EXTENDED EXPERIMENTAL PROCEDURES

Primary GBM Cell Cultures

GBM tumors were dissociated into single cell suspensions using a papain-based brain tumor dissociating kit (Miltenyi Biotec). Cells were then grown as gliomaspheres in serum-free neural stem cell medium as previously described (Wakimoto et al., 2009, 2012). From the same tumors, traditional GBM cells lines, grown as adherent monolayer in DMEM 10% FCS were derived. BMP4 induced differentiation was performed by addition of 100ng/ml of BMP4 (R&D) for 72 hr and removal of EGF and bFGF from medium (Piccirillo et al., 2006). A full description of our cellular model, including morphologic and genomic characterization, as well as differentiation assays has been published (Rheinbay et al., 2013; Wakimoto et al., 2009, 2012).

Immunofluorescence

Paraffin-embedded sectioned slides of human glioblastomas were deparaffinized and rehydrated according to standard protocols. Slides were blocked with 5% BSA for 2 hr followed by staining with directly conjugated antibodies (listed above) at 1:200 dilution in 5% BSA overnight at 4°C. Slides were imaged using an LSR710 scanning confocal microscope (Zeiss). Cells were fixed in 4% paraformaldehyde, permeabilized with 0.5% Triton X-100 (Sigma) and incubated at room temperature for two hours with antibodies for GFAP (R&D Systems, 1:200), mGalC (anti-Galactocerebroside, Millipore, 1:200), MAP2 (Cell Signaling Technology, 1:50), and Neuron Specific Beta-III Tubulin (Clone TuJ-1, R&D Systems, 1:200). Secondary antibodies: Alexa Fluor 536 Goat Anti-Rabbit (Invitrogen, 1:500), Alexa Fluor 488 Goat Anti-Mouse (Invitrogen, 1:500), or Alexa Fluor 546 Donkey Anti-Sheep (Invitrogen, 1:500). Coverslips were mounted with SlowFade Gold Antifade with DAPI (Invitrogen) and cells were visualized with an Olympus BX60 microscope.

Processing of ChIP-Seq Data

Read alignment to the hg19 reference genome, density map generation and peak calling for H3K27ac histone marks were performed as previously described (Rheinbay et al., 2013). Briefly, regions of enrichment were identified based on a 1 kb sliding window across the genome. An input experiment was used to account for copy-number variation in cancer genomes. Enriched windows were merged if the distance between them was less than 1kb. MACS (Zhang et al., 2008) was used to identify significant enrichment for transcription factor ChIP-Seq. For TF ChIP-Seq where two experiments were available (SOX2, OLIG2), high-confidence binding sites were identified as those that were present in both replicates. A peak was associated with a transcription start site (TSS) if it was present within 1.5 kb upstream or downstream of the TSS. IGV was used to generate and visualize ChIP-Seq density maps (Thorvaldsdóttir et al., 2013).

H3K27ac-Based Cell-Type Clustering

Regulatory sites enriched for H3K27ac in MGG4, 6, 8, TPCs and DGCs were collated into one comprehensive regulatory site "universe." Sites overlapping in one or more tumors were merged into a single site. Average H3K27ac density signal was calculated for each cell type with UCSC bigWigAverageOverBed. The distance metric between samples was calculated as One minus the pairwise Pearson correlation coefficient. Hierarchical clustering with complete linkage method was performed in R.

Motif Analyses

We used the HOMER software package (Heinz et al., 2010) to search for de novo enriched motifs. Comparison of de novo motifs with known motifs was performed with the Homer motif database augmented with motifs from (Jolma et al., 2013).

RNA Extraction and 3′ DGE RNA-Seq

To precisely quantify the gene expression, we used a 3′ DGE analysis pipeline. Briefly, to calculate expression values for each gene we scanned 500 basepair window within 10 kb of the annotated 3′ end of all genes, and counted reads that fell in the highest 500 basepair window across all libraries. To normalize across libraries we fit each individual libraries distribution of gene expression values into the same negative binomial distribution. Three replicates were acquired for each sample and condition. For comparative analyses, we used the edgeR package with general linear model (GLM) to identify differentially expressed genes between the three matched TPC/DGC pairs, and the MGG8 DGC empty (two replicates) and MGG8 POU3F2+SOX2+SALL2+OLIG2 iTPC isolated from mouse tumor (Robinson et al., 2010).

Regulatory Network Reconstruction

A list of "regulated genes" was defined as those genes that were at least 2-fold overexpressed in TPC over DGC and DGC empty plasmid control versus induced TPCs. Genes were assigned the smaller fold difference of the two comparisons. For each TF peak, a target was identified as a regulated gene within two gene loci up- and down-stream and 100 kb distance. In case where multiple genes fulfilled these criteria, the gene closest to the TF peak was chosen as presumed target. To eliminate spurious long-range association, we further removed all interactions between TFs and targets if all TF peak for this gene were located further than 50kb away from the TSS, so that only targets with at least one TF peak within 50kb, and possibly additional peaks within up to 100 kb remained. For the high-confidence stringent network displayed in Figure 6D, we additionally retained only protein-coding genes as

targets. A full list of targets, including non-coding RNAs and pseudogenes is included in Table S3. Cytoscape version 2.8.3 was used for visualization.

Immunoprecipitation and Western Blots

Immunoprecipitation (IP) using an antibody to SOX2 (R&D Systems) or RCOR2 (Abcam) was performed in 1.5 ml tubes with about 1 mg of protein, 2 mg of protein G Dynabeads (Lifetechnolgies) and 5ug of antibody for at least 4h at 4°C in the presence of protease inhibitors (Roche) and phosphatase inhibitors (Thermo Scientific) in a sample rotator. The beads were washed once with lysis buffer and twice with wash buffer then eluted in 1X sample buffer (Lifetechnologies) at 70°C for 10 min. Samples were then run on 4%–12% Bolt gels (Lifetechnologies) and transferred to PVDF membranes (BioRad). Western blots: membranes were blocked with Reliablot Block buffer (Bethyl) at 4°C and incubated with antibody to SALL2 (Bethyl) or LSD1 (Bethyl) overnight at 4°C. An HRP-linked secondary antibody (Bethyl) was incubated 4h at 4°C in Reliablot buffer. The membrane was then incubated for 1 min at room temperature with SpectraQuant-HRP CL reagent (BridgePath Scientific) and chemiluminescent images were collected on a BioRad ChemiDoc MP imaging system. The same general procedures was applied for Western blots with the following antibodies: SOX2 (R&D), OLIG2 (R&D), POU3F2 (Epitomics), SALL2 (Bethyl), SOX8 (Abcam), ASCL1 (Epitomics) and HEY2 (Abcam).

Overexpression and Reprogramming Experiments

Human cDNA for ASCL1, CITED1, HES6, HEY2, KLF15, OLIG1, OLIG2, POU3F2, RFX4, SALL2, SOX2 and SOX8 were cloned from GBM cells into a lentiviral plasmid (pLiV) and sequence verified. SOX1, SOX5, POU3F3, SOX21, and VAX2 were purchased (GeneCopoeia), as Gateway compatible pDONRvectors. Overexpression experiments were carried on the following way: GBM DGC were infected with cDNA expressing lentivirus; after 48 hr, the medium was changed to serum-free neural stem cells conditions and cells were monitored in those conditions for a 2-4 weeks period. Reprogramming experiments with 4 TFs were carried on stepwise and in a particular order as described in text, with each TF induction been separated by 2 weeks periods. For experiments using inducible constructs, corresponding cDNA were cloned into the pIND20 vector and induced with 0.1ug/ml doxycycline (Meerbrey et al., 2011).

SUPPLEMENTAL REFERENCES

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. Cell *152*, 327–339.

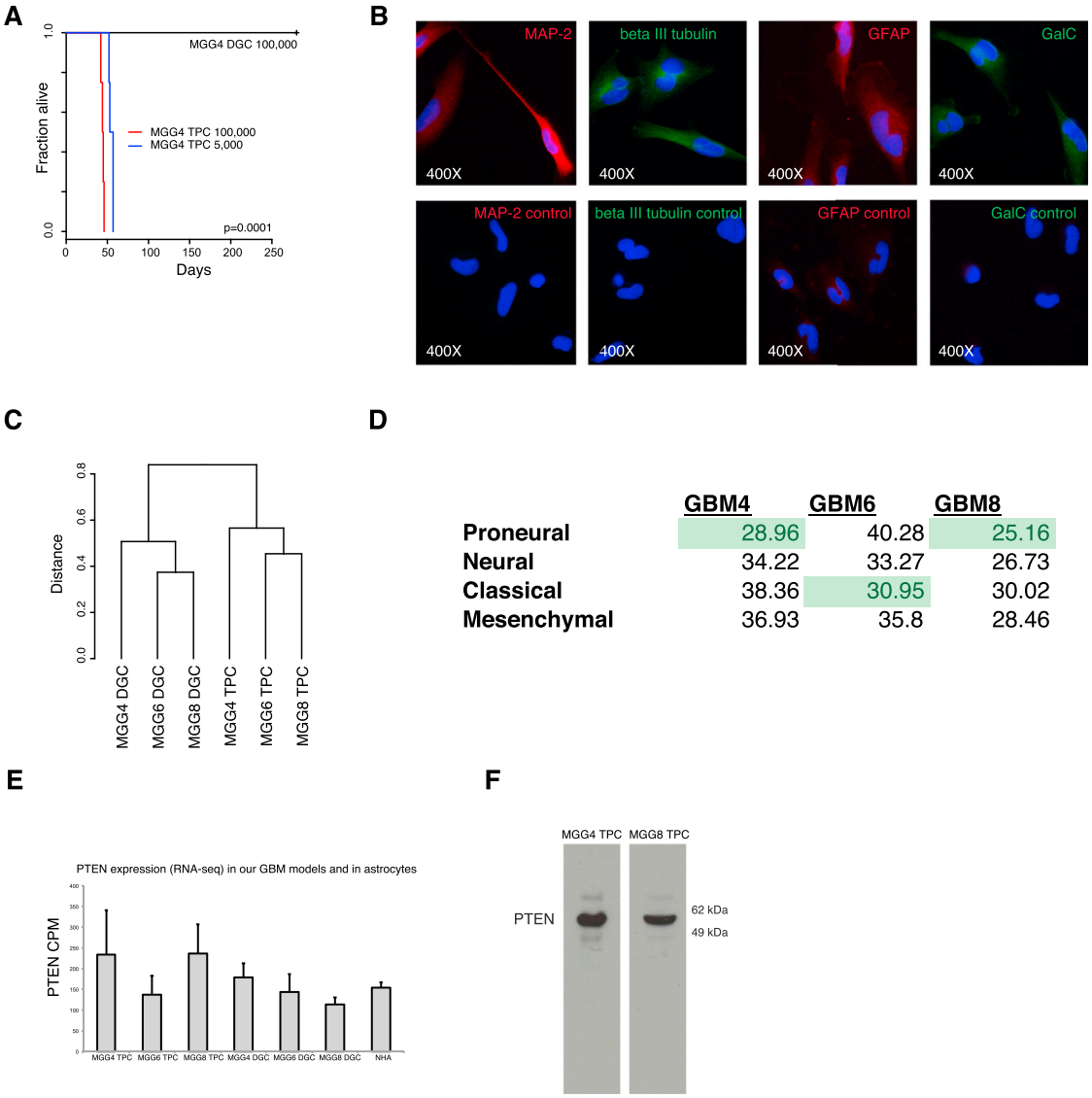
Meerbrey, K.L., Hu, G., Kessler, J.D., Roarty, K., Li, M.Z., Fang, J.E., Herschkowitz, J.I., Burrows, A.E., Ciccia, A., Sun, T., et al. (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc. Natl. Acad. Sci. USA *108*, 3665–3670.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Wakimoto, H., Mohapatra, G., Kanai, R., Curry, W.T., Jr., Yip, S., Nitta, M., Patel, A.P., Barnard, Z.R., Stemmer-Rachamimov, A.O., Louis, D.N., et al. (2012). Maintenance of primary tumor phenotype and genotype in glioblastoma stem cells. Neuro-oncol. *14*, 132–144.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.



	GBM4	GBM6	GBM8
Proneural	28.96	40.28	25.16
Neural	34.22	33.27	26.73
Classical	38.36	30.95	30.02
Mesenchymal	36.93	35.8	28.46

Figure S1. Additional Characterization of Our Models, Related to Figure 1

(A) GBM cells (MGG4) grown as gliomaspheres in serum-free conditions propagate tumor in vivo while serum-differentiated cells fail to do so.

(B) Serum-grown cells grow as adherent monolayers and express the differentiation markers GFAP (astroglial), beta III tubulin (neuronal), MAP-2 (neuronal) and GalC (oligodendroglial).

(C) Hierarchical clustering of H3K27ac ChIP-Seq signal separates GBM TPCs from DGCs (see methods).

(D) Distance of marker gene signature in TPCs to TCGA-defined centroids for each molecular subtype (Verhaak et al., 2010). Lower distance indicates greater similarity to respective subtype.

(E) Expression of PTEN by RNA-seq in our three matched lines of TPCs and DGCs shows expression levels comparable or higher to primary human astrocytes (NHA). Error bars indicate SEM based on three data points.

(F) Western blot for PTEN shows expression of the protein in MGG4 TPCs and MGG8 TPCs (Chen et al., 2010).

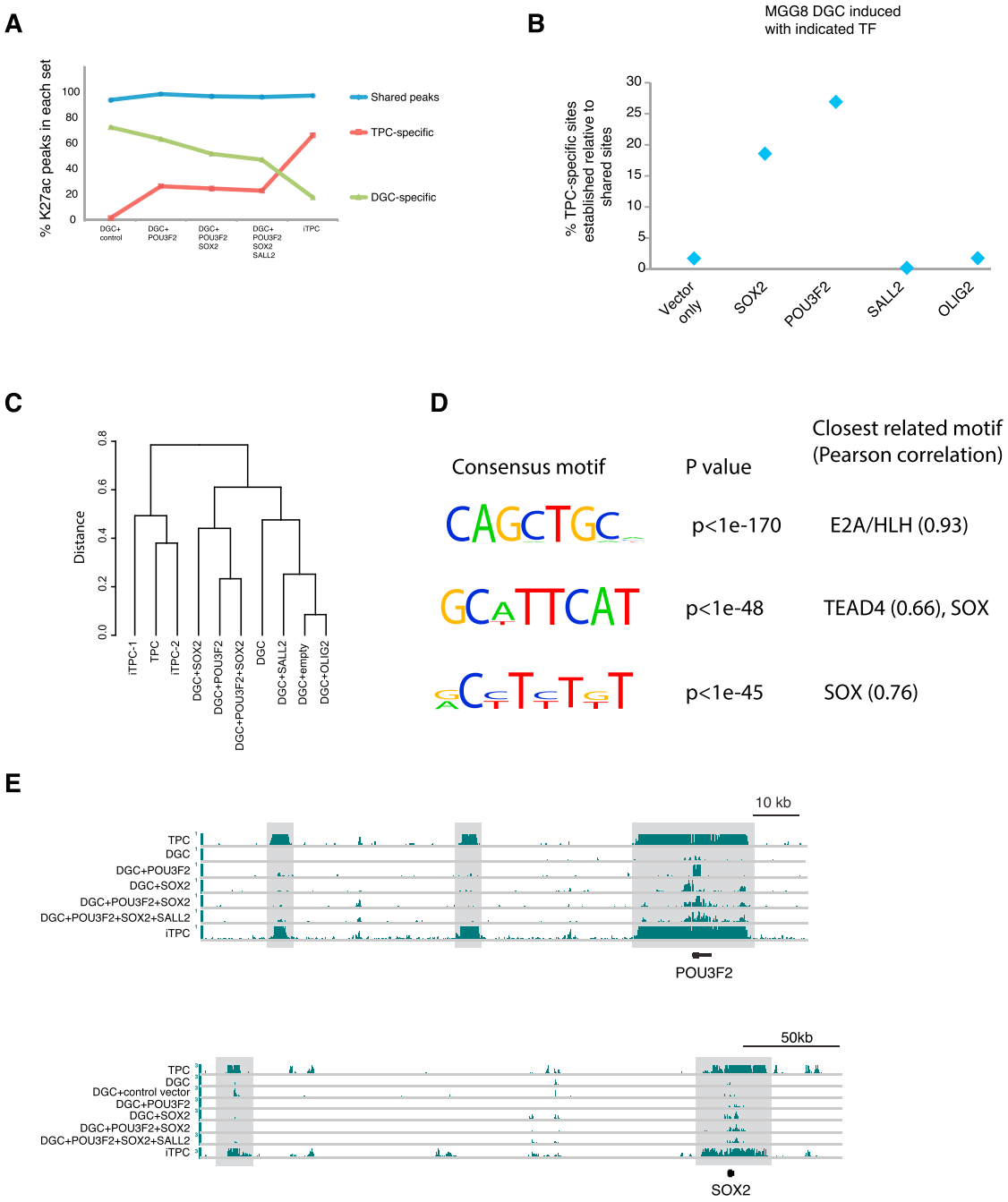


Figure S2. Reprogramming of the H3K27ac Epigenomic Landscape, Related to Figures 3 and 4

(A) Diagram shows percentage of H3K27ac peaks in the three sets of regulatory elements as defined in Figure 1E at different stages of reprogramming, showing a decrease of DGC specific and an increase of TPC-specific elements during reprogramming.

(B) Percentage of TPC-specific regulatory elements (relative to shared elements) that gain H3K27ac after single TF induction in DGCs. Only SOX2 and POU3F2 are capable of activating TPC-specific elements independently.

(C) Hierarchical clustering of H3K27ac ChIP-Seq tracks in MGG8 TPCs, DGCs and at different stages of reprogramming. iTPCs cluster with TPCs.

(D) De novo motif analysis of H3K27ac sites present in TPCs but not partially reprogrammed cells (POU3F2, SOX2, SALL2).shows enrichment of an HLH-class motif.

(E) Representative images of H3K27ac ChIP-Seq tracks at stages of reprogramming. SOX2 and POU3F2 are displayed as representative examples of loci that get activated during reprogramming.

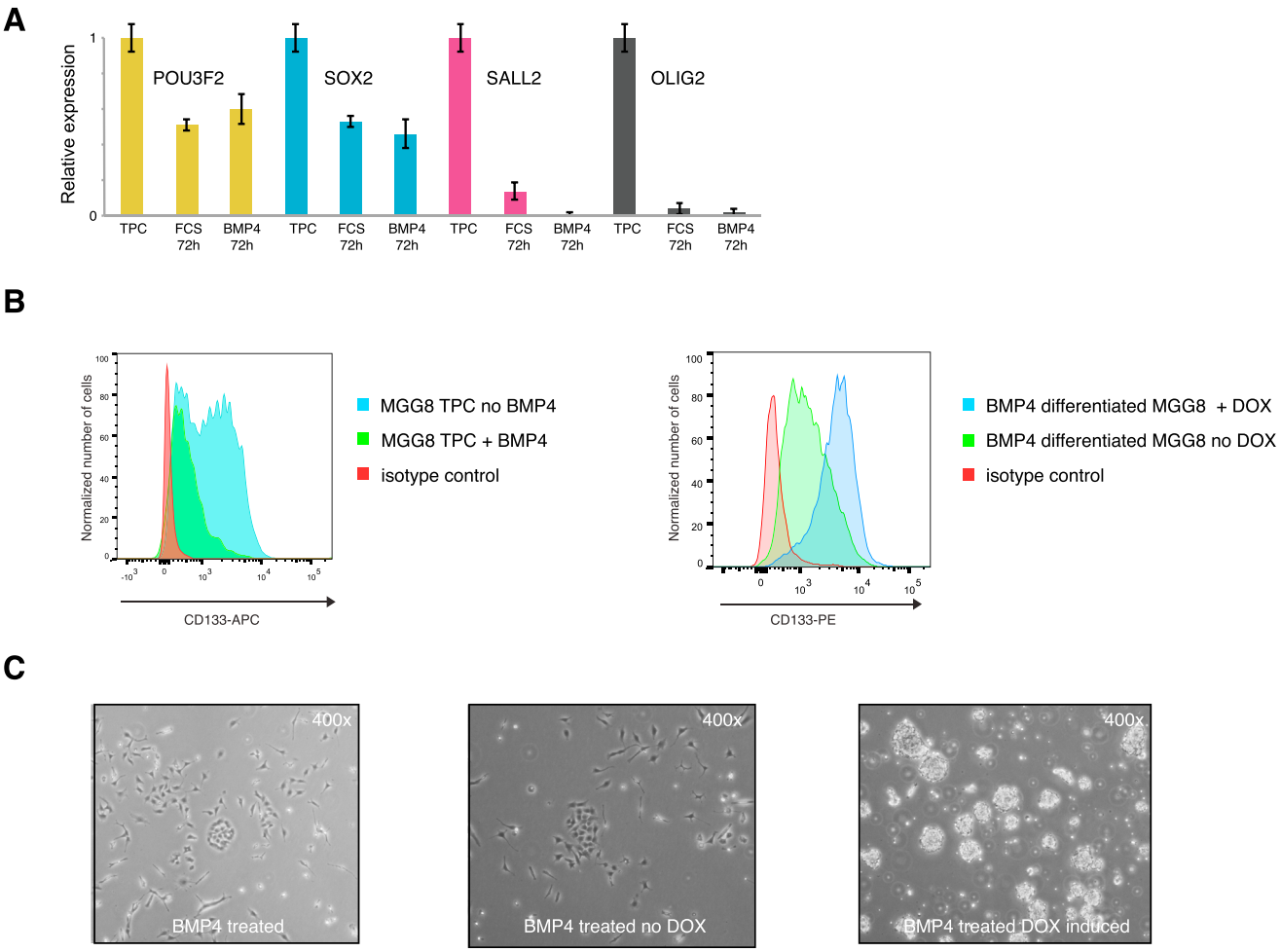


Figure S3. BMP4 Differentiation Downregulates Core TFs and Can Be Reversed by TF Induction, Related to Figure 3

(A) qRT-PCR measurements of mRNA for POU3F2, SOX2, OLIG2 and SALL2 in MGG8 TPCs, TPCs differentiated in serum for 72 hr (FCS 72h) and differentiated with BMP4 for 72 hr (BMP4 72h). Error bars indicate SEM based on three data points.

(B) Left: flow cytometry for CD133/isotype control in MGG8 TPC control or treated with BMP4. Right: flow cytometry for CD133/isotype control of BMP4-differentiated MGG8 TPCs infected with inducible lentiviruses encoding POU3F2, SOX2, OLIG2 and SALL2. Induction by doxycycline results in higher CD133 expression.

(C) Left: BMP4-differentiated MGG8 TPCs rapidly adhere and differentiate, as previously reported. Middle and Right: BMP4-differentiated MGG8 TPCs infected with inducible lentiviruses encoding POU3F2, SOX2, OLIG2 and SALL2 cultured in the absence or presence of doxycycline. Induction of TF expression generates spheres in vitro. Collectively, these data support a general role for these TFs in the stemness of GBM cells responding to different differentiation stimuli.

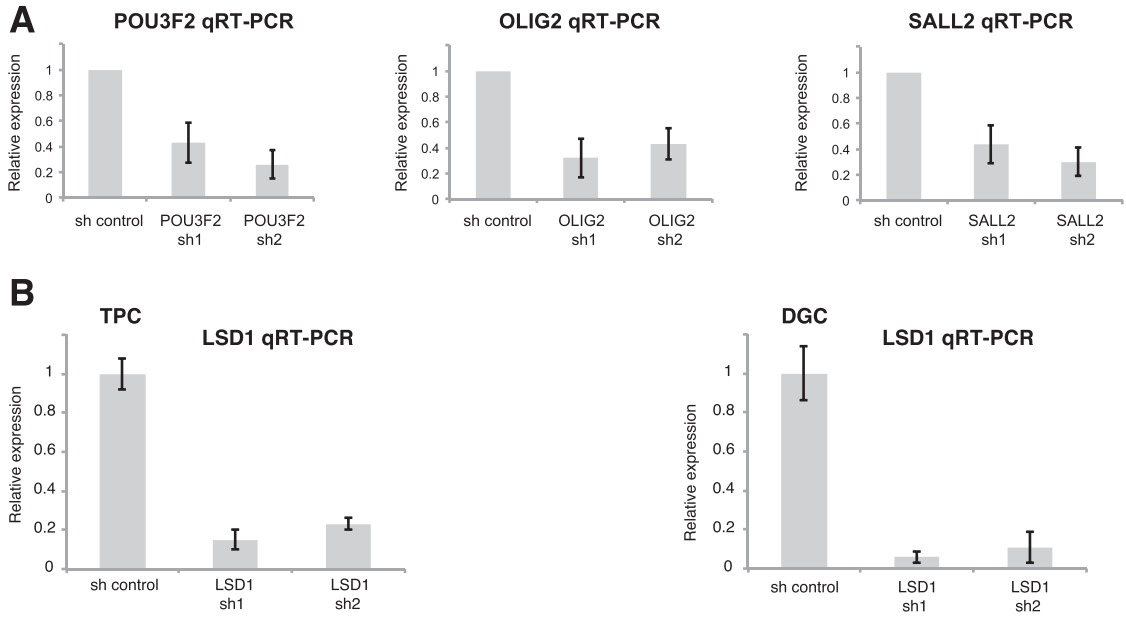


Figure S4. qRT-PCR Measurements of shRNA Knockdowns, Related to Figures 3 and 7

(A) qRT-PCR measurements of mRNA for POU3F2, OLIG2 and SALL2 in MGG4 TPCs infected with control lentivirus shRNA or with hairpins specifically targeting the corresponding mRNA, showing downregulation of each TF with two different hairpins. Error bars indicate SEM based on three data points.

(B) qRT-PCR measurements of mRNA for LSD1 in MGG4 TPCs and DGCs infected with control lentivirus shRNA or with hairpins specifically targeting LSD1, showing similar downregulation in TPCs and DGCs with two different hairpins.

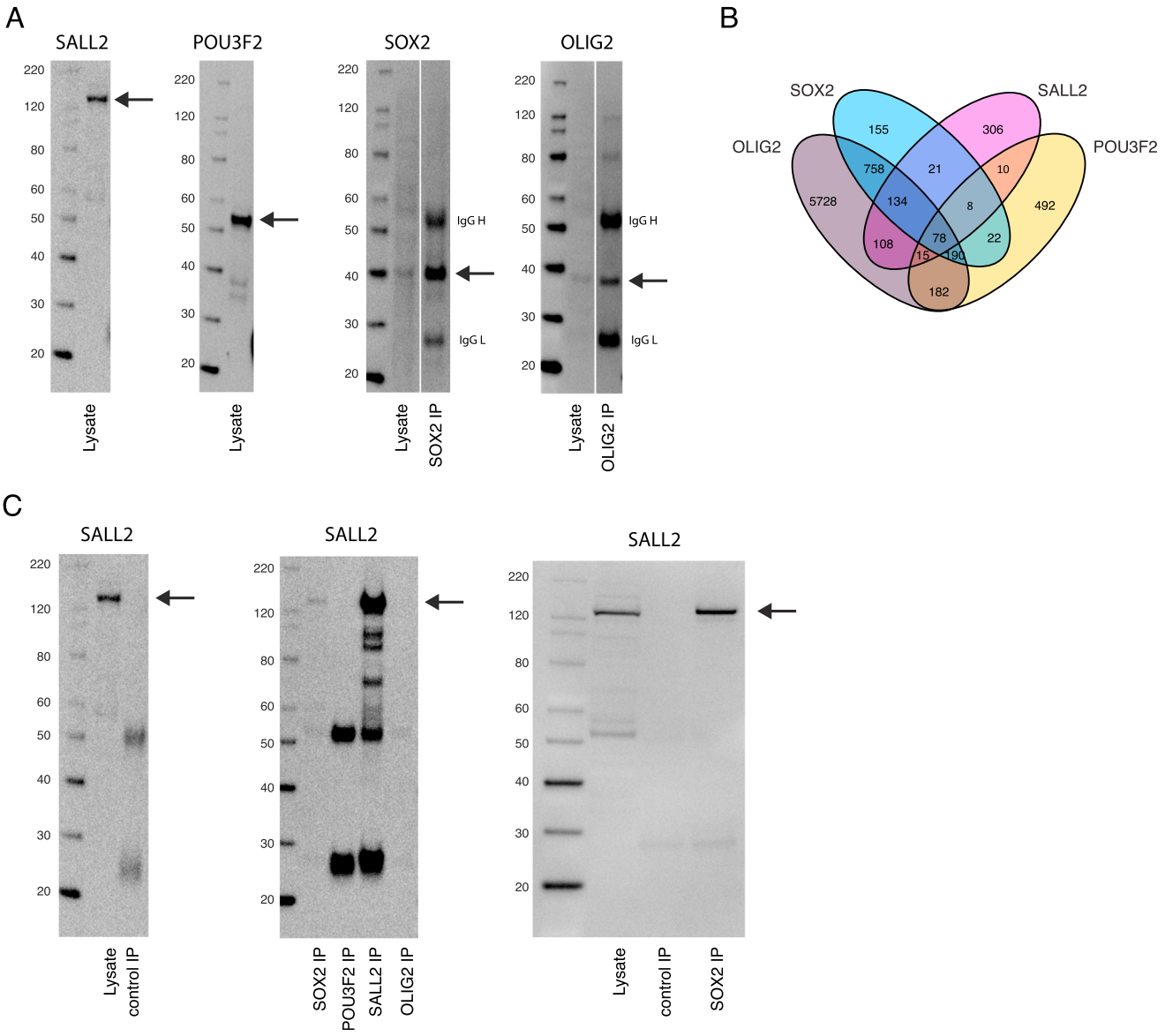


Figure S5. Western Blots, Immunoprecipitation, and Coimmunoprecipitation Assays of SOX2, POU3F2, SALL2, and OLIG2 in GBM TPC, Related to Figure 6

(A) Western blot and immunoprecipitation experiments using MGG8 TPC lysates show specificity of the antibodies for their corresponding TF. Arrow points to corresponding TF band.

(B) Venn diagram depicts numbers of TF peaks at regulatory elements and overlap among these sites.

(C) Western blot for SALL2 on MGG8 TPC lysate and after immunoprecipitation (control IgG, SOX2 IP, SALL2 IP, POU3F2 IP and OLIG2 IP) highlights interaction between SALL2 and SOX2 in GBM TPCs.