nature genetics

ARTICLES

https://doi.org/10.1038/s41588-019-0477-9

Systematic characterization of BAF mutations provides insights into intracomplex synthetic lethalities in human cancers

Sandra Schick[1,2], André F. Rendeiro[1], Kathrin Runggatscher[1], Anna Ringler[1], Bernd Boidol[1,2], Melanie Hinkel[1,2], Peter Májek[1], Loan Vulliard[1], Thomas Penz[1], Katja Parapatics[1], Christian Schmidl[1,3], Jörg Menche[1], Guido Boehmelt[4], Mark Petronczki[4], André C. Müller[1], Christoph Bock[1,5,6] and Stefan Kubicek[1,2]*

Aberrations in genes coding for subunits of the BRG1/BRM associated factor (BAF) chromatin remodeling complexes are highly abundant in human cancers. Currently, it is not understood how these mostly loss-of-function mutations contribute to cancer development and how they can be targeted therapeutically. The cancer-type-specific occurrence patterns of certain subunit mutations suggest subunit-specific effects on BAF complex function, possibly by the formation of aberrant residual complexes. Here, we systematically characterize the effects of individual subunit loss on complex composition, chromatin accessibility and gene expression in a panel of knockout cell lines deficient for 22 BAF subunits. We observe strong, specific and sometimes discordant alterations dependent on the targeted subunit and show that these explain intracomplex codependencies, including the synthetic lethal interactions SMARCA4–ARID2, SMARCA4–ACTB and SMARCC1–SMARCC2. These data provide insights into the role of different BAF subcomplexes in genome-wide chromatin organization and suggest approaches to therapeutically target BAF-mutant cancers.

The human SWItch/Sucrose Non-Fermentable (SWI/SNF) complexes (BRG1/BRM associated factor (BAF) complexes) are ATP-dependent chromatin remodelers that regulate DNA accessibility dynamically and thereby play important roles in essential cellular processes, such as transcription, DNA repair and replication. These large, polymorphic complexes consist of up to 15 subunits encoded by more than 29 genes[1,2]. On the basis of subunit composition, three major complexes can be distinguished: BAF; polybromo-associated BAF (PBAF); and non-canonical BAF (ncBAF/GBAF); they exist in multiple compositions since several subunit positions can be occupied alternatively by proteins encoded from paralogous genes. The catalytic ATPase encoded by *SMARCA2* or *SMARCA4* is essential for the function of BAF complexes in sliding nucleosomes along the DNA or evicting them from chromatin. Nearly the full remodeling activity of the entire complex is reconstituted in vitro when the ATPase subunits are bound to SMARCB1, SMARCC1 and SMARCC2 (ref. [3]). A further important complex position, alternatively occupied by ARID1A, ARID1B or PBAF-specific ARID2, is thought to recruit the complex to chromatin[4]. Other subunits harbor additional chromatin-binding domains, but much less is known about their contribution to complex function. Currently, no high-resolution structure of the BAF complexes with subunit assignment is available.

Mutations in BAF subunits have been observed with high frequency in human cancers[5,6]. Mostly, these are loss-of-function (LOF) mutations that result in loss of the mutated subunits at the protein level. All subunits can be affected, but mutation prevalence is cancer-dependent, suggesting context- and subunit-specific effects on complex function. Currently, it is unclear exactly how BAF mutations contribute to cancer initiation and/or progression and how they affect BAF complex composition and function. Since LOF mutations are not directly druggable, the focus has been on the identification of synthetic lethalities with BAF mutations. It has been shown that loss of SMARCA4 makes cells dependent on the *SMARCA2* gene[7–9], and ARID1B becomes essential following *ARID1A* mutation[10–12]. These data suggest that paralogous subunits occupying the same key complex positions exert, at least in part, redundant functions and can compensate for each other.

To systematically evaluate the impact of BAF subunit loss on complex composition, chromatin accessibility and transcription, we established and comprehensively characterized isogenic mutant cell lines for 22 targetable BAF subunits in HAP1 cells. In these cell lines, we further tested all intracomplex synthetic lethalities and identified ARID2 and ACTB as potential therapeutic targets for *SMARCA4*-mutant cancers as well as SMARCC2 in *SMARCC1*-mutant cancers and vice versa.

Results

An isogenic cell line panel for loss of individual BAF subunits. We established and comprehensively characterized a panel of isogenic HAP1 cell lines with individual knockouts for 22 targetable BAF subunits (Fig. 1a). The knockouts were validated at the gene

[1]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. [2]Christian Doppler Laboratory for Chemical Epigenetics and Antiinfectives, CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. [3]Regensburg Center for Interventional Immunology, University Regensburg and University Medical Center Regensburg, Regensburg, Germany. [4]Boehringer Ingelheim RCV GmbH & Co KG, Vienna, Austria. [5]Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria. [6]Max Planck Institute for Informatics, Saarbrücken, Germany. *e-mail: skubicek@cemm.oeaw.ac.at

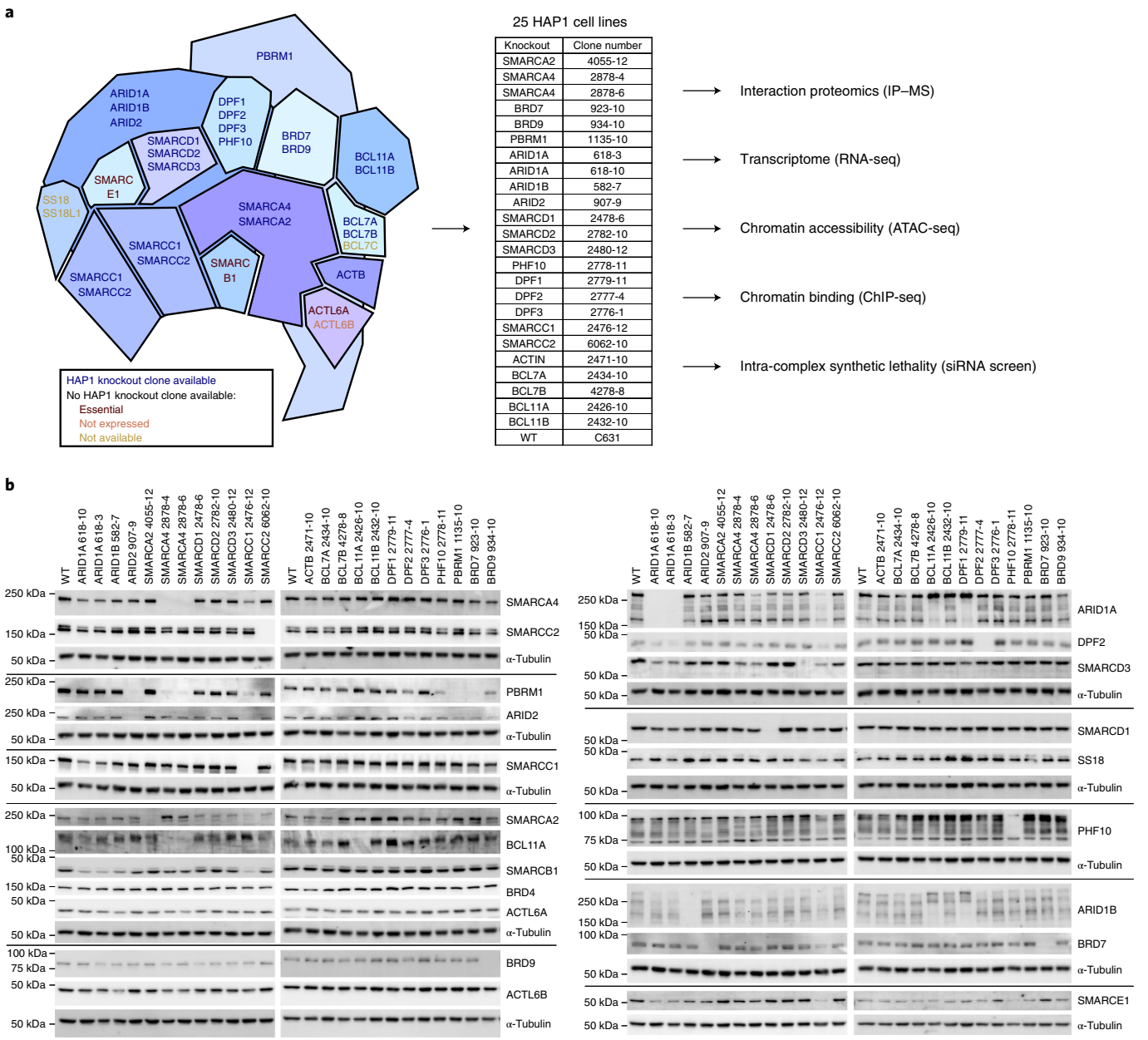


Fig. 1 | An isogenic cell line panel for loss of individual BAF subunits. a, Scheme illustrating which BAF subunits are included in the isogenic HAP1 knockout cell panel and which experiments were applied to these cells. **b**, Cropped immunoblots for HAP1 WT and different knockout cells stained for BAF subunits to confirm their knockouts and check for subunit dependencies.

level by mapping the mutation site (Supplementary Table 1) and at the protein level by immunoblot (Fig. 1b and Supplementary Note 3.1). In addition to loss of the targeted subunit, we also observed coregulation of other BAF members in some of these clonal cell lines, for example, between PBAF-specific subunits. The ARID2 knockout reduced the levels of PBRM1 and BRD7 and caused a moderate reduction of PHF10. BRD7^{KO}, PHF10^{KO} and SMARCA4^{KO} cells also had reduced PBRM1 levels.

BAF complex composition changes following knockout of single BAF-coding genes. To investigate the effects of loss of individual BAF-coding genes on complex composition, we developed a BAF immunoprecipitation approach with quantitative mass spectrometry as the readout. We performed immunoprecipitation for ARID1A, a BAF-specific subunit, and SMARCA4, a subunit incorporated into all BAF complex subtypes. These bait proteins were enriched in immunoprecipitations from the different cell lines, except in the respective knockout cells (Supplementary Fig. 1a). ARID1A levels were additionally reduced in SMARCC1^{KO} cells that expressed low levels of ARID1A. Furthermore, we did not observe enrichment for the mutually exclusive subunits ARID1B and ARID2 in the ARID1A immunoprecipitation nor for SMARCA2 in the SMARCA4 immunoprecipitation, suggesting paralog specificity of the antibodies used (Supplementary Fig. 1b,c). All other BAF and PBAF members were enriched in the SMARCA4 immunoprecipitation, except ACTL6B, BCL11A and BCL11B, which were not or only very lowly expressed in HAP1 cells. As expected, the PBAF-specific subunits PBRM1, PHF10, BRD7 and ARID2 were detected in the SMARCA4 immunoprecipitations but not in the ARID1A IPs. Also, BRD9 was only detected in the SMARCA4 immunoprecipitations, in line with

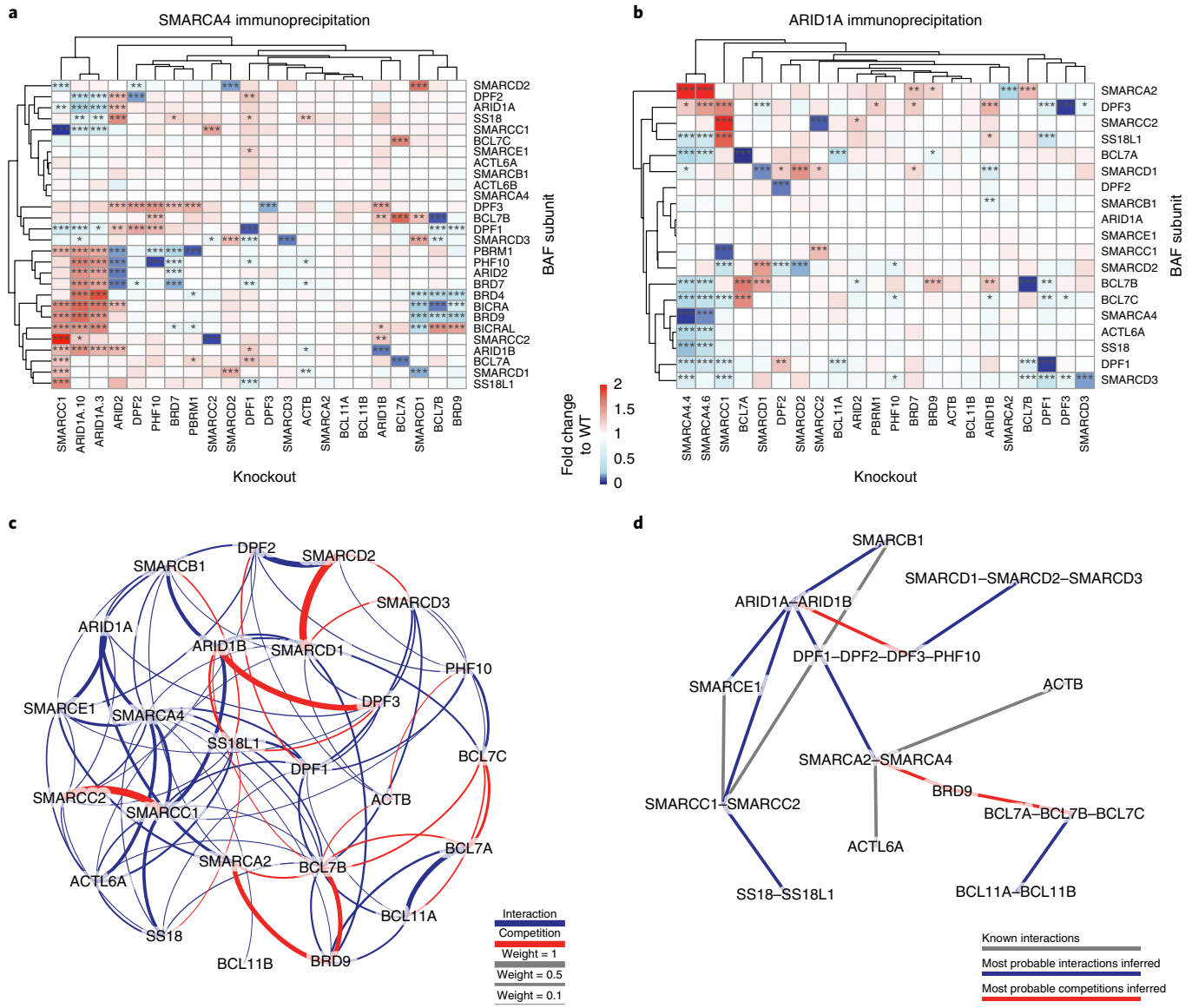


Fig. 2 | BAF complex composition changes following knockout of single BAF-coding genes. a, Heatmap showing enrichment of BAF subunits in SMARCA4 immunoprecipitation in the different HAP1 knockout cells relative to their enrichment in HAP1 WT cells. Hierarchical clustering of the knockouts and BAF subunits was done based on Euclidean distance ($n = 2$ biologically independent experiments). Significant changes were calculated as described in the Supplementary Note. The resulting false discovery rate values are indicated by an asterisk (*<1%, **<0.1%, ***<0.01%). **b**, Same as in **a** for the results of the ARID1A IP–MS. **c**, Weighted frequencies of all pairwise interactions (blue) or competitions (red) between the subunits of the BAF complexes as inferred from simulations based on the ARID1A IP–MS data. The weight of an interaction or competition is given by the normalized sum of the fitness of all simulated graphs where it was observed. A weight cutoff > 0.1095 was used, which corresponds to the minimal weight at which all subunits form a single connected graph. **d**, Putative organization of the BAF complexes obtained by combining interactions and competition classes with direct experimental evidence and the most probable interactions inferred from the simulations based on the ARID1A IP–MS data (using a weight cutoff > 0.42 to obtain a single connected graph).

its incorporation into GBAF/ncBAF complexes that do not contain any ARID subunit[13,14].

Analyses of the relative abundance of the detected subunits revealed knockout-specific alteration of complex compositions (Fig. 2a,b and Supplementary Fig. 1d). None of the tested knockouts led to a complete disruption of the BAF complexes in contrast to an observed dissociation of the SWI/SNF complex following loss of Snf5, Snf6 and Snf12 in yeast[15,16]. This is probably due to the evolution of modularity of the mammalian SWI/SNF complex. For example, *Snf12* has three orthologs in humans, namely *SMARCD1*, *SMARCD2* and *SMARCD3*. Our data suggest that the loss of either is compensated by increased incorporation of a paralog-encoded protein. *Snf6* does not have an ortholog in humans and *SMARCB1*, the ortholog of *Snf5*, could not be tested because the gene is essential in HAP1 cells[17].

Loss of one protein was often compensated by increased incorporation of its paralogous proteins into the complex, for example, for ARID1A–ARID1B–ARID2, SMARCA2–SMARCA4, SMARCC1–SMARCC2, SMARCD1–SMARCD2–SMARCD3, DPF1–DPF2–DPF3–PHF10 and BCL7A–BCL7B–BCL7C. This argues for the mutually exclusive incorporation of these subunits to fulfill the requirement to keep the respective complex position occupied.

While in most instances such compensatory effects occurred in a mutual fashion irrespective of which subunit was targeted, there are also cases where these mechanisms appear unidirectional.

Clustering of the subunits by their relative abundance in SMARCA4 immunoprecipitation revealed clear coregulation of the PBAF-specific subunits ARID2, PBRM1, PHF10 and BRD7. These subunits were all increased in complexes isolated from ARID1A^{KO} cells, whereas they were lost in ARID2^{KO} and to a lesser degree in BRD7^{KO} clones. Loss of PHF10 resulted in only a mild reduction of PBRM1 incorporation, whereas PBRM1 loss did not affect the level of any other PBAF-specific subunit in the complexes. These data suggest a clear hierarchy in the assembly of the PBAF complexes: ARID2>BRD7>PHF10>PBRM1. This observation is in line with a previous report showing that ARID2 is important for the stability of PBRM1 and the PBAF complexes and not vice versa[18]. Furthermore, we found that, besides ARID2^{KO} cells, BRD7^{KO} and to a smaller extent PHF10^{KO} cells also expressed less PBRM1 at the protein level and showed reduced enrichment of specific PBAF members in SMARCA4 immunoprecipitations, suggesting feedback mechanisms by destabilization of unincorporated proteins.

The multiple positions that can alternatively be occupied by different subunits suggest that these complexes can theoretically exist in thousands of different configurations; it is unknown which of these exist in human cells. To simulate which subunits directly lead to the increased ('interaction') and decreased ('competition') incorporation of other subunits into the BAF complexes, we implemented a genetic algorithm that tested all pairwise combinations using the immunoprecipitation–mass spectrometry (IP–MS) data as input. We first investigated the BAF complexes using the ARID1A IP–MS data (Fig. 2c,d and Supplementary Fig. 2a,b). The computational analysis confirmed very strong competition between SMARCD1–SMARCD2–SMARCD3 and SMARCC1–SMARCC2. It further suggested competition between ARID1B–DPF3 and SMARCA2–BRD9–BCL7B. We then repeated the analysis with just the SMARCA4 IP–MS data and with all data combined providing the input for all possible BAF and PBAF complexes that contain either SMARCA4 and/or ARID1A (Supplementary Fig. 2c–h). In addition to the competitions already found in the analysis of ARID1A-containing BAF complexes, these data revealed competition between BCL7A–BCL7B–BCL7C, DPF1–DPF2, DPF3–PBRM1 and ARID1A–ARID1B–ARID2. The analyses further provide strong support for the interaction between SMARCC1–ARID1A and the PBAF-specific subunits ARID2–BRD7–PHF10–PBRM1.

We performed the immunoprecipitation experiments under stringent conditions, but could still identify some interaction partners of the BAF complexes (Supplementary Fig. 3a). Among the most strongly enriched interaction partners found in the SMARCA4 immunoprecipitation were bromodomain-containing protein 4 (BRD4) and BRD4-interacting chromatin-remodeling complex-associated protein-(like) (BICRA and BICRAL). They were not identified in the ARID1A immunoprecipitations. Moreover, the interaction with all three proteins was lost in SMARCD1^{KO} cells, and BRD4 and BICRA enrichment was also reduced in BRD9^{KO} and BCL7B^{KO} cells (Fig. 2a and Supplementary Fig. 3b). Interestingly, a common feature of these three cell lines is decreased incorporation of BRD9 into the complexes, in line with the specificity of these proteins for the recently described GBAF/ncBAF complexes[13,14,19] and the presence of a functional SMARCD1–BRD9–BICRA (GLTCR1) module identified in fitness correlation network analyses from RNA interference (RNAi) screens[20]. Accordingly, we observed stronger enrichment of BRD4 and BICRA when we used an antibody against BRD9 for immunoprecipitation, compared to the ARID1A, ARID2 or common BAF subunit immunoprecipitations (Supplementary Fig. 3c).

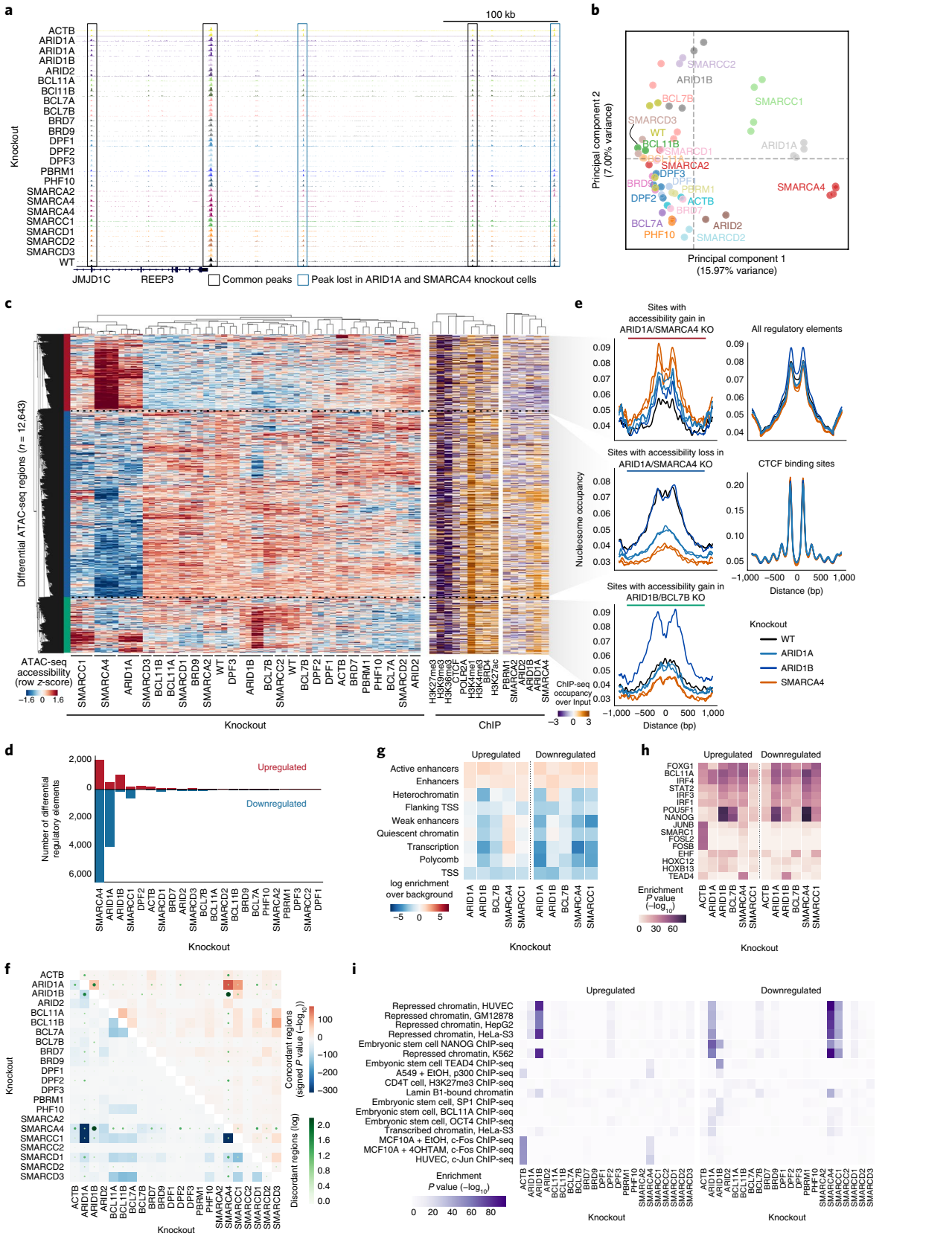
Knockout of single BAF-coding genes alters global chromatin accessibility. Next, we systematically investigated the functional consequences of the loss of individual BAF-coding genes on chromatin accessibility by applying the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) in HAP1 wild-type (WT) and knockout cells. The overall quality of the samples and the global distribution of open chromatin regions were highly comparable across all samples (Supplementary Fig. 4a–e and Supplementary Table 2). In general, chromatin accessibility was observed mainly at 5′ untranslated regions, transcription start sites (TSS) and enhancers, and at regions harboring histone marks that are associated with active chromatin (Supplementary Fig. 4f–h).

While many genomic sites showed similar accessibility in all cell lines, we also observed knockout-specific differences at specific loci (Fig. 3a). This prompted us to systematically assess global similarity in chromatin accessibility across the various knockout cells (Fig. 3b and Supplementary Fig. 4i). These analyses revealed that there are differences in chromatin accessibility for all knockout clones compared to WT cells with the strongest alterations in the SMARCC1^{KO}, ARID1A^{KO} and SMARCA4^{KO} clones. This observation was confirmed by the locus-specific analysis of regions significantly deregulated for each knockout. SMARCA4^{KO}, ARID1A^{KO} and SMARCC1^{KO} cells grouped together and showed, compared to WT cells, reduced accessibility at many regions across the genome (blue cluster), and accessibility gain at fewer other regions (red cluster) (Fig. 3c,d). This is in line with observations in colorectal cancer cells, where knockdown of ARID1A leads to reduced accessibility at many genomic loci[11]. In contrast, ARID1B^{KO} cells gained chromatin accessibility at numerous genomic loci compared to WT cells (green cluster). In line with these findings, the total open chromatin fraction was reduced in SMARCA4^{KO}, ARID1A^{KO} and SMARCC1^{KO} cells and increased in ARID1B^{KO} cells compared to WT cells (Supplementary Fig. 4j). By leveraging the paired-end ATAC-seq data, we also investigated how nucleosome positions in these regions changed (Fig. 3e). Sites with reduced accessibility in knockout lines gained nucleosome signal at the center of ATAC-seq peaks compared to WT cells and vice versa, providing functional implications of nucleosome positioning changes on chromatin accessibility.

To assess the chromatin context of the sites that changed accessibility, we performed chromatin immunoprecipitation (ChIP-seq)

Fig. 3 | Knockout of single BAF-coding genes alters global chromatin accessibility. a, Example browser tracks of ATAC-seq data. **b**, Principal component analysis (PCA) of the ATAC-seq data for individual replicates using all regions where a peak was called in any sample (154,533 regions). The knockout (KO) gene is indicated. **c**, Left: clustering of samples and regions significantly different between any knockout compared to WT ATAC-seq data. Hierarchical clustering was performed with Pearson correlation as the distance measure between accessibility values transformed with a *z*-score per row. Right: ChIP-seq enrichment for the indicated proteins at differential ATAC-seq regions in WT cells. **d**, Number of upregulated and downregulated regions in each knockout compared to WT cells. **e**, Nucleosome occupancy for selected cell lines, as determined with the NucleoATAC method, is shown around the indicated genomic sites. **f**, The percentage of regions changing in two knockouts concordantly compared to WT is shown in blue for downregulated regions and in red for upregulated regions. The number of disconcordantly changing regions is indicated with green dots. **g–i**, Enrichment in chromatin states (**g**), transcription factor motif (**h**) and ChIP-seq data from the ENCODE and BLUEPRINT consortia, as determined by the LOLA method (**i**), of upregulated and downregulated regions in the indicated knockout cells. **f**,**h**,**i**, The significance of the overlap was assessed with a one-sided Fisher's exact test; no adjustments were made for multiple comparisons. All panels were derived from at least $n=2$ biologically independent ATAC-seq experiments.

in HAP1 WT cells and associated them to the differential ATAC-seq regions (Fig. 3c and Supplementary Table 3). The sites of the blue and green clusters showed enrichment for members of the BAF chromatin complexes as well as for enhancer marks in HAP1 WT cells, indicating that preferentially BAF-bound active regulatory regions are changing accessibility when a specific BAF subunit is lost, and confirming a function of the BAF complexes not only at TSS, but also at distal regions[11,12,21,22]. On the other hand, red cluster sites were



enriched for inactive histone modifications such as H3K27me3 and H3K9me3 and not bound by the BAF complexes under WT conditions. However, these sites are bound by ARID1A-containing BAF complexes in SMARCA4^{KO} cells, which showed the strongest gain of chromatin accessibility in these regions (Supplementary Fig. 5). This suggests mistargeting of the aberrant BAF complexes to these loci under SMARCA4 knockout conditions. Together, the data provide strong evidence that intact BAF complexes containing SMARCA4, ARID1A and SMARCC1 are important for the recruitment of BAF complexes to cell-specific enhancer sites and their activation. Our data further suggest that BAF and PBAF complexes occupy different genomic loci (Supplementary Fig. 6). While the BAF complex is preferentially bound to active enhancers (marked with H3K4me1 and H3K27ac), PBAF-enriched sites are either active (marked with H3K4me3/H3K36me3/POLR2A) or inactive (marked with H3K27me3/H3K9me3) regions.

Analyses of shared and discordantly changing accessibility across the different cell lines illustrated again highly similar changes in SMARCA4^{KO}, SMARCC1^{KO} and ARID1A^{KO} cells that are distinct from the accessibility changes in other cell lines (Fig. 3f and Supplementary Fig. 7a). Furthermore, knockout of mutually exclusive subunits resulted in similar as well as discordant changes at certain genomic loci. For example, accessibility at some genomic loci is changed in the opposite direction in ARID1A^{KO} versus ARID1B^{KO} cells, in line with published data[23]. Accordingly, regions of altered accessibility in the two different knockouts showed opposite patterns for their localization across the genome, transcription factors (particularly pluripotency factors) and ChIP-seq signal enrichment (Fig. 3g–i and Supplementary Fig. 7b). Overall, the data suggest that structural compensation of mutually exclusive subunits may not necessarily rescue the function of the lost protein since they might have partly distinct regulatory roles.

Expression changes correlate with altered chromatin accessibility. To test whether the observed alterations in chromatin accessibility go along with changes in gene expression, we measured the transcriptome of all BAF-mutant HAP1 clones using RNA sequencing (RNA-seq; Supplementary Table 4). We first analyzed the expression pattern of BAF members across the different samples, which suggested a transcriptional coregulation of some subunits (Supplementary Fig. 8).

Investigating changes in global gene expression, knockout cells showing similar changes in chromatin accessibility also clustered together in regard to their transcriptomes (Fig. 4a,b). This was especially obvious for knockout cells with prominent accessibility changes, namely the SMARCA4^{KO}, ARID1A^{KO} and SMARCC1^{KO} cells, which also showed the most differentially expressed genes (Fig. 4c and Supplementary Fig. 9a). In general, nearly similar numbers of up- and downregulated genes were observed in each cell line. Despite overall coordination in expression changes between knockout cell lines, we also found a small number of genes with opposite directions of change, especially compared to SMARCA4^{KO} cells (Fig. 4d).

To test an association between ATAC-seq and RNA-seq data, we next compared the changes in ATAC-seq enrichment to the expression alterations of the associated genes and found a good correlation overall (Fig. 4e–g). This suggests a causative relationship between chromatin remodeling and the resulting cellular expression status. We then analyzed the transcriptome changes regarding enrichment for transcription factor binding sites and Gene Ontology terms. For example, genes downregulated in SMARCA4^{KO}, ARID1A^{KO} or ARID1B^{KO} cells were related to cell migration and transcription regulation (Fig. 4h and Supplementary Fig. 9b). In general, the down- and upregulated genes in ARID1A^{KO} and SMARCA4^{KO} cells were enriched for similar Gene Ontology terms (Supplementary Fig. 10).

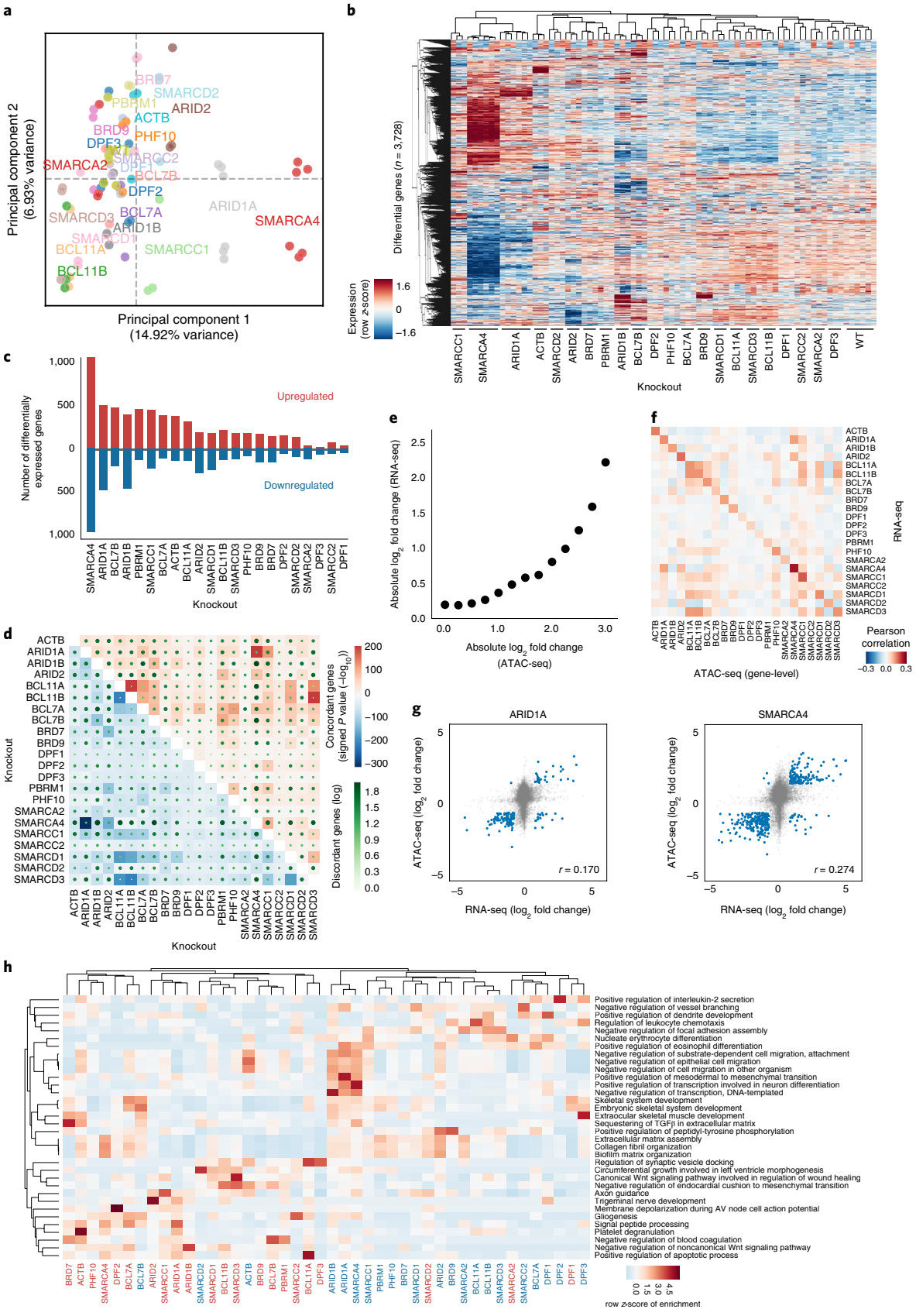
Systematic targeting of multiple BAF subunits identifies previously unknown intracomplex synthetic lethalities. We next applied the fully characterized cell line panel to identify intracomplex synthetic lethalities with BAF mutations. Therefore, we assembled pools of small interfering RNAs targeting all BAF subunits individually and all possible combinations of alternative subunits. We then transfected all HAP1 clones with these siRNA pools and measured viability after 5 d in cell culture (Supplementary Table 5). While transfection of the non-targeting control did not reduce the viability of the cells, knockdown of Polo-like kinase 1 (PLK1), whose inhibition has been shown to induce apoptosis in leukemic cells[24–26], led to cell death across all cell lines included in the screen (Supplementary Fig. 11a). This argues for similar efficiency of RNAi-mediated knockdown across all included cell lines. Unbiased clustering of cell viability data revealed that the sensitivities and resistances of the BAF-mutant cell lines can be explained in part by the alterations we observed in BAF complex composition and in part by chromatin and transcriptome changes (Fig. 5a,b and Supplementary Fig. 11a–d). For example, ARID2^{KO}, BRD7^{KO} and PBRM1^{KO} cells were most sensitive, while BCL7B^{KO} and ARID1B^{KO} clones, for which we had observed similar changes in gene expression, were resistant to further siRNA knockdown of many BAF subunits.

The three cell lines that showed strong loss of chromatin accessibility did not show co-clustering in the siRNA screen; however, they were all sensitive to knockdown of their respective paralog. SMARCA4^{KO} cells required SMARCA2 for their survival and ARID1A^{KO} cells were sensitive to ARID1B knockdown, both in line with previous reports[7–12]; however, SMARCC1^{KO} cells were also sensitive to loss of SMARCC2. Both SMARCC1 and SMARCC2 subunits, despite high sequence homology, are thought to be incorporated into BAF complexes simultaneously[27], but we observed the same signatures for these proteins as for the alternative subunits SMARCA2–SMARCA4 and ARID1A–ARID1B. Therefore, our data suggest the need to reevaluate whether these subunits are incorporated into BAF complexes as monomers, homodimers and/or heterodimers. In addition, we observed that SMARCA4^{KO} cells were sensitive to knockdown of ARID2, ACTB and SMARCB1. SMARCA2^{KO} cells had reduced viability following knockdown

Fig. 4 | Expression changes correlate with altered chromatin accessibility. a, PCA of RNA-seq data in HAP1 WT and knockout cells (knockout gene is indicated) using all expressed genes. **b**, Clustering of samples based on differentially expressed genes between any knockout and WT cells (knockout gene is indicated). Hierarchical clustering was performed with Pearson correlation as the distance measure between expression values transformed with a *z*-score per row. **c**, Number of upregulated and downregulated genes per knockout. **d**, The percentage of genes that change in two knockouts concordantly compared to WT is shown in blue for downregulated and in red for upregulated genes. Disconcordantly changing genes are indicated with green dots. Significance of overlap was assessed with a one-sided Fisher's exact test without adjustments for multiple comparisons. **e**, Relationship of expression and ATAC-seq signal aggregated per gene across all knockouts. Genes were binned into groups based on the ATAC-seq change. **f**, Pearson correlation of log fold changes in knockout versus WT cells (knockout gene indicated) between RNA-seq and ATAC-seq. **g**, Scatterplot and Pearson correlation coefficient for HAP1 ARID1A^{KO} and SMARCA4^{KO} cells between RNA-seq and ATAC-seq log fold changes relative to HAP1 WT cells. **h**, Overrepresentation enrichment analysis of Gene Ontology terms for upregulated or downregulated genes in different HAP1 knockout cells. A row *z*-score on the $-\log_{10}(P)$ of a one-sided Fisher's exact test (Enrichr software) is displayed without adjustments for multiple comparisons. The RNA-seq data of all panels are derived from $n = 3$ biologically independent experiments per clone.

of PBRM1 and DPF2^{KO} cells were sensitive to knockdown of SMARCA4 and ACTL6A.

We then assessed changes in viability of BAF-mutant cell lines relative to HAP1 WT cells in relation to changes in complex compositions measured by co-immunoprecipitation (Fig. 5b). We observed that subunits that are incorporated with increased abundance in BAF complexes after a particular knockout often constitute vulnerabilities of these cell lines. For example, in SMARCA4^{KO}



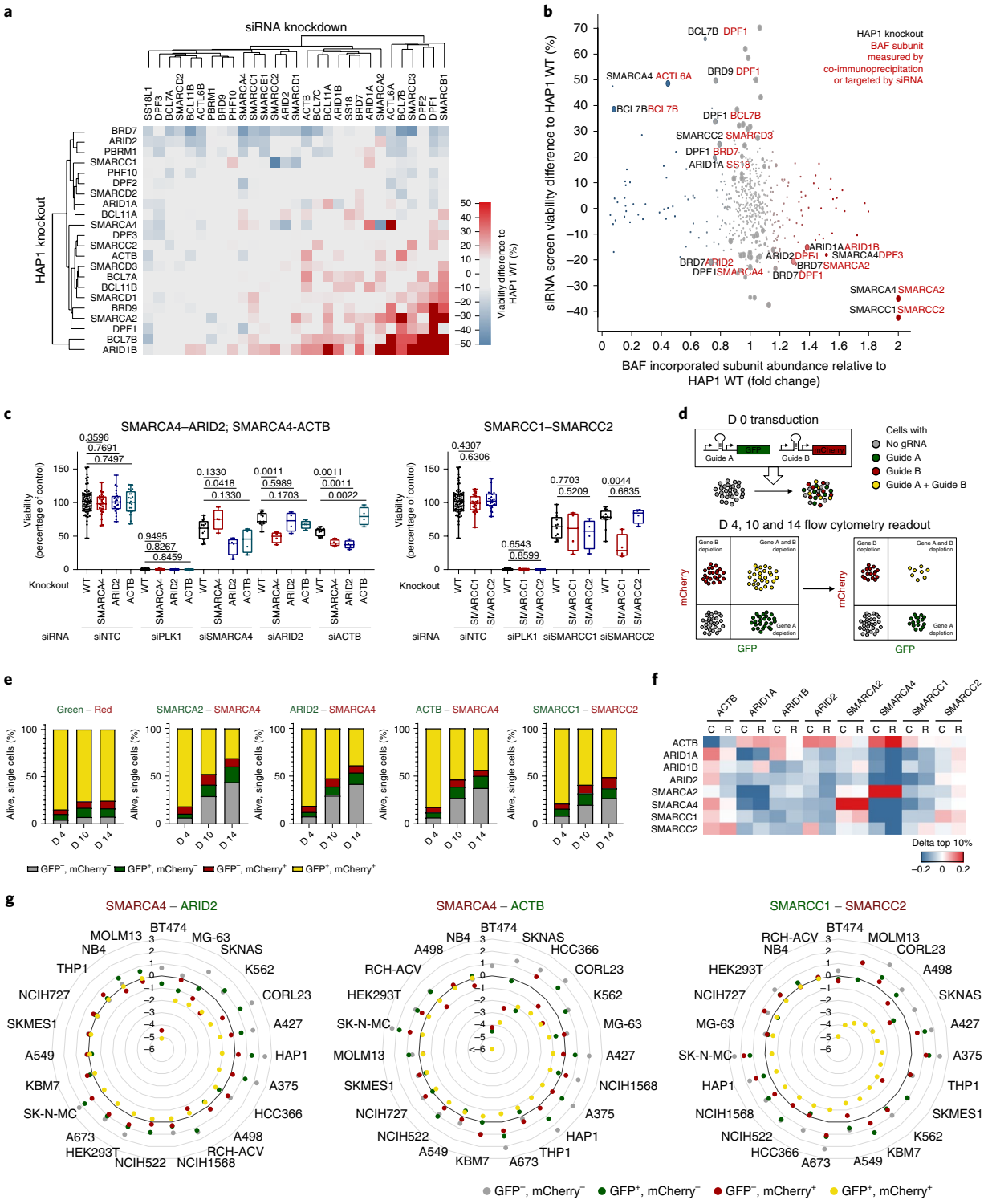


Fig. 5 | Systematic targeting of multiple BAF subunits identifies previously unknown intracomplex synthetic lethalities. a, Heatmap showing the viability difference between HAP1 knockout cell lines (*y* axis) and HAP1 WT cells following siRNA treatment (*x* axis). **b**, The effect on viability in different knockout cells (black) following siRNA treatment (red) (siRNA screen data, *y* axis) is compared to the change of subunit incorporation (red) in different knockout cells (black) relative to their incorporation in WT cells (IP–MS data, *x* axis). **c**, Boxplot showing viability 5 d after siRNA knockdown for three selected synthetic lethalities. A Mann–Whitney *U*-test was used to calculate significant viability difference in knockout cells relative to WT cells treated with the same siRNA ($n \geq 4$ independent measurements from two biologically independent experiments). The first and third quartiles are denoted by the lower and upper hinges. The whiskers extend to the smallest and largest values. **d**, Schematic explaining the CRISPR–Cas9 multicolor competition assay. **e**, Multicolor competition assay in HAP1 cells. The bar graphs show the percentages of the four different cell populations 4, 10 and 14 d after transduction of the gRNAs targeting the genes indicated above (green, red: vectors with the fluorescent marker only). **f**, Heatmap showing if cells expressing low levels of the gene (left) are more sensitive to CRISPR (C) or RNAi (R) targeting a BAF subunit (top) as cells that express high levels of this gene. Analyses were done using public data (depmap: CRISPR (Avana) Public 18Q4; combined RNAi). **g**, Multicolor competition assay results for three synthetic lethalities across various cell lines. $\log_2$(day 14/day 4) is displayed for each population. The targeted subunits are indicated above.

cell lines there is increased complex incorporation of SMARCA2; siRNA-mediated targeting of SMARCA2 specifically caused a growth disadvantage in SMARCA4^{KO} cells. The following subunits also showed increased BAF complex incorporation and evidence of requirements for cell viability: ARID1B in ARID1A^{KO} cells; SMARCC2 in SMARCC1^{KO} cells; DPF1 in ARID2^{KO} and BRD7^{KO} cells; SMARCA2 in BRD7^{KO} cells; and DPF3 in SMARCA4^{KO} cells. On the other hand, cell lines were often relatively more resistant to knockdown of subunits that showed reduced complex incorporation in a particular knockout. This is most prominently exemplified by the loss of ACTL6A in SMARCA4-mutant complexes and the resistance of SMARCA4^{KO} cells to ACTL6A knockdown, which goes along with the described direct binding of ACTL6A to SMARCA4 (ref. [28]). Additional examples are BCL7B and BRD7 in DPF1^{KO} cells, DPF1 in BRD9^{KO} and BCL7B^{KO} cells, and SMARCD3 in SMARCC2^{KO} cells.

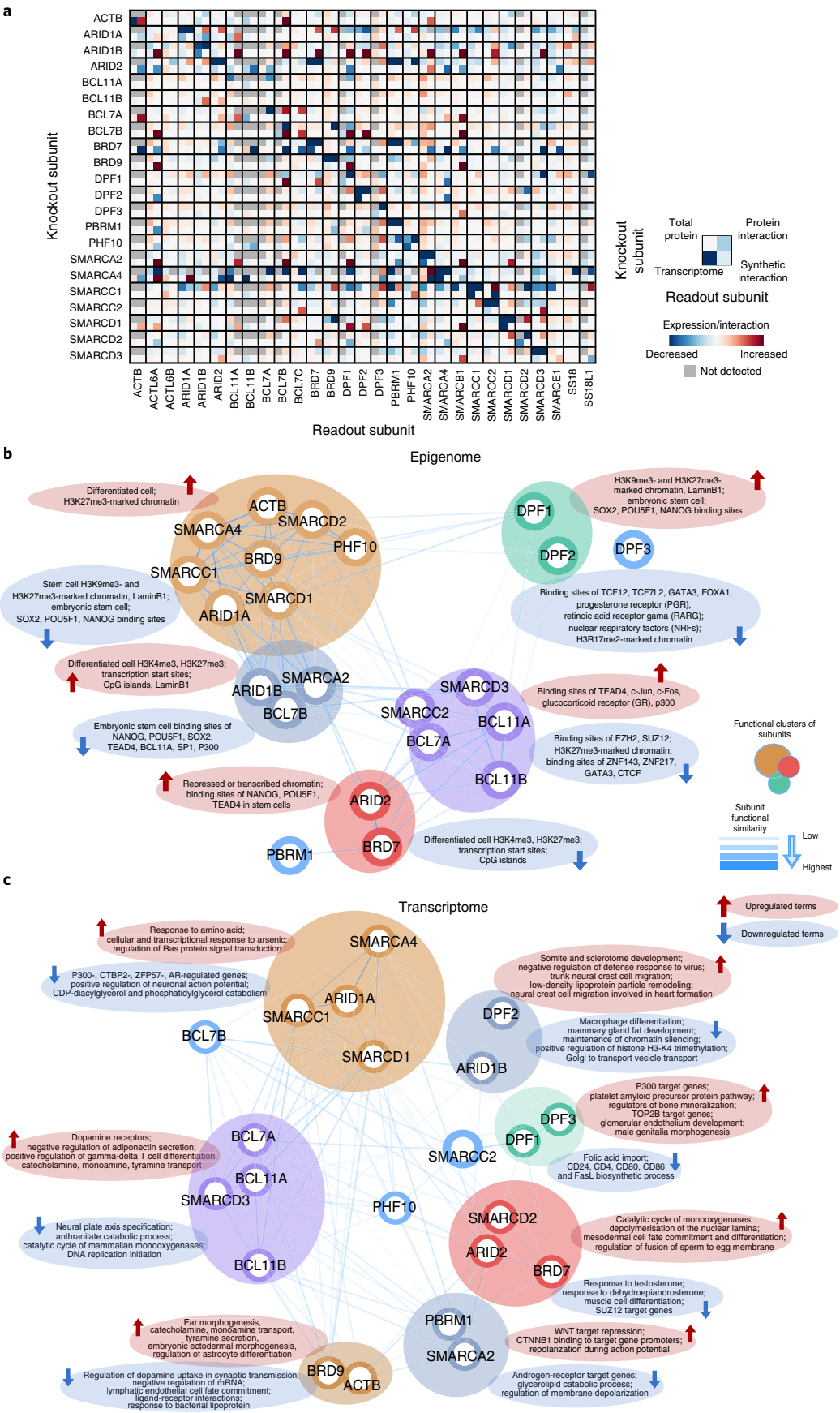
Intracomplex synthetic lethalities alter BAF complex composition. To choose the most robust intracomplex synthetic lethal pairs, we next analyzed the data for reciprocal synthetic lethalities observed in both cell line/ knockdown combinations. The following pairs stood out: SMARCA4–ACTB; SMARCA4–ARID2; and SMARCC1–SMARCC2 (Fig. 5c). Growth curves obtained by quantifying nuclei after Hoechst staining following siRNA treatment confirmed slower growth under the synthetic lethal conditions (Supplementary Fig. 11e). We next validated these synthetic interactions using CRISPR–Cas9-mediated gene targeting as an alternative approach to siRNA-mediated knockdown. Therefore, we applied a multicolor competition assay with flow cytometry as the readout (Fig. 5d and Supplementary Fig. 11f,g). While transduction of the vectors without any guide RNA (gRNA) barely showed a change of the different populations with time, the population of cells that contained both gRNAs of the new synthetic lethalities decreased over time nearly as strongly as when the well-described synthetic lethal genes *SMARCA4* and *SMARCA2* were targeted (Fig. 5e). These data independently validate and confirm all three synthetic lethalities under conditions of full genetic knockout.

We next explored the biochemical relationships between synthetic lethal subunits that may explain the phenotype. We used siRNA to knockdown the synthetic lethal subunit in the corresponding HAP1 knockout cell line for the pairs SMARCA4–ARID2, SMARCA4–ACTB and SMARCC1–SMARCC2 and analyzed cell cycle distribution and complex composition (Supplementary Fig. 12). When we targeted SMARCC1 in SMARCC2^{KO} cells and vice versa, we observed strong reductions of the global protein levels of the key BAF subunits SMARCA4, SMARCB1, ARID1A and SMARCD1 (Supplementary Fig. 12a). This is in line with similar observations in HEK293T cells[13]. Therefore, nearly complete destabilization of BAF complexes is probably causative for reduced viability and proliferation when both SMARCC1 and SMARCC2 are lost (Supplementary Fig. 12b). For the other two synthetic lethalities, we investigated the BAF complex compositions in the combined versus single perturbed conditions by SMARCC1 IP–MS (Supplementary Fig. 12c). The data revealed that both constitutive and acute loss of SMARCA4 resulted in increased ACTB incorporation into BAF complexes. Like other examples where increased complex incorporation results in increased dependency on a particular subunit, ACTB specifically also becomes essential to the viability of SMARCA4^{KO} cells. Analyzing the SMARCC1 IP–MS data further enabled us to assess the effects of SMARCA4 loss on the incorporation of PBAF-specific subunits, information that cannot be obtained from the SMARCA4 and ARID1A immunoprecipitations. Besides reduced incorporation of ACTL6, BCL7 and DPF1 into BAF complexes, SMARCA4 knockout resulted in substantial loss of PBRM1 expression and complex incorporation (Fig. 1b and Supplementary Fig. 12c), whereas the other PBAF subunits ARID2, BRD7 and PHF10 were present at normal or slightly increased levels. When ARID2 was depleted in SMARCA4^{KO} cells, these other PBAF-specific subunits were lost in addition to PBRM1. Overall, the observed changes in complex composition after the targeting of two subunits correlated well with the expected changes based on the network that we previously generated from the IP–MS data across the different knockout cells (Supplementary Fig. 12d). Using ARID2 immunoprecipitation, we further validated that HAP1 SMARCA4^{KO} cells contained PBAF complexes, but these contained less PBRM1 than WT cells (Supplementary Fig. 12e). Moreover, combined perturbation of SMARCA4 and ARID2 led to accessibility changes in regions related to cell death and gene expression alterations of cell cycle genes (Supplementary Fig. 13). In summary, these data suggest that, in the case of SMARCA4–ARID2 synthetic lethality, the effects of the single lost subunits add up to result in more severe perturbation of BAF complexes and consequent detrimental chromatin and gene expression changes.

Validation of intracomplex synthetic lethalities in other cell lines. Finally, we tested whether these synthetic lethalities are conserved across different cell types by first analyzing publicly available datasets on gene essentiality in cancer cell lines (DepMap: CRISPR (Avana) Public 18Q4) (Supplementary Fig. 14) (refs. [29–34]). Since only a few cancer cell lines included in the dataset harbor mutations in the synthetic lethal subunits SMARCC1, SMARCC2, ARID2 and ACTB, and the functional consequences of these mutations are often uncharacterized, we compared the sensitivity of the top 10% versus the bottom 10% BAF subunit-expressing cell lines to depletion of other synthetic lethal BAF subunits by CRISPR or RNAi (Fig. 5f). The analyses revealed high conservation of the SMARCA4–SMARCA2 synthetic lethality, frequent occurrence of ARID1A–ARID1B, SMARCC1–SMARCC2 and SMARCA4–ACTB synthetic lethality, while SMARCA4–ARID2 synthetic lethality was not widely conserved in other cell lines. These observation may be confounded by the arbitrary thresholds for high/low-expressing cells and the low variability of expression levels across cell lines for some BAF genes, such as *SMARCC1*, *SMARCC2* or *ARID2*. Therefore, we used the multicolor competition assay to directly test the three previously

Fig. 6 | Integrative view of BAF complex subunit dependencies and functional similarity. a, Joint visualization of the effects of single BAF subunit perturbation on the remaining complex subunits. Data from four assays have been used: (1) transcriptional changes compared to WT cells (log_2 fold change); (2) total protein changes relative to WT cells as measured by immunoblot; (3) subunit incorporation into BAF complexes relative to WT cells as measured by IP–MS; (4) synthetic interactions via siRNA knockdown in knockout cells compared to the WT background (percentage difference to WT). Measurements not detected or not performed are shown in gray. **b**,**c**, Functional similarity of BAF complex subunits assessed through similarity of enriched terms in knockout cells compared to WT cells. **b**, Differential chromatin regions (measured by ATAC-seq, $n = 2$ biologically independent experiments) were used. **c**, Differentially expressed genes (measured by RNA-seq, $n = 3$ biologically independent experiments) were used. The strength of similarity is reflected in the two-dimensional proximity between subunits and the width and transparency of the edges connecting them. The distances between subunits were derived using Pearson correlation of enrichment significance (one-sided Fisher's exact test; *P* value without adjustments for multiple comparisons). In the case of epigenome similarity, enriched gene sets from the LOLA tool were used; for the transcriptome, various ontology, pathway and perturbation terms from the Enrichr tool were used. The top enriched terms specific to each cluster of subunits were extracted by comparing the mean enrichment values across subunits in the same cluster versus all other subunits.

unknown synthetic lethality candidates in 23 different cell lines originating from multiple tissues (Fig. 5g, Supplementary Fig. 15a and Supplementary Table 1). Some of these cell lines (NB4, NCI-H727, A-498, HEK293T, RCH-ACV, NCI-H522) were mostly resistant to all perturbations, including the well-established SMARCA2–SMARCA4 synthetic lethality (Supplementary Fig. 15b–d),



possibly in part due to low editing efficiency in certain cell lines (Supplementary Note 2.1). The fractions of SMARCA4–ARID2 and SMARCA4–ACTB double-targeted cells were each depleted more than twofold over a 14-d period in approximately one-third of the cell lines tested. In contrast to the homozygous SMARCA4-mutant cell lines, depletion of ARID2 or ACTB alone did not reduce cell viability of heterozygous SMARCA4-mutant cells (COR-L23), suggesting that complete loss of SMARCA4 is necessary for these synthetic lethalities. These data validate the two synthetic lethalities but also highlight their dependency on context. For the SMARCC1–SMARCC2 pair, we observed strong synthetic lethality in 14 of the tested cell lines, to a similar level and in the same cell lines as the known SMARCA4–SMARCA2 vulnerability. Importantly, in SK-MES-1 cells that harbor an SNP in SMARCC1 resulting in very low SMARCC1 expression, targeting of SMARCC2 alone depletes cells nearly as efficiently as concomitant targeting of SMARCC1–SMARCC2. In summary, these data confirm that additional cell lines respond to combined SMARCA4–ARID2 and SMARCA4–ACTB targeting and indicate that SMARCC1–SMARCC2 is a highly conserved and strong synthetic lethality that can be exploited in cancers cells with low expression of either subunit.

Discussion

LOF mutations of different subunits of the BAF chromatin remodeling complexes have been detected in many human cancers; however, the cellular consequences of and differences between the loss of individual subunits are not fully understood. In the current study, we present a systematic investigation of single BAF subunit loss in human cells. The study provides a basis for understanding subunit dependencies at several levels (Fig. 6a), the role of single subunits in BAF function (Fig. 6b,c) and the cellular consequences that occur in BAF-mutant cancers. It suggests that preferred configurations of the human SWI/SNF complexes beyond BAF/PBAF/ncBAF exist and their distribution is altered by loss of single subunits. BAF mutations ultimately result in alterations of chromatin accessibility and gene expression, which are strongly dependent on the subunit that is lost. Most prominent are the global reduction of chromatin accessibility in ARID1A^{KO}, SMARCC1^{KO} and SMARCA4^{KO} cells and an increase in accessibility following mutation of ARID1B. This revealed similar cellular consequences of loss of subunits that occupy different positions and have different functions in the complex, while mutually exclusive subunits were functionally not necessarily completely redundant. By observing clonal cell lines, we have described the long-term changes following subunit loss that allow complex rewiring and homeostasis to select for viability and proliferation, processes that may also occur in carcinogenesis. Furthermore, we observed that chromatin accessibility 5 d after SMARCA4 knockdown is already very similar to the changes observed in SMARCA4^{KO} cells. Future work will extend these studies to include kinetic resolution immediately after subunit loss, thus allowing the investigation of processes related to tumor initiation.

In addition to the known intracomplex synthetic lethalities ARID1A–ARID1B (refs. [10–12]) and SMARCA4–SMARCA2 (refs. [7–9]), we identified SMARCC1–SMARCC2, SMARCA4–ACTB and SMARCA4–ARID2 as prominent synthetic lethal interactions. Integrative analyses of the effects on complex composition, chromatin accessibility and gene expression in the synthetic lethal conditions enabled us to provide mechanistic hypotheses for explaining these intracomplex codependencies. Future studies will further characterize these synthetic lethalities, including their dependence on the levels and catalytic activities of core ATPases. Finally and most importantly, steps must be taken to test whether these results can be translated in vivo. In addition to genetic models, the development of pharmacological tools to perturb specific BAF subunits will dramatically accelerate these approaches.

Overall, our study characterizes the impact of individual subunits on BAF complex composition, their functions and intracomplex dependencies, providing a basis and potential targets towards the goal of developing targeted treatments for BAF-mutant cancers.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0477-9.

Received: 1 June 2018; Accepted: 4 June 2019;
Published online: 19 August 2019

References

1. Wang, W. et al. Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15**, 5370–5382 (1996).
2. Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447 (2015).
3. Phelan, M. L., Sif, S., Narlikar, G. J. & Kingston, R. E. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* **3**, 247–253 (1999).
4. Chandler, R. L. et al. ARID1a-DNA interactions are required for promoter occupancy by SWI/SNF. *Mol. Cell Biol.* **33**, 265–280 (2013).
5. Hodges, C., Kirkland, J. G. & Crabtree, G. R. The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb. Perspect. Med.* **6**, a026930 (2016).
6. 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).
7. Hoffman, G. R. et al. Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *Proc. Natl Acad. Sci. USA* **111**, 3128–3133 (2014).
8. Wilson, B. G. et al. Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation. *Mol. Cell. Biol.* **34**, 1136–1144 (2014).
9. Oike, T. et al. A synthetic lethality-based strategy to treat cancers harboring a genetic deficiency in the chromatin remodeling factor BRG1. *Cancer Res.* **73**, 5508–5518 (2013).
10. Helming, K. C. et al. ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat. Med.* **20**, 251–254 (2014).
11. Kelso, T. W R. et al. Chromatin accessibility underlies synthetic lethality of SWI/SNFsubunits in ARID1A-mutant cancers. *eLife* **6**, e30506 (2017).
12. Mathur, R. et al. ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat. Genet.* **49**, 296–302 (2017).
13. Mashtalir, N. et al. Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell* **175**, 1272–1288.e20 (2018).
14. Alpsoy, A. & Dykhuizen, E. C. Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *J. Biol. Chem.* **293**, 3892–3903 (2018).
15. Dutta, A. et al. Composition and function of mutant Swi/Snf Complexes. *Cell Rep.* **18**, 2124–2134 (2017).
16. Sen, P. et al. Loss of Snf5 induces formation of an aberrant SWI/SNF complex. *Cell Rep.* **18**, 2135–2147 (2017).
17. Blomen, V. A. et al. Gene essentiality and synthetic lethality in haploid human cells. *Science* **350**, 1092–1096 (2015).
18. Yan, Z. et al. PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes. *Genes Dev.* **19**, 1662–1667 (2005).
19. Michel, B. C. et al. A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nat. Cell Biol.* **20**, 1410–1420 (2018).
20. Pan, J. et al. Interrogation of mammalian protein complex structure, function, and membership using genome-scale fitness screens. *Cell Syst.* **6**, 555–568.e7 (2018).
21. Wang, X. F. et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat. Genet.* **49**, 289–295 (2017).
22. Hodges, H. C. et al. Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. *Nat. Struct. Mol. Biol.* **25**, 61–72 (2018).
23. Raab, J. R., Resnick, S. & Magnuson, T. Genome-wide transcriptional regulation mediated by biochemically distinct SWI/SNF complexes. *PLoS Genet.* **11**, e1005748 (2015).
24. Valianou, M. et al. Pharmacological inhibition of Polo-like kinase 1 (PLK1) by BI-2536 decreases the viability and survival of hamartin and tuberin

deficient cells via induction of apoptosis and attenuation of autophagy. *Cell Cycle* **14**, 399–407 (2015).
25. Wang, N. N. et al. Molecular targeting of the oncoprotein PLK1 in pediatric acute myeloid leukemia: RO3280, a novel PLK1 inhibitor, induces apoptosis in leukemia cells. *Int. J. Mol. Sci.* **16**, 1266–1292 (2015).
26. Zhao, C. L. et al. Downregulation of PLK1 by RNAi attenuates the tumorigenicity of esophageal squamous cell carcinoma cells via promoting apoptosis and inhibiting angiogenesis. *Neoplasma* **62**, 748–755 (2015).
27. Wang, W. et al. Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev.* **10**, 2117–2130 (1996).
28. Zhao, K. et al. Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell* **95**, 625–636 (1998).
29. Barretina, J. et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
30. McFarland, J. M. et al. Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat. Commun.* **9**, 4610 (2018).
31. Stransky, N. et al. Pharmacogenomic agreement between two cancer cell line data sets. *Nature* **528**, 84–87 (2015).
32. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
33. McDonald, E. R.3rd et al. Project DRIVE: a compendium of cancer dependencies and synthetic lethal relationships uncovered by large-scale, deep RNAi screening. *Cell* **170**, 577–592.e10 (2017).
34. Tsherniak, A. et al. Defining a cancer dependency map. *Cell* **170**, 564–576.e16 (2017).

Acknowledgements

We thank the Biomedical Sequencing Facility, the Proteomics and Metabolomics Facility and the Platform Austria for Chemical Biology at CeMM for their support in generating and analyzing the next-generation sequencing, proteomics or screening data, respectively. We gratefully acknowledge Horizon Discovery for providing the HAP1 cell lines, and Boehringer Ingelheim, the Superti-Furga laboratory (CeMM) and Winter laboratory (CeMM) for providing various cancer cell lines. We acknowledge the experimental support provided by J. Block and D. Donertas. Research in the Kubicek laboratory is supported by the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development, the Austrian Science Fund (FWF) F4701 and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-CoG-772437). C.B. is supported by a New Frontiers Group award of the Austrian Academy of Sciences and by an ERC Starting Grant (European Union's Horizon 2020 research and innovation programme, grant no. 679146).

Author contributions

S.S., G.B., M.P. and S.K. planned the study and designed the experiments. S.S., K.R., M.H., T.P., K.P., C.S., A.R. and B.B. performed the experiments. A.F.R., P.M., L.V., S.S. and S.K. analyzed the data. S.S. and S.K. wrote the manuscript. S.K., M.P., G.B., A.C.M., C.B. and J.M. supervised the work. S.K. provided the funding.

Competing interests

G.B. and M.P. are employees of Boehringer Ingelheim RCV GmbH & Co KG.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0477-9.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to S.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Cell culture. HAP1 WT and knockout cells were cultured in IMDM (catalog no. 21980-032; Thermo Fisher Scientific) supplemented with 10% heat-inactivated FCS (catalog no.10500056; Thermo Fisher Scientific) and 1% penicillin-streptomycin (100 U ml^{-1} penicillin, 100 µg ml^{-1} streptomycin, catalog no. 15140-122; Thermo Fisher Scientific). For the multicolor competition assay, the following Cas9-expressing cells were used: A549 cells were cultured in Ham's F-12K medium (catalog no. 21127-022; Thermo Fisher Scientific) supplemented with 10% FCS; HEK293T, MG-63 and SK-N-AS cells were cultured in DMEM (catalog no. 41965-039; Thermo Fisher Scientific) with 10% FCS; KBM-7 and HAP1 cells were cultured in IMDM with 10% FBS; MOLM-13, NB4, RCH-ACV, A-673, SK-N-MC, BT-474, K-562, THP-1, A-498, NCI-H727 (NCI-H727 [H727] ATCC CRL-5815), A-427, A-375, COR-L23 (catalog no. 92031919-1VL; Sigma-Aldrich), NCI-H1568, NCI-H522 and HCC-366 were cultured in Roswell Park Memorial Institute 1640 medium (catalog no. 21875-034; Thermo Fisher Scientific) with 10% FCS; and SK-MES-1 (ATCC HTB-58; ATCC) in BioWhittaker Minimal Essential Medium Eagle (catalog no. 12–662F; Lonza) with 1% GlutaMAX (catalog no. 35050061; Thermo Fisher Scientific) and 10% FCS. A list of cell lines used in the study is provided in Supplementary Table 1.

siRNA screen. ON-TARGETplus SMARTpool siRNA (pool of four siRNAs targeting the same gene; Dharmacon; see Supplementary Table 1) were transferred to the screening plates using an Echo 550 liquid handler (Labcyte) and dried; 30–90 min before seeding the cells, the appropriate amount of transfection reagent in Opti-MEM I Reduced Serum Medium (catalog no. 31985070; Thermo Fisher Scientific) was added to the siRNAs. In particular, 500 HAP1 cells were seeded per well of a 96-well plate and 0.0125 µM siRNA was transfected with 0.075% DharmaFECT 1 (catalog no. T-2001-03; Dharmacon). Cell viability was measured 5 d after siRNA transfection using the CellTiter-Glo Luminescent Cell Viability Assay (catalog no. G7570; Promega). After equilibrating the plates and reagent to room temperature, the CellTiter-Glo reagent was added to the wells using a Multidrop Combi Reagent Dispenser (Thermo Fisher Scientific). The readout was performed after 20 min incubation using an EnVision 2104 Multilabel Plate Reader (PerkinElmer).

Immunoblot. Cell pellets were lysed by rotating at 4 °C for 1 h in radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris pH 8, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 5 mM EDTA, 1% glycerol, 2.5 mM $MgCl_2$, 2 mM Na_3VO_4) containing 1× cOmplete, EDTA-free Protease Inhibitor Cocktail (catalog no. 4693132001; Sigma-Aldrich). After centrifuging for 10 min, 4 °C, 13,000 r.p.m., the protein content of the supernatant was measured using a Bradford assay (catalog no. A6932; AppliChem). Equal amounts of protein were loaded on acrylamide gels using 4× SDS loading buffer (0.2 M Tris pH 6.8, 40% glycerol, 4% SDS, bromophenol blue, 0.04% β-mercaptoethanol) as well as a protein ladder (Precision Plus Protein All Blue Prestained Protein Standards, catalog no. 1610373; Bio-Rad Laboratories). After gel electrophoresis, the proteins were transferred to Immobilon-FL PVDF membranes (catalog no. IPFL00005; Merck) prewetted in 100% methanol. Transfer conditions were 1.5 h at 125 mA for 12% acrylamide gels and 1.5 h at 200 mA for 7% acrylamide gels. After blocking in tris-buffered saline with Tween 20 (TBST) + 1% casein, the membranes were incubated overnight with primary antibody in 5% milk in TBST (Supplementary Table 1). The next day, membranes were washed in TBST, incubated with fluorescently labeled secondary antibodies diluted in TBST + 1% casein for 1 h at room temperature, washed again in TBST and imaged on a ChemiDoc MP (Bio-Rad Laboratories). Image analysis was performed using the Image Lab software v6.0.1 (Bio-Rad Laboratories). Where chemiluminescence detection was used, proteins were transferred to Amersham Protran 0.45 µm nitrocellulose membranes (catalog no. 10600002; GE Healthcare); membranes were blocked in 5% milk in TBST and secondary antibodies were diluted in TBST. Membranes were then developed using the Clarity Western ECL Substrate (catalog no. 170-5060; Bio-Rad Laboratories) and imaged on the ChemiDoc MP.

Nuclear extraction and immunoprecipitation. Cells were collected by scraping and washed several times with cold PBS. Cell pellets were then resuspended in 3× volume of buffer N (300 mM sucrose, 10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 0.1 mM dithiothreitol (DTT), 0.75 mM spermidine, 0.15 mM spermine, 0.1% NP-40, 50 mM NaF, 1 mM Na_3VO_4, 1 mM phenylmethylsulfonyl fluoride (PMSF) and tosyl phenylalanyl chloromethyl ketone (TPCK) protease inhibitors), vortexed and incubated on ice for 5 min. The supernatant contained the cytoplasmic fraction. Nuclear pellets were resuspended in buffer N and pelleted by centrifugation two more times. Pellets were resuspended in 1× volume of buffer C420 (20 mM HEPES pH 7.9, 420 mM NaCl, 25% glycerol, 1 mM EDTA, 1 mM EGTA, 0.1 mM DTT, 50 mM NaF, 1 mM Na_3VO_4, 1 mM PMSF and TPCK protease inhibitors) and shaken in an Eppendorf ThermoMixer at 4 °C, 1,400 r.p.m. for 30 min. The salt concentration of the supernatant obtained after centrifugation was reduced to 150 mM by diluting the sample with 20 mM HEPES pH 7.9, 50 mM NaF, 1 mM Na_3VO_4 and 1 mM PMSF. NP-40 (0.1% final concentration), 20 U ml^{-1} benzonase nuclease (catalog no. 71205-3; Novagen) and 50 ng ml^{-1} RNase A (R4875; Sigma-Aldrich) was added and the samples incubated for 15 min at 4 °C. The protein amounts were quantified using a Bradford assay. Protein lysates were incubated overnight at 4 °C with SMARCA4 (recombinant anti-BRG1; catalog no. ab110641; Abcam), ARID1A (recombinant anti-ARID1A, catalog no. ab182560; Abcam), SMARCC1/BAF155 (catalog no. 11956S; Cell Signaling Technology), SMARCC2/BAF170 (catalog no. 12760; Cell Signaling Technology), ARID2 (PA5-35857) (lot no. RJ2288871K; Thermo Fisher Scientific), SMARCE1 (recombinant anti-BAF57/SMARCE1, catalog no. ab137081; Abcam), SMARCB1 (recombinant anti-SNF5/SMARCB1, catalog no. ab126734; Abcam) or BRD9 (catalog no. ab137245; Abcam) antibody. Samples were incubated for 3 h with washed Dynabeads Protein G (catalog no. 10004D; Thermo Fisher Scientific). Captured protein complexes were washed three times with buffer I (50 mM HEPES pH 7.9, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 50 mM NaF, 1 mM Na_3VO_4, 1 mM PMSF and TPCK protease inhibitors) and twice with buffer II (50 mM HEPES pH 7.9, 150 mM NaCl, 5 mM EDTA). Proteins were eluted from the beads using SDS buffer (50 mM HEPES pH 7.9, 150 mM NaCl, 5 mM EDTA, 2% SDS). For details of the mass spectrometry samples, conditions of the mass spectrometry runs and data analyses, see the Supplementary Note.

Multicolor competition assay. gRNA oligonucleotides were phosphorylated, annealed and ligated into BbsI-linearized pKLV2-U6gRNA5(BbsI)-PGKpuro2AZsG-W (plasmid no. 67975; Addgene) or pKLV2-U6gRNA5(BbsI)-PGKpuro2AmCherry-W (plasmid no. 67977; Addgene) vectors. One Shot Stbl3 Chemically Competent *Escherichia coli* (catalog no. C7373-03; Thermo Fisher Scientific) were used for transformation and QIAGEN kits for plasmid purification. Virus was produced with HEK293T cells transfected with psPAX2 (plasmid no. 12260; Addgene), pMD2.G (plasmid no. 12259; Addgene) and the gRNA plasmid using polyethylenimine (catalog no. 23966-1; Polysciences). Adherent target cells were seeded with the virus the day before transduction using 8 µg ml^{-1} Polybrene (catalog no. sc-134220; Santa Cruz Biotechnology). Suspension cells were seeded in medium with Polybrene; virus was added and the cells were spin-transduced for 45 min at 2,000 r.p.m. The final gRNAs are listed in Supplementary Table 1. Samples were measured with an LSRFortessa flow cytometer (BD Biosciences) using B/E Alexa Fluor 488 (green fluorescent protein (GFP)) and YG/D PE-Texas Red (mCherry) 4 d, 10 d and 14 d after transduction.

Cell cycle FACS analysis. 5-Ethynyl-2′-deoxyuridine (EdU) cell cycle analyses were performed with the Click-iT Plus EdU Alexa Fluor 488 Flow Cytometry Assay Kit (catalog no. C10633; Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, cells were treated for 1 h with 10 µM EdU and then collected by trypsinization. One million cells were fixed in Click-iT fixative for 15 min at room temperature, followed by a 15 min incubation at room temperature in Click-iT saponin-based permeabilization and wash reagent. The fixed and permeabilized cells were incubated for 1 h at 37 °C with the *In Situ* Cell Death Detection Kit (catalog no. 12156792910; Sigma-Aldrich). Then, the Click-iT reaction was performed according to the manufacturer's protocol in a total volume of 100 µl (2 µl copper protectant, 0.5 µl of the fluorescent dye picolyl azide, 0.2 µl secondary antibody, 10 µl reaction buffer additive, 87.3 µl PBS) for 1 h at room temperature in the dark. After two washes with Click-iT saponin-based permeabilization and wash reagent, cells were resuspended in PBS with 5 µg ml^{-1} 4,6-diamidino-2-phenylindole (DAPI). Samples were measured with an LSRFortessa using Horizon V450 (DAPI), B/E Alexa Fluor 488 (EdU) and YG/D PE-Texas Red (TMR-red).

ATAC-seq. ATAC-seq was performed according to Buenrostro et al.[35] with small adaptations. Briefly, 50,000 cells were resuspended in 25 µl transposase reaction mixture (0.05% digitonin, 1× tagmentation DNA buffer, 0.08% tagment DNA enzyme (Nextera DNA Library Prep Kit, catalog no. FC-121-1031; Illumina)) and incubated for 30 min at 37 °C, 300 r.p.m. Then, DNA was purified using the MinElute kit (catalog no. 28004; QIAGEN) and eluted in 11 µl elution buffer; 1 µl was used to determine the cycle number for PCR using a quantitative PCR approach. The remaining 10 µl were complemented with 1× NEBNext High-Fidelity 2X PCR Master Mix (catalog no. M0541; New England BioLabs), 1.25 µM index primer 1 and 1.25 µM index primer containing a barcode (Supplementary Table 1). PCR was performed: 5 min 72 °C, 30 s 98 °C, X cycles of 10 s 98 °C + 30 s 63 °C + 1 min 72 °C, 1 min 72 °C and then cleaned up using Agencourt AMPure XP beads (catalog no. A63880; Beckman Coulter Life Sciences). The libraries were checked on a Bioanalyzer instrument using High-Sensitivity DNA Chips (Agilent Technologies), quantified using a Qubit dsDNA HS Assay kit (catalog no. Q32854; Thermo Fisher Scientific) and pooled. Libraries were sequenced using the HiSeq 3000/4000 platform (Illumina) with the 75 base pair (bp) paired-end configuration.

RNA-seq. RNA was isolated using the RNeasy Mini kit (catalog no. 74106; QIAGEN). The amount of total RNA was quantified with the Qubit 2.0 Fluorometric Quantitation system (Thermo Fisher Scientific); and the RNA integrity number was determined with the Experion Automated Electrophoresis System (Bio-Rad Laboratories). RNA-seq libraries were prepared with the TruSeq Stranded mRNA LT sample preparation kit (Illumina) using Sciclone and Zephyr liquid handling robotics (PerkinElmer). The library amount was quantified with

the Qubit 2.0 Fluorometric Quantitation system and the size distribution was assessed with the Experion Automated Electrophoresis System. Libraries were pooled, diluted and sequenced with the HiSeq 3000/4000 platform with the 50 bp single-read configuration.

ChIP-seq. Cells were fixed in Pierce 1% formaldehyde (catalog no. 28908; Thermo Fisher Scientific) for 10 min at room temperature. The fixation was quenched using 0.125 M glycine, pH 2.5, for 5 min at 4 °C. After several washes with PBS, cells were collected by scraping. Cells were then resuspended in buffer L1 (50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, pH 8.0, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100) and incubated for 10 min on ice. After centrifugation, the pellet was resuspended and incubated for 5 min on ice in buffer L2 (200 mM NaCl, 1 mM EDTA, pH 8.0, 0.5 mM EGTA, pH 8.0, 10 mM Tris, pH 8.0). The pellet was resuspended in buffer L3 (1 mM EDTA, pH 8.0, 0.5 mM EGTA, pH 8.0, 10 mM Tris, pH 8.0, 100 mM NaCl, 0.1% Na-deoxycholate, 0.17 mM N-lauroyl sarcosine, 1× protease inhibitor) and washed twice in shearing buffer (10 mM Tris-HCl, pH 7.8, 1 mM EDTA, pH 8.0, 0.1% SDS). Then the chromatin was sheared in harsh shearing buffer (10 mM Tris-HCl, pH 7.8, 2 mM EDTA, pH 8, 0.25% SDS, 1× protease inhibitor) using a Covaris S220 (duty cycle: 5%; intensity: 4; peak incident power; 140 W; cycles per burst: 200; time: 30 min). After centrifugation, the supernatant was diluted 1:1.5 in equilibration buffer (10 mM Tris-HCl, pH 8.0, 223 mM NaCl, 1.66% Triton X-100, 0.166% Na-deoxycholate, 1 mM EDTA, pH 8, 1× protease inhibitor). Antibody incubation was performed overnight at 4 °C, followed by a 3-h incubation with Dynabeads Protein G. The beads were then washed twice in RIPA-LS (10 mM Tris-HCl, pH 8.0, 140 mM NaCl, 1 mM EDTA, pH 8.0, 0.1% SDS, 0.1% Na-deoxycholate, 1% Triton X-100), RIPA-HS (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0, 500 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate) and then RIPA-LiCl/Na-deoxycholate (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate). After transferring the beads in TE buffer to a new tube, the bound chromatin was eluted in elution buffer (1% SDS, 0.1 M $NaHCO_3$) twice for 20 min at room temperature. The eluted material was incubated with RNase for 30 min at 37 °C, with proteinase K for 2.5 h at 55 °C and decrosslinked overnight at 65 °C. DNA was then extracted using phenol–chloroform, precipitated and then dissolved in TE buffer. Libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (catalog no. E7645S; New England Biolabs) and sequenced using the Illumina HiSeq 3000/4000 platform with the 50 bp single-read configuration.

Histone and transcriptional repressor CTCF ChIP-seq data were generated using the ChIPmentation method as described previously[36].

ATAC-seq and ChIP-seq data preprocessing. Reads were trimmed using Skewer v0.1.124 (ref. [37]) and aligned to the GRCh37/hg19 assembly of the human genome using Bowtie 2 v2.2.9 (ref. [38]) with the '-very-sensitive' parameter. Duplicate reads were removed using sambamba v0.6.7 (ref. [39]); only properly paired reads with a mapping quality > 30 and alignment to the nuclear genome were kept. All downstream analyses were performed on the filtered reads. For visualization exclusively, we generated genome browser tracks with the genomeCoverageBed command in BEDTools v2.20.1 (ref. [40]) and normalized so that each value represents the read count per base pair per million mapped and filtered reads. This was done for each sample individually and for the merged replicates. ATAC-seq peak calling was performed with MACS2 v2.1.2 (ref. [41]) using the '-nomodel' and '-extsize 147' parameters; peaks overlapping blacklisted features as defined by the ENCODE project[42] were discarded. For ChIP-seq and ChIPmentation, library quality was assessed with the Phantompeakqualtools scripts[43]; we used HOMER findPeaks v4.8 (ref. [44]) in 'factor' mode to call peaks with matched immunoglobulin G controls as background. We used ChromHMM v1.18 (ref. [45]) to segment the genome into 12 states using 6 ChIP-seq histone marks (H3K27ac, H3K27me3, H3K36me3, H3K4me3, H3K4me1, H3K9me3) and ATAC-seq, creating HAP1-specific chromatin states.

RNA-seq data preprocessing. Base calls provided by the Illumina Realtime Analysis software v2.7.6 and v2.7.7 were converted into BAM format using Illumina2bam and demultiplexed using BamIndexDecoder (https://github.com/wtsi-npg/illumina2bam). Reads were trimmed with Trimmomatic v0.36 (ref. [46]) and aligned to the GRCh37/hg19 assembly of the human genome using Bowtie1 v1.1.2 (ref. [47]) with the following parameters: '-q -p 6 -a -m 100 -minins 0 -maxins 5000 -fr -sam -chunkmbs 200'. Duplicate reads were removed with the Picard MarkDuplicates utility with standard parameters before transcript quantification with BitSeq v0.7.5 (ref. [48]) using the Markov chain Monte Carlo method and standard parameters. To obtain gene-level quantifications, we assigned the expression values of its highest expressed transcript to each gene. Differential gene-level expression between each knockout and WT was performed using DESeq2 v1.20.0 (ref. [49]) from the raw count data with a significance threshold of 0.05.

Bioinformatic analyses. A consensus map of chromatin accessibility sites was created by merging a 1-kb window around the summit of ATAC-seq peaks from all samples using the BEDTools merge command[40]. The chromatin accessibility of each region in each sample was quantified using Pysam v0.15.2, counting the number of reads from the filtered BAM file that overlapped each region. To normalize read counts across samples, we performed quantile normalization using the normalize.quantiles function from the preprocessCore package v1.42.0 in R. We annotated each region with the identity of and distance to the nearest TSS and the overlap with Ensembl gene annotations (promoters were defined as the 2,500-bp region upstream of the TSS). Annotation with chromatin states was based on the 12-state ChromHMM genome segmentation on ChIP-seq marks.

Unsupervised analysis was performed with the scikit-learn library v0.19.2 (ref. [50]) for principal component analysis (PCA; sklearn.decomposition.PCA) applied to the chromatin accessibility values in the consensus map for all ATAC-seq samples. Supervised analysis was performed using DESeq2 (ref. [49]) comparing chromatin accessibility between each knockout and WT using the raw count data. We considered a region differential if it had a false discovery rate-adjusted $P < 0.01$ and absolute $\log_2$(fold change) > 1. Region set enrichment analysis was performed on the significant regions of each group using location overlap analysis (LOLA)[51] with its core databases: transcription factor binding sites from ENCODE[42]; tissue clustered DNase hypersensitive sites[52]; the CODEX 2016 database[53]; University of California, Santa Cruz Genome Browser annotation tracks[54]; the Cistrome 2015 database[55]; and data from the BLUEPRINT project[56]. Motif enrichment analysis was performed with HOMER findMotifs[44] and the AME tool from the MEME v4.10.2 suite[57] using 250 bp sequences centered on chromatin-accessible regions.

To assess the relationship between chromatin accessibility differences and nucleosome occupancy, data were pooled between replicates and we used NucleoATAC v0.3.4 (ref. [58]) with default parameters to output nucleosome signal, and call nucleosome dyads and nucleosome-free regions. To get a complementary view on the relationship between ATAC-seq signal and nucleosome position, we further split the ATAC-seq fragments in bins depending on their size: nucleosome free fragments up to 100 bp; and nucleosome fragments between 180 bp and 247 bp. We then quantified the bulk ATAC-seq signal, NucleoATAC smooth nucleosome occupancy signal, NucleoATAC predicted dyad positions and signal from nucleosome-free fragments and nucleosome-associated fragments in the sets of differential chromatin regions discovered with DESeq2, in the center of the CTCF binding sites and the center of all regions in the consensus chromatin accessibility region set.

In figures where the $\log_2$(fold change) of chromatin accessibility is reduced to one value per gene, we took the maximum or minimum value of change for the regulatory elements associated with a gene if all values agreed in the direction of change (positive or negative) by a two-thirds majority or otherwise the mean of all values.

To detect differential BAF- versus PBAF-bound sites, we used the ChIP-seq peaks from samples of BAF complex members (ARID1A, ARID1B, ARID2, PBRM1, SMARCA2 and SMARCA4) and created a consensus region set as for ATAC-seq data. The signal was quantified and normalized in the same fashion. We calculated the mean signal intensity and $\log_2$(fold change) between ARID2 and ARID1A ChIPs and standardized the fold change in 250 bins along the mean. We then fitted a bivariate Gaussian kernel (scipy.stats.gaussian_kde) on the two variables and selected differential sites as the ones in the 5th percentile of density, with an absolute normalized $\log_2$(fold change) > 1.5.

Nomenclature. For clarity, gene names were also used for proteins. A list of alternative names can be found in the Supplementary Note.

Statistics and reproducibility. The statistical tests used are described in the respective figure legends and Methods. Additional information regarding sample size and statistics is available in Supplementary Table 6.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Next-generation sequencing data have been deposited with the NCBI GEO (GSE108390). Mass spectrometry data have been deposited with the PRIDE archive PXD013102. The processed data used for the analyses are available at http://baf-complex.computational-epigenetics.org.

Code availability

Software used for the analyses is available at http://baf-complex.computational-epigenetics.org.

References

35. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 1–9 (2015).
36. Schmidl, C., Rendeiro, A. F., Sheffield, N. C. & Bock, C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* **12**, 963–965 (2015).
37. Jiang, H., Lei, R., Ding, S. W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
38. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

39. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
40. Quinlan, A. R. BEDTools: the Swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinformatics* **47**, 1–34 (2014).
41. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
42. Hoffman, M. M. et al. Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res.* **41**, 827–841 (2013).
43. Landt, S. G. et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* **22**, 1813–1831 (2012).
44. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
45. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* **28**, 817–825 (2010).
46. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
47. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
48. Glaus, P., Honkela, A. & Rattray, M. Identifying differentially expressed transcripts from RNA-seq data with biological variation. *Bioinformatics* **28**, 1721–1728 (2012).
49. 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).
50. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
51. Sheffield, N. C. & Bock, C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics* **32**, 587–589 (2016).
52. Sheffield, N. C. et al. Patterns of regulatory activity across diverse human cell types predict tissue identity, transcription factor binding, and long-range interactions. *Genome Res.* **23**, 777–788 (2013).
53. Sánchez-Castillo, M. et al. CODEX: a next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. *Nucleic Acids Res.* **43**, D1117–D1123 (2015).
54. Rosenbloom, K. R. et al. The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res.* **43**, D670–D681 (2015).
55. Liu, T. et al. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol.* **12**, R83 (2011).
56. Adams, D. et al. BLUEPRINT to decode the epigenetic signature written in blood. *Nat. Biotechnol.* **30**, 224–226 (2012).
57. Bailey, T. L. et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
58. Schep, A. N. et al. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res.* **25**, 1757–1770 (2015).

nature research

Corresponding author(s): Stefan Kubicek

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	- Viability measurement: 2104 EnVisionTM Multilabel Plate Reader (PerkinElmer) with Wallac EnVision Manager Version 1.12 - Western Blot: Bio-Rad ChemiDoc™ MP with Image LabTM Touch Software Version 2.3.0.07 - IP-MS: LTQ-Orbitrap Velos Agilent 1200 nano, Q Exactive Agilent 1200 nano, Fusion Lumos U3000 RSLCnano with Xcalibur version 4.0.0 and Tune 2.1.1565.24 - Flow cytometry: BD LSRFortessaTM flow cytometer (BD Biosciences) with BD FACSDiva Software Version 8.0.1 - NGS: Illumina HiSeq3000/4000 platform (HCS 3.3.76 and HCS HD 3.4.0.38), Illumina Realtime Analysis RTA v2.7.6 and v2.7.7
Data analysis	Codes used in the study are described in the methods section and a link to GitHub is further available under http://baf-complex.computational-epigenetics.org, where all the data and information is deposited. - Western Blot: Bio-Rad Image Lab software version 6.0.1 - IP-MS: Proteome Discoverer 2.2.0.388 platform, Percolator V3.0, isobar package version 1.30.0 - Flow cytometry: FlowJO_V10 - NGS: o BamIndexDecoder: 1.00 o bedtools: 2.20.1 o BitSeq: 0.7.5 o bowtie1: 1.1.2

o bowtie2: 2.2.9
o ChromHMM: 1.18
o DESeq2: 1.20.0
o HOMER: 4.8
o Illumina2bam: 0.03
o LOLA: 1.14.0
o MACS2: 2.1.2
o MEME: 4.10.2
o NucleoATAC: 0.3.4
o phantomPeakQualtools: ccdfd163d1fce94e9ba95fbfb5a4f1cc557547bb
o Picard: 2.6.0
o preprocessCore: 1.42.0
o pysam: 0.15.2
o sambamba: 0.6.7
o skewer: 0.1.124
o sklearn: 0.20.0
o trimmomatic: 0.36
- Structure analysis from IP-MS data:
o Julia: 0.6.2
o BioAlignments: 0.2.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data are available through the Supplementary Website (http://baf-complex.computational-epigenetics.org/). The ATAC-seq, ChIP-seq and RNA-seq data are available from NCBI GEO (accession number: GSE108390). The proteomics data is available via PRIDE: PXD013102.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	IP-MS and ATAC-seq were performed in two independent experiments (exception: IP-MS experiments shown in supplementary figure S3c and S12e were only one replicate), ChIP-seq data were one or more replicates depending on the used antibody and cell type, RNA-seq was measured in biological duplicate or triplicates, siRNA screen was performed in two biological independent experiments each with two or more independent samples. There was in general good correlation between replicates for the different experiments justifying the chosen sample size.
Data exclusions	Outliers were excluded in the siRNA screen data analysis as well as one clone for the SMARCA4 and ARID1A knock-out cells.
Replication	The experiments were performed in biological independent replicates that generally showed very high correlations. Furthermore, we included two different ARID1A and SMARCA4 knock-out clones in the study that revealed similar changes.
Randomization	Randomization was not relevant to this study as all experiments were performed with molecular biological techniques. Samples per replicate were processed together (exception: ATAC- and RNA-seq of SMARCC2 KO clone was performed later; HAP1 WT and SMARCC1 KO cells were repeated with the SMARCC2 KO clone in order to correct for batch effects). In IP-MS experiments, it was not possible to process all samples together. Here the samples were divided into three sets, trying to keep knock-outs of mutually exclusive genes in one set.
Blinding	Blinding was not relevant to this study as all experiments were performed with molecular biological techniques.

Reporting for specific materials, systems and methods

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
☒	☐ Unique biological materials	☐	☒ ChIP-seq
☐	☒ Antibodies	☐	☒ Flow cytometry
☐	☒ Eukaryotic cell lines	☒	☐ MRI-based neuroimaging
☒	☐ Palaeontology		
☒	☐ Animals and other organisms		
☒	☐ Human research participants		

Antibodies

Antibodies used

target company order number lot purpose used for
SMARCA2 Cell Signaling CST6889 1 Western Blot
SMARCA4 Santa Cruz SC-17796 G3013/A2216 Western Blot
BRD7 Santa Cruz SC-376180 A1513 Western Blot
BRD9 Abcam ab137245 GR257571-5 Western Blot, IP
PBRM1 Abcam ab196022 GR195933-3 Western Blot
ARID1A Cell Signaling CST12354 1 Western Blot
ARID1B Abcam ab57461 GR129020-2 Western Blot
ARID2 Santa Cruz SC-166117 J1515 Western Blot
SMARCD1 BD Biosciences BD611728 4003712 Western Blot
SMARCD3 Abcam ab194671 GR210283-2 Western Blot
PHF10 Abcam ab154637 GR116191-12 Western Blot
DPF2 Abcam ab128149 GR250670-3 Western Blot
SMARCC1 Cell Signaling CST11956 1,2 Western Blot, IP
SMARCC2 Cell Signaling CST12760 1 Western Blot, IP
ACTL6A Abcam ab131272 GR90068-11 Western Blot
ACTL6B Abcam ab140642 GR103982-6 Western Blot
SMARCE1 Abcam ab137081 GR106541-6 Western Blot, IP
BCL11A Abcam ab19487 GR221533-8 Western Blot
SMARCB1 Abcam ab126734 GR177618-2, Y1080106CS Western Blot, IP
BRD4 abcam ab128874 GR275920-39 Western Blot
SS18 Abcam ab179927 GR186578-5 Western Blot
RCC1 Santa Cruz SC-55559 C1307 Western Blot
Cas9 Novus Biologicals NBP2-36440 F-1 Western Blot
α-Tubulin Abcam ab7291 GR3197113-5 Western Blot
hFAB™ Rhodamine Anti-Tubulin Primary Antibody Bio-Rad 12004165 Western Blot
StarBright™ Blue 700 Goat Anti-Rabbit IgG Bio-Rad 12004161 Western Blot
StarBright™ Blue 700 Goat Anti-Mouse IgG Bio-Rad 12004158 Western Blot
StarBright™ Blue 520 Goat Anti-Mouse IgG Bio-Rad 12005866 Western Blot
anti-mouse IgG (HRP) Jackson ImmunoResearch 715-035-151 Western Blot
anti-rabbit IgG (HRP) Jackson ImmunoResearch 711-035-152 Western Blot
ACTB Abcam ab8227 GR288171-1 Western Blot
BCL7C Abcam ab126944 GR295371-1 Western Blot
SMARCD2 Abcam ab166622 GR163113-2 Western Blot
SMARCD2 Santa Cruz SC-102119 I1010 Western Blot
SMARCA4 Abcam ab110641 GR150844-22 ChIP, IP
SMARCA2 Cell Signaling CST11966 Lot1 ChIP
ARID1B Abcam Ab57461 GR12902-1 ChIP
ARID2 Thermo Fisher Scientific PA5-35857 RJ2288871K ChIP, IP
ARID2 Novus Biologicals NBP1-26615 A1 ChIP
PBRM1 Novus Biologicals NB100-79832 A2 ChIP
ARID1A Abcam ab182560 GR269670-9 ChIP, IP
IgG Santa Cruz SC-2025 ChIP
H3K27ac Diagenode pAb-196-050 (C15410196) A1723-0041D ChIP
H3K27me3 Millipore 07-449 ChIP
H3K4me1 Diagenode pAb-194-050 (C15410194) ChIP
H3K36me3 Diagenode pAb-192-050 A1847-001P ChIP
CTCF Millipore 07-729 ChIP
H3K9me3 Diagenode pAb-193-050 (C15410193) A1671-001P ChIP
H3K4me3 Diagenode pAb-003-050 (C15410003) A5051-001P ChIP

POLR2A (N-20) Santa Cruz SC-899 G2914 ChIP

Validation

The antibodies for the different BAF subunits were validated using Western Blot by staining WT and HAP1 knock-out cells (exception as no knock-out clones were available: ACTL6A, ACTL6B, SMARCE1, SMARCB1, BRD4, SS18).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

HAP1 cells were obtained from Horizon discovery; SK-MES-1 and NCIH727 cells were bought from ATCC; A549, HEK293T, MG-63, SKNAS, KBM7, HAP1, MOLM13, NB4, RCH-ACV, A673, SK-N-MC, BT474, K562, THP1, A498, A427, A375, NCIH1568, NCIH522, HCC366 cells were provided by Boehringer Ingelheim, the Superti-Furga, and Winter lab; CORL23 were bought from Sigma-Aldrich.

Authentication

The HAP1 KO cells were validated on the genetic level and on protein level. HAP1 ACTB, SMARCD2, DPF1, DPF3, BCL7A/B, BCL11A/B KO clones could not be validated by Western Blot due to the unavailability of highly specific antibodies. However, ACTB, SMARCD2, DPF1/3, and BCL7A/B KO cells showed reduced expression of the knocked-out gene relative to WT cells in the RNA-seq and/or IP-MS data.

Mycoplasma contamination

The cell lines were tested negative for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

None on the cell lines we use are on the current ICLAC version 9 list of commonly misidentified lines

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

NCBI GEO repository: GSE108390
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108390
The data is further available here: http://baf-complex.computational-epigenetics.org

Files in database submission

GSM2897151 ChIPmentation_HAP1_WT_CTCF_921-7_r1
GSM2897152 ChIPmentation_HAP1_WT_H3K27ac_1102-3_r1_t2
GSM2897153 ChIPmentation_HAP1_WT_H3K27ac_1102-3_r1_t3
GSM2897154 ChIPmentation_HAP1_WT_H3K27ac_1102-3_r1_t6
GSM2897155 ChIPmentation_HAP1_WT_H3K27ac_1102-3_r2
GSM2897156 ChIPmentation_HAP1_WT_H3K27ac_1102-3_r3
GSM2897157 ChIPmentation_HAP1_WT_H3K27ac_921-7_r2
GSM2897158 ChIPmentation_HAP1_WT_H3K27me3_1102-3_r1
GSM2897159 ChIPmentation_HAP1_WT_H3K27me3_921-7_r2
GSM2897160 ChIPmentation_HAP1_WT_H3K36me3_1102-3_r1
GSM2897161 ChIPmentation_HAP1_WT_H3K36me3_921-7_r1
GSM2897162 ChIPmentation_HAP1_WT_H3K36me3_921-7_r2
GSM2897163 ChIPmentation_HAP1_WT_H3K4me1_1102-3_r2
GSM2897164 ChIPmentation_HAP1_WT_H3K4me1_921-7_r1
GSM2897165 ChIPmentation_HAP1_WT_H3K4me1_921-7_r2
GSM2897166 ChIPmentation_HAP1_WT_H3K4me3_1102-3_r1
GSM2897167 ChIPmentation_HAP1_WT_H3K4me3_921-7_r1
GSM2897168 ChIPmentation_HAP1_WT_H3K4me3_921-7_r2
GSM2897169 ChIPmentation_HAP1_WT_H3K9me3_1102-3_r1
GSM2897170 ChIPmentation_HAP1_WT_H3K9me3_921-7_r1
GSM2897171 ChIPmentation_HAP1_WT_H3K9me3_921-7_r2
GSM2897172 ChIPmentation_HAP1_WT_IgG_1102-3_r1
GSM2897173 ChIPmentation_HAP1_WT_IgG_921-7_r1
GSM2897174 ChIPmentation_HAP1_WT_IgG_921-7_r2
GSM2897175 ChIP-seq_HAP1_WT_ARID1A_r1
GSM2897176 ChIP-seq_HAP1_WT_ARID1A_r2
GSM2897177 ChIP-seq_HAP1_WT_ARID1B_r1
GSM2897178 ChIP-seq_HAP1_WT_ARID2_NB_r1
GSM2897179 ChIP-seq_HAP1_WT_ARID2_PA5_r2
GSM2897180 ChIP-seq_HAP1_WT_BRD4_r1
GSM2897181 ChIP-seq_HAP1_WT_H3K27ac_r1
GSM2897182 ChIP-seq_HAP1_WT_IgG_r1
GSM2897183 ChIP-seq_HAP1_WT_IgG_r2
GSM2897184 ChIP-seq_HAP1_WT_Input_1x_r1
GSM2897185 ChIP-seq_HAP1_WT_Input_20_r1
GSM2897186 ChIP-seq_HAP1_WT_Input_25_r2
GSM2897187 ChIP-seq_HAP1_WT_Input_8x_r1

GSM2897188 ChIP-seq_HAP1_WT_PBRM1_r1
GSM2897189 ChIP-seq_HAP1_WT_PolII_r1
GSM2897190 ChIP-seq_HAP1_WT_SMARCA2_r1
GSM2897191 ChIP-seq_HAP1_WT_SMARCA4_r1
GSM2897192 ChIP-seq_HAP1_WT_SMARCA4_r2
GSM3755164 ChIP-seq_HAP1_ARID1A_ARID1A_r1
GSM3755165 ChIP-seq_HAP1_ARID1A_Input_r1
GSM3755166 ChIP-seq_HAP1_ARID1B_ARID1A_r1
GSM3755167 ChIP-seq_HAP1_ARID1B_Input_r1
GSM3755168 ChIP-seq_HAP1_BCL7B_ARID1A_r1
GSM3755169 ChIP-seq_HAP1_BCL7B_Input_r1
GSM3755170 ChIP-seq_HAP1_SMARCA2_ARID1A_r1
GSM3755171 ChIP-seq_HAP1_SMARCA2_Input_r1
GSM3755172 ChIP-seq_HAP1_SMARCA4_ARID1A_r1
GSM3755173 ChIP-seq_HAP1_SMARCA4_Input_r1
GSM3755174 ChIP-seq_HAP1_SMARCC1_ARID1A_r1
GSM3755175 ChIP-seq_HAP1_SMARCC1_Input_r1
GSM3755176 ChIP-seq_HAP1_SMARCC2_ARID1A_r1
GSM3755177 ChIP-seq_HAP1_SMARCC2_Input_r1

Genome browser session
(e.g. UCSC)

http://baf-complex.computational-epigenetics.org

Methodology

Replicates

6 replicates: H3K27ac
3 replicates: H3K4me1, H3K4me3, H3K9me1, H3K36me3, IgG (ChIPmentation)
2 replicates: H3K27me3, ARID1A, ARID2 (two different antibodies), SMARCA4, IgG (ChIP)
1 replicate: ARID1B, BRD4, CTCF, PBRM1, PolII, SMARCA2, ARID1A ChIP in knock-out cells

Sequencing depth

Average fastqc_total_pass_filter_reads = 44897837

Antibodies

SMARCA4 Abcam ab110641
SMARCA2 Cell Signaling CST11966
ARID1B Abcam Ab57461
ARID2 Thermo Fisher Scientific PA5-35857
ARID2 Novus Biologicals NBP1-26615
PBRM1 Novus Biologicals NB100-79832
ARID1A Abcam ab182560
IgG Santa Cruz Sc-2025
H3K27ac Diagenode pAb-196-050
H3K27me3 Millipore 07-449
H3K4me1 Diagenode pAb-194-050
H3K36me3 Diagenode pAb-192-050
CTCF Millipore 07-729
H3K9me3 Diagenode pAb-193-050
H3K4me3 Diagenode pAb-003-050

Peak calling parameters

For ChIP-seq and ChIPmentation, library quality was assessed with the phantomPeakQualtools/spp scripts, and we used HOMER findPeaks, version 4.8 in “factor” mode to call peaks with matched IgG controls as background.

Data quality

Raw sequencing quality was assessed with FASTQC, cross-correlation enrichment metrics were calculated with SPP (https://github.com/hms-dbmi/spp).

Software

SPP (https://github.com/hms-dbmi/spp; commit bd0ec7577afea637066ad608777e52c39e7c7904) was used for sample quality control. Custom code for additional analysis in ATAC-seq peaks was used and is provided in a Github repository (https://github.com/epigen/baf_complex).

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Multicolor competition assay: trypsinized adherent or collected suspension cells were resuspended in PBS with FBS. EdU cell cycle FACS: supernatant was collected and combined with trypsinized cells before fixation
Instrument	BD LSRFortessa flow cytometer (BD Biosciences)
Software	FlowJo was used to analyze the data.
Cell population abundance	Multicolor competition assay: Each sample contained untransduced (no color), single transduced (GFP or mCherry) or double transduced (GFP and mCherry) cells. The percentage of each population varied from sample to sample and changed during the time-course. EdU cell cycle FACS: The smallest population measured were cells in subG1 phase (low DAPI, low EdU) compared to cells in G1 or G2 phase (high DAPI, low EdU) or to cells in S phase (high DAPI, high EdU).
Gating strategy	Multicolor competition assay: The samples were gated for alive cells based on the FSC and SSC and then for single cells (SSC-H vs SSC-A). The gated cells were plotted for their fluorescence (YG_D PE Texas Red (mCherry) versus B/E Alexa Fluor 488 (GFP)) and four quadrant gating was used to determine the different populations. EdU cell cycle FACS: The samples were gated for alive cells based on the FSC and SSC and then for single cells (SSC-H vs SSC-A). The gated cells were plotted for their DNA content (V/C Horizon V450 (DAPI)) versus EdU incorporation (B/E Alexa Fluor 488 (EdU)).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.