LETTER

Insulator dysfunction and oncogene activation in *IDH* mutant gliomas

William A. Flavahan[1,2,3]*, Yotam Drier[1,2,3]*, Brian B. Liau[1,2,3], Shawn M. Gillespie[1,2,3], Andrew S. Venteicher[1,2,4], Anat O. Stemmer-Rachamimov[1], Mario L. Suvà[1,2] & Bradley E. Bernstein[1,2,3]

Gain-of-function *IDH* mutations are initiating events that define major clinical and prognostic classes of gliomas[1,2]. Mutant IDH protein produces a new onco-metabolite, 2-hydroxyglutarate, which interferes with iron-dependent hydroxylases, including the TET family of 5′-methylcytosine hydroxylases[3–7]. TET enzymes catalyse a key step in the removal of DNA methylation[8,9]. *IDH* mutant gliomas thus manifest a CpG island methylator phenotype (G-CIMP)[10,11], although the functional importance of this altered epigenetic state remains unclear. Here we show that human *IDH* mutant gliomas exhibit hypermethylation at cohesin and CCCTC-binding factor (CTCF)-binding sites, compromising binding of this methylation-sensitive insulator protein. Reduced CTCF binding is associated with loss of insulation between topological domains and aberrant gene activation. We specifically demonstrate that loss of CTCF at a domain boundary permits a constitutive enhancer to interact aberrantly with the receptor tyrosine kinase gene *PDGFRA*, a prominent glioma oncogene. Treatment of *IDH* mutant gliomaspheres with a demethylating agent partially restores insulator function and downregulates *PDGFRA*. Conversely, CRISPR-mediated disruption of the CTCF motif in *IDH* wild-type gliomaspheres upregulates *PDGFRA* and increases proliferation. Our study suggests that *IDH* mutations promote gliomagenesis by disrupting chromosomal topology and allowing aberrant regulatory interactions that induce oncogene expression.

The human genome is organized into topological domains that represent discrete structural and regulatory units[12]. Such domains are evident in genome-wide contact maps generated by high-throughput chromatin conformation capture (HiC) techniques[13], and have been termed 'topologically associated domains' or 'contact domains'[14–16]. Recent studies have strengthened the role of the CTCF insulator protein in creating chromatin loops and boundaries that partition such domains[15]. Genomic alterations that remove CTCF-associated boundaries allow aberrant enhancer-gene interactions and alter gene expression[17].

Since CTCF binding is methylation-sensitive[18,19], its localization might be altered by DNA hypermethylation in *IDH* mutant gliomas. We therefore used chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) to map CTCF binding genome-wide in 11 primary tumours and 4 glioma cell lines. Although CTCF binding patterns tend to be relatively stable, we detected highly overlapping subsets of CTCF sites that were lost in *IDH* mutants (Fig. 1a, b and Methods). Significantly more sites were commonly lost than gained (625 versus 300, $P < 10^{-12}$). Whole-genome bisulfite sequencing data from The Cancer Genome Atlas (TCGA)[10] was used to assess the methylation status of 625 loci with reduced CTCF binding in mutant tumours. We found that these loci have higher GC content, and exhibit significantly higher levels of DNA methylation in *IDH* mutant gliomas relative to *IDH* wild type (Fig. 1c, d).

We considered that altered DNA methylation and CTCF binding might disrupt topological domain boundaries and gene insulation in *IDH* mutant tumours. We collated a set of constitutive domain boundaries based on kilobase (kb)-resolution HiC maps[15]. We then examined published RNA-seq expression data for 357 normal brain tissue samples[20]. Consistent with previous studies[16], we found that genes in the same domain correlate across samples, but that genes separated by a boundary show lower correlation (Fig. 1e). We next incorporated expression data for 230 *IDH* mutant (218 *IDH1* mutant and 12 *IDH2* mutant) and 56 wild-type lower-grade gliomas, generated by TCGA[2]. Here again we found that the presence of an intervening boundary reduces correlation between neighbouring genes. We next scanned the genome for pairs of proximal genes separated by less than 180 kb (the average contact domain size[15]) that correlate much more strongly in *IDH* mutants than in wild-type gliomas (Fig. 1f and Methods). Remarkably, the resulting set is strongly enriched for gene pairs that cross domain boundaries (90% versus 69% expected at random; $P < 10^{-4}$). Conversely, gene pairs that correlate less strongly in *IDH* mutants are more likely to reside in the same domain (52% versus 31% expected at random; $P < 10^{-5}$). Notably, CTCF knockdown has been shown to increase cross-boundary interactions and decrease intra-domain interactions[21]. Thus, altered expression patterns in *IDH* mutant gliomas may reflect reduced CTCF binding and consequent disruption of domain boundaries and topologies.

We next sought to pinpoint specific boundaries that were disrupted by *IDH* mutation. For all pairs of genes separated by <1 megabase (Mb), we computed their correlation across mutant and wild-type *IDH* gliomas. We then scanned for loci in which cross-boundary gene pairs correlate more strongly in mutant tumours (false discovery rates (FDR) < 1%), while intra-domain gene pairs correlate less strongly (FDR < 1%). This analysis highlighted 203 domain boundaries (Fig. 2a, Supplementary Table 1 and Methods). The putatively disrupted boundaries exhibit higher DNA methylation and lower CTCF binding in *IDH* mutant compared with wild-type tumours (Extended Data Fig. 1). These data suggest that the methylator phenotype disrupts CTCF binding and domain boundaries, thereby affecting gene expression in *IDH* mutant gliomas.

We hypothesized that altered domain topologies might contribute to gliomagenesis by activating oncogenes that are normally insulated by domain boundaries. We therefore scanned the domains adjacent to the disrupted boundaries for genes with higher expression in *IDH* mutant than in wild-type gliomas (Fig. 2a). Genes in top-scoring domains include *PDGFRA* ($P < 10^{-21}$), an established glioma oncogene[22], and other candidate regulators of gliomagenesis (Supplementary Table 1).

The identification of *PDGFRA* as a potential target of epigenetic deregulation in *IDH* mutants was of particular interest, given its prominence as a glioma oncogene and established roles for PDGFA signalling in the normal brain. Although *PDGFRA* is a frequent target of genomic
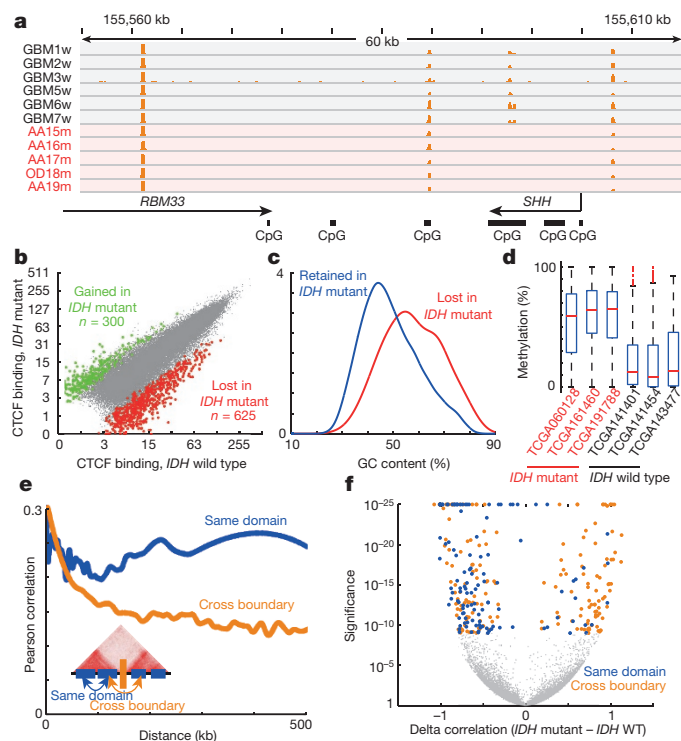
Figure 1 | CTCF binding and gene insulation compromised in *IDH* mutant gliomas. a, Binding profiles for the methylation-sensitive insulator CTCF are shown for a representative locus in *IDH* mutant and wild-type tumours, normalized by average signal. **b**, Scatterplot compares CTCF binding signals between *IDH* mutant (*y* axis) and wild-type (*x* axis) gliomas for all detected CTCF sites. A larger fraction of sites is commonly lost in all *IDH* mutants (*n* = 625) than gained (*n* = 300). **c**, Histogram compares GC content between CTCF sites that are lost or retained. **d**, Box plots show DNA methylation levels over lost CTCF sites, as determined by whole-genome bisulfite data for three *IDH* wild-type and three *IDH* mutant tumours. **e**, Plot depicts average correlation between gene pairs as a function of distance across RNA-seq profiles for human brain[20]. Gene pairs separated by a constitutive CTCF-bound boundary per HiC[15] have lower correlations. **f**, Volcano plot depicts the significance (*y* axis) of gene pairs that are more (or less) correlated in *IDH* mutant than in wild-type (WT) lower-grade gliomas. Gene pairs with significantly increased correlations in *IDH* mutants (right) tend to cross boundaries (orange), while those with decreased correlations (left) are more likely reside in the same domain (blue). These data indicate that *IDH* mutant, G-CIMP gliomas have reduced CTCF binding and altered expression patterns suggestive of defective gene insulation.

amplification and gain-of-function mutations in glioblastoma (15%), such alterations are rare in *IDH* mutant tumours[23,24]. Nonetheless, *IDH* mutant gliomas strongly express *PDGFRA* (Fig. 2b), and share the proneural transcriptional program characteristic of *PDGFRA*-amplified tumours[23,24]. Closer examination of the expression patterns in *IDH* mutant gliomas reveals a marked correlation between *PDGFRA* and *FIP1L1*, despite an intervening boundary (Fig. 2c). *FIP1L1* encodes an RNA-processing protein that is constitutively expressed in neural tissues, and particularly active in oligodendrocyte precursors, a putative glioma cell of origin[22] (Extended Data Fig. 2a). Moreover, combined expression of *PDGFRA* and *FIP1L1* is associated with poorer outcome in *IDH* mutant lower-grade gliomas (Extended Data Fig. 2b). This suggests that an aberrant interaction with this constitutive locus may drive *PDGFRA* expression in *IDH* mutant tumours.

We therefore investigated the topology of the region using kilobase-resolution HiC data[15]. In all six cell types examined, *PDGFRA* and *FIP1L1* reside in distinct domains, separated by one CTCF-anchored constitutive boundary (Fig. 3a and Extended Data Fig. 3). Our ChIP-seq data confirm that this boundary contains a strong CTCF-binding site over a canonical CTCF motif with a CpG dinucleotide
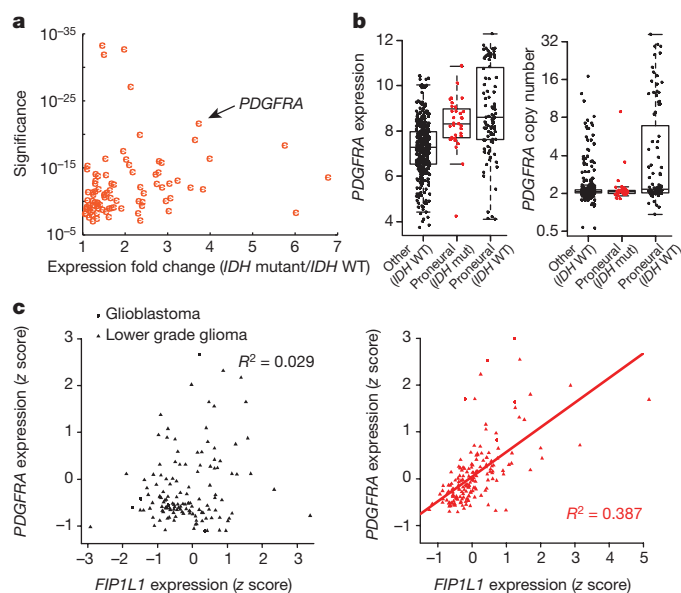

Figure 2 | Topological domain boundaries disrupted in *IDH* mutant gliomas. a, Scatterplot depicts significance of deregulated boundaries in *IDH* mutant tumours (*y* axis) against fold change of most upregulated gene in adjacent domains (*x* axis). *PDGFRA* is adjacent to a significantly deregulated boundary and upregulated in *IDH* mutants. **b**, Boxplots compare *PDGFRA* expression (left) or copy number (right) for 443 glioblastoma tumours, classified by *IDH* status and expression subtype[24]. *IDH* mutants (red) have increased *PDGFRA* expression, despite normal copy number. **c**, Plots compare *PDGFRA* (*y* axis) and *FIP1L1* (*x* axis) expression in *IDH* wild-type (left) and mutant (right) gliomas. The genes correlate specifically in *IDH* mutants, consistent with deregulation of the intervening boundary/insulator.

in a position previously linked to methylation-sensitivity[25] (Fig. 3b). Quantitative ChIP–PCR reveals that CTCF occupancy at this site is reduced between 30% and 50% in *IDH* mutant tumours and glioma-sphere models, relative to wild type (Fig. 3c, d). Moreover, the CpG in this motif becomes highly methylated in *IDH* mutants (Fig. 3e, f). This suggests that reduced CTCF binding may compromise the boundary flanking *PDGFRA* in *IDH* mutant, hypermethylated tumours.

To identify regulatory elements that might underlie *PDGFRA* induction, we mapped the enhancer-associated histone modification, histone H3 lysine 27 acetylation (H3K27ac), in glioma specimens and models. We identified a large enhancer ~50 kb upstream of *FIP1L1* with strong acetylation in wild-type and mutant tumours (Fig. 3a and Extended Data Fig. 4). In support of an enhancer identity, the element is enriched for H3 lysine 4 mono-methylation (H3K4me1), but lacks H3K4me3, and contains conserved motifs bound by the glioma master transcription factors OLIG2 and SOX2. Although this enhancer is normally insulated from *PDGFRA*, we reasoned that disruption of the intervening boundary might allow it to interact with the oncogene in *IDH* mutant gliomas. To test this, we used chromosome conformation capture (3C) to query the relative frequencies with which the *PDGFRA* promoter interacts with the *FIP1L1* enhancer, with an intragenic *PDGFRA* enhancer, or with nearby control sites (Fig. 3g). We fixed *IDH* mutant and wild-type glioma specimens and gliomaspheres, digested their chromatin with HinDIII, and performed proximity ligation to re-ligate physically interacting DNA sequences. We used quantitative PCR (qPCR) to measure ligation frequencies between elements, normalizing against control ligations performed with bacterial artificial chromosome DNA.

In wild-type gliomas, 3C revealed a strong interaction between the *PDGFRA* promoter and its intragenic enhancer, which are ~50 kb apart (Fig. 3j, k). In contrast, the *PDGFRA* promoter does not interact with the *FIP1L1* enhancer in wild-type tumours, consistent with retention of the intervening boundary (Fig. 3h, i). However, the interaction patterns
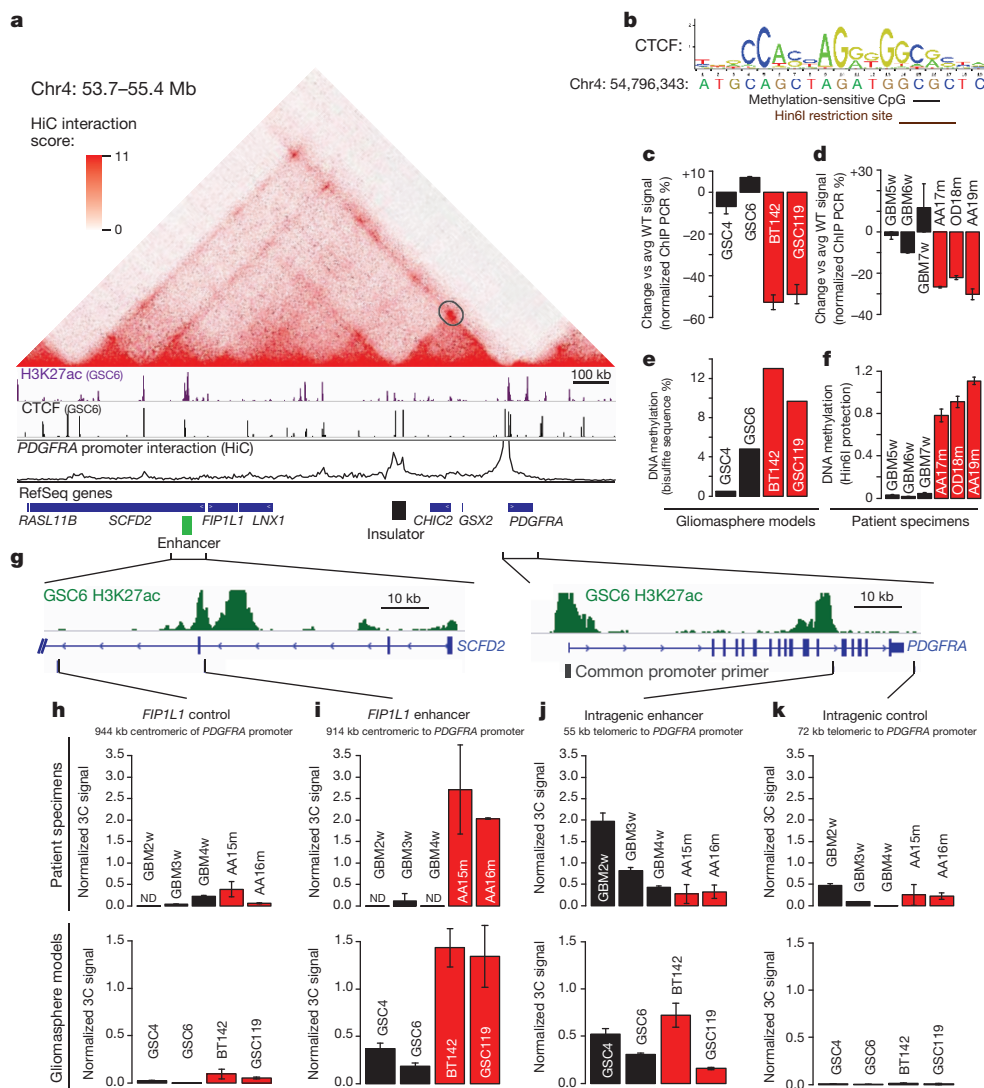
Figure 3 | Insulator loss allows *PDGFRA* to interact with a constitutive enhancer. a, Contact domain structure shown for a 1.7-Mb region containing *PDGFRA*. Heat depicts HiC interaction scores between triangulated loci in IMR90 cells[15]. Domains are visible as triangle-shaped regions of high interaction scores. Convergent CTCF sites anchor a loop that separates *PDGFRA* and *FIP1L1* (black circle). H3K27ac and CTCF profiles are aligned to the contact map. Interaction trace (below) depicts HiC signals between the *PDGFRA* promoter and all other positions in the region. Genes, *FIP1L1* enhancer (per H3K27ac) and insulator (per HiC and CTCF binding) are indicated. **b**, The right CTCF peak in the insulator contains a CTCF motif with a CpG at a methylation-sensitive position. **c, d**, ChIP-qPCR data show that CTCF occupancy over the boundary is reduced in *IDH* mutant (red) gliomas and models, relative to wild type

(black). **e**, Methylation levels of the CpG in the CTCF motif were measured in gliomaspheres by bisulfite sequencing, and plotted as a percentage of alleles protected from conversion. **f**, Methylation levels of the CpG in the CTCF motif were measured in glioma specimens by methylation-sensitive restriction, and plotted as relative protection. **g**, Expanded views of the *FIP1L1* enhancer locus and *PDGFRA* locus shown with H3K27ac tracks. Vertical black bars indicate the locations of the common *PDGFRA* promoter primer and four complementary primers tested in 3C. **h–k**, Plots show normalized 3C interaction frequencies between *PDGFRA* promoter and indicated regions. A strong interaction between the *PDGFRA* promoter and the *FIP1L1* enhancer is evident in *IDH* mutant tumours and models. ND, none detected. Bars and error bars in all panels reflect mean and s.d. of triplicate observations, respectively.

were markedly different in *IDH* mutant tumours. Here, 3C revealed a strong interaction between the *PDGFRA* promoter and the *FIP1L1* enhancer, despite a separation of ~900 kb (Fig. 3i). For comparison, this interaction is approximately fivefold stronger than that between the *PDGFRA* promoter and its intragenic enhancer. To confirm this interaction, we designed and normalized reciprocal probe and primers to compare the relative strength with which the *FIP1L1* enhancer interacts with nearby promoters and *PDGFRA* (Extended Data Fig. 5). Notably, we found that the interaction between *FIP1L1* enhancer and *PDGFRA* promoter in *IDH* mutant tumours is stronger than that between *FIP1L1* enhancer and *FIP1L1* promoter. This suggests that disruption of a boundary element by *IDH* mutation and hypermethylation allows a potent constitutive enhancer to interact aberrantly with, and upregulate, *PDGFRA*.

To test this model functionally, we considered whether perturbing the boundary alters *PDGFRA* expression in patient-derived gliomaspheres (Fig. 4a). First, we focused on the *IDH1* mutant astrocytoma model, BT142. In this mutant line, the CpG dinucleotide in the CTCF motif exhibits higher methylation than wild-type models (~13% versus ~2% per bisulfite sequencing), and CTCF binding is roughly threefold lower. Consistently, 3C reveals a strong interaction between the *FIP1L1* enhancer and the *PDGFRA* promoter that is specific to the mutant line (Fig. 3i), and *PDGFRA* is highly expressed.

We reasoned that a demethylating agent should reduce methylation at this CpG dinucleotide, allowing CTCF to bind and restore *PDGFRA* insulation. We therefore treated BT142 gliomaspheres with the DNA methyltransferase inhibitor 5-azacytidine (5-aza). 5-aza treatment reduced methylation of the CTCF motif by ~2.5-fold, increased CTCF
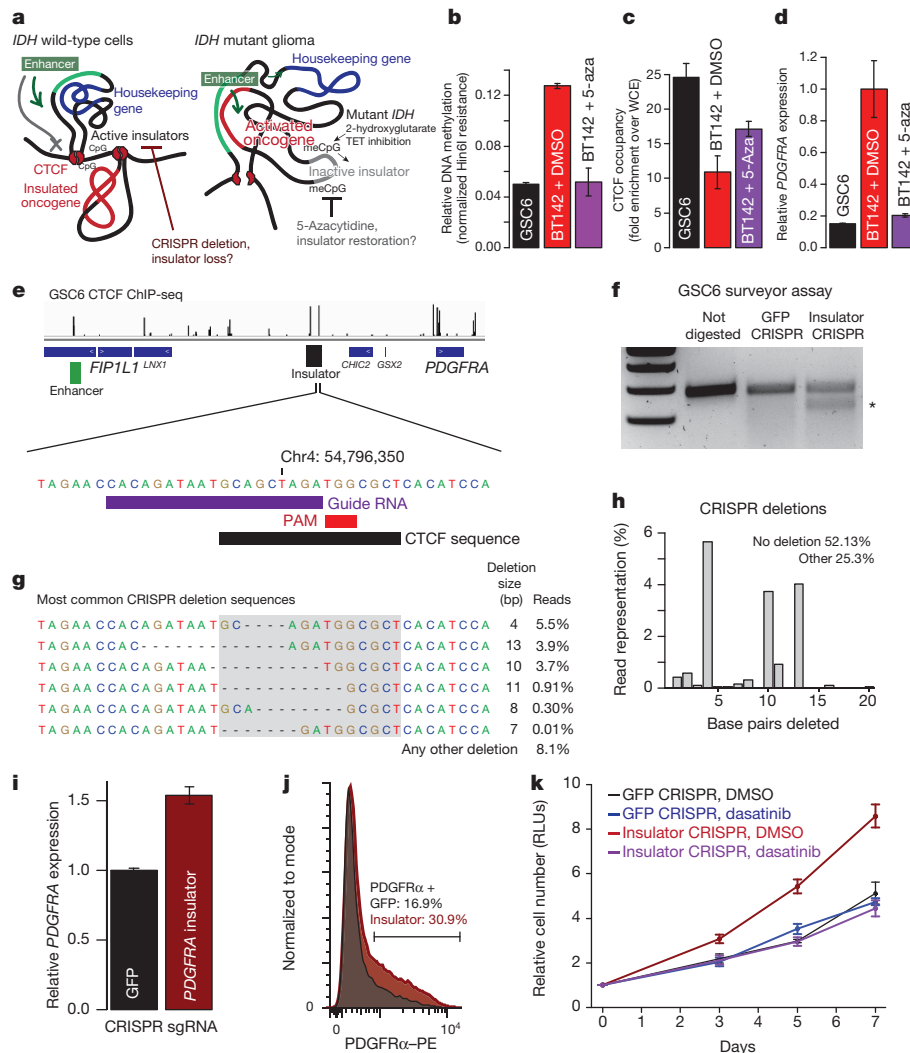
Figure 4 | Boundary methylation and CTCF occupancy affect *PDGFRA* expression and proliferation. a, Schematic depicts chromatin loops and boundaries in the *PDGFRA* locus. In *IDH* wild-type cells (left), intact boundary insulates oncogene. Disruption of the boundary by removing the CTCF motif should activate the oncogene. In *IDH* mutant cells (right), hypermethylation blocks CTCF, compromising the boundary and allowing enhancer to activate the oncogene. Demethylation should restore CTCF-mediated insulation. meCpG, methylated CpG. **b,** Plot compares CpG methylation in the CTCF motif in *IDH* wild-type gliomaspheres (black), *IDH1* mutant gliomaspheres (red), and *IDH1* mutant gliomaspheres treated with 5 μM 5-aza for 8 days (purple). **c,** Plot compares CTCF occupancy over the boundary. DMSO, dimethylsulfoxide; WCE, whole-cell extract. **d,** Plot compares *PDGFRA* expression. Demethylation restores *PDGFRA* insulation in *IDH1* mutant gliomaspheres. **e,** CTCF binding shown for the *FIP1L1/PDGFRA* region. Expanded view shows CTCF motif in the

insulator targeted for CRISPR-based deletion. sgRNA and PAM direct Cas9 nuclease to the motif. **f,** Surveyor assay detects target site alterations in GSC6 gliomaspheres infected with Cas9 and sgRNA (but not in control cells infected with GFP-targeting sgRNA). **g,** Sequencing of target site reveals the indicated deletions. CTCF motif disrupted on ~25% of alleles (compare to <0.01% in control). **h,** Plot depicts fraction of reads in insulator CRISPR cells with a deletion of indicated size. **i,** qPCR reveals increased *PDGFRA* expression in insulator CRISPR cells. **j,** Flow cytometry reveals ~2-fold greater PDGFRα in insulator CRISPR cells. PE, phycoerythrin. **k,** Plot depicts gliomasphere growth. Insulator CRISPR cells exhibit an approximately twofold increased proliferation, relative to control. This proliferation advantage is eliminated by PDGFRα inhibition. RLUs, relative light units. These results indicate that genetic or epigenetic disruption of the boundary compromises insulation of this oncogene. Bars and error bars in all panels reflect mean and s.d. of triplicate observations, respectively.

occupancy by ~1.7-fold and downregulated *PDGFRA* expression by ~5-fold (Fig. 4b–d). These results directly implicate DNA hypermethylation in compromising CTCF binding, boundary function and oncogene insulation in *IDH* mutant tumours.

Finally, we investigated whether genetic disruption of the CTCF motif could induce *PDGFRA* expression in wild-type gliomaspheres with an intact boundary (Fig. 4a). Here we focused on GSC6, a patient-derived glioblastoma model that contains an *EGFR* amplification, but is wild type for *IDH1*, *IDH2* and *PDGFRA*. We sought to disrupt the CTCF site in the boundary by CRISPR (clustered regularly interspaced short palindromic repeats)-based genome engineering[26,27] (Fig. 4e). We designed a short guide RNA (sgRNA) with a protospacer adjacent motif (PAM) within the CTCF motif. A single-vector lentiviral delivery system was used to infect GSC6 cells with a Cas9 expression

construct containing this insulator sgRNA or a control sgRNA (targeting green fluorescent protein, GFP). Surveyor assay confirmed target locus disruption in the insulator CRISPR condition (Fig. 4f). Direct sequencing of the target locus revealed that ~25% of alleles in the insulator CRISPR gliomaspheres contain a deletion within the CTCF motif expected to disrupt binding, compared to <0.1% in the GFP control (Fig. 4g, h).

We quantified *PDGFRA* expression in the genetically modified gliomaspheres. Reverse transcription PCR (RT–PCR) revealed an ~1.6-fold increase in *PDGFRA* messenger RNA in the insulator CRISPR cells, relative to control (Fig. 4i). Similarly, flow cytometry revealed an ~1.8-fold increase in the fraction of cells with PDGFRα surface expression (Fig. 4j). We conservatively estimate that CTCF motif disruption causes an ~3-fold increase in *PDGFRA* expression, given that

DNA level analysis indicates that less than 50% of insulator CRISPR cells were successfully edited.

Finally, we considered whether CRISPR-mediated boundary disruption and *PDGFRA* induction affects gliomasphere fitness. In support, the insulator CRISPR gliomaspheres have an approximately twofold growth advantage over the control GFP CRISPR gliomaspheres (Fig. 4k). This growth advantage is dependent on PDGFRα signalling, as it is abrogated by treatment with the PDGFR inhibitors dasatinib or crenolanib (Fig. 4k and Extended Data Fig. 6). Notably, *PDGFRA* expression in insulator CRISPR gliomaspheres increased further after extended culture (twofold increase compared with control), potentially owing to selection of effectively edited clones. The observation that genetic disruption of this CTCF boundary element induces *PDGFRA* expression and enhances proliferation provides strong support for our model that epigenetic disruption of this element offers similar growth advantage to *IDH* mutant gliomas.

In conclusion, we present a new epigenetic mechanism by which gain-of-function *IDH* mutations induce *PDGFRA* expression and thereby promote fitness in a subset of gliomas. We specifically find that, in addition to familiar effects on CpG islands, *IDH* mutations cause hypermethylation of CTCF binding sites genome-wide. This is associated with reduced CTCF binding and a global deregulation of boundary elements that partition topological domains. Disruption of a specific boundary bordering *PDGFRA* allows a potent enhancer to contact and activate this canonical glioma oncogene aberrantly.

Although disruption of this single boundary confers a growth advantage, it is unlikely to be the only mediator of *IDH* mutations in gliomas. The widespread disruption of CTCF binding and boundary element function could provide many opportunities for oncogene deregulation, and subsequent selection of proliferative progeny that inherit the altered epigenetic state. Insulator dysfunction may also be accompanied by promoter silencing events[28,29], and by alterations to other pathways affected by 2-hydroxyglutarate[7,30]. Conversely, disruption of chromosomal topology and oncogene insulation may be more generally relevant to methylator phenotypes observed in colorectal and renal cell carcinomas, leukaemia and other malignancies[28].

1. Parsons, D. W. *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science* **321,** 1807–1812 (2008).
2. The Cancer Genome Atlas Research Network. Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med.* **372,** 2481–2498 (2015).
3. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462,** 739–744 (2009).
4. Figueroa, M. E. *et al.* Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18,** 553–567 (2010).
5. Xu, W. *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19,** 17–30 (2011).
6. Lu, C. *et al.* IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature* **483,** 474–478 (2012).
7. Cairns, R. A. & Mak, T. W. Oncogenic isocitrate dehydrogenase mutations: mechanisms, models, and clinical opportunities. *Cancer Dis.* **3,** 730–741 (2013).
8. Pastor, W. A., Aravind, L. & Rao, A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nature Rev. Mol. Cell Biol.* **14,** 341–356 (2013).
9. Kohli, R. M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502,** 472–479 (2013).
10. Noushmehr, H. *et al.* Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* **17,** 510–522 (2010).
11. Turcan, S. *et al.* IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483,** 479–483 (2012).
12. Bickmore, W. A. & van Steensel, B. Genome architecture: domain organization of interphase chromosomes. *Cell* **152,** 1270–1284 (2013).
13. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289–293 (2009).
14. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485,** 376–380 (2012).
15. Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159,** 1665–1680 (2014).
16. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485,** 381–385 (2012).
17. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161,** 1012–1025 (2015).
18. Bell, A. C. & Felsenfeld, G. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405,** 482–485 (2000).
19. Hark, A. T. *et al.* CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* **405,** 486–489 (2000).
20. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348,** 648–660 (2015).
21. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl Acad. Sci. USA* **111,** 996–1001 (2014).
22. Sturm, D. *et al.* Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. *Nature Rev. Cancer* **14,** 92–107 (2014).
23. Brennan, C. W. *et al.* The somatic genomic landscape of glioblastoma. *Cell* **155,** 462–477 (2013).
24. Verhaak, R. G. *et al.* Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in *PDGFRA, IDH1, EGFR,* and *NF1. Cancer Cell* **17,** 98–110 (2010).
25. Wang, H. *et al.* Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* **22,** 1680–1688 (2012).
26. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157,** 1262–1278 (2014).
27. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnol.* **32,** 347–355 (2014).
28. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome – biological and translational implications. *Nature Rev. Cancer* **11,** 726–734 (2011).
29. Costello, J. F., Berger, M. S., Huang, H. S. & Cavenee, W. K. Silencing of *p16/CDKN2* expression in human gliomas by methylation and chromatin condensation. *Cancer Res.* **56,** 2405–2410 (1996).
30. Koivunen, P. *et al.* Transformation by the (*R*)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. *Nature* **483,** 484–488 (2012).

Author Contributions Conception and experimental design: W.A.F., Y.D., B.B.L., S.M.G., M.L.S. and B.E.B. Methodology and data acquisition: W.A.F., Y.D., B.B.L., S.M.G., A.S.V., A.O.S.-R., M.L.S. and B.E.B. Analysis and interpretation of data: W.A.F., Y.D. and B.E.B. Manuscript writing: W.A.F., Y.D. and B.E.B. W.A.F. and Y.D. contributed equally to this work.

Author Information Data generated for this study are available through the Gene Expression Omnibus (GEO) under accession number GSE70991. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.E.B. (bernstein.bradley@mgh.harvard.edu).

METHODS

No statistical methods were used to predetermine sample size.

Primary glioma specimens and gliomasphere models. Clinical samples GBM1w, GBM2w, GBM3w, GBM4w, GBM5w, GBM6w, GBM7w, AA15 m, AA16 m, AA17 m, OD18 m and AA19 m were obtained as frozen specimens from the Massachusetts General Hospital Pathology Tissue Bank, or received directly after surgical resection and flash frozen (Extended Data Table 1). All samples were acquired with Institutional Review Board approval, and were de-identified before receipt. GBM1w was obtained at autopsy; the remaining samples were surgical resections. *IDH* status was determined for all clinical samples by SNaPshot multiplex PCR[31]. *PDGFRA* status was confirmed by FISH analysis. Tissue (200–500 μg) was mechanically minced with a sterile razor blade before further processing.

Gliomaspheres were maintained in culture as described[32,33]. In brief, neurosphere cultures contain Neurobasal media supplemented with 20 ng ml^{-1} recombinant EGF (R and D Systems), 20 ng ml^{-1} FGF2 (R and D Systems), 1× B27 supplement (Invitrogen), 0.5× N2 supplement (Invitrogen), 3 mM L-glutamine, and penicillin/streptomycin. Cultures were confirmed to be mycoplasma-free via PCR methods. GSC4 and GSC6 gliomasphere lines were derived from *IDH* wild-type tumours resected at Massachusetts General Hospital, and have been previously described and characterized[32–34]. BT142 gliomasphere line (*IDH1* mutant)[35] was obtained from ATCC, and cultured as described above except 25% conditioned media was carried over each passage. BT142 G-CIMP status was confirmed by evaluating LINE methylation with the Global DNA Methylation Assay – LINE-1 kit (Active Motif), as described[36], and by methylation-sensitive restriction digests. GSC119 was derived from an *IDH1* mutant tumour (confirmed by SNaPshot) resected at Massachusetts General Hospital. We confirmed *IDH1* mutant status of GSC119 by RNA-seq (82 out of 148 reads overlapping the relevant position in the transcript correspond the mutant allele). The gliomasphere models were derived from tumours of the following types: GSC4 and GSC6: primary glioblastoma; BT142: grade III oligo-astrocytoma; GSC119: secondary glioblastoma, G-CIMP. Clinical specimens and models used in this study are detailed in Extended Data Table 1.

ChIP. ChIP-seq was performed as described previously[32]. In brief, cultured cells or minced tissue was fixed in 1% formaldehyde and snap frozen in liquid nitrogen and stored at −80 °C at least overnight. Sonication of tumour specimens and gliomaspheres was calibrated such that DNA was sheared to between 400 and 2,000 bp. CTCF was immunoprecipitated with a monoclonal rabbit CTCF antibody, clone D31H2 (Cell Signaling 3418). H3K27ac was immunoprecipitated with an antibody from Active Motif (39133). ChIP DNA was used to generate sequencing libraries by end repair (End-It DNA repair kit, Epicentre), 3′ A base overhang addition via Klenow fragment (NEB), and ligation of barcoded sequencing adapters. Barcoded fragments were amplified via PCR. Libraries were sequenced as 38-base paired-end reads on an Illumina NextSeq500 instrument or as 50-base single-end reads on a MiSeq instrument. Sequencing libraries are detailed in Extended Data Table 2. H3K27ac maps for GSC6 were previously deposited to the GEO under accession GSM1306340. Genomic data has been deposited into GEO as GSE70991.

For sequence analysis, identical reads were collapsed to a single paired-end read to avoid PCR duplicates. To avoid possible saturation, reads were downsampled to 5% reads collapsed as PCR duplicates, or 5 million fragments. Reads were aligned to hg19 using BWA, and peaks were called using HOMER. ChIP-seq tracks were visualized using Integrative Genomics Viewer (IGV, http://www.broadinstitute.org/igv/). To detect peaks lost in *IDH* mutants, we called signal over all peaks in a 100-bp window centred on the peaks. To control for copy number changes, we first called copy number profiles from input sequencing data using CNVnator[37]. We then removed all regions where at least one sample had a strong deletion (<0.25), and normalized by copy number. To account for batch effects and difference in ChIP efficiency, we quantile normalized each data set. Peaks were scored as lost or gained if the difference in signal between a given tumour and the average of the five wild-type tumours was at least twofold lower or higher, with a signal of at least 1 in all wild-type or *IDH* mutant tumours. Fisher exact test confirmed that the overlap between peaks lost in the *IDH* mutant tumours is highly significant ($P < 10^{-100}$).

GC content over CTCF peaks lost (or retained) in the *IDH* mutant glioma specimens was averaged over 200-bp windows centred on each peak lost in *IDH* mutant tumours. Methylation levels were quantified over these same regions for 3 *IDH* mutant and 3 *IDH* wild-type tumours, using TCGA data generated by whole genome bisulfite sequencing[10]. In brief, methylation levels (percentage) based on proportion of reads with protected CpG were averaged over all CpG di-nucleotides in these regions, treating each tumour separately.

Occupancy of the CTCF site in the boundary element adjacent to the *PDGFRA* locus was quantified by ChIP qPCR, using the following primers: PDGFRActcfF: 5′-GTCACAGTAGAACCACAGAT-3′; PDGFRActcfR: 5′-TAAGTATACTGGTCCTCCTC-3′. Equal masses of ChIP or input (WCE) DNA were used as input for PCR, and CTCF occupancy was quantified as a ratio between ChIP and WCE, determined by $2^{-\Delta Ct}$. CTCF peak intensity was further normalized as ratio to two invariant peaks, at *PSMB1* and *SPG11*, using the following primers: PSMB1ctcfF: 5′-CCTTCCTAGTCACTCAGTAA-3′; PSMB1ctcfR: 5′-CAGTGTTGACTCATCCAG-3′; SPG11ctcfF: 5′-CAGTACCAGCCTCTCTAG-3′; SPG11ctcfR: 5′-CTAAGCTAGGCCTTCAAG-3′.

Cross-boundary and intra-domain gene pair correlation analysis. RNA-seq data for 357 normal brain samples was downloaded from GTEx[20]. RNA-seq data and copy number profiles for lower grade gliomas were downloaded from TCGA[23,24]. Contact domains of IMR90, GM12878, K562 and NHEK cells were obtained from published HiC data[15]. Genes were assigned to the inner-most domain in which their transcription start site fell within. Gene pairs were considered to be in the same domain if they were assigned to the same domain in both GM12878 and IMR90. Gene pairs were considered to span a boundary if they were assigned to different domains in both GM12878 and IMR90, and separated by a CTCF-binding site in *IDH* wild-type tumours. Gene pairs that did not fit either criterion were excluded from this analysis. The plot of correlation vs distance for brain GTEx samples is based on Pearson correlations for all relevant pairs, smoothed by locally weighted scatterplot smoothing with weighted linear least squares (LOESS). To assess the bias in correlation differences, we computed the difference of Pearson correlations between wild-type and *IDH* mutant gliomas for all gene pairs separated by <180kb. In Fig. 1e, this difference in correlations is plotted against the significance of this difference (estimated by Fisher *z*-transformation). For each gene pair, we omitted samples with a deletion or amplification of one of the genes at or above threshold of the minimal arm level deletion or amplification (to avoid copy number bias). To ensure robustness, we also repeated the analysis using boundaries defined from HiC data for K562 and NHEK. This yielded similar results: 84% pairs gaining correlation cross boundary versus 71% expected ($P < 8 \times 10^{-3}$), 54% pairs losing correlation are within the same domain versus 29% expected ($P < 3 \times 10^{-8}$). Repeating the analysis with only the 14,055 genes that have expressed over 1 transcripts per million (TPM) in at least half the samples also yielded similar results (Extended Data Fig. 7): 92% pairs gaining correlation cross boundary versus 69% expected ($P < 2 \times 10^{-3}$), 73% pairs losing correlation are within the same domain versus 31% expected ($P < 8 \times 10^{-4}$).

Genomic scan for deregulated boundaries. To detect boundaries deregulated in *IDH* mutant gliomas, we scanned for gene pairs, separated by <1 Mb, with a significant difference in correlation between wild-type and *IDH* mutant tumours (Fisher *z*-transformation, FDR <1%). We omitted amplified or deleted samples as described above. To ensure robustness to noise from lowly expressed genes, we first filtered out 6,476 genes expressed <1 TPM in more than half of the samples (keeping 14,055 genes). We considered all domains and boundaries scored in IMR90 HiC data[13]. Gene pairs crossing a CTCF peak and an IMR90 boundary (that is, can be assigned to different domains) that were significantly more correlated in *IDH* mutant tumours were considered to support the loss of that boundary. Gene pairs not crossing a boundary (that is, can be assigned to the same domain) that were significantly less correlated in *IDH* mutant tumours were considered to support the loss of a flanking boundary. We collated a set of deregulated boundaries, supported by at least one cross-boundary pair gaining correlation and at least one intra-domain pair losing correlation. Each was assigned a *P* value equal to the product of both supporting pairs (best *P* value was chosen if there were more supporting pairs). If both boundaries of a domain were deregulated, or if the same pair of gene pairs (one losing and one gaining correlations) were supporting more than one boundary due to overlapping domains, the entries were merged (Supplementary Table 1). This definition allows every gene pair to be considered as potential support for a boundary loss. To quantify CTCF occupancy over these deregulated boundaries, we averaged the signal over all CTCF peaks located within a 1-kb window around the boundary, using copy number and quantile normalized CTCF signals. To quantify DNA methylation over the deregulated boundaries, we averaged DNA methylation signals from TCGA data in 200-bp windows as above. Figure 2a depicts significance of disrupted domains and the fold change of genes in them that are upregulated in *IDH* mutant tumours (compared to median expression in wild type). In addition to *PDGFRA*, top-ranking genes include *CHD4* ($P < 10^{-32}$), a driver of glioblastoma tumour initiation[38], *L1CAM* ($P < 10^{-8}$), a regulator of the glioma stem cells and tumour growth[39], and other candidate regulators (Supplementary Table 1).

To ensure robustness to cell-type-specific boundaries, we repeated the analysis with GM12878-, K562- and NHEK-defined boundaries. This yielded very similar results, and again highlighted *PDGFRA* as an overexpressed gene adjacent to a disrupted boundary.

TCGA correlation and outcome analysis. For the correlation of *FIP1L1* and *PDGFRA* expression, RNA-seq data from the TCGA lower grade glioma (LGG) and glioblastoma (GBM) data sets[2,24] were downloaded and segregated by *IDH* mutation status and subtype. Patients from the proneural subtype were divided by *IDH* mutation status, while patients from the mesenchymal, classical or neural subtypes (which had no *IDH* mutations) were classified as 'other'. For correlation

analysis, patients with copy number variation in either gene were excluded from the analysis to control for effects of co-amplification. For outcome analysis, LGG RNA-seq data and corresponding patient survival data was obtained from TCGA. Patients with sum *PDGFRA* and *FIP1L1* expression of at least one-half of one standard deviation above the mean were classified as 'high *PDGFRA* and *FIP1L1* expression' ($n = 17$), while all other patients were classified as 'low *PDGFRA* and *FIP1L1* expression' ($n = 201$). Data were plotted as Kaplan–Meier curves and statistically analysed via log–rank test.

HiC data analysis and visualization. HiC data[15] were downloaded from GEO. 5-kb resolution intra-chromosomal contact scores for chromosome 4 for the cell lines IMR90, NHEK, KBM7, K562, HUVEC, HMEC and GM12878 were filtered to the region between 53,700 and 55,400 kb. The average interaction score at each coordinate pair for all cell lines was calculated and used to determine putative insulator elements as local maxima at the interaction point of two domain boundaries. To determine the interactions of the *PDGFRA* promoter, the interaction scores of all points in the region with the *PDGFRA* promoter (chr4: 55,090,000) were plotted as a one-dimensional trace. To view the topological domain structure of the region, HiC interaction scores were visualized using Juicebox (http://www.aidenlab.org/juicebox/)[15]. Data shown is from the IMR90 cell line at 5-kb resolution, normalized to coverage.

DNA methylation quantification. DNA methylation was analysed in two ways. For gliomaspheres, genomic DNA was isolated via QiaAmp DNA minikit (Qiagen) and subjected to bisulfite conversion (EZ DNA Methylation Gold Kit, Zymo Research). Bisulfite-converted DNA specific to the CTCF-binding site (defined by JASPAR[40]) in the boundary adjacent to *PDGFRA* was amplified using the following primers forward: 5′-GAATTATAGATAATGTAGTTAGATGG-3′, reverse: 5′-AAATATACTA ATCCTCCTCTCCCAAA-3′. Amplified DNA was used to prepare a sequencing library, which was sequenced as 38-base paired-end reads on a NextSeq500. For tumours, limiting DNA yields required an alternative strategy for methylation analysis. Tumour genomic DNA was isolated from minced frozen sections of tumours by QiaAmp DNA minikit (Qiagen). Genomic DNA was digested using the methylation-sensitive restriction enzyme Hin6I (Thermo) recognizing the restriction site GCGC, or subjected to mock digestion. Protected DNA was quantified by PCR using the following primer set: PDGFRAinsF: 5′-CGTGAGCTGAATTGTGCCTG-3′, PDGFRAinsR: 5′-TGGGAGGACAGTTTAGGGCT-3′, normalizing to mock digestion.

3C analysis. 3C analysis was performed using procedures as described previously[41,42]. In brief, ~10 million cell equivalents from minced tumour specimens or gliomasphere cultures were fixed in 1% formaldehyde. Fixed samples were lysed in lysis buffer containing 0.2% PMSF using a Dounce pestle. Following lysis, samples were digested with HinDIII (NEB) overnight on a thermomixer at 37 °C rotating at 950 r.p.m. Diluted samples were ligated using T4 DNA ligase (NEB) at 16 °C overnight, followed by RNase and proteinase K treatment. DNA was extracted via phenol/chloroform/isoamyl alcohol (Invitrogen). DNA was analysed via TaqMan PCR using ABI master mix. Primers and probe were synthesized by IDT with the following sequences: common *PDGFRA* promoter: 5′-GGTCGTGCCTTTGTTTT-3′; *FIP1L1* control: 5′-CAGGGAAGAGAGGAAGTTT-3′; *FIP1L1* enhancer: 5′-TTAAGTAAGCAGGTAAACTACAT-3′; intragenic enhancer: 5′-AGCC TTTGCCTCCTTTT-3′; intragenic control: 5′-CCACAGGGAGAAGGAAAT-3′; intact promoter: 5′-CAAGGAATTCGTAGGGTTC-3′; probe: 5′-/56-FAM/ TTGTATGCG/ZEN/AGATAGAAGCCAGGGCAA/3IABkFQ/-3′. For the reciprocal *FIP1L1* enhancer interaction interrogation, the following primer sequences were used: common enhancer primer (as *FIP1L1* enhancer primer above): 5′-TTAAGTAAGCAGGTAAACTACAT-3′; *PDGFRA* promoter (as common *PDGFRA* promoter above): 5′-GGTCGTGCCTTTGTTTT-3′; *SCFD2* promoter: 5′-AATACATGGTCATGATGCTC-3′; *FIP1L1* promoter: 5′-AGGCATTGCTTAAACATAAC-3′; *FIP1L1* control: 5′-TTATTTGTAGT AGAGGTTACTGG-3′; *PDGFRA* control: 5′-ATGATAACACCACCATTCAG-3′; *FIP1L1* enhancer probe: 5′-/56-FAM/TATCCCAAC/ZEN/CAAATACAGGG CTTGG/3IABkFQ/-3′. To normalize primer signals, bacterial artificial chromosome (BAC) clones CTD-2022B5 and RP11-626H4 were obtained from Invitrogen. BAC DNA was purified via BACMAX DNA Purification kit (Epicentre) and quantified using two primer sets specific to the Chloramphenicol resistance gene: 1F: 5′-TTCGTCTCAGCCAATCCCTG-3′; 1R: 5′-TTTGCCCATG GTGAAAACGG-3′; 2F: GGTTCATCATGCCGTTTGTG-3′; 2R: 5′-CCACTCAT CGCAGTACTGTTG-3′. BAC DNA was subjected to a similar 3C protocol, omitting steps related to cell lysis, proteinase or RNase treatment. PCR signal from tumour and gliomasphere 3C was normalized to digestion efficiency and BAC primer signal.

Treatment with demethylating agent. BT142 cells were cultured in either 5 μM 5-azacytidine or equivalent DMSO (1:10,000) for 8 days, with drug refreshed every 2 days.

CRISPR/Cas9 insulator disruption. The following CRISPR sgRNAs were cloned into the LentiCRISPR vector obtained from the Zhang laboratory[43]: GFP: 5′-GAGCTGGACGGCGACGTAAA-3′; insulator: 5′-GCCACA GATAATGCAGCTAGA-3′. GSC6 gliomaspheres were mechanically dissociated and plated in 5 μg ml⁻¹ EHS laminin (Sigma) and allowed to adhere overnight, and then infected with lentivirus containing either CRISPR vector for 48 h. Cells were then selected in 1 μg ml⁻¹ puromycin for 4 days, with puromycin-containing media refreshed every 2 days. Genomic DNA was isolated and the region of interest was amplified using the PDGFRAins primer set described above. CRISPR-mediated disruption of this amplified DNA was confirmed via Surveyor Assay (Transgenomic), with amplified uninfected GSC6 genomic DNA being added to each annealing reaction as the unmodified control. To quantify the precise CRISPR alterations, genomic DNA from each construct was amplified using a set of primers closer to the putative deletion site as follows: forward: 5′-TTTGCAATGGGACACGGAGA-3′, reverse: 5′-AGAAATGTGTGGATGTGAGCG-3′. PCR product from these primers was used to prepare a library that was sequenced as 38-base paired-end reads on the Illumina NextSeq500.

***PDGFRA* qPCR.** Total RNA was isolated from CRISPR-infected GSC6 gliomaspheres (insulator or control GFP sgRNA) or BT142 gliomaspheres (5-aza-treated or control condition) using the RNeasy minikit (Qiagen) and used to synthesize cDNA with the SuperScriptIII system (Invitrogen). cDNA was analysed using SYBR mastermix (Applied Biosystems) on a 7500 Fast Real Time System (Applied Biosystems). *PDGFRA* expression was determined using the following primers: forward: 5′-GCTCAGCCCTGTGAGAAGAC-3′, reverse: 5′-ATTGCGGAATAACATCGGAG-3′, and was normalized to primers for ribosomal protein, large, P0 (*RPLP0*), as follows: forward: 5′-TCCCACTTGCTGAAAAGGTCA-3′, reverse: 5′-CCGACTCTTCCTTG GCTTCA-3′. Normalization was also verified by β-actin (*ACTB*), forward: 5′-AGAAAATCTGGCACCACACC-3′, reverse: 5′-AGAGGCGTACAGG GATAGCA-3′.

PDGFRα flow cytometry. Cells were incubated with PE-conjugated anti-PDGFRα (CD140a) antibody (Biolegend, clone 16A1) for 30 min at room temperature at the dilution specified in the manufacturer's protocol. Data was analysed and visualized with FlowJo software. Single live cells were selected for analysis via side and forward scatter, and viable cells were selected by lack of an unstained channel (APC) autofluorescence.

Cell growth assay. For the cell growth assay, 2,500 dissociated viable GSC6 cells expressing CRISPR and either GFP or insulator-targeting sgRNA (see above) was plated in 100 μl of media in an opaque-walled tissue culture 96-well plate, in 1 μM dasatinib, 500 nM crenolanib, or equivalent DMSO (1:10,000) as a vehicle control. Cell growth was analysed at days 3, 5 and 7 for dasatinib, or days 3, 7 and 10 for crenolanib, using CellTiter-Glo reagent (Promega) following the manufacturer's protocol. Data were normalized across days using an ATP standard curve.

31. Chi, A. S. *et al.* Prospective, high-throughput molecular profiling of human gliomas. *J. Neurooncol.* **110,** 89–98 (2012).
32. Rheinbay, E. *et al.* An aberrant transcription factor network essential for Wnt signaling and stem cell maintenance in glioblastoma. *Cell Rep.* **3,** 1567–1579 (2013).
33. Suvà, M. L. *et al.* Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell* **157,** 580–594 (2014).
34. Wakimoto, H. *et al.* Maintenance of primary tumor phenotype and genotype in glioblastoma stem cells. *Neuro Oncol.* **14,** 132–144 (2012).
35. Luchman, H. A. *et al.* An *in vivo* patient-derived model of endogenous *IDH1*-mutant glioma. *Neuro Oncol.* **14,** 184–191 (2012).
36. Lai, R. K. *et al.* Genome-wide methylation analyses in glioblastoma multiforme. *PLoS ONE* **9,** e89376 (2014).
37. Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21,** 974–984 (2011).
38. Chudnovsky, Y. *et al.* ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state. *Cell Rep.* **6,** 313–324 (2014).
39. Bao, S. *et al.* Targeting cancer stem cells through L1CAM suppresses glioma growth. *Cancer Res.* **68,** 6043–6048 (2008).
40. Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W. W. & Lenhard, B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32,** D91–D94 (2004).
41. de Laat, W. & Dekker, J. 3C-based technologies to study the shape of the genome. *Methods* **58,** 189–191 (2012).
42. Hagège, H. *et al.* Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature Protocols* **2,** 1722–1733 (2007).
43. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819–823 (2013).
44. Cahoy, J. D. *et al.* A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* **28,** 264–278 (2008).

a

b
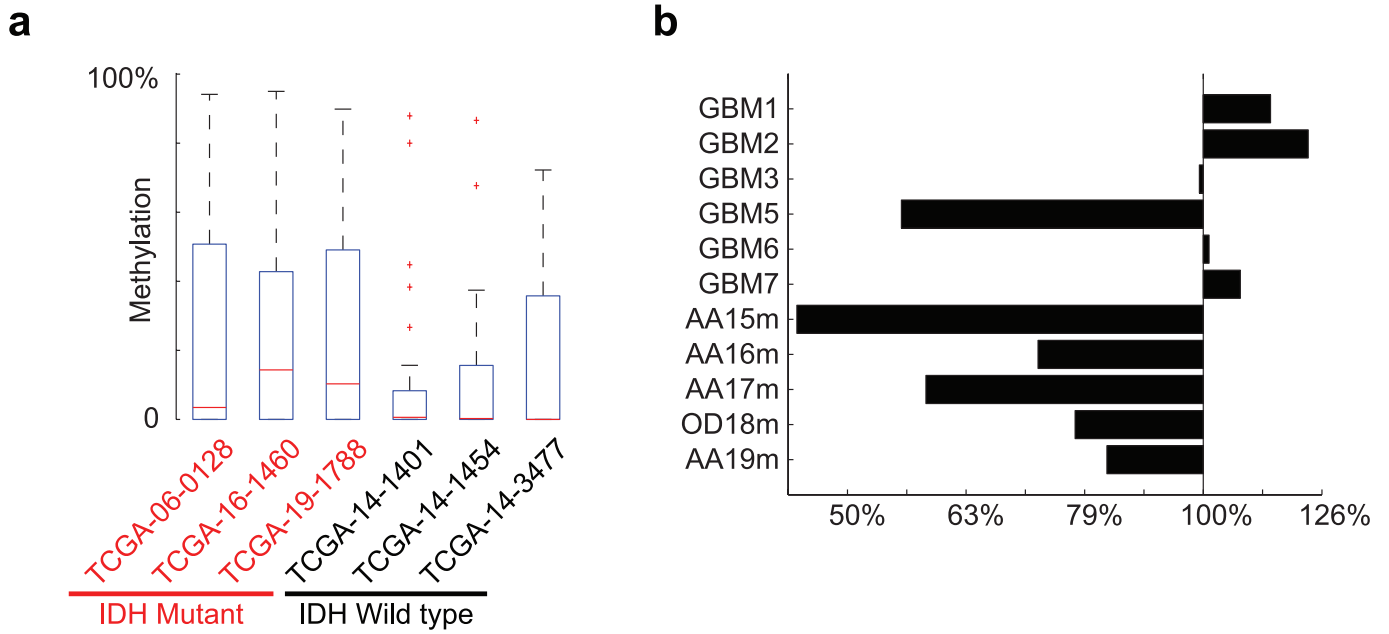
Extended Data Figure 1 | DNA methylation and CTCF binding at deregulated boundaries. a, Box plots show DNA methylation levels over CTCF sites (200-bp window centred on the peak) within boundaries predicted by gene pair correlation analysis to be disrupted. All CTCF sites located within a 1-kb window centred on a disrupted boundary were considered. Methylation levels were determined from whole-genome bisulfite data for three *IDH* mutant (red labels) and three *IDH* wild-type (black labels) tumours. **b**, Bars show average normalized ChIP-seq signal over all CTCF sites located inside a 1-kb window centred on a disrupted boundary.

a

Mouse brain cells (GSE9566)

b

TCGA, IDH1 Mutant LGG
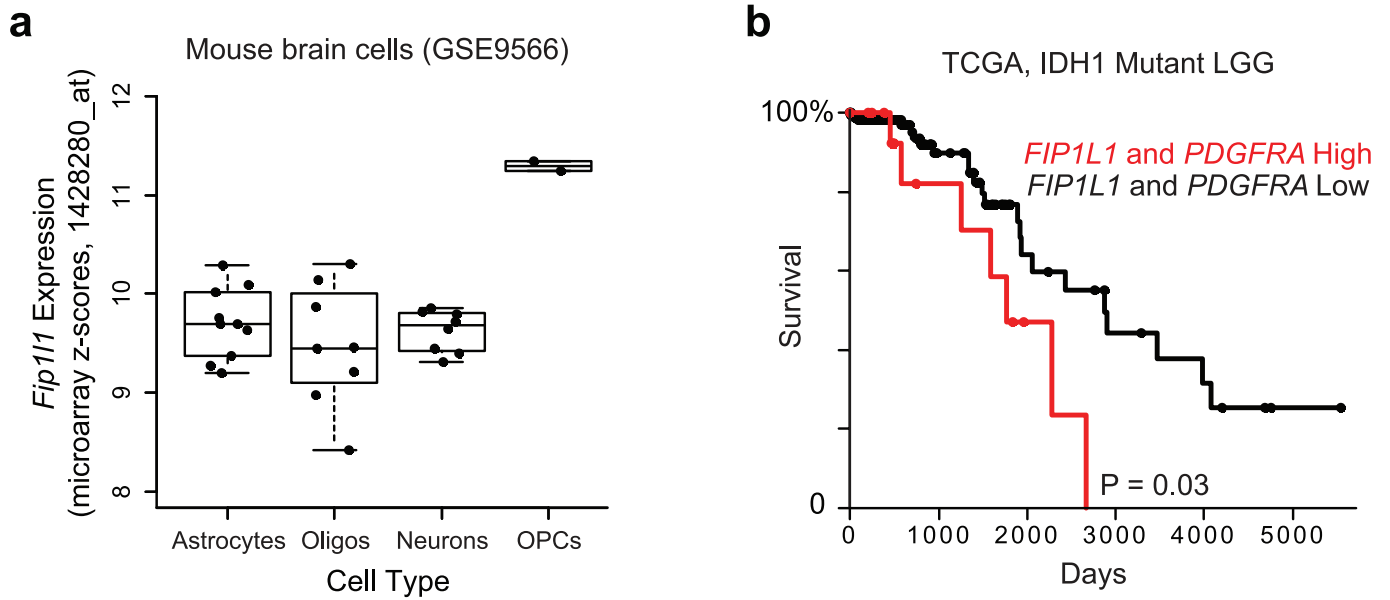
Extended Data Figure 2 | Expression of *Fip1l1* in mouse brain cells and survival effects of *PDGFRA* and *FIP1L1*. a, Expression of *Fip1l1* in a published data set for isolated mouse brain cell types[44]. **b**, Kaplan–Meier plot based on TCGA data[23] indicates that combined *FIP1L1* and *PDGFRA* expression is a negative prognostic factor in *IDH1* mutant lower-grade gliomas. Multivariate analysis including the known prognostic factor 1p/19q deletion diminished this effect into non-significance, suggesting that other predictors of survival may also have a role in this model.
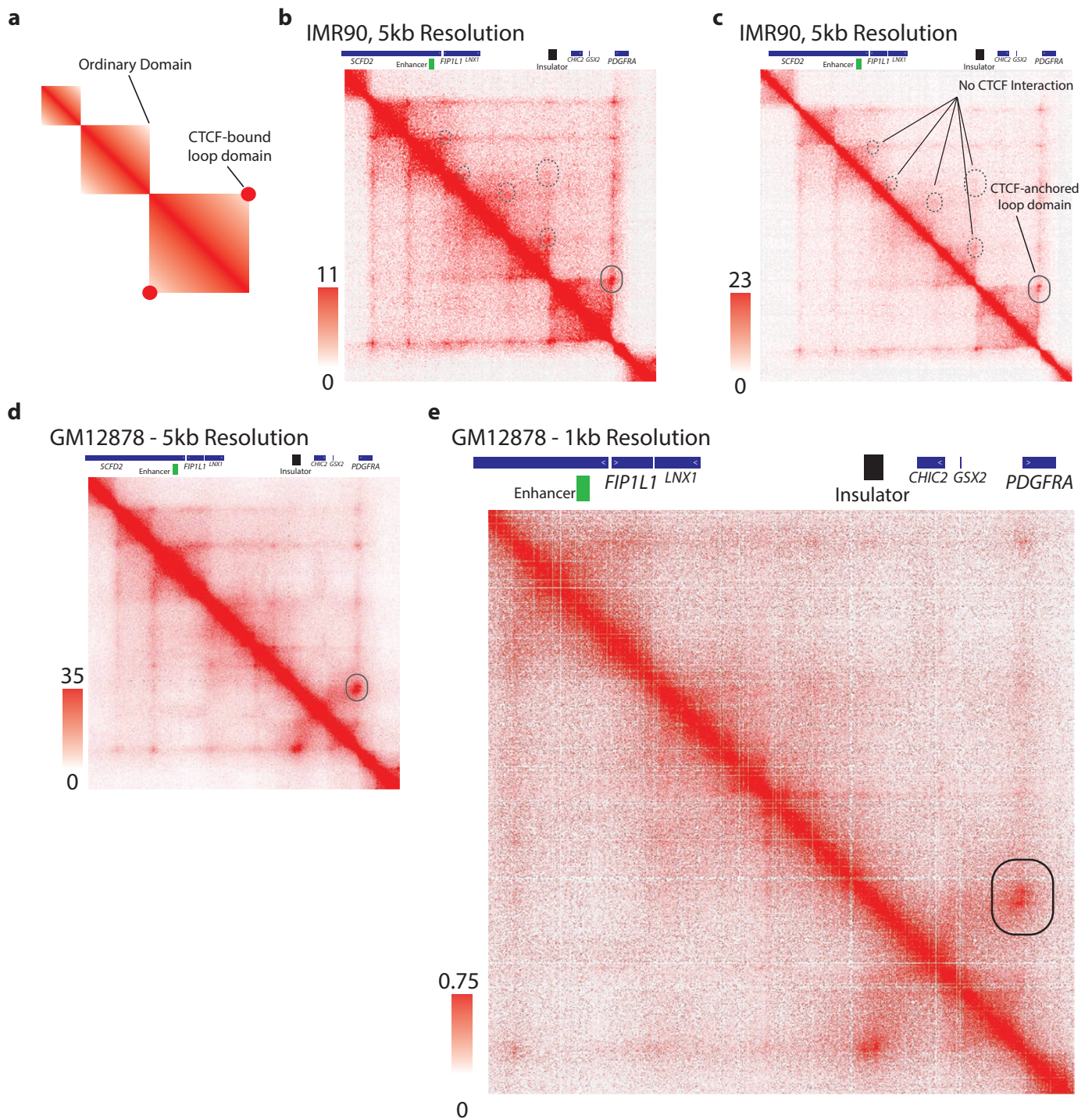
Extended Data Figure 3 | CTCF-anchored loop in the *PDGFRA* region. a, Schematic depiction of a HiC interaction signature of a CTCF-anchored loop domain, compared to an ordinary domain, as described previously[15]. CTCF-anchored loop domains are characterized by an increased interaction score at the apex of the domain, representing a CTCF–CTCF dimeric interaction. **b**, IMR90 HiC contact matrix for the *PDGFRA/FIP1L1* locus, as presented in Fig. 3a. Solid circle indicates CTCF dimer interaction point; dashed circles indicate lack of CTCF dimeric anchor signature. **c**, IMR90 HiC contact matrix as in **b**, but with an expanded heatmap scale, more clearly conveys the CTCF-anchored loop that insulates *PDGFRA*. **d**, **e**, HiC contact matrix for GM12878 cells for the same region confirms a single CTCF-anchored loop (solid circle) between *PDGFRA* and *FIP1L1*. These data support the significance of this specific boundary in locus topology and *PDGFRA* insulation.
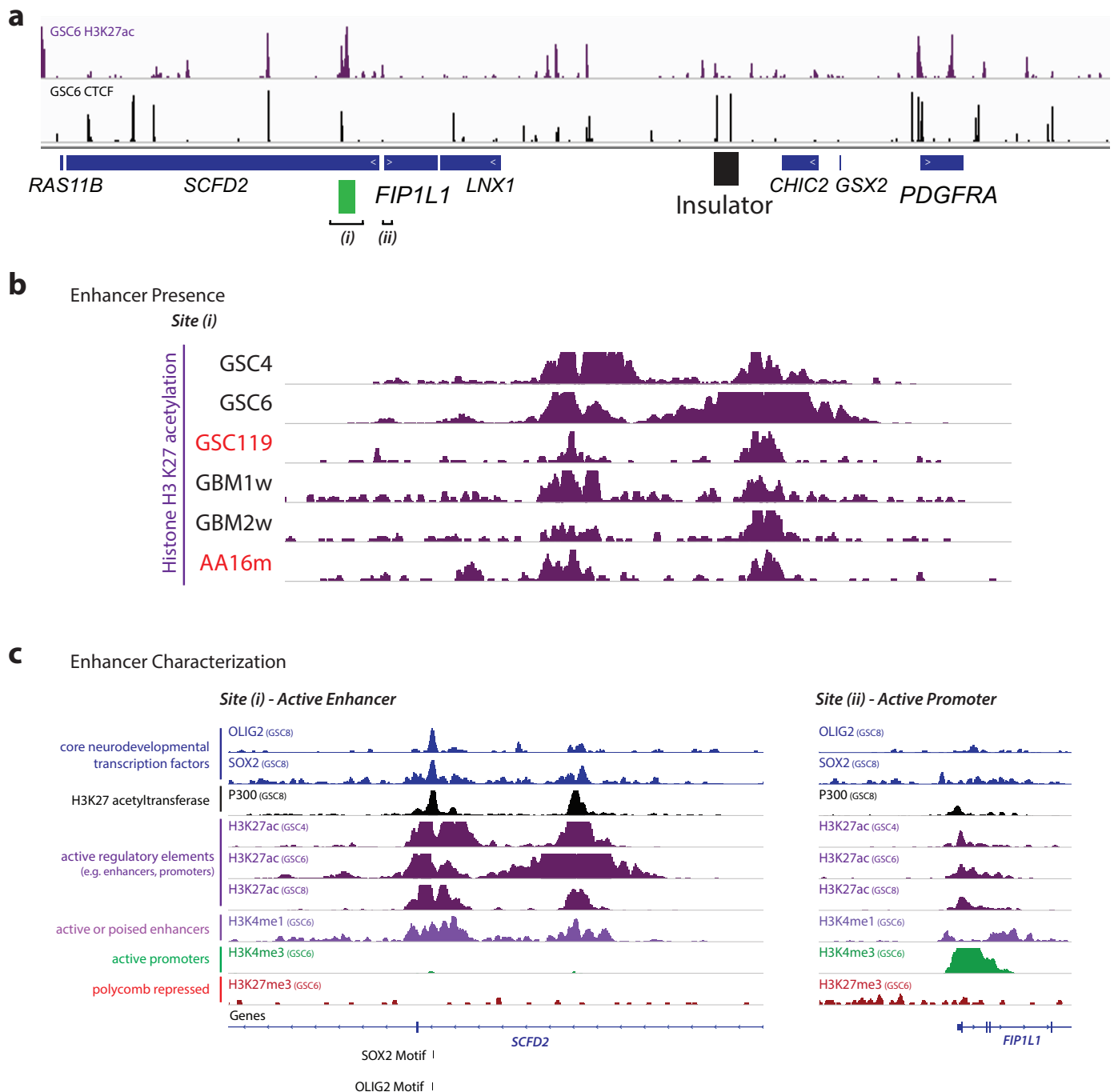
Extended Data Figure 4 | Characterization of the *FIP1L1* enhancer.
a, H3K27ac ChIP-seq track for GSC6 gliomaspheres reveals strong
enrichment over the *FIP1L1* enhancer. CTCF ChIP-seq track reveals
location of the boundary element insulator (as in Fig. 3a). *FIP1L1*
enhancer (i) and promoter (ii) are indicated. **b**, H3K27ac ChIP-seq tracks
for *IDH* mutant and wild-type gliomaspheres and glioma specimens reveal

enrichment over the *FIP1L1* enhancer. **c**, ChIP-seq tracks for glioma
master transcription factors and other histone modifications support the
enhancer identity of the element (H3K27ac, H3K4me1, SOX2, OLIG2;
lacks H3K4me3, lacks H3K27me3). By contrast, the *FIP1L1* promoter has
a distinct 'promoter-like' chromatin state.
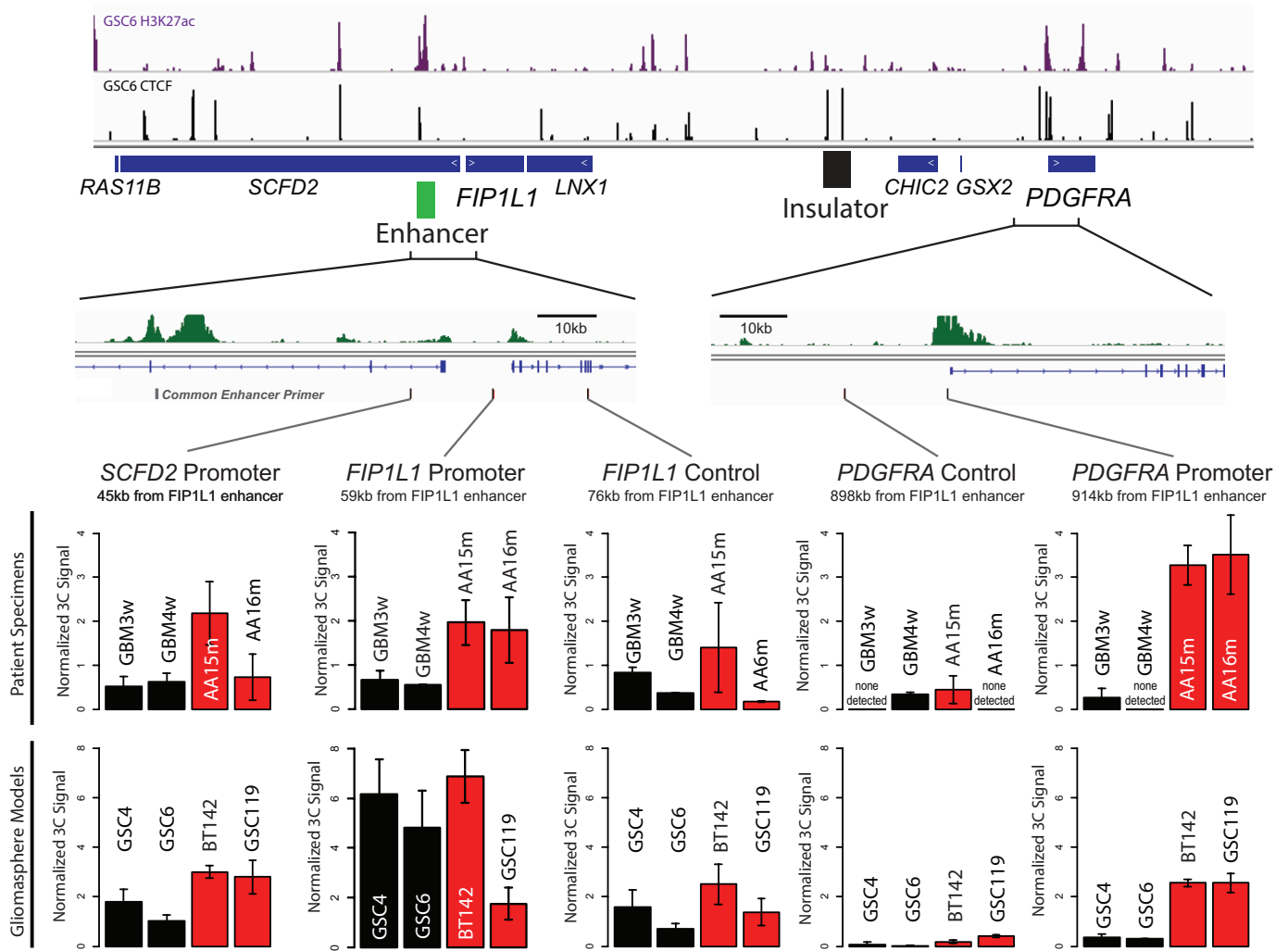
Extended Data Figure 5 | Interaction of the *FIP1L1* enhancer with nearby promoters and *PDGFRA* quantified by reciprocal 3C. Top, the H3K27ac, CTCF and genetic architecture of the *FIP1L1/PDGFRA* locus is indicated, highlighting the 3C strategy. Bottom, plots indicate the interaction signal of the indicated sites (black lines) with the common enhancer primer. The *FIP1L1* enhancer interacts with local promoters in wild-type and mutant tumours and models. In *IDH* wild-type gliomas, it shows essentially no interaction with the *PDGFRA* promoter. In *IDH* mutant gliomas, it interacts with the *PDGFRA* promoter with comparable strength to the local interactions, despite the much larger intervening distance (900 kb). Error bars reflect s.d.

GSC6 Growth in PDGFRA Inhibition

GFP, +DMSO
GFP, +500nM Crenolanib
Insulator, +DMSO
Insulator, +500nM Crenolanib
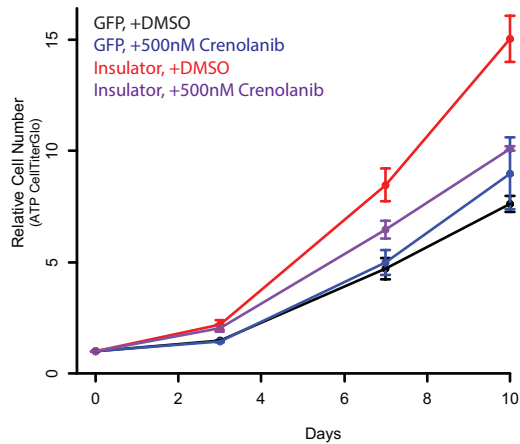
Extended Data Figure 6 | Crenolanib reverses the increased growth of *PDGFRA* insulator disrupted cells. Insulator CRISPR-infected gliomaspheres exhibit a roughly twofold increase in proliferation rate, compared to control sgRNA-infected gliomaspheres. This proliferative advantage is eliminated by treatment with the PDGFRα inhibitor crenolanib. Crenolanib and dasatinib both inhibit PDGFRα, but their other targets are non-overlapping. Hence, this sensitivity provides further support that *PDGFRA* induction drives the increased proliferation of the insulator CRISPR gliomaspheres. Error bars reflect s.d.
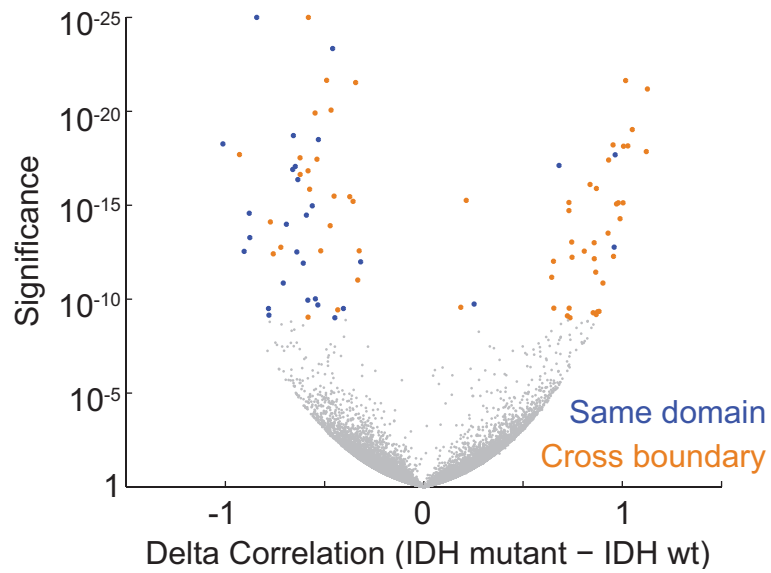
Extended Data Figure 7 | Signature of boundary deregulation in *IDH* mutant gliomas is robust. Volcano plot depicts the significance (*y* axis) of gene pairs that are either more or less correlated in *IDH* mutant than *IDH* wild-type gliomas. This plot was generated by repeating the analysis in the main text and shown in Fig. 1f, except that here the statistics were performed using only the 14,055 genes expressed at >1 TPM in at least half of the samples. This indicates that the boundary deregulation signature in *IDH* mutant gliomas is not sensitive to noise from lowly expressed genes.

Extended Data Table 1 | Clinical specimens and tumour models

Glioma	Tissue Type	Tissue Source	Source	IDH1 Status	PDGFRA Status	1p/19q Status	Grade	Disease
GBM1w	Autopsy Specimen	Banked	MGH	Wild Type	Amplified	Not tested	IV	Glioblastoma
GBM2w	Surgical Specimen	Banked	MGH	Wild Type	Wild Type	Not tested	IV	Glioblastoma
GBM3w	Surgical Specimen	Banked	MGH	Wild Type	Wild Type	Not tested	IV	Glioblastoma
GBM4w	Surgical Specimen	Banked	MGH	Wild Type	Wild Type	Not tested	IV	Glioblastoma
GBM5w	Surgical Specimen	Fresh	MGH	Wild Type	Wild Type	Not tested	IV	Glioblastoma
GBM6w	Surgical Specimen	Fresh	MGH	Wild Type	Wild Type	Not tested	IV	Glioblastoma
GBM7w	Surgical Specimen	Fresh	MGH	Wild Type	Wild Type	Not tested	IV	Glioblastoma
AA15m	Surgical Specimen	Banked	MGH	R132H	Wild Type	Intact	III	Anaplastic Astrocytoma
AA16m	Surgical Specimen	Banked	MGH	R132H	Wild Type	Intact	III	Anaplastic Astrocytoma
AA17m	Surgical Specimen	Fresh	MGH	R132H	Wild Type	Intact	III	Anaplastic Astrocytoma
OD18m	Surgical Specimen	Fresh	MGH	R132H	Wild Type	Lost	II	Oligodendroglioma
AA19m	Surgical Specimen	Fresh	MGH	R132H	Wild Type	Intact	III	Anaplastic Astrocytoma
GSC4	Gliomasphere	-	MGH	Wild Type	Wild Type	Intact	IV	Glioblastoma
GSC6	Gliomasphere	-	MGH	Wild Type	Wild Type	Intact	IV	Glioblastoma
BT142	Gliomasphere	-	ATCC	R132H	Wild Type	Intact	III	Anaplastic Oligoastrocytoma
GSC119	Gliomasphere	-	MGH	R132H	Wild Type	Intact	IV	Secondary Glioblastoma

Clinical information for glioma specimens and gliomasphere models is shown.

Extended Data Table 2 | Sequenced libraries characteristics

Sample Name	Experiment	Sequencing Depth	Sequencing Format	Sequencing Instrument	Total read number (millions)
GBM1w - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	19.3
GBM2w - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	17.6
GBM3w - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	20.2
GBM5w - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	30
GBM6w - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	35.1
GBM7w - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	36
AA15m - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	8.7
AA16m - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	23.7
AA17m - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	16.3
OD18m - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	9.2
AA19m - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	33
GSC4 - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	19.9
GSC6 - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	21.9
BT142 - CTCF	CTCF ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	16
GSC119 - CTCF	CTCF ChIP-seq	50 base pairs	Single end	Illumina Miseq	6.39
GBM1w - H3K27ac	H3K27ac ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	12.7
GBM2w - H3K27ac	H3K27ac ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	10.8
AA15m - H3K27ac	H3K27ac ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	11.8
GSC4 - H3K27ac	H3K27ac ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	9.7
GSC6 - H3K27ac	H3K27ac ChIP-seq	36 base pairs	Single end	Illumina Hiseq 2500	10.5
GSC119 - H3K27ac	H3K27ac ChIP-seq	38 base pairs	Paired end	Illumina NextSeq 500	9
GSC6 CRISPR - GFP sgRNA	Locus Sequencing	50 base pairs	Single end	Illumina Miseq	0.539
GSC6 CRISPR - insulator sgRNA	Locus Sequencing	50 base pairs	Single end	Illumina Miseq	0.639
GSC4 bisulfite	Bisulfite Sequencing	38 base pairs	Paired end	Illumina NextSeq 500	0.149
GSC6 bisulfite	Bisulfite Sequencing	38 base pairs	Paired end	Illumina NextSeq 500	0.149
BT142 bisulfite	Bisulfite Sequencing	38 base pairs	Paired end	Illumina NextSeq 500	0.149
GSC119 bisulfite	Bisulfite Sequencing	38 base pairs	Paired end	Illumina NextSeq 500	0.156

Pertinent statistics are listed for ChIP, genomic DNA and bisulfite-converted sequencing libraries.