Supplemental Figures, Tables, and Data
Figure S1, Related to Figure 1

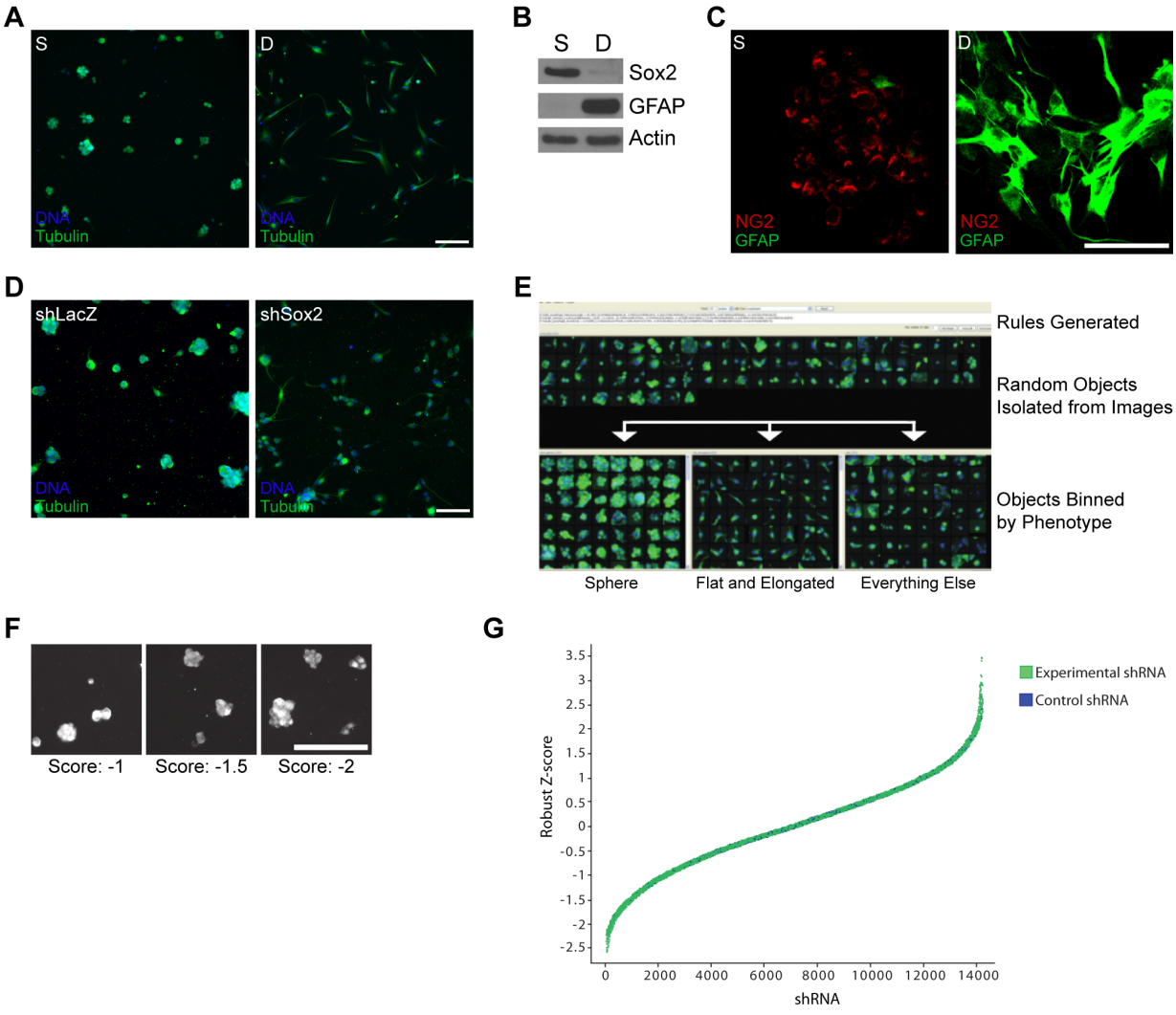

Figure S1. Image analysis of the screen quantifies morphologic phenotypes associated with forced human GBM TIC differentiation and identifies gene hits, Related to Figure 1. **(A)** Immunofluorescent images of 0308 glioma TICs grown in stem cell-promoting (S) conditions [NBE media(Lee et al., 2006)] or differentiated (D) for 1 week in serum-containing media demonstrate a shift from neurosphere formation to flat, elongated morphology upon forced differentiation. **(B)** Immunoblots of 0308 TICs grown in stem-cell or differentiating conditions show decreased expression of the stem cell marker Sox2 and increased expression of the astrocyte marker GFAP upon differentiation. **(C)** Immunofluorescent images of 0308 TICs grown in stem-cell or differentiating conditions show increased expression of GFAP and decreased expression of the stem cell marker NG2 upon differentiation. **(D)** Immunofluorescent images of 0308 TICs transduced with a control shRNA targeting *LacZ* or an shRNA targeting *SOX2* show that suppression of Sox2 leads to a shift from neurosphere formation to flat, elongated morphology, similar to that observed upon serum-mediated differentiation. **(E)** Using CellProfiler Analyst iterative machine learning to analyze phenotypes. After objects were identified and parsed from shRNA screen images using CellProfiler, several dozen objects were randomly selected, and then manually binned into one of 3 categories: Sphere, Flat and Elongated, or Everything Else. Based on the 147 measurements made on every object, and iterative training rounds, the software identified rules that distinguished between different morphologies. Rules were applied to every object identified by CellProfiler in every image from the loss of function screen in order to identify wells enriched for the differentiated phenotype. **(F)** Representative screen images with associated negative enrichment scores from the primary screen. **(G)** Robust Z-scores of relative cell number by gene-targeting shRNAs and control shRNAs. Gene-level hits were defined as those having at least two different shRNAs having robust Z-scores < -1.5. **(H)** Representative immunofluorescent images from the primary RNAi screen of 0308 glioma TICs, for shRNAs targeting the indicated genes. Stem cell-like neurosphere morphology is seen in TICs transduced with a control shRNA targeting *LacZ*, whereas a differentiated flat, elongated morphology is evident in TICs transduced with shRNAs targeting *NOTCH2*, *ZNF143*, *ING2*, *SALL3*, and *ZFHX4*. The gene-level scores from the morphology-based confirmation screen (Morphology score) and the relative cell number screen (Relative cell number score) are shown for each gene as indicated. Scale bars: **A, C, D,** and **H**, 50μm; **F**, 100μm.

Figure S2, Related to Figure 2

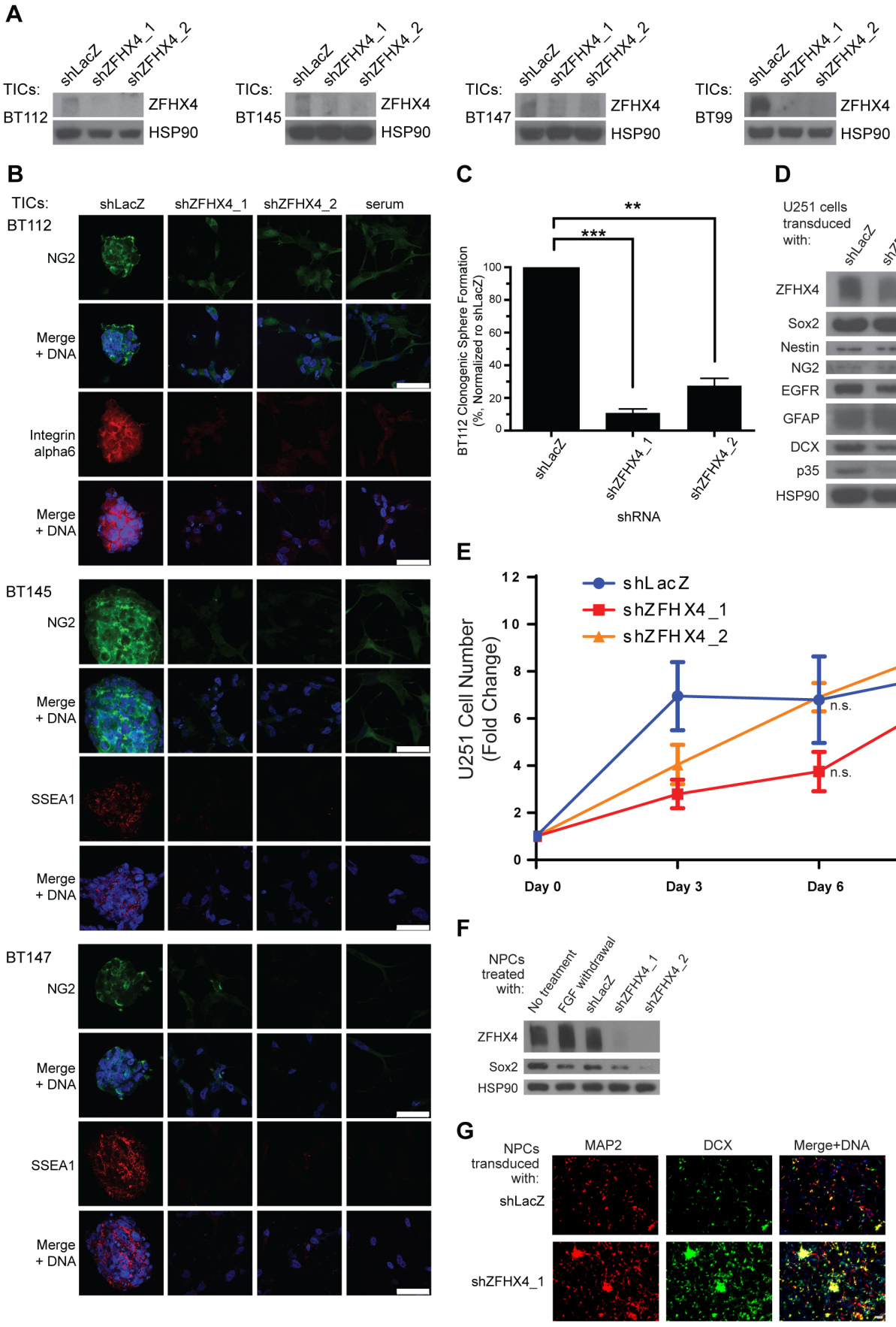

Figure S2. ZFHX4 suppression causes loss of stem cell-like features in human GBM TIC lines and normal NPCs but does not affect non-stem-like GBM cell lines, Related to Figure 2. **(A)** Immunoblots show depletion of ZFHX4 protein after transduction with both anti-*ZFHX4* shRNAs in BT112, BT145, BT147, and BT99 TIC lines. **(B)** Immunofluorescent images showing downregulation of the stem cell markers NG2 and either Integrin alpha6 or SSEA1 in the three TIC lines upon ZFHX4 suppression. Scale bars: 50μm. **(C)** ZFHX4 suppression in BT112 TICs decreases clonogenic sphere formation. *** p=0.0002 shLACZ vs. shZFHX4_1; ** p=0.0013 shLacZ vs. shZFHX4_2, Fisher's exact two-tailed t-tests. Error bars reflect standard deviations. n=3. **(D)** Immunoblots showing no decrease in expression of stem cell markers Sox2, Nestin, and NG2 or the GBM oncogene EGFR and no increase in expression of the astrocyte marker GFAP or the neuronal markers DCX and p35 in U251 cells in which ZFHX4 was suppressed. We also observed no significant effects of ZFHX4 suppression on proliferation or marker expression in the LNZ308 GBM cell line (data not shown). **(E)** ZFHX4 suppression in the U251 GBM cell line does not affect cell growth. n.s.: not significant; for fold change at day 6, p=0.1862 shLacZ vs. shZFHX4_1, p=0.9496 shLacZ vs. shZFHX4_2; for fold change at day 9, p=0.7473 shLacZ vs. shZFHX4_1, p=0.6090 shLacZ vs. shZFHX4_2, two-tailed t-tests. Error bars reflect standard deviations; n=3. **(F)** Immunoblots showing depletion of ZFHX4 and decreased expression of the stem cell marker Sox2 in human NPCs transduced with *ZFHX4*-targeting shRNAs. FGF withdrawal from the media served as a positive control for differentiation. **(G)** Immunofluorescent images showing increased expression of neuronal markers MAP2 and DCX in NPCs transduced with one *ZFHX4*-targeting shRNA. Scale bar: 100μm.

Figure S3, Related to Figure 3

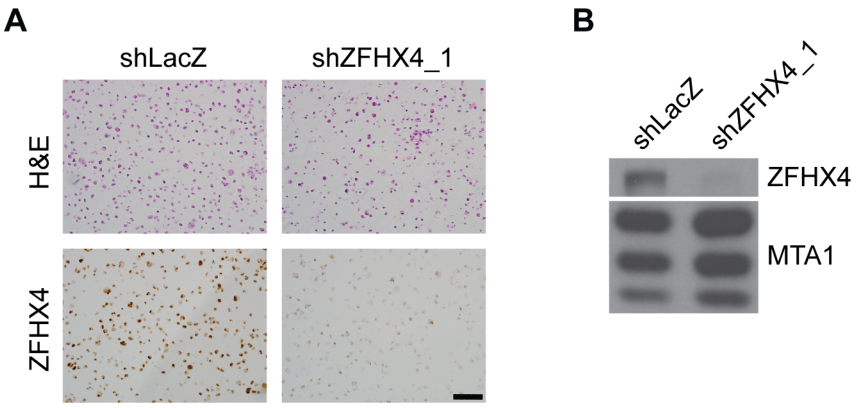

Figure S3. Validation of the ZFHX4 antibody used for immunohistochemical staining, Related to Figure 3. **(A)** Immunohistochemical (IHC) images of H&E and ZFHX4 staining of HistoGel (Thermo Scientific) embedded 0308 TICs transduced with a control *LacZ*-targeting shRNA or a *ZFHX4*-targeting shRNA, showing decreased staining in ZFHX4-depleted TICs by the rabbit anti-ZFHX4 antibody (Sigma) used for IHC staining of brain tissue, thus confirming that ZFHX4 is the target of the antibody. Scale bar: 100 μm. **(B)** Immunoblot of the same TICs transduced with the indicated shRNAs, confirming *ZFHX4* silencing.

Figure S4, Related to Figure 4

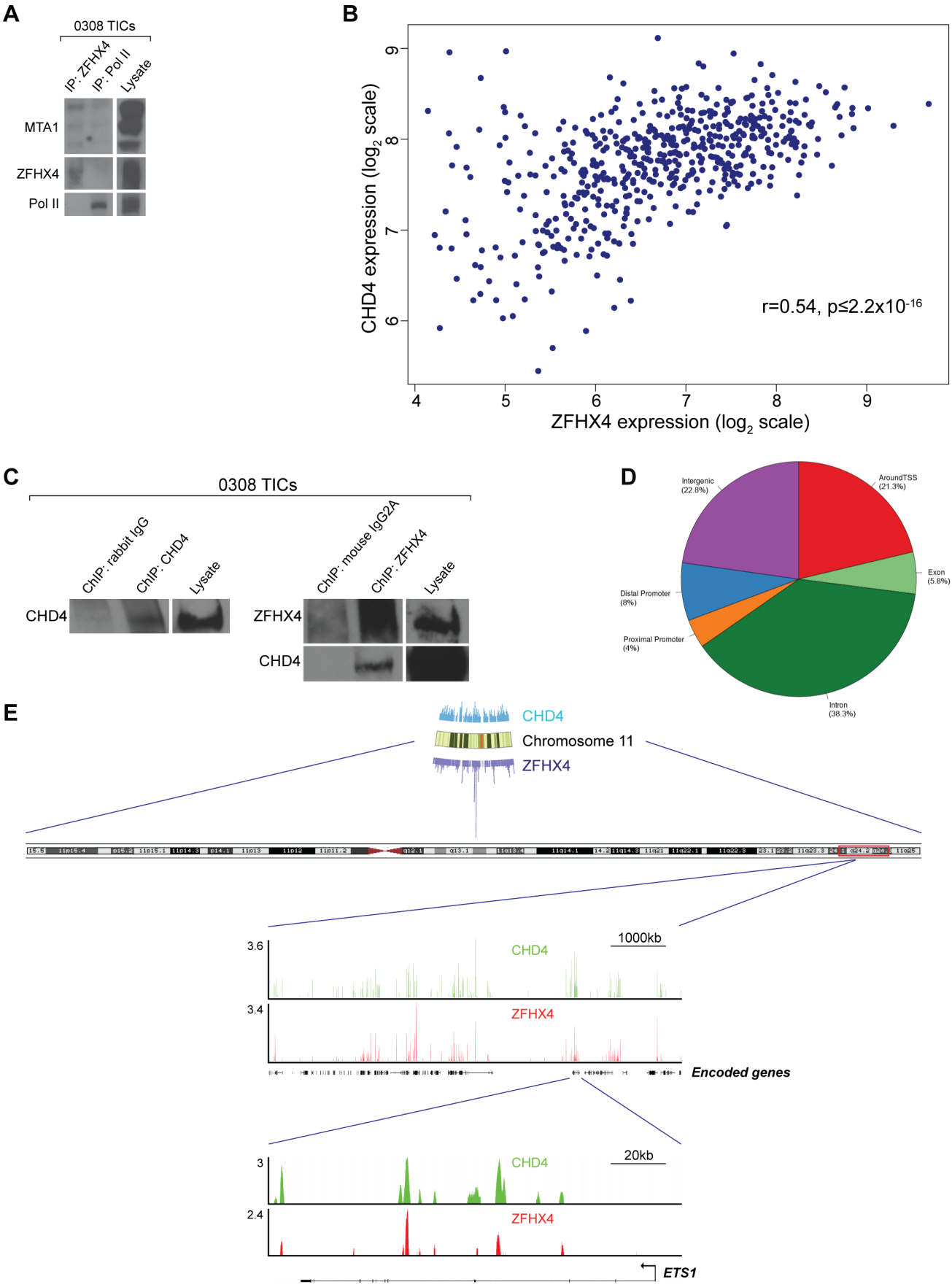

Figure S4. ZFHX4 interacts with NuRD complex component MTA1, and ZFHX4 and CHD4 correlate in expression in GBM patient samples and co-localize to sites throughout the genome, mainly in or near gene-coding regions, Related to Figure 4. **(A)** Immunoblots of endogenous ZFHX4 and RNA Polymerase II (Pol II) immunoprecipitates (IPs) and lysate from 0308 TICs, demonstrating modest but detectable enrichment of MTA1, a known member of the NuRD complex (Lai and Wade, 2011; Ramirez and Hagman, 2009), in the ZFHX4 IP. **(B)** Co-expression of ZFHX4 and CHD4 (at the RNA level) in 543 GBM samples from The Cancer Genome Atlas (TCGA (TCGA, 2008)). Scatterplot demonstrates a high and significant concordance of expression of ZFHX4 and CHD4 in primary GBM patient samples ($r=0.54$, $p \leq 2.2 \times 10^{-16}$, Pearson correlation test). **(C)** Immunoblots of endogenous CHD4, ZFHX4, and control chromatin immunoprecipitates (ChIPs) and lysates from 0308 TICs. **(D)** Relative frequencies of different genomic regions co-occupied by ZFHX4 and CHD4, based on ChIP with high-throughput DNA sequencing (ChIP-Seq) in 0308 TICs. TSS: transcription start site, position 0; transcription proceeds toward increasing numbers. AroundTSS, -250 base pairs (bp) to +250bp; Proximal Promoter, -500bp to -250bp; Distal Promoter, -10,000bp to -500bp; Intergenic: <-10,000bp; Exon and Intron, based on definitions of RefSeq genes (see Extended Experimental Procedures). **(E)** Top: section of whole-genome plot of ChIP-Seq mapped reads (Figure 4F) showing chromosome 11; middle and bottom: ChIP-Seq binding profiles (reads per million) for CHD4 and ZFHX4 at a section of chromosome 11 q24.1-q24.3 (note predominance of peaks in and near gene-coding regions) and at the *ETS1* locus, respectively, with the y-axis floor set to 1.

Table S1. Gene hits from screens for differentiated morphology and relative cell number, Related to Figure 1. (**Sheet 1, MorphologyHits**) Confirmed gene hits whose suppression alters TIC morphology. Confirmed genes were defined as those that scored at least as strongly as the serum-differentiated positive controls in the secondary screen, based on enrichment scores for the flat, elongated phenotype and ranking by the second-best scoring shRNA. Confirmed hits that are also found among the genes whose suppression decreases relative cell number (Sheet 2) are italicized. (**Sheet2, RelativeCellNumberHits**) Gene hits whose suppression decreases relative cell number. Relative cell number was measured using CellTiter-Glo Luminescent Assay (Promega), and gene hits were defined as those for which at least two different shRNAs had a robust Z-score < -1.5. Gene hits that are also found among the confirmed genes whose suppression alters TIC morphology (Sheet 1) are italicized.

Table S2. ZFHX4 and CHD4 physically interact, co-regulate a large-scale gene expression program, and co-localize to regions throughout the genome, Related to Figure 4. (**Sheet1, FLAG-ZFHX4_IP-mass_spec**) Most abundant proteins that specifically co-immunoprecipitate with FLAG-tagged ZFHX4 in HEK 293T cells, as identified by mass spectroscopy. (**Sheet2, GenesUp_shZFHX4&shCHD4**) Intersection of genes upregulated upon ZFHX4 or CHD4 suppression at both time points, including genes co-bound by ZFHX4 and CHD4 in ChIP-Seq. **** Gene identified in ChIP-Seq as co-bound by ZFHX4 and CHD4 (enrichment $p=1.5x10^{-13}$, chi-square test). (**Sheet3, GenesDn_shZFHX4&shCHD4**) Intersection of genes downregulated upon ZFHX4 or CHD4 suppression at both time points, including genes co-bound by ZFHX4 and CHD4 in ChIP-Seq. **** Gene identified in ChIP-Seq as co-bound by ZFHX4 and CHD4 (enrichment $p \leq 2.2x10^{-16}$, chi-square test). (**Sheet4, enriched_regions_ZFHX4**) Genomic regions enriched in ZFHX4 ChIP-Seq in 0308 TICs. (**Sheet5, enriched_regions_CHD4**) Genomic regions enriched in CHD4 ChIP-Seq in 0308 TICs. (**Sheet6,ZFHX4boundRegCobndByCHD4**) ZFHX4-occupied genomic regions (enriched in ZFHX4 ChIP-Seq) overlapped by at least 1bp of a CHD4-occupied region. (**Sheet7,CHD4boundRegCobndByZFHX4**) CHD4-occupied genomic regions (enriched in CHD4 ChIP-Seq) overlapped by at least 1bp of a ZFHX4-occupied region. (**Sheet8, enriched_genes_ZFHX4**) Genes enriched in ZFHX4 ChIP-Seq in 0308 TICs. Enriched gene defined as containing a ZFHX4-occupied region within the gene body or within 10kb of the start or stop of the gene. Note: for genes containing more than one ZFHX4-occupied region, only one region is shown here. (**Sheet9, enriched_genes_CHD4**) Genes enriched in CHD4 ChIP-Seq in 0308 TICs. Enriched gene defined as containing a CHD4-occupied region within the gene body or within 10kb of the start or stop of the gene. Note: for genes containing more than one CHD4-occupied region, only one region is shown here. (**Sheet10, co-bound_genes**) Co-bound genes. Co-bound genes defined as containing, within the gene body or within 10kb of the start or stop of the gene, a ZFHX4-occupied region overlapped by at least 1bp of a CHD4-occupied region.

Table S3. Candidate transcriptional targets of ZFHX4 identified using ARACNe, including those targets enriched among genes differentially expressed in CHD4-suppressed 0308 TICs, identified using GSEA, Related to Figure 4. ** Candidate ZFHX4 target that is part of a gene set significantly enriched among genes found to be differentially expressed in transcriptional profiling analysis of CHD4-depleted 0308 TICs (both shRNAs, both time points; enrichment p=0.002, non-parametric test with sample shuffling).

Data S1. Genome-wide binding tracks for ZFHX4 and CHD4 from ChIP-Seq in 0308 TICs (compressed ZIP file containing 2 binding track files in WIG format, which are viewable in the UCSC Genome Browser, http://genome.ucsc.edu/), Related to Figure 4.

Extended Experimental Procedures

Materials

Materials were obtained from the following sources: antibodies – anti-Sox2 (for immunoblotting) and anti-NG2 from R&D Systems; anti-GFAP from Dako; anti-MAP2 from Encor; anti-Actin, anti-Tubulin, anti-FLAG tag, and anti-ZFHX4 (for immunohistochemistry [IHC]) from Sigma; anti-Notch2 and anti-ZFHX4 (catalog #H00079776-M02 for immunoblotting, catalog #H00079776-M11 for immunoprecipitation) from Abnova; anti-Sox2 (for immunofluorescence), anti-HSP90, anti-EGFR, anti-Raptor, and anti-SSEA1 from Cell Signaling Technologies; anti-Sox2 (for IHC), anti-Nestin, anti-DCX, and anti-Integrin alpha6 from Millipore; anti-CHD4 from Abcam; anti-ING2, anti-SALL3, anti-ZNF143, anti-MTA1, anti-p35, anti-phospho-Histone H3, and anti-RNA Polymerase II from Santa-Cruz Biotechnology; anti-Ki-67 from Leica; anti-Olig2 as a gift from Charles Stiles; HRP-conjugated anti-mouse, anti-rabbit, anti-goat, and anti-guinea pig secondary antibodies from Santa-Cruz Biotechnology; Alexa Fluor-conjugated anti-mouse, anti-rabbit, anti-rat, and anti-guinea pig secondary antibodies, Hoechst 33342 reagent, Natural Mouse Laminin Protein, and Protein G Dynabeads from Invitrogen; Poly-D-Lysine coated 384-well and 6-well tissue culture plates from Becton Dickinson; Ultra-Low attachment cell culture flasks from Corning; Neurobasal medium, N-2 and B-27 supplements, Glutamax, and Albumax from Invitrogen; Neurocult NS-A Proliferation Kit (human) and mTeSR1 medium from STEMCELL Technologies; recombinant human EGF and bFGF from R&D Systems; accutase from Innovative Cell Technologies; Propidium Iodide, lactate, and biotin from Sigma; Neuroplex N2 and NS21 (-VitA) supplements from Gemini Bio; Dorsomorphin from Stemgent; CellTiter-Glo Luminescent Assay from Promega; HistoGel from Thermo Scientific; formalin from VWR; paraformaldehyde from Electron Microscopy Sciences; Borg Decloaker RTU solution and pressurized Decloaking Chamber from Biocare Medical; Prolong Gold Antifade reagent from Invitrogen. Dulbecco's Modified Eagle Medium (DMEM), fetal bovine serum (FBS), L-glutamine, Penicillin-Streptomycin (P-S) solution, and trypsin were provided by the Whitehead Institute; polybrene and puromycin were provided by the Broad Institute. Lentivirally delivered shRNAs for the primary and secondary screen as well as the follow-up studies were obtained from The RNAi Consortium (TRC) collection of the Broad Institute (Moffat et al., 2006). The following shRNA clones were used: shZFHX4_1: TRCN0000015182, hairpin sequence: 5'-CCGG-GCAACGAATGTGCCACTTCTT-CTCGAG-AAGAAGTGGCACATTCGTTGC-TTTTT-3', target sequence: GCAACGAATGTGCCACTTCTT; shZFHX4_2: TRCN0000015181, hairpin sequence: 5'-CCGG-CGGACAGTTCTTGCCATACTT-CTCGAG-AAGTATGGCAAGAACTGTCCG-TTTTT-3',target sequence: CGGACAGTTCTTG CCATACTT; shCHD4_1: TRCN0000021363, hairpin sequence: 5'-CCGG-GCGGGAGTTCAGTACCAATAA-CTCGAG-TTATTGGTACTGAACTCCCGC-TTTTT-3' target sequence: GCGGGAGTTCAGTACCAATAA; shCHD4_2: TRCN0000021361, hairpin sequence: 5'-CCGG-GCTGCTGACATCCTATGAATT-CTCGAG-AATTCATAGGATGTCAGCAGC-TTTTT-3', target sequence: GCTGCTGACATCCTATGAATT; shLACZ: TRCN0000072235, hairpin sequence: 5'-CCGG-CCGTCATAGCGATAACGAGTT-CTCGAG-AACTCGTTATCGCTATGAC GG-TTTTTG-3'; target sequence: CCGTCATAGCGATAACGAGTT.

Cell lines and tissue culture

The previously established 0308 GBM TIC line (Lee et al., 2006) was cultured in NBE medium (Neurobasal medium containing N-2 and B-27 supplements, EGF, bFGF, L-glutamine, and Penicillin-Streptomycin) as described (Lee et al., 2006). For establishment of BT112 (Mehta et al., 2011), BT145, and BT147 GBM TIC lines, GBM samples were acquired from patients undergoing neurosurgery at the Brigham and Women's Hospital on an IRB approved protocol as part of the DF/BWCC Living Tissue Bank program. Primary tissue was collected from surgery specimens and dissociated using the Miltenyi Biotec gentleMACS Dissociator and the Miltenyi Neural Tissue Dissociation Kit (P) containing papain but without further selection. Cells were then plated at a density of roughly 0.5-1x10^6 cells in 75 cm^2 Ultra-Low attachment cell culture flasks in NeuroCult NS-A Proliferation Kit (human) supplemented with Heparin (2mg/ml), human EGF and bFGF (20ng/ml each), and P-S. After establishment, the cells were cultured in the same medium on regular cell culture flasks. For differentiation experiments, TICs were transferred to DMEM containing 10% FBS and P-S (Lee et al., 2006). U251 and LNZ308 GBM cell lines were also cultured in DMEM containing 10% FBS and P-S. HEK 293T cells were cultured and used as described (Sancak et al., 2010).

Human neural precursor cells (NPCs) were derived from embryonic stem cell (ESC) line WIBR3 (Lengner et al., 2010) as follows. ESC cultures maintained on feeders were single-cell dissociated in PBS and plated on matrigel in mTeSR1 medium at a density of 10^5 cells/cm^2 (day 0). After 24 hours, the medium was changed to NBd [Neurobasal medium supplemented with pyruvate (1mM), Glutamax, lactate, biotin, Albumax (0.2%), Neuroplex N2 (0.5X) and NS21 (-VitA) (0.5X), adjusted to an osmolarity of 300 mOsm with NaCl]. On day 2, NBd was supplemented with 2.5uM Dorsomorphin, providing TGFβ inhibition, forcing neuro-ectodermal differentiation at the expense of other lineages. The medium was changed daily, using up to 10mL of medium in a 35mm dish (cells at this stage have very high metabolic needs resulting in lactic acidosis). On day 4, neural rosette structures emerged in the layers of differentiated cells covering the original pluripotent monolayer. By day 7-10, rosettes were confluent in the well, and Pax6+ cells represented >90% of the culture. These early NPCs were dissociated in PBS, taking care to avoid complete dissociation, in order to maintain seeding of rosettes structures. Dorsomorphin was removed at this stage. These first passages followed a split ratio of no more than 1:2. Around day 15, NBd was switched to a maintenance medium (NBm): Neuroplex N2 and NS21 were adjusted to 1X, and bFGF (10ng/mL) was added. NPCs obtained through this method were FGF-dependent and multipotent. To serve as a positive control for differentiation, NPCs were switched to medium containing a lower concentration of insulin and no bFGF.

Expression constructs

The pRK5-FLAG-Raptor plasmid has been described (Sancak et al., 2010). The pcZeo-FLAG-ZFHX4 plasmid was generated by constructing pHH207.5, which contained the entire coding region of ZFHX4, using cDNA plasmids (Hemmi et al., 2006). Then, FLAG-tagged full length ZFHX4 was generated using PCR and subcloning into pENTR11 vector (Invitrogen) and inserting it into pcDNA3.1-Zeo (Invitrogen) using the Gateway system (Invitrogen), yielding pcZeo-FLAG-ZFHX4.

Primary and secondary RNAi screens

In the primary screen, 0308 TICs were seeded in Poly-D-Lysine coated 384-well plates (500 cells / well) and transduced in 6 replicates with a library of 11,940 pLKO.1 lentiviral plasmid-based shRNAs (Moffat et al., 2006; Stewart et al., 2003), including 11,816 shRNAs targeting 2372 genes encoding all known and putative human transcription factors, chromatin-binding proteins, GTPases, and members of the PI3K-AKT-mTOR pathway (5-10 shRNAs/gene, median 5 shRNAs/gene), as well as 124 control shRNAs targeting LacZ, GFP, RFP, and Luciferase (1 shRNA/well; 1μg/mL polybrene used in transduction). The lentiviral library had been titered by The RNAi Consortium's RNAi Platform at the Broad Institute (Root et al., 2006) to achieve comparable viral titers for all constructs, and the transduction conditions were optimized for GBM TICs. Two replicates were selected with puromycin (0.5μg/mL) and the other four left unselected. Comparison of relative cell numbers in the two selected replicates and two of the unselected replicates after 5 days of selection using CellTiter-Glo (CTG) Luminescent Assay (according to manufacturer's instructions) showed approximately 85% or greater transduction efficiency across most wells. Empty wells were used to confirm puromycin selection. For each shRNA, an average CTG value was calculated from selected and unselected wells. After removing empty wells, the median and median absolute deviation (MAD) of the distribution of CTG scores was calculated and the robust Z-score was calculated for each shRNA. Robust Z-score, rather than Z-score, is the preferred method of normalization for RNAi screens, as it minimizes outlier effects (Birmingham et al., 2009). Only considering shRNAs with an infection efficiency greater than 75%, genes were considered hits if their second-best scoring (second most negative robust Z score) was less than -1.5. This requirement of at least two shRNAs yielding strong reductions in relative cell number was used in order to increase confidence that the decrease in relative cell number was due to on-target effects of suppression of the targeted gene.

The remaining two unselected replicates were stained with anti-Tubulin antibody and Hoechst 33342 and imaged on a Cellomics ArrayScan V (Thermo Scientific) automated microscope ("Acquire Only" application; 4 fields / well imaged in 2 channels at 10x magnification, with auto-focus before every field). To identify individual objects in each well, we created a CellProfiler (Carpenter et al., 2006) pipeline (available at http://cellprofiler.org/published_pipelines.shtml). For phenotype scoring, we used CellProfiler Analyst (Jones et al., 2009), a software package that uses machine-learning algorithms to distinguish between cellular phenotypes. We first manually binned a small number of randomly chosen

cellular objects from the screening set to one of three categories: Sphere, Flat and Elongated, or Everything Else. After we binned several dozen objects, the software used 147 measurements from CellProfiler to create a rule set (i.e., classifier) able to distinguish between the three phenotypic bins. We then requested random objects from each phenotype based on the tentative rules, corrected errors and asked the software to update the classifier based on the revised object samples. After several iterations of this procedure, the rules were optimized to correctly distinguish the phenotypes of interest; this optimized rule set was then used to classify all objects from the screen. To score individual wells (which contain numerous objects), the probability distribution of positive objects for a given phenotype was fit to a beta-binomial distribution across all wells, with parameters α and β. Enrichment in a given well for a given phenotype was calculated as the probability that a random draw from Beta($\alpha+n_{pos}$, $\beta+n_{neg}$) is larger than a draw from Beta(α, β), where n_{pos} and n_{neg} are the positive and negative counts, respectively. This approach provides a controlled estimate in a case when few cells are available for scoring.

To score genes, individual shRNAs were first ranked from strongest to weakest phenotype as denoted by their per-well enrichment scores. The rank of the second-strongest hairpin was used to rank-order the genes and served as the score for the gene. Use of the second-best rank prioritizes genes with at least two shRNAs showing strong phenotypic effects, increasing confidence that the phenotype is due to on-target effects of suppression of the targeted gene. Since the ranking of a gene is driven to some extent by the number of shRNAs targeting it, we normalized the gene score by dividing by the maximum possible score that all genes with a similar number of hairpins could have obtained by random chance. Using this second best ranking metric, we selected the top 116 genes for follow up. We also included genes in the follow up screen that did not meet this strict criterion: those that had a single shRNA with a differentiated phenotype enrichment score in the top 1.5% of shRNAs (6 genes) and those that scored in the top 1.5% by the alternatively used Kolmogorov–Smirnov-based method (23 additional genes) (Cheung et al., 2011; Luo et al., 2008). Together, we obtained 145 candidates whose suppression strongly alters TIC morphology.

The confirmation screen was performed in triplicate. Each replicate was composed of approximately 900 shRNAs targeting 145 of the top scoring genes from the primary screen (1 shRNA/well; median 5 shRNAs/gene), 32 unique control shRNAs distributed in 240 wells, and 36 wells treated with serum. Control wells were statistically treated like individual shRNAs, so a comparative "gene" score distribution was defined via permutation, to which individual genes could be compared. All analyses were performed using the GENE-E software package (Cheung et al., 2011) (http://www.broadinstitute.org/cancer/software/GENE-E/). The results of the confirmation screen were plotted as a relative frequency histogram of differentiation scores on a per-gene basis; the area under each of the curves is equal to 1.

GBM molecular subtype analysis on screen hits and TIC lines

To evaluate the subtype specific expression of screen hits, we examined the HT-HG-U133A expression data set from TCGA (TCGA, 2008) (http://cancergenome.nih.gov). Raw files of 543 TCGA glioblastoma samples were collected and preprocessed using the aroma package (Bengtsson et al., 2008). Of these 543 samples, 201 were included and classified by Verhaak *et al* (Verhaak et al., 2010). The other samples were classified using single sample Gene Set Enrichment Analysis (ssGSEA) (Barbie et al., 2009). In total, 141 samples were classified as Proneural, 96 as Neural, 146 as Classical and 160 as Mesenchymal. Significance Analysis of Microarrays (SAM) (Tusher et al., 2001) was used to test for differential gene expression in each subtype. Significant differentially expressed genes were identified by the cutoff of FDR ≤ 0.01 and fold change ≥ 1.5 for both up and down directions.

To assign GBM molecular subtype to each TIC line, transcriptional profiles from clinical GBM samples (Verhaak et al., 2010) and from 0308, BT112, BT145, BT147, and BT99 TIC lines [HG-U133A and HG-U133Plus2 platforms (Affymetrix), respectively] were normalized by robust multichip average (RMA) separately using the Affy package from Bioconductor. A custom probeset definition, such that there is one probeset per Entrez Gene ID, as defined by Dai *et al.* (Dai et al., 2005), was used to process the arrays, with 554 GBM subtype predictive genes (Verhaak et al., 2010) remaining post-processing. The data were then combined and quantile normalized for the predictive gene list (Verhaak et al., 2010). GBM subtype was assigned to each TIC line by hierarchical clustering using the Cluster

3.0 software. As an additional measure, ClaNC (classification to nearest centroids) software (Dabney, 2006) was used for verifying the assigned subtype.

Validation RNAi experiments

For validation and follow-up studies, GBM cell lines, TICs, or NPCs were seeded on uncoated, Poly-D-Lysine coated, or Matrigel-coated, respectively, 6-well plates (GBM cell lines and TICs: $1.5x10^5$ cells/well; NPCs: $2x10^6$ cells/well), transduced with individual control or gene-targeting lentiviral shRNAs (as above), selected with puromycin (GBM cell lines: 2 μg/mL; TICs: 0.5 μg/mL; NPCs: 1 μg/mL), and imaged in bright-field on a light microscope, collected using trypsin or accutase and counted for use in further *in vitro* or *in vivo* experiments, or lysed to extract protein or RNA (see below). For immunoblot-based assessment of marker expression in TICs (Figures 2B and 4D), the plates were additionally coated with Natural Mouse Laminin prior to cell seeding.

Immunofluorescent staining

TICs were plated on coverslips coated with Poly-D-Lysine and laminin in a 24-well plate ($8x10^4$ cells / well), transduced with *LacZ* or *ZFHX4* targeting shRNAs, and selected with puromycin (as above). Nine days after plating, cells were fixed in 4% paraformaldehyde and subjected to immunofluorescent staining as described (Sancak et al., 2010), using a solution of 0.3% Triton X-100 and 4% goat serum in PBS as the permeabilization/blocking agent and antibody diluent. Coverslips were mounted using Prolong Gold Antifade reagent and images were obtained using a Zeiss LSM 710 confocal microscope with ZEN software.

NPCs, after transduction and selection in 6-well plates, were transferred to Poly-D-Lysine and Natural Mouse Laminin coated 96-well plates (10^5 cells/well), fixed and subjected to immunofluorescent staining as above, and imaged on an Axiovert 200 epifluorescent microscope (Zeiss), with an Exfo Xcite120Q light source.

Immunoprecipitation, mass spectrometry, and immunoblotting

Immunoprecipitation (IP), mass spectrometry, and immunoblotting were performed as described (Sancak et al., 2010; Whyte et al., 2012). For FLAG tag IP, HEK 293T cells were transfected as described (Sancak et al., 2010) with pcZeo-FLAG-ZFHX4 or pRK5-FLAG-Raptor plasmid and lysed for whole cell extract (WCE) 72 hours post-transfection in lysis buffer (50mM HEPES [pH 7.4], 40 mM NaCl, 2 mM EDTA, 1.5 mM $NaVO_4$, 50 mM NaF, 10 mM pyrophosphate, 10 mM glycerophosphate, 1% Triton X-100, and one tablet of EDTA-free protease inhibitors (Roche) per 25 mL), and the WCE was used for direct immunoblotting or for FLAG IP followed by mass spectrometry or immunoblotting (Sancak et al., 2010). For endogenous IP from 0308 TICs, nuclear extracts were prepared and processed as described (Whyte et al., 2012). For CHD4 IP of ZFHX4-depleted cells, approximately 10^7 0308 TICs transduced with a *LacZ*-targeting shRNA or each of the two *ZFHX4*-targeting shRNAs were used.

Cell proliferation and viability assays

After shRNA transduction of 0308 TICs and puromycin selection for approximately 2.5 days, we re-suspended single cells using accutase and brief manual pipetting and plated them in 6-well Poly-D-Lysine coated plates (10^5 cells in 2mL NBE medium / well). We assessed cell number and viability using a Vi-Cell XR Cell Viability Analyzer (Beckman Coulter), which uses trypan blue exclusion. We replaced media every 2-3 days. To ensure we accounted for possible differences in plating efficiency, all cell numbers were normalized to day 0.

Cell cycle analysis

DNA content measurement for cell cycle analysis was performed using propidium iodide (Sigma 81845) as described (Darzynkiewicz and Juan, 2001), and analyzed using Flowjo software (Tree Star, Inc.).

Clonogenic neurosphere formation assay

After transduction, selection and resuspension of 0308 or BT112 TICs as above, we used fluorescence-activated high speed cell sorting (MoFlo XDP, Beckman Coulter) gating on 4',6-diamidino-

2-phenylindole (DAPI) to plate one viable cell per well in 384-well plates (each well pre-plated with 50 μl NBE, 50% of which was conditioned media supplemented with 50 ng/mL EGF and 50 ng/mL bFGF). The next day, we calculated plating efficiency by staining plates with Hoechst 33342 and visually inspecting each well at 40x magnification for the presence of a cell. Duplicate plates were similarly assessed two weeks later for the presence of spheres. The percentage of wells with spheres was calculated and normalized to plating efficiency and to sphere formation by control *LacZ*-targeting shRNA treated cells.

Xenograft tumorigenesis studies

BT112 TICs were transduced with a *LacZ*-targeting shRNA or each of the two *ZFHX4*-targeting shRNAs and selected with puromycin, then put into single-cell suspension using trypsin, resuspended in growth medium (see above) at a concentration of $5x10^4$ cells/μL, and intracranially injected into 6-8 week-old NOD.CB17-*Prkdc*^{scid}/NcrCrl female mice (Charles River Laboratories). Mice were anesthetized using a ketamine-xylazine mixture and injected with 2μL of the cell suspension (10^5 cells) in the striatum (2mm lateral, 2.5mm deep relative to bregma), then monitored every day and sacrificed upon exhibiting neurological symptoms or other evidence of morbidity. Animals that died due to non-cancer causes were censored at the time of death. Brains were excised from sacrificed animals, fixed in 4% paraformaldehyde, and processed for histology and IHC.

Xenograft studies were approved by the Dana-Farber Animal Care and Use Committee and conducted in compliance with the Animal Welfare Act Regulations and other Federal statutes relating to animals and experiments involving animals (institutional animal welfare assurance no. A-3023-01).

Histology and IHC

Paraffin embedding of murine brain tissue and generation of unstained and hematoxylin and eosin (H&E) stained slides were performed at the Harvard Medical School Rodent Histopathology Core. For validation of the ZFHX4 antibody used for IHC, approximately $9x10^6$ 0308 TICs transduced with a *LacZ*-targeting or *ZFHX4*-targeting shRNA were collected, resuspended and solidified in HistoGel (according to manufacturer's instructions), fixed in 10% formalin, and processed the same way as the murine brain tissue.

For Nestin, Ki-67, Sox2, and Olig2 IHC, antigen retrieval was done for 20 minutes in a steamer using Sodium Citrate. Antibodies were diluted using PBS and incubated at 4°C overnight. Dako Envision+ Kit (K4007 Mouse secondary, K4011 Rabbit secondary) was used for secondary and 3,3'-diaminobenzidine (DAB) staining.

For ZHFX4 IHC, antigen retrieval was performed with Borg Decloaker RTU solution in a pressurized Decloaking Chamber for 3 minutes. Blocking and antibody dilutions were carried out in 4% horse serum and 0.1% Tween in PBS. The Vectastain Elite ABC immunoperoxidase detection kit (Vector Labs) with DAB+ Substrate (Dako) was used for chromogenic visualization. All samples were processed identically with a DAB development step of 5.5 minutes.

IHC-based scoring of ZFHX4 expression

Two researchers scored, in a blinded manner, the murine brain sections subjected to ZHFX4 IHC. Scoring was based on a scale of 0-3, in which 0 indicated no visible chromogenic signal, 3 indicated saturation of signal in the nuclei, and 1 and 2 indicated faint and strong signal, respectively. For brains harboring tumors, an average score of multiple fields across the entire tumor area was obtained as a representative score for the tumor. For brain sections with no tumor detected, a representative score for the entire section was obtained.

Transcriptional profiling studies and analysis

We extracted total RNA at 3 days and 5 days after transduction of 0308 TICs with control, *CHD4*-, or *ZFHX4*-targeting shRNAs, from 3 (for CHD4) or 5 (ZFHX4) independent experiments. RNA was harvested using Trizol RNA isolation reagent (Life Technologies), DNase I digestion, and RNeasy MinElute Cleanup Kit (Qiagen) following manufacturer's protocols. We performed synthesis of complementary RNA from total RNA and hybridization/scanning of U133 plus 2.0 microarrays using GeneChip products (Affymetrix) as described in the GeneChip manual.

To identify the transcriptome alterations induced by ZFHX4 and CHD4 suppression, the TCGA CDF file was used to preprocess and normalize the raw expression data from the control and CHD4- or ZFHX4-depleted TICs, using the aroma package (Bengtsson et al., 2008). The ZFHX4 depletion signature was derived by comparing the gene expression profiles of ZFHX4-suppressed cells to those of control cells. Genes expressed at significantly higher or lower levels in ZFHX4-depleted cells were identified by the criterion of FDR ≤ 0.05. The same procedure was applied to the profiles of CHD4-suppressed cells to derive the CHD4 depletion signature. A chi-square test was used to test the significance of overlap of genes differentially expressed in the two gene sets. ZFHX4 and CHD4 expression correlation analysis in TCGA GBM samples (TCGA, 2008) was performed using a Pearson correlation test as described (Winslow et al., 2011).

Chromatin immunoprecipitation followed by DNA sequencing (ChIP-Seq)

20-100 million target cells were cross-linked at room temperature for 10 min with fresh 11% formaldehyde, quenched with 2.5M glycine, and processed through the DNA elution steps following the Agilent Technologies Mammalian ChIP-on-chip protocol (version 10.2, March 2011, http://www.chem.agilent.com/Library/usermanuals/Public/G4481-90010_MammalianProtocol_10.2.pdf), with the following modifications: sonication buffer composition 20 mM Tris-HCl , 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100; sonication in 3 mL was performed on a Bioruptor Standard (Diagenode, Denville, NJ) high setting at 4°C for 30 min total (cycles: 30 seconds on/30 seconds off); after sonication and centrifugation at 1400g, supernatant underwent immunoprecipitation with 10 ug antibody [anti-ZFXH4 (Abnova, H00079776-M11), mouse IgG2A (R&D Systems, AB003), anti-CHD4 (Abcam, ab72418), or rabbit polyclonal (Abcam, ab27478) antibodies] for four days for ZFHX4/mouse isotype control IP or overnight for CHD4/rabbit polyclonal control IP; protein G Dynabeads (Life Technologies, 1007D) were incubated with the IP lysates for 45 min and washed with sonication buffer x 2, buffer 1 (20 mM Tris HCl, 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100) x 2, buffer 2 (10 mM Tris HCl, 250 mM LiCl, 2 mM EDTA, 1% Igepal) x 2, and buffer 3 (10 mM Tris HCl, 1 mM EDTA, 50 mM NaCl) x 2. Immunoprecipitated DNA was quantified by picogreen and size was evaluated on a HighSense BioAnalyzer chip. Fragments between 100 and 600 bp were collected using an automated system (Pippin Prep, Sage Science) then end repaired, ligated and amplified for 15 cycles using reagents included in the Truseq DNA Sample Preparation kit from Illumina. Experimental conditions followed strictly the instructions of the manufacturer, with the exception of the adaptors being diluted 1/10 for the input DNA and 1/50 for all other samples. 7 barcoded libraries were run on 2 lanes of a Hiseq 2000 in a 50bp/50bp paired end run, using the TruSeq SBS Kit v3 (Illumina). The average number of read pairs per sample was 48 million.

ChIP-Seq data analysis

ChIP-Seq reads were aligned using the software Bowtie (Langmead et al., 2009) to NCBI build 36 (hg18) of the human genome with default settings. Sequences uniquely mapping to the genome with zero or one mismatch were used in further analysis. When multiple reads mapped to the same genomic position, a maximum of two reads mapping to the same position were used.

Analysis methods were derived from previously published methods (Whyte et al., 2012). Sequence reads from multiple flow cells for each IP target and/or biological replicates were combined. For all datasets, each read was extended 600bp towards the interior and 400bp towards the exterior of the sequenced fragment, based on the strand of the alignment. Across the genome, in 25 bp bins, the number of extended ChIP-Seq reads was tabulated. The 25bp genomic bins that contained statistically significant ChIP-Seq enrichment were identified by comparison to a Poissonian background model. Assuming background reads are spread randomly throughout the genome, the probability of observing a given number of reads in a genomic bin can be modeled as a Poissonian process in which the expectation can be estimated as the number of mapped reads multiplied by the number of bins (40 for all sequences datasets used here) into which each read maps, divided by the total number of bins available (we estimated 70% of the genome). Enriched bins within 200bp of one another were combined into regions.

The Poissonian background model assumes a random distribution of background reads, but we have observed significant deviations from this expectation. Some of these non-random events can be detected as sites of apparent enrichment in negative control DNA samples and can create many false

positives in ChIP-Seq experiments. To remove these regions, we compared genomic bins and regions that meet the statistical threshold for enrichment to a set of reads obtained from Illumina sequencing of DNA from whole cell extract (WCE) in matched cell samples. We required that enriched bins and enriched regions have five-fold greater ChIP-Seq density in the specific IP sample, compared with the control sample, normalized to the total number of reads in each dataset. This served to filter out genomic regions that are biased to having a greater than expected background density of ChIP-Seq reads.

The whole-genome plot of ZFHX4 and CHD4 ChIP-Seq mapped reads was made using Circos software (http://circos.ca). Chromosome Y had extremely low sequence coverage and was not shown in the plot. ZFHX4/CHD4 co-bound regions were defined as having an overlap of at least 1bp between a ZFHX4-occupied region and a CHD4-occupied region.

Assessing significance of co-binding

The total number of genomic regions available for occupancy by ZFHX4 or CHD4 was estimated by dividing the total Illumina sequence coverage of DNA from WCE (2,043,862,849bp) by the average sequence length of all ZFHX4-occupied and CHD4-occupied regions (622bp; Table S2 Sheets 4 and 5), yielding 3,284,154 regions. Based on the findings of 19,261 ZFHX4-occupied regions, 32,078 CHD4-occupied regions, and 5594 ZFHX4-occupied regions co-bound by CHD4 (Table S2 Sheets 4-6), enrichment of ZFHX4-bound regions for co-occupancy by CHD4 was calculated by dividing the fraction of ZFHX4-bound regions co-occupied by CHD4 (5594/19,261=0.29) by the fraction of all available regions bound by CHD4 (32,078/3,284,154=0.0098), and p-value was calculated using a two-tailed Fisher's exact test.

Assigning ChIP-Seq enriched regions to genes

The complete set of RefSeq genes was downloaded from the UCSC table browser (http://genome.ucsc.edu/cgi-bin/hgTables) on June 1, 2010. For all datasets, genes with enriched regions within the gene body or within 10kb of the start or stop of the gene were called bound. Co-bound genes were defined as containing, within the gene body or within 10kb of the start or stop, a ZFHX4-bound region co-bound by CHD4.

Identification of candidate transcriptional targets of ZFHX4 using ARACNe

ARACNe (Margolin et al., 2006) is an information-theoretic algorithm for the dissection of transcriptional interactions. It uses a large compendium of gene expression profiles to identify statistically significant gene-gene coregulation by mutual information. It then eliminates indirect relationships, in which two genes are coregulated through one or more intermediate genes, using well-established data transmission theory called data processing inequality (DPI). To generate a GBM-specific transcriptional network, we used 273 gene expression profiles (HG-U133A GeneChip arrays, Affymetrix) from 262 primary patient glioblastoma samples, 10 control samples, and 1 cell line control sample obtained from The Cancer Genome Atlas (TCGA (TCGA, 2008)), resulting in approximately 270,000 interactions. This network was generated using the bootstrap version of the ARACNe (Margolin et al., 2006) algorithm, which helps to reduce false negative connections in the network. The ARACNe algorithm was run using the following parameters: p-value=10^{-7}, dpi=0, and 100 bootstraps. The resulting interaction network for 1173 total transcription factors was parsed to identify 119 interactions for ZFHX4, 113 of which were represented in the CHD4 suppression transcriptional profiling arrays and were therefore testable for enrichment.

Gene set enrichment analysis of ZFHX4 targets in *CHD4* suppression signature

Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) uses a Kolmogorov-Smirnov (KS)-like test to determine whether genes in a given gene set (in this case, 113 candidate transcriptional targets of ZFHX4, identified using ARACNe) are over-represented among the genes in a reference signature (in this case, the ranked list of genes, from most to least differentially expressed, in *CHD4*-suppressed 0308 TICs [both day 3 and day 5 post-transduction, both shRNAs] compared to control TICs expressing an anti-*LacZ* shRNA). Genes to the left of the highest point of the enrichment score curve are called the *leading edge.* The significance of enrichment is estimated using a non-parametric test with sample shuffling, as follows. First, from the total set of samples (n_1 *CHD4*-

suppressed samples and n_2 control samples, for a total population of n_1+n_2), n_1 samples are randomly selected and assigned to the *CHD4*-suppressed group, and the rest to control. Second, differential expression is computed for these randomly assigned *CHD4*-suppressed and control samples, and the reference signature is generated as above. Third, enrichment of the 113 candidate ZFHX4 targets in this random reference signature is calculated using GSEA with a KS test as above, and compared to enrichment in the actual reference signature. The use of 1,000 such sample shuffles allows the p-value to be estimated with an accuracy of up to $1x10^{-3}$.

Color editing of bright-field images

Bright-field images of TICs (Figures 1F, 2A, and 4D) were edited in Adobe Photoshop CS5 as follows. First, color mode was set to grayscale using the command “Image – Mode – Grayscale.” Second, contrast was adjusted using the command “Image – Auto Contrast.” All images were processed identically.

Supplemental References

Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C.*, et al.* (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108-112.

Bengtsson, H., Simpson, K., Bullard, J., and Hansen, K. (2008). aroma.affymetrix: A generic framework in R for analyzing small to very large Affymetrix data sets in bounded memory. Tech Report, Department of Statistics, University of California, Berkeley *745*.

Birmingham, A., Selfors, L.M., Forster, T., Wrobel, D., Kennedy, C.J., Shanks, E., Santoyo-Lopez, J., Dunican, D.J., Long, A., Kelleher, D.*, et al.* (2009). Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods *6*, 569-575.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J.*, et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol *7*, R100.

Cheung, H.W., Cowley, G.S., Weir, B.A., Boehm, J.S., Rusin, S., Scott, J.A., East, A., Ali, L.D., Lizotte, P.H., Wong, T.C.*, et al.* (2011). Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. Proc Natl Acad Sci U S A *108*, 12372-12377.

Dabney, A.R. (2006). ClaNC: point-and-click software for classifying microarrays to nearest centroids. Bioinformatics *22*, 122-123.

Dai, M., Wang, P., Boyd, A.D., Kostov, G., Athey, B., Jones, E.G., Bunney, W.E., Myers, R.M., Speed, T.P., Akil, H.*, et al.* (2005). Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res *33*, e175.

Darzynkiewicz, Z., and Juan, G. (2001). DNA Content Measurement for DNA Ploidy and Cell Cycle Analysis. In Current Protocols in Cytometry, J.P. Robinson, ed. (Somerset, NJ, John Wiley & Sons, Inc.), pp. 7.5.1–7.5.24.

Hemmi, K., Ma, D., Miura, Y., Kawaguchi, M., Sasahara, M., Hashimoto-Tamaoki, T., Tamaoki, T., Sakata, N., and Tsuchiya, K. (2006). A homeodomain-zinc finger protein, ZFHX4, is expressed in neuronal differentiation manner and suppressed in muscle differentiation manner. Biol Pharm Bull *29*, 1830-1835.

Jones, T.R., Carpenter, A.E., Lamprecht, M.R., Moffat, J., Silver, S.J., Grenier, J.K., Castoreno, A.B., Eggert, U.S., Root, D.E., Golland, P.*, et al.* (2009). Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. Proc Natl Acad Sci U S A *106*, 1826-1831.

Lai, A.Y., and Wade, P.A. (2011). Cancer biology and NuRD: a multifaceted chromatin remodelling complex. Nat Rev Cancer *11*, 588-596.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lee, J., Kotliarova, S., Kotliarov, Y., Li, A., Su, Q., Donin, N.M., Pastorino, S., Purow, B.W., Christopher, N., Zhang, W.*, et al.* (2006). Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. Cancer Cell *9*, 391-403.

Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Muffat, J.*, et al.* (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell *141*, 872-883.

Luo, B., Cheung, H.W., Subramanian, A., Sharifnia, T., Okamoto, M., Yang, X., Hinkle, G., Boehm, J.S., Beroukhim, R., Weir, B.A.*, et al.* (2008). Highly parallel identification of essential genes in cancer cells. Proc Natl Acad Sci U S A *105*, 20380-20385.

Margolin, A.A., Wang, K., Lim, W.K., Kustagi, M., Nemenman, I., and Califano, A. (2006). Reverse engineering cellular networks. Nat Protoc *1*, 662-671.

Mehta, S., Huillard, E., Kesari, S., Maire, C.L., Golebiowski, D., Harrington, E.P., Alberta, J.A., Kane, M.F., Theisen, M., Ligon, K.L.*, et al.* (2011). The central nervous system-restricted transcription factor Olig2 opposes p53 responses to genotoxic damage in neural progenitors and malignant glioma. Cancer Cell *19*, 359-371.

Moffat, J., Grueneberg, D.A., Yang, X., Kim, S.Y., Kloepfer, A.M., Hinkle, G., Piqani, B., Eisenhaure, T.M., Luo, B., Grenier, J.K.*, et al.* (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell *124*, 1283-1298.

Ramirez, J., and Hagman, J. (2009). The Mi-2/NuRD complex: a critical epigenetic regulator of hematopoietic development, differentiation and cancer. Epigenetics *4*, 532-536.

Root, D.E., Hacohen, N., Hahn, W.C., Lander, E.S., and Sabatini, D.M. (2006). Genome-scale loss-of-function screening with a lentiviral RNAi library. Nat Methods *3*, 715-719.

Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A.L., Nada, S., and Sabatini, D.M. (2010). Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. Cell *141*, 290-303.

Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, I.S.Y., Hahn, W.C., Sharp, P.A.*, et al.* (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. Rna *9*, 493-501.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

TCGA (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature *455*, 1061-1068.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.

Verhaak, R.G., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P.*, et al.* (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98-110.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221-225.

Winslow, M.M., Dayton, T.L., Verhaak, R.G., Kim-Kiselak, C., Snyder, E.L., Feldser, D.M., Hubbard, D.D., DuPage, M.J., Whittaker, C.A., Hoersch, S.*, et al.* (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. Nature *473*, 101-104.