Transcription elongation factors represent in vivo cancer dependencies in glioblastoma

Tyler E. Miller1,2,3,4, Brian B. Liu5,6,7, Lisa C. Wallace1, Andrew R. Morton1,4, Qi Xie1, Deobrat Dixit1, Daniel C. Factor1, Leo J. Y. Kim1,2,3, James J. Morrow1,4, Qiulian Wu1, Stephen C. Mack1,2, Christopher G. Hubert1,2, Shawn M. Gillespie5,6,7, William A. Flavahan1,2, Thomas Hoffmann8, Rohit Thummalapalli5,6,7, Michael T. Hemann9, Patrick J. Paddison10, Craig M. Horbinski11,12, Johannes Zuber8, Peter C. Scacheri4, Bradley E. Bernstein5,6,7, Paul J. Tesar4 & Jeremy N. Rich1,2

Glioblastoma is a universally lethal cancer with a median survival time of approximately 15 months1. Despite substantial efforts to define druggable targets, there are no therapeutic options that notably extend the lifespan of patients with glioblastoma. While previous work has largely focused on in vitro cellular models, here we demonstrate a more physiologically relevant approach to target discovery in glioblastoma. We adapted pooled RNA interference (RNAi) screening technology2–4 for use in orthotopic patient-derived xenograft models, creating a high-throughput negative-selection screening platform in a functional in vivo tumour microenvironment. Using this approach, we performed parallel in vivo and in vitro screens and discovered that the chromatin and transcriptional regulators needed for cell survival in vivo are non-overlapping with those required in vitro. We identified transcription pause–release and elongation factors as one set of in vivo-specific cancer dependencies, and determined that these factors are necessary for enhancer-mediated transcriptional adaptations that enable cells to survive the tumour microenvironment. Our lead hit, JMJD6, mediates the upregulation of in vivo stress and stimulus response pathways through enhancer-mediated transcriptional pause–release, promoting cell survival specifically in vivo. Targeting JMJD6 or other identified elongation factors extends survival in orthotopic xenograft mouse models, suggesting that targeting transcription elongation machinery may be an effective therapeutic strategy for glioblastoma. More broadly, this study demonstrates the power of in vivo phenotypic screening to identify new classes of ‘cancer dependencies’ not identified by previous in vitro approaches, and could supply new opportunities for therapeutic intervention.

Chromatin regulators have emerged as a promising class of druggable targets for cancer therapy5,6,8–11. Chromatin regulation is often context-specific2–4,7, suggesting that the microenvironment mediates cancer cell response to inhibition of specific chromatin regulators. Therefore, we developed an in vivo RNAi screening strategy to enable the identification of chromatin regulators that are crucial for the survival of glioblastoma cells within a functional tumour microenvironment (Fig. 1a). Using an advanced short hairpin RNA (shRNA) delivery vector2–4 (Extended Data Fig. 1a–c), glioblastoma patient-derived xenograft (PDX) cells (Supplementary Table 1) were transduced with a pooled library containing 1,586 inducible shRNAs targeting 406 known chromatin and transcriptional regulators (2–4 shRNAs per gene) and controls, at efficiencies to achieve a single retroviral integration per cell. Transduced cells were selected by a constitutive green fluorescent reporter using fluorescence activated cell sorting (FACS) and used in concurrent in vivo and in vitro screens. In each screen, transduced cells were split into an induced arm and an uninduced control arm. Cells in the induced arm were treated with doxycycline, which induced shRNA expression and a second fluorescent reporter, dsRED. After 3 weeks, induced cells (dsRED+) or uninduced control cells were sequenced and shRNA representation was quantified. For the in vivo screen, 61 mice were implanted with cells and randomly assigned to the control or induced arm. Multiple mice were grouped as single biological replicates, providing the fold coverage necessary to obtain reproducible results (Extended Data Fig. 2a), and enabling successful in vivo negative-selection screening in a solid tumour model (see Methods).

Genes crucial for glioblastoma cell survival were prioritized by calculating depletion scores for each shRNA, based on the decrease in shRNA frequency in the induced arm compared to the uninduced control arm. Positive hits were defined as expressed genes that were targets of at least two non-overlapping shRNAs that effectively mediated cellular depletion (Fig. 1b and Supplementary Table 2). In vivo hits outnumbered the in vitro hits and, surprisingly, there was almost no overlap between hits that caused cell depletion in vitro versus in vivo (Fig. 1c). Genes that caused cell depletion in both screens were restricted to the positive-control gene RPA3 (Extended Data Fig. 2b) and two genes essential for transcription and maintenance of DNA methylation, POLR2B and DNMT1. Differences in molecular dependencies were not explained by expression of the hits, as there were no significant differences in the in vivo expression of the hits compared to their in vitro expression (Extended Data Fig. 3a, b). Collectively, these primary screen results reveal unique molecular dependencies for glioblastoma cells in vivo.

Gene Ontology analysis of the hits revealed significantly enriched molecular classes that were microenvironment-specific. In vitro-specific hits were enriched for genes that promote cellular metabolism and macromolecule biogenesis, whereas in vivo-specific hits were enriched for genes controlling transcriptional elongation (Fig. 1d and Extended Data Fig. 3c, d). Nearly all mediators of transcriptional pause–release and elongation included in the screen scored as in vivo-specific hits, including recently annotated regulators, junomji C-domain–containing protein 6 (JMJD6)12, the DOT1 like histone lysine methyltransferase (DOT1L) complex13, and ring finger protein 20 (RNF20)14 (Fig. 1e). Several primary hits were validated in secondary in vivo survival assays (Extended Data Fig. 3e–h).

1Department of Stem Cell Biology and Regenerative Medicine, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA. 2Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland, Ohio 44195, USA. 3Department of Pathology, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA. 4Department of Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA. 5Department of Neurological Surgery, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60615, USA. 6Department of Pathology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60614, USA. 7Epigenomics Program, Broad Institute, Cambridge, Massachusetts 02142, USA. 8Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. 9Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC), 1030 Vienna, Austria. 10The Koch Institute for Integrative Cancer Research at MIT, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. 11Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. 12Department of Neurological Surgery, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611, USA. 13Department of Pathology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60615, USA. 14Present address: Division of Regenerative Medicine, Department of Medicine, University of California at San Diego, San Diego, California 92093, USA.
To investigate the mechanisms underlying the selective dependency of glioblastoma cells on transcriptional pause–release and elongation as an in vivo-specific target, a schematic diagram depicting screen. Dox, doxycycline; GSC, glioma stem cell; NGS, next-generation sequencing. b, Plot of score of second-best shRNA targeting each gene in each screen as calculated by RIGER. The score increases depicting screen. Dox, doxycycline; GSC, glioma stem cell; NGS, next-generation sequencing. c, Schematic of transcription elongation machinery, highlighting in vivo-specific hits. TFs, transcription factors; TSS, transcription start site.

Figure 2 | Transcription of pause-controlled programs is upregulated in the in vivo tumour microenvironment. a, Workflow for global analysis of glioblastoma cells. b, Cellular programs enriched by GSEA in cells grown in each condition represented using Enrichment Map. Node size represents the number of genes overlapped. c, Representative GSEA enrichment plots. FDR calculated by GSEA software. d, Principle component analysis of matched glioblastoma (GBM) cells in primary tumours, intracranial (IC) tumours and cell culture. NSC, neural stem cells. e, Fold change of the H3K27ac signal at enhancers of genes with a greater than 2.5-fold change in mRNA expression between conditions, or a 0.9–1.1-fold change (stable genes). P-values were determined by two-sided Mann–Whitney U-test. RPKM, reads per kilobase per million.

an intracranial xenograft model as compared to cell culture conditions (Extended Data Fig. 4a, b and Supplementary Table 3).

Gene set enrichment analysis (GSEA) coupled to Enrichment Map visualization was used to annotate differentially enriched biological pathways (Fig. 2b, c, Extended Data Fig. 4 and Supplementary Table 4a–f). Cancer cells cultured in serum-free conditions, in which nutrients and space are in abundant supply, were enriched for transcriptional programs of proliferation. By contrast, transcriptional programs associated with stress response, signalling response and other stimulus response pathways were enriched in intracranial tumours, in which nutrients and space are less abundant. The stimulus response pathways upregulated in the intracranial tumour environment include
pause-controlled pathways consisting of genes with a strong reliance on transcription pause–release and elongation for their expression. 13,18–22 Of the 55 genes that were upregulated more than 2.5-fold in tumour cells grown intracranially in both the proneural GBM528 and mesenchymal GBM3565 models, many were transcription factors and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled genes, such as EGR1 and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled pathways consisting of genes with a strong reliance on transcription pause–release and elongation for their expression. 13,18–22 Of the 55 genes that were upregulated more than 2.5-fold in tumour cells grown intracranially in both the proneural GBM528 and mesenchymal GBM3565 models, many were transcription factors and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled genes, such as EGR1 and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled pathways consisting of genes with a strong reliance on transcription pause–release and elongation for their expression. 13,18–22 Of the 55 genes that were upregulated more than 2.5-fold in tumour cells grown intracranially in both the proneural GBM528 and mesenchymal GBM3565 models, many were transcription factors and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled genes, such as EGR1 and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled pathways consisting of genes with a strong reliance on transcription pause–release and elongation for their expression. 13,18–22 Of the 55 genes that were upregulated more than 2.5-fold in tumour cells grown intracranially in both the proneural GBM528 and mesenchymal GBM3565 models, many were transcription factors and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled genes, such as EGR1 and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled pathways consisting of genes with a strong reliance on transcription pause–release and elongation for their expression. 13,18–22 Of the 55 genes that were upregulated more than 2.5-fold in tumour cells grown intracranially in both the proneural GBM528 and mesenchymal GBM3565 models, many were transcription factors and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled genes, such as EGR1 and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled pathways consisting of genes with a strong reliance on transcription pause–release and elongation for their expression. 13,18–22 Of the 55 genes that were upregulated more than 2.5-fold in tumour cells grown intracranially in both the proneural GBM528 and mesenchymal GBM3565 models, many were transcription factors and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled genes, such as EGR1 and signalling molecules regulated by polymerase II (Pol II) pausing, including pause-controlled pathways consisting of genes with a strong reliance on transcription pause–release and elongation for their expression.

Enhancers act together with transcription factors to drive transcriptional changes through transcriptional pause–release and elongation. Leveraging chromatin immunoprecipitation followed by high-throughput sequencing (ChIP–seq) of the dynamic enhancer mark Lys27 acetylation of histone 3 (H3K27ac), transcriptional changes in tumour cells grown intracranially were reflective of genome-wide alterations in enhancers in two glioblastoma models (Fig. 2e). Nearly 20% of all enhancer elements, including ‘super-enhancer’ loci, were condition-specific, and expression of their putative target genes, as determined by the nearest expressed gene, corresponded to the condition-dependent changes in the enhancer signal (Extended Data Fig. 6 and Supplementary Table 5). Together, these data show that the microenvironment regulates the epigenome to transform the glioblastoma cell state by differentially activating enhancers and their target genes.

Of the in vivo–specific hits, we prioritized the 12 genes encoding transcription elongation factors for further consideration as therapeutic targets. Using large, independent datasets of primary patient tumours, we correlated expression of each of the 12 transcription elongation factors with the expression of the 55 genes that were consistently upregulated in vivo across both PDX tumour models, which we hypothesized are main drivers of the transcriptional programs needed by tumours to adapt and survive in the stressful and dynamic in vivo microenvironment. JMJD6 was the most positively correlated...
hit across all datasets, including the Ivy Glioblastoma Atlas Project (Ivy GAP) dataset, which provides intratumour microenvironment-specific expression24 (Fig. 3a and Extended Data Fig. 7a–d), suggesting that JMJD6 regulates many of the genes that are important for the survival of glioblastoma cells in vivo.

To assess further the regulatory role of JMJD6 in patient tumours, we correlated the expression of all genes in The Cancer Genome Atlas (TCGA) glioblastoma tumours individually with the expression of JMJD6 (Extended Data Fig. 7e and Supplementary Table 6). Genes positively correlated with JMJD6 were enriched in pause-controlled programs, similar to those upregulated in intracranial tumours, whereas genes negatively correlated with JMJD6 were enriched in metabolic programs, similar to those upregulated in cell culture (Extended Data Fig. 7f and Supplementary Table 4g). These findings indicate a potential mechanism for the in vivo specificity of JMJD6 in the screen, and provide evidence that JMJD6 may control transcriptional pause–release in primary glioblastoma tumours.

To explore the clinical significance of JMJD6, we analysed gene expression in primary tumours and found that both JMJD6 mRNA (Fig. 3b) and protein (Fig. 3c and Extended Data Fig. 7g) were highly expressed in gliomas, and increased with tumour grade. These data, along with the robust in vivo depletion of tumour cells harbouring JMJD6 shRNA in the primary screen (Extended Data Fig. 7h), provide further evidence that JMJD6 constitutes a strong lead target for further evaluation.

In HEK293T and HeLa cells, JMJD6 acts with bromodomain containing 4 (BRD4) as a key activator of enhancer-mediated pause–release at genes controlled by Pol II pausing25 (Extended Data Fig. 8a). To determine whether JMJD6 localizes to enhancers in glioblastoma in vivo, we performed JMJD6 ChIP–seq of tumour cells in intracranial tumours. Globally, JMJD6 was distributed throughout the genome, but was strongly enriched at enhancers and promoters (Extended Data Fig. 8b, c and Fig. 3d). Furthermore, in the PDX models, target genes of JMJD6-bound enhancers were enriched for genes upregulated in vivo (Extended Data Fig. 8d, e) and for genes that positively correlated with JMJD6 in primary patient tumours (Extended Data Fig. 8f, g). These results indicate that JMJD6 may regulate the expression of genes targeted by JMJD6-bound enhancers through enhancer-mediated pause–release, both in the intracranial tumour environment in our PDX models, as well as in patient tumours.

To investigate, we conducted ChIP–seq of RNA Pol II in GBM528 and GBM3565 cells both in vivo and in vitro. Transcription pause–release and elongation was measured by the pausing index, which is the ratio of the Pol II density surrounding the transcriptional start sites to the density of Pol II over the gene body22. The higher the pausing index, the more paused the gene transcript. Genes that were consistently upregulated in vivo in both PDX tumours models, such as EGR1, were transcriptionally paused in vitro and released in vivo (Fig. 3e and Extended Data Fig. 8h, i). To interrogate globally whether JMJD6 activity at enhancers promotes pause–release in vivo, we calculated the in vivo pausing index for all expressed genes with or without enhancers. Genes regulated by JMJD6-bound enhancers had lower pausing indexes, or increased levels of pause–release activity, compared to genes regulated by enhancers not bound by JMJD6 (Fig. 3f). This effect was specific to JMJD6 binding at enhancers, as genes bound by JMJD6 at locations other than an enhancer did not demonstrate increased pause–release.

We then evaluated the activity of JMJD6-bound enhancers relative to enhancers not bound by JMJD6 in vivo. Building upon recent evidence that active enhancers are transcribed and that activity can be estimated by RNA Pol II binding25, we measured relative RNA Pol II binding at JMJD6-bound enhancers versus enhancers not bound by JMJD6, and found that JMJD6-bound enhancers had significantly higher levels of Pol II binding (Fig. 3g). Measurements of Pol II binding at JMJD6 sites, including outside of enhancers, revealed that Pol II binding was highest at JMJD6-bound enhancers. Collectively, these data provide evidence that JMJD6 binding is associated with increased enhancer activity and promotes pause–release in human glioblastoma cells within the tumour microenvironment.

We next sought to determine the potential preclinical value of JMJD6 as a therapeutic target. Targeting JMJD6 with an inducible shRNA targeting the Renilla protein (not expressed in human cells). e, f, Parallel in vitro proliferation assay (e) and in vivo survival assay (f) of cells from d. P values (c, f) determined by Mantel–Cox log-rank test. Error bars denote ± s.d. of at least triplicates.
JMD6-deficient cells showed a notable survival advantage compared to mice with control cells implanted, with over 25% of mice in the JMD6-deficient group being tumour-free after 100 days (Fig. 4c). Similar results were found using an independent PDX model (Extended Data Fig. 9e–g).

To validate other screen results using in vivo survival studies, we performed similar experiments targeting DOT1L, a recently discovered mediator of transcription elongation and top scoring hit from the primary screen, and DPY30, another top scoring hit from the primary screen that is not known to be involved in transcription pause–release or elongation. Constitutive shRNA-mediated knockdown of these genes in glioblastoma cells from three independent PDX models caused no proliferation defect in vitro (Fig. 4d, e and Extended Data Fig. 9h–q). However, mice bearing cells with knockdown of DOT1L or DPY30 survived significantly longer than mice implanted with cells expressing non-targeting shRNAs (Fig. 4f and Extended Data Fig. 9r–v). These results further demonstrate the power of in vivo screening to identify in vivo-specific cancer dependencies, and indicate that many of the identified hits from the primary screen are strong candidates for future investigation as potential therapeutic targets in glioblastoma.

Pharmaceutical approaches to cancer drug discovery typically involve high-throughput screening of established cell lines cultured in vitro to reveal individual targetable oncogenes that predominantly regulate cancer cell proliferation. However, drugs developed to modulate these targets have thus far achieved limited success in patients, especially for glioblastoma. This therapeutic roadblock prompted us to develop and validate a new in vivo functional screening strategy for glioblastoma that recapitulates most stressors and stimuli of the tumour microenvironment. Our approach revealed several in vivo-specific biological targets for glioblastoma, including JMD6, indicating that cancer cells are dependent on unique molecular effectors for growth and survival depending on the extrinsic factors in their microenvironment. Specifically, we found that glioblastoma cells in vivo, but not in vitro, were dependent on Pol II pause–release and transcription elongation machinery for survival (Extended Data Fig. 10). This machinery is necessary to upregulate pause-controlled stress and signalling response pathways that promote cell survival.13,18,20,21 This finding unexpectedly revealed that in the primary tumour, targeting the microenvironment-induced stress response mechanisms of the cancer cell may be a more effective therapeutic strategy than targeting cell growth, the main target of traditional chemotherapies. Overall, our results demonstrate feasibility for direct target identification in vivo in solid tumours and suggest new avenues for therapeutic development.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 November 2015; accepted 5 June 2017.
Published online 5 July 2017.

Stupp, R. et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol. 10, 459–466 (2009).

Zuber, J. et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 476, 524–528 (2011).

Feilmann, C. et al. An optimized microRNA backbone for effective single-copy RNAi. Cell Reports 5, 1704–1713 (2013).

Dawson, M. A. & Kouzarides, T. Cancer epigenetics: from mechanism to therapy. Cell 150, 12–27 (2012).

Heintz, N. & Dhaenens, D. Chromatin proteins and modifications as drug targets. Nature 502, 480–488 (2013).

Polak, P. et al. Cell-of-origin chromatin organization shapes the mutational landscape of cancer. Nature 518, 360–364 (2015).

Adam, R. C. et al. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature 521, 365–370 (2015).

Brown, J. D. et al. NF-κB directs dynamic super enhancer formation in inflammation and atherosclerosis. Mol. Cell 56, 219–231 (2014).
METHODS

Human glioblastoma specimen culture conditions. All human glioblastoma tissues were obtained from excess surgical materials from consented patients after review from a neuropathologist and used in accordance with an approved protocol by the Institutional Review Board at Cleveland Clinic. As previously described27, glioblastoma cells were derived immediately after dissociation of primary patient tumour or after transient xenograft passage. For all in vitro studies, glioblastoma cells were cultured in Neurobasal medium (Gibco) with B27 (without vitamin A, Thermofisher), basic fibroblast growth factor (20 ng ml\(^{-1}\), R&D) and epidermal growth factor (20 ng ml\(^{-1}\), R&D). Primary glioblastoma models were validated to be unique by short-tandem repeat analysis at multiple xenograft passage numbers (analysis conducted by ATCC and Duke University, Cell Line Authentication Service). Short-tandem repeat results are available upon request. All cells used were derived from primary patient tissues, and are not included in the database of commonly misidentified cell lines maintained by the IACUC. All cell lines are routinely tested for mycoplasma contamination and were negative.

Patient-derived orthotopic xenografts. All animal procedures were performed in accordance with Cleveland Clinic Institutional Animal Care and Use Committee (IACUC) approved protocols. For intracranial tumours, size cannot be measured directly. Therefore, specific neurological signs indicating the presence of brain tumours (such as seizures, ataxia and lethargy) are monitored, in accordance with the IACUC approved protocol, and mice were always euthanized when these signs became apparent. The number of animals included in each of the described studies was based on extensive past experience in the development and use of glioblastoma xenograft models by our group. Each study was designed to minimize unnecessary animal use, optimize statistical power, and account for known variance in each model system. For all in vivo experiments, human glioblastoma cells were intracranially implanted into nude mice ( athymic BALB/c) to serve as a model for a primary glioblastoma. To ensure the presence of glioblastoma xenografts, we used a second plating of glioblastoma cells in each xenograft, which was performed 14 days post-implantation. The second xenograft was performed with the glioblastoma cells grown in serum-containing medium. If glioblastoma xenografts were not observed at the second xenograft, the mice were excluded from the study. The number of animals used was 10 per group for the induced arm and 10 per group for the control arm. Doxycycline (1 mg \(\mu\)g ml\(^{-1}\) water (2 mg ml\(^{-1}\) doxycycline and 2% sucrose, Sigma-Aldrich) for the duration the shRNA library was used. The representation of shRNAs from each cell population was monitored by calculating depletion scores for each shRNA, based on the decrease in shRNA expression levels in cells actively expressing an shRNA at the time of collection, and markedly reduces leakiness. It also contains an enhanced design of the miR30 backbone to allow for more efficient shRNA processing. In addition, the second reporter is inducible and is expressed in conjuction with the shRNA, which allows for selection of cells actively expressing an shRNA at the time of collection, and markedly reduces the number of cells included in the final analysis that contained silenced shRNAs. We shuttled the library of 1,586 shRNAs (shERWOOD Epigenetics-related genes library, transOMIC) into the RT3REVIR vector using previously published methods4. To produce virus, retroviral plasmid library was transfected into HEK293T Phoenix packaging cells as previously described2. Chloroquine (25 \(\mu\)M, Sigma-Aldrich) was added to enhance plasmid stability.

Functional screening assays. Approximately 60 million patient-derived glioblastoma cells were transduced with the pooled inducible shRNA library at 1.1% transduction efficiency to ensure that cells were transduced with only 1 shRNA per cell. Protamine sulfate (Sigma-Aldrich) was used for transduction. Successfully transduced cells were selected by a constitutive Venus fluorescent reporter using FACS. Cells were allowed to recover and expand for three passages. The same population of cells was used simultaneously for a screen completed in triplicate in standard serum-free cell culture conditions and for a screen completed in an in vivo intracranial xenograft mouse model.

In vivo and in vitro inducible shRNA screens. To investigate the effects of targeting epigenetic regulators in glioblastoma cells within the tumour microenvironment, we modified an inducible in vivo RNAi screening system, previously used for screening in haematological malignancies2,29, for use in solid tumours (Fig. 1a). The advanced inducible shRNA delivery vector has high fidelity expression and dual fluorescent reporters to reduce the noise and bias that weaken traditional shRNA screens28,29 (Extended Data Fig. 1a–c). Our shRNA library contained 1,586 shRNAs targeting 406 known chromatin and transcriptional regulator genes (2–4 shRNAs per gene), with positive and negative control shRNAs. We shuttled the library of 1,586 shRNAs (shERWOOD Epigenetics-related genes library, transOMIC) into the RT3REVIR vector using previously published methods4. To produce virus, retroviral plasmid library was transfected into HEK293T Phoenix packaging cells as previously described2. Chloroquine (25 \(\mu\)M, Sigma-Aldrich) was added to enhance plasmid stability.

In vivo and in vitro inducible shRNA screens. To investigate the effects of targeting epigenetic regulators in glioblastoma cells within the tumour microenvironment, we modified an inducible in vivo RNAi screening system, previously used for screening in haematological malignancies2,29, for use in solid tumours (Fig. 1a). The advanced inducible shRNA delivery vector has high fidelity expression and dual fluorescent reporters to reduce the noise and bias that weaken traditional shRNA screens28,29 (Extended Data Fig. 1a–c). Our shRNA library contained 1,586 shRNAs targeting 406 known chromatin and transcriptional regulator genes (2–4 shRNAs per gene), with positive and negative control shRNAs. We shuttled the library of 1,586 shRNAs (shERWOOD Epigenetics-related genes library, transOMIC) into the RT3REVIR vector using previously published methods4. To produce virus, retroviral plasmid library was transfected into HEK293T Phoenix packaging cells as previously described2. Chloroquine (25 \(\mu\)M, Sigma-Aldrich) was added to enhance plasmid stability.

In vivo and in vitro inducible shRNA screens. To investigate the effects of targeting epigenetic regulators in glioblastoma cells within the tumour microenvironment, we modified an inducible in vivo RNAi screening system, previously used for screening in haematological malignancies2,29, for use in solid tumours (Fig. 1a). The advanced inducible shRNA delivery vector has high fidelity expression and dual fluorescent reporters to reduce the noise and bias that weaken traditional shRNA screens28,29 (Extended Data Fig. 1a–c). Our shRNA library contained 1,586 shRNAs targeting 406 known chromatin and transcriptional regulator genes (2–4 shRNAs per gene), with positive and negative control shRNAs. We shuttled the library of 1,586 shRNAs (shERWOOD Epigenetics-related genes library, transOMIC) into the RT3REVIR vector using previously published methods4. To produce virus, retroviral plasmid library was transfected into HEK293T Phoenix packaging cells as previously described2. Chloroquine (25 \(\mu\)M, Sigma-Aldrich) was added to enhance plasmid stability.

In vivo and in vitro inducible shRNA screens. To investigate the effects of targeting epigenetic regulators in glioblastoma cells within the tumour microenvironment, we modified an inducible in vivo RNAi screening system, previously used for screening in haematological malignancies2,29, for use in solid tumours (Fig. 1a). The advanced inducible shRNA delivery vector has high fidelity expression and dual fluorescent reporters to reduce the noise and bias that weaken traditional shRNA screens28,29 (Extended Data Fig. 1a–c). Our shRNA library contained 1,586 shRNAs targeting 406 known chromatin and transcriptional regulator genes (2–4 shRNAs per gene), with positive and negative control shRNAs. We shuttled the library of 1,586 shRNAs (shERWOOD Epigenetics-related genes library, transOMIC) into the RT3REVIR vector using previously published methods4. To produce virus, retroviral plasmid library was transfected into HEK293T Phoenix packaging cells as previously described2. Chloroquine (25 \(\mu\)M, Sigma-Aldrich) was added to enhance plasmid stability.

In vivo and in vitro inducible shRNA screens. To investigate the effects of targeting epigenetic regulators in glioblastoma cells within the tumour microenvironment, we modified an inducible in vivo RNAi screening system, previously used for screening in haematological malignancies2,29, for use in solid tumours (Fig. 1a). The advanced inducible shRNA delivery vector has high fidelity expression and dual fluorescent reporters to reduce the noise and bias that weaken traditional shRNA screens28,29 (Extended Data Fig. 1a–c). Our shRNA library contained 1,586 shRNAs targeting 406 known chromatin and transcriptional regulator genes (2–4 shRNAs per gene), with positive and negative control shRNAs. We shuttled the library of 1,586 shRNAs (shERWOOD Epigenetics-related genes library, transOMIC) into the RT3REVIR vector using previously published methods4. To produce virus, retroviral plasmid library was transfected into HEK293T Phoenix packaging cells as previously described2. Chloroquine (25 \(\mu\)M, Sigma-Aldrich) was added to enhance plasmid stability.

In vivo and in vitro inducible shRNA screens. To investigate the effects of targeting epigenetic regulators in glioblastoma cells within the tumour microenvironment, we modified an inducible in vivo RNAi screening system, previously used for screening in haematological malignancies2,29, for use in solid tumours (Fig. 1a). The advanced inducible shRNA delivery vector has high fidelity expression and dual fluorescent reporters to reduce the noise and bias that weaken traditional shRNA screens28,29 (Extended Data Fig. 1a–c). Our shRNA library contained 1,586 shRNAs targeting 406 known chromatin and transcriptional regulator genes (2–4 shRNas per gene), with positive and negative control shRNAs. We shuttled the library of 1,586 shRNAs (shERWOOD Epigenetics-related genes library, transOMIC) into the RT3REVIR vector using previously published methods4. To produce virus, retroviral plasmid library was transfected into HEK293T Phoenix packaging cells as previously described2. Chloroquine (25 \(\mu\)M, Sigma-Aldrich) was added to enhance plasmid stability.
Sigma-Aldrich) was added to cells in the induced arm at time of plating and cells were maintained on doxycycline for 21 days, which is the same amount of time that the last of the intracranial tumours were collected for the in vivo screen. Cells in the uninhibited control arm were also maintained for 21 days in culture without doxycycline. Cells had a doubling time of 1.5 days and were passaged every 3–4 days to prevent cells from becoming over-confluent. During passage, at least 2 million cells were re-seeded in the plate to maintain library representation. At the conclusion of the screen Venus-ΔRED+ cells (1.5 million for each replicate) were collected by FACs from the induced arm, while Venus-ΔRED- cells (1.5 million for each replicate) were collected by FACs for the uninduced arm.

**shRNA amplification and screening.** Each mouse was processed and analysed separately. Genomic DNA was isolated and sequenced as described with slight modification. Genomic DNA was isolated by two rounds of phenol extraction using PhaseLock tubes (5prime) followed by isopropanol precipitation. Deep sequencing libraries were generated by PCR amplification of shRNA guide strands using barcoded primers that tag the product with standard Illumina adapters (p7+loop, 5′-CGCACATAGCCAGTGATCGA-GNNNN (4 nucleotide barcode)-TAGTGAAGCAGACAGTGA-3′; p5+mir3′, 5′-AATGTACAGGGCCGACCACCGATGATGTGAAATGTGCGCAGG-3′). Libraries were sequenced on the HiSeq 2500 platform at the Cleveland Clinic Genomics Core Facility. Libraries were sequenced using a primer that reads in reverse into the guide strand (mir36EcorRSeq, 5′-TAGGCCCTTTGATTTCCGGAGC TATGGGCA-3′).

**Sequencing analysis and shRNA scoring.** Sequence processing was performed using two custom workflows at https://usegalaxy.org (see Supplementary Information). Annotated workflows are stored in the https://usegalaxy.org published workflows repository and can be accessed and used using the following links: Workflow 1: https://usegalaxy.org/u/tyleremiller/w/shrna-pipeline1. It is used first to generate an output needed for workflows 2: Workflow 2: https://usegalaxy.org/u/tyleremiller/w/shrna-pipeline2. Raw read counts were converted to reads per million (RPM) to control for variations in the total shRNA reads in each sample. Results from multiple mice were randomly pooled together to form replicates (n = 3 for the induced arm, with each replicate containing 13–14 mice, while n = 2 for the control arm, with each replicate containing 10 mice). To provide a sufficient baseline for detecting shRNA depletion (negative selection) in experimental samples, we aimed to acquire an average of more than 1,000 reads per shRNA in each replicate, which would require 1.6 × 106 reads per replicate. In practice, we achieved over 1.5 × 106 reads for all replicates (nearly 10,000-fold coverage). Pooling the mice was essential for achieving proper fold coverage and highly correlated replicates. The average correlation score between individual mice from the induced groups was R = 0.24, while the average correlation value between the three combined replicates was R = 0.80.

shRNAs were scored using RIGER, and extension of the GENE-E package (Broad Institute) (28). Median RPM value for each replicate was used for analysis. The signal-to-noise ratio of replicates was used to calculate individual shRNA score based on their ability to deplete cells in the induced arm compared to the control arm and second-best shRNA score was used to rank genes. Expressed genes with a shRNA depletion score greater than 0.74 were considered hit.

**Gene Ontology enrichment analysis and visualization.** Gene Ontology analysis for screen hits (Fig. 1 and Extended Data Fig. 3) was conducted using background correction to understand what was enriched above what would be expected for a list comprised of chromatin modifiers. Specifically, the gene list of in vivo or in vitro hits was imported into gProfiler (29) to generate enrichment scores for all Gene Ontology gene sets according to recommended settings for gProfiler (http://baderlab.org/Software/EnrichmentMap/gProfilerTutorial), with the exception that we used the full list of genes included in the screen as a background list (‘Gene list as a stat. background’ feature) to control for bias towards chromatin modifiers.

**Analysis when starting with whole transcriptome data.** For Gene Ontology analysis when analysing whole transcriptome data (Fig. 2 and Extended Data Fig. 4), data were imported into GSEA (30) software to generate enrichment scores for gene sets in Hallmark, C2.all and C5.all MSigDB datasets. Exact settings are listed in Supplementary Table 4. Cytoscape (v3.2.1) and the Enrichment Map (31) plug-in was used to generate networks for gene sets enriched with a FDR cut-off of <0.05. For GSEA analysis in Extended Data Fig. 7, ranked lists were generated from whole transcriptome expression data of GBM528 or GBM3565 cells grown in vivo versus in vitro (for Extended Data Fig. 8, d), or from the correlations of genes with H3K27ac peak enrichment using the GSEA database (for Extended Data Fig. 8f, g). These ranked lists were run against a gene set consisting of genes targeted by JMJD6-bound enhancers in GBM528 or GBM3565 cells. Normalized enrichment scores and FDR values were generated by GSEA software.

**Global characterization of glioblastoma cells grown in vivo and in vitro.** To characterize the cell state of human tumour cells grown in vivo and compare it to those grown in culture, we used the same experimental setup as was done in the screen (schematic in Fig. 2a). We took patient-derived glioblastoma cells growing in vitro and split them in half. We injected half into 20 mice per glioblastoma model and maintained the other half in culture in independently maintained replicates. Multiple PDX models representing different molecular glioblastoma subtypes were included in the analysis, including our primary screen model, the proneural model GBM528, and the mesenchymal model GBM3565 (Supplementary Table 1). We also analysed data obtained from an independent laboratory from two other cerebral tumour models, which were grown in vivo (32) and in vitro (33). In addition, these two models had matched data from the corresponding primary tumour (34). For the in vivo arm, each mouse was intracranially implanted with 500,000 cells. When a mouse began to show overt neurological signs, which occurred between 2 and 3 weeks, it was euthanized, the tumour was collected. Macroscopically dissected, dissociated to single cells (Tumour Dissociation Kit, human and GentleMACS Octo Dissociator with Heaters, Miltenyi Biotech) and depleting of any remaining mouse cells using magnetic-activated cell sorting (Mouse Cell Depletion Kit, Miltenyi Biotech). We used two rounds of mouse cell depletion to ensure only human tumour cells were used in the analysis. We were consistently able to obtain greater than 98% purity of human cells with this method. Each mouse was processed independently. For RNA-seq samples, cells from 2–3 mice were pooled together to form each replicate. For ChIP–seq samples, many mice were pooled together for ChIP–seq as cell number was a limiting factor. Cells grown in culture were maintained for 3 weeks and then harvested for RNA-seq and ChIP–seq.

**RNA-seq.** RNA was extracted from human glioblastoma cells with TRIzol (ThermoFisher), separated using Phase Lock Gel tubes (5 Prime), and purified using the mirNAeasy kit (Qiagen) according to the manufacturer’s protocol. Total RNA was prepared for sequencing by Beckman-Coulter Genomics using the Illumina TruSeq Stranded Total RNA Library Prep Kit according to the manufacturer’s protocol. RNA-seq libraries were sequenced on the Illumina HiSeq 2500 platform by Beckman-Coulter Genomics. For gene expression analysis, reads were aligned to the hg19 genome build (retrieved from cufflinks.cbcb.umd.edu, currently available at http://cole-trapnell-lab.github.io/cufflinks/igenome_table/index.html), using TopHat v2.0.6. The distribution of alignments was analysed using the CollectRNASeqMetrics module of Picard v1.89 (http://picard.sourceforge.net/). Fragments per kilobase of transcript per million mapped reads (FPKM) values for known genes were calculated using Cufflinks (35) v2.0.2 provided with the GTF file via the -G (known genes only) option. FPKM values were quantile normalized. Differential expression testing was performed using Cuffdiff v2.0.2; however, all FPKM values provided are those calculated by Cufflinks. To generate an expressed genes list, an average of replicates for each condition was calculated and genes with FPKM values greater than 0.25 in both in intracranial sample or culture sample were considered expressed. Genes that did not meet this expression cut-off (replicate average FPKM < 0.25 in intracranial and culture conditions) were removed as not expressed. Expressed genes were tabled by converting FPKM values > 0.25 to 0.25.

**Exome sequencing.** DNA was collected from GBM528, GBM3565 and GBMcw1919 cells using standard techniques. The Case Genomics Core processed the DNA using the Illumina Nextera Rapid Capture Exome kit and sequenced on the Illumina HiSeq 2500 platform at greater than 100 × coverage. Data were processed by the Case Genomics Core to generate variant call format (VCF) files, and clinical variant calls were made using Omicia Opal Software (http://www.omicia.com/).

**ChIP–seq.** ChIP was performed as previously described (36). For histone modification and transcription factor ChIP–seq, 2 million cells (H3K27ac), 15–20 million cells (JMID6), or 5 million cells (Pol2) were crosslinked in PBS plus 1% formaldehyde for 10 min at 25 °C, quenched for 5 min with 125 mM glycine, washed twice in cold PBS with protease inhibitors (complete PI, Roche), and stored at −80 °C. In GBM, formaldehyde-fixed cells were lysed and sheared (Tekmar S200) on wet ice. The sheared chromatin was cleared and incubated overnight at 4 °C with the following antibodies: H3K27ac (Active Motif, M0474, JMID6 (Abcam, ab64575, lot GR54735-1), and total Pol II (Santa Cruz, sc-899-X, lot H0510). Antibody–chromatin complexes were immunoprecipitated with protein G magnetic Dynal beads (Life Technologies), washed, eluted, reverse crosslinked and treated with RNAse A followed by proteinase K. ChIP DNA was purified using Ampure XP beads (Beckmann Coulter) and then used to prepare sequencing libraries for sequencing with the Next-Seq Illumina genome analyser.

**Peak calling.** Reads were aligned to hg19 using Burrows–Wheeler Aligner (BWA) (37) with the following parameters: H3K27ac, -histone -tagThreshold 50 was used for enhancer analysis. JMID6, -factor.

**Enhancer landscape analysis.** To generate enhancer loci lists for each condition, H3K27ac ChIP–seq peak files were filtered to remove all peaks overlapping...
ENCODE blacklist regions for functional genomics analysis (https://sites.google.com/site/anshulkundaje/projects/blacklists) as well as peaks with any overlap occurring within ±1 kb from transcription start sites of all annotated RefSeq genes to exclude promoters. To call target genes of enhancers, enhancer loci were mapped to the nearest expressed gene within the glioblastoma cells. An expressed gene had to be within 200 kb for an enhancer loci to be mapped to a gene. The expressed gene list was the same as described in the RNA-seq methods section above. Peaks were visualized with the Integrative Genomics Viewer (IGV, Broad Institute).

Condition-specific enhancers. H3K27ac ChIP–seq enhancer lists from in vivo intrachromosomal xenograft and in vitro cultured cells were merged to create a single peak file. RPKM values within merged peaks were calculated. In vivo–specific or in vitro–specific enhancers were called as peaks with threefold increased or decreased RPKM values in vivo relative to in vitro, respectively.

Enhancer data presentation. Heatmaps, aggregate plots and motif analysis were conducted using the Cistrome galaxy instance. Heatmaps were created using the Heatmap tool, with 5,000 bp upstream and downstream, 200 bp step and saturation of 0.05. Aggregative plots were created using the aggregation plot tool, using a span of 3000 bp and resolution of 200 bp.

Super-enhancers. In vivo and in vitro super-enhancers were identified using the dynamicEnhancer software (retrieved from https://github.com/BradnerLab/dynamicEnhancer). Enhancer data presentation.

Genomic distribution of JMJD6. To call the genomic distribution of JMJD6 binding (Extended Data Fig. 8c), each JMJD6 ChIP–seq binding peak was assigned to the underlying genomic element. Enhancers were defined as described earlier; areas surrounding TSSs were defined as ±1 kb of all annotated TSSs in hg19, and other elements were defined by hg19 annotation. Each peak was assigned to a single element. In the cases in which a JMJD6 binding peak overlapped with two elements (for example, enhancer and intron), priority was assigned in the following order: enhancer; TSS ±1 kb; 3′ UTR, 5′ UTR, exons; introns; and intergenic.

Enrichment analysis of JMJD6 binding. To determine enrichment scores of JMJD6 peaks at a particular genomic element (for example, exons) (Fig. 3d), a null binding distribution of JMJD6 peaks for each element was generated by randomizing the peaks throughout the genome 1,000 times using bedtools shuffle. A z-score enrichment value at each genomic element was calculated using the z-score enrichment value at each genomic element was calculated using the 1/(size of peak in bp/1,000))/number of aligned

Validation studies of in vivo screen hits. We used the constitutive shRNA vector SGEPl, which is a constitutive version of the lentiviral vector used above. It allows for puromycin selection and has a constitutive GFP fluorescent reporter. For validation of 9 primary screen hits (Extended Data Fig. 3e–h), including the positive control, patient-derived glioblastoma GBM528 cells were transduced with individual shRNAs targeting primary screen hits, or 2 independent negative controls (totaling 11 independent transduced populations, each transduced with a single shRNA). Infected populations of GBM528 cells were selected by 1 μg ml−1 puromycin for 3 days and immediately implanted intracranially into age-matched female NSG mice; 4 mice were used for each experimental and control group. All mice were monitored daily until development of neurological signs, at which time they were euthanized.

In parallel, we conducted more in-depth validation of a top scoring transcription elongation factor in the primary screen, DOT1L, as well as another top screen hit that was not a transcription elongation factor, DPY30 (Fig. 4c, d and Extended Data Fig. 9i–x). We used the top two scoring shRNAs from the screen for each gene and used three independent glioblastoma models, GBM3565, GBM528 and GBMcw1919. As above, we transduced cells with shRNAs targeting DOT1L, DPY30 or a non-targeting shRNA. Infected populations were selected by 1 μg ml−1 puromycin for 5 days and knockdown of target was confirmed by qRT–PCR. To determine the effect of target knockdown on cells in vitro and in vivo, the same populations of cells collected at the same time point were used in in vitro proliferation assays and for in vivo survival studies in intracranial PDX models. For in vivo studies, cells were plated in 96-well plates on Geltrex as above and maintained in standard serum-free media. Cell proliferation was measured using Cell-Titer Glow (Promega). For in vivo studies, cells were implanted intracranially into age-matched NSG mice. All mice were monitored daily until development of neurological signs, at which time they were euthanized.

shRNA sequences. JMJD6.100S: 5′-TGCTGTGGAGCAGACCGAAAGCAGAAGCAGATGGATAGGAGCACACAGTATCTGCGGATTCGTTGTGGCTCTACATGGCAGA-3′

3′-AACCTTTACATCTTTTG-5′.

The CRISPR design tool from Broad Institute (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design) was used to design the guide RNA (gRNA). Oligonucleotides were purchased from Fisher, and annealed and cloned into LentiCRISPR v2 plasmid, which was a gift from E. Zhang (Addgene plasmid 52961). The oligonucleotides used were as follows: JMJD6: 5′-ACACCTCTATCTCTACTGATAGGACCGAGGGAATTAAC-3'; 5′-TGCTGTGGAGCAGACCGAAAGCAGAAGCAGATGGATAGGAGCACACAGTATCTGCGGATTCGTTGTGGCTCTACATGGCAGA-3'; 5′-ACACCTCTATCTCTACTGATAGGACCGAGGGAATTAAC-3'.

Experimental design for complete knockout in clonal populations. For complete knockout studies (Fig. 4a–c), GBM528 cells were transduced with one of two independent CRISPR–Cas9 constructs targeting JMJD6 or a non-targeting control and

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
selected for integration of the lentiviral construct by puromycin. Single cells were expanded in vitro to obtain clonal populations and knockout was confirmed by western blot (see details below). Two clonal populations per sgRNA were subjected to parallel in vitro proliferation assays and in vivo survival assays. For in vitro studies, cells were plated in 96-well plates on Gelrett as above and maintained in standard serum-free media. Cell proliferation was measured using AlamarBlue cell viability reagent (ThermoFisher). For in vivo studies, cells were intracranially implanted into age-matched female NSG mice. Five mice for each clone, or ten mice for each sgRNA construct were used. All mice were monitored daily until development of neurological signs, at which time they were euthanized.

**Experimental design for knockout in bulk populations without clonal selection.**

Owing to the knockout efficiency being extremely high in the experiment above, we used CRISPR-mediated knockout of JMJD6 on a population of cells without clonal selection to confirm our results in another PDX glioblastoma model (Extended Data Fig. 9e–g). GBM21191 cells were transduced as above, selected by puromycin, subjected to western blot to confirm population based knockdown, and then assayed in vitro and in vivo as above. For in vivo survival studies, five age-matched female NSG mice were used for each sgRNA construct.

**Western blotting for JMJD6 to determine knockout.** Cells were collected and lysed in hypotonic buffer with non-ionic detergent (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.5% NP-40; 50 mM NaF with protease inhibitors), incubated on ice for 15 min, and cleared by centrifugation at 10,000 × g for 10 min. Protein concentration was determined using the Bradford assay (Bio-Rad Laboratories). Equal amounts of protein were mixed with reducing Laemmli loading buffer, boiled, and electrophoresed in NuPAGE Gels (ThermoFisher), and then transferred to PVDF membranes (Millipore). Blocking was performed for 30 min with 5% nonfat milk in TBS plus 0.1% Tween 20 (TBST) and blotting performed with primary JMJD6 antibody (Antibody Motif, 61494) for 16 h at 4 °C.

**Tissue processing, histology and imaging.** Tissue processing for histological sections was performed as reported previously36. Primary antibodies used: GFP to stain for Venus (1:250; Aves Labs GFP-1020), mCherry to stain for dRRED (1:250; Abcam Ab167453) and human nuclear antigen to stain for human cells (1:250; NovusBio clone 235-1 NBP2-34525V3). Species-specific Alexa-Fluor-conjugated secondary antibodies were used for detection (1:500; ThermoFisher).

**Tissue microarray immunohistochemistry and analysis.** In brief, de-identified tissue microarrays were constructed from gliomas after obtaining University of Kentucky Institutional Review Board Approval. Three 2-mm diameter cores per tumour were obtained, with each core embedded in a separate tissue microarray block. A total of 104 cases comprised the TMAs, including 9 nonneoplastic controls (cortical dysplasias), 9 grade II astrocytomas, 11 grade III astrocytomas, 12 anaplastic oligodendrogliomas, 16 grade II oligodendrogliomas, and 47 grade IV glioblastomas. Immunohistochemistry was performed for each core as described previously37, but using an antibody towards JMJD6 (Abcam, ab64575). In brief, each core was semi-quantitative on a relative scale from 0 to 3, with 0 = negative and 3 = strongest. Results from all three cores were averaged together to produce a final score for a tumour. Results were plotted based on WHO grade and differences were calculated via Mann–Whitney–Wilcoxon test.

**Retrospective analysis of gene expression in human gliomas.** Gene expression correlations across primary patient glioblastoma tumours, expression of individual genes or gene signatures in primary patient gliomas, and patient survival were determined through analysis of the Allen Institute Ivy GAP (http://glioblastoma.alleninstitute.org), the National Cancer Institute Repository for Molecular Brain Neoplasia Data (REMBRANDT, https://caintegrator.nci.nih.gov/rembrandt/), or the TCGA (https://tcga-data.nci.nih.gov/tcga/). Normalized IypGFP (http://glioblastoma.alleninstitute.org) and REMBRANDT datasets were downloaded from GlioVis (http://gliovis.bioinfo.cnio.es/), and TCGA glioblastoma RNAseqV2 dataset was downloaded from the Broad Institute GDAC Firehose via the TCGA2STAT package on R38.

**Correlation analysis of gene expression and gene signature scores.** We created a signature of the 55 genes upregulated in vivo more than 2.5-fold in both GBM528 and GBM3565 models (Fig. 3a and Extended Data Fig. 7a–d). Gene signature scores within IypGFP samples were analysed via single-sample GSEA on GenePattern and normalized as z-scores across all samples in the given dataset39,40. Correlation between gene signature z-scores and median-centred gene expression was determined as Pearson coefficient (r) values. Associated FDR-adjusted P values were calculated using the Benjamini–Hochberg procedure. MatLab (MathWorks) was used to determine correlation values and P values.

**Individual gene correlation with JMJD6.** To test correlation of all genes with JMJD6 across glioblastoma tumours (Extended Data Fig. 7e), TCGA RNAseqV2 data were downloaded and analysed using R (http://www.r-project.org). The function ggpairs, available through the GGally package, was used to generate plots and determine pairwise correlation coefficients (http://CRAN.R-project.org/package=GGally). All glioblastoma tumours in the TCGA with RNA-seq data available were used.

**Survival analysis.** For gene expression changes, high and low groups were defined as above and below the median, respectively. Kaplan–Meier curves were generated and log-rank (Mantel–Cox) analysis was performed using GraphPad Prism software (GraphPad Software).

**Direct comparison of TCGA RNA-seq gene expression data to our RNA-seq data.** To evaluate directly the expression/activity level of genes in an existing glioblastoma gene expression databases (Extended Data Fig. 5d, e), we used the RNAseqV2 TCGA database of glioblastoma tumours that were characterized by mRNA sequencing. FPKM data for RNAseqV2 were downloaded from https://gdc-portal.ncl.nih.gov/.

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

**Data availability.** The datasets generated during and analysed during the current study are included within the published manuscript (and Supplementary Information) or have been deposited in Gene Expression Omnibus (RNA-seq and ChIP-seq data) under accession number GSE74529. All other data are available from the corresponding authors upon reasonable request.

27. Bao, S. et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature 444, 756–760 (2006).
28. Goecks, J., Nekrutenko, A., Taylor, J. & The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. 11, R86 (2010).
29. Reimand, J., Arak, T. & Vilo, J. g:Profiler—a web server for functional interpretation of gene lists (2011 update). Nucleic Acids Res. 39, W307–W315 (2011).
30. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1110 (2009).
31. Trapnell, C. et al. Transcript assembly and quantification of RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28, 511–515 (2010).
32. Suvà, M. L. et al. Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. Cell 157, 580–594 (2014).
33. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transformer. Bioinformatics 25, 1754–1760 (2009).
34. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589 (2010).
35. Liu, T. et al. Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol. 12, R83 (2011).
36. Miller, T. E. et al. Lgr5 marks post-mitotic, lineage restricted cerebellar granule neurons during postnatal development. PLoS One 9, e114433 (2014).
37. Gilbert, M. R. et al. Autophagy and oxidative stress in gliomas with IDH1 mutations. Acta Neuropathol. 127, 221–233 (2014).
38. Wang, Y. W., Allen, G. J. & Liu, Z. TCGA2STAT: simple TCGA data access for integrated statistical analysis in R. Bioinformatics 32, 952–954 (2016).
39. Barbie, D. A. et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462, 108–112 (2009).
Extended Data Figure 1 | shRNA delivery vector performance.

a, Top, schematic of the RT3REVIR shRNA delivery vector. Middle, once integrated into cells, a constitutive PGK promoter drives Venus-fluorescent reporter and rtTA through an IRES element, creating an all-in-one inducible vector. Bottom, when doxycycline is introduced to cells, it binds to rtTA and drives activity of the third generation TET-inducible promoter. This drives the dsRED fluorescent reporter and shRNA expression. In addition, it drives higher rtTA transcription through the IRES element, creating a positive-feedback loop that increases rtTA expression in the cell resulting in higher expression of inducible elements.

b, The inducible shRNA delivery vector displays almost no unintended induction. Representative FACS plots from the parallel screen of cells infected with RT3REVIR with and without doxycycline treatment in vitro (left) and in vivo (right). c, RT3REVIR robustly expresses shRNAs and depletes cells expressing cell-lethal shRNAs in a competitive proliferation assay. Left, representative FACS plots over time of cells infected with a positive control shRNA against RPA3 and induced. Right, quantification of fluorescent cells in the representative competitive proliferation assay. Empty denotes cells with vector that had no shRNA; REN.713 denotes cells with vector containing negative-control shRNA targeting the Renilla protein. The percentage of cells actively expressing the shRNA within the total infected population from a single dish is shown.
Extended Data Figure 2 | Importance of combining multiple mice to achieve increased reproducibility. a, Correlation (r) values between individual mice (41 pairwise comparisons for induced mice and 20 pairwise comparisons for control uninduced mice) or of grouped replicates containing multiple mice (3 pairwise comparisons for triplicate induced arm replicates and 1 pairwise comparison for the duplicate control uninduced replicates). b, Positive control gene RPA3 was effectively depleted from cell populations in both in vivo and in vitro screens using grouped replicates for the in vivo screen. Four shRNAs targeting RPA3 were included in the shRNA screening library. At least two out of four shRNAs achieved a RIGER depletion score of 2.0 or greater.
Extended Data Figure 3 | See next page for caption.
Extended Data Figure 3 | Validation of in vivo screen results.

a, b, Average mRNA expression of intracranial-specific (a) or cell culture-specific (b) hits in vivo and in vitro. *P* values calculated by two-sided Mann–Whitney *U*-test and were greater than 0.05. c, d, Screen hits were analysed for enrichment of Gene Ontology gene sets, using the screened library gene list as background to control for bias towards chromatin modifiers. c, All results with an FDR value of 0.05 or lower for in vivo-specific hits are presented. d, Top 10 results with an FDR value of 0.05 or lower for in vitro-specific hits are presented. Significance calculated by Benjamini–Hochberg FDR.

e–h, Validation of elongation factor hits by in vivo survival assays. shRNAs from the primary screen were used to transduce GBM528 cells using a constitutive expression vector. Primary screen hits that led to an increase survival with knockdown, with at least 2 out of 4 mice surviving longer than all 9 negative control mice, were considered validated. e, f, Target mRNA knockdown by qRT–PCR (e) and Kaplan–Meier survival curve (f) of validated hits. Black lines and bars represent two independent negative-control shRNAs. Purple and pink lines and bars represent validated in vivo-specific hits; blue lines and bars represent validated common hits found in both in vivo and in vitro screen. g, h, Target mRNA knockdown by qRT–PCR (g) and Kaplan–Meier survival curve (h) of primary screen hits that did not validate (green lines and bars).
Extended Data Figure 4 | Independent models confirm stimulus-controlled and stress response programs upregulated in vivo and in primary tumours. a, b, Genes with an average expression change of more than 2.5-fold between conditions in GBM528 (a) and GBM3565 (b) cells, as determined by RNA-seq. GBM528 heatmap associated with Fig. 2b, c. c, d, Cellular programs enriched by GSEA in cells grown in each condition. c, Representation of all enriched programs in vitro and in vivo using Enrichment Map. d–g, As in b–d, GSEA was performed using data generated from two independent models by Lee and colleagues on cells grown in vitro versus in vivo intracranial xenograft tumours (e, g), or in vitro versus the primary glioblastoma from which the cells were derived (f, h). Cellular programs enriched in cells grown in each condition are presented using Enrichment Map. FDR values calculated by GSEA software.
Extended Data Figure 5 | Transcription factors and signalling molecules that drive stimulus-controlled programs consistently upregulated in vivo. a, Transcription factors upregulated in GBM528 cells upon growth in vivo. Values are mean FPKM ± s.d. from biological duplicates. b, Cell signalling programs regulated by pause control that are enriched for upregulated transcription factors in a. FDR values calculated by MSigDB for enrichment against all genes. c, 55 genes upregulated more than 2.5-fold in both GBM528 and GBM3565 upon growth in vivo. d, Expression of those 55 genes in vivo and in vitro in GBM528 and GBM3565 compared to expression in primary glioblastoma tumours from the TCGA RNaseqV2 database. Data were FPKM quantile normalized across all datasets before plotting. e, All genes in datasets shown to confirm normalization.
Extended Data Figure 6 | Epigenomic regulation of glioblastoma cells is microenvironment-specific. a–h, Global enhancer landscape of GBM528 (a) and GBM3565 (d) cells in both conditions and microenvironment-specific enhancers. b–h, Browser track examples (b, e), aggregate plots (c, f), and gene expression fold change of target genes (nearest expressed gene) (g, h) of microenvironment-specific enhancer loci from (a, d). i, Super-enhancers identified in vivo in GBM528. j, Browser track examples of condition-specific super-enhancers. k, Super-enhancers specific to each condition were identified. l, Expression of condition-specific super-enhancer target genes. Box plot P values calculated by two-sided Mann–Whitney U-test.
Extended Data Figure 7 | See next page for caption.
Extended Data Figure 7 | Prioritization of JMJD6 as lead target.

a, Top, expression across all Ivy GAP samples of gene signature of 55 genes upregulated in vivo in both PDX models. Middle, elongation factor hits. Bottom, corresponding histological tumour structure and TCGA molecular subtype of each sample represented. The signature and JMJD6 expression is highest in hypoxic regions, which also corresponds to more mesenchymal-like regions of the tumour. b, Expression correlation dot plot for JMJD6 data represented in a, c, d. Expression correlation of each elongation factor hit with gene signature of 55 genes upregulated in vivo in both PDX models across all REMBRANDT (c) and TCGA (d) glioblastoma tumours (bulk tumour expression. Ivy GAP data in a and Fig. 3a is microenvironment-specific expression. P value (a–d) by FDR-adjusted Benjamini–Hochberg procedure. e, JMJD6 mRNA-seq correlation with each gene across TCGA glioblastoma tumours. f, Example plots from GSEA using the gene correlations in e and a pre-rank list. g, Representative images from tissue microarray analysis of JMJD6 protein expression in Fig. 3c. h, Primary screen results for the four shRNAs targeting JMJD6. Only two out of the four shRNAs were represented in the library at appreciable levels, and both led to a RIGER depletion score of greater than 2. Values are median RPM ± s.d. of three biological replicates for induced populations and two biological replicates for the uninduced population.
Extended Data Figure 8 | JMJD6 regulates enhancer mediated pause–release in glioblastoma. a, Known role of JMJD6 in transcription pause–release. In HEK293T and HeLa cells, JMJD6 acts with BRD4 as a key activator of enhancer–mediated pause–release at genes controlled by Pol II pausing. Upon enhancer activation, JMJD6 demethylates 7SK RNA releasing positive transcription elongation factor (P–TEFb) inhibition from the 7SK/HEXIM complex. b, Browser tracks of JMJD6 at enhancers and TSSs. c, Left, global distribution of genomic elements as determined by the hg19 reference genome and H3K27ac ChIP–seq. Right, global distribution of JMJD6 binding peaks per genomic element as determined by JMJD6 ChIP–seq. Enrichments shown in Fig. 3d. d–g, GSEA enrichment plots of genes with JMJD6–bound enhancers in the GBM528 (d) or GBM3565 (e) PDX model against differential expression of genes between in vivo and in vitro conditions (expression from Fig. 2b for GBM528 and Extended Data Fig. 4a for GBM3565). GSEA enrichment plots of genes with JMJD6–bound enhancers in the GBM528 (f) or GBM3565 (g) PDX model against gene correlations with JMJD6 in TCGA tumours (correlations from Extended Data Fig. 7e). h, i, Distribution of the pausing index of the common in vivo upregulated genes from Extended Data Fig. 5c for which the pausing index could be determined in GBM528 (h) and GBM3565 (i). All P values calculated by two-sided Mann–Whitney U–test.
Extended Data Figure 9 | Validation of JMJD6 and other hits in multiple PDX models of glioblastoma. a, JMJD6 mRNA expression by qRT–PCR after inducible shRNA knockdown of JMJD6. b, c, In vitro proliferation (b) and in vivo survival (c) compared to uninduced and induced non-targeting controls. Data are mean ± s.d. of three technical replicates. d, Endpoint tumours collected from the induced arm of c stained to show human tumour cells (human nuclear antigen) that harbour a JMJD6 shRNA (Venus+ dsRED−). The vast majority of tumour cells at endpoint had silenced the shRNA (Venus+ dsRED−). Scale bar, 200 μm. e, CRISPR-mediated knockout of JMJD6 in a bulk population of GBMcw1919 cells in vitro. f–g, Parallel in vitro proliferation assay (f) and in vivo survival assay (g) of cells from e. h–l, Constitutive shRNA knockdown of DOT1L and DPY30 in vitro. m–v, Parallel in vitro proliferation assays (m–q) and in vivo survival assays (r–v) of cells from h–l, respectively. Data are mean ± s.d. of at least triplicates.
Extended Data Figure 10 | Summary figure. Overview summary of the results. The *in vivo* tumour microenvironment, both in primary glioblastoma tumours and intracranial xenograft tumours, is complex and stressful for cells. Tumour cells must appropriately interact with and respond to a large number of other cells, both cancerous and non-cancerous, to survive. They also must activate response pathways to survive in the face of reduced nutrient availability, including hypoxic and low glucose conditions, and in the face of increased cell stress due to immune regulators, and debris and signalling from apoptotic cells. Thus, slower growth is seen as the cells expend energy on responding to these microenvironmental stimuli to survive. Owing to the large number of pause-controlled genes needed to respond to the cell stresses *in vivo* appropriately, cells are dependent on transcriptional pause–release and elongation. By contrast, cell culture conditions are optimized to reduce cell stress and drive growth by providing a surplus of all required nutrients for cell growth. Cells are largely homogenous and cancerous. Together, these *in vitro* conditions lead to rapid cell growth and little need for pause-controlled pathways that respond to environmental stimuli and stress. Therefore, *in vitro* cells are not as dependent on transcriptional pause–release and elongation for growth and survival. We discovered this by conducting parallel screens in the *in vivo* and *in vitro* environments, and found that there were many genes necessary for survival *in vivo* that were not necessary, or less important, in cell culture. We focused on the transcriptional pause–release and elongation factors as those fell out as enriched, but there are many other targets to explore.