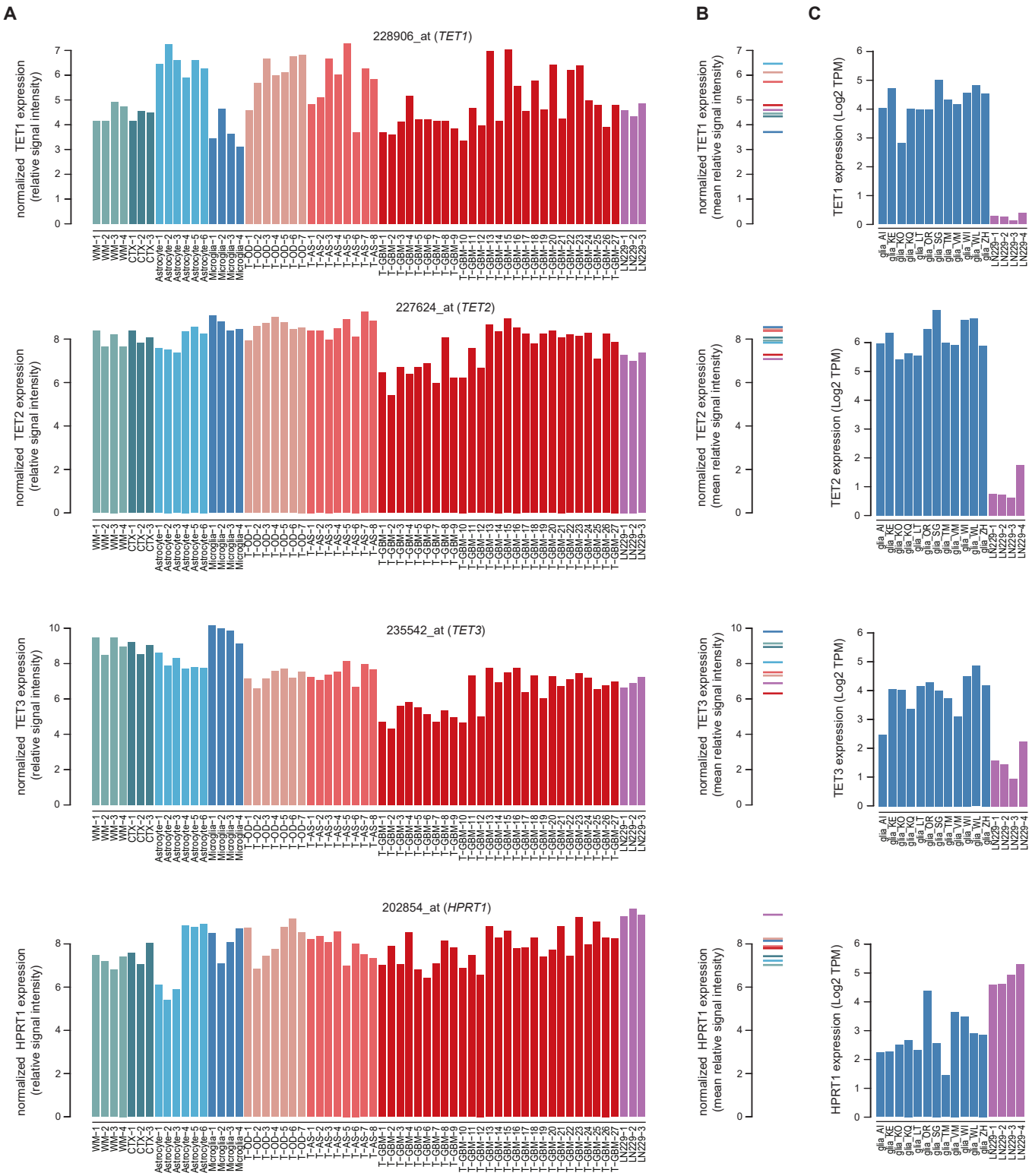
Supplementary Figure S2 - Biological validation in an independent cohort of samples confirms that loss of *TET3* 5hmC and gain of 5mC are a frequent phenomena in brain tumors



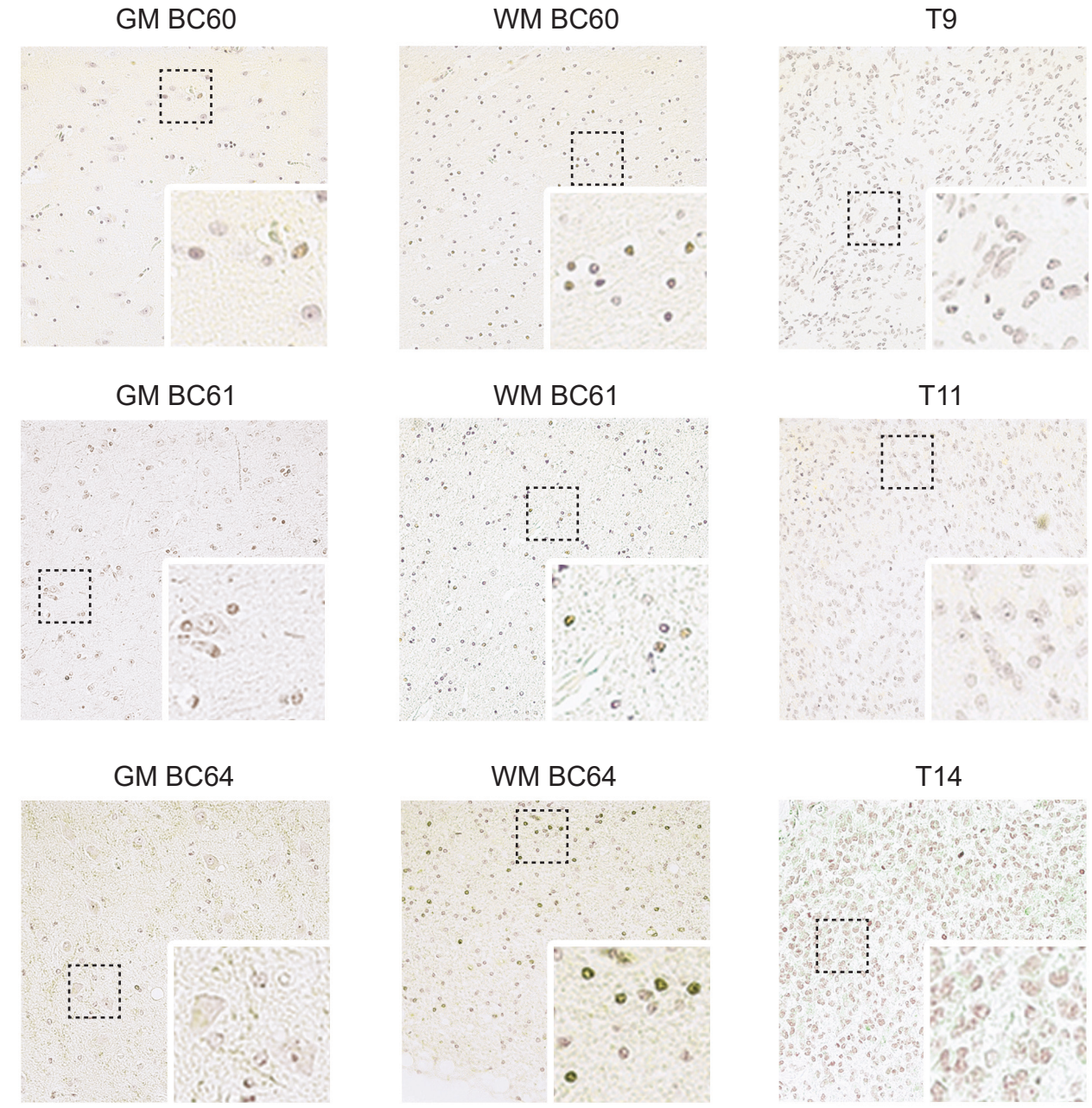
Supplementary Figure S3 - Epigenetic profiling of *TET* genes in brain tissue and glioblastoma multiforme



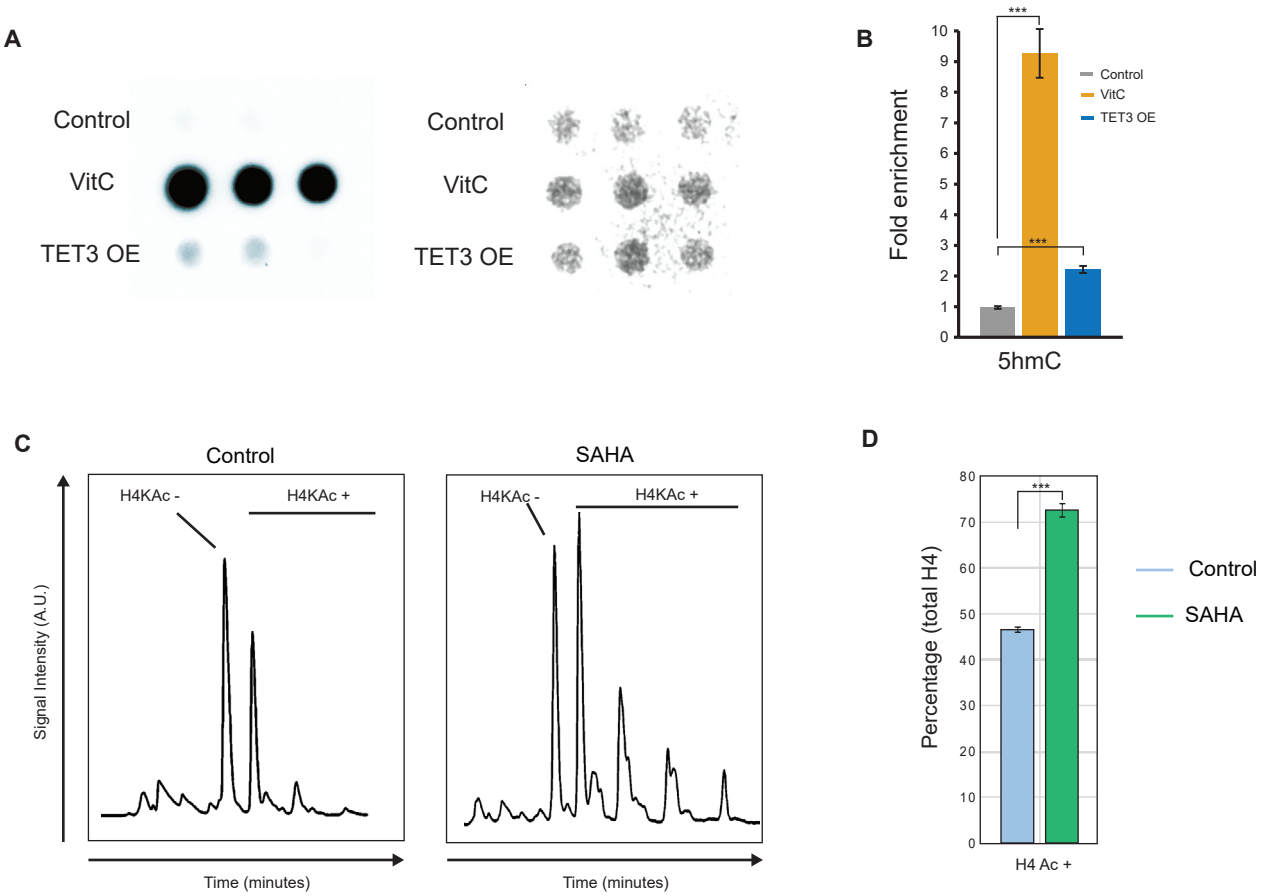
Supplementary Figure S4 - Computational validation using TCGA-GBM data indicates that *TET3* levels are reduced in human glioblastoma.



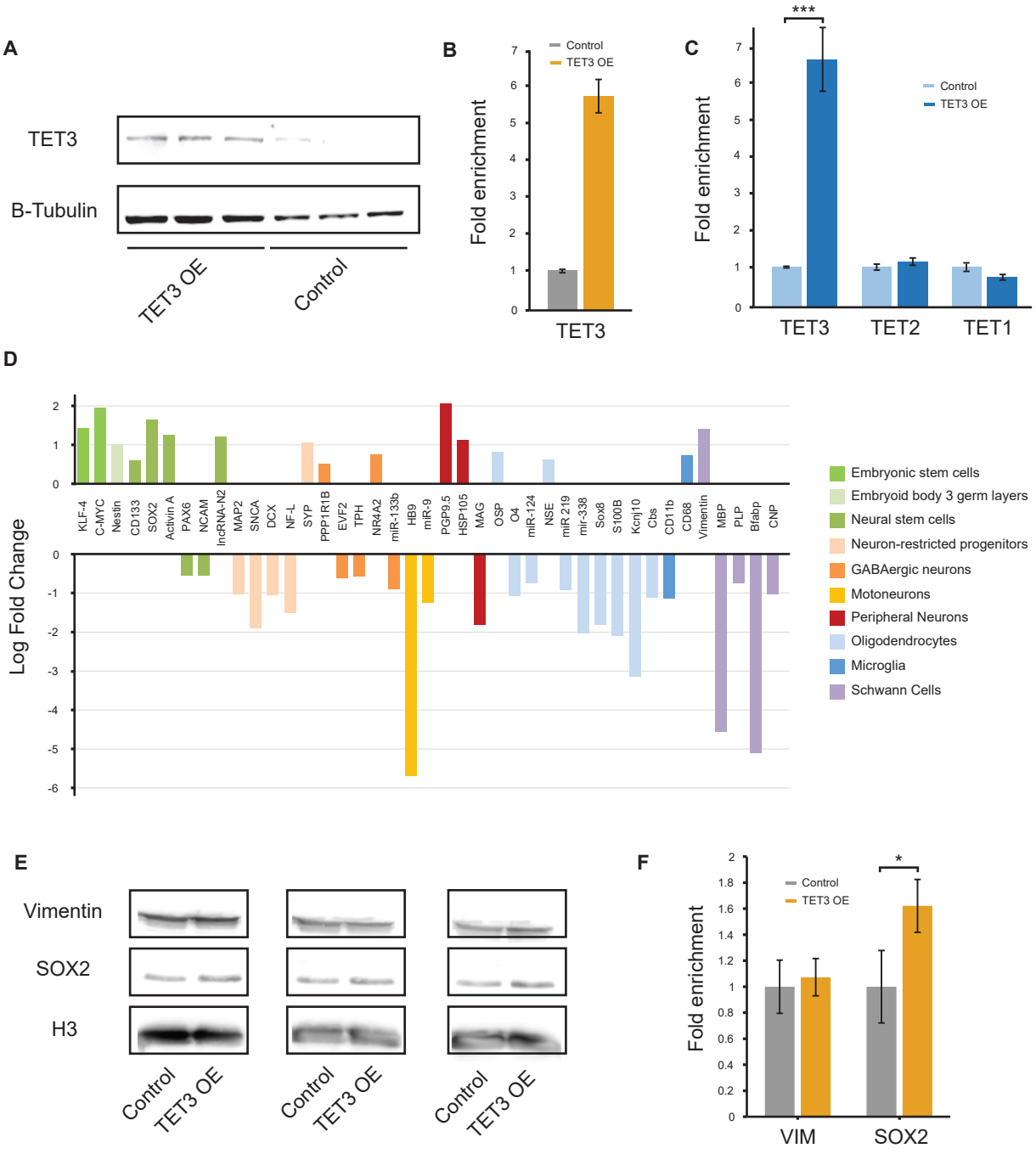
Supplementary Figure S5 - Expression of *TET* genes is compromised in glioma samples



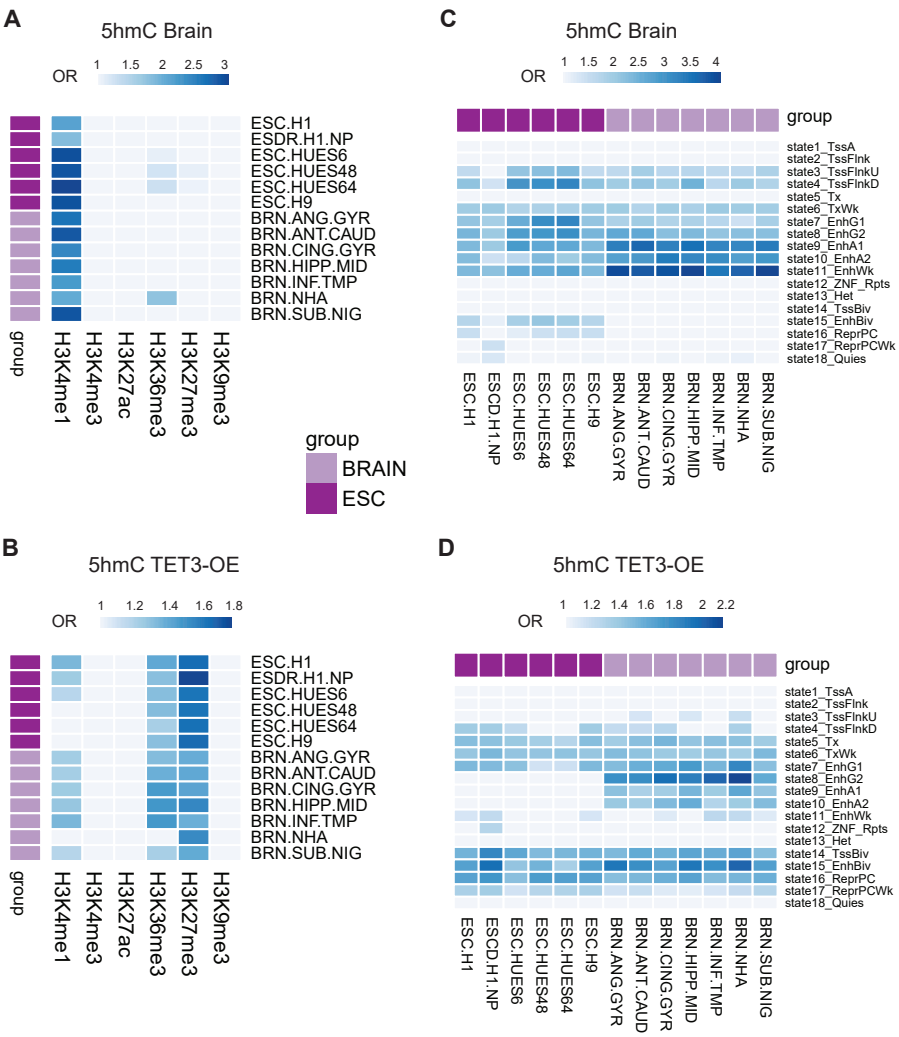
Supplementary Figure S6 - TET3 protein expression is diminished in glioblastoma samples.



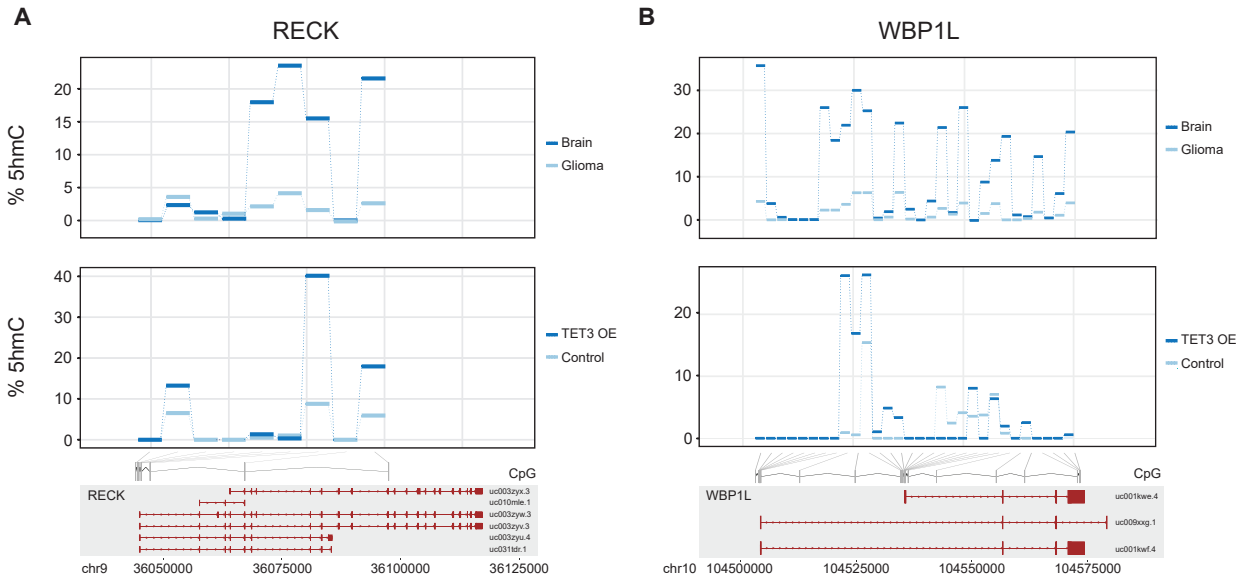
Supplementary Figure S7 - Effectiveness of epigenetic drug treatments in LN229 cells.



Supplementary Figure S8 - Characterization of the differentiation status of stably-transfected LN229 cells.



Supplementary Figure S9 - Chromatin signatures of 5hmC in brain and TET3 transfected glioblastoma cells



Supplementary Figure S10 - Partial restoration of 5hmC levels of *RECK* and *WBP1L* genes upon TET3 overexpression