SMARCB1 is required for widespread BAF complex–mediated activation of enhancers and bivalent promoters

Robert T Nakayama1,2,8, John L Pulice1,3,9, Alfredo M Valencia1,4,9, Matthew J McBride1,4, Zachary M McKenzie1, Mark A Gillespie5, Wai Lim Ku6, Mingxiang Teng7, Kairong Cui6, Robert T Williams1, Seth H Cassel1,8, He Qing1, Christian J Widmer1, George D Demetri2, Rafael A Irizarry7, Keji Zhao6, Jeffrey A Rankin5,8 & Cigall Kadoch1,3

Perturbations to mammalian SWI/SNF (mSWI/SNF or BAF) complexes contribute to more than 20% of human cancers, with driving roles first identified in malignant rhabdoid tumor, an aggressive pediatric cancer characterized by biallelic inactivation of the core BAF complex subunit SMARCB1 (BAF47). However, the mechanism by which this alteration contributes to tumorigenesis remains poorly understood. We find that BAF47 loss destabilizes BAF complexes on chromatin, absent significant changes in complex assembly or integrity. Rescue of BAF47 in BAF47-deficient sarcoma cell lines results in increased genome-wide BAF complex occupancy, facilitating widespread enhancer activation and opposition of Polycomb-mediated repression at bivalent promoters. We demonstrate differential regulation of mSWI/SNF complexes, BAF and PBAF complexes, enhancers and promoters, respectively, suggesting that each complex has distinct functions that are perturbed upon BAF47 loss. Our results demonstrate collaborative regulation of mSWI/SNF-mediated gene activation, identifying functions that are co-opted or abated to drive human cancers and developmental disorders.

Chromatin regulation is critical for the maintenance of timely and appropriate gene expression, with epigenetic regulators playing key parts in both normal development and oncogenesis1. Chromatin remodeling complexes regulate DNA accessibility via alteration of nucleosome positioning and/or occupancy in an ATP-dependent manner2. One of the most well-characterized chromatin remodeling complexes is the mSWI/SNF (BAF) complex, first identified in yeast3 and subsequently characterized in Drosophila4 and mammals5. Specialized BAF complex subunit configurations are critical in pluripotency6,7 and neural differentiation8, as well as the development of several other adult tissue types9. Various epigenetic modifiers and chromatin remodeling complexes, including BAF complexes, have been shown to localize to active promoters and enhancers in embryonic stem cells10,11; however, the roles for specific complexes in the establishment and maintenance of enhancer and promoter states are not well understood.

Evidence for a driving role of BAF complex alterations in cancer was first documented in malignant rhabdoid tumor (MRT), a highly aggressive pediatric cancer characterized by biallelic inactivation of the SMARCB1 gene, which encodes the core BAF complex subunit BAF47 (also known as INI1 and hSNF5), in addition to both in a locus-specific manner26,27 and in a global regulatory manner through upregulation of Ezh2 (ref. 28). In synovial sarcoma, the oncogenic SS18–SSX fusion has been demonstrated in additional cancer types, including atypical teratoid/rhabdoid tumor (AT/RT; ~100%)14 and epithelioid sarcomas (EpS; >90%)15. Additionally, BAF47 alterations are implicated in the development of meningiomas16, schwannomatosis17, and Coffin–Siris syndrome18. MRTs are genomically stable sarcomas with extremely low mutational burden19,20, and conditional biallelic inactivation of Smarcb1 in mice leads to the most rapid tumorigenesis documented for a single gene deletion, with median onset at 11 weeks21. Recent exome-sequencing studies have demonstrated that genes encoding BAF complex subunits are mutated in >20% of human cancers22, including gain-of-function perturbations such as the SS18–SSX oncogenic fusion, a hallmark of ~100% of synovial sarcomas23. The clear link between SMARCB1 deletion and MRT suggests MRT as a uniquely powerful disease setting in which to understand BAF complex alterations across human cancer.

Dynamic opposition between BAF complexes and Polycomb repressive complexes was first demonstrated genetically in Drosophila24, and has since been shown to govern critical processes in both normal development and disease25. In mammals, this opposition has been suggested to occur both in a locus-specific manner26,27 and in a global regulatory manner through upregulation of Ezh2 (ref. 28). In synovial sarcoma, the oncogenic SS18–SSX fusion has been demonstrated

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.
to direct BAF complexes to new genomic loci such as SOX2, opposing Polycomb-mediated repression and leading to oncogene activation
23. More recently, mechanistic studies have demonstrated that BAF–Polycomb complex opposition occurs on chromatin in a rapid, ATP-dependent manner, with loss of BAF47 leading to significant diminution in the ability of BAF complexes to oppose Polycomb-mediated repression
29. PRC2 complexes have particularly critical roles in maintaining bivalent gene promoters, marked dually by trimethylation of histone H3 Lys4 (H3K4me3) and H3K27me3 (refs. 30,31). The bivalent state of a given locus is maintained by a balance between activating trithorax (Trx) proteins (such as mixed-lineage leukemia proteins (MLLs)) and repressive Polycomb group (PcG) proteins (such as PRC2 components)
32, with several BAF complex subunits also categorized as Trx-group proteins
24. Loss of PRC2 leads to activation of tissue-specific bivalent promoters but not monovalent promoters marked by H3K27me3 only
33. Preclinical studies (and now early-stage clinical studies) using EZH2 inhibitors in BAF47-deficient sarcoma model systems have begun to show promise
34, suggesting that this dynamic opposition is a critical mediator of oncogenesis in BAF47-deficient sarcomas.

We sought to understand how loss of BAF47, a core BAF complex subunit, affects the stability, targeting, and gene expression regulation of BAF complexes in sarcomas. We determined that loss of BAF47 destabilizes the association of BAF complexes on chromatin without greatly impairing complex stability or assembly. Rescue of BAF47 in MRT and EpS cell lines drives a major gain of genome-wide BAF complex occupancy and enhancer state activation across the genome. In addition, we found that rescue of BAF47 targets BAF complexes to bivalent promoters, enabling opposition of Polycomb-mediated repression to resolve bivalent promoters to activation. The enhancer activation and resolution of bivalent promoters observed are collaborative with respect to gene expression, suggesting dual complementary roles in for BAF47-mediated tumor suppression. These data suggest two defining functions of BAF complexes that can be singly or collaboratively perturbed in BAF complex–altered cancers and developmental disorders.

RESULTS
BAF47 loss decreases chromatin affinity of BAF complexes
BAF47 is stable in its association with the BRG1 ATPase subunit in over 2 M urea treatment
23. To determine the effect of BAF47 loss on the integrity and subunit stability of BAF complexes, we infected G401 MRT cells with a lentiviral vector encoding full-length BAF47 or an empty vector control (Fig. 1a). Nuclear protein levels of core BAF subunits, both total and BAF-bound, were largely unchanged upon rescue of BAF47 (Fig. 1b and Supplementary Fig. 1a). To confirm this, we generated SMARCB1/∆∆ HEK293T cells using CRISPR-Cas9-mediated gene editing knockout and again did not observe changes in either total or BAF complex–bound protein levels (Fig. 1c). Silver stain analyses of anti-BRG1 and anti-BAF250A IPs from nuclear protein showed highly similar banding patterns in both conditions (Fig. 1d and Supplementary Fig. 1b). To complement this, we used low-stringency anti-BRG1 affinity purification–proteomic mass spectrometry (Fig. 1e and Supplementary Table 1). The abundance of peptides corresponding to most BAF complex subunits was roughly equivalent in both conditions, with the exception of the BAF45A, BAF45D and BAF60C subunits, which showed increases upon BAF47 rescue, possibly indicating their direct tethering to BAF47. Prior studies examining changes in complex subunit composition upon loss of the BAF47 subunit have presented conflicting results; some showed dissociation of BAF complexes
35,36 and others suggested no changes to BAF complex assembly
29,37–39, probably owing to differences in chromatin-bound protein purification methods and effects of super-stoichiometric protein abundance resulting from strong overexpression. Additionally, the harsh denaturing detergents, such as SDS and sodium deoxycholate, used in some of these prior experiments can disrupt protein–protein complex interactions and/or reduce the antibody pulldown efficacy in solution (Supplementary Fig. 1c,f).

To examine changes to the biochemical stability and size of BAF complexes, we performed 10–30% glycerol gradient–based density sedimentation analyses of nuclear extracts from G401 cells treated with empty vector or BAF47 lentiviral vector (below referred to as empty or BAF47 conditions, respectively) (Fig. 1f–g and Supplementary Fig. 1c). We found that BAF47 fully incorporated into BAF complexes upon re-expression (Fig. 1g, fractions 13–15), and subunits corresponding to both BAF and PBAF complexes shifted up by approximately 1–2 fractions, in accordance with the expected gain in complex mass resulting from BAF47 and associated subunits (Fig. 1c). Select subunits, including BRG1 and BAF60A, showed a greater spread across gradient fractions in the absence of BAF47; however, BRG1–bound subunit stability was largely retained, as demonstrated by urea denaturation experiments (Supplementary Fig. 1d). Collectively, these results demonstrate that loss of BAF47 does not affect global protein abundance or complex incorporation of most BAF complex subunits, nor does it render the complex wholly unstable in solution, as might have been predicted given its high degree of evolutionary conservation and highly penetrant loss-of-function phenotype.

As BAF chromatin remodeling complexes contain several DNA- and histone-binding domains, we sought to determine whether BAF47 loss alters the stability of BAF complexes on chromatin. We used NaCl-based differential salt extraction to determine the relative affinity of BAF complex proteins on chromatin in G401 cells in either empty or BAF47 conditions (Fig. 1h and Supplementary Fig. 2a). We found that BAF47–deficient complexes in G401 cells dissociate from chromatin at NaCl concentrations between 150 and 300 mM whereas subunits dissociate from chromatin at 500–1,000 mM NaCl treatment upon BAF47 rescue (Fig. 1i and Supplementary Fig. 2b,c). We assessed PRC2 subunits EZH2 and SUZ12 as controls and observed no changes in chromatin dissociation between conditions (Supplementary Fig. 2d). Results were similar in comparing BAF complex chromatin affinity in wild-type and SMARCB1/∆∆ HEK293T cells (Supplementary Fig. 2e–g). These results indicate that the primary biophysical consequence of BAF47 loss is decreased affinity of BAF complexes for chromatin, suggesting alterations in their genome-wide chromatin occupancy and regulatory capacity.

BAF47 rescue drives widespread BAF complex occupancy gain
To examine the effect of BAF47 rescue on BAF complex targeting and gene regulation, we infected two MRT cell lines, TCTC1240 and G401, with lentiviral vectors (empty control or encoding BAF47) (Fig. 2a) and performed chromatin IP of BAF complexes followed by sequencing (ChIP-seq) using antibodies to two core BAF complex subunits, BRG1 and BAF155. We observed a gain of genome-wide BAF complex occupancy upon rescue of BAF47 in TCTC1240 cells (Fig. 2b–d and Supplementary Fig. 3a–e). We found that gained (BAF47–only) BAF complex sites (defined as shared BRG1–BAF155 sites) in TCTC1240 cells were disproportionately localized to promoter-distal regions, as compared to conserved BAF complex sites empty–BAF47 shared (Fig. 2e and Supplementary Fig. 3f,g). Gained BAF complex sites were selectively enriched for unique motifs such as the AP-1 motif (Fig. 2f and Supplementary Fig. 3h). In addition, we
evolved, implicating an evolutionarily recent cell of origin for MRTs. These results demonstrate a widespread gain of BAF complex chromatin occupancy driven by BAF47 rescue and, further, show that BAF47–rescued BAF complex sites are distinct in localization and predicted functional properties from conserved BAF complex sites.

We next sought to determine the sensitivity of BAF47-deficient sarcoma cell lines spanning MRT (four lines), EpS (four lines), and

Figure 1 BAF47 confers BAF complex stability on chromatin without affecting intra-complex subunit stability. (a) Schematic for rescue experiments in BAF47-deficient cell lines. (b) Nuclear extract inputs and anti-BRG1 IP from G401 nuclear extracts in empty and BAF47 conditions. (c) Nuclear extract input and IPs for IgG, BRG1, and BAF47 in control and BAF47/KIA1 HEK293T cells. (d) Silver stain analysis of control IgG and anti-BRG1 IPs in G401 (empty) and BAF47 conditions. (e) Anti–BRG1 IP–mass spectrometry proteomics in G401 cells in empty and BAF47 conditions for BAF complex subunits. (f) Density sedimentation analyses using 10–30% glycerol gradients (10 ml; 0.5 ml/fraction) on nuclear extracts from G401 MRT cells in the empty and BAF47 conditions. (g) Schematic for differential salt extraction experiments in G401 cells in empty or BAF47 conditions (left) and immunoblot analysis of BAF complex subunits in differential salt extraction experiments (right). (i) Relative densitometry from differential salt extraction. Error bars, mean ± s.e.m. for n = 3 biological replicates. Experiments were repeated at least twice for each cell line.
AT/RT (two lines), to BAF47 rescue (Supplementary Table 2). In all MRT and AT/RT cell lines assessed, we found marked proliferative arrest upon BAF47 rescue; however, this occurred in only one of the EpS cell lines, with the three others showing no significant proliferative arrest upon BAF47 rescue (Fig. 2h and Supplementary Fig. 4).

To validate our genome-wide findings from TTC1240 cells and to decouple changes from BAF47 reintroduction and subsequent proliferative suppression, we performed ChIP-seq for BRG1 and BAF155 in G401 cells as well as in the HS-ES-2M and VA-ES-BJ EpS lines. We observed a similar gain in BAF complex occupancy upon BAF47 rescue to promoter-distal sites, irrespectively of the cell line used (Fig. 2i–k and Supplementary Fig. 5). These data demonstrate that reintroduction of the BAF47 subunit drives a consistent, widespread gain of genome-wide BAF complex occupancy across distinct BAF47-deficient sarcoma subtypes, independently of sensitivity to BAF47-mediated growth suppression.
BAF complex–mediated enhancer activation requires BAF47

To determine the effect of gained BAF complex occupancy on the histone landscape, we performed ChIP-seq studies for H3K4me3, H3K4me1, and acetylation of H3K27 (H3K27ac). Notably, we found substantial gains in H3K27ac and H3K4me1, but very minor changes in H3K4me3 levels in TTC1240 cells (Fig. 3a), suggesting that gained BAF complex occupancy predominantly determines both enhancer state and enhancer activation⁴⁰. We found that this activation was specific to distal enhancer sites (Fig. 3b-d) and observed similar gains in enhancer activation across all cell lines studied (Supplementary Fig. 6). In addition, we noted the presence of enhancer sites that retained BAF complex occupancy and activation irrespectively of BAF47 status, suggesting alternate activators at these sites. We found a strong correlation between the log₂ fold change in occupancy of BRG1 and H3K27ac (Pearson correlation coefficient (PCC) = 0.82) over all TTC1240 BAF sites (in both control and BAF47 conditions), and a lower but strong correlation with H3K4me1 (PCC = 0.56), likely owing to mark abundance; we observed minimal correlations with H3K4me3 (PCC = 0.15) (Fig. 3e and Supplementary Fig. 7a-c). These results indicate that in addition to BAF complex–mediated enhancer state and activation, the levels of BAF complex occupancy correspond directly to degree of enhancer activation.

Connecting BAF47-mediated enhancer gain to gene expression, we find that the number of gained distal BAF sites associated with a target gene correlated with greater gene activation in TTC1240 cells (Fig. 3f) and G401 cells (Supplementary Fig. 7d). Given that clusters of enhancers mediate the greatest degrees of gene activation, we sought to determine whether BAF47 activates only enhancers or super-enhancers. We found that this widespread enhancer activation upon rescue of BAF47 occurs at both typical enhancers (12,875) and super-enhancers (283) and that both typical and super-enhancers are retained in the absence of BAF47 (Fig. 3g and Supplementary Fig. 7c,f), in contrast to previous reports⁴⁵.

We then performed chromosome conformation capture followed by massively parallel sequencing (Hi-C) in VA-ES-BJ cells to determine whether BAF47 rescue affects global chromatin topology independently

Figure 3 Gain of BAF complex occupancy drives widespread enhancer activation. (a) Heat maps of BRG1, H3K4me3, H3K4me1, and H3K27ac sites in TTC1240 cells in empty and BAF47 conditions over all BRG1–BAF155 shared sites in the TTC1240+BAF47 condition. Heat maps are ranked by BRG1 occupancy in empty condition. (b–d) Metagene plots of BRG1–BAF155 sites in the TTC1240+BAF47 condition split by promoter-proximal (≤5 kb from TSS; left) and promoter-distal (>2 kb from TSS; right) for H3K4me3 (b), H3K4me1 (c), and H3K27ac (d) occupancy. (e) Correlation plot of log₂(fold change) (log₂(FC)) for BRG1 and H3K27ac over all BRG1–BAF155 sites (70,777) in empty or BAF47 conditions. (f) Gained promoter-distal BAF complex sites assigned to nearest gene (genes were categorized on the basis of number of gained distal sites) versus log₂(FC) in expression. (g) Example tracks of BRG1, BAF155, H3K4me3, H3K4me1, H3K27ac, and RNA-seq at the TGM2 locus in TTC1240 cells. n = 3 biological replicates. Experiments were repeated at least twice for each cell line.
Enhancer activation is mediated by BAF, not PBAF, complexes

Mammalian SWI/SNF complexes exist in two different assemblies: canonical BAF complexes and polybromo-containing BAF (PBAF) complexes. The two are defined by distinct subunits (Fig. 4a) but share the BAF47 core subunit. Given that both BAF and PBAF complexes remain intact in the absence of BAF47 (Fig. 4b), we performed ChIP-seq for SS18 (BAF complex–specific) and BAF200 (ARID2; PBAF complex–specific) to determine how rescue of BAF47 influences the targeting of each complex. SS18-marked complexes exhibited substantially greater gains in chromatin occupancy than BAF200 over gained BAF complex sites, mirroring those of BRG1 and BAF155 (Fig. 4c). SS18-marked BAF complexes showed a dramatic gain of occupancy at distal sites, whereas gains in BAF200-marked PBAF complexes were almost entirely restricted to proximal sites (Fig. 4d–f and Supplementary Fig. 9a). We measured log_{2} fold changes in BRG1, SS18, and BAF200 occupancy upon BAF47 reintroduction and found that BRG1 and SS18 showed high correlation (PCC = 0.88), whereas BRG1 and BAF200 exhibited a substantially more modest correlation (PCC = 0.49) (Fig. 4g,h and Supplementary Fig. 9b). BRG1 and BAF200 showed stronger correlations at proximal sites (0.56) than at distal sites (0.43), whereas BRG1 and SS18 correlations were similar (proximal PCC = 0.82, distal PCC = 0.86), suggesting a greater role for retargeting of PBAF complexes to of proliferative arrest, which has been suggested as a mechanism of oncogenesis in other cancers1,2. While we found that BAF47 had no significant impact on global genome architecture (Supplementary Fig. 8a–e), we identified new promoter–enhancer interactions at gained enhancers such as CDKN1A (Supplementary Fig. 8f), possibly due to downstream effects of enhancer activation. These results collectively suggest that BAF47 has a key role in mediating activation of constituent enhancers at both typical and super-enhancer clusters, with large clusters of de novo–gained BAF complex target sites promoting greatest gene activation.
proximal sites (Supplementary Fig. 9c,d). These results demonstrate a disproportionate targeting of BAF complexes to enhancers and PBAF complexes to promoters, with BAF47 driving widespread enhancer activation by BAF complexes (Fig. 4i).

BAF47 rescues BAF complex–mediated resolution of bivalency

Loss of the opposition between BAF complexes and Polycomb repressive complex 2 (PRC2) has been extensively implicated in MRT, suggesting mechanisms of global regulatory opposition and locus-specific opposition (e.g., at the p16INK4A locus)\(^26\). However, to date, BAF–PRC2 complex opposition has not been studied at a genome-wide level. We performed ChIP-seq for SUZ12 (a core PRC2 subunit) and H3K27me3 and found that over all promoters, occupancy of BAF and PBAF complexes correlated with H3K4me3 occupancy as well as gene expression, whereas H3K27me3 and SUZ12 exhibited highest occupancy at nonexpressed genes (Fig. 5a and Supplementary Fig. 10a–h). This suggests that BAF and PBAF complexes have a maintenance role at active promoters, even in the absence of BAF47. We found a set of bivalent promoters in TTC1240 cells, marked by both H3K4me3 and H3K27me3, with 3,022 (12.57%) genes categorized as bivalent (Fig. 5a and Supplementary Fig. 10i,j). Gene Ontology (GO) term analysis of genes with bivalent promoters strongly enriched for genes involved in kidney and neural development, likely reflecting the initiation of cell lineage–specific regulation (Fig. 5b). We also performed ChIP-seq analyses for SUZ12 and H3K27me3 in G401 cells and found, similarly to our results in TTC1240 cells, that 2,470 (10.27%) genes were bivalent. Notably, the majorit (1,902) of these bivalent genes were shared between G401 and TTC1240 cell lines (Fig. 5c and Supplementary Fig. 10k,l), suggesting a concordant, lineage-specific set of bivalent genes in MRT cell lines.

Given that recent studies have demonstrated efficacy of EZH2 inhibitors in MRT cell lines\(^34\) and that EZH2 inhibitors may work through activating bivalent genes\(^35\), we sought to determine whether BAF47 rescue results in altered bivalent promoter regulation. We found a marked increase in target gene occupancy upon rescue of BAF47, in both BAF and PBAF complexes (Fig. 5d and Supplementary Fig. 11a). Notably, while only 8.03% of conserved BRG1 target genes were bivalent in TTC1240 cells, 30.13% of gained BRG1 target genes were bivalent (Fig. 5e). This results in an increase in BAF complex occupancy at bivalent promoters, from 29.2% to 68.6%.
of BAF47, the greatest percentage increase of any promoter category (Supplementary Fig. 11b). We next aimed to assess whether the number of bivalent genes was affected by rescue of BAF47. We found that 506 (16.7%) genes that were bivalent in TTC1240 cells in control conditions were no longer bivalent in BAF47 conditions (Fig. 5f), suggesting a role for BAF complex–mediated activation at these specific sites upon rescue with BAF47.

Examining all BRG1–BAF155 sites in TTC1240 cells in BAF47 conditions, we found that rescue of BAF47 led to a gain in BAF complex occupancy and a decrease in H3K27me3 occupancy, and that this occurred predominantly at promoter proximal sites (Supplementary Fig. 11c–e). Over all bivalent promoters in TTC1240 cells, we observed a major decrease in H3K27me3 but no changes in H3K4me3 (Fig. 5g and Supplementary Fig. 11f), suggesting the regulation of bivalency is due solely to regulation of Polycomb-mediated repression, in contrast to previous findings.26 Over bivalent promoters, we observed a gain in both BAF and PBAF complex occupancy (Fig. 5h and Supplementary Fig. 11f). Notably, log₂ fold change values over proximal TTC1240 BRG1–BAF155 sites showed opposition of BRG1 and H3K27me3 (PCC = −0.26) and of SS18 and H3K27me3 (PCC = −0.24), but opposition was greater between BAFO and H3K27me3 (PCC = −0.40) (Fig. 5i and Supplementary Fig. 11g–i). We obtained...
Collaborative activation of bivalent promoters and enhancers

Having identified two mechanisms through which BAF47 affects gene activation, by both BAF and PBAF complexes, we sought to determine the relative contribution of each mechanism in gene regulation and subsequent tumor suppression. We performed RNA-seq on G401 and TTC1240 cell lines with either empty vector or BAF47 (Supplementary Fig. 12). We found that bivalent genes were overrepresented in the 2,635 significantly regulated genes in TTC1240 cells, with 20.80% marked as bivalent (Fig. 6a). We found that 13.99% of all bivalent genes were upregulated by BAF47, the largest proportion of any gene category, and that bivalent genes showed net upregulation when compared to H3K4me3-only genes (Fig. 6b). We observed similar effects in G401 cells (Supplementary Fig. 14c,d).

We identified a set of 642 genes that were significantly and concordantly regulated in both G401 and TTC1240 cells (Fig. 6c,d). Gene-set enrichment analysis (GSEA) and GO term analyses showed that upregulated genes have critical roles in kidney development and epithelial–mesenchymal transition (EMT), suggesting these as key pathways involved in BAF47-mediated tumor suppression (Fig. 6e–h). This would correspond with mechanisms of mesenchymal–epithelial transition (MET) in sarcomagenesis and roles for BAF47 in EMT of pancreatic cancer. We did not observe the previously suggested downregulation of EZH2 or upregulation of CDKN2A genes in TTC1240, suggesting this is specific to certain cell lines (Supplementary Fig. 14i,j). Increases in BAF complex occupancy and decreases in PRC2-mediated repression were specific to the promoters of upregulated genes upon BAF47 rescue (Supplementary Fig. 15a–f). These results suggest that while MRTs are heterogeneous and multi-origin tumors, critical cellular processes such as EMT may be regulated by BAF47-containing BAF complexes and altered in sarcomagenesis driven by BAF47 loss.

We next sought to determine how regulation of enhancers and bivalent promoters cooperate to control gene regulation. We categorized genes by the number of conserved (empty-BAF47) and gained (BAF47-only) distal BAF complex sites and found overrepresentation of bivalent genes among those with greater numbers of gained distal BAF complex sites (Fig. 6g and Supplementary Fig. 14k), and these genes corresponded to the upregulation described above (Fig. 3f). This collaborative activation is exemplified at the CTGF and FN1 loci, at which these activated genes gain (i) promoter occupancy that resolves bivalency to activation and (ii) numerous distal BAF complex–activated enhancers (Fig. 6h,i). We demonstrate here that these enhancer and promoter regulatory functions are hallmarks of BAF47-mediated gene regulation and tumor suppression, and that a critical collaborative role for these two distinct BAF complex functions is lost in BAF47-deficient cancers.

DISCUSSION

Mutations in the genes encoding BAF complexes are recurrent in more than 20% of human cancers and in several developmental disorders. Here we show that loss of a core BAF complex subunit, BAF47, dramatically impairs the chromatin affinity and regulation of BAF complexes without impairing in-solution assembly or subunit stability (Fig. 7a). We show that BAF complexes have a critical role in mediating enhancer state and activation, as well as in resolving bivalent promoters to activation through opposition of Polycomb-mediated repression, and that these activities collaborate in BAF47-driven gene activation and tumor suppression in MRT cell lines (Fig. 7b). Our data suggest a broad-reaching role for BAF complexes in directing cell state through regulation of developmental enhancers and bivalent promoters that may contribute to numerous cancers and developmental disorders characterized by BAF complex perturbations.

Understanding the role of BAF47 in the stability and assembly of BAF complexes has been a challenge in the field, with some studies suggesting little change in complex composition and others suggesting dramatic loss of stability upon BAF47 loss. We did not find global changes in BAF complex stability or assembly, in agreement with recent yeast studies, which show largely retained SWI/SNF assembly but loss of ATPase function upon loss of Snf5 (ref. 46). The structural integrity of BAF47-deficient residual BAF complexes shown here supports previous observations of BRG1 dependency in MRT cell lines and suggests therapeutic avenues for targeting these intact residual BAF complexes in MRT. Our results suggest that this dependency is due to a retained and required regulation by BAF and PBAF complexes, largely at active promoters, in the absence of BAF47.

The opposition between BAF and Polycomb complexes is a critical mechanism that has been extensively implicated, through genetic and...
Mechanistic studies, in cancer.23,26,28,29 Our results reaffirm the BAF complex as a Trx-group protein through opposition of Polycomb-mediated repression. This study substantiates how BAF complexes, within the existing dynamics of Trx or PcG proteins, drive resolution of bivalent promoters toward activation.32 These results help to explain the observed efficacy of EZH2 inhibitors in BAF47-deficient sarcomas and suggest a mechanistic synergy between EZH2 inhibition and BAF47 rescue, particularly in the control of bivalent promoters demonstrate a broad role for BAF complexes at both bivalent promoters and enhancers, which, when perturbed, may explain the outsized role for this complex in human disease.

We establish here that BAF47 drives a widespread gain of BAF complex occupancy that mediates enhancer state and activation, such that BAF47 restores the functional ability of the BAF complex to bind to and activate distal regulatory sites. As such, recruitment by transcription factors or other regulators fails to create BAF complex targeting and activation in the absence of BAF47, as previously shown with Arid1a loss promoting regeneration.32 Enhancers have co-occupancy of numerous chromatin remodelers and regulators30,48, and our results suggest a dynamic regulation of enhancers by both activating (BAF complex) and repressive (NuRD complex)49 remodelers that may be occurring in a manner similar to histone acetyltransferases and histone deacetylases.50 Interruption of this regulatory dynamic via BAF47 loss may then decommission a large number of enhancers in the MRT genome, further supporting a critical pioneering role for BAF complexes at enhancers.

In summary, our studies demonstrate that reintroduction of BAF47 in BAF47-deficient cells triggers a dual gain of BAF complex–mediated activation at enhancers and bivalent promoters. We demonstrate that the enhancer activation and resolution of bivalency to activation are collaborative, leading to activation of key genes involved in cell fate determination and tumor suppression. Further studies will be required to determine the ordering and relative contributions of these functions in the tumor suppression pathway. Our data suggest multiple defining roles for the BAF complex in chromatin activation at enhancers and bivalent promoters, each of which could be independently or collaboratively perturbed in other BAF complex–driven cancers. Taken together, these data have important implications for the outsized contribution of BAF complex aberrations in human malignancy and developmental disorders.

URLs. GSEA, http://www.broadinstitute.org/gsea.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

Acknowledgments

We thank members of the Molecular Biology Core Facility at the Dana-Farber Cancer Institute, especially Z. Herbert, for expertise and technical assistance with ChIP-seq and RNA-seq data sets. We thank G.R. Crabtree, A. Kuo, and W. Wang for anti-BRGG1 (I1) and anti-BAF155 rabbit polyclonal antibodies used in ChIP-seq experiments. The authors are grateful to T.J. Triche (Children's Hospital Los Angeles), P. Houghton (The University of Texas Health Science Center), H. Kawashima (Niigata University Graduate School of Medical and Dental Sciences), H. Iwasaki (Fukuoka University), H. Gotoh (Kanagawa Childrens Medical Center), and T. Tskahara (Sapporo Medical University) for providing the MRT and EpS cell lines. This work was supported in part by awards from the NIH DP2 New Innovator Award 1DP2CA195762-01 (C.K.), the American Cancer Society Research Scholar Awards RSG-14-051-01-DMC (C.K.), the Pew-Stewart Scholars in Cancer Research Grant (C.K.), the Alex's Lemonade Stand Foundation Young Investigator Award (C.K.), the Rare Cancer Research Foundation (C.K.), the WWWWW Foundation (QuadW) (R.T.N.), and the Uehara Memorial Foundation (R.T.N.). Additionally, this work was supported by the US National Institutes of Health (NIH) grant 5 T32 GM095450-04 (A.M.V.) and National Institute of General Medical Sciences (NIGMS) award T32GM007753 (S.H.C.).

Author Contributions

R.T.N., J.L.P., A.M.V., and C.K. conceived of and designed the experiments. R.T.N., A.M.V., M.J.M., Z.M.M., R.T.W., S.H.C., H.Q., and C.J.W. performed experiments. J.L.P. performed all bioinformatics analyses and statistical calculations. W.I.K., K.C., and K.Z. performed and analyzed chromatin capture experiments. M.A.G., A.M.V., M.J.M., Z.M.M., R.T.W., S.H.C., H.Q., and C.J.W. performed experiments. M.T. and R.A.I. provided advice for bioinformatics analyses. G.D.D. provided valuable conceptual advice and guidance. J.L.P., R.T.N., and C.K. wrote the paper.

Competing Financial Interests

The authors declare competing financial interests: details are available in the online version of the paper.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Dawson, M.A. & Kouzarides, T. Cancer epigenetics: from mechanism to therapy. Cell 150, 12–27 (2012).
2. Clapier, C.R. & Cairns, B.R. The biology of chromatin remodeling complexes. Annu. Rev. Biochem. 78, 273–304 (2009).
3. Peterson, C.L. & Herskowitz, I. Characterization of the yeast SWI1/SWI2 and SWI3 genes, which encode a global activator of transcription. Cell 68, 573–583 (1992).
4. Tamkun, J.W. et al. brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell 68, 561–572 (1992).
5. Hnisz, D., Imbalzano, A.N., Khavari, P.A., Kingston, R.E. & Green, M.R. Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. Nature 370, 477–481 (1994).
6. Ho, L. et al. An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc. Natl. Acad. Sci. USA 106, 5181–5186 (2009).
7. Singhal, N. et al. Chromatin-remodeling components of the BAF complex facilitate reprogramming. Cell 141, 943–955 (2010).
8. Lessard, J. et al. An essential switch in subunit composition of a chromatin remodeling complex during neural development. Neuron 55, 201–215 (2007).
9. Ho, L. & Crabtree, G.R. Chromatin remodelling during development. Nature 463, 474–484 (2010).
10. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. Cell 155, 934–947 (2013).
11. de Dieuveuelt, M. et al. Genome-wide nucleosome specificity and function of chromatin remodelers in ES cells. Nature 530, 113–116 (2016).
12. Versteeghe, I. et al. Truncating mutations of NSF5/SNFI in aggressive paediatric cancer. Nature 394, 203–206 (1998).
13. Chun, H.J. et al. Genome-wide profiles of extracranial malignant rhabdoid tumors reveal heterogeneity and dysregulated developmental pathways. Cancer Cell 29, 394–406 (2016).
14. Biegel, J.A. et al. Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. Cancer Res. 59, 74–79 (1999).
15. Modena, P. et al. SMARCB1/INI1 tumor suppressor gene is frequently inactivated in epithelial sarcomas. Cancer Res. 65, 4012–4019 (2005).
16. Christiania, L. et al. Germline SMARCB1 mutation and somatic NF2 mutations in familial multiple meningiomas. J. Med. Genet. 48, 93–97 (2011).
17. Hulsebos, T.J. et al. Germline mutation of INI1/SMARCB1 in familial schwannomatosis. Am. J. Hum. Genet. 80, 805–810 (2007).
18. Tsuchiya, Y. et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat. Genet. 44, 376–378 (2012).
19. McKenna, E.S. et al. Loss of the epigenetic tumor suppressor SNF5 leads to cancer without genomic instability. Mol. Cell. Biol. 30, 6223–6233 (2008).
20. Lawrence, M.S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499, 214–218 (2013).
21. Roberts, C.W., Leroux, M.M., Fleming, M.D. & Orkin, S.H. Highly penetrant, rapid schwannomatosis. Nature 463, 474–484 (2010).
22. Kadoch, C. et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nat. Genet. 45, 592–601 (2013).
23. Kadoch, C. & Crabtree, G.R. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. Cell 153, 71–85 (2013).
24. Kennison, J.A. & Tankum, J.W. Dosage-dependent modifiers of Polycomb and antennapedia mutations in Drosophila. Proc. Natl. Acad. Sci. USA 85, 8136–8140 (1988).
25. Kadoch, C., Copeland, R.A. & Keilhack, H. PRC2 and SWI/SNF chromatin remodeling complexes in health and disease. Biochemistry 55, 1600–1614 (2016).
26. Kia, S.K., Gorski, M.M., Giannakopoulos, S. & Verrijzer, C.P. SWI/SNF mediates Polycomb eviction and epigenetic reprogramming of the INK4d–ARF–INK4a locus. Mol. Cell. Biol. 28, 3457–3464 (2008).
27. Ho, L. et al. esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating Polycomb function. Nat. Cell Biol. 13, 903–913 (2011).
28. Wilson, B.G. et al. Epigenetic antagonism between Polycomb and SWI/SNF complexes during oncogenic transformation. Cancer Cell 18, 316–328 (2010).
29. Kadoch, C. et al. Dynamics of BAF–Polycomb complex opposition on heterochromatin in normal and oncogenic states. Nat. Genet. 49, 213–222 (2017).
30. Bernstein, B.E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125, 315–326 (2006).
31. Mikkelsen, T.S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448, 553–560 (2007).
32. Voigt, P., Tee, W.W. & Reinberg, D. A double take on bivalent promoters. Genes Dev. 27, 1318–1338 (2013).
33. Jadhav, U. et al. Acquired tissue-specific promoter bivalency is a basis for PRC2 necessity in adult cells. Cell 165, 1389–1400 (2016).
34. Knutson, S.K. et al. Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. Proc. Natl. Acad. Sci. USA 110, 7922–7927 (2013).
35. Wang, X. et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. Nat. Genet. 49, 289–295 (2017).
36. Wei, D. et al. SNF5/INI1 deficiency redefines chromatin remodeling complex composition during tumor development. Mol. Cancer Res. 12, 1574–1585 (2014).
37. Doan, D.N. et al. Loss of the INI1 tumor suppressor does not impair the expression of multiple BRG1-dependent genes or the assembly of SWI/SNF enzymes. Oncogene 23, 3462–3473 (2004).
38. Jamshidi, F. et al. The genomic landscape of epithelioid sarcoma cell lines and tumours. J. Pathol. 238, 63–73 (2016).
39. Wang, X. et al. Oncogenesis caused by loss of the SNF5 tumor suppressor is dependent on activity of BRG1, the ATPase of the SWI/SNF chromatin remodeling complex. Cancer Res. 69, 8094–8101 (2009).
40. Creyghton, M.P. et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA 107, 21931–21936 (2010).
41. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science 351, 1454–1458 (2016).
42. Flavahan, W.A. et al. Insulator dysfunction and oncogene activation in IDH1 mutant gliomas. Nature 529, 110–114 (2016).
43. Middeljans, E. et al. SS18 together with animal-specific factors defines human BAF-type SWI/SNF complexes. PLoS One 7, e33834 (2012).
44. Yang, J. et al. Mesenchymal to epithelial transition in sarcomas. Eur. J. Cancer 50, 593–601 (2014).
45. Genovese, G. et al. Synthetic vulnerabilities of mesenchymal subpopulations in pancreatic cancer. Nature 542, 362–366 (2017).
46. Sen, P. et al. Loss of Snf5 induces formation of an aberrant SWI/SNF complex. Cell Rep. 18, 2135–2147 (2017).
47. Sun, X. et al. Suppression of the SWI/SNF component Arid1a promotes mammalian regeneration. Cell Stem Cell 18, 456–466 (2016).
48. Morris, S.A. et al. Overlapping chromatin-remodeling systems collaborate genome wide at dynamic chromatin transitions. Nat. Struct. Mol. Biol. 21, 73–81 (2014).
49. Yildirim, O. et al. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell 147, 1498–1510 (2011).
50. Wang, Z. et al. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell 138, 1019–1031 (2009).
ONLINE METHODS

Cell lines and tissue culture. Eight MRT cell lines and nine EpS cell lines were used in this study (Supplementary Table 2). Of these, four cell lines were purchased from ATCC (GA01, GA02, A204, and VA-ES-B1), and three were from RIKEN (HS-ES-1, HS-ES-2R, and HS-ES-2M). TTYC1240, TM87-16, and STM19-01 were a generous gift from T.J. Triche (Children’s Hospital Los Angeles); BT12 and BT16 were from P. Houghton (The University of Texas Health Science Center at San Antonio); NEPS was from H. Kawashima (Niigata University Graduate School of Medical and Dental Sciences); FU-EPS-1 and SFT-8606 were from H. Iwasaki (Fukuoka University); YCUS-5 was from H. Gotoh (Kanagawa Children’s Medical Center); ESX was from T. Tsukahara (Sapporo Medical University). Cell lines were cultured either in DMEM/F12, RPMI 1640, or DMEM ( Gibco), supplemented with 10% FBS, 1% GlutaMAX ( Gibco) and 1% penicillin–streptomycin ( Gibco).

Vector and cloning information. C-terminal V5-tagged BAF47 (SMARCB1) constitutive expression in MRT and EpS cell lines was achieved using lentiviral infection of an EF1α-driven expression vector (modified from Clontech, dual promoter EF-1α-MCS-PGK-Blast), selected with blasticidin (1 µg/µl).

Lentiviral generation. Lentivirus was produced by PEI (Polysciences Inc.) transfection of HEK293T LentiX cells (Clontech) with gene delivery vector co-transfected with packaging vectors papax2 and pMD2.G as previously described.21 Supernatants were harvested 72 h after transfection and centrifuged at 20,000 r.p.m. for 2 h at 4 °C. Virus-containing pellets were resuspended in PBS and placed on cells dropwise. Selection of lentivirally infected cells was achieved with either blasticidin or puromycin, both used at 2 µg/ml.

Nuclear extract. Nuclear extract (NE) preparation and IP studies were performed as described previously4. Briefly, trypsinized cells were incubated in buffer A (25 mM HEPES, pH 7.6, 5 mM MgCl2, 25 mM KCl, 0.05 mM EDTA, 10% glycerol and 0.1% NP40 with protease inhibitor (Roche), 1 mM DTT) for 10 min, and the pellets were resuspended in 600 µl buffer C (10 mM HEPES, pH 7.6, 3 mM MgCl2, 100 mM KCl, 0.5 mM EDTA and 10% glycerol with protease inhibitor, 1 mM DTT and 1 mM PMSF) with 67 µl 3 (NH4)2SO4 for 20 min. The lysates were spun down in an ultracentrifuge at 10,000 r.p.m. at 4 °C for 10 min. Nuclear extracts were precipitated with 200 µg (NH4)2SO4 on ice for 20 min and purified as pellets by ultracentrifugation at 10,000 r.p.m. at 4 °C for 10 min. The pellets were resuspended in IP buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA and 1% Triton X-100 with protease inhibitor, 1 mM DTT and 1 mM PMSF) or IP buffer (same as above, except 300 mM NaCl). To analyze the localization of the protein, NE-PER Nuclear and Cytoplasmic Extraction Reagents (#78833, Thermo Scientific) were used according to the manufacturer’s protocol. The details of the antibodies used for immunoblotting are presented in Supplementary Table 3.

Immunoprecipitation. For immunoprecipitation, 150–300 µg of nuclear extract was incubated with 1.25 µg of antibody in IP buffer overnight. Samples were incubated with Dynabeads (Thermo Scientific) for 2 h. Beads were washed three times with IP buffer and twice with BC100 (20 mM HEPES, 100 mM KCl, 0.2 mM EDTA, 10% glycerol), and eluted with 20 µl of sample buffer (NuPage LDS buffer (1×, Thermo Scientific) and 100 mM DTT).

SMARCB1 knockout by CRISPR–Cas9 genome editing. The SMARCB1 locus was targeted by the In1 CRISPR–Cas9 KO Plasmid and In1 HDR Plasmid (Santa Cruz Biotechnology sc–423027; sc–423027-HDR) in HEK293T Lenti-X cells (Clontech) following the manufacturer’s protocol. Specifically, 5 million HEK293T cells were co-electroporated with two plasmids (2 µg DNA/plasmid) using the Amaxa Biosystems Nucleofector I and Amaxa Cell Line Nucleofector Kit V. After nucleofection, cells were expanded for 48 h and GFP+/RFP+ cells (expressing both the KO and HDR plasmids) were single-cell sorted by FACS. Single-cell clones were expanded further and screened through immunoblotting for identification of successful knockouts.

Two-dimensional liquid chromatography–mass spectrometry IP proteomics. Four samples for each cell type were prepared (two IgG controls and two replicates of anti-BRG1 IPs) for mass spectrometry analysis (label-free quantification). Eluted proteins from each condition were processed simultaneously to reduce sample variability. Proteins were reduced, alkylated, digested with trypsin, and desalted using offline C18 reversed-phase chromatography. Purified peptides were separated by online C18 reversed-phase chromatography and analyzed in a top10 collision-induced dissociation data-dependent manner using an LTQ-Velos mass spectrometer.51

Two-dimensional liquid chromatography–mass spectrometry IP proteomics data processing and analysis. Data analysis was performed with MaxQuant software52, supported by a database search engine for peptide identification (human IPI). Label-free quantification algorithms were added to MaxQuant by extracting isotope patterns for each peptide in each run.

Density sedimentation analyses. Nuclear extract (500 µg) was resuspended in 200 µl 0% glycerol HEMG buffer and carefully overlaid onto a 10 ml 10–30% glycerol (in HEMG buffer) gradient prepared in a 14 × 89 mm polyallomer centrifuge tube (331327, Beckman Coulter). Tubes were centrifuged in an SW40 rotor at 4 °C for 16 h at 40,000 r.p.m. Fractions (0.5 ml) were collected and used in analyses.

Urea denaturation studies. Nuclear extracts (150 µg) were subjected to partial urea denaturation in 0.25 to 5.0 M urea (in IP buffer), for 15 min at room temperature before anti-BRG1 IP. The co-precipitated proteins were analyzed by immunoblotting.

Differential salt extraction. Cell types were grown under standard conditions. Following collection of 5 × 106 cells, cells were subjected to elution buffer (50 mM Tris-HCl at pH 7.5, 1 mM EDTA, 0.1% NP40) supplemented with protease inhibitor mixture (Roche) and 1 mM PMSF, incubated on ice for 5 min, and centrifuged. Supernatant was collected, and the pellet was subjected to elution buffer with 150 mM NaCl. This process was repeated sequentially with increasing concentrations of NaCl to collect 0, 150, 300, 500, and 1,000 mM NaCl soluble fractions. Each fraction, including the total (5 × 106 cells in elution buffer) and pellet (material remaining following 1,000 mM NaCl extraction) fractions, was prepared in SDS (final concentration of 1%), quantified by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific), and analyzed (1.5 µg of protein) by immunoblot. Quantitative densitometry analyses were performed with the Li-Cor Odyssey Imaging System (Li-COR Biosciences).

ChIP-seq data collection. Cells were harvested following 48-h exposure to the lentivirus and 5-d selection with 10 µg/ml of blasticidin for chromatin immunoprecipitation (ChIP) experiments. ChIP experiments were performed per standard protocols (Millipore) with minor modifications. Briefly, cells were cross-linked for 10 min with 1% formaldehyde at 37 °C. Five million fixed cells were used per chromatin immunoprecipitation experiment. This reaction was subsequently quenched with 125 mM glycine for 5 min. Antibodies used for ChIP studies are listed in Supplementary Table 3.

RNA data collection. Cells were harvested following 48-h exposure to the lentivirus and either 1-d (day 3 after infection) or 5-d (day 7 after infection) selection with 10 µg/ml of blasticidin for RNA-seq experiments. RNA-seq samples were prepared in biological duplicate (independent lentiviral production, infection, selection, and cell culture). All RNA was obtained using the RNeasy Mini Kit (Qiagen).

Proliferation experiments. 20,000 cells were plated after 48-h exposure to the lentivirus and 48-h selection with 10 µg/ml of blasticidin, with day 0 denoting the day cells were plated after infection and selection. The numbers of viable cells in three wells were measured using Vi-CELL Cell Viability Analyzer (Beckman Coulter) on days 1, 3, 5, and 7.
Library prep and sequencing for ChIP-seq and RNA-seq. All library prep and sequencing (75-bp single end on Illumina NextSeq 500) was performed in the Molecular Biology Core Facilities at the Dana-Farber Cancer Institute.

Sequence data processing. ChIP-seq reads were mapped to the human reference genome (hg19) using Bowtie2 (ref. 53) version 2.1.0 with parameters –k 1. RNA-seq reads were mapped to the human reference genome (hg19) using STAR38 version 2.3.1 with default parameters. All sequence data is deposited in the Sequence Read Archive under GSE90634. Summary statistics for sequencing experiments are presented in Supplementary Table 4.

ChIP-seq data analysis. Peaks were called against input reads using MACS2 (ref. 55) version 2.1.0 at q = 1 × 10−3. Narrow peak calls were used for BRG1, BAF155, SS18, BAF200, and H3K4me3, and broad peak calls were used for H3K27ac, H3K4me1, H3K27me3, and SUZ12. Peaks were filtered to remove peaks that overlap with ENCODE blacklisted regions, as well as peaks mapped to unmappable chromosomes (only chr1-22, X, Y included). Duplicate reads were removed using SAMTools rmdup for all downstream analyses. ChIP-seq track densities were generated per million mapped reads with MACS2 2.1.0 using parameters –b –SPMR.

BAF complex sites were determined in each condition using the bedtools overlap of the BRG1 and BAF155 sites, and this peak set was used in downstream analyses for determination of BAF complex targeting. Conserved sites were determined as sites with peaks overlapping in both empty and BAF47 conditions; gained sites were determined using sites only in the BAF47 condition. Venn diagrams were generated with the R statistical package, using the minimum number of overlapping regions for resolving multiple peak overlaps.

Metagene read densities were generated using HTSeq56, with fragment length extended to 200 bp to account for the average 200-bp fragment size selected in sonication, using the center of peak calls from MACS2. Total read counts for each region were normalized to the number of mapped reads to give reads per million mapped reads. Metagene plots were generated using average read density across all sites indicated for each condition. Heat maps were generated using the same HTSeq read densities as metagene plots; sites were then ranked by mean ChIP-seq signal for the epitope and condition indicated in each figure. Heat maps were visualized using Python matplotlib with a midpint of 0.5 reads per million for the heat map color scale to set the threshold for visualization.

To generate plots of log2 fold change for ChIP-seq reads, the peak sets for the BAF complex (BRG1–BAF155 overlapping sites) in the empty and BAF47 conditions were merged using bedtools merge, generating a total of 70,777 BRG1–BAF155 sites in TCC1240. ChIP-seq read counts for each BAF complex site were generated using Rsubread featureCounts, and read counts in each peak region were normalized per million mapped reads. Input RPM values for each region in each condition were subtracted from each ChIP epitope in that condition, and values with higher input enrichment than ChIP enrichment were set to 0. Log2 fold change values were determined for each ChIP epitope using the normalized RPM values above, with a pseudocount of 0.1. Pairwise correlation was determined using a Pearson correlation coefficient between normalized fold change values for each pair of ChIP experiments.

For motif enrichment analysis, 500-bp core sequences centered on peak centers were submitted to MEME-Chip analysis37. Conservation scores were calculated using bedtools map –o mean to generate the average PhyloP score for each 500-bp core sequence as in motif analysis, using the PhyloP 46-way vertebrate conservation score from UCSC26,59. Determination of super-enhancers was performed using ROSEP60,61 with a union peak set of H3K27ac in empty and BAF47 conditions in TCC1240, run with H3K27ac ChIP-seq files in empty and BAF47 conditions to determine typical and super-enhancer designations in each condition.

Distance to TSS for ChIP-seq peaks was determined using bedtools closest with a distance threshold for visualization. Indicated in each figure. Heat maps were visualized using Python matplotlib. Sites were determined as sites with peaks overlapping in both empty and BAF47 conditions; gained sites were determined using sites only in the BAF47 condition. Venn diagrams were generated with the R statistical package, using the minimum number of overlapping regions for resolving multiple peak overlaps.

Metagene read densities were generated using HTSeq56, with fragment length extended to 200 bp to account for the average 200-bp fragment size selected in sonication, using the center of peak calls from MACS2. Total read counts for each region were normalized to the number of mapped reads to give reads per million mapped reads. Metagene plots were generated using average read density across all sites indicated for each condition. Heat maps were generated using the same HTSeq read densities as metagene plots; sites were then ranked by mean ChIP-seq signal for the epitope and condition indicated in each figure. Heat maps were visualized using Python matplotlib with a midpint of 0.5 reads per million for the heat map color scale to set the threshold for visualization.

To generate plots of log2 fold change for ChIP-seq reads, the peak sets for the BAF complex (BRG1–BAF155 overlapping sites) in the empty and BAF47 conditions were merged using bedtools merge, generating a total of 70,777 BRG1–BAF155 sites in TCC1240. ChIP-seq read counts for each BAF complex site were generated using Rsubread featureCounts, and read counts in each peak region were normalized per million mapped reads. Input RPM values for each region in each condition were subtracted from each ChIP epitope in that condition, and values with higher input enrichment than ChIP enrichment were set to 0. Log2 fold change values were determined for each ChIP epitope using the normalized RPM values above, with a pseudocount of 0.1. Pairwise correlation was determined using a Pearson correlation coefficient between normalized fold change values for each pair of ChIP experiments.

For motif enrichment analysis, 500-bp core sequences centered on peak centers were submitted to MEME-Chip analysis37. Conservation scores were calculated using bedtools map –o mean to generate the average PhyloP score for each 500-bp core sequence as in motif analysis, using the PhyloP 46-way vertebrate conservation score from UCSC26,59. Determination of super-enhancers was performed using ROSEP60,61 with a union peak set of H3K27ac in empty and BAF47 conditions in TCC1240, run with H3K27ac ChIP-seq files in empty and BAF47 conditions to determine typical and super-enhancer designations in each condition.

Distance to TSS for ChIP-seq peaks was determined using bedtools closest with a distance threshold for visualization. Indicated in each figure. Heat maps were visualized using Python matplotlib. Sites were determined as sites with peaks overlapping in both empty and BAF47 conditions; gained sites were determined using sites only in the BAF47 condition. Venn diagrams were generated with the R statistical package, using the minimum number of overlapping regions for resolving multiple peak overlaps.

RNA-seq data analysis. RPMK values for samples were generated using GGold4d version 1.1.0. All error bars represent mean ± s.e.m. Significance was assessed using the R package DESeq (ref. 65) using raw read counts generated with Rsu6rd featureCounts against the hg19 refFlat annotation. Significantly changing genes were assessed with a Bonferroni-corrected P < 1 × 10−3 and a twofold gene expression change (log2FC > 1) to determine set of significantly changing genes. GSEA was performed using the GSEA Primerank function of the JAVA program as described previously66. Rank files for GSEA were generated using RPKMs for duplicate RNA-seq in each cell line, removal of short RNAs, filtering for expressed genes (minimum RPMK value for four samples > 1), averaging replicates of each condition, then doing a log; fold change comparison with a pseudocount of 1 in each condition, i.e., log2((RPKMBAF47 + 1)/(RPKMempty + 1)). Log2 fold change values for RNA-seq in figures were identical to those used for GSEA, except nonexpressed genes were included in the analysis. Two-tailed t-tests were used to determine significance of differences for each.

RNA-seq tracks were generated using bedtools genomecov –split –sce with the mapped read count to generate tracks normalized per million mapped reads. All RNA-seq tracks visualized are day 7 after infection using a representative example in each condition. For analysis of significantly changed genes, we used day 3 RNA-seq to capture primary effects of tumor suppression (to the extent possible) before downstream regulation.

Hi-C experimental method. Cells were cross-linked with 1% formaldehyde for 10 min at room temperature. The Hi-C assay was performed as previously described67 with the following modifications. After cross-linking, the cells were lysed and digested with CviQI + CviAII + BfaI for 30 min. Enzyme-digested DNA ends were repaired and labeled with biotin-14-DATP with Klenow enzyme (large fragment). Proximity-based ligation of chromatin ends was performed using T4 DNA ligase overnight at 16 °C. DNA was reverse cross-linked at 65 °C for 6 h and purified by phenol–chloroform extraction, which was followed by treatment with T4 DNA polymerase to remove biotin from unligated DNA ends. DNA was sheared to 300–500 bp by sonication. Biotinylated DNA was enriched using streptavidin beads, and sequencing libraries were generated as described previously.

Hi-C read mapping. Sequencing data were obtained from an Illumina HiSeq 2000 machine. The paired-end tags (PEs) from Hi-C libraries were mapped to the human reference genome (hg18) using Bowtie2 (ref. 53). Only one uniquely mapped PET were considered at each genome coordinate, since the mapped PETs with the same coordinate on the genome were considered to be PCR replicates derived from the same original DNA fragments. The uniquely aligned PETs for the two biological replicates (empty and BAF47) were subjected to further analysis.

Hi-C filtering and heatmap analysis. The interaction matrices were binned using the following settings: bin size = 50 kb, bin step = 5 kb, bin mode = mean. First, we generated the matrices from the Hi-C interaction data using bin step of 5 kb. Second, the self-circularized restriction fragments were filtered by setting the diagonal elements of the matrices to 0. Third, we removed rows or columns of the matrices if the sum of elements in the row or column was 0. Fourth, we followed the approach in Hnisz et al.41 to calculate the Z-score matrices of the interaction matrices. We detected and flagged the elements of the interaction matrices if their corresponding Z-score values were greater than 21, which were considered as outlier pixel/interactions. We then took the union of all outlier pixel/interactions across all the interaction matrices and set them to 0. Fifth, the matrices were balanced according to the KR normalization
method\textsuperscript{68}, similarly to the study by Rao et al.\textsuperscript{67}. Sixth, we recovered the interaction matrices with bin size = 50 kb by combining every 10 bins into one bin with bin mode = mean. The differential heat map was the difference between the matrices in the two conditions.

**Analysis for topological associating domains (TADs).** Each chromosome was separated into 50-kb bins, and an interaction matrix of each chromosome was generated for Hi-C data. The interaction matrix was normalized by KR normalization\textsuperscript{68}. The normalized interaction matrix was used as input for identifying TAD by Armatus\textsuperscript{69}.

**Data availability.** The sequencing data sets generated and/or analyzed in the current study have been deposited in the Gene Expression Omnibus (GEO) repository under accession GSE90634. A Life Sciences Reporting Summary is available.

51. Olsen, J.V. et al. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. Mol. Cell. Proteomics 8, 2759–2769 (2009).
52. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372 (2008).
53. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).
54. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).
55. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).
56. Anders, S., Pyl, P.T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).
57. Machanick, P. & Bailey, T.L. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics 27, 1696–1697 (2011).
58. Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R. & Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 20, 110–121 (2010).
59. Siepel, A. et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15, 1034–1050 (2005).
60. Lovén, J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 320–334 (2013).
61. Whyte, W.A. et al. Master transcription factors and Mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319 (2013).
62. Chepelev, I., Wei, G., Wangsa, D., Tang, Q. & Zhao, K. Characterization of genome-wide enhancer–promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. Cell Res. 22, 490–503 (2012).
63. Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucleic Acids Res. 43, D1049–D1056 (2015).
64. Feng, J. et al. GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. Bioinformatics 28, 2782–2788 (2012).
65. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).
66. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550 (2005).
67. 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).
68. Knight, P.A. & Ruiz, D. A fast algorithm for matrix balancing. IMA J. Numer. Anal. 33, 1029–1047 (2013).
69. Filipova, D., Patro, R., Duggal, G. & Kingsford, C. Identification of alternative topological domains in chromatin. Algorithms Mol. Biol. 9, 14 (2014).
Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

1. Sample size

Describe how sample size was determined.

No sample size calculation was performed; sample sizes were determined based on number of available cell lines and all experiments were performed in >/= 2 biological replicates.

2. Data exclusions

Describe any data exclusions.

No data were excluded from this analysis.

3. Replication

Describe whether the experimental findings were reliably reproduced.

All attempts at replication were successful in each experiment reported in the study.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Allocation was not relevant to this study.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Blinding was not relevant to this study.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

☑ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)

☑ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.

☐ A statement indicating how many times each experiment was replicated

☐ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)

☐ A description of any assumptions or corrections, such as an adjustment for multiple comparisons

☐ The test results (e.g. p values) given as exact values whenever possible and with confidence intervals noted

☐ A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)

☐ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.
## Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

| All codes and software used to analyze the data are described in the Methods section in full detail. |

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The Nature Methods guidance for providing algorithms and software for publication may be useful for any submission.

## Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

| All unique materials used are readily available from the authors or from standard commercial sources (ATCC, Abcam, SCBT, Cell Signaling, etc). |

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

| All antibodies used in this study are described in full detail in the Methods section and in Supplementary Table 3. |

10. Eukaryotic cell lines

   a. State the source of each eukaryotic cell line used.
   b. Describe the method of cell line authentication used.
   c. Report whether the cell lines were tested for mycoplasma contamination.
   d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

| All cell lines were confirmed mycoplasma-negative. |

## Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

| N/A |

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

| N/A |