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

Robert T. Nakayama[1,2,9], **John L. Pulice**[1,3,9], **Alfredo M. Valencia**[1,4,9], **Matthew J. McBride**[1,4], **Zachary M. McKenzie**[1], **Mark A. Gillespie**[5], **Wai Lim Ku**[6], **Mingxiang Teng**[7], **Kairong Cui**[6], **Robert T. Williams**[1], **Seth H. Cassel**[1,8], **He Qing**[1], **Christian J. Widmer**[1], **George D. Demetri**[2], **Rafael A. Irizarry**[7], **Keji Zhao**[6], **JeffRanish**[5], and **Cigall Kadoch**[1,3,*]

[1]Department of Pediatric Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA

[2]Ludwig Center at Dana-Farber/Harvard and Center for Sarcoma and Bone Oncology, Department of Medical Oncology, Harvard Medical School, Boston, MA, USA

[3]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[4]Program in Chemical Biology, Harvard University, Cambridge, MA, USA

[5]Institute for Systems Biology, Seattle, WA, USA

[6]Systems Biology Center, NHLBI, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD, USA

[7]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA

[8]Medical Scientist Training Program, Harvard Medical School, Boston, MA, USA

Abstract

Perturbations to mammalian SWI/SNF (BAF) complexes contribute to over 20% of human cancers, with driving roles first identified in malignant rhabdoid tumor (MRT), 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

*Corresponding Author: Cigall Kadoch, Ph.D., Assistant Professor, Dept. of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 450 Brookline Avenue; Dana 620, Boston, MA 02215, Phone: 617-632-3789, Cigall_Kadoch@dfci.harvard.edu.
[9]These authors contributed equally to this work.
ORCID: 0000-0002-4058-5985

on chromatin, absent significant changes in intra-complex 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 by BAF and PBAF complexes at enhancers and promoters, respectively, suggesting distinct functions of each complex which are perturbed upon BAF47 loss. Our results demonstrate collaborative mechanisms of mSWI/SNF-mediated gene activation, identifying functions that are coopted or abated to drive human cancers and developmental disorders.

Introduction

Chromatin regulation is critical for the maintenance of timely and appropriate gene expression, with epigenetic regulators playing key roles both in normal development and oncogenesis[1]. Chromatin remodeling complexes regulate DNA accessibility via alteration of nucleosome positioning and/or occupancy in an ATP-dependent manner[2]. One of the most well-characterized chromatin remodeling complexes is the mammalian SWI/SNF (BAF) complex, first identified in yeast[3] and subsequently characterized in Drosophila[4] and mammals[5]. Specialized BAF complex subunit configurations have been demonstrated to be critical in pluripotency[6,7], neural differentiation[8], as well as the development of several other adult tissue types[9]. Various epigenetic modifiers and chromatin remodeling complexes, including BAF complexes, have been shown to localize to active promoters and enhancers in ES cells[10,11], however, the roles for specific complexes in the establishment and maintenance of promoter and enhancer 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[12], in which the *SMARCB1* gene, which encodes the core BAF complex subunit BAF47 (also known as INI1, hSNF5), undergoes biallelic inactivation in ~98% of MRT cases[12,13]. BAF47 loss has since been shown to be the hallmark genetic alteration in additional cancer types including atypical teratoid/ rhabdoid tumors (AT/RT; ~100%)[14] and epithelioid sarcomas (EpS; >90%)[15]. Additionally, BAF47 mutations are implicated in the development of meningiomas[16], schwannomatosis[17], and Coffin-Siris syndrome[18]. MRTs are genomically stable sarcomas with extremely low mutational burden[19,20], and conditional biallelic inactivation of *Smarcb1* in a mouse model leads to the most rapid tumorigenesis documented for a single gene deletion, with median onset at 11 weeks[21]. Recent exome sequencing studies have demonstrated that genes encoding BAF complex subunits are mutated in >20% of human cancers[22], including gain-of-function perturbations such as the SS18-SSX oncogenic fusion hallmark to ~100% of synovial sarcomas[23]. The clear link between *SMARCB1* deletion and MRT suggests MRT as a uniquely powerful disease setting in which to understand BAF complex mutations across human cancer.

Dynamic opposition between BAF complexes and polycomb repressive complexes was first demonstrated genetically in Drosophila[4,24], and has since been shown to govern critical processes in both normal development and disease[25]. In mammals, this opposition has been suggested to occur in a locus-specific manner[26,27], and in a global regulatory manner

through upregulation of *Ezh*[28]. In synovial sarcoma, the oncogenic SS18-SSX fusion has been demonstrated 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 play particularly critical roles in maintaining bivalent gene promoters, marked dually by H3K4me3 and H3K27me3[30,31]. The bivalent state of a given locus is maintained by a balance between activating trithorax (Trx) proteins (such as MLLs) and repressive polycomb group (PcG) proteins (such as PRC2 components)[32], with several BAF complex subunits also categorized as trithorax-group proteins[24]. Loss of PRC2 leads to activation of tissue-specific bivalent promoters but not monovalent promoters marked by H3K27me3 only[33]. Pre-clinical studies (and now early-stage clinical studies) using EZH2 inhibitors in BAF47-deficient sarcoma model systems have begun to show promise[34], suggesting this dynamic opposition as 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 driven by loss of BAF47. 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 find that rescue of BAF47 targets BAF complexes to bivalent promoters, enabling opposition of polycomb-mediated repression to resolve bivalent promoters to activation. Finally, we demonstrate that the observed enhancer activation and resolution of bivalent promoters are collaborative with respect to gene expression, suggesting dual complementary roles for BAF47-mediated tumor suppression. These data suggest two defining functions of BAF complexes that can be singly or collaboratively perturbed in BAF complex-mutated cancers and developmental disorders.

Results

BAF47 loss decreases chromatin affinity of intact BAF complexes

BAF47 is a core BAF complex subunit that is stable in its association with the BRG1 ATPase subunit in over 2M urea treatment[23]. To determine the effect of BAF47 loss on the integrity and subunit stability of BAF complexes, we lentivirally infected the G401 MRT cell line with either 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). BAF47 rescue was accompanied by minimal changes to BAF complex-bound protein levels of most core subunits in anti-BRG1 immunoprecipitations (IPs) (Fig. 1b, Supplementary Fig. 1a). To confirm this, we generated *BAF47* $^{-/-}$ HEK293T cells using CRISPR/Cas9-mediated 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 demonstrated highly similar banding patterns in both conditions (Fig. 1d, Supplementary Fig. 1b). To complement this, we used low-stringency

anti-BRG1 affinity purification/ proteomic mass-spectrometry (Fig. 1e, Supplementary Table 1). Peptide abundance corresponding to most BAF complex subunits were roughly equivalent in both conditions, with the exception of the BAF45A/D and BAF60C subunits, which exhibited 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 been conflicting, with some showing dissociation of BAF complexes[35,36] and others suggesting no changes to BAF complex assembly[29,37-39], likely due to differences in chromatin-bound protein purification methods and effects of super-stoichiometric protein abundance resulting from strong overexpression. Additionally, harsh denaturing detergents, such as sodium dodecyl sulfate (SDS) and sodium deoxycholate, used in some of these experiments can disrupt protein-protein complex interactions and/or reduce the antibody pulldown efficacy in solution (Supplementary Fig. 1e-f).

In order to examine changes to the biochemical stability and size of BAF complexes, we performed 10-30% glycerol gradient-based density sedimentation analyses in G401 nuclear extracts in both empty vector and BAF47 conditions (Fig. 1f-g, Supplementary Fig. 1c). BAF47 fully incorporates into BAF complexes upon re-expression (fractions 13-15), and subunits corresponding to both BAF and PBAF complexes shift up by approximately 1-2 fractions, in accordance with the expected gain in complex mass resulting from BAF47 and associated subunits (from Fig. 1e). Select subunits, including BRG1 and BAF60A, exhibit a greater spread across gradient fractions in the absence of BAF47, however BRG1-bound subunit stability was largely retained as demonstrated by urea desaturation 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 may have been predicted given its high degree of evolutionary conservation and its highly penetrant loss-of-function phenotype.

As BAF chromatin remodeling complexes contain several DNA- and histone-binding domains, we sought to determine if 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 containing either empty vector control or BAF47 (Fig. 1h, *left*, Supplementary Fig. 2a). We found that BAF47-deficient complexes in G401 cells dissociate from chromatin between 150-300 mM NaCl, while upon BAF47 rescue, subunits dissociate from chromatin between 500-1000 mM NaCl treatment (Fig. 1h, right, Fig. 1i, Supplementary Fig. 2b-c). As controls, we assessed PRC2 subunits EZH2 and SUZ12, 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 BAF47-knockout 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 chromatin occupancy and regulatory capacity.

BAF47 rescue drives a widespread gain in BAF complex occupancy

To examine the effect of BAF47 rescue on BAF complex targeting and gene regulation, we lentivirally infected two MRT cell lines, TTC1240 and G401, with empty vector or BAF47, as above (Fig. 2a), and performed chromatin immunoprecipitation of BAF complexes followed by sequencing (ChIP-seq) using antibodies to two core BAF complex subunits, BRG1 and BAF155. We observed a striking gain of genome-wide BAF complex occupancy upon rescue of BAF47 in TTC1240 cells (Fig. 2b-d, Supplementary Fig. 3a-e). We found that gained (BAF47-only) BAF complex sites (defined as shared BRG1-BAF155 sites) in TTC1240 cells are disproportionately localized to promoter-distal regions, as compared to conserved (empty-BAF47) BAF complex sites (Fig 2e, Supplementary Fig. 3f-g). Gained BAF complex sites were selectively enriched for unique motifs such as the AP-1 motif (Fig. 2f, Supplementary Fig. 3h). In addition, we examined the 46-way vertebrate PhyloP evolutionary conservation scores at all conserved and gained BAF complex sites upon BAF47 reintroduction in TTC1240 cells, and found that gained sites were much less evolutionarily conserved than conserved sites (Fig. 2g). This pattern was observed at both proximal and distal BAF complex sites (Supplementary Fig. 3i-j) and suggests that sites which lose BAF complex regulation upon BAF47 deletion in MRT are more recently evolved, implicating an evolutionarily recent cell of origin from which MRTs arise. 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 (4), EpS (4), and AT/RT (2) types, to BAF47 rescue (Supplementary Table 2). We found that all MRT and AT/RT cell lines assessed exhibited marked proliferative arrest upon BAF47 rescue, however this occurred in only one of the four EpS cell lines, with three EpS cell lines showing no significant proliferative arrest upon BAF47 rescue (Fig. 2h, 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, HS-ES-2M, and VA-ES-BJ cells, and observed a similar gain in BAF complex occupancy upon BAF47 rescue to promoter-distal sites, irrespective of the cell line used (Fig. 2i-k, 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, independent of sensitivity to BAF47-mediatedgrowth suppression.

BAF47 is required for BAF complex-mediated enhancer activation

To determine the effect of gained BAF complex occupancy on the histone landscape, we performed ChIP-seq studies for H3K4me3, H3K4me1, and H3K27ac marks. Notably, we found significant gains in H3K27ac and H3K4me1, but very minor changes in H3K4me3 levels (Fig. 3a), suggesting that gained BAF complex occupancy predominantly determines both enhancer state and enhancer activation[40]. We find that this activation is 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 which retain BAF complex occupancy and activation irrespective 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 (PCC=0.82) over all TTC1240 BAF sites (in both Empty and BAF47 conditions), with a lower but strong correlation with H3K4me1 (PCC=0.56), likely due to relative antibody enrichment, while we observed minimal correlations with H3K4me3 (PCC=0.15) (Fig. 3e, Supplementary Fig. 7a-c). These results indicate that in addition to BAF complex- mediated enhancer state and activation, the levels of BAF complex occupancy directly correspond 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 if BAF47 activates only enhancers or super-enhancers. We found that significant enhancer activation upon rescue of BAF47 occurs at both typical enhancers (12875) and super-enhancers (283), and that both typical and super-enhancers are retained in the absence of BAF47 (Fig. 3g, Supplementary Fig. 7e-f), in contrast to previous reports[35].

We then performed chromosome conformation capture followed by massively parallel sequencing (Hi-C) in VA-ES-BJ cells to determine if BAF47 rescue affects global chromatin topology independent of proliferative arrest, as this has been a suggested mechanism of oncogenesis in other cancers[41,42]. While we found that BAF47 had no significant impact on global genome architecture (Supplementary Fig. 8a-e), we did identify new promoter-enhancer interactions at gained enhancers such as *CDKN1A* (Supplementary Fig. 8f), likely due to downstream effects of enhancer activation. These results collectively suggest that BAF47 plays 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.

Enhancer activation is mediated by BAF not PBAF complexes

Mammalian SWI/SNF complexes exist in two different assemblies, canonical BAF complexes and PBAF complexes, defined by distinct subunits (Fig. 4a), with BAF47 as a shared core subunit in both. 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)[43] and BAF200/ARID2 (PBAF complex-specific) to determine how rescue of BAF47 influences each complex. We found that SS18 exhibits substantially more retargeting than BAF200, with SS18 retargeted in a similar manner as BRG1 and BAF155, while BAF200 exhibited only modest retargeting over all TTC1240+BAF47 BRG1-BAF155 shared sites (Fig. 4c). We find that SS18-marked BAF complexes exhibit a dramatic gain of occupancy at distal sites, whereas gains in BAF200-marked PBAF complexes are nearly entirely restricted to proximal sites (Fig. 4d-f, Supplementary Fig. 9a). Using log2 fold changes in BRG1, SS18, and BAF200 occupancy upon BAF47 reintroduction, we find that BRG1 and SS18 exhibit high correlation (PCC = 0.88), while BRG1 and BAF200 exhibit a substantially more modest correlation (PCC = 0.49) (Fig. 4g-h, Supplementary Fig. 9b). BRG1 and BAF200

exhibit stronger correlations at proximal (0.56) than distal (0.43) sites, whereas BRG1 and SS18 correlations are similar (proximal PCC=0.82, distal PCC=0.86), suggesting a greater role for retargeting of PBAF complexes to 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 and not PBAF complexes, as PBAF complexes are not significantly targeted to these sites (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 both mechanisms of global regulatory opposition[28] and locus-specific opposition (e.g. at the p16^{INK4A} locus)[26]. However, to date, BAF-polycomb complex opposition has not been studied at a genome-wide level. We performed ChIP-seq for SUZ12 (a core PRC2 subunit) and H3K27me3, and find that over all promoters, occupancy of BAF and PBAF complexes correlates with H3K4me3 occupancy as well as gene expression, whereas H3K27me3 and SUZ12 exhibit highest occupancy at non-expressed genes (Fig. 5a, Supplementary Fig. 10a-h). This suggests that BAF and PBAF complexes play 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 3022 (12.57%) genes categorized as bivalent (Fig. 5a, Supplementary Fig. 10i-j). GO term analysis of genes with bivalent promoters strongly enriches for genes involved in kidney and neural development, likely reflecting initiation of cell lineage-specific regulation (Fig. 5b). We also performed ChIP-seq analyses for SUZ12 and H3K27me3 in G401 cells, and similar to TTC1240 cells, found that 2470 (10.27%) genes are bivalent; interestingly, the large majority (1902) of these bivalent genes were shared between G401 and TTC1240 cell lines (Fig. 5c, 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 likely work through activating bivalent genes[33], we sought to determine whether BAF47 rescue resulted in altered bivalent promoter regulation. We find a significant increase in target gene occupancy upon rescue of BAF47 by both BAF and PBAF complexes (Fig. 5d, Supplementary Fig. 11a). Interestingly we find that while only 8.03% of conserved BRG1 target genes are bivalent, 30.13% of gained BRG1 target genes are bivalent in TTC1240 cells (Fig. 5e). This results in an increase in BAF complex occupancy at bivalent promoters from 29.2% to 68.6% upon rescue of BAF47, the greatest percent increase of any promoter category (Supplementary Fig. 11b). We next aimed to assess whether the number of bivalent genes is affected by rescue of BAF47. We find that 506 (16.7%) bivalent genes in the TTC1240-Empty setting are no longer bivalent in the BAF47 condition (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+BAF47, we find that rescue of BAF47 leads to a gain in BAF complex occupancy and a decrease in H3K27me3 occupancy, and that this is predominantly at promoter proximal sites (Supplementary Fig. 11c-e). Over all

bivalent promoters in TTC1240, we observed a significant decrease in H3K27me3 absent changes in H3K4me3 (Fig. 5g, Supplementary Fig. 11f), suggesting the regulation of bivalency is solely due 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, Supplementary Fig. 11f). Interestingly, log2 fold change values over proximal TTC1240 BRG1-BAF155 sites show opposition of BRG1 and H3K27me3 (PCC=-0.26) and SS18 and H3K27me3 (PCC=-0.24), but opposition is greater between BAF200 and H3K27me3 (PCC=-0.40) (Fig. 5i, Supplementary Fig. 11g-i). While we see similar results in G401 cells (Supplementary Fig. 12), opposition of polycomb-mediated repression is more modest in EpS cell lines, which contain fewer bivalent promoters (Supplementary Fig. 13), suggesting that the contribution of bivalency to oncogenesis may vary by tissue of origin. Our results demonstrate that BAF complexes, both BAF and PBAF complex assemblies, play a critical role in resolving bivalent promoters to activation in development, a process significantly impaired by BAF47 loss which leads to reformation of bivalency and repression of key developmental and lineage-specific differentiation genes.

Collaborative activation of bivalent promoters and enhancers

Having identified two mechanisms by which BAF47 affects gene activation, by both BAF and PBAF complexes, we wanted 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. 14a-b). We find that bivalent genes are overrepresented in the 2635 significantly-regulated genes in TTC1240 cells, with 20.80% marked as bivalent (Fig. 6a). We find that 13.99% of all bivalent genes are upregulated by BAF47, the largest proportion of any gene category, and that bivalent genes exhibit clear net upregulation as compared to H3K4me3-only genes (Fig. 6b), with these effects similarly observed in G401 (Supplementary Fig. 14c-d).

We identified a set of 642 genes that are significantly and concordantly regulated in both G401 and TTC1240 cells (Fig. 6c-d). GSEA and GO term analyses show that upregulated genes play critical roles in kidney development and epithelial-mesenchymal transition (EMT), likely suggesting key pathways implicated in BAF47-mediated tumor suppression (Fig. 6e-f, Supplementary Fig. 14e-h), which corresponds with mechanisms of MET in sarcomagenesis[44] along with roles for BAF47 in EMT of pancreatic cancer[45]. We do not observed ownregulation of *EZH2* and upregulation of *CDKN2A* genesin both cell lines, as previously suggested (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 BAF47 loss-driven sarcomagenesis.

We 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 a clear overrepresentation of bivalent genes among those with greater numbers of gained distal BAF complex sites(Fig.

6g, Supplementary Fig. 14k), and these genes correspond to the upregulation shown previously (Fig. 3f). This collaborative activation is exemplified at the *CTGF* and *FN1* loci, at which these activated genes gain (1) promoter occupancy that resolves bivalency to activation, and (2) numerous distal BAF complex-activated enhancers (Fig. 6h-i). We demonstrate here that these enhancer and promoter regulatory functions are hallmark to BAF47-mediated gene regulation and tumor suppression, with a critical collaborative role for these two distinct BAF complex functions lost in BAF47-deficient cancers.

Discussion

Mutations in the genes encoding BAF complexes are recurrent in over 20% of human cancers, as well as 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 significantly impairing in-solution assembly or subunit stability (Fig. 7a). We show that BAF complexes play 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 contradicting studies suggesting either little change in complex composition[29,37-39] or dramatic loss of stability[35,36] upon BAF47 loss. We do not find global changes in BAF complex composition or assembly, in agreement with recent yeast SWI/SNF studies[46]. The structural integrity of BAF47-deficient residual BAF complexes shown here indeed supports previous observations of BRG1 dependency in MRT cell lines[39], and suggests therapeutic avenues for targeting these intact residual BAF complexes in MRT. Our results suggest 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 complexes 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 trithorax-group protein through opposition of polycomb-mediated repression. This study substantiates how BAF complexes, within the existing dynamics of trithorax (Trx) or polycomb (PcG) proteins, drive resolution of bivalent promoters toward activation or repression, respectively[32]. These results begin to explain the observed therapeutic 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. We demonstrate a broad-spanning 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 that BAF47 drives a widespread gain of BAF complex occupancy that mediates enhancer state as well as enhancer activation. Our results suggest the BAF complex plays a

pioneering role in governing enhancer state and activation, such that BAF47 restores the functional ability of the BAF complex to bind to and activate these sites. As such, recruitment by transcription factors or regulators fails to create BAF complex targeting and activation in the absence of BAF47, as previously shown with *Arid1a* loss promoting regeneration[47]. Enhancers have co-occupancy of numerous chromatin remodelers/regulators[10,48], and our results suggest a dynamic regulation of enhancers by both activating (BAF complex) and repressive (NuRD complex)[49] remodelers that may be occurring similar to HATs and HDACs[50]. Interruption of this regulatory dynamic via BAF47 loss could 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 observed 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 widespread implications for the outsized contribution of BAF complex aberrations in human malignancy and developmental disorders.

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 (G401, G402, A204 and VA-ES-BJ), and three were from RIKEN (HS-ES-1, HS-ES-2R and HS-ES-2M). TTC1240, TM87-16 and STM91-01 were a generous gift from Prof. Timothy J. Triche (Children's Hospital Los Angeles, Los Angeles, CA), BT12 and BT16 were from Dr. Peter Houghton (The University of Texas Health Science Center at San Antonio, TX), NEPS was from Dr. Hiroyuki Kawashima (Niigata University Graduate School of Medical and Dental Sciences, Niigata), FU-EPS-1 and SFT-8606 were from Prof. Hiroshi Iwasaki (Fukuoka University, Fukuoka), YCUS-5 was from Dr. Hiroaki Gotoh (Kanagawa Children's Medical Center, Kanagawa), and ESX was from Dr. Tomohide Tsukahara (Sapporo Medical University, Sapporo). Cell lines were cultured either in DMEM/F12, RPMI1640, or DMEM medium (Gibco, Grand Island, NY, USA), supplemented with 10% fetal bovine serum, 1% Glutamax (Gibco) and 1% Penicillin-Streptomycin (Gibco).

Vector/cloning information

BAF47 (*SMARCB1*) constitutive expression in MRT and EpS cell lines was achieved using lentiviral infection of an EF1alpha-driven expression vector (modified from Clonetech, dual Promoter EF-1a-MCS-PGK-Blast), selected with blasticidin (1ug/ul).

Lentiviral Generation

Lentivirus was produced by PEI (Polysciences Inc.) transfection of HEK293T LentiX cells (Clontech) with gene delivery vector co-transfected with packaging vectors pspax2 and pMD2.G as previously described[23]. Supernatants were harvested 72h post-transfection and centrifuged at 20,000 rpm for 2h at 4°C. Virus containing pellets were resuspended in PBS and placed on cells dropwise. Selection of lentivirally-infected cells was achieved with either blasticydin or puromycin, both used at 2μg/ml.

Nuclear extract

Nuclear extract (NE) preparation and immunoprecipitation (IP) studies were performed as described previously in Ho et al. (2009). Briefly, the trypsinized cells were incubated in Buffer A (25 mM HEPES pH 7.6, 5 mM $MgCl_2$, 25 mM KCl, 0.05 mM EDTA, 10% glycerol and 0.1% NP40 with protease inhibitor (Roche), 1 mM DTT and 1 mM phenylmethylsulfonyl fluoride (PMSF)) for 10 minutes and the pellets were resuspended in 600 μl of Buffer C (10 mM HEPES pH 7.6, 3 mM $MgCl_2$, 100 mM KCl, 0.5 mM EDTA and 10% glycerol with protease inhibitor, 1 mM DTT and 1 mM PMSF) with 67 μl of 3 M $(NH_4)_2SO_4$ for 20 minutes. The lysates were spun down using ultracentrifuge at 10,000 rpm at 4°C for 10 minutes. Nuclear extracts were precipitated with 200 μg of $(NH_4)_2SO_4$ on ice for 20 minutes and finally purified as pellets by ultracentrifugation at 10,000 rpm at 4°C for 10 minutes. The pellets were resuspended in IP Buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 1 mM EDTA and 1% Triton-X100 with protease inhibitor, 1 mM DTT and 1 mM PMSF) for the subsequent experiments. For the RIPA and IP Buffer comparison experiments performed, nuclear pellets were resuspended in either RIPA buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 1% Nonidet P-40, 0.5% Sodium Deoxycholate, and 0.1% SDS, complete 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 manufacture'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. Then each sample is incubated with Dynabeads (Thermo Scientific) for two hours. 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 Ini1 CRISPR/Cas9 KO Plasmid and Ini1 HDR Plasmid (Santa Cruz Biotechnology sc-423027; sc-423027-HDR) in HEK293T Lenti-X cells (Clonetech) following the manufacturer's protocol. Specifically, five 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 hours and GFP+/RFP+ cells expressing both the KO/HDR plasmids were

single-cell sorted through FACS (fluorescence-activated cell sorting). Single-cell clones were expanded further and screened through immunoblot for identification of successful knockouts.

2D/LC/MS IP Proteomics

4 samples for each cell type were prepared (2 IgG controls and 2 replicates of anti-BRG1IPs) for mass spec analysis (label free quantitation). Eluted proteins from each condition were processed simultaneously to reduce sample variability. Proteins were reduced, alkylated, and digested with trypsin, and desalted using offline C18 reversed-phase chromatography. Purified peptides were separated by online C18 reversed-phase chromatography then analyzed with a top10 CID data-dependent manner using an LTQ-Velos mass spectrometer[51].

2D/LC/MS IP Proteomics Data Processing and Analysis

Data analysis was performed with MaxQuant software[52], supported by a database search engine for peptide identification (human IPI). Label-free quantitation 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 ml of 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, Brea, CA, USA). Tubes were centrifuged in an SW40 rotor at 4°C for 16 hr at 40,000 rpm. Fractions (0.5 ml) were collected and used in analyses.

Urea Denaturation Studies

Nuclear extracts (150 μg) were subjected to partial urea denaturation, ranging from 0.25 to 5.0 M urea (in IP buffer), for 15 min at room temperature (RT) prior to anti-BRG1 IP. The co-precipitated proteins were analyzed by immunoblot.

Differential Salt Extraction

Cell types were grown under standard conditions and following collection of 5×10^7 cells, suspended in 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 minutes, and centrifuged. Supernatant was collected and pellet was suspended in elution buffer with 150 mM NaCl. This process was repeated sequentially with increasing concentration of NaCl to collect 0, 150, 300, 500, and 1000 mM NaCl soluble fractions. Each fraction, including total fraction (5×10^6 cells in elution buffer) and pellet fraction (material remaining following 1000 mM NaCl extraction), 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, Lincoln, NE, USA).

ChIP-seq Data Collection

Cells were harvested following 48-hour exposure to the lenti-virus and 5-day selection with 10 μg/ml of Blasticidin for chromatin immunoprecipitation (ChIP) experiments. ChIP experiments were performed per standard protocols (Millipore, Billerica, MA) 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-hour exposure to the lenti-virus and either 1-day (day 3 post-infection) or 5-day (day 7 post-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 produced using the RNeasy Mini Kit (Qiagen).

Proliferation experiments

20,000 cells were plated following 48-hour exposure to the lenti-virus and 48-hour 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, Brea, CA) on Day 1, 3, 5 and 7.

Library Prep and Sequencing for ChIP-seq and RNA-seq

All library prep and sequencing (75bp single end on Illuminia 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[53] version 2.1.0 with parameters –k 1. RNA-seq reads were mapped to the human reference genome (hg19) using STAR[54] 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[55] version 2.1.0 at q=1e-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 condition, gained sites were determined using sites only in the BAF47 condition. Venn diagrams were generated using the R statistical package, using the minimum number of overlapping regions for resolving multiple peak overlaps.

Metagene read densities were generated using HTSeq[56], with fragment length extended to 200bp to account for the average 200bp fragment size selected in sonication, using the center of peak calls from MACS2. Total read counts for each region were normalized the number of mapped reads to give reads per million mapped reads. Metagene plots were generated using average read densities across all sites indicated for each condition. Heatmaps 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. Heatmaps were visualized using Python matplotlib with a midpoint of 0.5 reads per million for the heatmap 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 70777 BRG1-BAF155 sites in TTC1240. 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, 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, 500bp core sequences centered on peak centers were submitted to MEME-ChIP analysis[57]. Conservation scores were calculated using bedtools map –o mean to generate the average PhyloP score for each 500bp core sequence as in motif analysis, using the PhyloP 46-way vertebrate conservation score from UCSC[58,59]. Determination of super-enhancers was performed using ROSE[60,61] with a union peak set of H3K27ac in empty and BAF47 conditions in TTC1240, 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 function with hg19 refFlat TSS annotation, with small RNA genes (MIR and SNO) removed. Target genes were determined using TSS sites within 2kb of a peak. For visualization of promoters, the same promoter set used for target gene analysis was used, except identical promoters with multiple annotations were only included once. Number of genes counts each unique gene annotation once, whereas number of promoters counts all varied TSS sites that are annotated as such, so the number of bivalent genes and bivalent promoters are distinct numbers reflecting the same set of sites. Target genes of distal BAF sites were determined using the distance to TSS calculations above, filtered for peaks >2k and 50kb from their assigned

target gene, as most enhancer-promoter interactions have been shown to occur within a distance of 50kb[62]. Genes were then binned by the number of distal BAF sites in each condition, and this was used for RNA-seq and promoter state analyses. GO Term analysis was performed on the target gene sets using biological processes annotation[63], with a significance threshold of 1e-3.

RNA-seq Data Analysis

RPKM values for samples were generated using GFold[64] version 1.1.0. All error bars represent Mean±SEM. Significance was assessed using the R package DESeq2[65] using raw read counts generated with Rsubread feature Counts against the hg19 refFlat annotation. Significantly changing genes were assessed with a Bonferri-corrected p-value of less than 1e-3 and a two-fold gene expression change ($|\log2FC|>1$) to determine set of significantly changing genes. GSEA was performed using the GSEA Preranked function of the JAVA program (http://www.broadinstitute.org/gsea) as described previously[66]. 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 RPKM value for four samples >= 1), averaging replicates of each condition, then doing a log2 fold change comparison with a pseudocount of 1 in each condition, i.e. $\log2((RPKM^{BAF47} + 1)/(RPKM^{Empty} + 1))$. Log2 fold change values for RNA-seq in figures were identical to those used for GSEA analysis except non-expressed genes were included in the analysis. Two-tailed t tests were used to determine significance of difference for each.

RNA-seq tracks were generated using bedtools genomecov –split –scale with the mapped read count to generate tracks normalized per million mapped reads. All RNA-seq tracks visualized are Day 7 post-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 before downstream regulation to the extent possible.

Hi-C Experimental Method

Cells were cross-linked with 1% formaldehyde for 10min at room temperature. The Hi-C assay was performed as previously described[67] with the following modifications. Following cross-linking, the cells were lysed and digested with CviQ I + CviA II + Bfa I for 30min. Enzymes digested DNA ends were repaired and labeled with biotin-14-dATP with Klenow enzyme (large fragment). The proximity-based ligation of chromatin ends was performed using T4 DNA Ligase for overnight at 16°C. DNA was reverse cross-linked at 65°C for 6 hours 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 300-500bps 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 (PETs) from Hi-C libraries were mapped to the human reference genome (hg18) using bowtie2[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.

Hi-C Filtering and Heatmap Analysis

The interaction matrices were binned using a binsize=50 kb, binstep=5kb, binmode=mean. First, we generated the matrices from the Hi-C interaction data using binstep of 5kb. Second, the self-circularized restriction fragments were filtered by setting the diagonal elements of the matrices to be zero. Third, we removed the rows/cols of the matrices if the sum of elements in the rows/cols were zero. 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 be zero. Fifth, the matrices were balanced according to the KR normalization method[68] which was similar to the study by Rao *et al.*[67]. Sixth, we recovered the interaction matrices with binsize=50kb by combining every 10 bins into one bin with binmode=mean. The differential heatmap was the subtraction between the matrices in the two conditions.

Analysis for Topological Associating Domains

Each chromosome was separated into 50kb bins and interaction matrix of each chromosome was generated for Hi-C data. The interaction matrix was normalized by KR normalization[68]. The normalized interaction matrix was used as input for identifying TAD by Armatus[69].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell. 2012; 150:12–27. [PubMed: 22770212]

2. Clapier CR, Cairns BR. The biology of chromatin remodeling complexes. Annu Rev Biochem. 2009; 78:273–304. [PubMed: 19355820]

3. Peterson CL, Herskowitz I. Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell. 1992; 68:573–83. [PubMed: 1339306]

4. Tamkun JW, et al. brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell. 1992; 68:561–72. [PubMed: 1346755]

5. Kwon H, Imbalzano AN, Khavari PA, Kingston RE, Green MR. Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. Nature. 1994; 370:477–81. [PubMed: 8047169]

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 U S A. 2009; 106:5181–6. [PubMed: 19279220]

7. Singhal N, et al. Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. Cell. 2010; 141:943–55. [PubMed: 20550931]

8. Lessard J, et al. An essential switch in subunit composition of a chromatin remodeling complex during neural development. Neuron. 2007; 55:201–15. [PubMed: 17640523]

9. Ho L, Crabtree GR. Chromatin remodelling during development. Nature. 2010; 463:474–84. [PubMed: 20110991]

10. Hnisz D, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013; 155:934–47. [PubMed: 24119843]

11. de Dieuleveult M, et al. Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. Nature. 2016; 530:113–6. [PubMed: 26814966]

12. Versteege I, et al. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. Nature. 1998; 394:203–6. [PubMed: 9671307]

13. Chun HJ, et al. Genome-Wide Profiles of Extra-cranial Malignant Rhabdoid Tumors Reveal Heterogeneity and Dysregulated Developmental Pathways. Cancer Cell. 2016; 29:394–406. [PubMed: 26977886]

14. Biegel JA, et al. Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. Cancer Res. 1999; 59:74–9. [PubMed: 9892189]

15. Modena P, et al. SMARCB1/INI1 tumor suppressor gene is frequently inactivated in epithelioid sarcomas. Cancer Res. 2005; 65:4012–9. [PubMed: 15899790]

16. Christiaans I, et al. Germline SMARCB1 mutation and somatic NF2 mutations in familial multiple meningiomas. J Med Genet. 2011; 48:93–7. [PubMed: 20930055]

17. Hulsebos TJ, et al. Germline mutation of INI1/SMARCB1 in familial schwannomatosis. Am J Hum Genet. 2007; 80:805–10. [PubMed: 17357086]

18. Tsurusaki Y, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet. 2012; 44:376–8. [PubMed: 22426308]

19. McKenna ES, et al. Loss of the epigenetic tumor suppressor SNF5 leads to cancer without genomic instability. Mol Cell Biol. 2008; 28:6223–33. [PubMed: 18710953]

20. Lawrence MS, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013; 499:214–8. [PubMed: 23770567]

21. Roberts CW, Leroux MM, Fleming MD, Orkin SH. Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. Cancer Cell. 2002; 2:415–25. [PubMed: 12450796]

22. Kadoch C, et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nat Genet. 2013; 45:592–601. [PubMed: 23644491]

References for Methods

23. Kadoch C, Crabtree GR. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. Cell. 2013; 153:71–85. [PubMed: 23540691]

24. Kennison JA, Tamkun JW. Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila. Proc Natl Acad Sci U S A. 1988; 85:8136–40. [PubMed: 3141923]

25. Kadoch C, Copeland RA, Keilhack H. PRC2 and SWI/SNF Chromatin Remodeling Complexes in Health and Disease. Biochemistry. 2016; 55:1600–14. [PubMed: 26836503]

26. Kia SK, Gorski MM, Giannakopoulos S, Verrijzer CP. SWI/SNF mediates polycomb eviction and epigenetic reprogramming of the INK4b-ARF-INK4a locus. Mol Cell Biol. 2008; 28:3457–64. [PubMed: 18332116]

27. Ho L, et al. esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. Nat Cell Biol. 2011; 13:903–13. [PubMed: 21785422]

28. Wilson BG, et al. Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. Cancer Cell. 2010; 18:316–28. [PubMed: 20951942]

29. Kadoch C, et al. Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states. Nat Genet. 2017; 49:213–222. [PubMed: 27941796]

30. Bernstein BE, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006; 125:315–26. [PubMed: 16630819]

31. Mikkelsen TS, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature. 2007; 448:553–60. [PubMed: 17603471]

32. Voigt P, Tee WW, Reinberg D. A double take on bivalent promoters. Genes Dev. 2013; 27:1318–38. [PubMed: 23788621]

33. Jadhav U, et al. Acquired Tissue-Specific Promoter Bivalency Is a Basis for PRC2 Necessity in Adult Cells. Cell. 2016; 165:1389–400. [PubMed: 27212235]

34. Knutson SK, et al. Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. Proc Natl Acad Sci U S A. 2013; 110:7922–7. [PubMed: 23620515]

35. Wang X, et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. Nat Genet. 2017; 49:289–295. [PubMed: 27941797]

36. Wei D, et al. SNF5/INI1 deficiency redefines chromatin remodeling complex composition during tumor development. Mol Cancer Res. 2014; 12:1574–85. [PubMed: 25009291]

37. Doan DN, 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. 2004; 23:3462–73. [PubMed: 14990991]

38. Jamshidi F, et al. The genomic landscape of epithelioid sarcoma cell lines and tumours. J Pathol. 2016; 238:63–73. [PubMed: 26365879]

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. 2009; 69:8094–101. [PubMed: 19789351]

40. Creyghton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A. 2010; 107:21931–6. [PubMed: 21106759]

41. Hnisz D, et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science. 2016; 351:1454–8. [PubMed: 26940867]

42. Flavahan WA, et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature. 2016; 529:110–4. [PubMed: 26700815]

43. Middeljans E, et al. SS18 together with animal-specific factors defines human BAF-type SWI/SNF complexes. PLoS One. 2012; 7:e33834. [PubMed: 22442726]

44. Yang J, et al. Mesenchymal to epithelial transition in sarcomas. Eur J Cancer. 2014; 50:593–601. [PubMed: 24291235]

45. Genovese G, et al. Synthetic vulnerabilities of mesenchymal subpopulations in pancreatic cancer. Nature. 2017; 542:362–366. [PubMed: 28178232]

46. Sen P, et al. Loss of Snf5 Induces Formation of an Aberrant SWI/SNF Complex. Cell Rep. 2017; 18:2135–2147. [PubMed: 28249160]

47. Sun X, et al. Suppression of the SWI/SNF Component Arid1a Promotes Mammalian Regeneration. Cell Stem Cell. 2016; 18:456–66. [PubMed: 27044474]

48. Morris SA, et al. Overlapping chromatin-remodeling systems collaborate genome wide at dynamic chromatin transitions. Nat Struct Mol Biol. 2014; 21:73–81. [PubMed: 24317492]

49. Yildirim O, et al. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell. 2011; 147:1498–510. [PubMed: 22196727]

50. Wang Z, et al. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell. 2009; 138:1019–31. [PubMed: 19698979]

51. Olsen JV, et al. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. Mol Cell Proteomics. 2009; 8:2759–69. [PubMed: 19828875]

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. 2008; 26:1367–72. [PubMed: 19029910]

53. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9:357–9. [PubMed: 22388286]

54. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29:15–21. [PubMed: 23104886]

55. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]

56. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015; 31:166–9. [PubMed: 25260700]

57. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011; 27:1696–7. [PubMed: 21486936]

58. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010; 20:110–21. [PubMed: 19858363]

59. Siepel A, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005; 15:1034–50. [PubMed: 16024819]

60. Loven J, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013; 153:320–34. [PubMed: 23582323]

61. Whyte WA, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013; 153:307–19. [PubMed: 23582322]

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. 2012; 22:490–503. [PubMed: 22270183]

63. Gene Ontology, C. Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015; 43:D1049–56. [PubMed: 25428369]

64. Feng J, et al. GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. Bioinformatics. 2012; 28:2782–8. [PubMed: 22923299]

65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. [PubMed: 25516281]

66. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–50. [PubMed: 16199517]

67. Rao SS, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014; 159:1665–80. [PubMed: 25497547]

68. Knight PA. A fast algorithm fo rmatrix balancing. IMA J Numer Anal. 2012

69. Filippova D, Patro R, Duggal G, Kingsford C. Identification of alternative topological domains in chromatin. Algorithms Mol Biol. 2014; 9:14. [PubMed: 24868242]
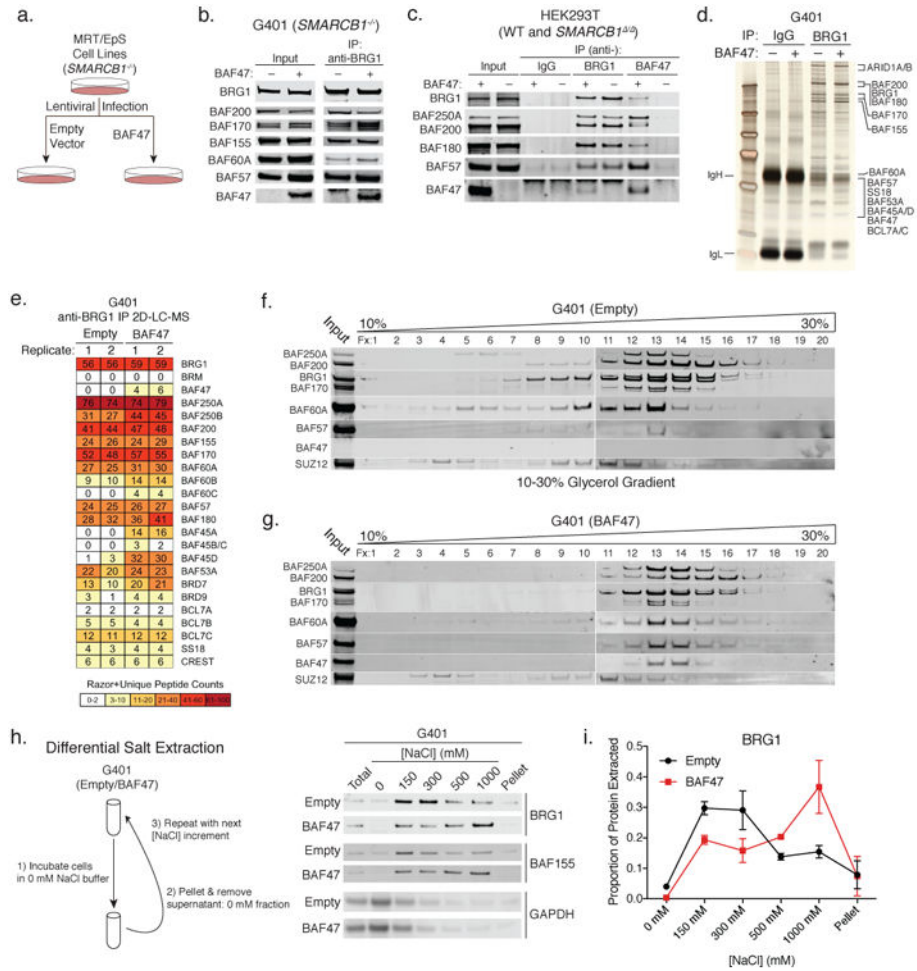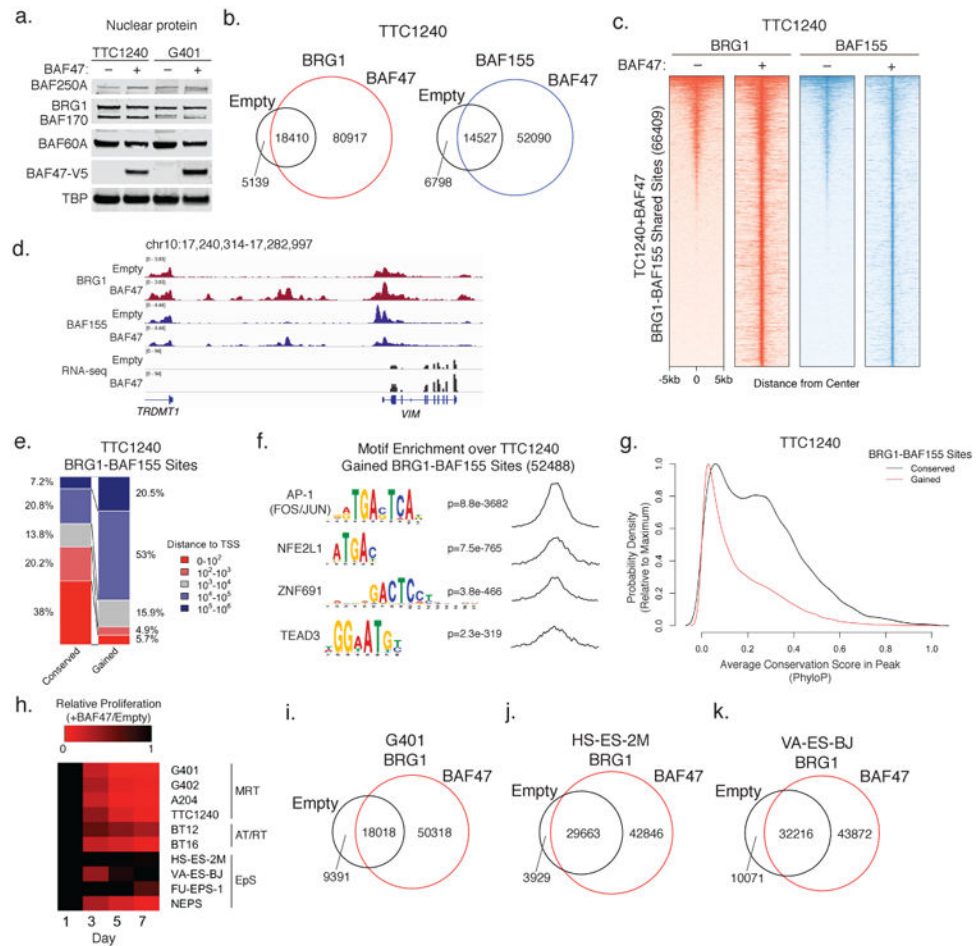
Figure 1.

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 vector and BAF47 conditions. (c) Nuclear extract input and IPs for IgG, BRG1, and BAF47 in control and BAF47 / (knockout) HEK293T cells. (d) Silver stain analysis of control IgG and anti-BRG1 IPs in G401 empty vector and BAF47 conditions. (e) anti-BRG1 IP-mass spectrometry proteomics in G401 empty vector and BAF47 conditions for BAF complex subunits. (f-g) Density sedimentation analyses using 10-30% glycerol gradients (10m; 0.5ml/ fx) on nuclear extracts from G401 MRT cells in (f) the empty vector control and (g) BAF47 conditions. (h) (left) Schematic for differential salt extraction experiments in G401 empty vector or BAF47 conditions; (right) Immunoblot analysis of BAF complex subunits in differential salt extraction experiments. (i) Relative densitometry from differential salt extraction demonstrates gained stability of BAF complexes on chromatin in the BAF47 condition. Error bars = mean ± SEM for n=2 biological replicates.

Figure 2.
Rescue of BAF47 drives a genome-wide gain in BAF complex chromatin occupancy (a) Input blot for TTC1240 and G401 cell nuclear extracts in empty vector and BAF47 conditions. (b) Venn diagram of (left) BRG1 and (right) BAF155 peaks in empty vector and BAF47 conditions in TTC1240 cells. (c) Heatmaps of BRG1 and BAF155 occupancy in TTC1240 empty vector and BAF47 conditions over all BRG1-BAF155 shared sites in the TTC1240+BAF47 condition. (d) Example BRG1, BAF155, and RNA-seq tracks at CDKN1A enhancers in TTC1240 cells. (e) Distance to closest transcription start site (TSS) for conserved (empty-BAF47) and gained (BAF47-only) BRG1-BAF155 sites in TTC1240 cells. (f) Centrimo plots for top four centrally enriched motifs at gained BRG1-BAF155 sites in TTC1240 cells. (g) Average sequence conservation (PhyloP) of conserved and gained BRG1-BAF155 sites. (h) Proliferation analyses of MRT, EpS, and AT/RT cell lines; values shown are relative proliferation between BAF47 and empty vector conditions at noted days. (i-k) Venn diagrams of BRG1 peaks in empty vector and BAF47 conditions in (i) G401, (j) HS-ES-2M, and (k) VA-ES-BJ cell lines.
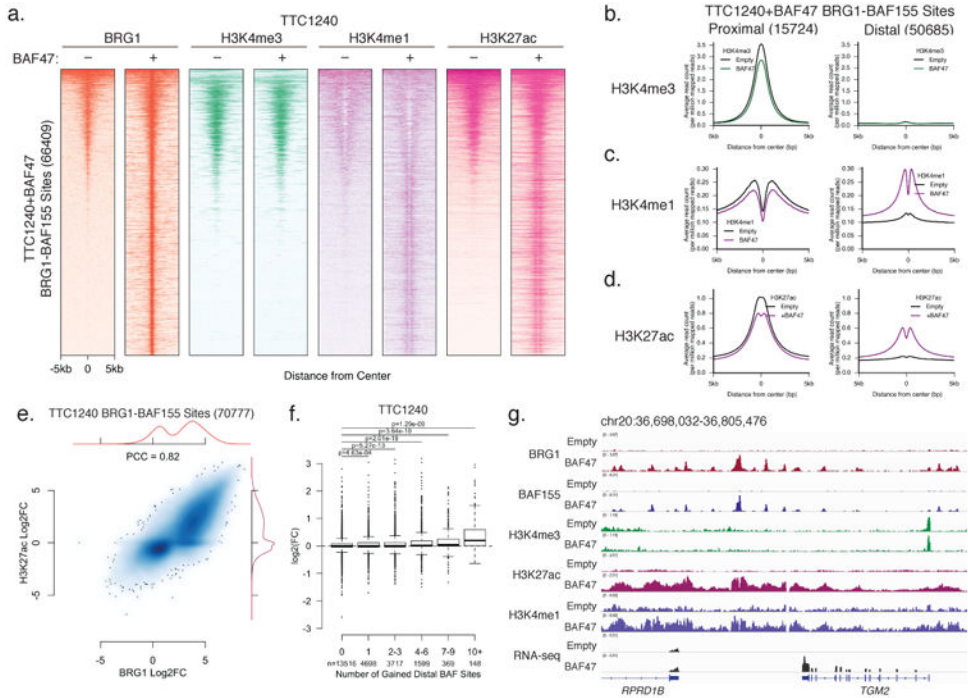
Figure 3.
Gain of BAF complex occupancy drives widespread enhancer activation. (a) Heatmaps of BRG1, H3K4me3, H3K4me1, and H3K27ac sites in TTC1240 empty vector and BAF47 conditions over all BRG1-BAF155 shared sites in the TTC1240+BAF47 condition. Heatmaps are ranked by BRG1 occupancy in empty condition. (b-d) Metagene plots of BRG1-BAF155 sites in the TTC1240+BAF47 condition split by (left) promoter-proximal (2kb from TSS), and (right) promoter-distal (>2kb from TSS) for (b) H3K4me3, (c) H3K4me1, and (d) H3K27ac occupancy. (e) Correlation plot of log2(fold change) for BRG1 and H3K27ac over all BRG1-BAF155 sites (70777) in empty or BAF47 conditions. (f) Gained promoter-distal BAF complex sites assigned to nearest gene (genes were categorized based on number of gained distal sites) versus log2(fold change) in expression. (g) Example tracks of BRG1, BAF155, H3K4me3, H3K4me1, H3K27ac, and RNA-seq at the TGM2 locus in TTC1240 cells.
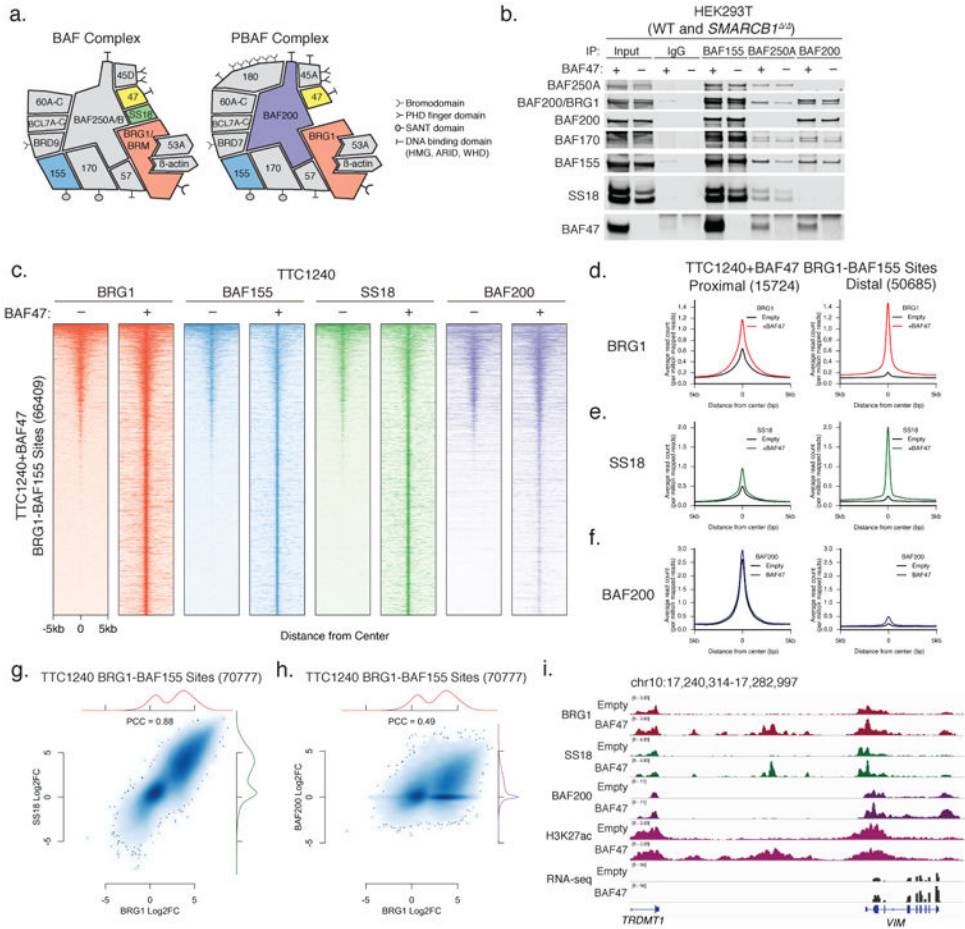
Figure 4.
Enhancer activation upon BAF47 rescue is specific to BAF but not PBAF complexes. (a) Schematic of BAF and PBAF complex subunits with subunits targeted for ChIP-seq in respective colors. (b) Input and IPs for IgG control, BAF155, BAF250A, and BAF200 from HEK293T nuclear extracts in naive and BAF47 −/− conditions. (c) Heatmap of BRG1, BAF155, SS18, and BAF200 in TTC1240 empty vector and BAF47 conditions over all BRG1-BAF155 shared sites in the TTC1240+BAF47 condition. (d-f) Metagene plots of BRG1-BAF155 shared sites in TTC1240+BAF47 split by (left) promoter-proximal (2kb from TSS), and (right) promoter-distal (>2kb from TSS) for (d) BRG1, (e) SS18, and (f) BAF200 occupancy. (g-h) Correlation plot of log2(fold change) for (g) BRG1 and SS18, and (h) BRG1 and BAF200, over all TTC1240 BRG1-BAF155 sites in empty or BAF47 conditions. (i) Example ChIP-seq tracks for BRG1, SS18, BAF200, H3K27ac, and RNA-seq at the VIM locus in TTC1240 cells.
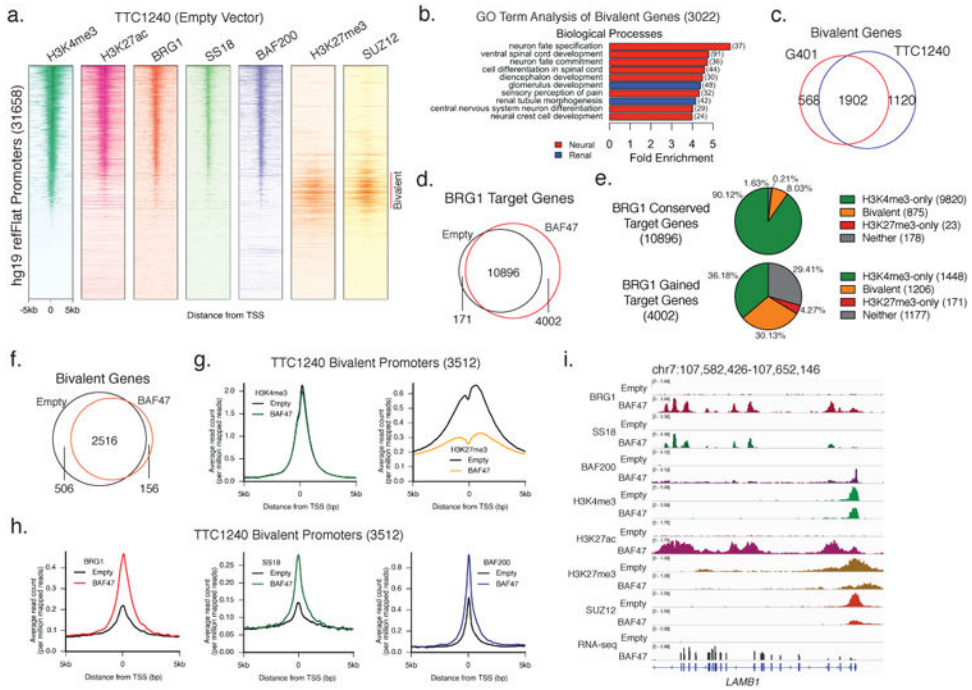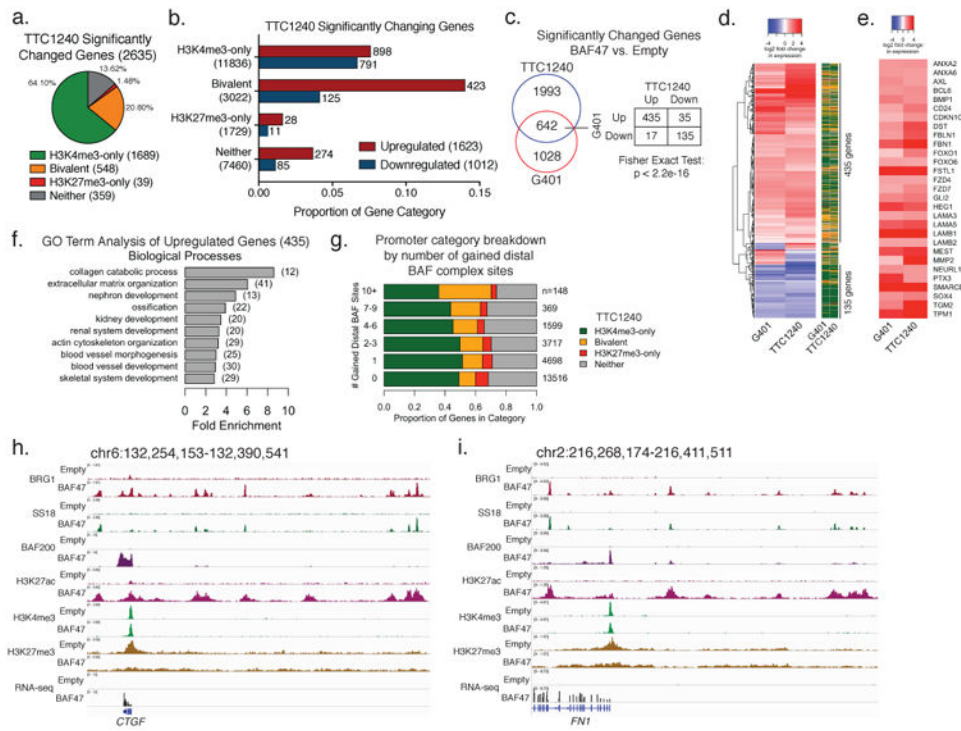
Figure 5.
Resolution of bivalent promoters to activation by BAF complex-mediated opposition of polycomb-mediated repression. (a) Heatmaps of H3K4me3, BRG1, SS18, BAF200, H3K27ac, H3K27me3, and SUZ12 across all hg19 promoters in empty vector condition in TTC1240 cells, ranked by H3K4me3 occupancy. (b) GO term analysis of bivalent genes in TTC1240 cells. (c) Overlap of bivalent genes in G401 and TTC1240 cells. (d) Overlap of BRG1 target genes in empty and BAF47 conditions in TTC1240 cells. (e) Distribution of (left) conserved and (right) gained BRG1 target genes in TTC1240 cells. (f) Overlap of bivalent genes in empty and BAF47 conditions in TTC1240 cells. (g-h) Metagene plots of (g) H3K4me3 and H3K27me3, as well as (h) BRG1, SS18, and BAF200, over all 3512 bivalent promoters in TTC1240 cells. (i) Example tracks at the LAMB1 bivalent promoters demonstrate resolution of bivalent promoters to activation upon gain of BAF complex occupancy in TTC1240 cells.

Figure 6.
Collaborative gene activation by BAF complex-mediated enhancer activation and polycomb opposition at bivalent promoters (a) Distribution of significantly-regulated genes in TTC1240 cells. (b) Directional regulation of significantly changed genes in TTC1240 cells, with y-axis indicating proportion of all genes in each category. (c) (left) Overlap of significantly-changed genes in G401 and TTC1240 cell types in empty vector and BAF47 conditions; (right) genes significantly-regulated in both cell lines show significant concordance (p < 2.2e-16, Fisher exact test). (d) Heatmap of 642 significantly changed genes in both G401 and TTC1240 cells. Right bar indicates promoter status of each gene in each cell line using colors from (a). (e) Heatmap of selected genes that are significantly-upregulated by BAF47 in both G401 and TTC1240 cells. (f) GO term analysis of significantly upregulated genes in both G401 and TTC1240. (g) Genes categorized by number of distal gained (BAF47-only) BAF complex sites, broken down by promoter status of genes in each category. n = number of genes in each group. (h-i) Example ChIP-seq tracks of BRG1, SS18, BAF200, H3K27ac, H3K4me3, H3K27me3, and RNA-seq for (h) CTGF and (i) FN1 shows collaborative gene activation via enhancer activation and polycomb opposition at bivalent promoters.
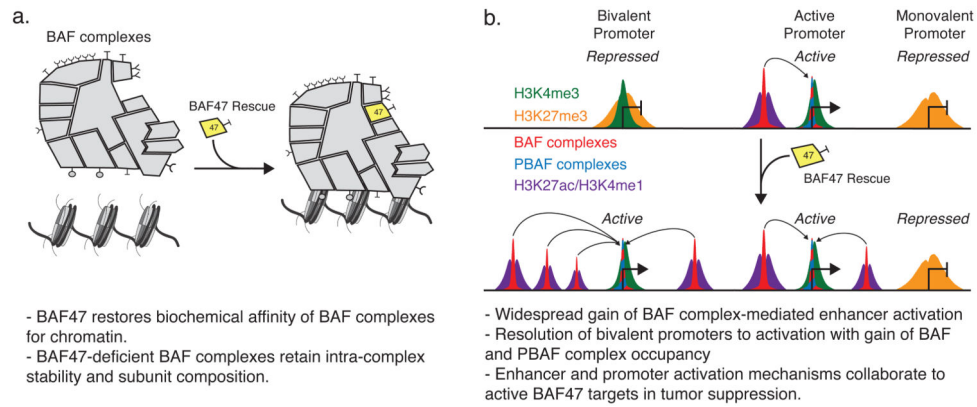
a.

BAF complexes

BAF47 Rescue

- BAF47 restores biochemical affinity of BAF complexes for chromatin.
- BAF47-deficient BAF complexes retain intra-complex stability and subunit composition.

b.

Bivalent Promoter — *Repressed*
Active Promoter — *Active*
Monovalent Promoter — *Repressed*

H3K4me3
H3K27me3
BAF complexes
PBAF complexes
H3K27ac/H3K4me1

BAF47 Rescue

Active *Active* *Repressed*

- Widespread gain of BAF complex-mediated enhancer activation
- Resolution of bivalent promoters to activation with gain of BAF and PBAF complex occupancy
- Enhancer and promoter activation mechanisms collaborate to active BAF47 targets in tumor suppression.

Figure 7.
BAF47 restores BAF complex affinity and functional regulation of chromatin (a) BAF47 rescues the biochemical association of BAF complexes with chromatin, absent major changes in subunit composition or intra-complex stability. (b) Rescue of BAF47 leads to a widespread gain in BAF complex occupancy, mediating enhancer activation and opposition of polycomb-mediated repression at bivalent promoters.