LETTER

https://doi.org/10.1038/s41586-019-1646-9

Spliceosomal disruption of the non-canonical BAF complex in cancer

Daichi Inoue[1,2,10], Guo-Liang Chew[3,4,10], Bo Liu[1], Brittany C. Michel[5,6,7], Joseph Pangallo[3,4], Andrew R. D'Avino[6,7], Tyler Hitchman[1], Khrystyna North[3,4,8], Stanley Chun-Wei Lee[1], Lillian Bitner[1], Ariele Block[1], Amanda R. Moore[1], Akihide Yoshimi[1], Luisa Escobar-Hoyos[1], Hana Cho[1], Alex Penson[1], Sydney X. Lu[1], Justin Taylor[1], Yu Chen[1,9], Cigall Kadoch[6,7], Omar Abdel-Wahab[1,9]* & Robert K. Bradley[3,4,8]*

SF3B1 is the most commonly mutated RNA splicing factor in cancer[1–4], but the mechanisms by which *SF3B1* mutations promote malignancy are poorly understood. Here we integrated pan-cancer splicing analyses with a positive-enrichment CRISPR screen to prioritize splicing alterations that promote tumorigenesis. We report that diverse *SF3B1* mutations converge on repression of BRD9, which is a core component of the recently described non-canonical BAF chromatin-remodelling complex that also contains GLTSCR1 and GLTSCR1L[5–7]. Mutant SF3B1 recognizes an aberrant, deep intronic branchpoint within *BRD9* and thereby induces the inclusion of a poison exon that is derived from an endogenous retroviral element and subsequent degradation of *BRD9* mRNA. Depletion of BRD9 causes the loss of non-canonical BAF at CTCF-associated loci and promotes melanomagenesis. BRD9 is a potent tumour suppressor in uveal melanoma, such that correcting mis-splicing of *BRD9* in *SF3B1*-mutant cells using antisense oligonucleotides or CRISPR-directed mutagenesis suppresses tumour growth. Our results implicate the disruption of non-canonical BAF in the diverse cancer types that carry *SF3B1* mutations and suggest a mechanism-based therapeutic approach for treating these malignancies.

SF3B1 is subject to recurrent missense mutations at specific residues in myeloid leukaemia[1,2] and lymphoid leukaemia[3,8] as well as in solid tumours, at rates of up to 14–29% for uveal melanoma (UVM)[9–12] and 65–83% for myelodysplastic syndromes with ring sideroblasts[1,2]. Consistent with the critical role of SF3B1 in the recognition of 3′ splice sites[13], several previous studies have reported that *SF3B1* mutations induce widespread usage of abnormal 3′ splice sites[10,14,15]. Although many mis-spliced genes have been identified in *SF3B1*-mutant samples, few of these have been functionally implicated in driving disease.

We hypothesized that effectors of the pro-tumorigenic consequences of *SF3B1* mutations might appear as pan-cancer targets of mutant SF3B1. We accordingly identified mis-spliced events that were shared between erythroleukaemic (K562) and UVM (MEL270) cells that expressed wild-type *SF3B1* or the most-common *SF3B1* mutation, $SF3B1^{K700E}$. A compact set of 40 events exhibited concordant splicing changes, and was sufficient to infer *SF3B1* mutational status across 249 samples from patients with chronic lymphocytic leukaemia, myelodysplastic syndromes and UVM (Fig. 1a, Extended Data Fig. 1a, Supplementary Tables 1–3).

We designed a single-guide RNA (sgRNA) library that targeted both pan-cancer and cancer-type-specific targets of mutant SF3B1, focusing on genes for which *SF3B1* mutations are predicted to cause mis-splicing that triggers nonsense-mediated RNA decay (NMD) (Fig. 1b, Supplementary Table 4). We tested whether the knockout of any of these genes promoted the transformation of Ba/F3 cells (a mouse cell line with a wild-type spliceosome, with a requirement for IL-3 that can be overcome by oncogenic lesions) (Fig. 1c). In addition to the positive control *Pten*, our screen revealed that the loss of *Brd9* promoted the transformation of Ba/F3 cells (Fig. 1d, Extended Data Fig. 1b–d, Supplementary Tables 5, 6). *Brd9* was a notable hit because *BRD9* exhibited notable mis-splicing in all cohorts of patients with *SF3B1*-mutant cancer (Fig. 1e). *Brd9* knockout conferred cytokine independence to mouse 32Dcl3 cells, and growth advantage to human cancer cells with a wild-type spliceosome derived from UVM, cutaneous melanoma, and pancreatic cancer (Extended Data Fig. 1d–f). By contrast, acute myeloid leukaemia cells with rearranged *MLL* (also known as *KMT2A*) required BRD9 for growth (Extended Data Fig. 1g), as previously reported[16].

SF3B1 mutations cause the exonization of a *BRD9* intronic sequence, which results in the inclusion of a poison exon that interrupts the open reading frame of *BRD9*. This *BRD9* poison exon is derived from a primate-specific endogenous retroviral element, explaining its absence from mice (Extended Data Fig. 1h, i). We confirmed that the inclusion of the poison exon was induced by the expression of endogenous or ectopic mutant SF3B1 in K562 and NALM-6 cells, whereas *SF3B1* knockdown in *SF3B1* wild-type cells had no effect (Extended Data Fig. 1j–m). The poison exon was included in an *SF3B1*-mutation-dependent manner in diverse cell lines and in samples of chronic lymphocytic leukaemia, myelodysplastic syndromes and UVM that bear 19 different *SF3B1* mutations—but not in healthy tissues (Extended Data Fig. 1m–p, Supplementary Table 7).

The inclusion of the *BRD9* poison exon triggered NMD and reduced the half-life of *BRD9* mRNA and steady-state levels of full-length BRD9 protein (Extended Data Fig. 1q–w). Patients with *SF3B1* mutations exhibited reduced total levels of *BRD9* mRNA relative to patients with wild-type *SF3B1* (Extended Data Fig. 1x). We tested whether the inclusion of the poison exon could result in the production of C-terminally truncated BRD9 by knocking an N-terminal haemagglutinin tag into the *BRD9* locus in MEL270 and K562 cells that transgenically express wild-type or mutant SF3B1 (Extended Data Fig. 2a–c). Mutant SF3B1 suppressed levels of full-length BRD9 protein, without generating a truncated BRD9 protein (Fig. 1f).

SF3B1 mutations promote the use of cryptic 3′ splice sites[10,14,15], probably by altering the normal role of SF3B1 in branchpoint recognition[17]. We therefore mapped the *BRD9* branchpoints used in K562, MEL270 and T47D (breast cancer) cells that express mutant SF3B1 (Fig. 2a, Extended Data Fig. 2d–f). The inclusion of the poison exon was associated with an unusually close branchpoint (close branchpoints are rare and normally inefficiently recognized[18]). Mutating the aberrant branchpoint abolished poison exon recognition (Fig. 2b, Extended Data

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [2]Department of Hematology-Oncology, Institute of Biomedical Research and Innovation, Foundation for Biomedical Research and Innovation at Kobe, Kobe, Japan. [3]Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [4]Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [5]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [6]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [7]Biomedical and Biological Sciences Program, Harvard Medical School, Boston, MA, USA. [8]Department of Genome Sciences, University of Washington, Seattle, WA, USA. [9]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [10]These authors contributed equally: Daichi Inoue, Guo-Liang Chew. *e-mail: abdelwao@mskcc.org; rbradley@fredhutch.org

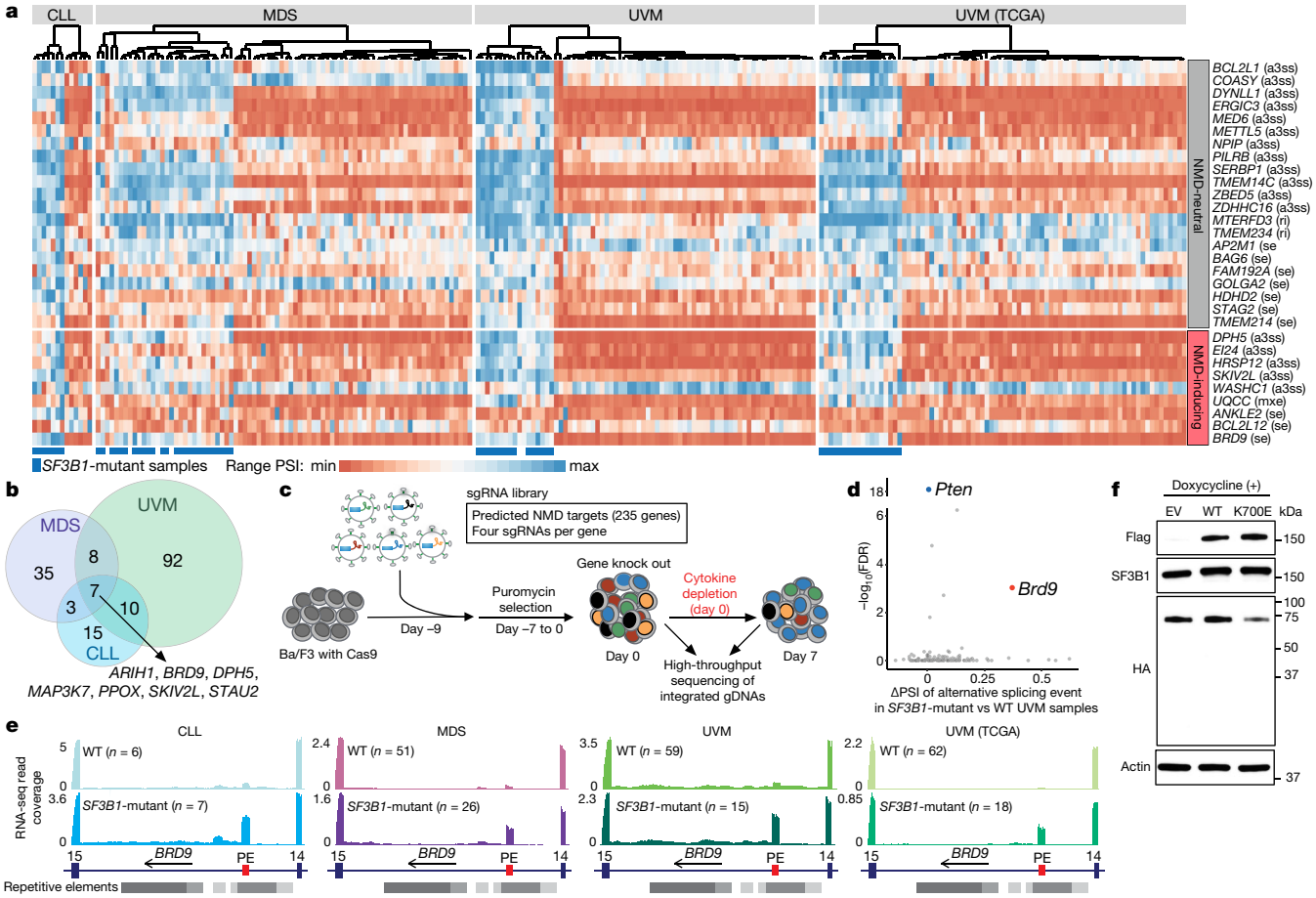


Fig. 1 | *BRD9* mis-splicing causes BRD9 loss and proliferative advantage in *SF3B1*-mutated cancers. a, Unsupervised clustering of patient samples on the basis of events that are differentially spliced in UVM (MEL270) and myeloid leukaemia (K562) cells that express $SF3B1^{K700E}$ versus wild-type (WT) *SF3B1*. a3ss, alternative 3′ splice site; CLL, chronic lymphocytic leukaemia (data are from ref. [15]); MDS, myelodysplastic syndromes (data are from ref. [27]); mxe, mutually exclusive exons; PSI, percentage spliced in (fraction of mRNA that corresponds to the mutant SF3B1-promoted isoform), with per-event and per-cohort range normalization; ri, retained intron; se, skipped exon; TCGA, The Cancer Genome Atlas. Data for UVM are from ref. [10] (middle) or the TCGA (right). *FAM192A*, *HRSP12*, *MTERFD3*, *NPIP* and *UQCC* are also known as *PSME3IP1*, *RIDA*, *MTERF2*, *NPIPA1* and *UQCC1*, respectively. **b**, Genes for which mutant SF3B1 promotes an isoform predicted to trigger NMD (alternative 3′ splice site and skipped exon events only) in one or more cohorts. **c**, CRISPR–Cas9-based positive-selection screen targeting genes for which mutant SF3B1 promotes an isoform predicted to trigger NMD. **d**, Per-gene scatter plot comparing CRISPR screen enrichment (*y* axis) to differential splicing in TCGA cohort of patients with UVM (*x* axis). *Pten* was used as a positive control. $n = 6$ biologically independent experiments. Per-gene significance computed with two-sided correlation-adjusted mean rank gene set (CAMERA) test. The false-discovery rate (FDR) was computed using the Benjamini–Hochberg method. **e**, *BRD9* RNA sequencing (RNA-seq) read coverage in patient samples. *n*, number of patients. PE, *BRD9* poison exon; 14 and 15, flanking constitutive exons. Repetitive elements from RepeatMasker[28]. **f**, Western blot for N-terminally haemagglutinin (HA)-tagged endogenous BRD9 in MEL270 cells transduced with empty vector (EV) or doxycycline-inducible Flag–SF3B1(WT) or Flag–SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments.

Fig. 2g). Consistent with the lack of an obvious polypyrimidine tract upstream of the poison exon, neither *U2AF1* nor *U2AF2* knockdown compromised poison exon recognition, whereas introducing a poly(Y) tract resulted in robust inclusion of the poison exon even in wild-type cells (Fig. 2b, Extended Data Fig. 2h–j). Finally, we identified a putative exonic splicing enhancer that was essential for inclusion of the poison exon (Fig. 2c, Extended Data Fig. 2k). We confirmed that the aberrant branchpoint, lack of a polypyrimidine tract, and exonic splicing enhancer were essential for poison exon recognition in the context of $SF3B1^{R625H}$, the most common mutation of *SF3B1* in UVM (Extended Data Fig. 2l–n). Disrupting the 3′ splice site and/or exonic splicing enhancer of the poison exon with CRISPR-directed mutagenesis markedly increased the levels of BRD9 protein in UVM cells with mutated *SF3B1* (Fig. 2d, Extended Data Fig. 2o, Supplementary Table 8), but had no effect on *BRD9* splicing or expression in cells with wild-type *SF3B1* (Extended Data Fig. 2p–r).

Several studies have recently described BRD9 as part of a non-canonical (nc) BAF complex, which is biochemically distinct from canonical BAF and polybromo-associated BAF[5–7] (Fig. 3a). Although ncBAF is not recurrently mutated in cancer—unlike canonical BAF and polybromo-associated BAF (Extended Data Fig. 3a)—our data suggested that ncBAF is nonetheless frequently disrupted via *SF3B1* mutations.

We investigated the consequences of BRD9 loss by *SF3B1* mutations for ncBAF function. Immunoprecipitation and mass spectrometry to identify the chromatin-associated interaction partners of BRD9 in K562 cells specifically recovered ncBAF components (Extended Data Fig. 3b, c, Supplementary Table 9). We confirmed these results by immunoblotting against shared and complex-specific components of canonical BAF, polybromo-associated BAF and ncBAF in K562 and UVM cells (Fig. 3b, Extended Data Fig. 3d). Expression of mutant, but not wild-type, SF3B1 reduced the levels of BRD9 protein and abolished interactions between BRG1 and GLTSCR1 while leaving interactions between BRG1 and BAF155 intact, which indicates that *SF3B1* mutations specifically perturb ncBAF rather than disrupting all BAF complexes (Fig. 3c, Extended Data Fig. 3e). Chemical degradation of BRD9[19] or *BRD9* knockout similarly reduced the BRG1–GLTSCR1 interaction (Fig. 3c, Extended Data Fig. 3f). We next identified the BRD9 domains that are necessary for ncBAF formation by generating 3×Flag–BRD9 deletion mutants

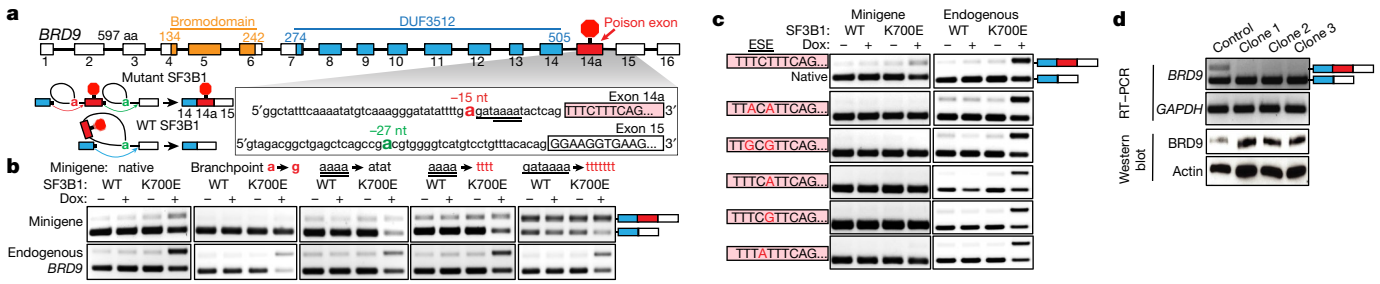


Fig. 2 | Mutant SF3B1 recognizes an aberrant, deep intronic branchpoint within *BRD9*. **a**, *BRD9* gene structure and protein domains. Inset illustrates the branchpoints used when the poison exon is included (top) or excluded (bottom). Single and double underlining indicates sequence motifs that were subsequently mutated. aa, amino acid. **b**, PCR with reverse transcription (RT–PCR) analysis of inclusion of the *BRD9* poison exon in a minigene (top) or endogenous (bottom) context, following transfection of minigenes with the illustrated mutations into MEL270 cells with doxycycline (dox)-inducible Flag–SF3B1(WT) or Flag–SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. Native, no mutations. **c**, As in **b**, but for minigene mutations (shown in red) at the 5′ end of the poison exon. ESE, exonic splicing enhancer. **d**, RT–PCR (top) illustrating the loss of inclusion of the *BRD9* poison exon, and corresponding western blot (bottom) in MEL202 ($SF3B1^{R625G}$) clones following CRISPR–Cas9 targeting of the poison exon. Indels are illustrated in Extended Data Fig. 2o. Control, unedited cells. Representative images from $n = 2$ (RT–PCR) and $n = 3$ (western blot) biologically independent experiments.

and testing for interactions with GLTSCR1 and GLTSCR1L. These experiments revealed that the DUF3512 domain of BRD9 mediates its interactions with GLTSCR1 and GLTSCR1L (Extended Data Fig. 3g–h).

We next determined how *SF3B1* mutations altered ncBAF localization to chromatin. We mapped the genome-wide binding of the pan-BAF component BRG1, and the ncBAF-specific components BRD9 and GLTSCR1, in MEL270 cells that express wild-type or mutant SF3B1. We additionally performed the same chromatin immunoprecipitation with sequencing (ChIP–seq) experiments after treatment with dimethylsulfoxide (DMSO) or a BRD9 degrader to identify BRD9-dependent effects. BRD9 and GLTSCR1 exhibited substantial co-localization, consistent with their mutual requirement for ncBAF formation, and were found at a subset of the loci bound by BRG1 (Fig. 3d). BRD9 and GLTSCR1 bound to promoters, gene bodies, and probable enhancers, with focal binding at promoters relative to BRG1 (Fig. 3e, Extended Data Fig. 4a). CTCF motifs exhibited notable co-localization with GLTSCR1, but only modest co-localization with BRG1 (Fig. 3f, Extended Data Fig. 4b).

We then tested how the depletion of BRD9, induced by $SF3B1^{K700E}$ or by chemical degradation of BRD9, altered ncBAF localization. We defined the genomic loci bound by GLTSCR1 in all samples as constitutive sites. Conversely, we defined genomic loci bound by GLTSCR1 in both control (wild-type SF3B1 or DMSO) but not BRD9-depleted

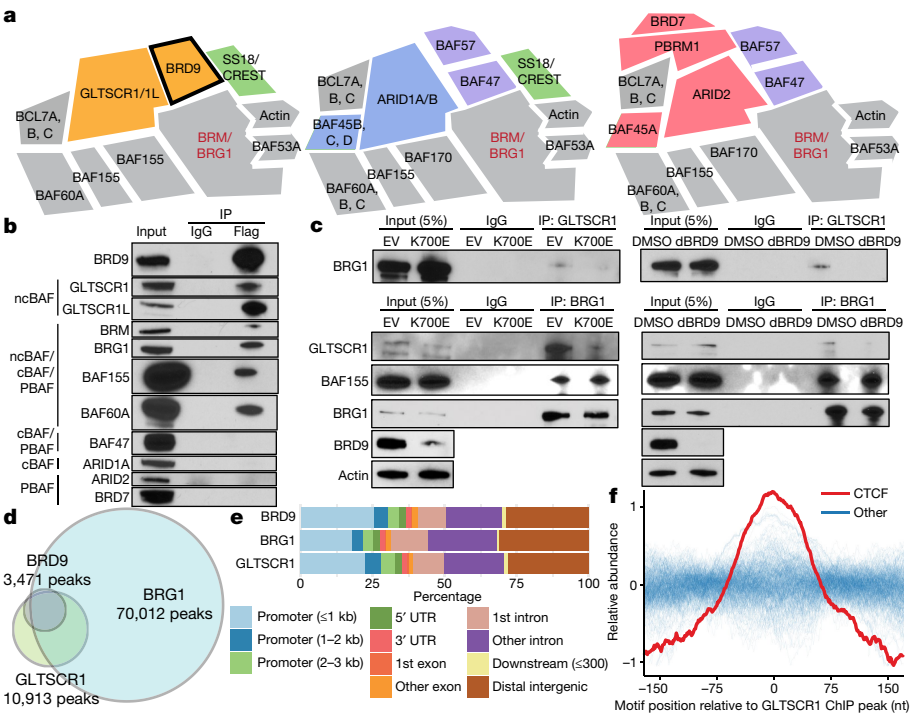


Fig. 3 | BRD9 loss perturbs the formation and localization of the ncBAF complex. **a**, Schematic of non-canonical BAF (ncBAF; left), canonical BAF (cBAF; middle) and polybromo-associated BAF (PBAF; right) complexes[5,6]. Solidus denotes one of the proteins is present; comma denotes one or more of the proteins are present or that mutually exclusive inclusion of proteins may occur. **b**, Cross-linking and immunoprecipitation (IP) with IgG or Flag followed by immunoblotting in K562 cells that express 3×Flag–BRD9. Representative images from $n = 3$ biologically independent experiments. **c**, Immunoprecipitation with GLTSCR1 or BRG1 antibody followed by immunoblotting in MEL270 cells that express exogenous SF3B1(K700E) (left) or were treated with BRD9 degrader[19] (dBRD9) (right). Representative images from $n = 3$ biologically independent experiments. **d**, Overlap of consensus BRD9, BRG1 and GLTSCR1 ChIP–seq peaks called in both MEL270 control samples (DMSO and ectopic expression of wild-type SF3B1). **e**, Genomic localization of BRD9-, BRG1-, and GLTSCR1-bound loci in **d**. UTR, untranslated region. **f**, Distributions of transcription factor-binding motifs at GLTSCR1-bound loci (20 nucleotide (nt) rolling mean). $n = 401$ transcription factors analysed.

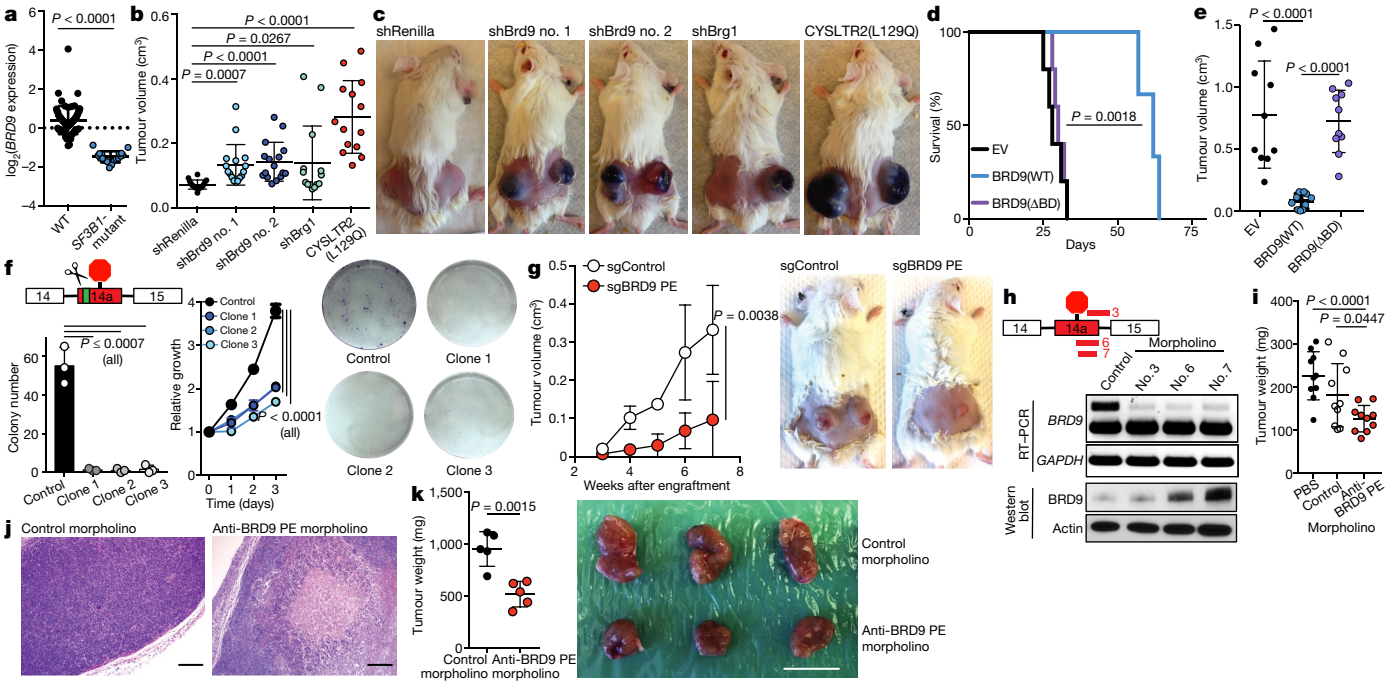


Fig. 4 | BRD9 is a therapeutically targetable tumour suppressor in melanoma. a, *BRD9* expression (*z*-score normalized) in TCGA UVM samples with ($n = 18$) or without ($n = 62$) *SF3B1* mutations. *P* value calculated by two-sided *t*-test. **b**, Tumour volume 49 days after subcutaneous engraftment of Melan-a cells transduced with the indicated shRNAs into SCID mice. $n = 16$, 16, 16, 14 and 14 tumours per group (left to right). Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test. **c**, Representative mice from **b** at day 63. **d**, Survival of SCID mice engrafted with MEL270 cells that express empty vector, full-length wild-type BRD9 or a BRD9 bromodomain-deletion mutant (ΔBD). $n = 5$ mice per group. *P* value calculated by log-rank test. **e**, Tumour volume from experiments shown in **d**, 21 days after engraftment. $n = 10$ tumours per group. Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test. **f**, Colony number (left) and representative images (right) of MEL202 cells ($SF3B1^{R625G}$) without (control) or with (clone 1, clone 2 and clone 3) CRISPR–Cas9-induced disruption of the *BRD9* poison exon. Indels are illustrated in Extended Data Fig. 2o. $n = 3$ biologically independent experiments. Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test at day 3 (middle). **g**, Tumour volume (left) and representative images (right) of mice engrafted with control or clone 1 cells from **f**. $n = 6$ tumours per group. Error bars, mean ± s.d. *P* value calculated by two-sided *t*-test at week 7. **h**, ASO design (top), RT–PCR (middle) and western blot (bottom) for BRD9. MEL202 cells ($SF3B1^{R625G}$) were treated with a non-targeting (control) or targeting morpholino at 10 μM for 24 h. Representative images from $n = 3$ biologically independent experiments. **i**, Tumour weight following 16 days of in vivo treatment of MEL202-derived xenografts ($SF3B1^{R625G}$) with PBS or a non-targeting (control) or poison-exon-targeting (no. 6) morpholino (12.5 mg kg^{-1}, every other day to a total of 8 intratumoral injections). $n = 10$ tumours per group. Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test. **j**, Haematoxylin and eosin (H&E) images of tumours from **i**. Scale bars, 200 μm. Representative images from $n = 3$ biologically independent histological analyses. **k**, Tumour weight (left) and representative images (right) following in vivo morpholino treatment of a patient-derived rectal melanoma xenograft ($SF3B1^{R625C}$). Scale bar, 1 cm. $n = 5$ tumours per group. *P* value calculated by two-sided *t*-test. Error bars, mean ± s.d.

(mutant SF3B1 or BRD9 degradation) samples as BRD9-sensitive sites (Extended Data Fig. 4c). GLTSCR1 peaks were more sensitive to BRD9 loss than were BRG1 peaks, and CTCF motifs were uniquely enriched in BRD9-sensitive loci ($P < 1 \times 10^{-8}$) versus constitutive GLTSCR1-bound loci (Extended Data Fig. 4d). CTCF was similarly highly enriched at BRG1-bound loci that were BRD9-sensitive ($P < 1 \times 10^{-55}$) (Extended Data Fig. 4e, f). We conclude that the BRD9 loss induced by *SF3B1* mutations causes specific loss of ncBAF at CTCF-associated loci.

We identified genes with BRD9-sensitive ncBAF binding in their promoters or enhancers and found that BRD9 loss in UVM preferentially affects genes involved in apoptosis and cell growth, adhesion and migration (Extended Data Fig. 4g). To understand how BRD9 loss altered gene expression, we identified genes with promoters that exhibited BRD9-sensitive ncBAF binding and that were differentially expressed in patients with UVM with mutant versus wild-type *SF3B1*. Loss of ncBAF binding was associated with promotion as well as repression of gene expression, suggesting that ncBAF—similar to other SWI/SNF complexes—has both activating and repressive roles[20] (Extended Data Fig. 4h–j).

Several recent studies have reported that BRD9 is required for the survival of some cancer types, particularly cancers with mutations that affect polybromo-associated BAF and canonical BAF[6,16,21]. Because *BRD9* loss conferred a proliferative advantage to types of cancer with recurrent *SF3B1* mutations (Fig. 1d, Extended Data Fig. 1), we wondered whether normalizing the levels of BRD9 might suppress the growth of *SF3B1*-mutant cells.

As *SF3B1* is recurrently mutated in uveal (Fig. 4a), mucosal and cutaneous melanomas, we first tested whether *BRD9* loss induced melanomagenesis in vivo. We transduced non-tumorigenic mouse melanocytes (Melan-a cells), which require oncoprotein expression for sustained growth, with a non-targeting short hairpin RNA (shRNA) (Extended Data Fig. 5a), doxycycline-inducible shRNAs targeting *Brd9* or *Brg1* (also known as *Smarca4*) or a cDNA encoding the oncoprotein CYSLTR2(L129Q) (as a positive control[22]). Knockdown of either *Brd9* or *Brg1* resulted in potent tumour growth, augmented melanocyte pigmentation, and expression of melanocyte-lineage-specific genes in vivo (Fig. 4b, c, Extended Data Fig. 5b–g).

We next tested whether *Brd9* expression influences metastasis. *Brd9* knockdown significantly increased the number of pulmonary metastatic foci following intravenous injection of cells from a mouse model of melanoma (B16) or of human UVM (92.1) cells into mice (Extended Data Fig. 6a–f). By contrast, restoring *Brd9* expression in established tumours in vivo, by withdrawing doxycycline, suppressed tumour growth (Extended Data Fig. 6g, h). Similarly, ectopic expression of full-length BRD9, but not the bromodomain- or DUF3512-deletion mutants, suppressed the growth of UVM cell lines and xenografts (Fig. 4d, e, Extended Data Fig. 6i–k). These data demonstrate that loss of *Brd9* promotes cell transformation, tumour maintenance, and metastatic progression, and that the bromodomain and DUF3512 domain of BRD9 are essential for its anti-proliferative effects.

We sought to understand how BRD9 loss promotes melanoma tumorigenesis. We identified BRD9-bound genes that exhibited dysregulated expression in samples from patients with UVM with mutated versus wild-type *SF3B1*, and in isogenic UVM cells with or without mutant SF3B1 and with or without forced loss of BRD9. *HTRA1*, a known tumour suppressor in melanoma[23,24], was the most downregulated gene in UVM (Extended Data Fig. 7a–c). HTRA1 was suppressed by mutant SF3B1 expression and BRD9-degradation treatment of UVM cells with wild-type *SF3B1*, and mutagenesis of the *BRD9* poison exon increased levels of HTRA1 in UVM cells with mutated *SF3B1* (Extended Data Fig. 7d, e). *HTRA1* is bound by ncBAF in UVM, and this binding is reduced by mutant SF3B1 (Extended Data Fig. 7f). *HTRA1* knockdown promoted the growth of UVM cells with wild-type *SF3B1*, and ectopic expression of *HTRA1* suppressed the growth of UVM cells with mutated *SF3B1* (Extended Data Fig. 7g–k). These data suggest that perturbation of ncBAF-dependent regulation of *HTRA1* contributes to the pro-tumorigenic effects of BRD9 loss.

We next tested whether correcting *BRD9* mis-splicing suppressed tumorigenesis. CRISPR-based mutagenesis of the poison exon markedly slowed the growth of cells with mutated *SF3B1*, but not of wild-type cells, both in vitro and in vivo (Figs. 2d, 4f, g, Extended Data Figs. 2o–r, 8). We then designed antisense oligonucleotides (ASOs) to block the inclusion of the *BRD9* poison exon (Fig. 4h, Extended Data Fig. 9a). We treated *SF3B1*-mutated cells with a non-targeting (control) or poison-exon-targeting ASO, and measured *BRD9* splicing, BRD9 protein levels, and cell growth. Each targeting ASO prevented the inclusion of the poison exon, increased the level of BRD9 protein, and suppressed cell growth relative to the control ASO (Fig. 4h, Extended Data Fig. 9b). The relative abilities of each ASO to restore BRD9 protein levels and suppress cell growth were strongly correlated, consistent with on-target effects.

We therefore tested whether ASO treatment slowed tumour growth in vivo. We treated *SF3B1*-mutated xenografts (derived from MEL202 cells) with each ASO via intratumoral injection for 16 days. Treatment with the poison-exon-targeting ASO—but not with the non-targeting ASO—corrected *BRD9* mis-splicing, significantly reduced tumour growth, and induced tumour necrosis (Fig. 4i, j, Extended Data Fig. 9c–f). We observed a similar ASO efficacy in a patient-derived xenograft model of rectal melanoma with the $SF3B1^{R625C}$ mutation (Fig. 4k, Extended Data Fig. 9g–i, Supplementary Table 7). By contrast, when we performed an identical experiment with a patient-derived xenograft model of UVM that lacked an *SF3B1* mutation, treatment with the poison-exon-targeting ASO had no effect (Extended Data Fig. 9j–l, Supplementary Table 7). We conclude that correcting *BRD9* mis-splicing restores the tumour suppressor activity of BRD9 in cancers with *SF3B1* mutations.

Although recognition of the *BRD9* poison exon requires mutant SF3B1, *BRD9* mis-splicing and ncBAF disruption may also have roles in cancers with wild-type *SF3B1*. We identified significant pan-cancer expression correlations between *BRD9* and many genes that encode RNA-binding proteins, as well as six additional *BRD9* isoforms that are predicted to trigger NMD that are expressed in cancers with wild-type *SF3B1* and are predictive of *BRD9* expression (Extended Data Fig. 10, Supplementary Table 10). A recent study has also identified a promoter polymorphism associated with decreased *GLTSCR1* (also known as *BICRA*) expression as a common risk allele for acute myeloid leukaemia[25].

As we observed *BRD9* mis-splicing in a range of cancer types that carry distinct *SF3B1* mutations, targeting *BRD9* mis-splicing could be a productive pan-cancer therapy. Although ASO treatment merely restored *BRD9* mRNA and BRD9 protein to normal levels, we nonetheless observed a strong suppression of tumour growth. The functional effect of correcting *BRD9* mis-splicing was particularly notable given that the UVM models used here contain hundreds of other mis-splicing events and multiple pro-tumorigenic mutations. Given recent clinical successes with treating spinal muscular atrophy and other diseases with ASOs[26], the tumour-suppressive effects of correcting *BRD9* mis-splicing suggest that oligonucleotide-based therapy may prove similarly promising for treating cancers with spliceosomal mutations.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1646-9.

Received: 12 January 2019; Accepted: 30 August 2019;
Published online 9 October 2019.

1. Yoshida, K. et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
2. Papaemmanuil, E. et al. Somatic *SF3B1* mutation in myelodysplasia with ring sideroblasts. *N. Engl. J. Med.* **365**, 1384–1395 (2011).
3. Wang, L. et al. *SF3B1* and other novel cancer genes in chronic lymphocytic leukemia. *N. Engl. J. Med.* **365**, 2497–2506 (2011).
4. Seiler, M., et al. Somatic mutational landscape of splicing factor genes and their functional consequences across 33 cancer types. *Cell Rep.* **23**, 282–296.e284 (2018).
5. 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).
6. 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).
7. Gatchalian, J. et al. A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells. *Nat. Commun.* **9**, 5139 (2018).
8. Quesada, V. et al. Exome sequencing identifies recurrent mutations of the splicing factor *SF3B1* gene in chronic lymphocytic leukemia. *Nat. Genet.* **44**, 47–52 (2011).
9. Furney, S. J. et al. *SF3B1* mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov.* **3**, 1122–1129 (2013).
10. Alsafadi, S. et al. Cancer-associated *SF3B1* mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat. Commun.* **7**, 10615 (2016).
11. Harbour, J. W. et al. Recurrent mutations at codon 625 of the splicing factor *SF3B1* in uveal melanoma. *Nat. Genet.* **45**, 133–135 (2013).
12. Martin, M. et al. Exome sequencing identifies recurrent somatic mutations in *EIF1AX* and *SF3B1* in uveal melanoma with disomy 3. *Nat. Genet.* **45**, 933–936 (2013).
13. Cretu, C. et al. Molecular architecture of SF3b and structural consequences of its cancer-related mutations. *Mol. Cell* **64**, 307–319 (2016).
14. DeBoever, C. et al. Transcriptome sequencing reveals potential mechanism of cryptic 3′ splice site selection in *SF3B1*-mutated cancers. *PLoS Comput. Biol.* **11**, e1004105 (2015).
15. Darman, R. B. et al. Cancer-associated SF3B1 hotspot mutations induce cryptic 3′ splice site selection through use of a different branch point. *Cell Rep.* **13**, 1033–1045 (2015).
16. Hohmann, A. F. et al. Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. *Nat. Chem. Biol.* **12**, 672–679 (2016).
17. Gozani, O., Potashkin, J. & Reed, R. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol. Cell. Biol.* **18**, 4752–4760 (1998).
18. Pineda, J. M. B. & Bradley, R. K. Most human introns are recognized via multiple and tissue-specific branchpoints. *Genes Dev.* **32**, 577–591 (2018).
19. Remillard, D. et al. Degradation of the BAF complex factor BRD9 by heterobifunctional ligands. *Angew. Chem. Int. Edn Engl.* **56**, 5738–5743 (2017).
20. Martens, J. A. & Winston, F. Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Curr. Opin. Genet. Dev.* **13**, 136–142 (2003).
21. Brien, G. L. et al. Targeted degradation of BRD9 reverses oncogenic gene expression in synovial sarcoma. *eLife* **7**, e41305 (2018).
22. Moore, A. R. et al. Recurrent activating mutations of G-protein-coupled receptor CYSLTR2 in uveal melanoma. *Nat. Genet.* **48**, 675–680 (2016).
23. Baldi, A. et al. The HtrA1 serine protease is down-regulated during human melanoma progression and represses growth of metastatic melanoma cells. *Oncogene* **21**, 6684–6688 (2002).
24. Chien, J. et al. Serine protease HtrA1 associates with microtubules and inhibits cell migration. *Mol. Cell. Biol.* **29**, 4177–4187 (2009).
25. Walker, C. J. et al. Genome-wide association study identifies an acute myeloid leukemia susceptibility locus near *BICRA*. *Leukemia* **33**, 771–775 (2019).
26. Stein, C. A. & Castanotto, D. FDA-approved oligonucleotide therapies in 2017. *Mol. Ther.* **25**, 1069–1075 (2017).
27. Pellagatti, A. et al. Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. *Blood* **132**, 1225–1240 (2018).
28. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0, http://www.repeatmasker.org (2013–2015).

METHODS

Sample sizes for xenograft experiments were chosen on the basis of published studies of known oncogenic drivers of relevant models (for example, expression of the oncoprotein CYSLTR2(L129Q) in Melan-a cells). Mice were randomly assigned to experimental groups. The data presented did not require the use of blinding.

Cell lines and tissue culture. All cell lines underwent short-tandem repeat testing (ATCC) and Memorial Sloan Kettering integrated mutation profiling of actionable cancer targets (MSK IMPACT) genetic analysis[29] to evaluate for spliceosome-gene mutational status and status of recurrently mutated genes in cancer. HEK293T cells were grown in DMEM with 10% FCS. Ba/F3 cells and Melan-a cells (provided by D. Bennett) were grown in RPMI with 10% FCS with 1 ng/ml IL-3 (PeproTech, 213-13) and 200 nM TPA (Sigma-Aldrich), respectively, unless noted otherwise. The K562 and NALM-6 isogenic cell lines (engineered to express *SF3B1*^{K700E}, *SF3B1*^{K666N} or *SF3B1*^{K700K} (wild-type control for genome engineering) from the endogenous *SF3B1* locus) were cultured in RPMI with 10% FCS and their generation has previously been described[30]. MEL270, MEL285 and RN2 cell lines were cultured in RPMI with 10% FCS. MEL202, 92-1 and SK-MEL30 cells were grown in RPMI with 10% FCS and 1% GlutaMAX (Gibco). UPMD1 and UPMD2 cells were grown in Ham F-12 with 10% FCS. CFPAC1 cells were cultured in IMDM with 10% FCS. KPC, Miapaca2 and B16 cells were cultured in DMEM with 10% FCS. Panc 05.04 cells were grown in RPMI with 20% FCS and 20 units per millilitre human recombinant insulin. T47D cells were cultured in RPMI1640 supplemented with 10% fetal bovine serum (Corning), 100 μg/ml penicillin, 100 mg/ml streptomycin (Corning), and 4 mM glutamine. All of the cell culture media included penicillin (100 U/ml) and streptomycin (100 μg/ml).

Primary human samples and human patient-derived xenograft models. Studies were approved by the Institutional Review Boards (IRBs) of Memorial Sloan Kettering Cancer Center (MSK), informed consent was obtained from all subjects (under MSK IRB protocol 06-107) and studies were conducted in accordance to the Declaration of Helsinki protocol. Patients provided samples after their informed consent, and samples of primary human de-identified chronic lymphocytic leukaemia derived from whole peripheral blood or bone marrow mononuclear cells were used. Patient-derived xenograft models were performed using tumour biopsies from de-identified patients under MSK IRB protocol 14-191. Genomic alterations in melanoma tumour biopsies and chronic lymphocytic leukaemia cells were analysed using the MSK IMPACT[29] assay or FoundationOne Heme[31] assay, both as previously described. Patient samples were anonymized by the Hematologic Oncology Tissue Bank of MSK (for chronic lymphocytic leukaemia samples) and the MSK Antitumour Assessment Core Facility (for patient-derived xenograft samples).

Mice. All mice were housed at Memorial Sloan Kettering Cancer Center (MSKCC). All mouse procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC. All mouse experiments were performed in accordance with a protocol approved by the MSKCC Institutional Animal Care and Use Committee (11-12-029). SCID mice (Jackson Laboratories stock no. 001303) were used for all human cell line xenografts and NSG mice (Jackson Laboratories stock no. 005557) were used for patient-derived xenografts. For all mouse experiments, the mice were monitored closely for signs of disease or morbidity daily and were killed when they showed a volume of the visible tumour formation above 1 cm^3, failed to thrive, experienced weight loss >10% total body weight or showed open skin lesions, bleeding or any signs of infection. These limits were not exceeded in any experiments.

HA-tag knock-in into endogenous *BRD9*. The following guide RNA sequence targeting the *BRD9* transcriptional start site was selected using the optimized CRISPR design tool (http://crispr.mit.edu): CGAGTGGCGCTCGTCCTACG. DNA oligonucleotides were purchased from IDT and cloned into the px458-GFP vector. For homologous recombination, we purchased a custom IDT Ultramer 197-bp repair template (single-stranded donor DNA) with the HA sequence (TACCCATACGATGTTCCAGATTACGCT) directly following the *BRD9* start codon. This 197-bp fragment contained two silent mutations, one to remove the PAM site (AGG > AAG) and another to introduce an XhoI restriction enzyme (CTCGGG > CTCGAG) site upstream of the HA tag. The 197-bp fragment also contained 83 bp of homology to the *BRD9* 5′ UTR upstream of the HA tag and 87 bp of homology to the *BRD9* exon 1 downstream of the start codon. Five micrograms of the targeting construct and 500 nM of the repair template were nucleofected into K562 *SF3B1*^{K700E} cells and MEL270 cells using the Lonza Nucleofector V kit and Program T-003 on the nucleofector device. Nucleofected cells were single-cell-sorted on the basis of GFP positivity 48 h after nucleofection. Clones were screened for the presence of successful HA insertion by *BRD9* exon 1 PCR and subsequent restriction enzyme digestion with XhoI and direct Sanger sequencing. A single positive clone containing the HA coding sequencing was selected to carry out further studies. The sgRNAs were CACCCGAGTGGCGCTCGTCCTACG (top) and AAACCGTAGGACGAGCGCCACTCG (bottom); the single-strand donor DNA was CCAGGGGGCGGTGGCGGCCAAGGTCCGACCGGGTGC CAGCTGTTCCCAGCCCCCGCCTCGAGCCCGCCGCCGGCGCCGCCATGT ACCCATACGATGTTCCAGATTACGCTGGCAAGAAGCACAAGAAGCACA AGGCCGAGTGGCGCTCGTCCTACGAAGGTGAGGCGGCGGCGCTTTGTG ACGCGCGGCGGGCGGGG; the PCR primers were forward (fwd) AGCG AGCTCGGCAACCTCG and reverse (rev) CTTCAGGACTAGCTTTAGAGGC; the Sanger sequence primer was rev TGCAGCCTCGAACCCAGAAC

Overexpression of *SF3B1* cDNA in K562 cells. Two micrograms of PiggyBac Transposase construct (CMV-PB-Transposase-IRES-TK-HSV) and 6 μg of wild-type *SF3B1* (ITR-CAG-Flag-SF3B1^{WT}-IRES-Puro-ITR) or *SF3B1*^{K700E} (ITR-CAG-Flag-SF3B1^{K700E}-IRES-Puro-ITR) cDNA constructs were electroporated into 2 × 10^6 cells (in 200 μl volume) using the Amaxa Nucleofector Protocol (Program T-003) according to manufacturer instructions (Lonza). Puromycin selection (1 μg/ml) was initiated 4 days after electroporation to select for cells that successfully incorporated the constructs. Sanger sequencing was performed to confirm successful integration of the cDNA plasmid using the following primers: fwd, TCCAATCAAAGATCTTCTTCCAA and rev, GAGCAGGTTTCTGCAACGAT.

RT–PCR and quantitative RT–PCR. Total RNA was isolated using RNeasy Mini or Micro kit (Qiagen). For cDNA synthesis, total RNA was reverse-transcribed to cDNA with SuperScript VILO cDNA synthesis kit (Life Technologies). The resulting cDNA was diluted 10–20 fold before use. Quantitative RT–PCR (qRT–PCR) was performed in 10-μl reactions with SYBR Green PCR Master Mix. All qRT–PCR analysis was performed on an Applied Biosystems QuantStudio 6 Flex Cycler (ThermoFisher Scientific). Relative gene expression levels were calculated using the comparative C_T method. Primers used in RT–PCR reactions were: *BRD9* (human) fwd, GCAATGACATACAATAGGCCAGA and rev, GAGCTGCCTGTTTGCTCATCA; *Brd9* (mouse) fwd, TTGGAG ATGGAAGTCTGCTCT and rev, GCAACTTGCTAGACAGTGAACT; *BRD9* poison exon (human) fwd, AGCTCTGTTCTGGAGTTCATG and rev, CTGAAG AAACTCATAGGGGTCGTG; *Brd9* poison exon (mouse) fwd, GGCCCT GTTCTGGACTTCATG and rev, CTGAAGGAATTCATAAGGGTCGTG; *BRD9* poison exon inclusion for small interfering RNA (siRNA) experiment (human) fwd, CAGCAGCTCTGTTCTGGAGT and rev: CCTGAAAGAAACCAGAGAGCTG; *BRD9* poison exon exclusion for siRNA experiment (human) fwd, CAGCAGCTCTGTTCTGGAGT and rev, TCACCTTCCCCAGAGAGCTG; *EPB49* (also known as *DMTN*) cassette exon inclusion (human) fwd, GCCTGCAGAACGGAGAGG and rev: ACCACTAGCATTTCATAGGGATAGATCT; *EPB49* cassette exon exclusion (human) fwd, GCCTGCAGATCTATCCCTATGAAAT and rev, CTCAAGCCGCATCCGATCC; *BRD9* poison exon for mRNA half-life experiment (human) fwd, GTTGGGGACACCCTAGGAGA, rev (exclusion-specific), CTTCACCTTCCCCAGAGAGC and rev (inclusion-specific), CCCTGAAAGAAACCAGAGAGC; 18S rRNA (human) fwd CTAC CACATCCAAGGAAGCA and rev, TTTTTCGTCACTACCTCCCCG; *Mitf* (mouse) fwd, CCAACAGCCCTATGGCTATGC and rev: CTGGGCA CTCACTCTCTGC; *Dct* (mouse) fwd, GTCCTCCACTCTTTTACAGACG and rev, ATTCGGTTGTGACCAATGGGT; *Pmel* (mouse) fwd, GAGCTTCCTTCCCGTGCTT and rev, TGCCTGTTCCAGGTTTTAGTTAC; *Tyrp1* (mouse) fwd, CCCCTAGCCTATATCTCCCTTTT and rev, TACCATCGT GGGGATAATGGC; *GAPDH* (human) fwd, GGAGCGAGATCCCTCCAAAAT and rev, GGCTGTTGTCATACTTCTCATGG; and *Gapdh* (mouse) fwd, AGGTCGGTGTGAACGGATTTG and rev, TGTAGACCATGTAGTTGAGGTCA.

mRNA stability assay. For mRNA half-life measurement using qRT–PCR, K562 and NALM-6 cells with isogenic *SF3B1*^{K700E} mutations were infected with anti-*UPF1* shRNAs or control shRNA, and treated with 2.5 μg/ml actinomycin D (Life Technologies) and collected at 0, 2, 4, 6 and 8 h (using methods as previously described[32]). *BRD9* poison exon inclusion or exclusion and 18S rRNA mRNA levels were measured by qRT–PCR.

Western blotting. For western blotting, the following antibodies to the following proteins were used: BRD9 (Bethyl Laboratories A303-781A and Active Motif 61538), SF3B1/Sap-155 (MBL D221-3), Flag-M2 (Sigma-Aldrich F-1084), β-actin (Sigma-Aldrich A-5441), GLTSCR1 (Santa Cruz Biotechnology sc-515086), GLTSCR1L (Thermo Fisher Scientific PA5-56126), BRM (Bethyl Laboratories A303-015A), BRG1 (Santa Cruz Biotechnology sc-17796), BAF155 (Santa Cruz Biotechnology sc-48350), BAF60A (Santa Cruz Biotechnology sc-135843), BAF47 (Santa Cruz Biotechnology sc-166165), ARID1A (Santa Cruz Biotechnology sc-373784), ARID2 (Santa Cruz Biotechnology sc-166117), BRD7 (Thermo Fisher Scientific PA5-49379), U2AF2 (Bethyl Laboratories A303-665A), U2AF1 (Bethyl Laboratories A302-080A), histone H3 (Abcam ab1791), HTRA1 (R&D Systems MAB2916-SP). All primary antibodies for western blotting were diluted to a final concentration of 1:500 to 1,000, in either 5% BSA (Sigma-Aldrich) in 0.05% TBS–Tween 20 (TBS-T) or 5% skim milk in 0.05% TBS-T. Nuclear extracts were quantified using BCA and 1 mg protein (1 mg ml^{-1} in immunoprecipitation buffer supplemented with protease inhibitors) was used per immunoprecipitation.

Proteins were incubated for 3 h with 2–5 μg of antibody or with protein A/G PLUS-Agarose (Santa Cruz Biotechnology sc-2003) with rotation at 4 °C. After washing three times with Pierce IP Lysis Buffer (Thermo Fisher Scientific 87787), immunoprecipitated proteins were eluted with Pierce Lane Marker Reducing Sample Buffer (Thermo Fisher Scientific 39000) and loaded onto 4–12% Bis-Tris NuPAGE Gels (Life Technologies).

ChIP. For ChIP–seq studies in MEL270 cells, antibodies to endogenous BRG1 (Abcam EPNCIR111A, lot no. GR3208604-8), GLTSCR1 (Santa Cruz SC-240516, lot no. A2313) and BRD9 (Abcam ab137245) were used, and ChIP was performed as previously described in detail[6]. MEL270 cells transduced with empty vector, doxycycline-inducible wild-type *SF3B1* cDNA, or doxycycline-inducible *SF3B1*^{K700E} cDNA in the backbone of pInducer20, were treated with doxycycline (1 μg/ml) plus BRD9 degrader[19] (250 nM) or DMSO for 72 h before crosslinking.

Mass spectrometry. For anti-Flag–BRD9 ChIP followed by mass spectrometry, K562 cells transduced with empty vector or 3×Flag-tagged BRD9 were grown in RPMI with 10% FCS. Ten million cells were crosslinked according to the manufacturer's instruction (Active Motif) and as previously described[33]. Cells were fixed with 1% methanol-free formaldehyde (Sigma, F-8775) for 8 min and quenched with 0.125 M glycine (Sigma, G-7403). Chromatin was isolated by the addition of lysis buffer, followed by disruption with a Dounce homogenizer. Lysates were sonicated and the DNA sheared to an average length of 300–500 bp. Genomic DNA (input) was prepared by treating aliquots of chromatin with RNase, proteinase K and heat for reverse-crosslinking, followed by ethanol precipitation. Pellets were resuspended and the resulting DNA was quantified on a NanoDrop spectrophotometer. Extrapolation to the original chromatin volume allowed quantification of the total chromatin yield. An aliquot of chromatin (150 μg) was precleared with protein G agarose beads (Invitrogen). Proteins of interest were immunoprecipitated using 15 μg of antibody against Flag and protein G magnetic beads. Protein complexes were washed and trypsin was used to remove the immunoprecipitate from beads and digest the protein sample. Protein digests were separated from the beads and purified using a C18 spin column (Harvard Apparatus). The peptides were vacuum-dried using a SpeedVac. Digested peptides were analysed by liquid chromatography and tandem mass spectrometry on a Thermo Scientific Q Exactive Orbitrap mass spectrometer in conjunction with a Proxeon Easy-nLC II HPLC (Thermo Scientific) and Proxeon nanospray source.

Protein identifications were accepted if they contained at least one identified peptide. Proteins that contained similar peptides and could not be differentiated on the basis of tandem mass spectrometry analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Protein peptide evidence is specified in Supplementary Table 9. The final list was generated by taking all proteins with a spectral count of five and above from each replicate reaction and comparing them in a Venn diagram against IgG control replicates. Proteins unique to both experimental replicates were then applied to the PANTHER database for protein ontology results.

shRNA experiments. Cells were transduced with a doxycycline-inducible LT3GEPIR lentiviral vector, T3G-GFP-mirE-PGK-Puro-IRES-rtTA3[34], expressing shRNAs against *BRD9* or a non-targeting shRNA against *Renilla*. shRNAs were induced with the addition of doxycycline (2.0 μg/ml, Sigma Aldrich). All shRNAs were designed using the SplashRNA algorithm[35]. The shRNA sequences are: *BRD9* shRNA no. 1 (human, shBRD9_352): TTTATTATCATTGAATATCCAG; *BRD9* shRNA no. 2 (human, shBRD9_353): TTTTATTATCATTGAATATCCA; *Brd9* shRNA no. 1 (mouse, shBrd9_511): TTTATTATCATTGAATACCCAG; *Brd9* shRNA no. 2 (mouse, shBrd9_512): TTTTATTATCATTGAATACCCA; *HTRA1* shRNA no. 1 (human, shHTRA1_1192): TTTTTAATCTTATCAGATGGGA; *HTRA1* shRNA no. 2 (human, shHTRA1_1669): TGAACAAACAAAATGGCAGTCA; and *HTRA1* shRNA no. 3 (human, shHTRA1_1898): TTCTATCTACGCATTGTATCGA.

siRNA transfections. K562 cells were transfected with a non-targeting control siRNA (Dharmacon, D-001810-01, target sequence: UGGUUUACAUGUCGACUAA), an siRNA pool against *U2AF1* (Dharmacon ON-TARGETplus SMARTpool, L-012325-01) or an siRNA pool against *U2AF2* (Dharmacon ON-TARGETplus SMARTpool, L-012380-02) using the Nucleofector II device from Lonza with the Cell Line Nucleofector Kit V (program T16). RNA and protein were extracted 48 h after transfection. cDNA was produced using 1 μg of RNA and the Superscript III first strand synthesis system (Thermo Fisher, 18080051).

In vitro competition assay. For the competition assay to evaluate the cellular effect of sgRNA against *BRD9* or *Brd9*, cell lines were transduced with LentiCas9-Blast (Addgene no. 52962) and then single-cell-sorted into 96-well plates. Among these clones, we used single clones with strong Cas9 expression, which was confirmed by western blotting. Cas9-expressing cells were lentivirally transduced with iLenti-guide-GFP vectors subcloned with the gRNA sequences, in which sgRNA expression was linked to GFP expression. The percentage of GFP-expressing cells was then measured over time after infection using BD LSRFortessa. The GFP-positive rates in living cells at each point compared to that of day 2 were calculated. Similarly, to evaluate the cellular effect of *BRD9* fragment cDNA, *HTRA1* cDNA or shRNAs against *BRD9* or *HTRA1*, GFP-positive rates were measured after transducing pMIGII-backbone plasmids (cDNA) and LT3GEPIR plasmids (shRNA), respectively, into melanoma cell lines. The sgRNAs were as follows: *BRD9* sgRNA no. 1 (human): ACTCCAGTTACTATGATGAC, *BRD9* sgRNA no. 2 (human): AGAGAGGGAGCACTGTGACA, *BRD9* sgRNA no. 3 (human): AGATACCGTGTACTACAAGT, *Brd9* sgRNA no. 1 (mouse): ATTAACCGGTTTCTCCCGGG, *Brd9* sgRNA no. 2 (mouse): GGAACACT GCGACTCAGAGG, *Brd9* sgRNA no. 3 (mouse): ACTTGCTAGACAGTGAACTC, and control (scrambled sgRNA): ACGGAGGCTAAGCGTCGCAA.

CRISPR enrichment screening for NMD targets. First, Ba/F3 cells were transduced with LentiCas9-Blast (Addgene no. 52962) and single-cell-sorted into 96-well plates. Among these clones, we used a single clone with strong Cas9 expression. The sgRNA library of NMD targets in SF3B1-mutant cells were amplified and packaged as lentivirus. The library includes 4 sgRNAs against each target gene (a total of 274 genes), 100 control sgRNAs and positive-control sgRNAs against *Pten*. Ba/F3 cells were transduced with the lentivirus-carrying sgRNA library produced by 293FT cells, and puromycin selection (2 μg/ml) was performed in IL-3-containing medium for 7 days. Then, we washed out IL-3 (on day 0) and the surviving cells were collected 7 days after IL-3 depletion (day 7). Cell pellets were lysed and genomic DNA was extracted (Qiagen), and quantified by Qubit (Thermo Scientific). A quantity of gDNA covering a 1,000× representation of gRNA was PCR-amplified using Q5 high-fidelity polymerase (NEB cat. no. M0491) to add Illumina adapters and multiplexing barcodes. Amplicons were quantified by Qubit and Bioanalyzer (Agilent) and sequenced on Illumina HiSeq 2500. Sequencing reads were aligned to the screened library; counts were computed for each gRNA and counts for each sgRNA were compared between days 0 and 7 after cytokine depletion. For the probe-level analysis, we fitted a negative binomial generalized log-linear model and performed a likelihood ratio test with glmFit and glmLRT in the Bioconductor edgeR package. For the gene-level analysis, we used the CAMERA test as implemented in edgeR[36–38]. FDR values were computed using the Benjamini–Hochberg method.

CRISPR-directed mutations. Cas9-expressing MEL202 and MEL270 cells were transduced with iLenti-guide-Puro vector targeting the 5′ end of the *BRD9* poison exon. The sgRNA used to induce mutations was: AAAATACTCAGTTTCTTTCA. Single-cell sorting was performed into a 96-well plate using a BD FACSAria III cell sorter after puromycin selection (2.0 μg/ml for 7 days). The mutations caused by Cas9 and the sgRNA were confirmed by PCR amplification, followed by Sanger sequencing. The PCR primers were fwd, TGTTGGGTCAGGAAGAGACTTG and rev, CCATGGACTGAACGGATTCC; and the Sanger sequencing primer was fwd, TGTTGGGTCAGGAAGAGACTTG.

Colony-forming assays. Single-cell suspensions were prepared from MEL202 and MEL270 cells with or without CRISPR-mediated insertions and deletions (indels), 3,000 cells from each cell line were plated in triplicates in 6-well treated plates and colonies were enumerated 10 days later. After 10 days, colonies were fixed with 3.7% paraformaldehyde for 5 min and stained in a solution of 0.05% crystal violet for 30 min at room temperature and washed in PBS and tap water.

In vitro cell viability assays. Cells were seeded in white flat-well 96-well plates (Costar) at a density of 1,000 cells per well. ATP luminescence readings were taken every 24 h after seeding, using Cell Titer Glo (Promega) according to the manufacturer's instructions.

***BRD9* minigene assay.** We identified putative exonic splicing enhancers with SpliceAid 2[39]. We used SF3B1 wild-type MEL270 and T47D cells transduced with Flag–SF3B1(WT) or Flag–SF3B1(K700E) (in the backbone of pInducer20 (Addgene no. 44012)). After drug selection with neomycin (Thermo Fisher Scientific 10131027), the selected cells were treated with 1 μg /ml doxycycline (Sigma D9891). The *BRD9* minigene construct was generated by inserting the DNA fragment containing the *BRD9* genomic sequence from exon 14 to exon 15 in between the BamHI and AgeI restriction sites in the FRE5 plasmid (Addgene 62377) via Gibson assembly. *BRD9* minigene mutagenesis was performed with the Agilent QuikChange II site-directed mutagenesis kit with specific primers according to the manufacturer's directions. For transient transfection experiments, cells were seeded into a 24-well plate one day before transfection of *BRD9* minigene constructs in the presence of X-tremeGENE HP DNA transfection reagent (Roche) according to the manufacturer's directions. Forty-eight hours after transfection, cells were collected and RNA was extracted using Qiagen RNeasy mini kit. Minigene-derived and endogenous BRD9 transcripts were analysed by RT–PCR using specific primers. Primers and oligonucleotides used in RT–PCR reactions were: cloning fwd GGACCCAGTACCAGGATCCGTTGGGGACACCCTAGGAG and rev CTTGGAGGAGCGCACACCGGTCAGGTGGTGCTGGTCCCTC; RT–PCR fwd (minigene) GGATTACAAGGATGACGATGAC, fwd (endogenous) CATGAAGCCTCCAGATGAAG and rev (common) CTTCGTCGT CTCATCCAAGTTC; mutagenesis (AAAA to TTTT) fwd CAAAGGGATATA

TTTTGAGATTTTTTACTCAGTTTCTTTCAGG and rev CCTGAAAG AAACTGAGTAAAAAATCTCAAAATATATCCCTTTG; mutagenesis (AAAA to ATAT) fwd CAAAGGGATATATTTTGAGATATATTACTCAGTTTCTTTCAGG and rev CCTGAAAGAAACTGAGTAATATATCTCAAAATATATCCCTTTG; mutagenesis (GATAAAA to TTTTTTT) fwd GTCAAAGGGATATATTTT GATTTTTTTTACTCAGTTTCTTTCAGG and rev ; mutagenesis (TTTCT to TTACA at exon 14a) fwd GAGATAAAATACTCAGTTACATTCAG GGCCTGCCATCTATC and rev GATAGATGGCAGGCCCTGAATG TAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTGCG at exon 14a) fwd GAGATAAAATACTCAGTTGCGTTCAGGGCCTGCCATCTATC, and rev GATAGATGGCAGGCCCTGAACGCAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTTCA at exon 14a) fwd GAGATAAAATAC TCAGTTTCATTCAGGGCCTGCCATCTATC and rev GATAGATGGCAG GCCCTGAATGAAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTTCG at exon 14a) fwd GAGATAAAATACTCAGTTTCGTTCAGG GCCTGCCATCTATC, and rev GATAGATGGCAGGCCCTGAAC GAAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTTAT at exon 14a) fwd GAGATAAAATACTCAGTTTATTTCAGGGCCTGCCATCTATC and rev GATAGATGGCAGGCCCTGAAATAAACTGAGTATTTTATCTC; and mutagenesis (branchpoint A to G) fwd GTCAAAGGGATATATTTTGG GATAAAATACTCAGTTTCTTTC and rev GAAAGAAACTGAGTAT TTTATCCCAAAATATATCCCTTTGAC.

Lariat sequencing. To map the branchpoints that were used when the *BRD9* poison exon was included or excluded, RT–PCR was performed to amplify branchpoint-spanning fragments from lariat RNAs arising during normal (poison exon exclusion) or aberrant (poison exon inclusion) splicing of *BRD9* pre-mRNA. In brief, SuperScript III reverse transcriptase (Invitrogen) and a primer complementary to the intronic sequences downstream of the 5′ splice sites were used to generate cDNA from lariat RNAs. Branchpoint-spanning fragments were then amplified from lariat RNAs by nested PCR with pairs of outer primers (with the RT primer being the reverse primer) and inner primers. The forward primers were complementary to sequences about 200–300 nucleotides upstream of the 3′ splice sites and the reverse primers were complementary to sequences downstream of the 5′ splice sites (Supplementary Table 11). The PCR products were cloned into the pGEM-T vector (Promega) and sequenced by Sanger sequencing.

Melanoma transplant model. We resuspended stably transduced Melan-a, MEL270 and MEL202 cells (1M cells) with doxycycline-inducible shRNAs, cDNAs or sgRNAs in 100 μl of a 1:1 mix of medium and Matrigel (BD Biosciences), and subcutaneously and bilaterally injected the mix into the flanks of 7-week-old female CB17-SCID mice (Taconic). For doxycycline-regulated shRNA induction, we used doxycycline-containing diets (625 mg/kg diet, Envigo). To assess tumour growth, at least five mice per group were injected for a total of ten tumours per group. No randomization or blinding was used in the analysis of tumour growth. Tumours were measured with callipers every seven days. Growth curves were visualized with Prism GraphPad 8.0. Tumour volume was calculated using the formula; Volume = π(length)(width)(height)/6.

In vitro morpholino transfection. To deliver morpholinos into cultured cell lines, we followed the manufacturer's instruction (GeneTools). In brief, we used 6 μM Endo-Porter after adding morpholinos (final concentration of 10 μM). RNA and proteins are collected 48 h after delivery. Morpholino target sequence no. 3 was TAATGAGGCAAGTCCAGTCCCGCTT; no. 6 was AAAGAGGGGAT AATGAGGCAAGTCC; and no. 7 was GGGATAATGAGGCAAGTCCAGTCCC.

In vivo morpholino treatment. Treatment with morpholinos was started when the tumour volume in mice reached 100–200 mm^3. Cohorts were treated intratumorally with 12.5 mg/kg scrambled or poison-exon-targeting Vivo-Morpholinos (AAAGAGGGGATAATGAGGCAAGTCC, GeneTools) dissolved in 50 μl PBS, every 2 days for 8 doses in total. The mice were dissected 24 h after the final treatment. For the patient-derived xenograft model, patient-derived rectal melanoma cells (*SF3B1*^{R625C}) and UVM cells (*SF3B1* wild type) were serially transplanted into SCID mice and treated similarly.

In vivo metastasis model. For lung experimental metastasis, B16 and 92.1 melanoma cells retrovirally transduced with shRNAs targeting *Renilla*, *BRD9* (no. 1 and no. 1) or *Brd9* (no. 1 or no. 2) in MLS-E vector (sorted using GFP) were trypsinized, resuspended in PBS and then 0.4 M cells in 0.2 ml PBS were injected via the lateral tail vein using a 27-gauge needle. Mice were killed 14 days after injection and tissues were isolated and fixed in 4% paraformaldehyde (Thermo Fisher Scientific). For evaluation of metastatic colonization of the lung using 92.1 human UVM cells, the burden of metastatic cells was evaluated using GFP expression by flow cytometry as well as anti-GFP immunohistochemistry 14 days following tail-vein injection of 0.4 M cells into NOD–SCID *Il2rg*$^{-/-}$ mice.

Histological analysis. Tissues were fixed in 4% paraformaldehyde, processed routinely in alcohol and xylene, embedded in paraffin, sectioned at 5-μm thickness and stained with H&E. Immunohistochemistry was performed on a Leica Bond RX automated stainer (Leica Biosystems). Following heat-induced epitope retrieval at pH 6.0, the primary antibody against Ki67 (Vector VP-K451) was applied, followed by application of a polymer detection system (DS9800, Novocastra Bond Polymer Refine Detection, Leica Biosystems) in which the chromogen was 3,3-diaminobenzidine tetrachloride (DAB) and the counterstain was haematoxylin. Photomicrograph examination of all H&E and immunohistochemistry slides were performed using a Zeiss Axioskop imaging.

***BRD9* expression correlates.** The cor.test (in R) was used to calculate Spearman's ρ and the P value associated with the correlation of *BRD9* expression with the expression of each coding gene across all samples within each cohort from the TCGA. Analysis was restricted to coding genes that are not on the same chromosome arm as *BRD9* (chromosome 5p) to remove potential confounding effects of local correlations. Coding genes with $P < 0.01$ in at least 10 cancer types were ranked by their absolute mean value of ρ (computed across all TCGA cohorts) and classified as RNA-binding if they were annotated with the 'RNA-binding' Gene Ontology term (GO: 0003723).

***BRD9* alternative splicing.** Potential NMD-targeted isoforms of *BRD9* were identified as follows: we queried the MISO v.2.0 alternative splicing annotation[40] for exon skipping and competing splice site events within the *BRD9* gene locus, restricted to those events with evidence of alternative splicing based on spliced junction reads (described in 'Genome annotation, RNA-seq read mapping, and estimation of gene and isoform expression'), assigned open reading frames for the isoforms resulting from each alternative splicing event based on the *BRD9* isoform with the longest open reading frame, and classified isoforms as predicted NMD substrates if they contained a termination codon >50 nt upstream of a splice junction.

Robust linear modelling of *BRD9* expression on the basis of the identified alternatively spliced isoforms of *BRD9* that are predicted NMD substrates was performed for each TCGA cohort with the rlm function in the MASS package in R. Relative expression of each isoform in each sample was estimated from RNA-seq data across all TCGA cohorts as described in 'Genome annotation, RNA-seq read mapping, and estimation of gene and isoform expression'. A z-score normalization was performed across all samples for each isoform in each cohort before model fitting. The resulting coefficients from the fitted models were subsequently used to predict *BRD9* expression from *BRD9* NMD-targeted isoform expression.

RNA-seq library preparation. RNA-seq libraries were prepared from TRIzol-isolated (Thermo Fisher cat. no. 15596026) RNA using the Illumina TruSeq RNA Library Prep Kit v.2 (Illumina cat. no. RS-122-2001/2). K562 libraries were sequenced at MSKCC with 101-bp single-end reads. MEL270 libraries were sequenced by the FHCRC Genomics Shared Resource with 2 × 51-bp paired-end reads.

ChIP–seq library preparation. ChIP–seq libraries were prepared and sequenced as previously described[6] by the Molecular Biology Core Facilities at the Dana-Farber Cancer Institute with 75-bp single-end reads.

Genomic analysis of SWI–SNF complex members from TCGA. Mutational analysis of genes encoding members of the SWI–SNF complex was performed as previously described[41].

Genome annotation, RNA-seq read mapping, and estimation of gene and isoform expression. RNA-seq reads were processed for gene expression and isoform ratio quantification as previously described[42]. In brief, RNA-seq reads were aligned to the hg19/GRCh37 assembly of the human genome using a gene annotation created by merging the UCSC knownGene gene annotation[43], Ensembl v.71.1 gene annotation[44] and MISO v.2.0 isoform annotation[40]. Read alignment and expression estimation were performed with RSEM v.1.2.4[45], Bowtie v.1.0.0[46] and TopHat v.2.1.1[47]. Isoform ratios were quantified with MISO v.2.0[40]. Gene expression estimates were normalized by applying the trimmed mean of M values method[48] to coding genes. Statistical tests for differential gene and isoform expression were performed for single-sample comparisons with Wagenmakers' Bayesian framework[49] and for sample group comparisons with the Mann–Whitney U-test. RNA-seq read-coverage plots (for example, Fig. 1e) represent reads normalized by the number of reads mapping to all coding genes in each sample (per million).

RNA-seq coverage plots. RNA-seq coverage plots were made using the UCSC Genome Browser[50] and/or the ggplot2 package in R[51]. Repetitive elements were annotated by RepeatMasker[28].

Cluster analysis. Unsupervised clustering of chronic lymphocytic leukaemia, myelodysplastic syndrome and UVM samples (Fig. 1a) was based on the 40 events that were differentially spliced in isogenic UVM (MEL270 cells) as well as myeloid leukaemia (K562 cells) cells expressing *SF3B1*^{K700E} versus wild-type *SF3B1*, restricted to the 30 of these events that had sufficient read coverage in all cohorts for clustering.

ChIP–seq data analysis. ChIP–seq reads were mapped to the genome by calling Bowtie v.1.0.0[46] with the arguments '-v 2 -k 1 -m 1–best–strata'. Peaks were called using MACS2 v.2.1.1.20160309[52] against input control libraries with $P < 10^{-5}$ and subsequently filtered to remove peaks contained within ENCODE blacklisted regions[53] and the mitochondrial genome. Subsequent data analysis was performed with Bioconductor in the R programming environment[54]. Consensus

peaks between samples were called using the soGGI package v.1.14.0[55]. Peaks were annotated using the ChIPseeker package v.1.18.0[56]. Potential transcription factor binding in a 300-nucleotide region around the centre of consensus peaks was scored using the TFBSTools package v.1.20.0[57], with models taken from the HOCOMOCO v.11 human core collection[58] and applied with a threshold of $P < 10^{-4}$. The highest scores for each consensus peak region were collated for each transcription factor. A two-sided Mann–Whitney *U*-test was used to assess the significance of the difference in scores between constitutive and sensitive peaks for each transcription factor.

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

Data availability

RNA-seq and ChIP–seq data generated as part of this study were deposited in the Gene Expression Omnibus (accession number GSE124720). RNA-seq data from published studies were downloaded from CGHub (TCGA UVM[59]), EMBL-EBI ArrayExpress (Illumina Human BodyMap 2.0: E-MTAB-513), the Gene Expression Omnibus (accession numbers GSE72790 and GSE114922 for chronic lymphocytic leukaemia[15] and myelodysplastic syndromes[27], respectively), or directly obtained from the authors (for UVM[10]). Gel source data can be found in Supplementary Fig. 1. Other data that support the findings of this study are available from the corresponding authors upon reasonable request.

29. Cheng, D. T. et al. Memorial Sloan Kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J. Mol. Diagn.* **17**, 251–264 (2015).
30. Seiler, M. et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **24**, 497–504 (2018).
31. Leeksma, A. C. et al. Clonal diversity predicts adverse outcome in chronic lymphocytic leukemia. *Leukemia* **33**, 390–402 (2019).
32. Kim, E. et al. *SRSF2* mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell* **27**, 617–630 (2015).
33. Mohammed, H. et al. Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. *Cell Rep.* **3**, 342–349 (2013).
34. Fellmann, C. et al. An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).
35. Pelossof, R. et al. Prediction of potent shRNAs with a sequential classification algorithm. *Nat. Biotechnol.* **35**, 350–353 (2017).
36. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
37. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
38. Wu, D. & Smyth, G. K. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res.* **40**, e133 (2012).
39. Piva, F., Giulietti, M., Burini, A. B. & Principato, G. SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. *Hum. Mutat.* **33**, 81–85 (2012).
40. Katz, Y., Wang, E. T., Airoldi, E. M. & Burge, C. B. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods* **7**, 1009–1015 (2010).
41. Mashtalir, N., et al. Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell* **175**, 1272–1288.e1220 (2018).
42. Dvinge, H. et al. Sample processing obscures cancer-specific alterations in leukemic transcriptomes. *Proc. Natl Acad. Sci. USA* **111**, 16802–16807 (2014).
43. Meyer, L. R. et al. The UCSC genome browser database: extensions and updates 2013. *Nucleic Acids Res.* **41**, D64–D69 (2013).
44. Flicek, P. et al. Ensembl 2013. *Nucleic Acids Res.* **41**, D48–D55 (2013).
45. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
46. 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).
47. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105–1111 (2009).
48. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
49. Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H. & Grasman, R. Bayesian hypothesis testing for psychologists: a tutorial on the Savage-Dickey method. *Cognit. Psychol.* **60**, 158–189 (2010).
50. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
51. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, 2016).
52. Zhang, Y. et al. Model-based analysis of ChIP–seq (MACS). *Genome Biol.* **9**, R137 (2008).
53. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
54. Huber, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).
55. Dharmalingam G. & Carroll, T. soGGi: Visualise ChIP–seq, MNase-seq and motif occurrence as aggregate plots summarised over grouped genomic intervals. R package version 1.14.0, https://rdrr.io/bioc/soGGi/ (2018).
56. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
57. Tan, G. & Lenhard, B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32**, 1555–1556 (2016).
58. Kulakovskiy, I. V. et al. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018).
59. Robertson, A.G., et al. Integrative analysis identifies four molecular and clinical subsets in uveal melanoma. *Cancer cell* **32**, 204–220.e215 (2017).
60. Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010).
61. McLean, C. Y. et al. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

Acknowledgements We are grateful for the support of the MSK RNAi core facility for help with the CRISPR screens performed in the study, the MSK anti-tumor assessment core facility for help with patient-derived xenograft experiments and the Genomics Shared Resource of the Fred Hutchinson/University of Washington Cancer Consortium (P30 CA015704). D.I., S.C.-W.L., A.Y. and O.A.-W. are supported by the Leukemia & Lymphoma Society. D.I. is supported by grants from Lydia O'Leary Memorial Pias Dermatological Foundation and Kobayashi Foundation for Cancer Research. A.Y. is supported by grants from the Aplastic Anemia and MDS International Foundation (AA&MDSIF) and the Lauri Strauss Leukemia Foundation. S.X.L. is supported by a Conquer Cancer Foundation and ASCO Young Investigator Award, an Aplastic Anemia & Myelodysplastic Syndrome International Foundation research award, as well as an AACR Lymphoma Research Fellowship. G.-L.C. is a Mahan Fellow. O.A.-W. is supported by the Pershing Square Sohn Cancer Research Alliance, the Henry & Marilyn Taub Foundation and the Starr Cancer Consortium. R.K.B. is a Scholar of The Leukemia and Lymphoma Society (1344-18) and is supported in part by the US National Institutes of Health (R01 DK103854). O.A.-W. and R.K.B. are supported by the Evans MDS Foundation, the US National Institutes of Health (R01 HL128239) and the Department of Defense Bone Marrow Failure Research Program (BM150092 and W81XWH-12-1-0041). The results shown here are in part based upon data generated by the TCGA research network (https://cancergenome.nih.gov/).

Author contributions D.I., G.-L.C., B.L., B.C.M., C.K., O.A.-W. and R.K.B. designed the study. D.I., B.L. and J.P. performed minigene assays. D.I., B.C.M. and C.K. performed ChIP–seq experiments. A.R.D., G.-L.C. and C.K. performed ChIP–seq analyses. G.-L.C., K.N., A.P. and R.K.B. performed computational analyses of the CRISPR screen, RNA splicing and gene expression data. D.I. and S.X.L. performed the CRISPR screen. D.I. and B.L. generated and validated the anti-BRD9 PE morpholinos in in vitro assays. D.I., L.B., A.B., S.C.-W.L., A.Y. and H.C. performed the mouse experiments. A.R.M., D.I. and L.E.-H. performed experiments using pancreatic cancer cell lines. T.H. and Y.C. provided melanoma cell lines. D.I., G.-L.C., O.A.-W. and R.K.B. prepared the manuscript with help from all co-authors. J.T. and O.A.-W. provided clinically annotated samples from the MSK Hematology/Oncology Tissue Bank and from the MSK Antitumor Assessment Core Facility. O.A.-W. and R.K.B. provided funding and study supervision.

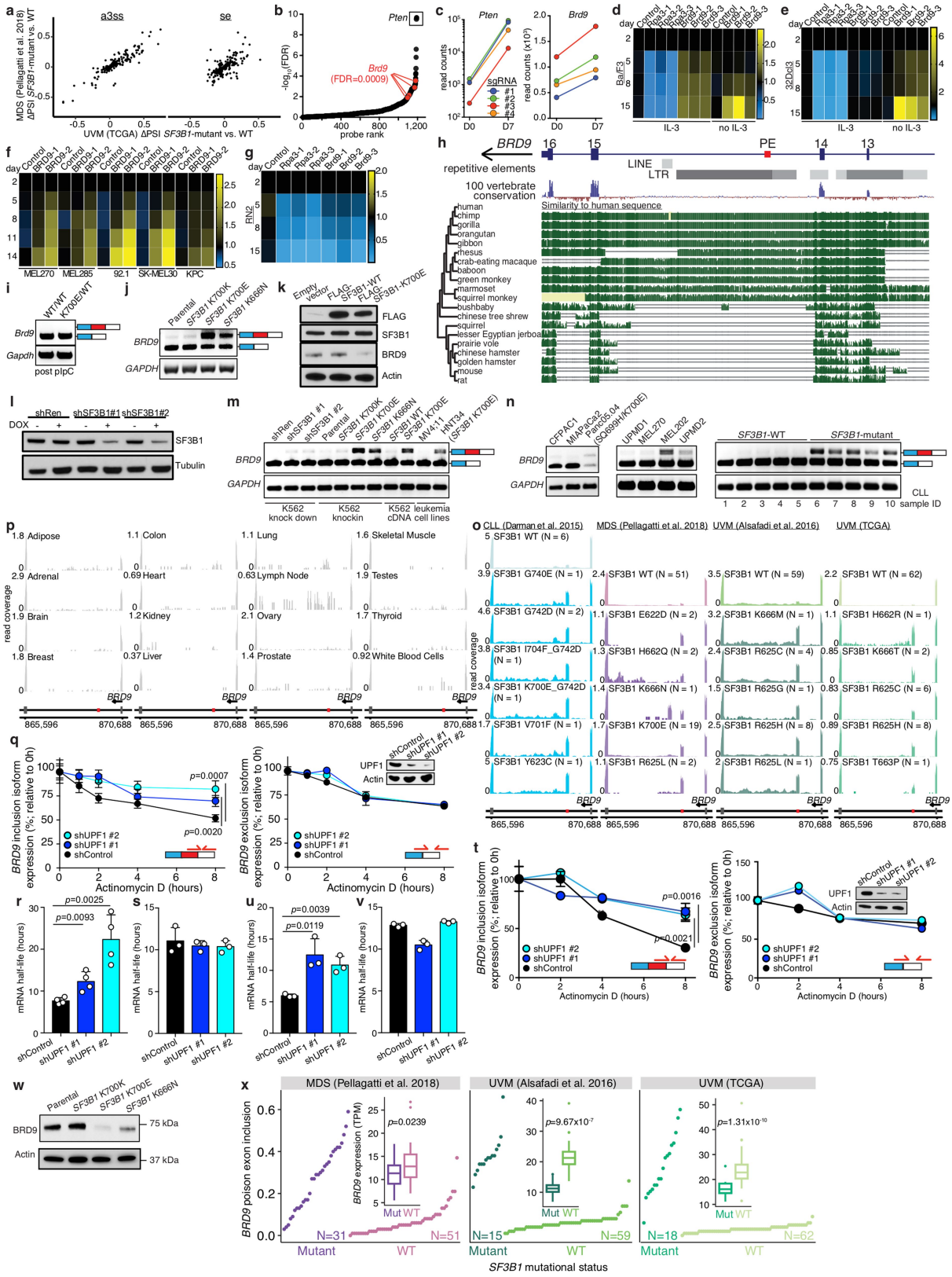
Competing interests C.K. is a Scientific Founder, fiduciary Board of Directors member, Scientific Advisory Board member, consultant and shareholder of Foghorn Therapeutics, none of which are related to the current manuscript. O.A.-W. has served as a consultant for H3 Biomedicine, Foundation Medicine, Merck and Janssen; O.A.-W. has received personal speaking fees from Daiichi Sankyo. O.A.-W. has received previous research funding from H3 Biomedicine unrelated to the current manuscript. D.I., O.A.-W. and R.K.B. are inventors on a provisional patent application submitted by the Fred Hutchinson Cancer Research Center that covers BRD9 activation in cancer.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1646-9.

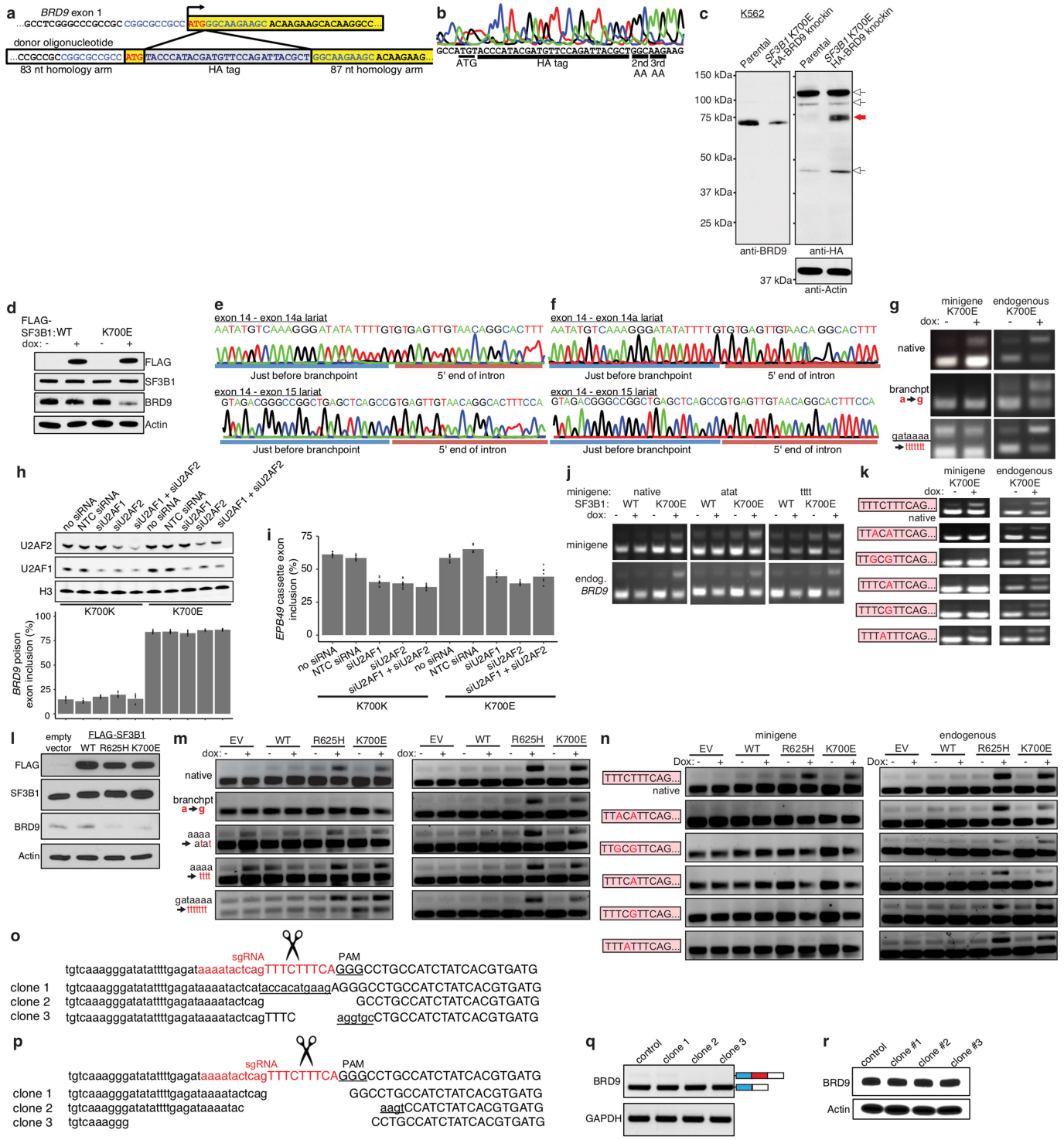
Correspondence and requests for materials should be addressed to O.A. or R.K.B.

Peer review information *Nature* thanks Boris Bastian, Rotem Karni and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



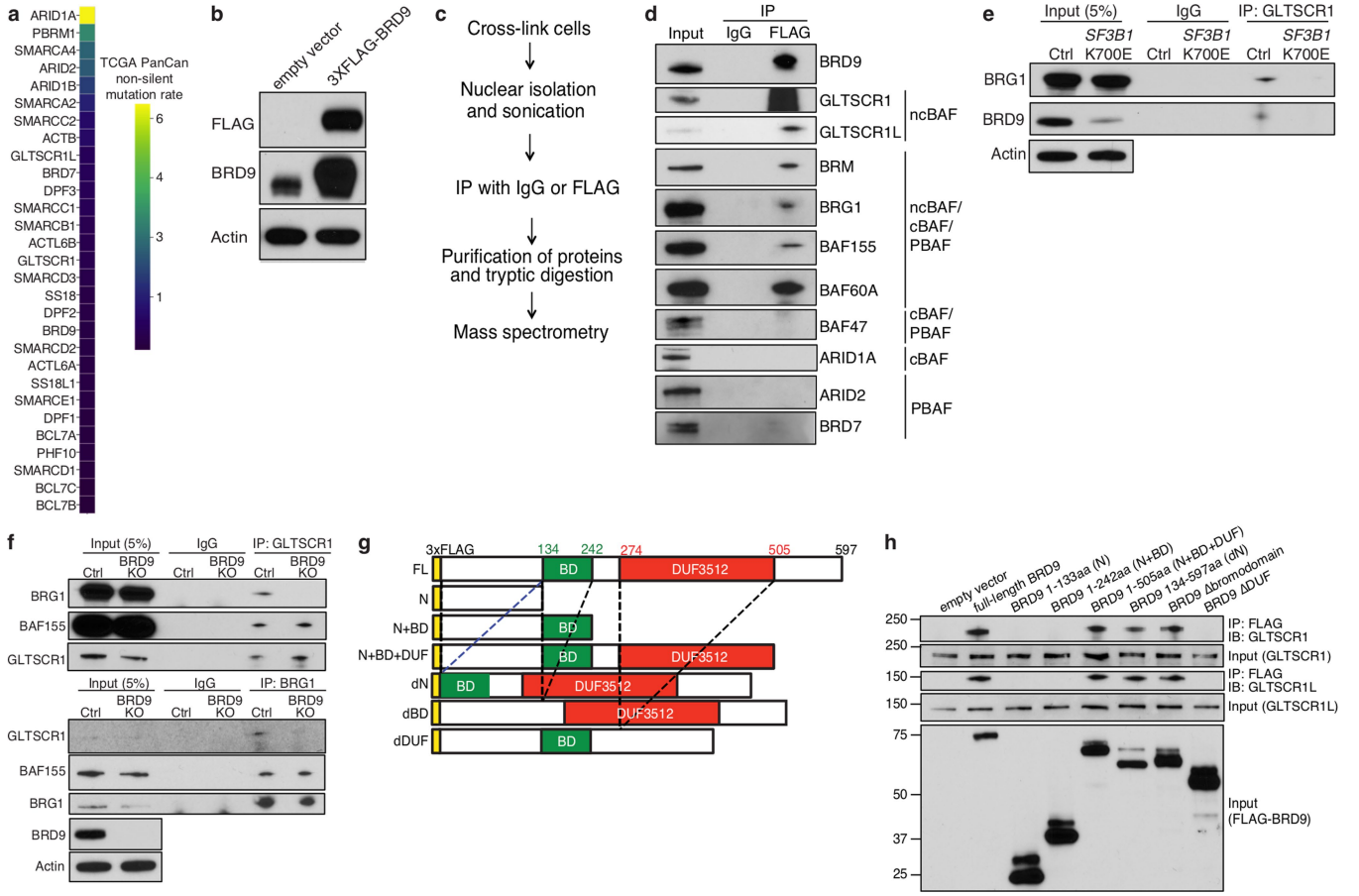
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | *BRD9* is mis-spliced in *SF3B1*-mutated human cells, and BRD9 loss confers a proliferative advantage. a, Scatter plots comparing differential splicing (ΔPSI) between patients in the TCGA UVM cohort with *SF3B1* mutations or wild-type *SF3B1*[59] (*x* axis) and patients from a myelodysplastic syndromes cohort with *SF3B1* mutations or wild-type *SF3B1*[27] (*y* axis). Events were classified as alternative 3′ splice sites or skipped exons. **b**, Rank plot for the $-\log_{10}$(FDR) associated with each sgRNA in our CRISPR–Cas9 positive-selection screen. sgRNAs targeting the positive control (*Pten*) and *Brd9* are highlighted. For the probe-level (per-sgRNA) analysis, we fitted a negative binomial generalized log-linear model and performed a likelihood ratio test. FDR values were computed using the Benjamini–Hochberg method. **c**, Read counts for sgRNAs targeting the positive control (*Pten*) or *Brd9*. D0 and D7 indicate days following withdrawal of IL-3. **d**, Heat map summarizing the results of a competition assay to measure the effect of each indicated sgRNA on the growth of Cas9-expressing Ba/F3 cells. Cell growth was computed with respect to cells treated with a non-targeting (control) sgRNA and the percentages of GFP^+ cells on day 14 were normalized to the percentages on day 2. The illustrated values correspond to the mean computed value over $n = 3$ biological replicates. *Rpa3* sgRNAs were used as a negative control. **e**, As in **d**, but for 32Dcl3 cells. **f**, As in **d**, but for the indicated melanoma and pancreatic ductal adenocarcinoma cells. **g**, As in **d**, but for RN2 cells. **h**, Sequence conservation of the *BRD9* poison exon locus as estimated by phyloP[60]. Conservation and repetitive element annotation is from the UCSC Genome Browser[43]. **i**, RT–PCR analysis of *Brd9* poison exon inclusion using whole bone marrow cells from *Mx1-cre Sf3b1*$^{WT/WT}$ (WT/WT) and *Mx1-cre Sf3b1*$^{K700E/WT}$ (K700E/WT) mice. Three weeks after pIpC treatment, RT–PCR was performed with mouse primers corresponding to those used to assay *BRD9* poison exon inclusion in human cells. Representative images from $n = 2$ technically independent replicates. **j**, RT–PCR analysis to confirm mutant-SF3B1-dependent inclusion of the *BRD9* poison exon in isogenic NALM-6 cell lines engineered to contain the indicated mutations. *SF3B1*^{K700K} is a wild-type control for genome engineering. Representative images from $n = 2$ technically independent replicates. **k**, Western blot for Flag, SF3B1 and BRD9 in K562 cells overexpressing N-terminally Flag-tagged wild-type *SF3B1* or *SF3B1*^{K700E} cDNAs, or an empty vector; this panel corresponds to the cells evaluated in **m**. Representative images from $n = 2$ biologically independent replicates. **l**, Western blot for SF3B1 in K562 cells treated with doxycycline-inducible *SF3B1*-targeting shRNAs or a non-targeting control shRNA (shRen); this panel corresponds to cells evaluated in **m**. Representative images from $n = 2$ technically independent replicates. **m**, RT–PCR illustrating the specificity of *BRD9* poison exon inclusion for *SF3B1*-mutated cells in the indicated cell lines. These include K562 cells treated with control shRNA (shRen) or *SF3B1*-targeting shRNAs (the columns labelled 'K562 knock-down'); knock-in of the *SF3B1*^{K700K}, *SF3B1*^{K700E} or *SF3B1*^{K666N} mutation into the endogenous locus of *SF3B1* (the columns labelled 'K562 knock-in'); or overexpression of wild-type *SF3B1* or *SF3B1*^{K700E} cDNA (the columns labelled 'K562 cDNA'). The two right-most lanes show acute myeloid leukaemia cell lines with wild-type *SF3B1* (MV4;11) or a naturally occurring endogenous *SF3B1*^{K700E} mutation (HNT34 cells; the columns labelled 'leukaemia cell lines'). Representative images from $n = 3$ biologically independent experiments. **n**, As in **m**, but for the indicated pancreatic ductal adenocarcinoma cell lines (left), UVM cell lines (centre) and a cohort of patients with chronic lymphocytic leukaemia (right). CFPAC1 and MIA PaCa2 cells lack *SF3B1* mutations; Panc05;04 cells carry *SF3B1*$^{Q699H/K700E}$; UPMD1 and MEL270 cells lack *SF3B1* mutations; MEL202 and UPMD2 cells carry *SF3B1*^{R625G} and *SF3B1*^{Y623H} mutations, respectively. Sample identifiers for patients with chronic lymphocytic leukaemia correspond to the genotypes shown in Supplementary Table 7. Representative images from $n = 2$ technically independent experiments (left and centre) and $n = 3$ biologically independent experiments (right). **o**, RNA-seq read coverage plots of the *BRD9* poison exon locus from patient samples with the indicated *SF3B1* genotypes. All *SF3B1*-mutated samples exhibit *BRD9* poison exon inclusion. **p**, As in **o**, but for the indicated tissues from healthy donors (from BodyMap 2.0). **q**, qRT–PCR measurement of the half-lives of the poison exon inclusion (left) and exclusion (right) isoforms in isogenic K562 *SF3B1*^{K700E} cells treated with the indicated shRNAs and actinomycin D to inhibit transcription. NMD inhibition via *UPF1* knockdown stabilizes the inclusion, but not exclusion, isoform. Red arrows indicate primers used to specifically detect the two isoforms. $n = 2$ biologically independent experiments and $n = 2$ technically independent experiments for the inclusion isoform; $n = 3$ technically independent experiments for the exclusion isoform. *P* value was calculated by two-sided *t*-test at 8 h. **r**, Bar graph illustrating the estimated poison exon inclusion isoform half-life in the indicated conditions from the data in **q**. Error bars, mean + s.d. $n = 2$ biologically independent experiments and $n = 2$ technically independent experiments. *P* value was calculated by two-sided *t*-test. **s**, As in **r**, but for the exclusion isoform. Error bars, mean + s.d. $n = 3$ technically independent experiments. *P* value was calculated by two-sided *t*-test. **t**, As in **q**, but for NALM-6 *SF3B1*^{K700E} cells. $n = 3$ technically independent experiments for the inclusion isoform and the exclusion isoform. *P* value was calculated by two-sided *t*-test at 8 h. **u**, As in **r**, but for NALM-6 *SF3B1*^{K700E} cells. $n = 3$ technically independent experiments. Error bars, mean + s.d. *P* value was calculated by two-sided *t*-test. **v**, As in **s**, but for NALM-6 *SF3B1*^{K700E} cells. Error bars, mean + s.d. $n = 3$ technically independent experiments. *P* value was calculated by two-sided *t*-test. **w**, Western blot for BRD9 in NALM-6 cells with or without knock-in of an *SF3B1* mutation. Actin, loading control. Representative images from $n = 3$ biologically independent experiments. **x**, Rank plot of *BRD9* poison exon inclusion (scale of 0 to 1; top) and box plot of gene expression (inset) for patients stratified by *SF3B1* mutational status (data are from cohorts of patients with myelodysplastic syndromes or UVM, as in Fig. 1a). *SF3B1* mutations were strongly associated with high poison exon inclusion and low *BRD9* expression. Boxes illustrate 1st and 3rd quartiles, with whiskers extending to 1.5× interquartile range. *P* value computed with one-sided Mann–Whitney *U*-test.

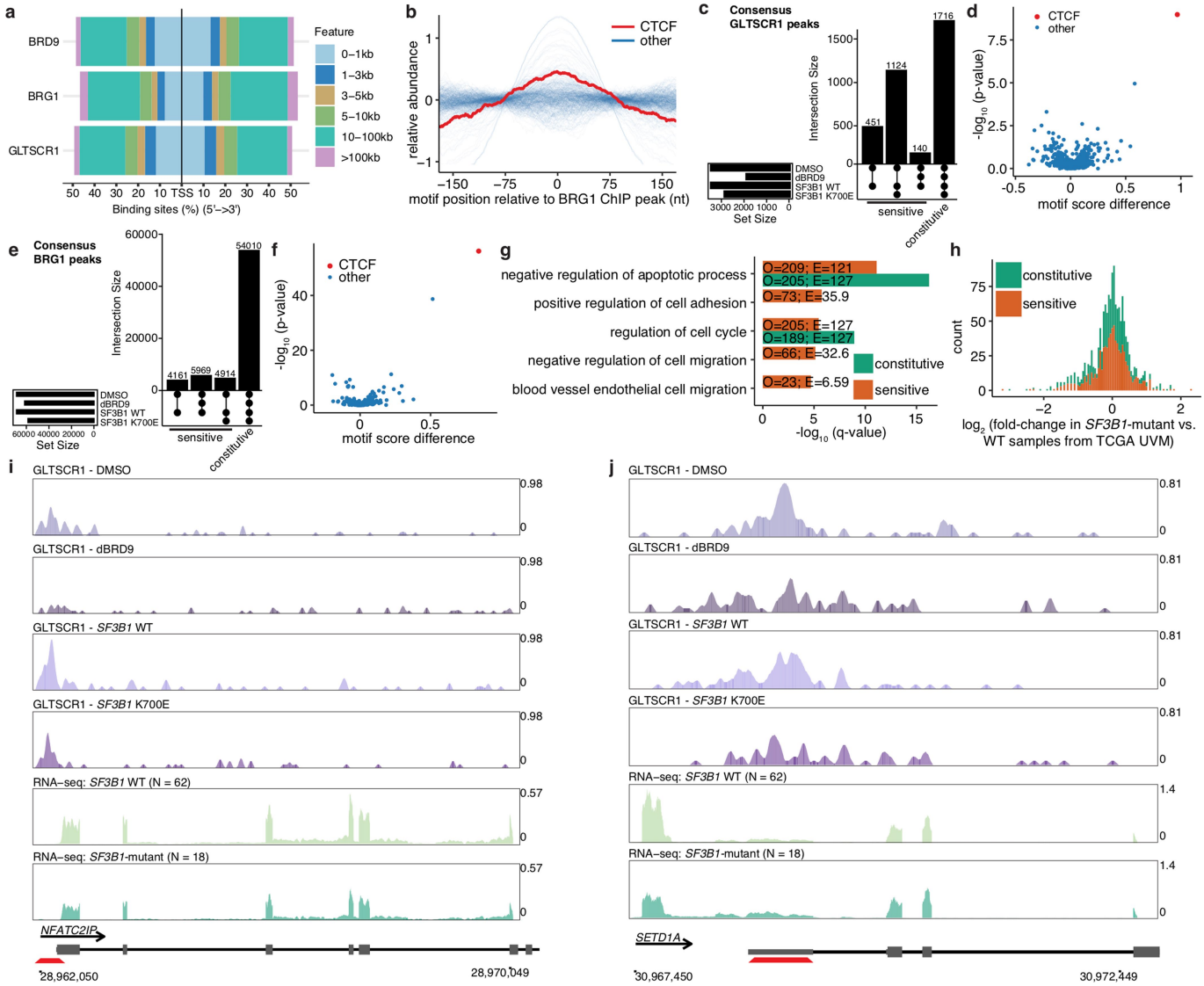


Extended Data Fig. 2 | See next page for caption.

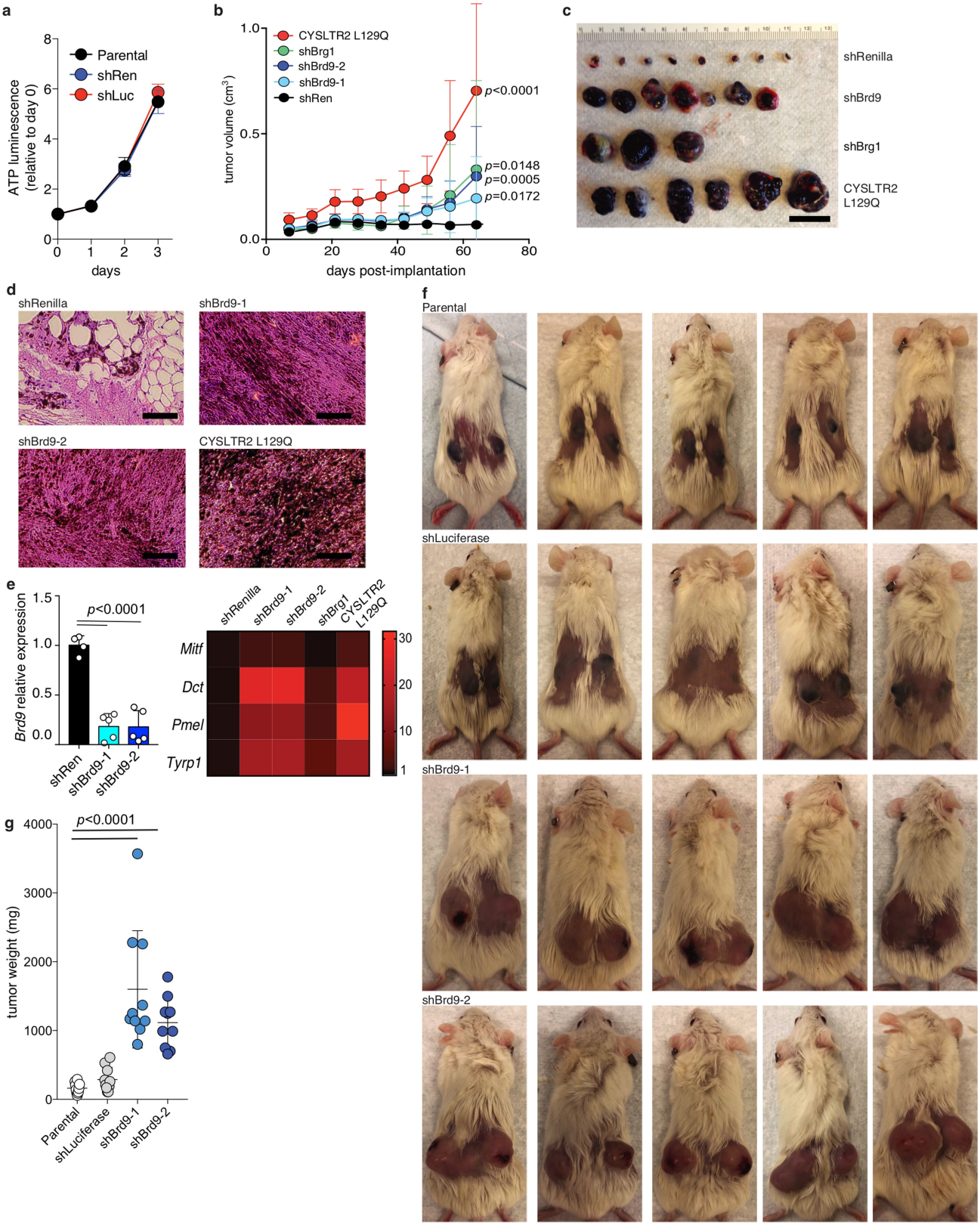
Extended Data Fig. 2 | Mutant SF3B1 recognizes an aberrant branchpoint within *BRD9* to promote poison exon inclusion, causing loss of full-length BRD9 protein. a, Schematic illustrating the strategy for knock-in of an HA tag into the endogenous *BRD9* locus. The single-stranded donor DNA contained a 197-nt fragment, including 83 nt homologous to the *BRD9* 5′ UTR (upstream of the HA tag) and 87 nt homologous to *BRD9* exon 1 (downstream of the start codon). **b**, Sanger sequencing of genomic DNA validating successful HA tag knock-in in K562 *SF3B1*^{K700E} cells. Representative images from $n = 2$ biologically independent experiments. **c**, Western blot with anti-BRD9 (left), anti-HA (right, top) or anti-actin (right, bottom) used to probe K562 *SF3B1*^{K700E} cells carrying an endogenously N-terminally HA-tagged BRD9. White arrows, non-specific bands. Red arrow, expected size of BRD9 protein. Representative images from $n = 3$ biologically independent experiments. **d**, Western blot for Flag, SF3B1 and endogenous BRD9 protein in MEL270 cells with doxycycline-inducible Flag-tagged wild-type SF3B1 or Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **e**, Sanger sequencing of cDNA arising from reverse transcription of lariats arising from inclusion (top) (exon 14–exon 14a splicing) or exclusion (bottom) (exon 14–exon 15 splicing) of the *BRD9* poison exon in MEL270 cells with doxycycline-inducible Flag-tagged wild-type SF3B1 (bottom) or Flag-tagged SF3B1(K700E) (top). The branchpoints are illustrated in Fig. 2a. Representative images from $n = 3$ biologically independent experiments. **f**, As in **e**, but for T47D cells. Representative images from $n = 3$ biologically independent experiments. **g**, As in Fig. 2b, but for the indicated minigene mutagenesis in T47D cells with doxycycline-inducible Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **h**, Western blot of U2AF2, U2AF1 and histone H3 in K562 cells transfected with siRNAs against *U2AF1* and/or *U2AF2* (top) and bar plot illustrating mean *BRD9* poison exon inclusion as measured by quantitative PCR (qPCR) following siRNA knockdown of *U2AF1* and/or *U2AF2* (bottom). Experiment performed with $n = 1$ biologically independent replicate for siRNA transfection, $n = 1$ technically independent replicate for western blot and $n = 3$ technically independent replicates for RT–PCR. Poison exon inclusion was computed over all $n = 3 \times 3$ (9) combinations of technical replicates for RT–PCR for the inclusion and exclusion isoforms. Bars illustrate mean inclusion. **i**, *EPB49* cassette exon inclusion as measured by qPCR following siRNA knockdown of *U2AF1* and/or *U2AF2*. As the *EPB49* cassette exon is U2AF-dependent, this experiment serves as a positive control for the functional efficacy of *U2AF1* and *U2AF2* knockdown. $n = 3$ technically independent experiments. Cassette exon inclusion was computed over all $n = 3 \times 3$ (9) combinations of technical replicates for RT–PCR for the inclusion and exclusion isoforms. Bars illustrate mean inclusion. **j**, As in Fig. 2b, but for the indicated minigene mutagenesis in T47D cells with doxycycline-inducible Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **k**, As in Fig. 2c, but for the indicated minigene mutagenesis in T47D cells with doxycycline-inducible Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **l**, Western blot for Flag, SF3B1, BRD9 and actin in MEL270 cells expressing an empty vector or N-terminally Flag-tagged wild-type *SF3B1*, *SF3B1*^{R625H} or *SF3B1*^{K700E} cDNA. Representative images from $n = 3$ biologically independent experiments. **m**, RT–PCR analysis of *BRD9* splicing in MEL270 cells expressing doxycycline-inducible empty vector, wild-type SF3B1, SF3B1(R625H) or SF3B1(K700E). The left column illustrates minigene splicing and the right column illustrates endogenous *BRD9* splicing. Representative images from $n = 3$ biologically independent experiments. **n**, As in **m**, but for the illustrated minigene mutations at the 5′ end of the poison exon. Representative images from $n = 3$ biologically independent experiments. **o**, Mutations generated at the 5′ end of the *BRD9* poison exon by CRISPR–Cas9-mediated indels in MEL202 cells (*SF3B1*^{R625G}). The PAM sequence is illustrated with uppercase, underlined nucleotides. Red nucleotides hybridize to the sgRNA. Substitutions are illustrated with lowercase, underlined nucleotides. **p**, As in **o**, but for MEL270 cells. Representative images from $n = 3$ biologically independent experiments. **q**, As in Fig. 2d top, but for MEL270 cells. Representative images from $n = 3$ biologically independent experiments. **r**, As in Fig. 2d bottom, but for MEL270 cells. Representative images from $n = 3$ biologically independent experiments.



Extended Data Fig. 3 | BRD9 loss impairs ncBAF complex formation. **a**, Mutation rate observed across TCGA cohorts for canonical BAF, polybromo-associated BAF and ncBAF components. **b**, Western blot confirming Flag-tagged BRD9 protein expression in 3×Flag–BRD9-expressing K562 cells. Representative images from $n = 3$ biologically independent experiments. **c**, Experimental workflow for using rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME)[33] for purification and identification of the chromatin-associated interactions partners of BRD9. **d**, Cross-linking and immunoprecipitation with IgG or Flag followed by probing with the indicated antibodies. Data from 3×Flag–BRD9-expressing MEL270 cells. Representative images from $n = 3$ biologically independent experiments. **e**, Immunoprecipitation of GLTSCR1 followed by western blotting with the indicated antibodies in *SF3B1*^{K700E} knock-in NALM-6 cells. Representative images from $n = 3$ biologically independent experiments. **f**, Immunoprecipitation of GLTSCR1 (top) or BRG1 (bottom) followed by blotting with the indicated antibodies in K562 cells with CRISPR-mediated knockout (KO) of BRD9. Representative images from $n = 3$ biologically independent experiments. **g**, Schematic of the BRD9 full-length (FL) protein and the deletion mutants that we constructed. BD, bromodomain; DUF, domain of unknown function; N, amino acids 1–133 of BRD9; N + BD, amino acids 1–242 of BRD9; N + BD + DUF, amino acids 1–505 of BRD9; dN, amino acids 134–597 of BRD9; dBD, bromodomain-deletion mutant of BRD9; dDUF, DUF-domain deletion mutant of BRD9. **h**, Immunoprecipitation with Flag following by probing (immunoblot, IB) for GLTSCR1 or GLTSCR1L in 293T cells expressing 3×Flag-tagged versions of the indicated deletion mutants. Deletion mutants illustrated in **g**. Representative images from $n = 3$ biologically independent experiments.

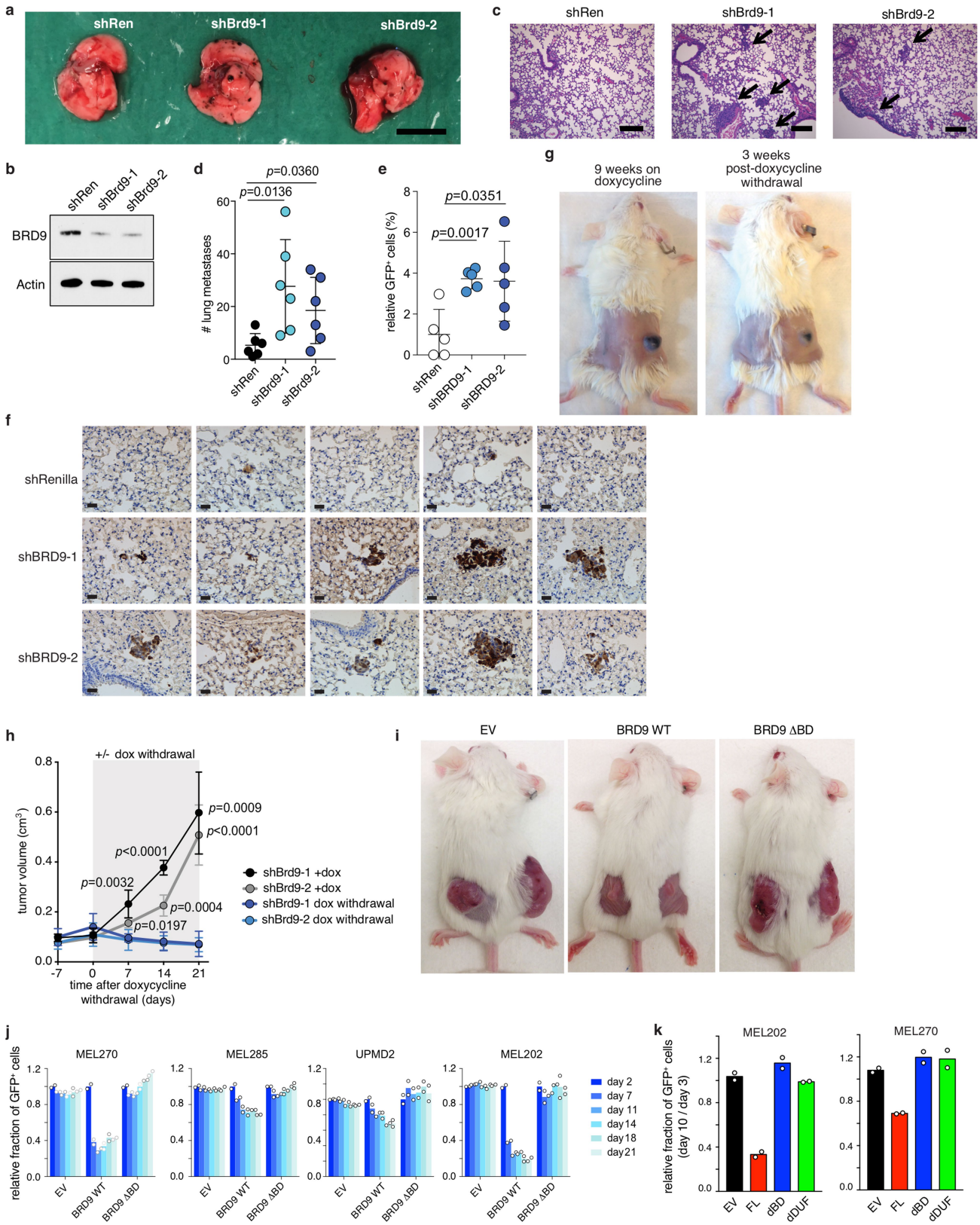


Extended Data Fig. 4 | BRD9 loss drives relocalization of GLTSCR1 away from CTCF-associated loci. a, As in Fig. 3e, but illustrating relative positions with respect to transcription start sites (TSSs). **b**, As in Fig. 3f, but for motifs at BRG1-bound loci. $n = 401$ transcription factors analysed. **c**, UpSet plots depicting the overlap of consensus GLTSCR1-bound loci in MEL270 cells with the indicated treatments. **d**, Volcano plot illustrating the difference in the mean motif scores at BRD9-sensitive versus constitutive GLTSCR1-bound loci for the transcription factors in Fig. 3f, as well as associated statistical significance. $n = 401$ transcription factors analysed. *P* values computed with a two-sided Mann–Whitney *U*-test. **e**, As in **c**, but for BRG1-bound loci. **f**, As in **d**, but for BRG1-bound loci. $n = 401$ transcription factors analysed. **g**, Selected enriched annotation terms from a Genomic Regions Enrichment of Annotations Tool (GREAT) analysis[61] of genes near BRD9-sensitive and constitutive GLTSCR1-bound loci. Plot illustrates $-\log_{10}$(FDR), computed with a one-sided binomial test and corrected for multiple testing using the Benjamini–Hochberg procedure. O and E, numbers of genes that were observed and expected, respectively. **h**, Differences in gene expression in *SF3B1*-mutated versus wild-type samples in the TCGA UVM cohort for genes with GLTSCR1-bound promoters identified in MEL270 cells. Colours indicate the responsiveness of peaks to BRD9 loss. **i**, Read coverage from GLTSCR1 ChIP–seq (MEL270 cells) and RNA-seq (TCGA UVM cohort) around *NFATC2IP*. Red trapezoid indicates GLTSCR1 binding in the promoter, with reduced binding upon treatment with BRD9 degrader or expression of $SF3B1^{K700E}$. *NFATC2IP* was significantly differentially expressed in UVM samples with *SF3B1* mutations relative to wild-type samples. Vertical axis scales were rendered comparable by normalizing ChIP–seq read coverage to mapped library size and RNA-seq read coverage to mapped library size, restricted to coding genes. ChIP–seq experiment performed for $n = 1$ biologically independent replicate. **j**, As in **i**, but for *SETD1A*. *SETD1A* was significantly differentially expressed in UVM samples with *SF3B1* mutations relative to wild-type samples.



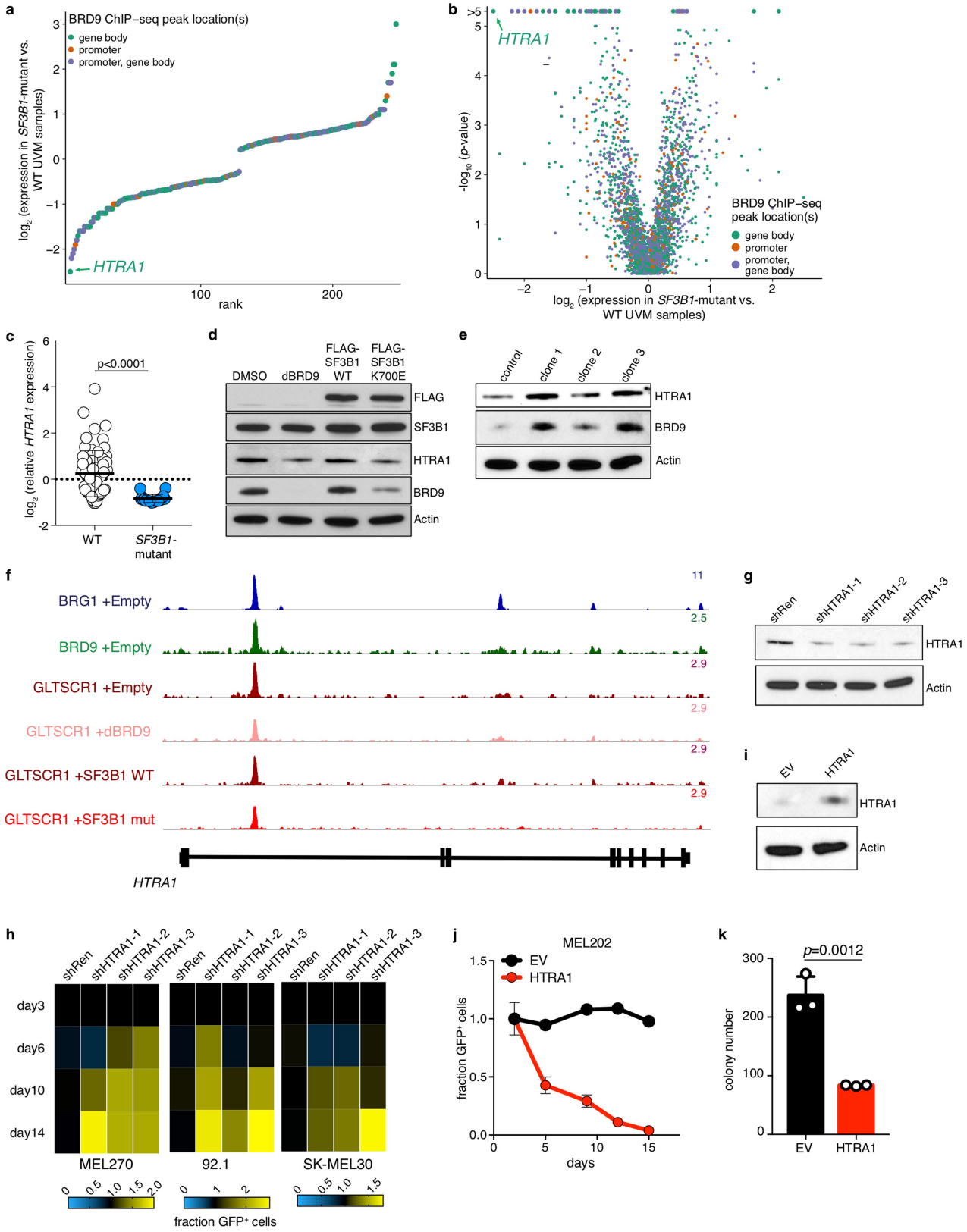
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | BRD9 is a potent tumour suppressor in UVM. a, In vitro growth curves of Melan-a cells treated with two non-targeting shRNAs (against *Renilla* (shRen) and luciferase (shLuc)) versus parental, un-manipulated Melan-a cells. $n = 3$ biologically independent experiments per group. Data are presented as mean ± s.d. **b**, Tumour volume in SCID mice subcutaneously injected with Melan-a cells expressing a control shRNA (against *Renilla*), shRNA against *Brd9* (*Brd9* shRNA no. 1 and no. 2), shRNA against *Brg1* or cDNA encoding CYSLTR2(L129Q) ($n = 8$ mice per group). Data are presented as mean ± s.d. *P* value at day 64 was calculated compared to *Renilla* shRNA group with a two-sided *t*-test. **c**, Representative images of the dissected melanomas from **b**. **d**, H&E images of melanomas from **b**. Scale bars, 100 μm. Representative images from $n = 3$ biologically independent experiments. **e**, qRT–PCR measuring expression of *Brd9* (left) and melanoma-associated genes (*Mitf*, *Dct*, *Pmel* and *Tyrp1*) of melanomas from **a**. $n = 4$ (*Renilla* shRNA) and $n = 5$ (*Brd9* shRNA no. 1 and no. 2) biologically independent experiments. Data are presented as mean ± s.d. *P* value was calculated by two-sided *t*-test. **f**, Images of mice transplanted with parental, un-manipulated Melan-a cells or Melan-a cells transduced with a non-targeting shRNA or *Brd9*-targeting shRNA. Cells were subcutaneously engrafted into SCID mice and tumour volume was estimated 36 days after transplant. **g**, Volumes of tumours from **f** at day 36. Data are presented as mean ± s.d. $n = 10$ tumours per group. *P* value was calculated relative to the parental group by a two-sided *t*-test.



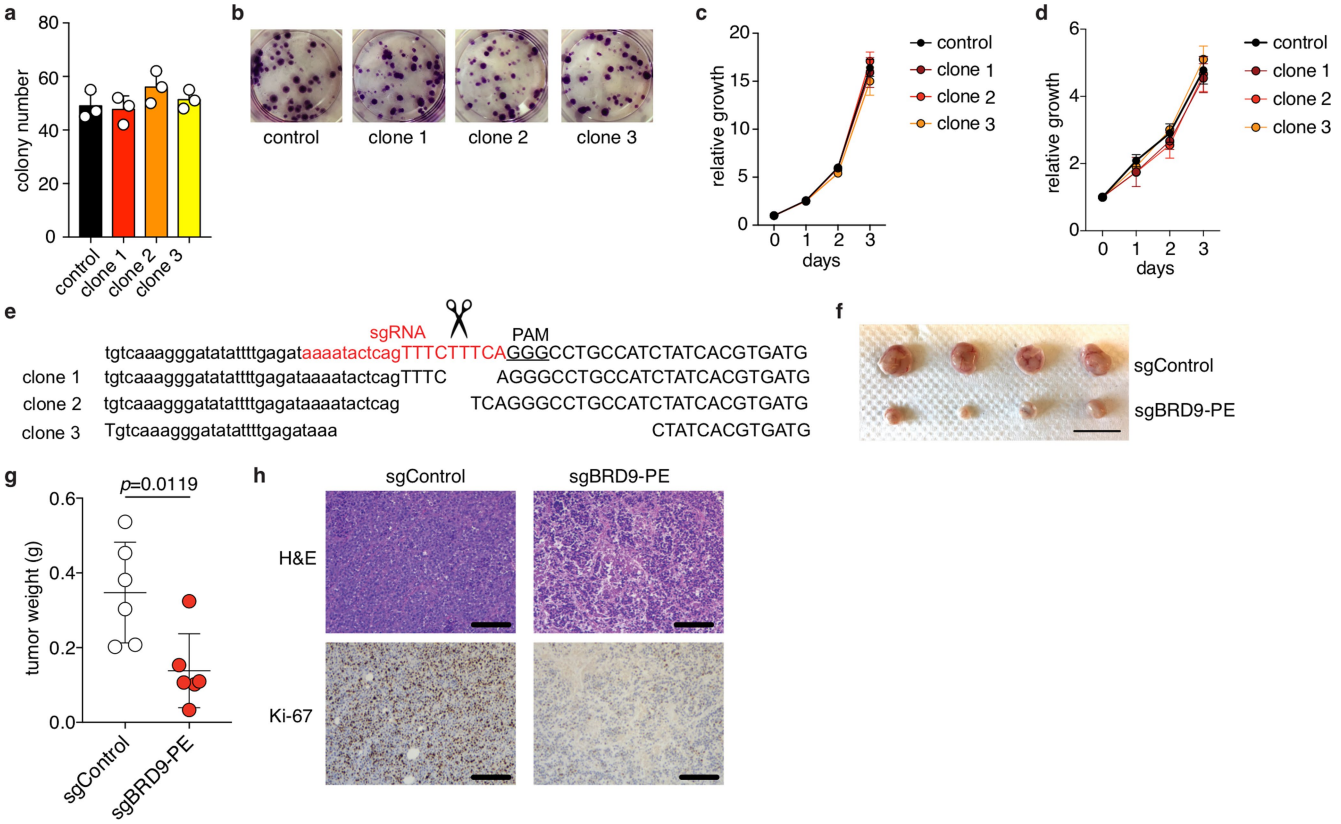
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | BRD9 is a potent tumour suppressor in UVM. **a**, Representative images of pulmonary metastatic foci produced 14 days after intravenous injection of B16 cells with or without *Brd9* shRNA (MLS-E vector). Scale bar, 5 mm. **b**, Western blot of endogenous BRD9 in B16 cells immediately before injection. Actin, loading control. The experiment was repeated three times with similar results. **c**, H&E sections of lung metastases. Arrows indicate metastatic foci. Scale bar, 100 μm. The experiment was repeated three times with similar results. **d**, Numbers of pulmonary B16 metastases identified in the experiments from **a**. $n = 6$ mice per group. *P* value was calculated relative to the *Renilla* shRNA group by a two-sided *t*-test. **e**, Relative percentages of GFP^+ 92.1 cells with or without *BRD9* shRNA (MLS-E vector), assessed by flow cytometric analysis of lung tissue in recipient NSG (NOD–SCID $Il2rg^{-/-}$) mice 14 days after intravenous injection by tail vein. The signal was normalized by dividing by the average percentage of GFP^+ cells in the *Renilla* shRNA (control) group. $n = 5$ biologically independent experiments per group. *P* value was calculated relative to the *Renilla* shRNA group by a two-sided *t*-test. **f**, Anti-GFP immunohistochemistry for sections of lung metastases from the experiment in **e**. Scale bar, 200 μm. The experiment was repeated three times with similar results. **g**, Representative images of tumours derived from transplantation of Melan-a cells transduced with doxycycline-inducible *Brd9* shRNA. Doxycycline was administered for nine weeks (left) and followed by doxycycline withdrawal for three weeks (right). **h**, Tumour volume for the experiment in **g**. $n = 4$ mice per group. The experiment was repeated twice with similar results. *P* value was calculated relative to the parental group by a two-sided *t*-test at day 7, day 14 and day 21. **i**, Representative images of recipient mice engrafted with MEL270 cells transduced with empty vector, full-length BRD9 (BRD9 WT) or a bromodomain-deletion mutant of BRD9 (BRD9 ΔBD) at day 12. $n = 5$ mice per group. **j**, Results of a competition assay to measure the effects of expression of the indicated cDNAs on growth of the indicated melanoma cells. Transduced cells were identified by co-expression of GFP (pMIGII vector). The percentage of GFP^+ cells was tracked over 21 days and normalized to the GFP percentage on day 2. Data are presented as mean ± s.d. $n = 2$ biologically independent experiments per group. **k**, Results of a competition assay to measure the effects of expression of the indicated cDNAs on growth of the indicated melanoma cells. Transduced cells were identified by the co-expression of GFP (pMIGII vector). The percentage of GFP^+ cells was tracked over 10 days and normalized to the GFP percentage on day 3. $n = 2$ biologically independent experiments per group.

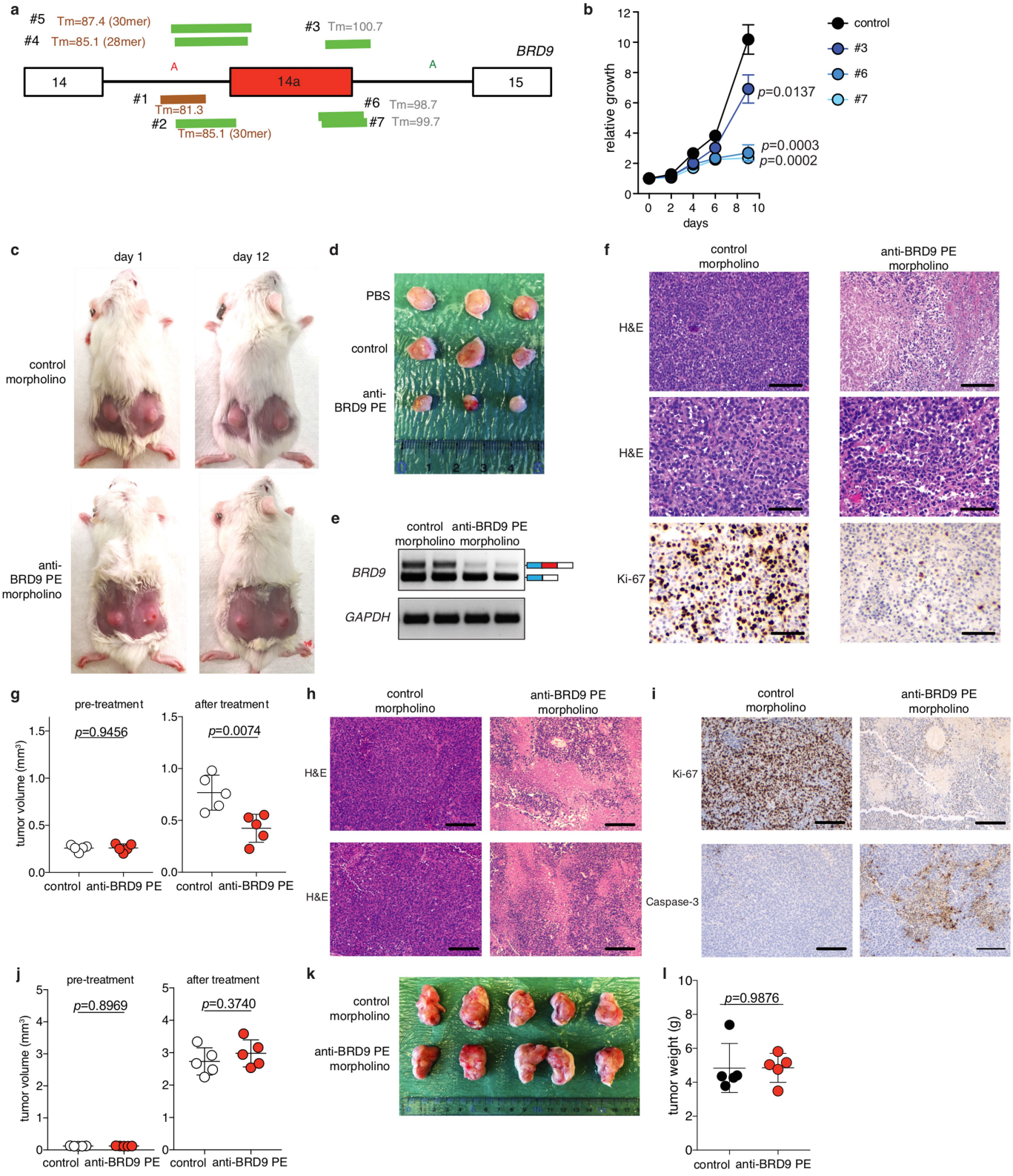


Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | BRD9 regulates *HTRA1* expression to promote UVM tumorigenesis. a, Rank plot illustrating fold change (expressed in $\log_2$) of each significantly differentially expressed gene identified by comparing samples from patients in the TCGA UVM cohort with mutated versus wild-type *SF3B1*. The plot is restricted to genes with BRD9 ChIP–seq peaks within their promoters or gene bodies in the absence of perturbations to BRD9 (MEL270 cells treated with DMSO or following ectopic expression of wild-type SF3B1). $n = 3{,}122$ genes analysed, of which $n = 248$ met the significance ($P < 0.001$) and expression (median expression in both wild-type and mutant samples > 2 transcripts per million) thresholds, and so are illustrated here. *P* value was computed using a two-sided Mann–Whitney *U*-test. **b**, As in **a**, but a volcano plot additionally illustrating the *P* value associated with the comparison between *SF3B1*-mutated and wild-type *SF3B1* samples. $n = 3{,}122$ genes analysed and illustrated. *P* value computed with two-sided Mann–Whitney *U*-test. **c**, *HTRA1* expression in samples from patients in the TCGA UVM cohort with ($n = 18$) or without ($n = 62$) *SF3B1* mutations. Expression is *z*-score normalized across all samples. Data are presented as mean ± s.d. *P* value computed with two-sided *t*-test. **d**, Western blot for Flag, SF3B1, HTRA1, BRD9 and actin in MEL270 cells (wild-type *SF3B1*), treated with DMSO, BRD9 degrader, Flag–SF3B1(WT) or Flag–SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **e**, Western blot for HTRA1, BRD9 and actin in MEL202 cells ($SF3B1^{R625G}$) following CRISPR–Cas9-mediated mutagenesis of the *BRD9* poison exon (as shown in Extended Data Fig. 2o). Representative images from $n = 3$ technically independent experiments. **f**, Read coverage for BRG1, BRD9 and GLTSCR1 ChIP–seq at the *HTRA1* locus in MEL270 cells (shown in **d**) treated with an empty vector, BRD9 degrader, or $SF3B1^{WT}$ or $SF3B1^{K700E}$ cDNAs ($n = 1$ ChIP–seq experiment performed for each condition). **g**, Western blot for HTRA1 and actin in MEL270 cells treated with shRNAs against *HTRA1* or with a non-targeting control shRNA (against *Renilla*). Representative images from $n = 3$ biologically independent experiments. **h**, Heat map summarizing the results of a competition assay to measure the effect of each indicated shRNA on the growth of Cas9-expressing UVM cell lines with wild-type *SF3B1*. Cell growth was computed with respect to cells treated with a non-targeting control shRNA (against *Renilla*) and the percentage of GFP^+ cells at day 14 was normalized to that at day 2. The illustrated values correspond to the mean values computed over $n = 3$ biologically independent experiments. **i**, Western blot for HTRA1 and actin in MEL202 cells ($SF3B1^{R625G}$) following stable overexpression of an empty vector or HTRA1 (both with an MSCV-IRES-GFP vector). Representative images from $n = 3$ biologically independent experiments. **j**, Ratio of GFP^+ to GFP^- MEL202 cells ($SF3B1^{R625G}$) from a competition experiment in which GFP^+ cells from **i** were seeded at an initial 1:1 ratio with GFP^- control cells. Data are presented as mean of $n = 2$ biologically independent experiments. **k**, Colony number of MEL202 cells expressing empty vector or *HTRA1* cDNA from **i**, following 10 days of growth in soft agar. Data are presented as mean of $n = 3$ biologically independent experiments.

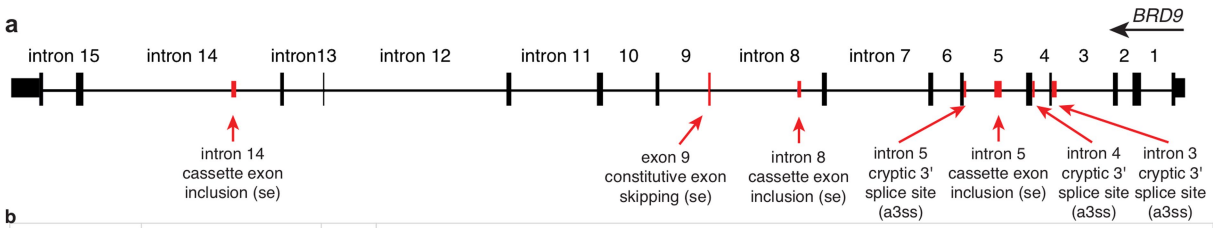


Extended Data Fig. 8 | CRISPR–Cas9-mediated mutagenesis of the *BRD9* poison exon corrects *BRD9* aberrant splicing and abrogates growth of *SF3B1*-mutated melanoma. a, Colony number for MEL270 cells (wild-type *SF3B1*) without (control) or with (clone 1, clone 2 and clone 3) CRISPR–Cas9-induced indels that disrupted *BRD9* poison exon recognition. Data presented as mean ± s.d. $n = 3$ biologically independent experiments per group. **b**, Representative images from **a**. **c**, Proliferation of the clones described in **a**. $n = 3$ biologically independent experiments per group. **d**, Proliferation of MEL285 cells (wild-type *SF3B1*) without (control) or with (clones 1, 2 and 3) CRISPR–Cas9-induced indels that disrupted *BRD9* poison exon recognition. $n = 3$ biologically independent experiments per group. **e**, Mutations generated at the 5′ end of the *BRD9* poison exon by CRISPR–Cas9-mediated indels in clones 1, 2 and 3 of MEL285 cells from **d**. The PAM sequence is illustrated with uppercase, underlined nucleotides. Red nucleotides hybridize to the sgRNA. **f**, Representative images of dissected tumours from recipient mice transplanted with CRISPR–Cas9-modified MEL202 clones. **g**, Tumour weight for the tumours illustrated in **f**. Data presented as mean ± s.d. $n = 6$ biologically independent experiments per group. *P* value was calculated relative to the control shRNA group by a two-sided *t*-test. **h**, H&E staining, as well as Ki-67 immunohistochemistry images, for the tumours illustrated in **f**. Representative images from $n = 3$ independent histological analyses.



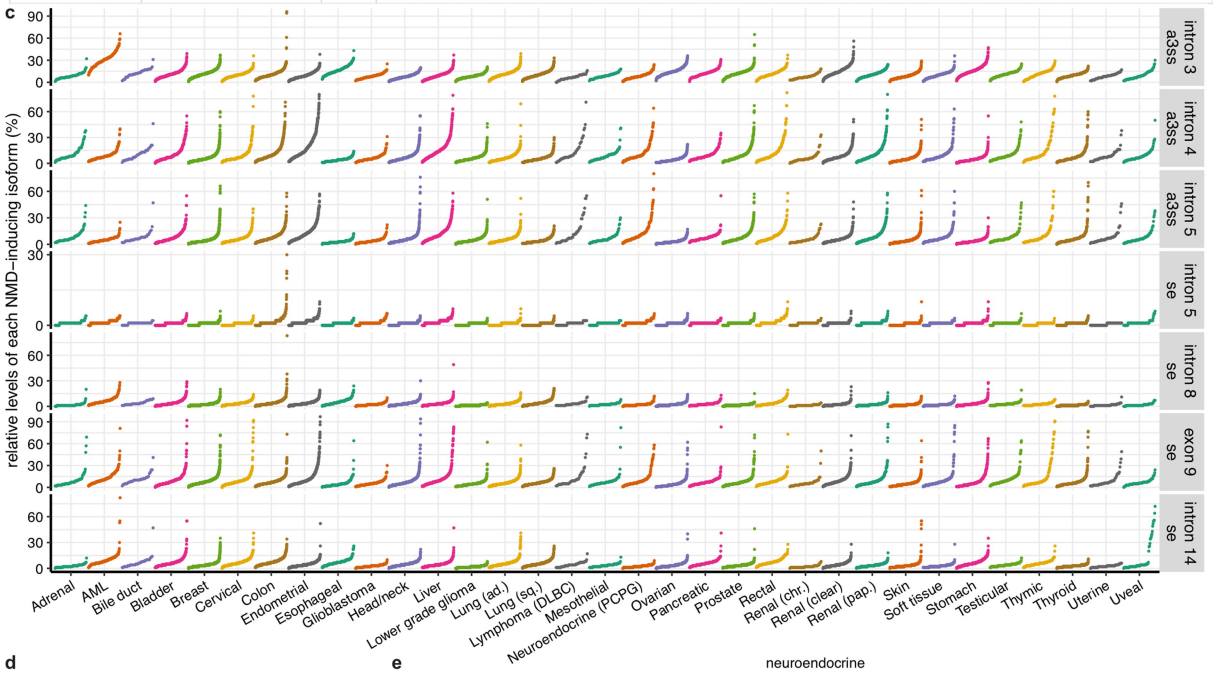
Extended Data Fig. 9 | See next page for caption.

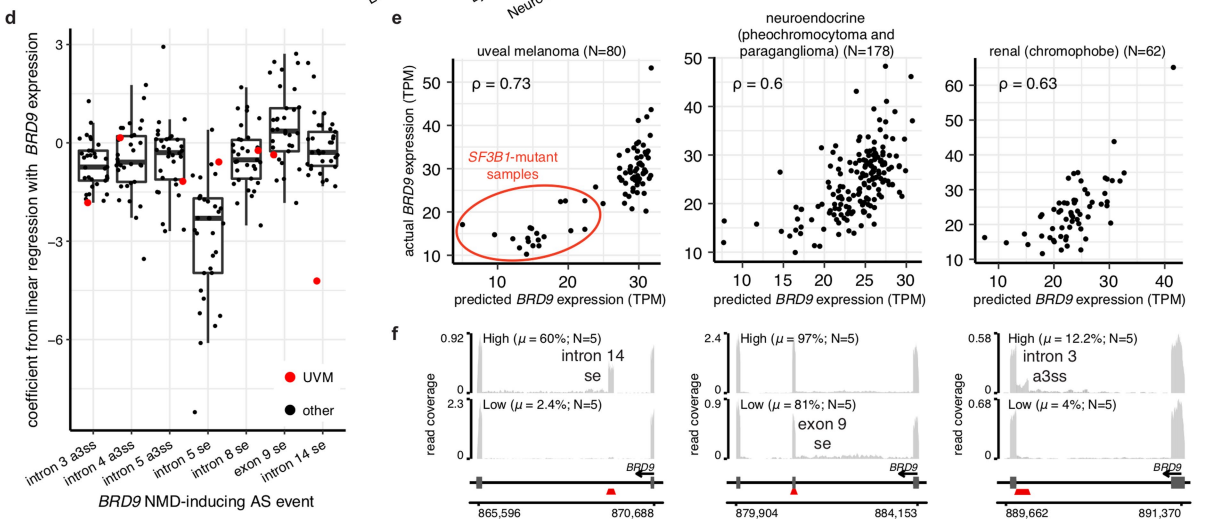
Extended Data Fig. 9 | Correcting *BRD9* mis-splicing in *SF3B1*-mutated xenografts with ASOs suppresses tumour growth. a, Cartoon representation of the *BRD9* loci targeted by each designed morpholino. Melting temperature (T_m) is shown. Lengths of target sequences are indicated in parentheses; if these are not indicated, then the length is 25 nt. **b**, Growth of MEL202 cells (*SF3B1*^{R625G}) treated with 10 μM of control non-targeting (control) or *BRD9* poison-exon-targeting morpholinos (no. 3, no. 6 and no. 7). $n = 3$ biologically independent experiments per group. *P* values at day 9 were calculated relative to the control group by a two-sided *t*-test. **c**, Representative images of recipient mice xenografted with MEL202 cells and treated with PBS or morpholinos in vivo. Each tumour was analysed after in vivo treatment with PBS, control morpholino or morpholino against the *BRD9* poison exon (no. 6) (12.5 mg kg^{-1}, every other day to a total of 8 intratumoral injections). $n = 10$ tumours per group. **d**, Representative images of dissected tumours from the experiment described in **c**. **e**, RT–PCR results of tumours from **c** to evaluate *BRD9* splicing. The experiment was repeated three times with similar results. **f**, Representative H&E and Ki-67 staining images of tumours from **c**. Scale bars, 100 μm (top), 50 μm (middle and bottom). The experiment was repeated three times with similar results. **g**, Estimated tumour volume for recipient mice transplanted with a patient-derived xenograft model of *SF3B1*^{R625C} rectal melanoma and treated with in vivo morpholinos (control or morpholino against *BRD9* poison exon, no. 6) (12.5 mg kg^{-1}, every other day to a total of 8 intratumoral injections). $n = 5$ mice per group. Estimated tumour volumes before and after treatment are shown. Data are presented as mean ± s.d. *P* values were calculated relative to the control group by a two-sided *t*-test. **h**, Representative H&E staining images of tumours from **g**. The experiment was repeated three times with similar results. **i**, Representative Ki-67 staining images of tumours from **g**. The experiment was repeated three times with similar results. **j**, Estimated tumour volume for recipient mice transplanted with a patient-derived xenograft model of UVM (wild-type *SF3B1*), treated with in vivo morpholinos (control or morpholino against *BRD9* poison exon, no. 6) (12.5 mg kg^{-1}, every other day to a total of 8 intratumoural injections). $n = 5$ mice per group. Estimated tumour volumes before and after treatment are shown. Data are presented as mean ± s.d. *P* values weres calculated relative to the control group by a two-sided *t*-test. **k**, Representative images of dissected tumours from **j**. **l**, Tumour weight for tumours from **k**. $n = 5$ mice per group. *P* value was calculated relative to the control group by a two-sided *t*-test.



b

intron 3 a3ss	chr5:889763-889881	prox.	TTCTCGGTGGAACTTAAAATGCTGTGAGACACCAGACAGACAGATACTGTGAACTTGGAGCTCTCTAATGAAGGGATACCAAAGTCTTGTATTCAATTTTTTTTTCCTTAAATTGTCAG
intron 4 a3ss	chr5:889276-889280	dist.	AAAAG
intron 5 se	chr5:888350-888522	inc.	GGTTTGCTGACTCACAGAAACACAACAGAACAGACTTAAAGGAGAGTGGGAGAAAAGGAAACTCGATGAGAGGATCAGGCAGCGCCTGAGGGTAGATGTTAAGAGCACTCCATGCCCGAAAGCCCTCTGTATTCCCTGCTTCGGGCCCCAGTCTGGCCATTCCTAGCAACAAA
intron 5 a3ss	chr5:887583-887586	dist.	GCAG
intron 8 se	chr5:883466-883565	inc.	GAGACGTTCTAGACCCAGTGTCTTCAGTGGAAACAGAGCACAGGCTGTCTGCAGTCTCAGGCTTCCTAGCAGCCCTGATAAGAGAAGTAGAGCAATGCAG
exon 9 se	chr5:881222-881297	exc.	ATGGGCTATCTGAAGAGGAACGGGGACGGGAGCCTGCTCTACAGCGTGGTCAACACGGCCGAGCCGGACGCTGATG
intron 14 se	chr5:869360-869519	inc.	TTTCTTTCAGGGCCTGCCATCTATCACGTGATGCTGTACTAAAGTCGGATCGGAATTTGGTGTCTTATTACCACGGAGTCTGCTTTGTCAACCTTCAGATTTCTGTTTTAAAGTTAATGCTGGTCAGCTGTGCCTGAATTCCAAAAGAGGGGATAATGAG





Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Use of multiple, distinct NMD isoforms of *BRD9*. **a**, *BRD9* gene structure illustrating constitutive *BRD9* exons and alternative splicing events that are predicted to induce NMD. *SF3B1* mutations promote inclusion of the BRD9 poison exon in intron 14. **b**, Genomic coordinates (hg19/GRCh37 assembly) of each NMD-inducing event illustrated in **a**, as well as genomic sequence of each alternatively spliced region highlighted in red in **a**. The third column indicates the specific isoform that is a predicted NMD substrate. Prox, intron-proximal competing 3′ splice site; dist, intron-distal competing 3′ splice site; inc, exon inclusion; exc, exon exclusion. **c**, Rank plot illustrating levels of each NMD-inducing isoform relative to total *BRD9* mRNA levels for each sample in each indicated TCGA cohort. Boxes illustrate first and third quartiles, with whiskers extending to 1.5× interquartile range. **d**, Box plot illustrating the distribution of coefficients estimated by fitting a linear model to predict *BRD9* gene expression on the basis of relative levels of each NMD-inducing isoform. The relative levels of NMD-inducing isoforms illustrated in **c**, as well as *BRD9* gene expression estimates for each sample, were used to construct an independent linear model with robust regression for each TCGA cohort. The coefficients resulting from this model fitting procedure are illustrated in the box plot, in which each dot corresponds to the coefficient associated with the corresponding NMD-inducing event for a single TCGA cohort. Coefficients for the TCGA UVM cohort are highlighted in red. The coefficients are typically negative (as expected for NMD-inducing isoforms), with the exception of constitutive exon 9 skipping, for which the coefficients are generally positive—as expected for an event in which NMD is induced when a constitutive exon is excluded. The *SF3B1*-mutation-responsive poison exon in intron 14 dominates the fit for UVM, as expected. $n = 33$ TCGA cohorts analysed and illustrated. **e**, Scatter plots comparing actual (y axis) and predicted (x axis) *BRD9* expression levels for three TCGA cohorts. Each dot corresponds to a single sample. ρ, Spearman's correlation between actual and predicted values. **f**, RNA-seq read coverage plots for patient samples from the TCGA cohorts illustrated in **e** for representative alternative splicing events illustrated in **a**. Each coverage plot illustrates data averaged over the $n = 5$ patient samples from the vertically matched cohort in **e** that exhibit the lowest or highest relative expression of the NMD-inducing isoform. μ, mean relative expression of the illustrated NMD-inducing isoform, computed over each group of samples.

nature research

Corresponding author(s): Omar Abdel-Wahab
Robert K. Bradley

Last updated by author(s): Jul 6, 2019

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.

Statistics

For all statistical analyses, confirm that the following items are present in the 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
☐	☒ A statement on 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 statistical parameters 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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Publicly available software was used in this study. Specific programs are RSEM (v1.2.4), Bowtie (v1.0.0), TopHat (v2.0.8b), MISO (v2.0), and Bioconductor (v3.7) within the R (v3.5.1) programming environment.

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. 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

RNA-seq and ChIP-seq data generated as part of this study were deposited in the Gene Expression Omnibus (accession number GSE124720). RNA-seq data from published studies were downloaded from CGHub (TCGA UVM57), EMBL-EBI ArrayExpress (Illumina Human BodyMap 2.0: E-MTAB-513), the Gene Expression Omnibus (accession numbers GSE72790, GSE114922) (CLL: Darman et al. 201515, MDS: Pellagati et al. 201858), or directly obtained from the authors (UVM: Alsafadi et al. 201610). Gel source data can be found in Supplementary Fig. 1. Other data that support this study's findings are available from the authors upon reasonable request.

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	Sample sizes for xenograft experiments were chosen based on published studies of known oncogenic drivers of relevant models (e.g., expression of the oncoprotein CYSLTR2 L129Q in Melan-a cells).
Data exclusions	No data were excluded.
Replication	Attempts at replication were successful.
Randomization	Animals were randomly assigned to experimental groups.
Blinding	The data presented did not require the use of blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For Western blotting, the following antibodies to the following proteins were used: BRD9 (Bethyl Laboratories; A303-781A and Active Motif; 61538), SF3B1/Sap-155 (MBL; D221-3), Flag-M2 (Sigma-Aldrich; F-1084), β-actin (Sigma-Aldrich; A-5441), GLTSCR1 (Santa Cruz Biotechnology; sc-515086), GLTSCR1L (Thermo Fisher Scientific; PA5-56126), BRM (Bethyl Laboratories; A303-015A), BRG1 (Santa Cruz Biotechnology; sc-17796), BAF155 (Santa Cruz Biotechnology; sc-48350), BAF60A (Santa Cruz Biotechnology; sc-135843), BAF47 (Santa Cruz Biotechnology; sc-166165), ARID1A (Santa Cruz Biotechnology; sc-373784), ARID2 (Santa Cruz Biotechnology; sc-166117), BRD7 (Thermo Fisher Scientific; PA5-49379), U2AF2 (Bethyl Laboratories; A303-665A), U2AF1 (Bethyl Laboratories; A302-080A), Histone H3 (Abcam; ab1791), HTRA1 (R&D systems; MAB2916-SP). For ChIP-seq studies in MEL270 cells, antibodies to endogenous BRG1 (Abcam EPNCIR111A, Lot # GR3208604-8), GLTSCR1 (Santa Cruz SC-240516, Lot # A2313), and BRD9 (Abcam, ab137245) were used.
Validation	All antibodies were validated by the supplier for human samples, and were checked in the lab by Western blotting on cell lysate and by comparing to the manufacturer's or in-house results.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Ba/F3, K562, NALM6, TF1, HEK293T, MCF10A, Miapaca2, CFPAC1, B16, and HEK293T cells were obtained from the American Type Culture Collection (ATCC) . The isogenic K562 and NALM6 cell lines with or without SF3B1 mutations were generated at Horizon Discovery. Melan-a cells were provided by Dr. Dorothy Bennett (PMID: 3102392); MEL202, MEL270, UPMD1, UPMD2, and 92-1 were obtained from Dr. Boris Bastian (PMID: 22236444). KPC cells were obtained from Dr. Ben Stanger (PMID: 21436454). SK-MEL30 cells and RN2 cells were provided by Dr. David B. Solit (PMID: 21725359) and Dr. Scott W. Lowe (PMID: 21131983), respectively.

Authentication	An aliquot of each cell lines were authenticated using ATCC fingerprinting. Otherwise, the cells were submitted for short tandem repeat (STR) profiling and MSK-IMPACT (integration mutation profiling of actionable cancer targets) for mutational status at MSKCC to confirm their authenticity.
Mycoplasma contamination	All cell lines are frequently tested for mycoplasma contamination. Cell lines used in this study were verified to be mycoplasma negative before undertaking any experiments with them.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	6-8 week SCID and NGS mice were purchased from The Jackson Laboratory (stock #001303 and stock #005557, respectively).
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All animals were housed at Memorial Sloan Kettering Cancer Center (MSKCC). All animal procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC. All mouse experiments were performed in accordance with a protocol approved by the MSKCC Institutional Animal Care and Use Committee (11-12-029).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Patients with chronic lymphocytic leukemia (CLL) and uveal and mucosal melanoma seen at Memorial Sloan Kettering Cancer Center who provided anonymized peripheral blood mononuclear cells (for CLL) and tissue biopsies (for melanoma).
Recruitment	Patients with CLL and uveal and mucosal melanoma seen at Memorial Sloan Kettering Cancer Center (MSK) who consented to MSK IRB protocol 06-107 (for CLL) and IRB 14-191 (for melanoma samples) were eligible for inclusion regardless of race, gender, ethnicity or other characteristics.
Ethics oversight	Studies were approved by the Institutional Review Boards of Memorial Sloan Kettering Cancer Center (MSK), informed consent was obtained from all subjects (under MSK IRB protocol 06-107) and conducted in accordance to the Declaration of Helsinki protocol. Patients provided samples after their informed consent and primary human de-identified CLL samples derived from whole peripheral blood or BM mononuclear cells were utilized. PDX models were performed using tumor biopsies from de-identified patients under MSK IRB protocol 14-191. Genomic alterations in melanoma tumor biopsies and CLL cells were analyzed using MSK IMPACT28 assay or FoundationOne Heme30 assay, both as previously described. Patient samples were anonymized by the Hematologic Oncology Tissue Bank of MSK (for CLL samples) and the MSK Antitumor Assessment Core Facility (for PDX samples).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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.*	The data (including related RNA-seq data) has been deposited in the Gene Expression Omnibus under accession number GSE124720 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124720)
Files in database submission	The above accession includes related RNA-seq data. The relevant ChIP-seq files (.fastq, .bigwig and associated peak calls) can be obtain at the link above: GSM3544057 MEL270.dBRD9.BRG1 GSM3544058 MEL270.dBRD9.GLTSCR1 GSM3544059 MEL270.dBRD9.input GSM3544061 MEL270.DMSO.BRD9 GSM3544062 MEL270.DMSO.BRG1 GSM3544063 MEL270.DMSO.GLTSCR1 GSM3544064 MEL270.DMSO.input GSM3544066 MEL270.SF3B1_K700E.BRG1 GSM3544067 MEL270.SF3B1_K700E.GLTSCR1 GSM3544068 MEL270.SF3B1_K700E.input

GSM3544070 MEL270.SF3B1_WT.BRD9
GSM3544071 MEL270.SF3B1_WT.BRG1
GSM3544072 MEL270.SF3B1_WT.GLTSCR1
GSM3544073 MEL270.SF3B1_WT.input

Genome browser session
(e.g. UCSC)

https://genome.ucsc.edu/s/chewguoliang/SF3B1%20BRD9%20MEL270%20ChIP

Methodology

Replicates

The ncBAF complex was pulled-down with two distinct antibodies against BRD9 and GLTSCR1 in two separate control and ncBAF perturbation experiments that used orthogonal methods to perturb ncBAF formation. Analyses used the intersection of data from these distinct experiments.

Sequencing depth

Sequencing depth, as well as mapping rates are described below:

Sample Reads_total Reads_mapped Reads_unmapped Reads_multimapped
MEL270.dBRD9.GLTSCR1 21663452 17327309 (79.98%) 3317111 (15.31%) 1019032 (4.70%)
MEL270.dBRD9.input 18584217 16915436 (91.02%) 761948 (4.10%) 906833 (4.88%)
MEL270.dBRD9.BRG1 20392804 18016161 (88.35%) 1654723 (8.11%) 721920 (3.54%)
MEL270.DMSO.BRD9 19785290 15734102 (79.52%) 2934586 (14.83%) 1116602 (5.64%)
MEL270.DMSO.GLTSCR1 19281052 15911493 (82.52%) 2501078 (12.97%) 868481 (4.50%)
MEL270.DMSO.input 19743934 17876186 (90.54%) 899111 (4.55%) 968637 (4.91%)
MEL270.DMSO.BRG1 21373496 18958611 (88.70%) 1645389 (7.70%) 769496 (3.60%)
MEL270.SF3B1_K700E.GLTSCR1 20098772 15780486 (78.51%) 3347543 (16.66%) 970743 (4.83%)
MEL270.SF3B1_K700E.input 18866831 17097903 (90.62%) 832672 (4.41%) 936256 (4.96%)
MEL270.SF3B1_K700E.BRG1 19879373 17670132 (88.89%) 1491006 (7.50%) 718235 (3.61%)
MEL270.SF3B1_WT.BRD9 15700266 12196453 (77.68%) 2546295 (16.22%) 957518 (6.10%)
MEL270.SF3B1_WT.GLTSCR1 19755923 16270285 (82.36%) 2565754 (12.99%) 919884 (4.66%)
MEL270.SF3B1_WT.input 19401125 17523685 (90.32%) 895478 (4.62%) 981962 (5.06%)
MEL270.SF3B1_WT.BRG1 19408244 17112467 (88.17%) 1589600 (8.19%) 706177 (3.64%)

Antibodies

BRG1 (Abcam EPNCIR111A, Lot # GR3208604-8)
GLTSCR1 (Santa Cruz SC-240516, Lot # A2313)
BRD9 (Abcam, ab137245)

Peak calling parameters

Narrow peaks were called using the callpeak function from MACS2 v2.1.1.20160309 against matched input samples, using default parameters and a p-value cutoff of 1e-5.

Data quality

For all samples, a p-value cutoff of 1e-5 against input was used. All peaks were called at a q-value of < 0.017. For each sample, the number of peaks with a fold-change > 5, and the total number of peaks called is as follows:
MEL270.dBRD9.BRG1 68229 94156
MEL270.dBRD9.GLTSCR1 2982 4504
MEL270.DMSO.BRD9 13390 20591
MEL270.DMSO.BRG1 73624 101272
MEL270.DMSO.GLTSCR1 4453 6564
MEL270.SF3B1_K700E.BRG1 60362 85463
MEL270.SF3B1_K700E.GLTSCR1 6939 11118
MEL270.SF3B1_WT.BRD9 17434 25208
MEL270.SF3B1_WT.BRG1 64209 89486
MEL270.SF3B1_WT.GLTSCR1 5794 8678

Software

ChIP-seq reads were mapped to the genome by calling Bowtie v1.0.048 with the arguments '-v 2 -k 1 -m 1 --best --strata'. Peaks were called using MACS2 v2.1.1.2016030952 against input control libraries with p < 10-5 and subsequently filtered to remove peaks contained within ENCODE blacklisted regions and the mitochondrial genome. Subsequent data analysis was performed with Bioconductor in the R programming environment. Consensus peaks between samples were called using the soGGI package v1.14.0. Peaks were annotated using the ChIPseeker package v1.18.0.