SUPPLEMENTAL INFORMATION

This section includes Supplemental Figures and Figure Legends; Supplemental Table 1; Supplemental Methods; Supplemental References; and an Excel file with 2,577 differentially expressed genes in our NSCs titled “RNA-seq clusters”.

SUPPLEMENTAL FIGURES and FIGURE LEGENDS

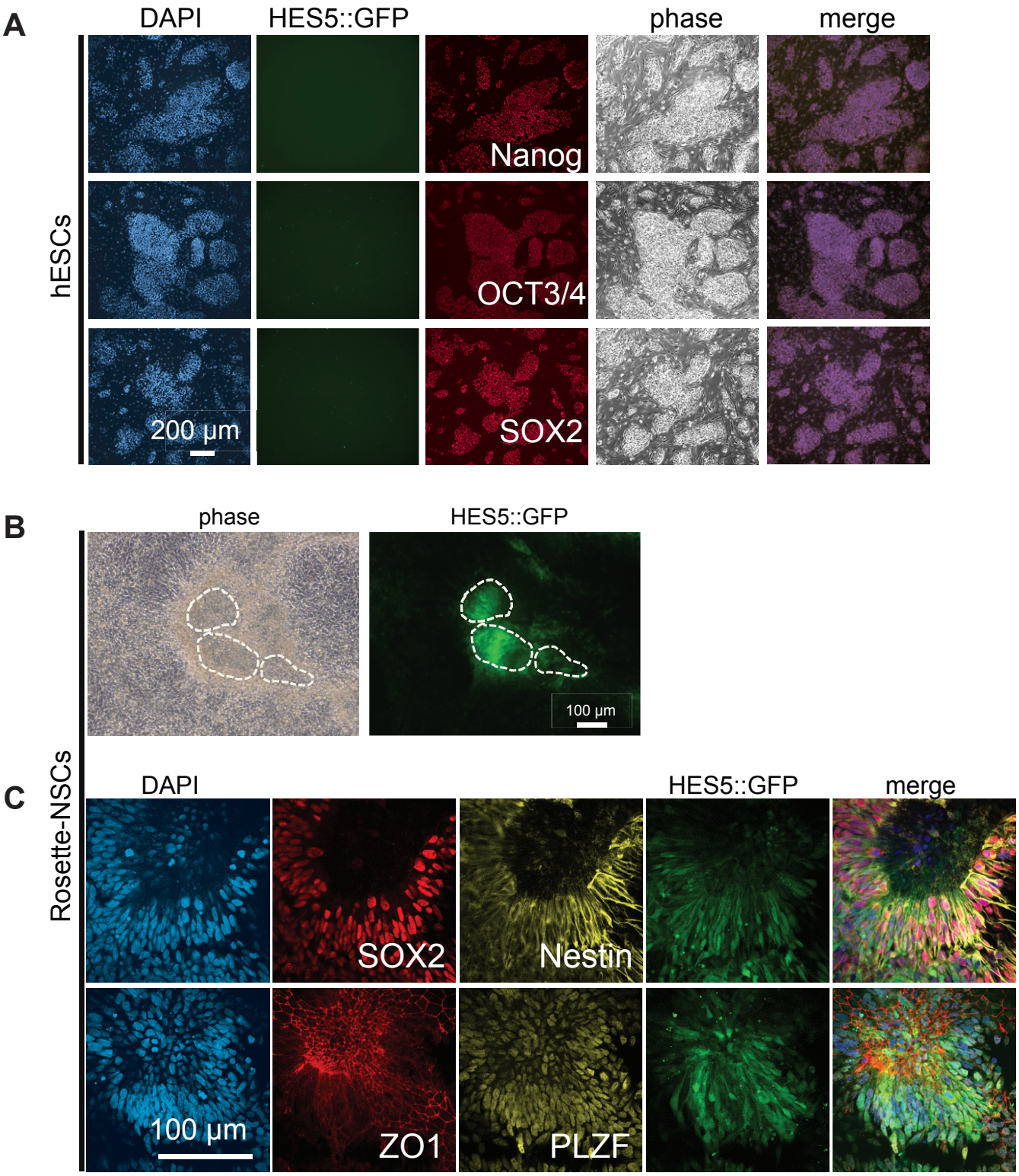

Figure S1. Derivation of rosette-NSCs from hESCs, Related to Figure 1.

A. hESCs grown in colonies on mouse embryonic fibroblast feeders were stained for pluripotency markers Nanog, OCT3/4 and SOX2. GFP from the *HES5::GFP* reporter (Placantonakis et al., 2009) was not expressed, as expected.

B. Live fluorescence microscopy shows rosette NSCs forming columnar, HES5::GFP+ structures *in vitro*.

C. Rosette NSCs are positive for markers ZO1, SOX2, Nestin and PLZF.

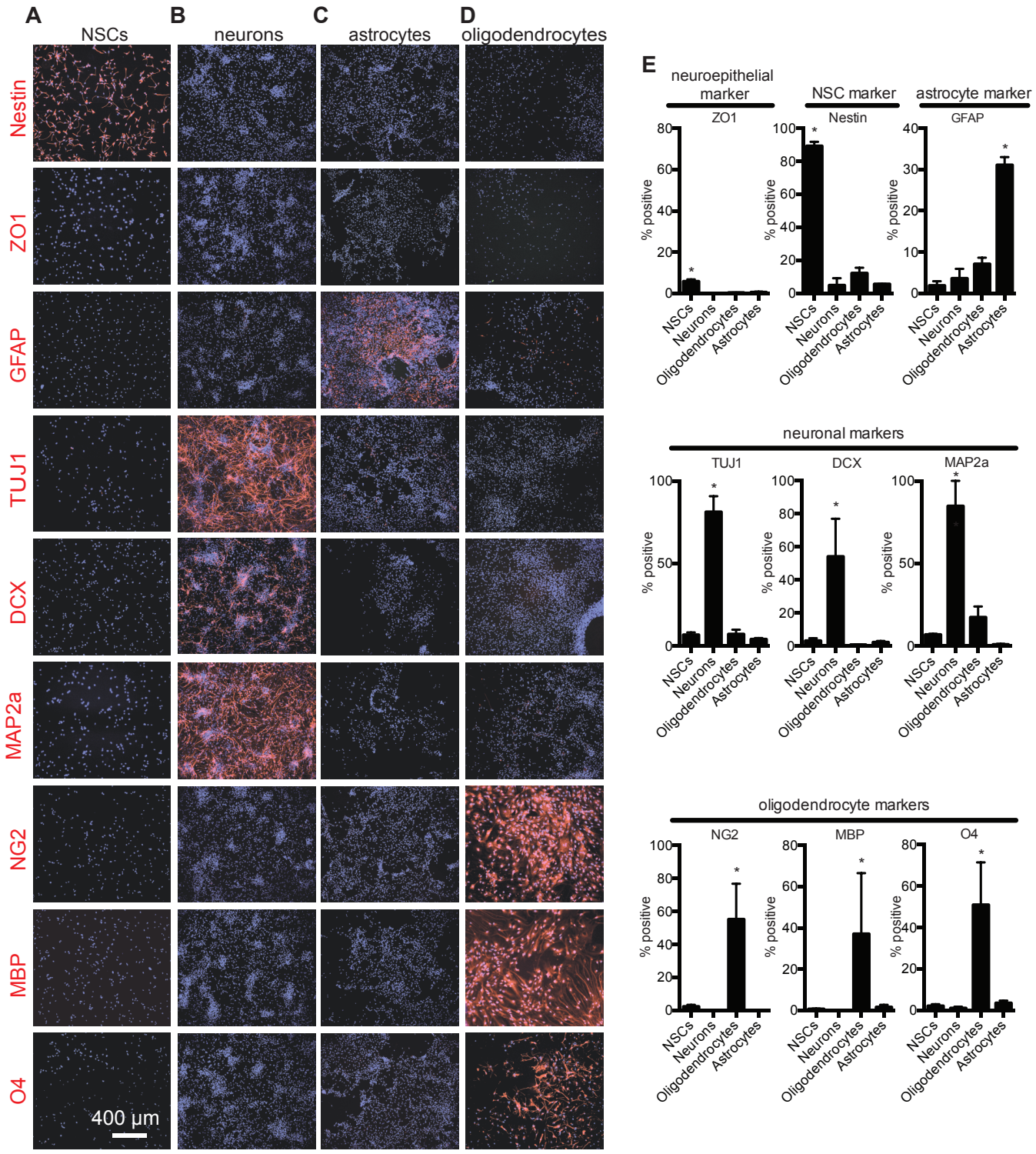

Figure S2. Neuroglial differentiation of NSCs, Related to Figure 1.

- **A.** hESC-derived NSCs stained positively for Nestin. Nuclei were counterstained with DAPI.
- **B.** NSC-derived neurons were positive for TUJ1, MAP2a and DCX.
- **C.** NSC-derived astrocytes stained positively for GFAP.
- **D.** NSC-derived oligodendrocytes are positive for NG2, MBP and O4.
- **E.** Quantification of neuroglial markers in NSCs and NSC-derived cell types (n = 3/condition). The neuroepithelial marker ZO-1 marked ~5% of NSCs, and <~1% of differentiated progeny (ANOVA $F_{(3,8)}$=21.49, p=0.0003). Nestin stained ~90% of NSCs, and <10% of differentiated progeny (ANOVA $F_{(3,8)}$=257.7, p<0.0001). Astrocytic cultures were enriched for GFAP (~30%) (ANOVA $F_{(3,8)}$=53.86, p=0.0042), while other conditions had <10% GFAP+ cells.

Neurons were ~85% positive for TUJ1 (ANOVA $F_{(3,8)}$=55.41, p=0.016) and MAP2a (ANOVA $F_{(3,8)}$=21.48, p=0.0003) and ~55% positive for DCX (ANOVA $F_{(3,8)}$=5.122, p=0.0288). Oligodendrocytes were enriched for NG2 (~55% positive) (ANOVA $F_{(3,8)}$=6.28, p=0.0169), MBP (~37% positive) (t-test, p=0.05) and O4 (~50% positive) (ANOVA $F_{(3,8)}$=5.578, p=0.0232). *p<0.05, *post hoc* Tukey's test.

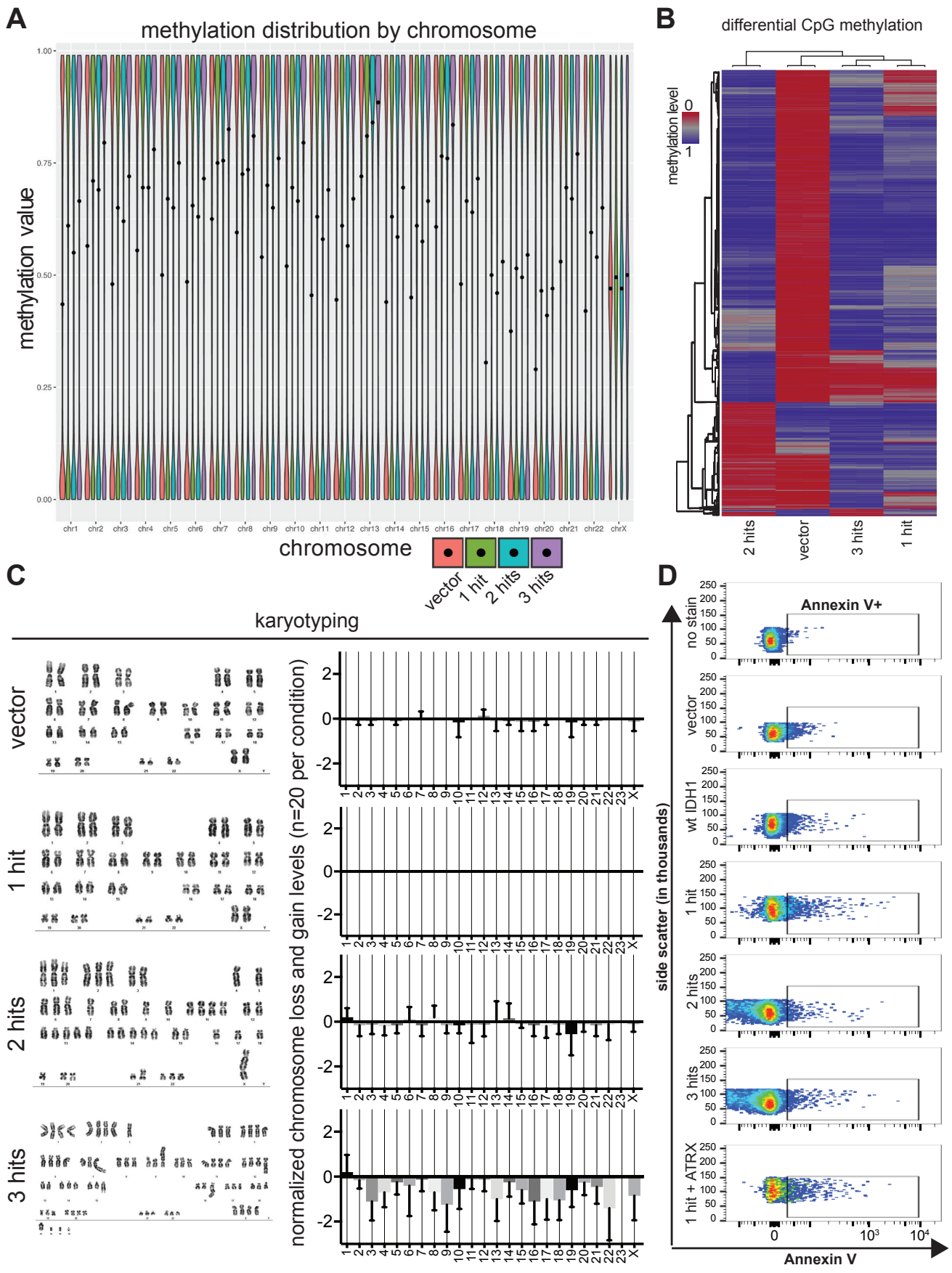
A
methylation distribution by chromosome
methylation value
1.00
0.75
0.50
0.25
0.00
chr1 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9 chr10 chr11 chr12 chr13 chr14 chr15 chr16 chr17 chr18 chr19 chr20 chr21 chr22 chrX
chromosome
vector
1 hit
2 hits
3 hits
B
differential CpG methylation
methylation level
0
1
2 hits
vector
3 hits
1 hit
C
karyotyping
vector
1 hit
2 hits
3 hits
normalized chromosome loss and gain levels (n=20 per condition)
2
0
-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 X
D
Annexin V+
no stain
vector
wt IDH1
1 hit
2 hits
3 hits
1 hit + ATRX
250
200
150
100
50
0
side scatter (in thousands)
0
10^3
10^4
Annexin V

Figure S3. DNA methylation, karyotypic and apoptotic profiles of NSCs, Related to Figures 2 and 3.

- **A.** Violin plots with marked mean values of β (methylation) status by chromosome number.
- **B.** Hierarchical clustering of the top 5000 most variable CpG loci amongst the four NSC lines profiled reveals each oncogenic hit is capable of differentially modulating the methylation status of select CpG sites.
- **C.** Representative karyotypes of vector, 1-hit, 2-hit and 3-hit NSCs (n=20 nuclei/condition). Corresponding bar graphs represent normalized lost or gained chromosome counts, which represent losses or gains relative to the corresponding aneuploidy.

  Karyotypes shown:

  vector: 46 XX (normal female)

  1 hit: 46 XX (normal female)

  2 hit: 50,tas(X;X)(q28;p22.3),+1,+2,-4,-5,+6,-7,-9,-10,+11,+11,+13,+13,+13,+14,+14,+14,-16,-17,-18,-19,tas(21;21)(p13;p13),+22

  3 hit: 79,XXXX,-3,-3,-3,-4,-7,tas(7;10)(q36;q34.3),-8,-9, add(9)(p24), dic(12;?)(q22;?),-13,-13,14,-15,der(15)t(1;15)(p13;q22),-16, dic(16;?)(q24;?),-17,-18,-19,-21,-22,+4ma
- **D.** Flow cytometry dot-plots of NSCs in stained for Annexin V. See **Figure 3F,G** for corresponding quantifications and histograms.

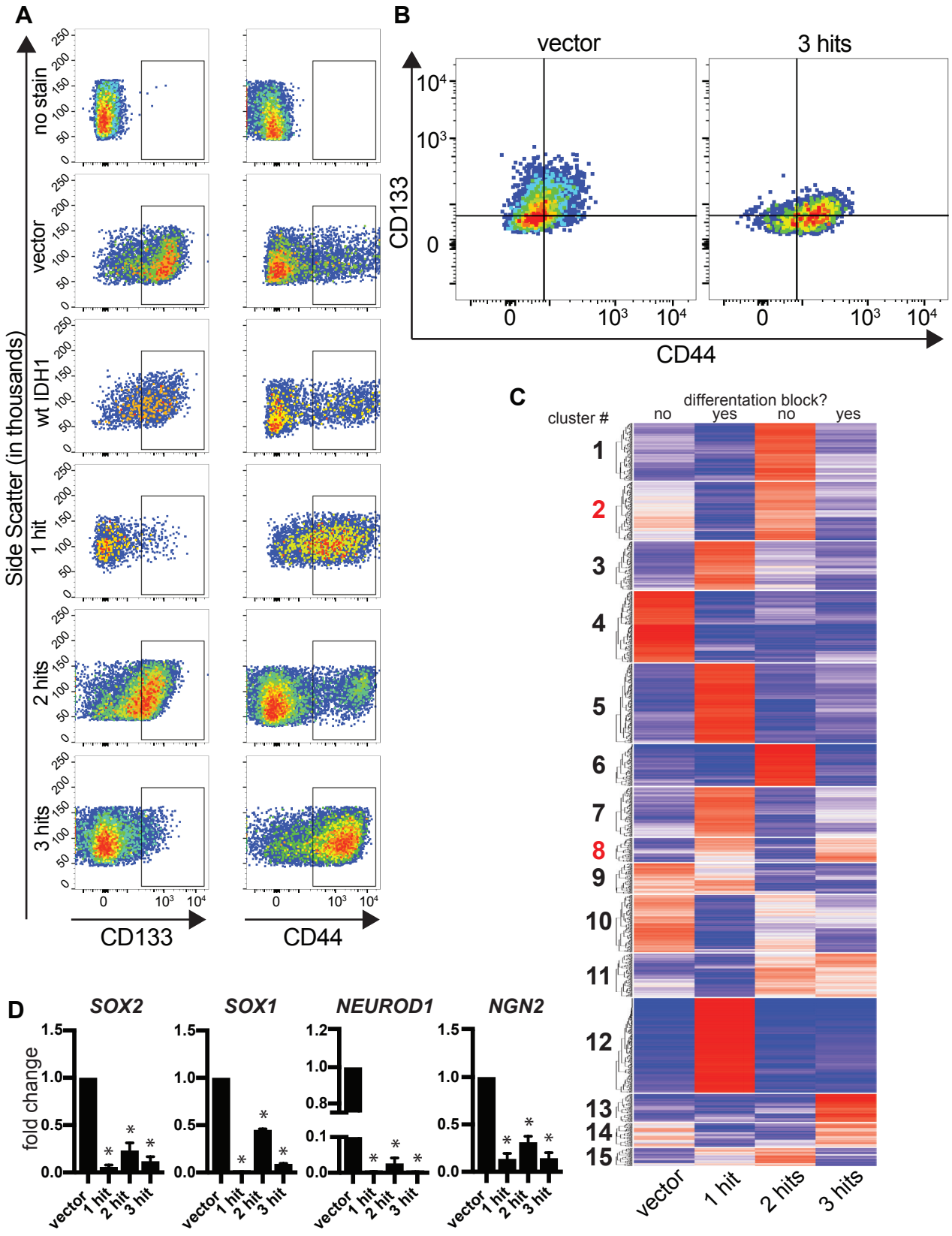
A
no stain
vector
wt IDH1
1 hit
2 hits
3 hits
Side Scatter (in thousands)
CD133
CD44
B
vector
3 hits
CD133
CD44
C
differentation block?
cluster #
no
yes
no
yes
vector
1 hit
2 hits
3 hits
D
SOX2
SOX1
NEUROD1
NGN2
fold change
vector
1 hit
2 hit
3 hit

Figure S4. Surface markers and gene expression profiles of NSCs, Related to Figure 4.

- **A.** Flow cytometry dot-plots of NSCs in self-renewing conditions stained for CD44 or CD133. See **Figure 4B** for corresponding quantifications and histograms.
- **B.** Flow cytometry dot-plots of NSCs in self-renewing conditions co-stained for CD44 or CD133. See **Figure 4C** for corresponding quantifications.
- **C.** Genes from RNA-seq (**Figure 4G**) were clustered into 15 k-means groups based on gene expression and represented on a heatmap without clustering of columns, which represent NSCs with vector, 1, 2 or 3-hits. Two of the 15 groups had genes with expression patterns that strongly correlated or inversely correlated with the ability of NSCs to differentiate.
- **D.** *SOX2* mRNA is downregulated ~10-fold in 1-hit NSCs, ~5 fold in 2-hit NSCs and ~10-fold in 3-hit NSCs (n=3, ANOVA $F_{(3,8)} = 85.6$, $p<0.0001$). *SOX1* mRNA is downregulated >300-fold in 1-hit NSCs, ~2 fold in 2-hit NSCs and ~100-fold in 3-hit NSCs (n=3, ANOVA $F_{(3,8)} = 5809$, $p<0.0001$). *NEUROD1* mRNA is downregulated >500-fold in 1-hit NSCs, ~50 fold in 2-hit NSCs and >500-fold in 3-hit NSCs (n=3, ANOVA $F_{(3,8)} = 3926$, $p<0.0001$). *NGN2* mRNA is downregulated ~10-fold in 1-hit NSCs, ~3 fold in 2-hit NSCs and ~10-fold in 3-hit NSCs (n=3, ANOVA $F_{(3,8)} = 64.1$, $P<0.0001$). *$p<0.05$, *post hoc* Tukey's test.

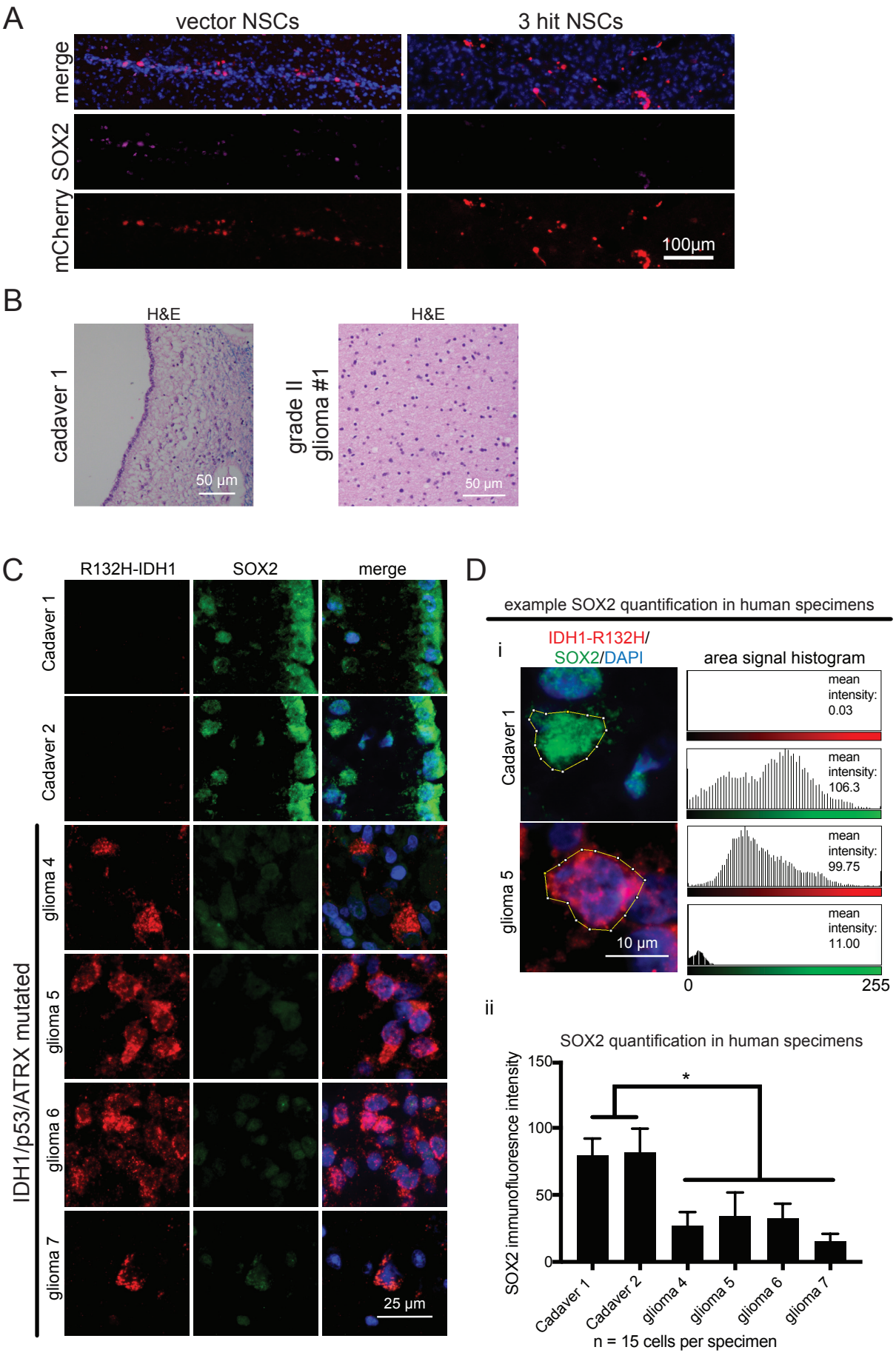
A
vector NSCs
3 hit NSCs
merge
mCherry SOX2
100µm
B
H&E
H&E
cadaver 1
grade II glioma #1
50 µm
50 µm
C
R132H-IDH1
SOX2
merge
Cadaver 1
Cadaver 2
IDH1/p53/ATRX mutated
glioma 4
glioma 5
glioma 6
glioma 7
25 µm
D
example SOX2 quantification in human specimens
i
IDH1-R132H/
SOX2/DAPI
area signal histogram
Cadaver 1
glioma 5
10 µm
mean intensity: 0.03
mean intensity: 106.3
mean intensity: 99.75
mean intensity: 11.00
0
255
ii
SOX2 quantification in human specimens
SOX2 immunofluoresnce intensity
150
100
50
0
*
Cadaver 1
Cadaver 2
glioma 4
glioma 5
glioma 6
glioma 7
n = 15 cells per specimen

Figure S5. SOX2 expression in NSCs *in vivo* and in human glioma samples, Related to Figure 4.

- **A.** SOX2 and mCherry immunofluorescence of human NSCs (vector or 3-hits), four-weeks post injection into NOD.SCID mouse brains.
- **B.** H&E of human cadaveric SVZ (normal brain parenchyma) and a representative H&E stain of a WHO grade II IDH mutated astrocytoma.
- **C.** R132H-IDH1 and SOX2 immunofluorescence microscopy of two SVZ regions from cadavers and four additional IDH1/P53/ATRX mutated WHO grade II astrocytomas.
- **D.** Quantification of SOX2 Immunofluorescence intensity from **Figure S5C**. (i) Example of quantification of individual SVZ cells and mutant IDH1-positive cells for SOX2 intensity. (ii) SOX2 quantification of SVZ cells and mutant IDH1 glioma cells from human specimens (n=2 cadavers, n=4 gliomas, n=15 cells per specimen, ANOVA $F_{(5,84)} = 68.76$, $p<0.0001$, $^*p<0.05$, *post hoc* Tukey's test).

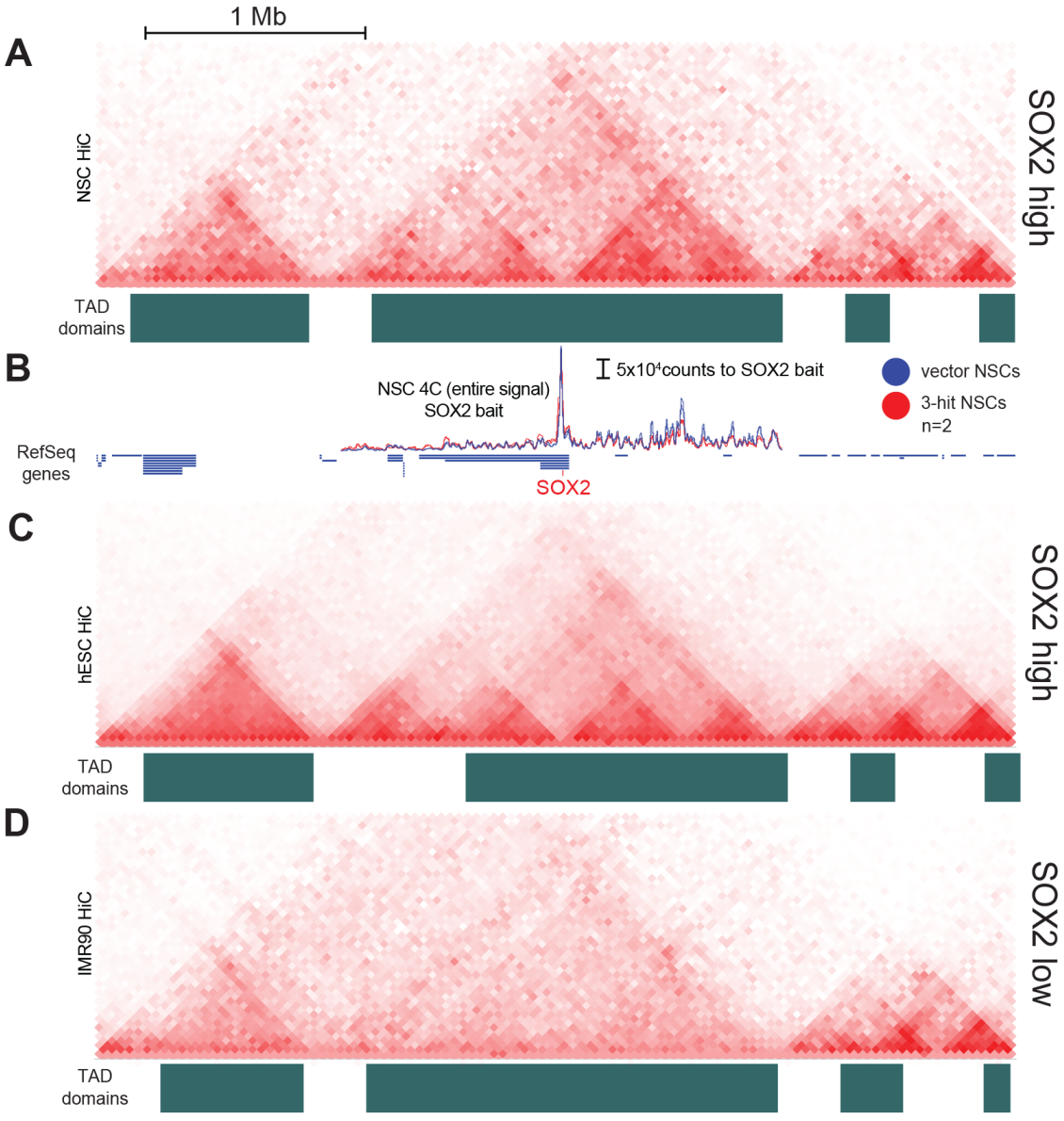

Figure S6. HiC and 4C profiles around the *SOX2* locus, Related to Figure 6.

A. 4 Mb HiC plot of H1-NSCs with aligned TAD domain tracks below.

B. 4C plots with vector and 3-hit NSCs (n=2 each) with RefSeq genes. *SOX2* is highlighted in red. The large majority of 4C interactions is downstream of *SOX2*.

C,D. hESC (**C**) and IMR90 (**D**) HiC plots with accompanying TAD domains.

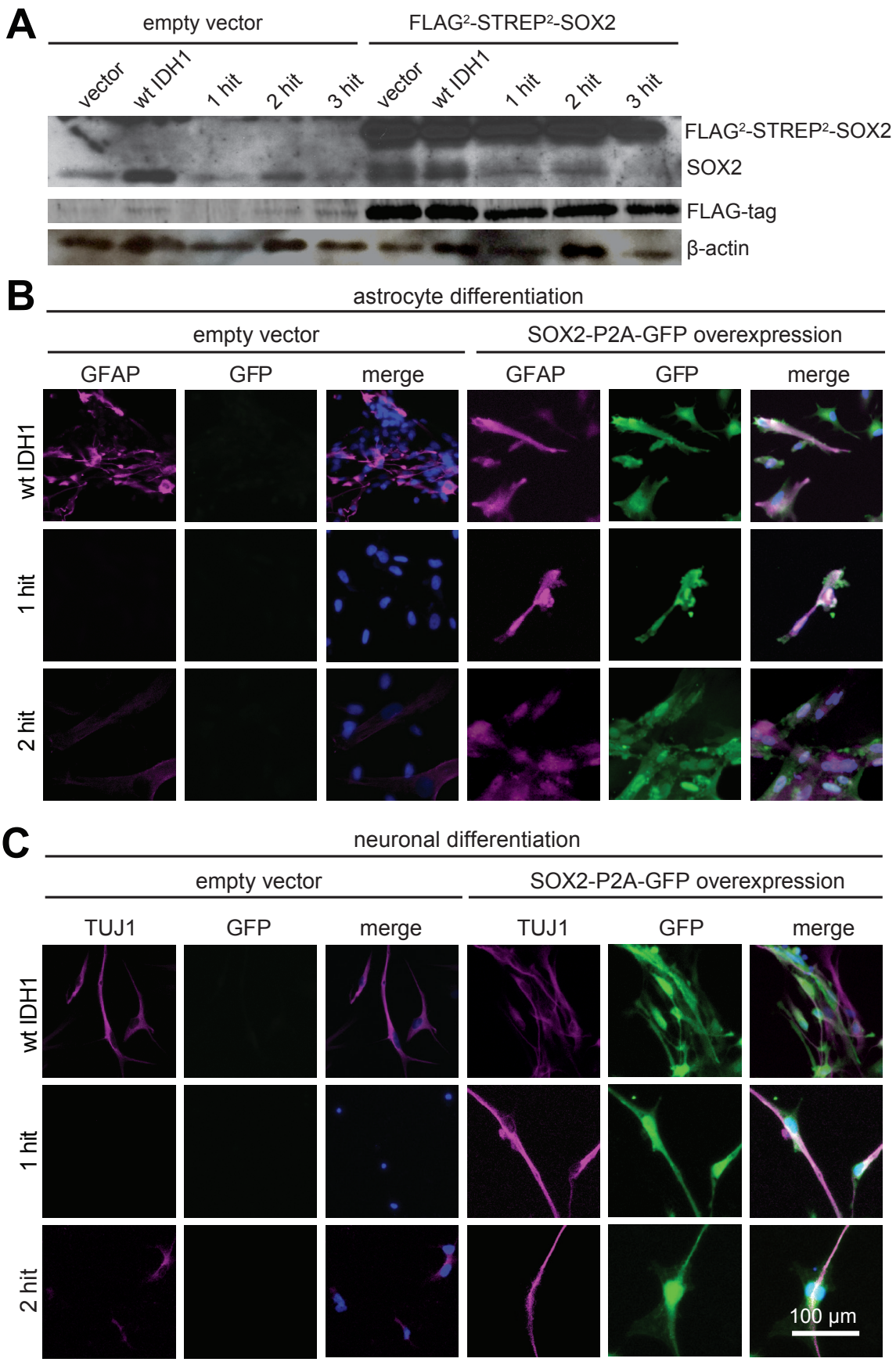
A
empty vector
FLAG²-STREP²-SOX2
vector
wt IDH1
1 hit
2 hit
3 hit
vector
wt IDH1
1 hit
2 hit
3 hit
FLAG²-STREP²-SOX2
SOX2
FLAG-tag
β-actin
B
astrocyte differentiation
empty vector
SOX2-P2A-GFP overexpression
GFAP
GFP
merge
GFAP
GFP
merge
wt IDH1
1 hit
2 hit
C
neuronal differentiation
empty vector
SOX2-P2A-GFP overexpression
TUJ1
GFP
merge
TUJ1
GFP
merge
wt IDH1
1 hit
2 hit
100 µm

Figure S7. Ectopic SOX2 rescues the differentiation block, Related to Figure 7.

- **A.** Immunoblot of NSC lines transduced with an empty or SOX2-expressing vector (**Figure 7A**). The tagged $FLAG^2$-$STREP^2$-SOX2 protein is larger than the endogenous SOX2. FLAG-tagged SOX2 was only present in lines transduced with $FLAG^2$-$STREP^2$-SOX2. β-Actin was used as a loading control.
- **B.** Additional immunofluorescent GFAP stains of cells from the astrocyte differentiation experiment outlined in **Figure 7B.** GFAP images are merged with the GFP and DAPI.
- **C.** Additional immunofluorescent TUJ1 stains of cells from the neuronal differentiation experiment outlined in **Figure 7C**. Tuj1 images are merged with the GFP and DAPI.

SUPPLEMENTAL TABLE 1. Patient demographics and molecular characteristics, Related to Figure 4.

patient	histology	WHO grade	IDH	P53	ATRX	1p/19q	sex	age
cadaver 1	normal	n/a	n/a	n/a	n/a	n/a	F	66
cadaver 2	normal	n/a	n/a	n/a	n/a	n/a	M	66
patient 1	astrocytoma	II	R132H	loss	loss	intact	F	35
patient 2	astrocytoma	II	R132H	H179R	loss	unknown	F	15
patient 3	astrocytoma	II	R132H	unknown	unknown	intact	F	29
patient 4	astrocytoma	II	R132H	unknown	loss	intact	M	37
patient 5	astrocytoma	II	R132H	R273C	loss	intact	M	59
patient 6	astrocytoma	II	R132H	unknown	loss	intact	F	50
patient 7	astrocytoma	II	R132H	R273C	loss	intact	F	37

SUPPLEMENTAL METHODS

Generation and differentiation of human NSCs

Human Embryonic Stem Cell (hESC) culture: hESCs (WA09 line) with a HES5::GFP reporter (Placantonakis et al., 2009) were cultured as colonies in hESC medium (1:1 DMEM/F12 (Invitrogen), 20% KSR (Invitrogen), 5 ng/ml FGF2 (R&D), 2 mM glutamine (Invitrogen), 0.1 mM non-essential amino acids (Invitrogen), 0.1 mM β-Mercaptoethanol (Invitrogen) on a layer of mouse embryonic fibroblasts (MEFs; GlobalStem) plated 24 hours earlier on 0.1% (w/v) gelatin. Human ESC culture was as previously described (Thomson et al., 1998). Cultures were routinely groomed of differentiating colonies and fed daily. Passaging was performed with dispase (Invitrogen) or by mechanically picking colonies. All hESC culture and lines were approved for use under ESCRO (protocol #14-00267) at New York University School of Medicine. Human ESC to Rosette-NSC differentiation was performed as described (Elkabetz et al., 2008; Placantonakis et al., 2009). Briefly, hESCs were passaged as colonies via dispase digestion into KSR medium as floating serum-free embryoid bodies. The medium was gradually switched to N2 medium (DMEM/F12 with N2 supplement from Invitrogen) in the presence of 100 ng/mL Noggin (R&D) and 10 µM SB51329 (Tocris). Rosette-NSC to NSC differentiation was modified and adapted from prior protocols (Conti et al., 2005; Koch et al., 2009; Tabar et al., 2005). Rosette-NSCs were mechanically picked, dissociated in trypsin, and plated at $1x10^5$ cells/cm^2 on poly-L-ornithine (Sigma) and laminin (Sigma)-coated plates in Neural Stem Cell Medium (NSCM) composed of DMEM/F12 (Invitrogen), N2 (Invitrogen), 20 µg/mL Insulin (Invitrogen), 1:1000 B27 (Invitrogen), 1.6 g/L glucose (Fisher) and 20 ng/mL EGF and FGF2 (R&D Systems). NSC lines were generated from the hESC WA09 line several times. Cells were passaged at 1:2 ratios and maintained at high densities for 30 days before being used for downstream experiments. For neuronal differentiation, NSCs were exposed to 20 ng/mL BDNF (R&D) and 200 nM Ascorbic Acid (Sigma) in N2 medium. Medium and growth factors were replaced every two days for 30 days before analysis. Astrocytic differentiation was performed with either 10 ng/mL of BMP4 (R&D) or 2% FBS (Invitrogen), with fresh medium and growth factors every two days for 30 days. Cell lines used in experiments were generated three times, as outlined in **Figure 1A-C**. Cell lines were infected (MOI = 1) first with an mCherry-containing lentivirus that harbored wild-type IDH1, mutant IDH1 or mCherry alone (vector), and were FACS-sorted. These cell lines were then infected (MOI = 1) with either P53 shRNA lentiviruses or ATRX shRNA lentiviruses.

Flow cytometry, immunostaining and immunoblotting

To perform flow cytometry cells were dissociated with 0.5% Trypsin (Invitrogen) and washed with NSCM before live staining. CD133 (1:10 dilution) and CD44 (1:150 dilution) staining was performed with APC or FITC conjugated antibodies (Miltenyi). Flow cytometry was performed on a LSRII analyzer (BD Biosciences) and analyzed using FloJo software. For cell cycle analysis, cells were incubated in NSCM for 45 minutes in 1 µg/mL Hoechst dye, and analyzed via flow cytometry. Data for cell cycle analysis was performed using built-in FloJo software that defines and gates the cell cycle phases.

For immunostaining, cells were washed of medium with PBS and incubated in 4% (v/v) paraformaldehyde for 30 minutes for fixation. Cells were then washed in 0.1% (v/v) Triton-X100 in PBS three times and incubated for 1 hour at room temperature with blocking solution, which consisted of 10% (w/v) BSA, 0.1% Triton-X100 in PBS. Primary antibodies were then incubated overnight at 4°C. Samples were washed again three times with 0.1% (v/v) Triton-X100 in PBS. Secondary antibodies (Alexa flour, Invitrogen) were added at 1:500 for 45 minutes at room temperature protected from light. After three more washes in 0.1% (v/v) Triton-X100 in PBS, cells were incubated with 1 µg/mL DAPI for 10 minutes and washed again before visualization. A list of the primary antibodies used in this study is shown below.

To quantify immunofluorescence intensity, all material to be quantified was handled, stained and imaged using the same methodology for consistency. To quantify staining intensity differences, ImageJ's plugin, Color Histogram, was used after defining single cells within an image.

To generate lysates for immunoblots, cells were washed in PBS, collected via cell scrapers and suspended in lysis solution (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% (w/v) SDS, 10% (v/v) glycerol, and 1x Roche protease inhibitor cocktail). Whole lysates were boiled for 20 minutes and spun down at 18,000 g to remove debris. Quantification prior to SDS-PAGE was performed using DC protein assay (Biorad). Equivalent amounts of lysates were used for immunoblotting. After transfer, primary antibodies were incubated overnight at 4°C with gentle shaking. Secondary antibodies, either conjugated to HRP (Life technologies) or to fluorophores (Licor), were used according to manufacturer protocols for final visualization.

Antibodies and dilutions used in this study. IHC: immunohistochemistry

Manufacturer	Catalog #	Primary Antibody	Immuno-fluorescence	Immuno-blot	Flow cytometry
Origene	TA500610	IDH1	1:80	1:2000	
Histonova	DIA-H09	R132H-IDH1	1:80 or 1:50	1:500	
Genetex	GTX101448	Hsp90-beta		1:1000	
Invitrogen	339100	ZO-1	1:50		
R&D systems	AF2018	SOX2	1:50 1:100 for IHC	1:500	
Santa Cruz	sc-5279	OCT3/4	1:100		
Santa Cruz	sc-22839	PLZF	1:100		
R&D systems	AF1997	Nanog	1:100		
Neuromics	MO15012	Nestin	1:500		
DAKO	Z0334	GFAP	1:500		
Santa Cruz	sc-8066	DCX	1:200		
R&D systems	MAB1326	O4	1:100		
Biolegend	845502	TUJ1	1:500		
abcam	ab13970	GFP	1:1000		
Clontech	632496	mCherry	1:500		
Santa Cruz	sc-69879	ß-Actin		1:2000	
Bethyl Labs	A301-045A	ATRX	1:50	1:500	
Santa Cruz	sc-126	P53		1:2000	
Milipore	AB5320	NG2	1:200		
Miltenyi	130-098-110	CD44			1:150
Miltenyi	130-098-829	CD133			1:10
Santa Cruz	PG-M3	PML	1:100		
Generated by Dr. Susan Smith's lab (Cook et al., 2002)	n/a	TRF1	1:1000		

Lentivirus production

HEK-293T cells (Clontech) were brought to 95% confluence in 10% FBS (Invitrogen) in DMEM (Invitrogen) and transfected with pLP1, pLP2 and pVSVG, along with any given lentiviral transfer vector using lipofectamine 2000 (Invitrogen). Medium was collected at 24 hours and 48 hours and pooled. Viral supernatant was mixed with lenti-concentrator-X (Clontech) and precipitated. Viral pellets were suspended in N2 medium and stored in -80°C. Viral stocks were titered using qPCR viral titer kits (ABM) or using flow cytometry of transduced cells.

Extreme Limiting Dilution Assay (ELDA)

NSCs were passaged into limiting dilutions (100 cells to 1 cell) in low-attachment 96-well dishes with 100 µL of NSCM with 20 ng/uL of EGF and FGF2 per well. After two weeks of replenishing medium and growth factors every two days, wells were assessed for sphere presence. Statistics and sphere forming ability were calculated using available ELDA analysis software outlined (Hu and Smyth, 2009).

***In vivo* xenografts**

Male NOD.SCID mice (6–8 weeks) were stereotactically injected with $2.5x10^5$ NSCs in the frontal lobe, as previously described (Bayin et al., 2014). Procedures were performed according to IACUC protocol number 160403 at NYU School of Medicine. NSCs were analyzed 4 weeks post-injection. Images were processed using ImageJ software.

Quantitative-PCR

Quantitative-PCR was performed according to the manufacturer using the Cells-to-C_t kit (Invitrogen), Taqman Gene Expression kit (Invitrogen), and the StepOne Real Time PCR System (Invitrogen) for gene expression profiling. Following completion of the qPCR, extracted C_t values from the StepOne Software (v2.1) were assessed to calculate the normalized target gene expression level using the $\Delta\Delta C_t$ method to determine relative fold-change. The following TaqMan (Invitrogen) primers were used: *TP53*: Hs01034249_m1; *ATRX*: Hs00230877_m1; *SOX1*: Hs01057642_s1; *SOX2*: Hs01053049_s1; *NEUROD1*: Hs00159598_m1; *NEUROG2* (NGN2): Hs00702774_s1; and *HPRT1*: Hs01053049_s1. For ChIP-qPCR, chromatin from vector and 3-hit NSCs (n=2) was extracted using EZ-Magna ChIP A/G Chromatin Immunoprecipitation Kit

(Millipore #17-10086) and immunoprecipitated with CTCF antibody (Millipore # 07-729) according to manufacturer protocols. Purified DNA was subject to qPCR using the primers shown below. Input DNA was used as a normalization control.

Primers for CTCF ChIP-qPCR

CTCF site (chr3)	**sequence of qPCR primers** 5' to 3'
i	GTGATGGGGCTTGCTCTTTG
i	AAAACCCAGCTGCTTTTCCC
ii	AACCCGACTCAGATGTGAAGC
ii	AGCAGGATTCCAGCGATTTC
iii	AAAGCAATAGTCTTGGAAAAGCG
iii	AGGTTGAGAAGATCTGGGGAAG
iv	TGTTTGATCTGGGACTGCGTAG
iv	CGTCCCGGCTTGCTAAATATG
v	ACAGTGGGCATTTTCAGTGC
v	ACAAATCTTGGCAGGTCCTG

Methylation profiling

Genomic DNA was extracted from $5x10^5$ cells in biological duplicates for vector, 1-hit, 2-hit and 3-hit NSCs using the DNeasy Blood and Tissue Kit (Qiagen). The Qubit 2.0 Fluorometer (Invitrogen Life Technologies) was used for DNA quantification with the Qubit dsDNA BR Assay Kit (Invitrogen Life Technologies). DNA was subject to bisulfite conversion with the EZ-96 DNA Methylation Kit (Zymo Research). Global methylation profiling was performed using the Infinium HumanMethylation450 BeadChip Kit. BeadChip arrays were analyzed using the R package RnBeads verson 1.0.0 (Assenov et al., 2014). For comparison to data generated in this study, the raw data for TCGA's low-grade glioma samples were downloaded (Cancer Genome Atlas Research et al., 2015), and processed simultaneously in an identical manner.

For targeted bisulfite sequencing, genomic DNA extracted using the DNeasy Blood and Tissue Kit (Qiagen) was made ready for methylation analysis through bisulfite conversion of the vector and 3-hit NSC DNA with the EpiMark Bisulfite Conversion Kit (NEB). Corresponding primers (shown below) were used to amplify bisulfite-converted segments of the genome with the EpiMark Taq polymerase (NEB). PCR products were run on an agarose gel, extracted and cloned into the pLVX-N1 (Clontech) vector using standard molecular cloning techniques. Ligations were transfected into One Shot Stbl3 chemically competent E. coli (Invitrogen). Ten clones of each condition were sent for Sanger sequencing.

Primers for CTCF motif bisulfite sequencing

CTCF site (chr3)	**sequence** 5' to 3' (with 5' NotI or EcoR1 overhangs for sub-cloning) note: Y and R nucleotides are used in these primers
i	ACATGCGCGGCCGCTATGGGGGTAGAATAAGTAGGAAATATTA
i	ACATGCGAATTCAAACCCAACTACTTTTCCCAATACTA
ii	ACATGCGCGGCCGCGTGTTTAGYGTTTTTTGTTGGATATAGTAGG
ii	ACATGCGAATTCAAAACAATAAAAAAATTTTCTCTAACCCCAAAAC
iii	ACATGCGCGGCCGCTGGAAAAGYGAATTTATTAGTAGGGGG
iii	ACATGCGAATTCACRCTCTTCAAAATACAACACTATTTTC
iv	ACATGCGCGGCCGCGGGGTTAGYGTTGAGTTGTAGATTTGG
iv	ACATGCGAATTCCRAAACCCCCTCCTCCATCCC
v	ACATGCGCGGCCGCGTGGGTATTTTTAGTGTTGTTAATAGTGTTTATG
v	ACATGCGAATTCTAAATCCRCCAACAAAAAAAATTAAAATAATTAAC

RNA-seq analysis

Vector, 1-hit, 2-hit and 3-hit NSCs had RNA extracted and purified with the RNeasy Micro Kit (Qiagen) in duplicates. Libraries were generated using TrueSeq mRNA preparation kits (Illumina). Libraries were multiplexed and sequenced using 50-nucleotide paired-end reads in an Illumina HiSeq 2500 sequencer. Sequencing results were analyzed using the Tuxedo suite (Trapnell et al., 2012). RNA-seq data used for the transcriptional comparison of NSCs in this study was from 239 LGGs in the TCGA (Cancer Genome Atlas Research et al., 2015). All tumors had annotated molecular data, such as 1p/19q, P53 and ATRX status. The 500 most variable genes amongst the 239 tumors from the TCGA studied were used for differential analysis alongside our RNA-seq data. The RNA-seq data from this study was re-processed using the

same pipeline outlined by the TCGA in order to facilitate comparison (Cancer Genome Atlas Research et al., 2015). Data was visualized using the R software suite (www.r-project.org).

Circularized Chromosome Conformation Capture (4C)

To generate a 4C library for sequencing, we modified a recent protocol (Rocha et al., 2016) as follows: To capture interactions from the *SOX2* promoter, the following primer pairs were used: 5'-TCCACTCTCTTGTTGGGATC-3', and 5'-CTTCTAGTCGGGACTGTGAG-3' with DpnII and Csp6I for restriction digests. For the generation of 4C libraries, 10^7 vector NSCs and 3-hit NSCs in duplicates were collected, immediately fixed, and processed accordingly. Illumina adapters were added to the primers before generation of 4C libraries for multiplexed sequencing. 4C libraries were subjected to Agilent Tapestation quality control and quantified with qPCR. Libraries were then sequenced on Illumina HiSeq 2500 sequencers with 50-nucleotide single-end reads. 4C-Seq data were aligned and processed as described previously using 4C-ker (Raviram et al., 2016). Differential analysis was performed in the near-bait region (2 Mb surrounding the bait) using an adjusted p value of 0.001.

Track alignments and visualizations

Methylation data (450k illumina arrays) was visualized on Integrative Genome Browser. For NSC methylation array data, replicates were averaged to a single track. Methylation data from the TCGA was averaged into a single track divided by subtype (wild-type IDH, n=53; and mutant IDH1 with 1p/19q intact, n=157). CTCF ChIP-seq data was obtained using GEO accession GSE70991 (Flavahan et al., 2016). We selected ChIP-seq data from two grade III mutant IDH and two grade IV wild-type IDH datasets that were generated and processed at the same institution (MGH) in order to allow comparison of ChIP-seq signal. All tracks were plotted on the same intensity scale.

ChromHMM and roadmap epigenome tracks: We visualized the 25-state imputed chromHMM model that uses integrative data from 127 cell lines that had at least five different chromatin marks used to annotate the epigenome (Roadmap Epigenomics et al., 2015). ChromHMM data was visualized using the WashU epigenome browser and rows were manually reorganized to group ESCs, iPSCs and neural tissues together.

HiC data: Publicly available Hi-C data from lung fibroblasts (IMR90), H1 hESCs and NSCs were downloaded from Gene Expression Omnibus (GSE35156 and GSE52457) (Dixon et al., 2015; Dixon et al., 2012).

Metabolite mass spectrometry

For metabolite extraction and sample preparation, media was changed 24 hours and 1 hour prior to harvest. At harvest, a second plate of cells with equal density was used to count the total number of cells used for harvest. Cells were placed on ice, washed twice with ice cold PBS and once with ice-cold MS-grade ddH_2O (Fisher). Cells were rapidly placed on a bed of dry ice and 80% (v/v) MeOH was added to each well for 15 minutes. Cells were scraped off and lysates were spun down at 14,000g for 20 minutes at 4°C. Ssupernatants were saved as extracts. The extract was dried in a SpeedVac, and then dissolved in 100 µL 80% (v/v) methanol with 1% formic acid. Undissolved particles were removed by filtering with C18 StageTips prior to liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. For all LC-MS/MS analysis, an LTQ-Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific) equipped with a nanoelectrospray ionization source (Jamie Hill Instrument Services) was used. An Eksigent nanoLC system (Eksigent Technologies) equipped with a self-packed 100 µm inner diameter x 150 mm reverse phase column (Magic C18AQ, 5µm, New Objective) was coupled to the mass spectrometer. To analyze the LC-MS/MS data, the raw files were converted to mzXML format using MSConvert (ProteoWizard) and uploaded to XCMS online platform for analysis (www.xcmsonline.scripps.edu). Data were processed using XCMS's default settings. An in-house software platform was developed for automated metabolite identification using MS/MS data. We identified metabolites by matching experimental MS/MS spectra to a reference MS/MS library obtained from the Human Metabolome Database (HMDB). Identification and quantitation of metabolites were confirmed by manual inspection of raw MS data evidence.

Immunohistochemistry staining of human specimens

Unconjugated goat anti-human SOX2 antibody (R&D Systems Cat# AF2018 RRID: AB_355110), generated against a recombinant human SOX2, was used for immunohistochemistry. Chromogenic immunohistochemistry was performed on a Ventana Medical Systems Discovery XT instrument with online deparaffinization using Ventana's reagents and detection kits unless otherwise noted (Ventana Medical Systems Tucson, AZ USA). Four µm formalin-fixed, paraffin-embedded (FFPE) tissue sections were deparaffinized, rehydrated and epitope retrieved in a 1200-Watt microwave oven at 100% power in 10 mM sodium citrate buffer, pH 6.0 for 20 minutes. Sections were allowed to cool for 30 minutes and then rinsed in distilled water. SOX2 was diluted 1:100 in Dulbecco's phosphate buffered saline (Thermo Fischer Gibco Cat# 14190-136) and incubated for 10 hours at room temperature. Primary antibody was detected with biotinylated horse anti-goat, diluted 1:200 in Dulbecco's PBS (Vector Laboratories Cat# BA-9500 Lot# S0206 RRID: AB_2336123) and incubated for 30 minutes at 40° C. This was followed by application of streptavidin-horseradish-peroxidase conjugate. The complex was visualized with 3,3 diaminobenzidene and enhanced with copper sulfate. Slide were washed in distilled water, counterstained with hematoxylin,

dehydrated and mounted with permanent medium. Negative controls were incubated with Dulbecco's phosphate buffered saline instead of primary antibody.

Immunofluorescence microscopy of human specimens

For immunofluorescence staining of human tissues, unconjugated mouse anti-human Isocitrate Dehydrogenase-1 (IDH1, Dianova Cat# DIA-H09 Lot# 16829/04, RRID:AB_2335716) clone H09, raised against the human IDH-1 R132H point mutation and goat anti-human Sex Determining Region Y, Box 2 (SOX2, R&D Systems Cat# AF2018 Lot# RRID: AB_355110), generated against recombinant human SOX2, were used for immunofluorescence (Camelo-Piragua et al., 2010; Canham et al., 2014). Sequential duplex immunofluorescence was performed on a Ventana Medical Systems Discovery XT instrument using Ventana's reagents and detection kits unless otherwise noted (Ventana Medical Systems). In brief, 5 μm formalin-fixed, paraffin-embedded (FFPE) tissue sections were collected on plus slides and heated at 60° C in a convection oven for 1 hour. Slides were deparaffinized in xylene (3 changes), rehydrated through graded alcohols (3 changes 100% ethanol, 3 changes 95% ethanol) and rinsed in distilled water. Epitope retrieval was performed in a 1200-Watt microwave oven at 100% power in 10 mM sodium citrate buffer, pH 6.0 for 20 minutes. Sections were allowed to cool for 30 minutes and then rinsed in distilled water. Slides were blocked with 10% hydrogen peroxides for 15 minutes. Anti-R132H IDH1 antibody was diluted 1:50 in Ventana Discovery Diluent (Ventana Medical Systems Cat#760-018) and incubated for 3 hours at 37° C. It was detected with anti-mouse horseradish peroxidase-conjugated multimer incubated for 16 minutes and then visualized with tyramide-conjugated rhodamine. Slides were washed with instrument buffer for 5 minutes followed by application of anti-SOX2 antibody, diluted 1:50 in tris Buffered saline (25 mM Tris, 0.15 mM NaCL, pH7.2, Thermo-Fischer) and incubated for 10 hours at room temperature. SOX2 was detected with Alexa-Flour 488 conjugated donkey anti-goat antibody (Molecular Probes Cat# A-11055 Lot# 1463163 RRID: AB_142672) diluted 1:100 in TBS and incubated for 1 hour at 37° C. Completed slides were washed in soapy water followed by rinsing in distilled water. Slides were counterstained with 100 ng/ml DAPI and cover-slipped with Prolong Gold Anti-fade medium (Molecular Probes). Negative controls consisted of substituting primary antibody with diluent. Duplex negative control (crossover reactivity of donkey anti-goat against R132H IDH1) was verified by eliminating IDH1 anti-mouse multimer incubation during sequence 1 and substituting diluent during the primary SOX2 incubation of sequence 2. Negative controls showed no signal in either channel.

Karyotyping

NSC cultures with vector, 1 hit, 2 hits or 3 hits were treated with colcemid at final concentration of 0.1 μg/mL. Following 30-60 min incubation, cells were trypsinized according to standard procedures, incubated in 0.075M KCl at 37°C for 10 minutes and fixed in methanol-acetic acid (3:1). The fixed cell suspension was then dropped onto slides, stained in 0.08 μg/ml DAPI in 2xSSC (saline sodium citrate) buffer for 3 minutes and mounted in antifade solution (Vectashield, Vector Labs). The stained slides were scanned using a Zeiss Axioplan 2i epifluorescence microscope equipped with a megapixel CCD camera (CV-M4+CL, JAI) controlled by Isis 5.2 imaging software. For each sample, a minimum of 20 metaphases were captured. All metaphases were fully karyotyped and described as per International System for Human Cytogenetic Nomenclature (ISCN) 2013.

SUPPLEMENTAL RERERENCES

Assenov, Y., Muller, F., Lutsik, P., Walter, J., Lengauer, T., and Bock, C. (2014). Comprehensive analysis of DNA methylation data with RnBeads. Nat Methods *11*, 1138-1140.

Bayin, N. S., Modrek, A. S., Dietrich, A., Lebowitz, J., Abel, T., Song, H. R., Schober, M., Zagzag, D., Buchholz, C. J., Chao, M. V., and Placantonakis, D. G. (2014). Selective lentiviral gene delivery to CD133-expressing human glioblastoma stem cells. PloS one *9*, e116114.

Camelo-Piragua, S., Jansen, M., Ganguly, A., Kim, J. C., Louis, D. N., and Nutt, C. L. (2010). Mutant IDH1-specific immunohistochemistry distinguishes diffuse astrocytoma from astrocytosis. Acta Neuropathol *119*, 509-511.

Cancer Genome Atlas Research, N., Brat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., Cooper, L. A., Rheinbay, E., Miller, C. R., Vitucci, M.*, et al.* (2015). Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. The New England journal of medicine *372*, 2481-2498.

Canham, M., Charsou, C., Stewart, J., Moncur, S., Hoodless, L., Bhatia, R., Cong, D., Cubie, H., Busby-Earle, C., Williams, A.*, et al.* (2014). Increased cycling cell numbers and stem cell associated proteins as potential biomarkers for high grade human papillomavirus+ve pre-neoplastic cervical disease. PloS one *9*, e115379.

Conti, L., Pollard, S. M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q. L., Cattaneo, E., and Smith, A. (2005). Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. PLoS Biol *3*, e283.

Cook, B. D., Dynek, J. N., Chang, W., Shostak, G., and Smith, S. (2002). Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2 at human telomeres. Mol Cell Biol *22*, 332-342.

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W.*, et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N. D., Tabar, V., and Studer, L. (2008). Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev *22*, 152-165.

Flavahan, W. A., Drier, Y., Liau, B. B., Gillespie, S. M., Venteicher, A. S., Stemmer-Rachamimov, A. O., Suva, M. L., and Bernstein, B. E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature *529*, 110-114.

Hu, Y., and Smyth, G. K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J Immunol Methods *347*, 70-78.

Koch, P., Opitz, T., Steinbeck, J. A., Ladewig, J., and Brustle, O. (2009). A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. Proc Natl Acad Sci U S A *106*, 3225-3230.

Placantonakis, D. G., Tomishima, M. J., Lafaille, F., Desbordes, S. C., Jia, F., Socci, N. D., Viale, A., Lee, H., Harrison, N., Tabar, V., and Studer, L. (2009). BAC transgenesis in human embryonic stem cells as a novel tool to define the human neural lineage. Stem Cells *27*, 521-532.

Raviram, R., Rocha, P. P., Muller, C. L., Miraldi, E. R., Badri, S., Fu, Y., Swanzey, E., Proudhon, C., Snetkova, V., Bonneau, R., and Skok, J. A. (2016). 4C-ker: A Method to Reproducibly Identify Genome-Wide Interactions Captured by 4C-Seq Experiments. PLoS Comput Biol *12*, e1004780.

Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J.*, et al.* (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317-330.

Rocha, P. P., Raviram, R., Fu, Y., Kim, J., Luo, V. M., Aljoufi, A., Swanzey, E., Pasquarella, A., Balestrini, A., Miraldi, E. R.*, et al.* (2016). A Damage-Independent Role for 53BP1 that Impacts Break Order and Igh Architecture during Class Switch Recombination. Cell Rep *16*, 48-55.

Tabar, V., Panagiotakos, G., Greenberg, E. D., Chan, B. K., Sadelain, M., Gutin, P. H., and Studer, L. (2005). Migration and differentiation of neural precursors derived from human embryonic stem cells in the rat brain. Nat Biotechnol *23*, 601-606.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols *7*, 562-578.