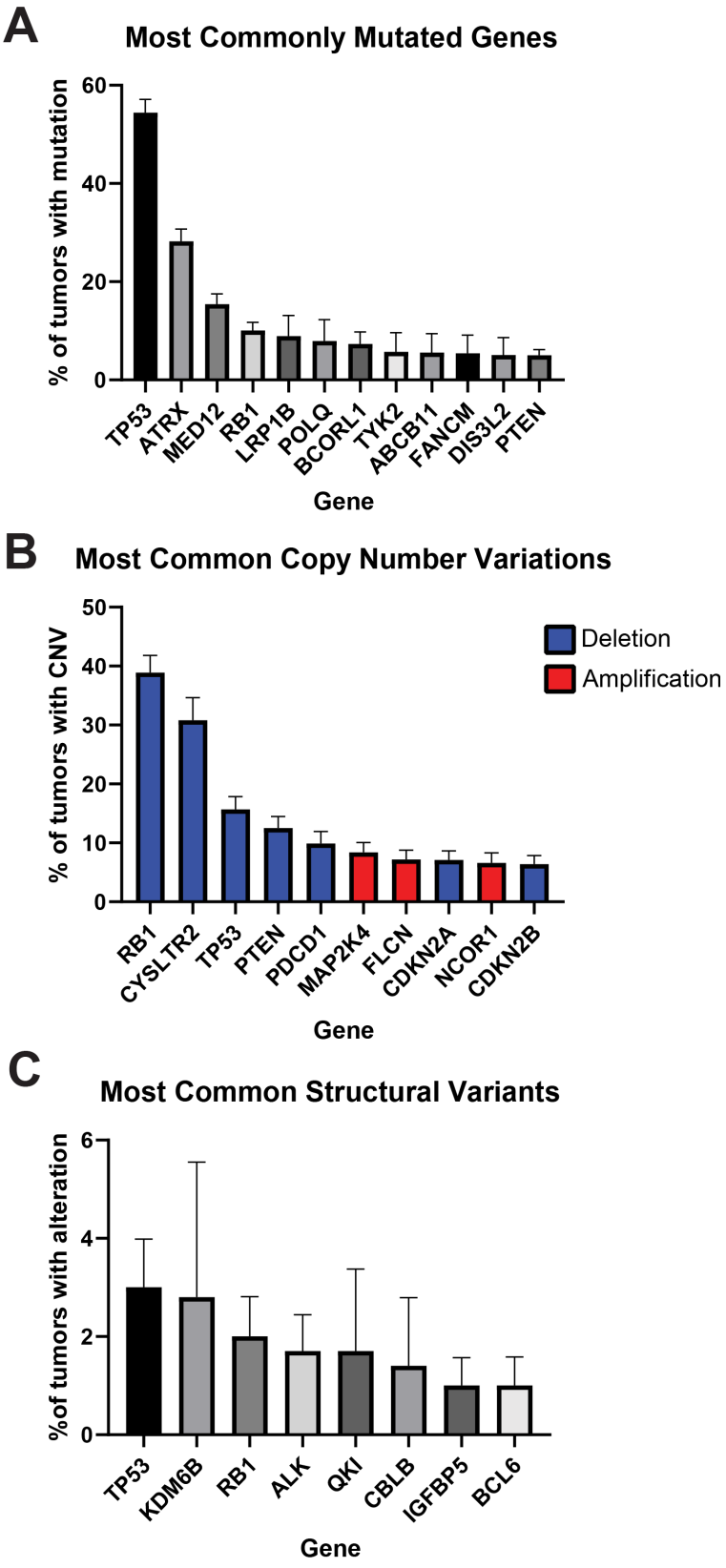
Supplemental Figures

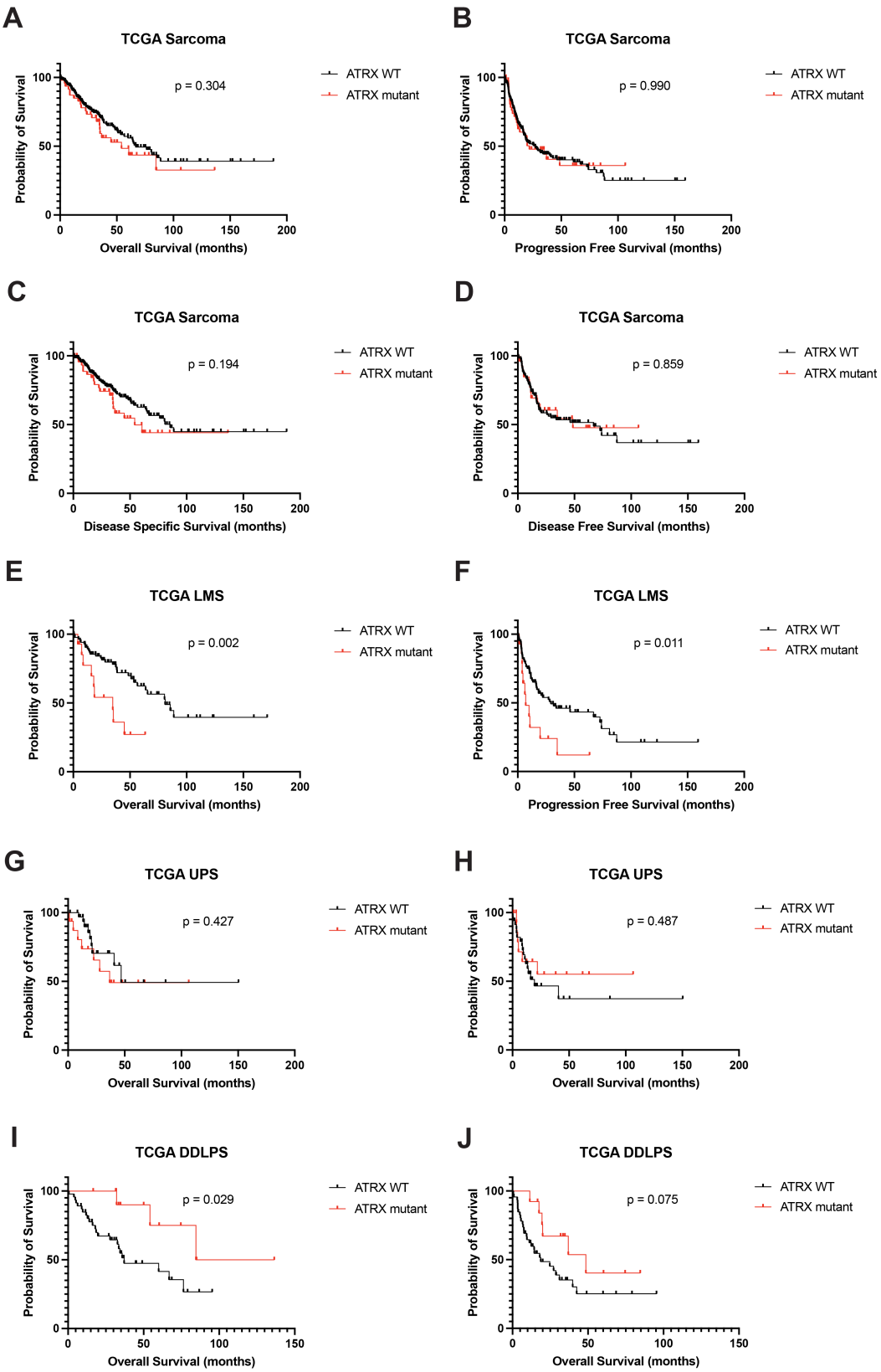
Supplemental Figure S1. Most commonly altered genes in sarcoma in GENIE database.

Bar graphs show the most common genes with mutations (A), copy number variations (B), and structural variants (C) in sarcoma samples in the GENIE database. Percent of profiled samples are shown. A gene must have been profiled in at least 10% of samples within the group (e.g. ST URS) to be included on the graph. Only those genes that were altered in at least > 5% of profiled sarcomas within the group are shown. In (B), deletions are denoted with blue bars, while amplifications are denoted with red bars.


A
Most Commonly Mutated Genes
% of tumors with mutation
60
40
20
0
TP53
ATRX
MED12
RB1
LRP1B
POLQ
BCORL1
TYK2
ABCB11
FANCM
DIS3L2
PTEN
Gene
B
Most Common Copy Number Variations
% of tumors with CNV
50
40
30
20
10
0
Deletion
Amplification
RB1
CYSLTR2
TP53
PTEN
PDCD1
MAP2K4
FLCN
CDKN2A
NCOR1
CDKN2B
Gene
C
Most Common Structural Variants
%of tumors with alteration
6
4
2
0
TP53
KDM6B
RB1
ALK
QKI
CBLB
IGFBP5
BCL6
Gene

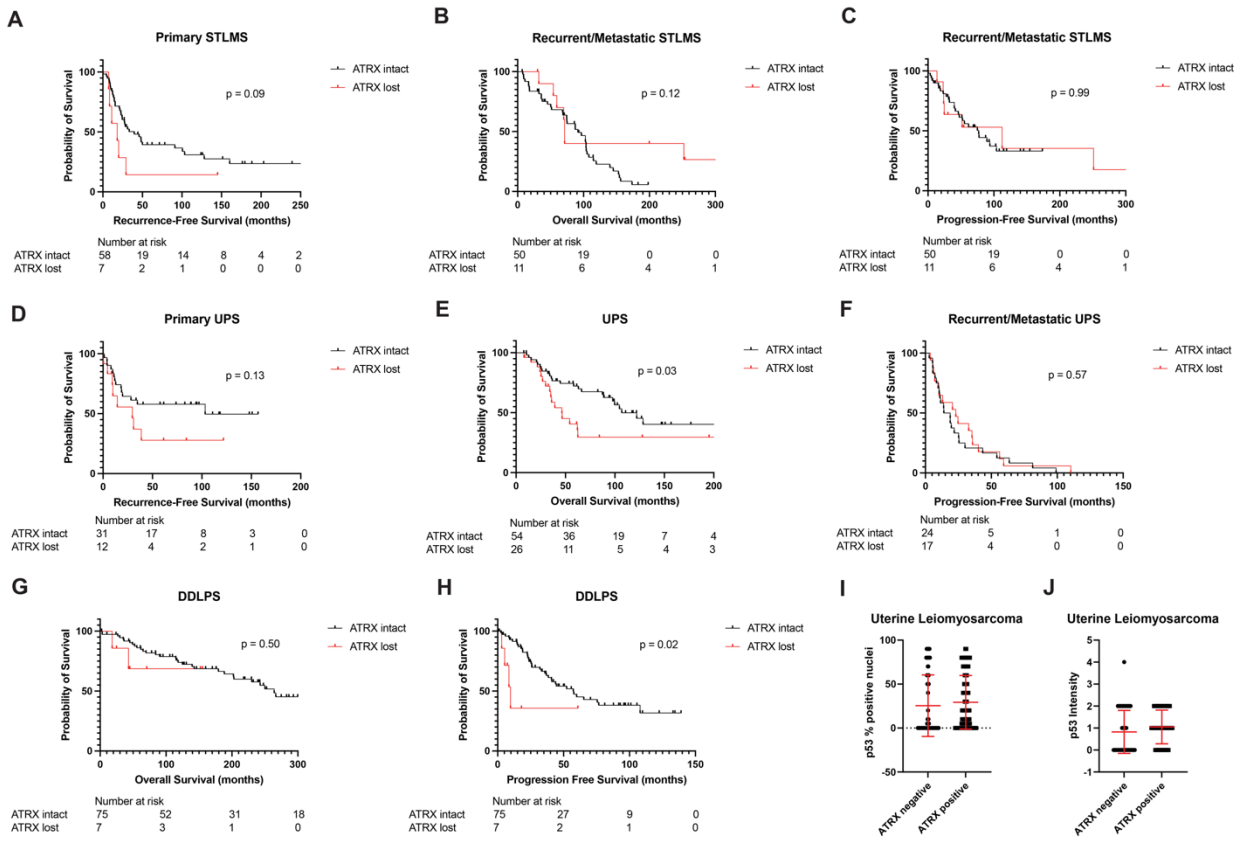
Supplemental Figure S2. Association of ATRX alteration with survival in TCGA sarcoma dataset.

(A-D) Overall survival (A), progression-free survival (B), disease-specific survival (C), and disease-free survival (D) in the entire sarcoma TCGA cohort based on *ATRX* alteration (red lines for mutant, black for wild type). (E-F) Overall survival (E) and progression-free survival (F) for just the patients with LMS in the TCGA sarcoma study. (G-H) Overall survival (G) and progression-free survival (H) for just the patients with UPS in the TCGA sarcoma study. (I-J) Overall survival (I) and progression-free survival (J) for just the patients with DDLPS in the TCGA sarcoma study.


A
TCGA Sarcoma
p = 0.304
ATRX WT
ATRX mutant
Probability of Survival
Overall Survival (months)
B
TCGA Sarcoma
p = 0.990
ATRX WT
ATRX mutant
Probability of Survival
Progression Free Survival (months)
C
TCGA Sarcoma
p = 0.194
ATRX WT
ATRX mutant
Probability of Survival
Disease Specific Survival (months)
D
TCGA Sarcoma
p = 0.859
ATRX WT
ATRX mutant
Probability of Survival
Disease Free Survival (months)
E
TCGA LMS
p = 0.002
ATRX WT
ATRX mutant
Probability of Survival
Overall Survival (months)
F
TCGA LMS
p = 0.011
ATRX WT
ATRX mutant
Probability of Survival
Progression Free Survival (months)
G
TCGA UPS
p = 0.427
ATRX WT
ATRX mutant
Probability of Survival
Overall Survival (months)
H
TCGA UPS
p = 0.487
ATRX WT
ATRX mutant
Probability of Survival
Overall Survival (months)
I
TCGA DDLPS
p = 0.029
ATRX WT
ATRX mutant
Probability of Survival
Overall Survival (months)
J
TCGA DDLPS
p = 0.075
ATRX WT
ATRX mutant
Probability of Survival
Overall Survival (months)

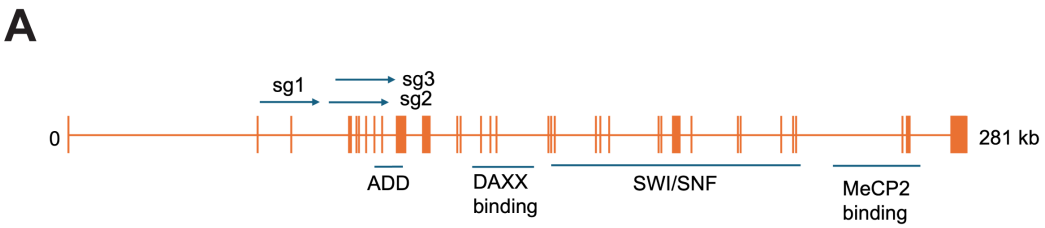
Supplemental Figure S3. ATRX loss is associated with worse outcomes in multiple sarcoma subtypes.

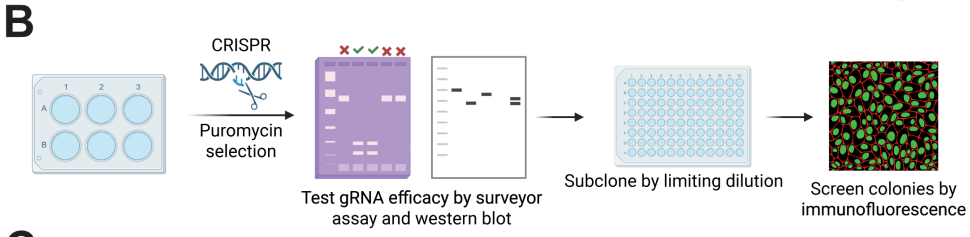
(A) Recurrence-free survival for primary localized STLMS. (B) Overall survival for recurrent/metastatic STLMS. (C) Progression-free survival for recurrent/metastatic STLMS. (D) Recurrence-free survival for primary localized UPS. (E) Overall survival for recurrent/metastatic UPS. (F) Progression-free survival for recurrent/metastatic UPS. (G) Overall survival for DDLPS. (H) Progression-free survival for DDLPS. (I) Percent of tumor cells with TP53 positive nuclei. (J) TP53 intensity by IHC.

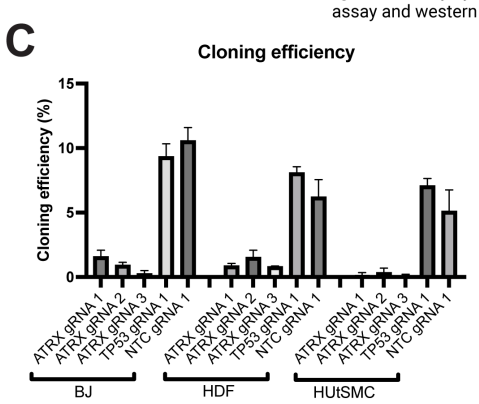


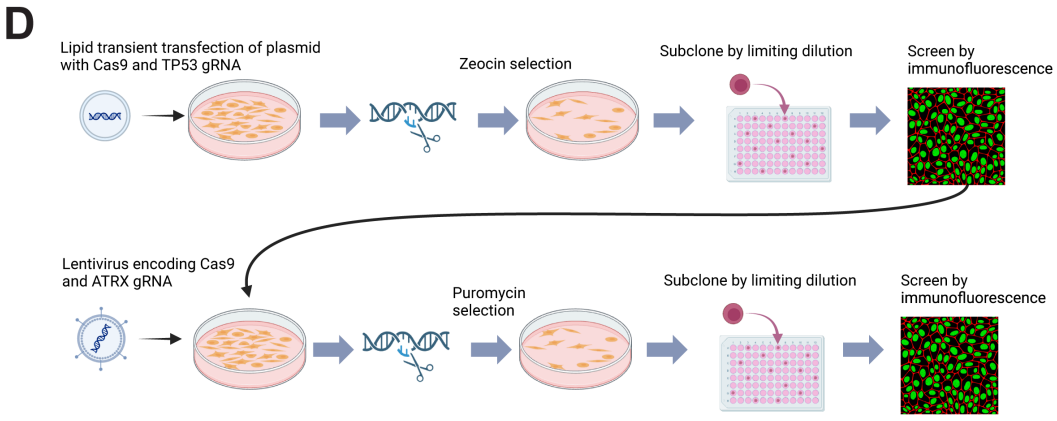
Supplemental Figure S4. Generation of ATRX/TP53 double knockout cells.

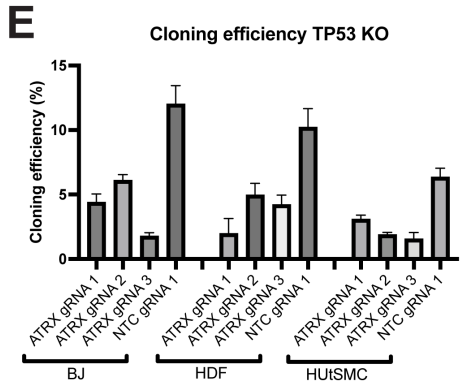
(A) *ATRX* is located at chromosome Xq21.1, spans 281kb, has 40 exons with ~10 different protein-coding transcripts, though the *ATRX* canonical sequence (P46100-1 in Uniprot) has 2492 amino acids, 35 exons. (B) Schematic of CRISPR and assessment of cloning efficiency. CRISPR was used to knockout *ATRX*. Several gRNAs were tested by surveyor assay and western blotting. Pools of cells with successful knockout of *ATRX* were subcloned by limiting dilution. Colonies were expanded and screened by immunofluorescent staining for *ATRX*. (C) Cloning efficiency for *ATRX* knockout cells. Cells were plated by limiting dilution with an expected cell number of either 0.5 or 2 per well in 96-well plates with 3 replicate plates per condition and cell number. Cloning efficiency was calculated as the number of wells with colonies divided by the number of expected colonies based on the number of cells plated. (D) Schematic of sequential gene knockout using CRISPR. Due to the poor cloning efficiency and inability to generate monoclonal *ATRX* knockout cells, we first engineered *TP53* knockout cells and subsequently knocked out *ATRX*. (E) Cloning efficiency for *ATRX* knockouts in *TP53* knockout genetic background. Panels B and D made with Biorender.


A
sg1
sg3
sg2
0
281 kb
ADD
DAXX binding
SWI/SNF
MeCP2 binding


B
CRISPR
Puromycin selection
Test gRNA efficacy by surveyor assay and western blot
Subclone by limiting dilution
Screen colonies by immunofluorescence


C
Cloning efficiency
Cloning efficiency (%)
15
10
5
0
ATRX gRNA 1
ATRX gRNA 2
ATRX gRNA 3
TP53 gRNA 1
NTC gRNA 1
BJ
HDF
HUtSMC

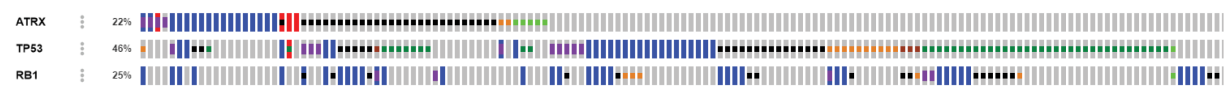

D
Lipid transient transfection of plasmid with Cas9 and TP53 gRNA
Zeocin selection
Subclone by limiting dilution
Screen by immunofluorescence
Lentivirus encoding Cas9 and ATRX gRNA
Puromycin selection
Subclone by limiting dilution
Screen by immunofluorescence


E
Cloning efficiency TP53 KO
Cloning efficiency (%)
15
10
5
0
ATRX gRNA 1
ATRX gRNA 2
ATRX gRNA 3
NTC gRNA 1
BJ
HDF
HUtSMC

Supplemental Figure S5. Co-occurrence of *ATRX* mutation with *TP53* and *RB1* alterations.

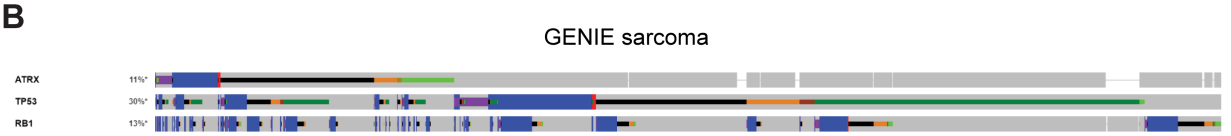
Oncoprint plots demonstrating co-occurrence of *ATRX* alterations with *TP53* and *RB1* alterations in TCGA (A) and GENIE (B) sarcoma cohorts.

A



A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value ▲	Tendency
TP53	RB1	128	62	8	55	>3	<0.001	**<0.001**	**Co-occurrence**
ATRX	TP53	111	25	86	31	0.678	0.131	0.164	Co-occurrence
ATRX	RB1	152	38	45	18	0.678	0.164	0.164	Co-occurrence

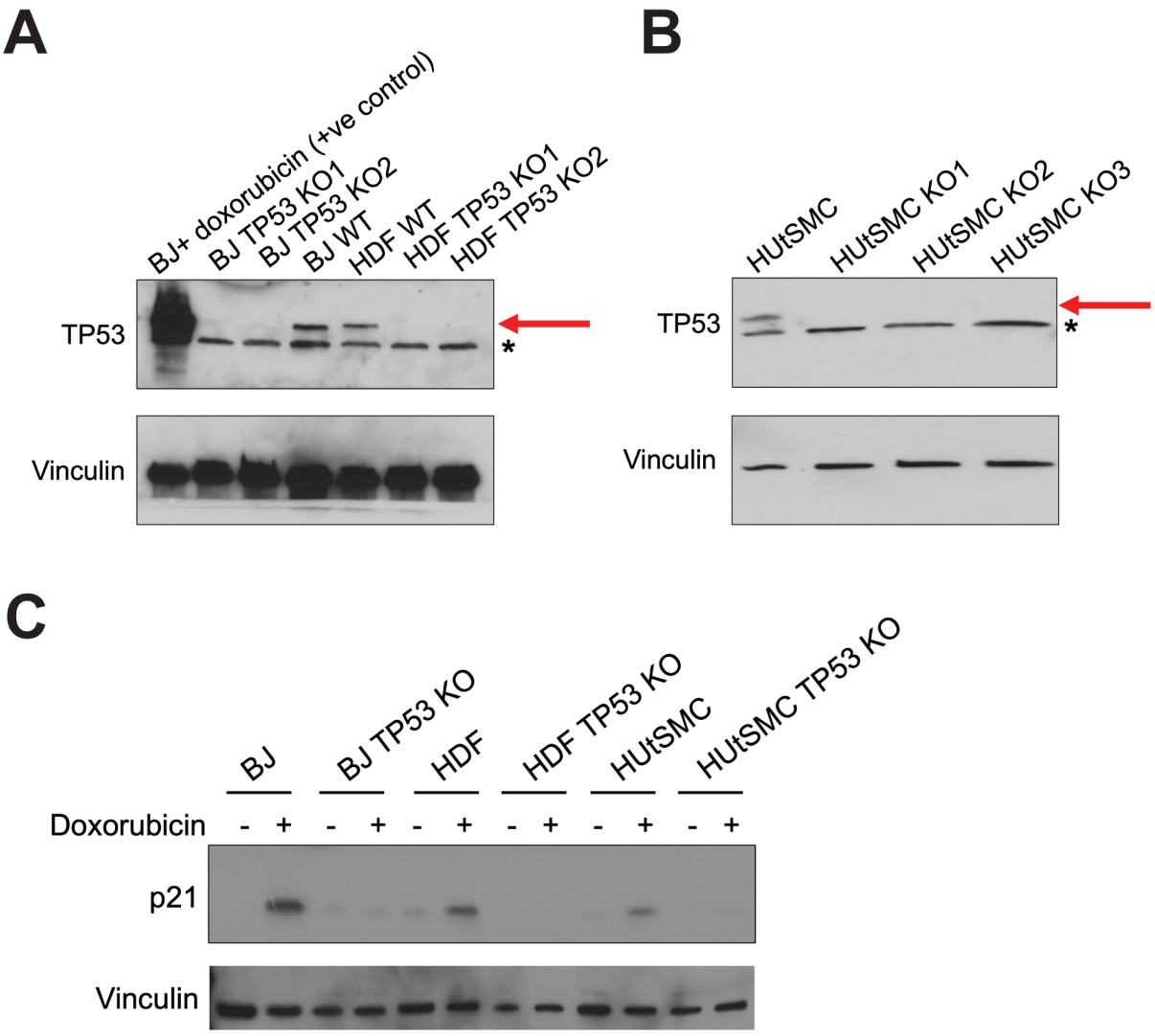
B



A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value ▲	Tendency
TP53	RB1	3144	923	161	564	>3	<0.001	**<0.001**	**Co-occurrence**
ATRX	TP53	3132	173	1120	367	2.569	<0.001	**<0.001**	**Co-occurrence**
ATRX	RB1	3753	314	499	226	2.436	<0.001	**<0.001**	**Co-occurrence**

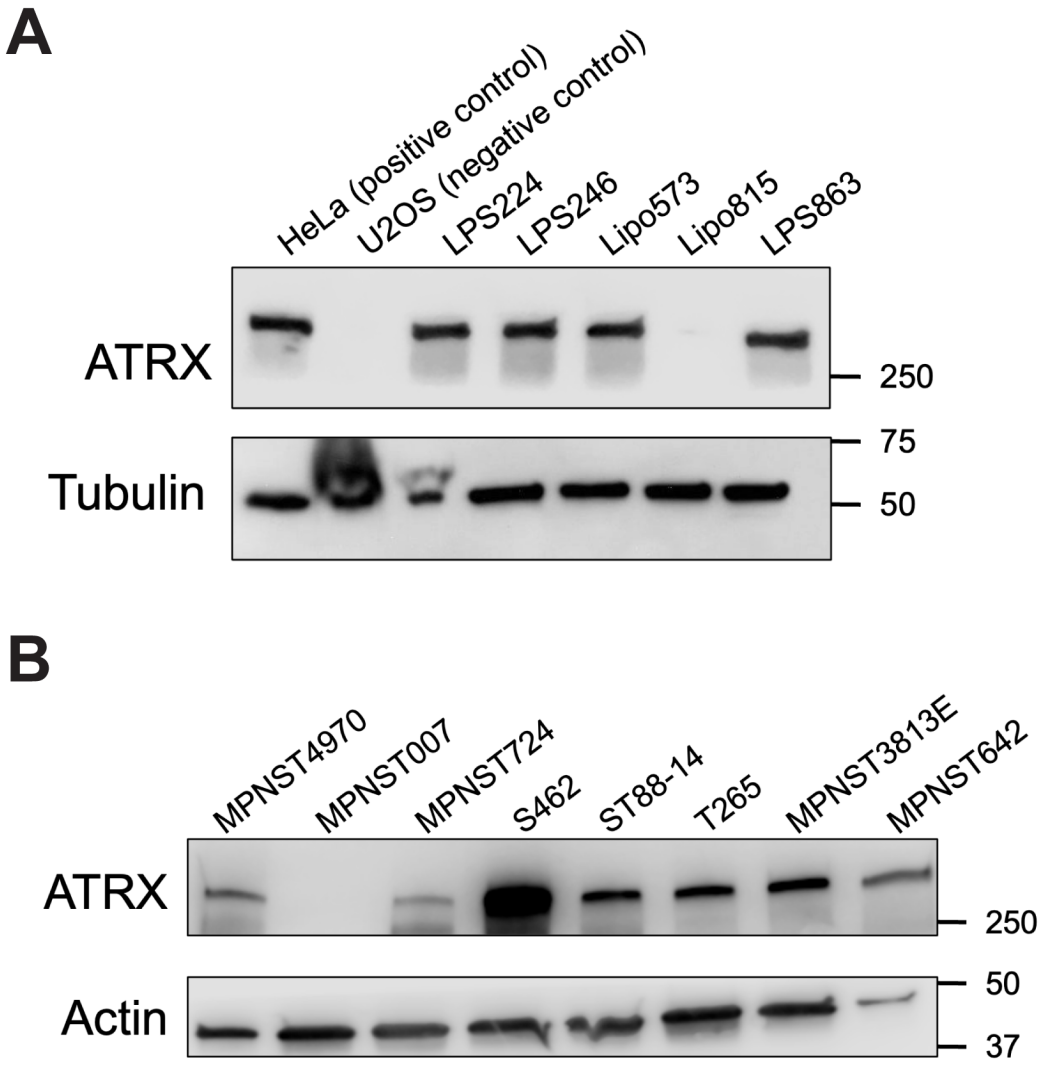
Supplemental Figure S6. Validation of p53 knockout.

(A-B) Western blotting for TP53 in BJ and HDF fibroblasts (A) and HUtSMC (B). The red arrow indicates p53; the lower molecular weight band is non-specific. (C) To confirm p53 functional status, we blotted for a downstream p53 effector, p21, after exposure to the DNA damaging agent doxorubicin.



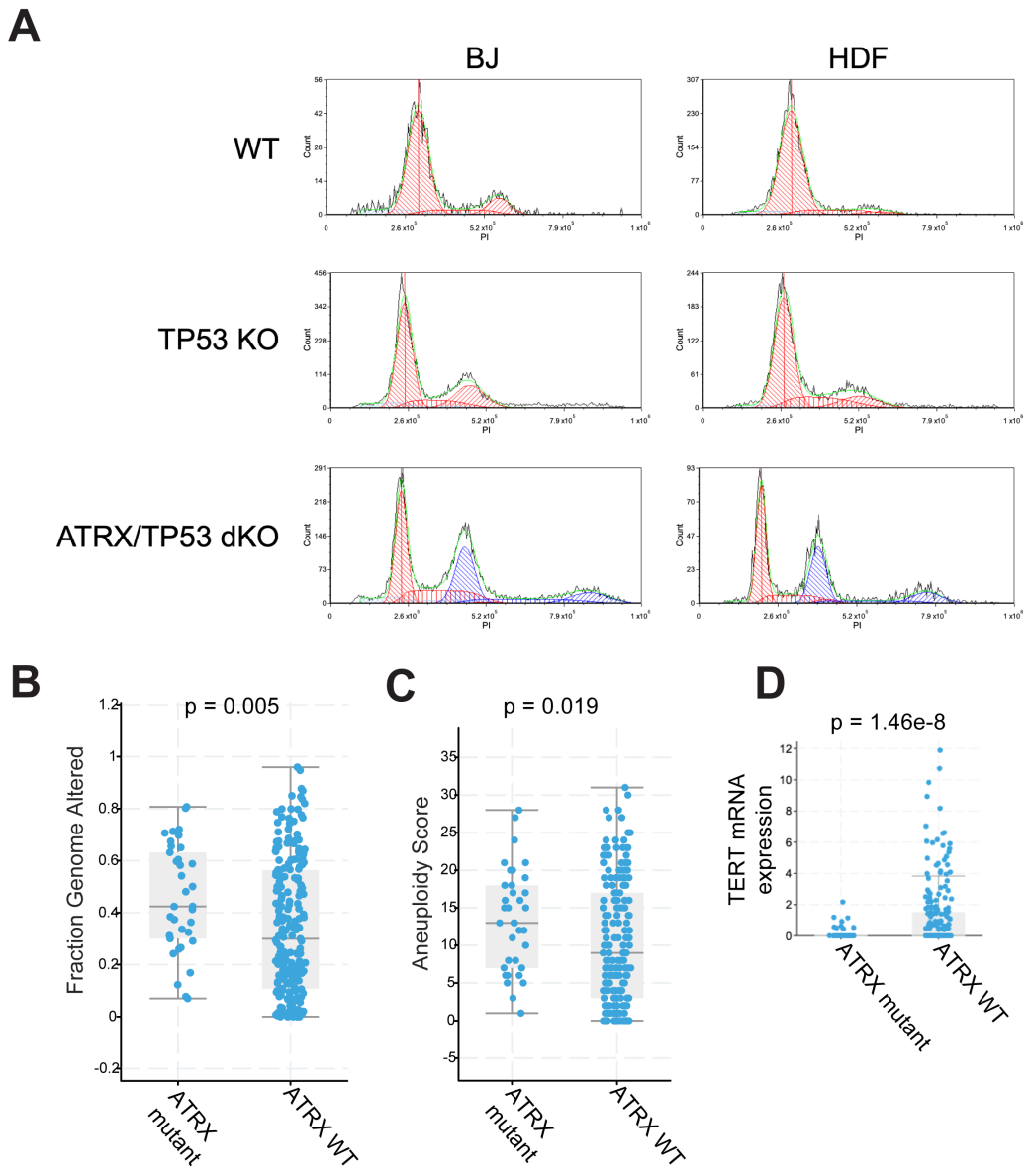
Supplemental Figure S7. Identification of ATRX knockout cell lines.

Western blotting for ATRX in a panel of DDLPS cell lines (A) and MPNST cell lines (B).



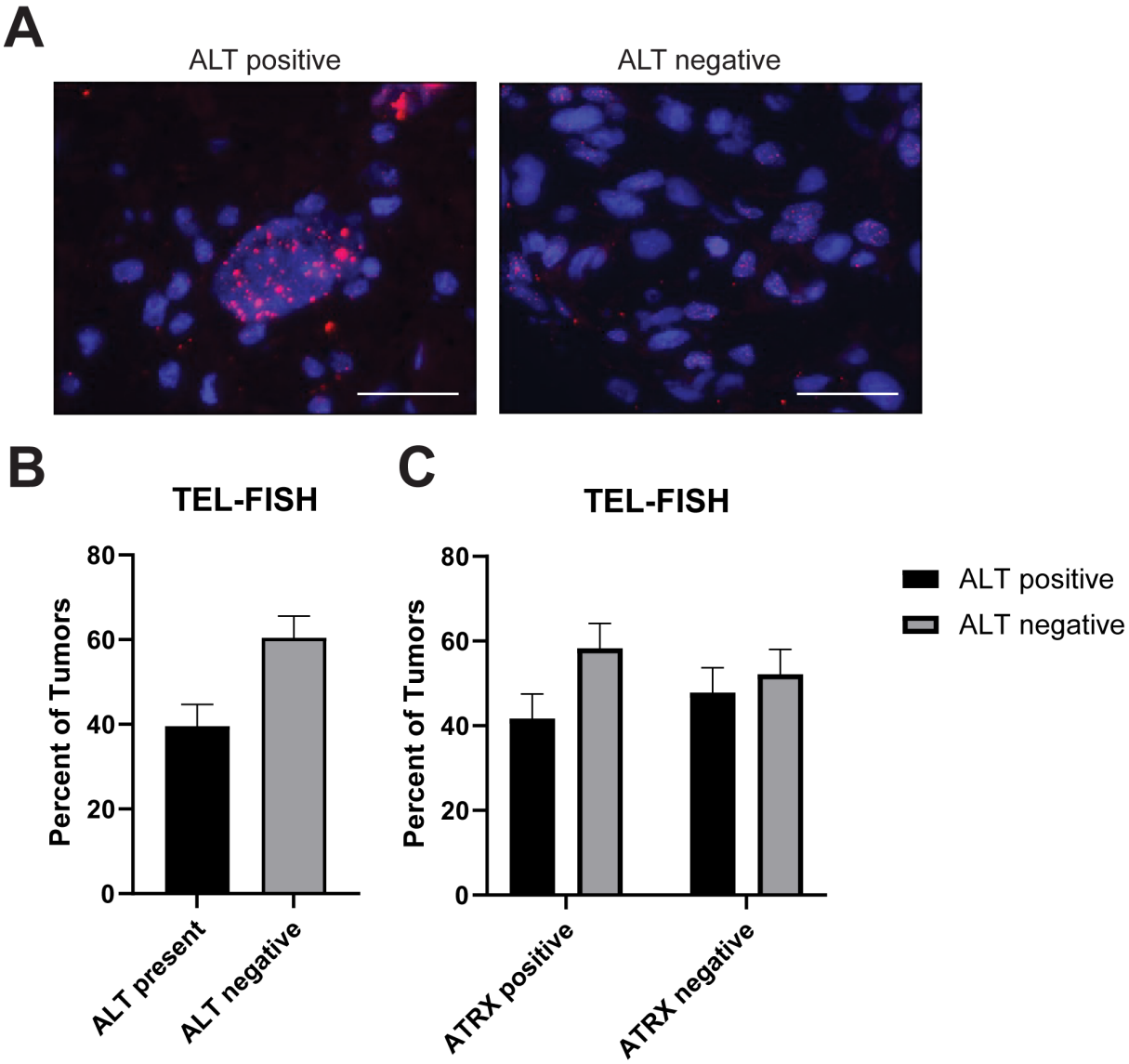
Supplemental Figure S8. ATRX loss is associated with genomic instability.

(A) Representative histograms from flow cytometry for cell cycle analysis after propidium iodide staining. (B) Fraction of the genome altered in *ATRX* mutant versus WT sarcomas in TCGA. (C) Aneuploidy score in *ATRX* mutant versus WT sarcomas in TCGA. (D) TERT mRNA expression in *ATRX* mutant versus WT sarcomas in TCGA. Plots in B-D made in cBioPortal.



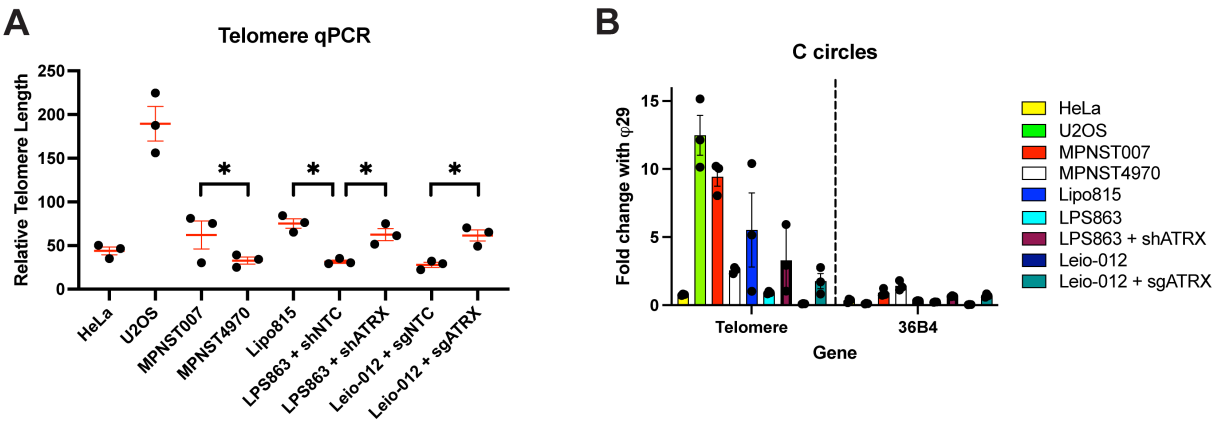
Supplemental Figure S9. ATRX loss in human UPS is not associated with ALT.

(A) Representative images of TEL-FISH showing ALT positive (top) versus negative (bottom) tumors. Scale bars = 50 μm. (B) Percent of tumors with ALT presence or negativity. (C) Percent of ALT positivity based on ATRX expression by IHC.



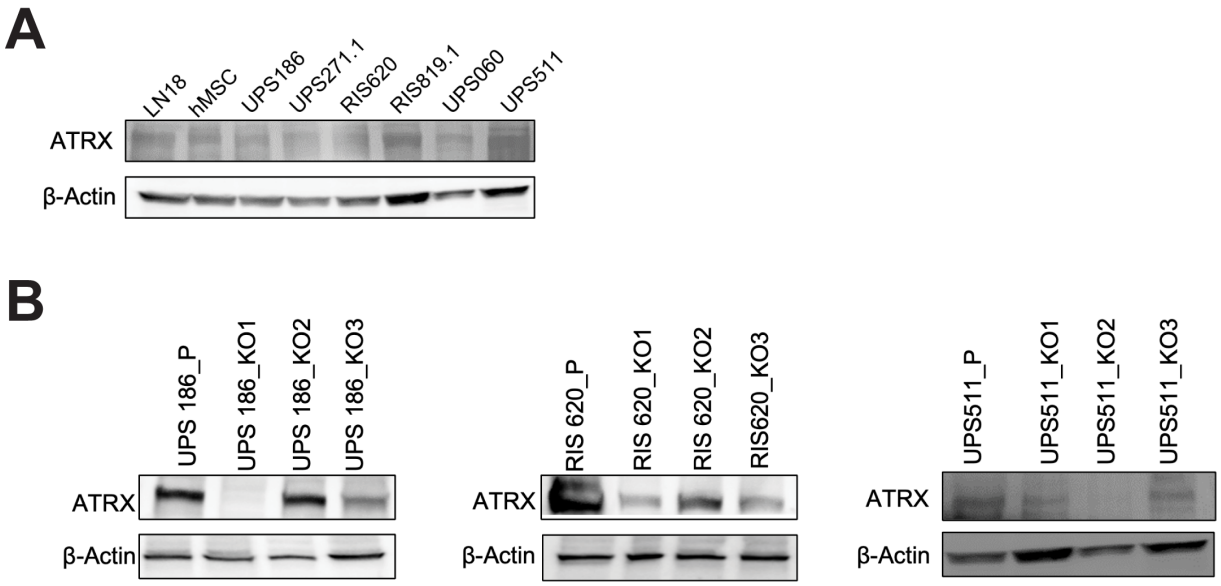
Supplemental Figure S10. ATRX loss is associated with ALT in sarcoma cell lines.

(A) Relative telomere length, which was calculated by using the following formula: RTL = $2^{-(CTtelomere-CT36B4)}$, where RTL is the relative telomere length, $CT_{telomere}$ is the cycle threshold values of telomere and CT_{36B4} is the cycle threshold values of a single-copy gene, 36B4. (B) C circle assay assessed by telomere and 36B4 quantitative PCR with and without the φ29 polymerase, which amplifies C circles. *P value < 0.05.



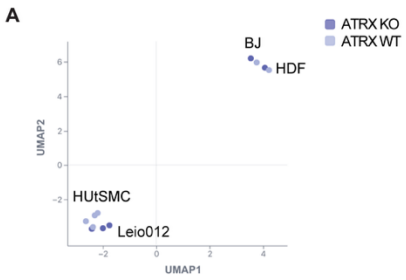
Supplemental Figure S11. Depletion of ATRX with CRISPR in UPS cell lines.

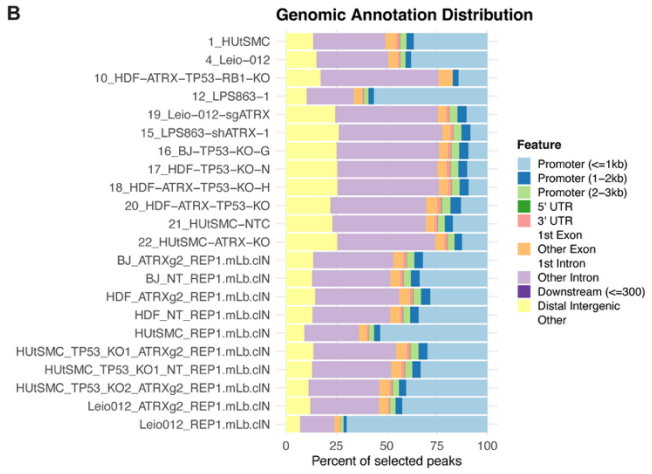
(A) Western blot of ATRX in untreated cell lines. UPS186, UPS271.1, RIS620, RIS819.1, UPS060, and UPS511 are patient-derived UPS cell lines. LN18 is a glioblastoma cell line, hMSC is human mesenchymal stem cells, and these served as positive controls. (B) Western blotting for ATRX in polyclonal populations of UPS cell lines following expression of 3 different all-in-one CRISPR plasmids. Three different gRNAs were screened. P indicates parental control (untreated cells).

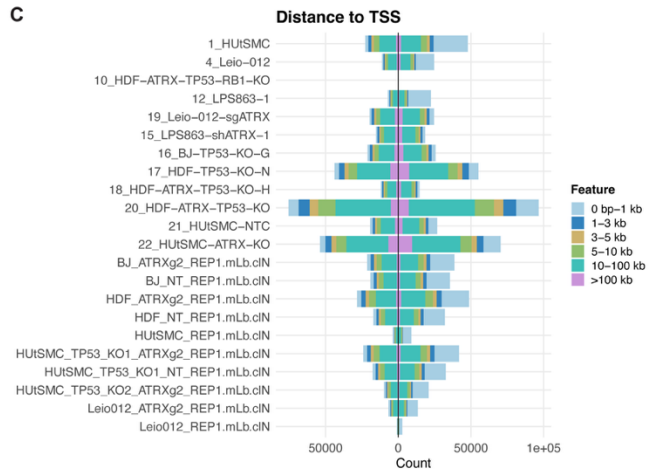


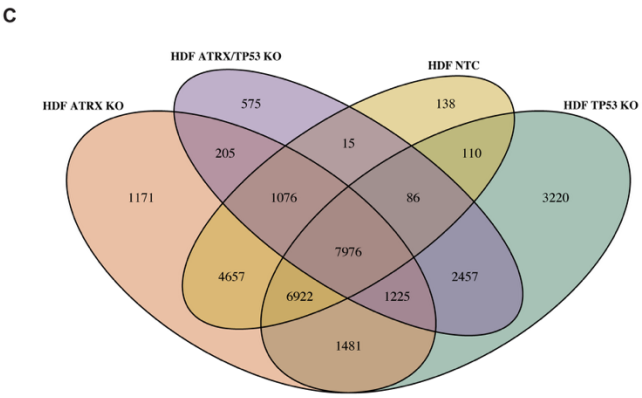
Supplemental Figure S12. Chromatin remodeling in ATRX KO cells.

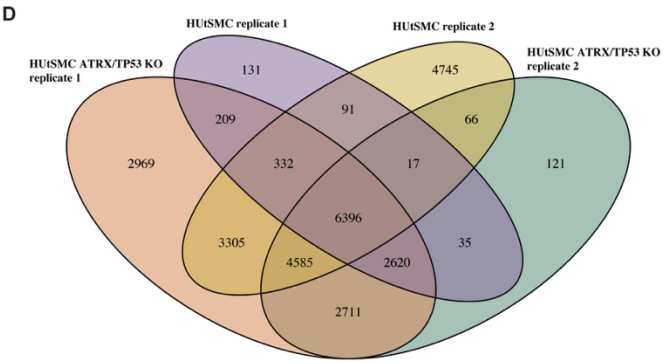
(A) UMAP analysis utilizing ATAC-seq data. *ATRX* KO conditions are colored in dark purple and *ATRX* WT conditions are colored in light purple. (B) Genomic annotation distribution of all the samples included analyzed by ATAC-seq. (C) Distance to the transcription start site (TSS) of the peaks identified by ATAC-seq in all the samples analyzed. (D-F) Sample correlation Venn diagrams of the HDF samples (D), HUtSMC samples (E), and Leio-012 samples (F) analyzed by ATAC-seq.

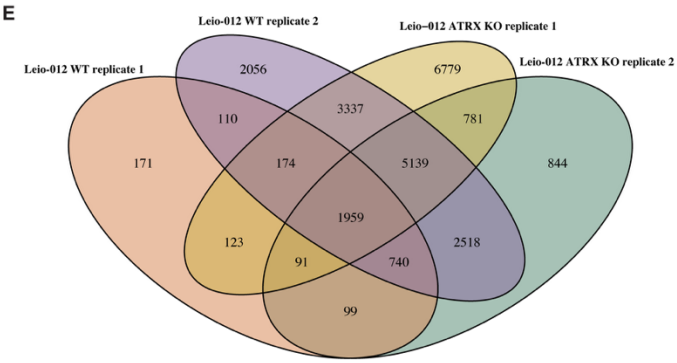

A
ATRX KO
ATRX WT
BJ
HDF
HUtSMC
Leio012
UMAP1
UMAP2


B
Genomic Annotation Distribution
1_HUtSMC
4_Leio−012
10_HDF−ATRX−TP53−RB1−KO
12_LPS863−1
19_Leio−012−sgATRX
15_LPS863−shATRX−1
16_BJ−TP53−KO−G
17_HDF−TP53−KO−N
18_HDF−ATRX−TP53−KO−H
20_HDF−ATRX−TP53−KO
21_HUtSMC−NTC
22_HUtSMC−ATRX−KO
BJ_ATRXg2_REP1.mLb.clN
BJ_NT_REP1.mLb.clN
HDF_ATRXg2_REP1.mLb.clN
HDF_NT_REP1.mLb.clN
HUtSMC_REP1.mLb.clN
HUtSMC_TP53_KO1_ATRXg2_REP1.mLb.clN
HUtSMC_TP53_KO1_NT_REP1.mLb.clN
HUtSMC_TP53_KO2_ATRXg2_REP1.mLb.clN
Leio012_ATRXg2_REP1.mLb.clN
Leio012_REP1.mLb.clN
0
25
50
75
100
Percent of selected peaks
Feature
Promoter (<=1kb)
Promoter (1−2kb)
Promoter (2−3kb)
5' UTR
3' UTR
1st Exon
Other Exon
1st Intron
Other Intron
Downstream (<=300)
Distal Intergenic
Other


C
Distance to TSS
1_HUtSMC
4_Leio−012
10_HDF−ATRX−TP53−RB1−KO
12_LPS863−1
19_Leio−012−sgATRX
15_LPS863−shATRX−1
16_BJ−TP53−KO−G
17_HDF−TP53−KO−N
18_HDF−ATRX−TP53−KO−H
20_HDF−ATRX−TP53−KO
21_HUtSMC−NTC
22_HUtSMC−ATRX−KO
BJ_ATRXg2_REP1.mLb.clN
BJ_NT_REP1.mLb.clN
HDF_ATRXg2_REP1.mLb.clN
HDF_NT_REP1.mLb.clN
HUtSMC_REP1.mLb.clN
HUtSMC_TP53_KO1_ATRXg2_REP1.mLb.clN
HUtSMC_TP53_KO1_NT_REP1.mLb.clN
HUtSMC_TP53_KO2_ATRXg2_REP1.mLb.clN
Leio012_ATRXg2_REP1.mLb.clN
Leio012_REP1.mLb.clN
50000
0
50000
1e+05
Count
Feature
0 bp−1 kb
1−3 kb
3−5 kb
5−10 kb
10−100 kb
>100 kb

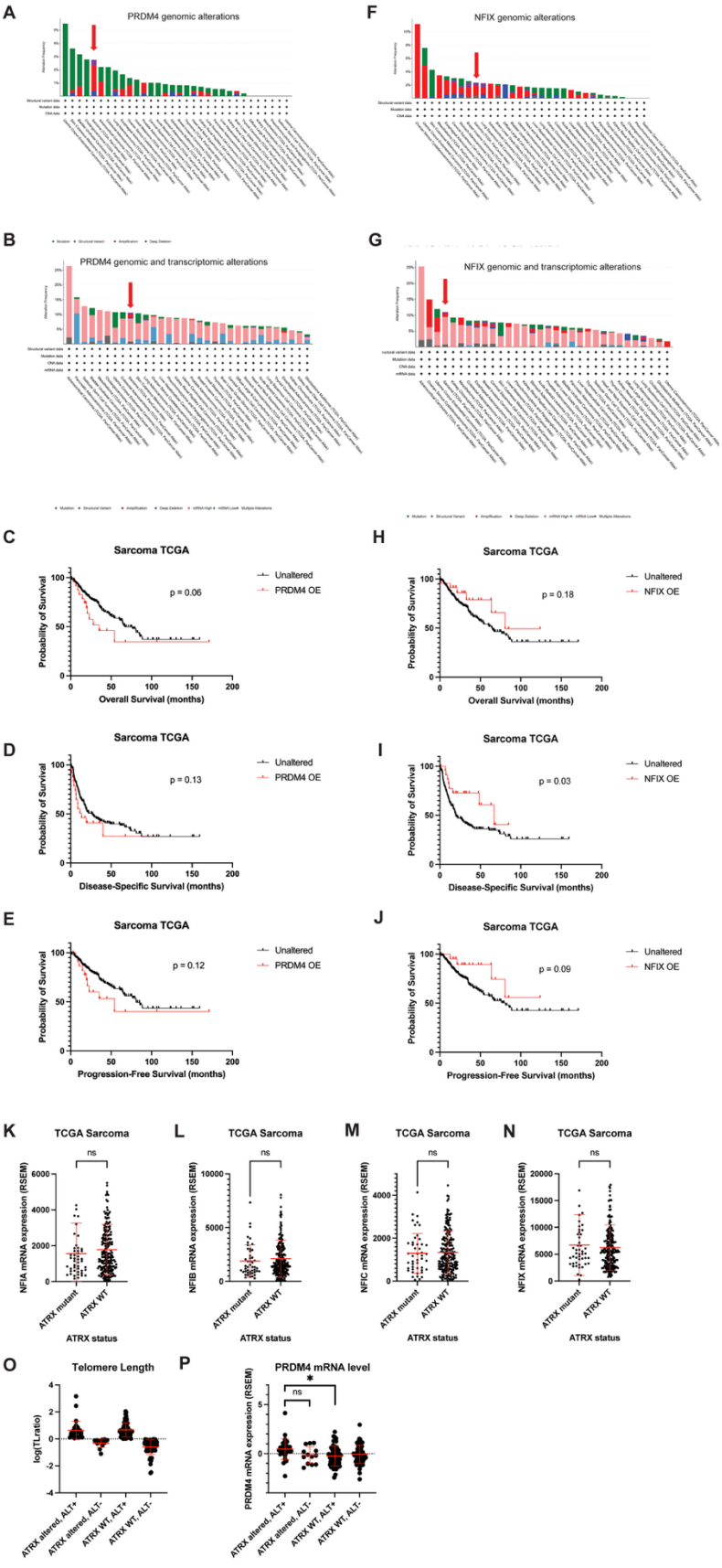

C
HDF ATRX/TP53 KO
HDF NTC
HDF ATRX KO
HDF TP53 KO
575
138
15
205
110
1171
1076
86
3220
7976
4657
2457
6922
1225
1481


D
HUtSMC replicate 1
HUtSMC replicate 2
HUtSMC ATRX/TP53 KO replicate 1
HUtSMC ATRX/TP53 KO replicate 2
131
4745
91
209
66
2969
332
17
121
6396
3305
35
4585
2620
2711


E
Leio-012 WT replicate 2
Leio−012 ATRX KO replicate 1
Leio-012 WT replicate 1
Leio-012 ATRX KO replicate 2
2056
6779
3337
110
781
171
174
5139
844
1959
123
2518
91
740
99

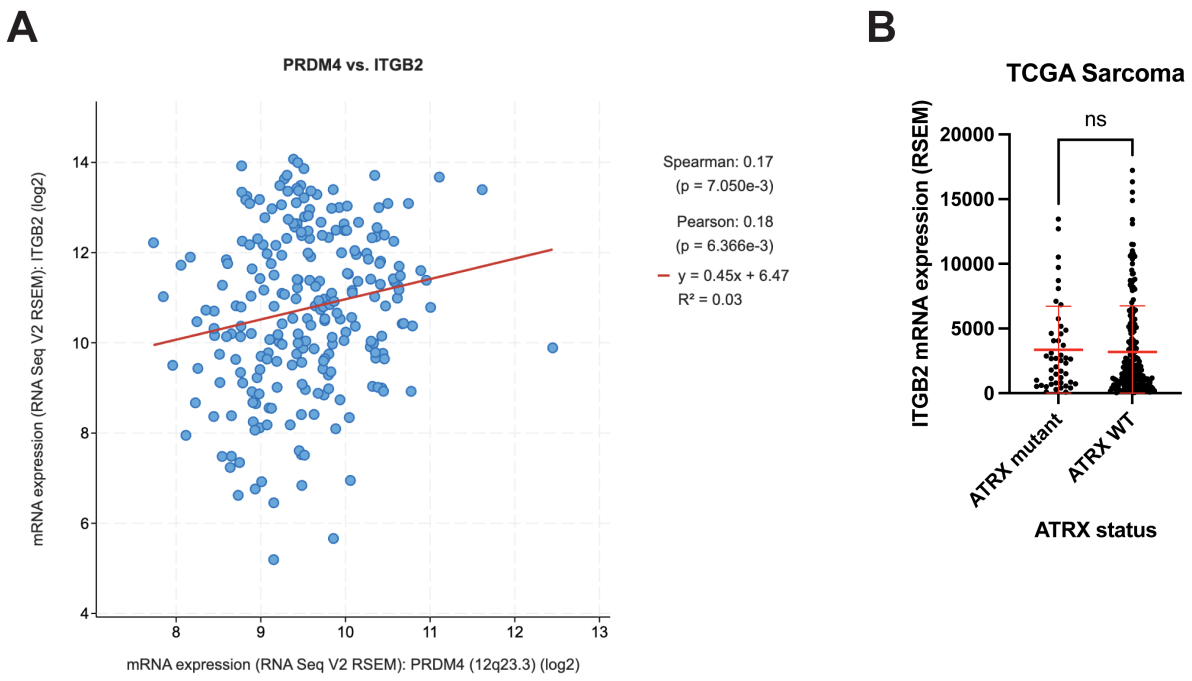
Supplemental Figure S13. Alterations in *PRDM4* and *NFIX* in sarcoma TCGA.

(A-B) PRDM4 genomic alterations (A) and combined genomic and transcriptomic alterations (B) by cancer subtype in TCGA. The red arrow highlights the sarcoma dataset. (C-E) Overall survival (C), disease-specific survival (D), and progression-free survival (E) based on the presence or absence of PRDM4 overexpression (OE). (F-G) NFIX genomic alterations (F) and combined genomic and transcriptomic alterations (G) by cancer subtype in TCGA. The red arrow highlights the sarcoma dataset. (H-J) Overall survival (H), disease-specific survival (I), and progression-free survival (J) based on the presence or absence of NFIX overexpression (OE). (K-N) Gene expression of the 4 NFI genes based on ATRX expression. (O-P) Sarcomas in TCGA dataset were divided into 4 groups based on ATRX alterations (mutant or WT) and ALT (positive or negative). Panel O plots log of telomere length ratio (tumor to matched normal), and panel P plots PRDM4 mRNA expression in each of the 4 groups. NS = not significant. *P < 0.01. P values on Kaplan Meier curves are from log-rank tests.


A
PRDM4 genomic alterations
F
NFIX genomic alterations
B
PRDM4 genomic and transcriptomic alterations
G
NFIX genomic and transcriptomic alterations
C
Sarcoma TCGA
p = 0.06
Unaltered
PRDM4 OE
Probability of Survival
Overall Survival (months)
H
Sarcoma TCGA
p = 0.18
Unaltered
NFIX OE
Probability of Survival
Overall Survival (months)
D
Sarcoma TCGA
p = 0.13
Unaltered
PRDM4 OE
Probability of Survival
Disease-Specific Survival (months)
I
Sarcoma TCGA
p = 0.03
Unaltered
NFIX OE
Probability of Survival
Disease-Specific Survival (months)
E
Sarcoma TCGA
p = 0.12
Unaltered
PRDM4 OE
Probability of Survival
Progression-Free Survival (months)
J
Sarcoma TCGA
p = 0.09
Unaltered
NFIX OE
Probability of Survival
Progression-Free Survival (months)
K
TCGA Sarcoma
ns
NFIA mRNA expression (RSEM)
ATRX mutant
ATRX WT
ATRX status
L
TCGA Sarcoma
ns
NFIB mRNA expression (RSEM)
ATRX mutant
ATRX WT
ATRX status
M
TCGA Sarcoma
ns
NFIC mRNA expression (RSEM)
ATRX mutant
ATRX WT
ATRX status
N
TCGA Sarcoma
ns
NFIX mRNA expression (RSEM)
ATRX mutant
ATRX WT
ATRX status
O
Telomere Length
log(TL ratio)
ATRX altered, ALT+
ATRX altered, ALT-
ATRX WT, ALT+
ATRX WT, ALT-
P
PRDM4 mRNA level
*
ns
PRDM4 mRNA expression (RSEM)
ATRX altered, ALT+
ATRX altered, ALT-
ATRX WT, ALT+
ATRX WT, ALT-

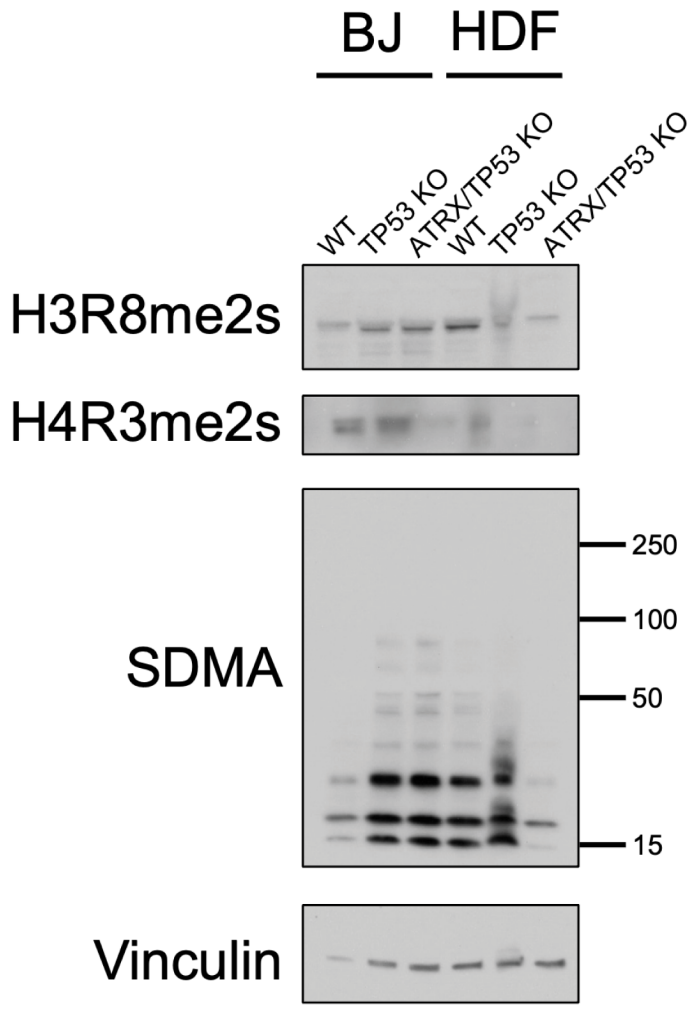
Supplemental Figure S14. Correlation with ITGB2.

(A) Correlation of *ITGB2* with *PRDM4* in sarcoma TCGA cohort. (B) *ITGB2* expression in *ATRX* mutant versus WT in TCGA sarcoma database.



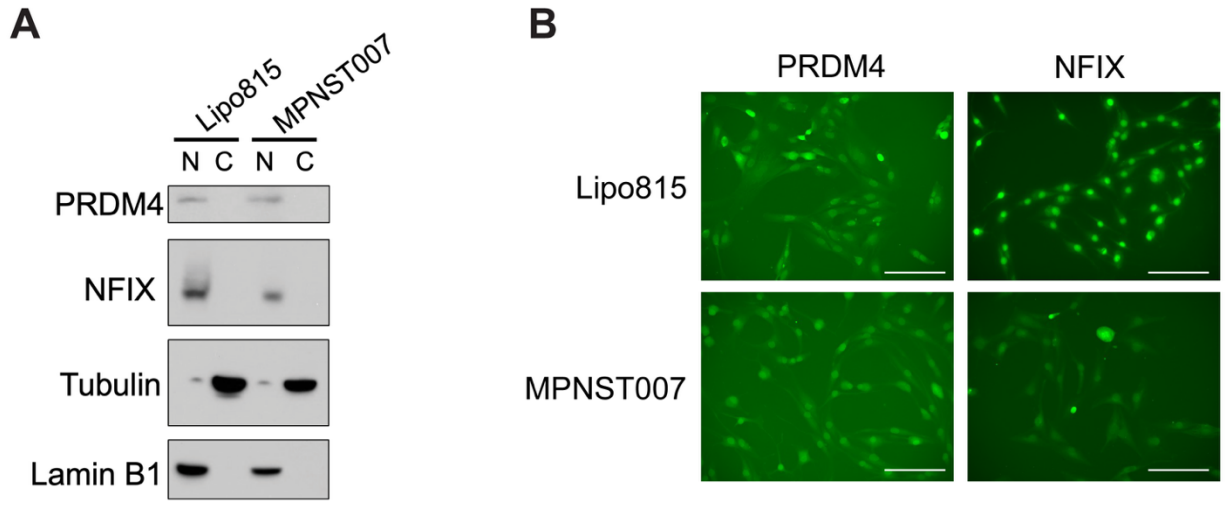
Supplemental Figure S15. Analysis of PRMT5 targets.

PRDM4 interacts with PRMT5. Western blots showing PRMT5 products in the indicated BJ and HDF genotypes. Nuclear extracts were used for H3R8me2s and H4R3me2s blots. whole cell extracts were used for SDMA blot.



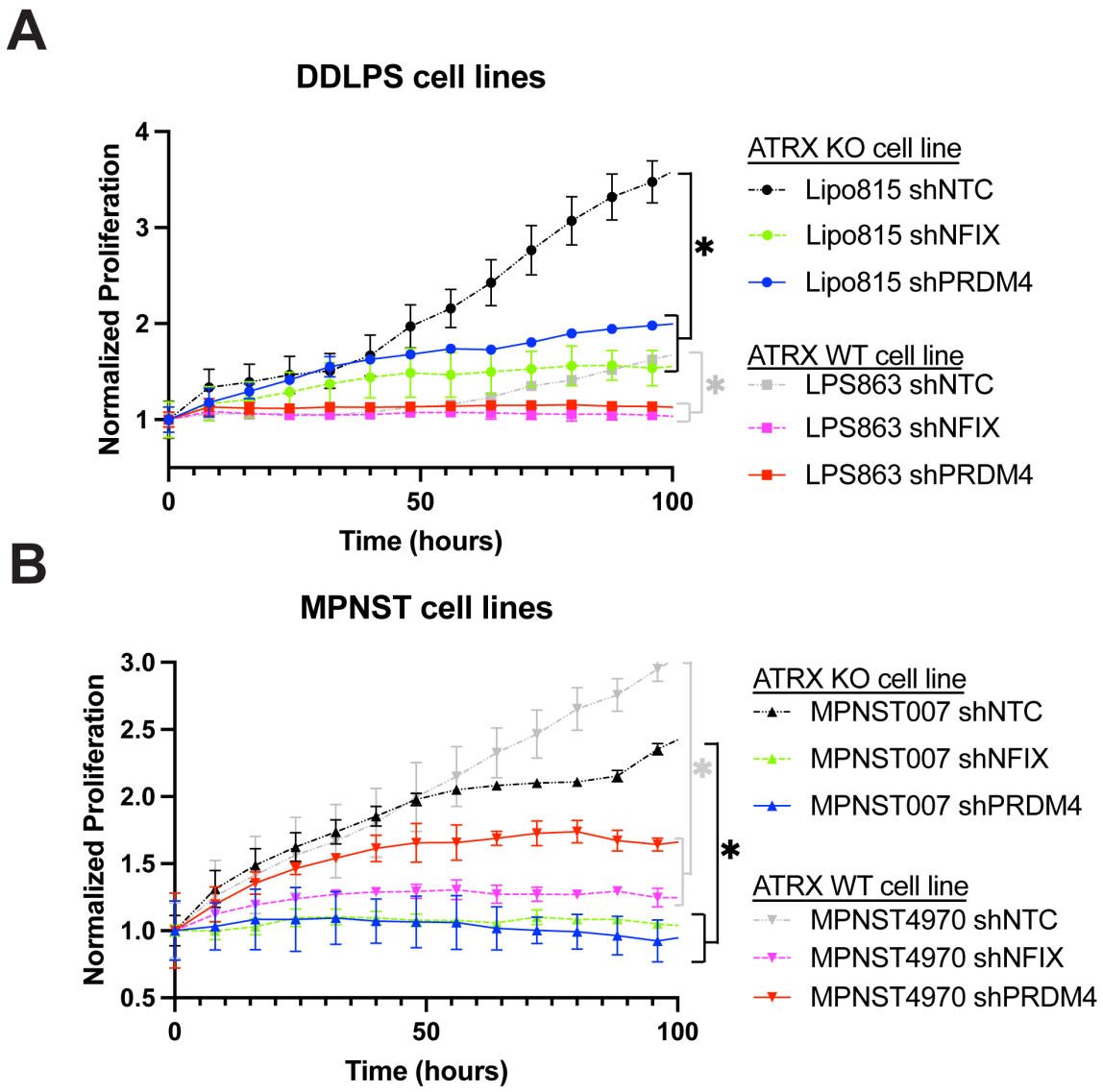
Supplemental Figure S16. Validation of PRDM4 and NFIX antibodies.

(A) Nuclear and cytoplasmic fractions were separated and used to blot for PRDM4 and NFIX. Tubulin was used as a cytoplastic loading control, and lamin B1 was used as a nuclear loading control. (B) Immunofluorescent staining for PRDM4 and NFIX, demonstrating nuclear localization. Scale bar = 50 μm.



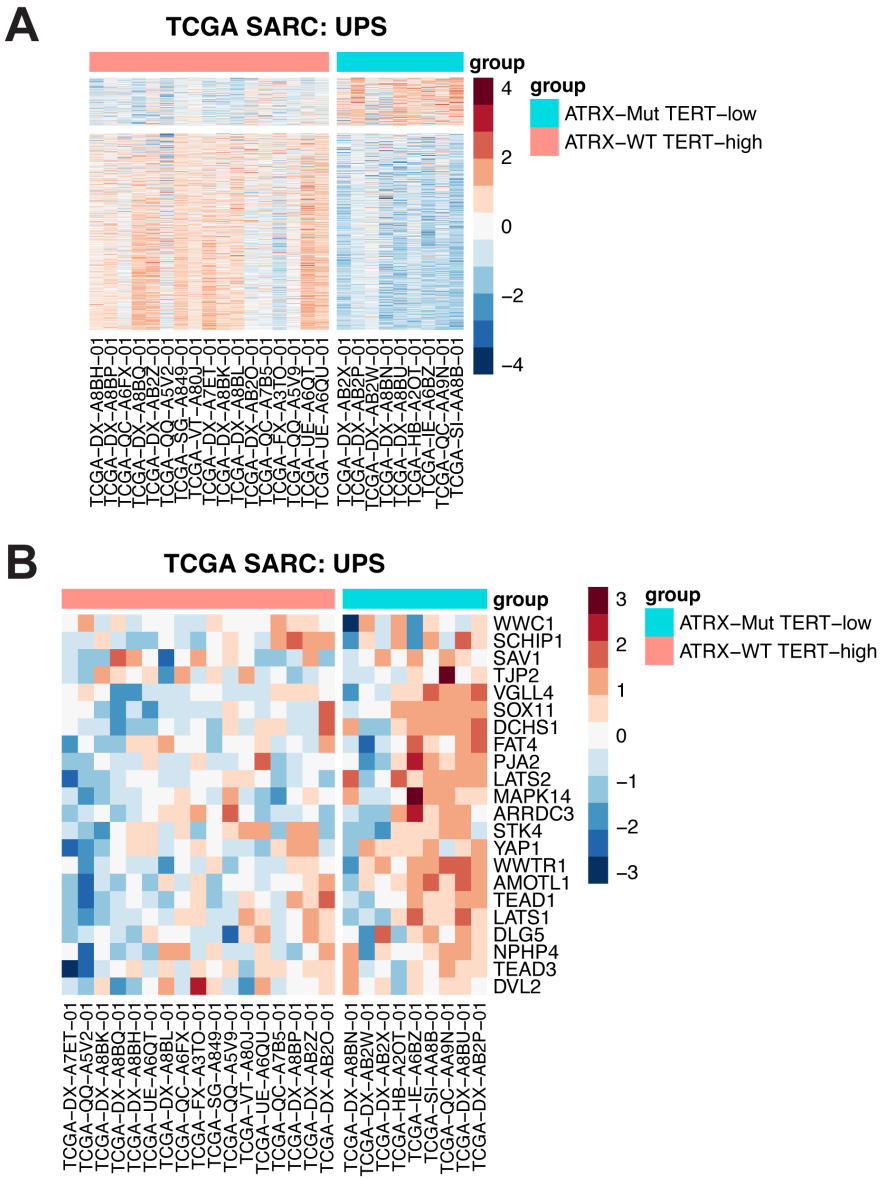
Supplemental Figure S17. Proliferation with PRDM4 and NFIX depletion.

Cells were transduced with lentivirus encoding shRNA targeting either PRDM4, NFIX, or a non-targeting control (NTC). This was performed in ATRX deficient and intact DDLPS cell lines (A) and MPNST cell lines (B). Results from this figure are summarized in Figure 6D.



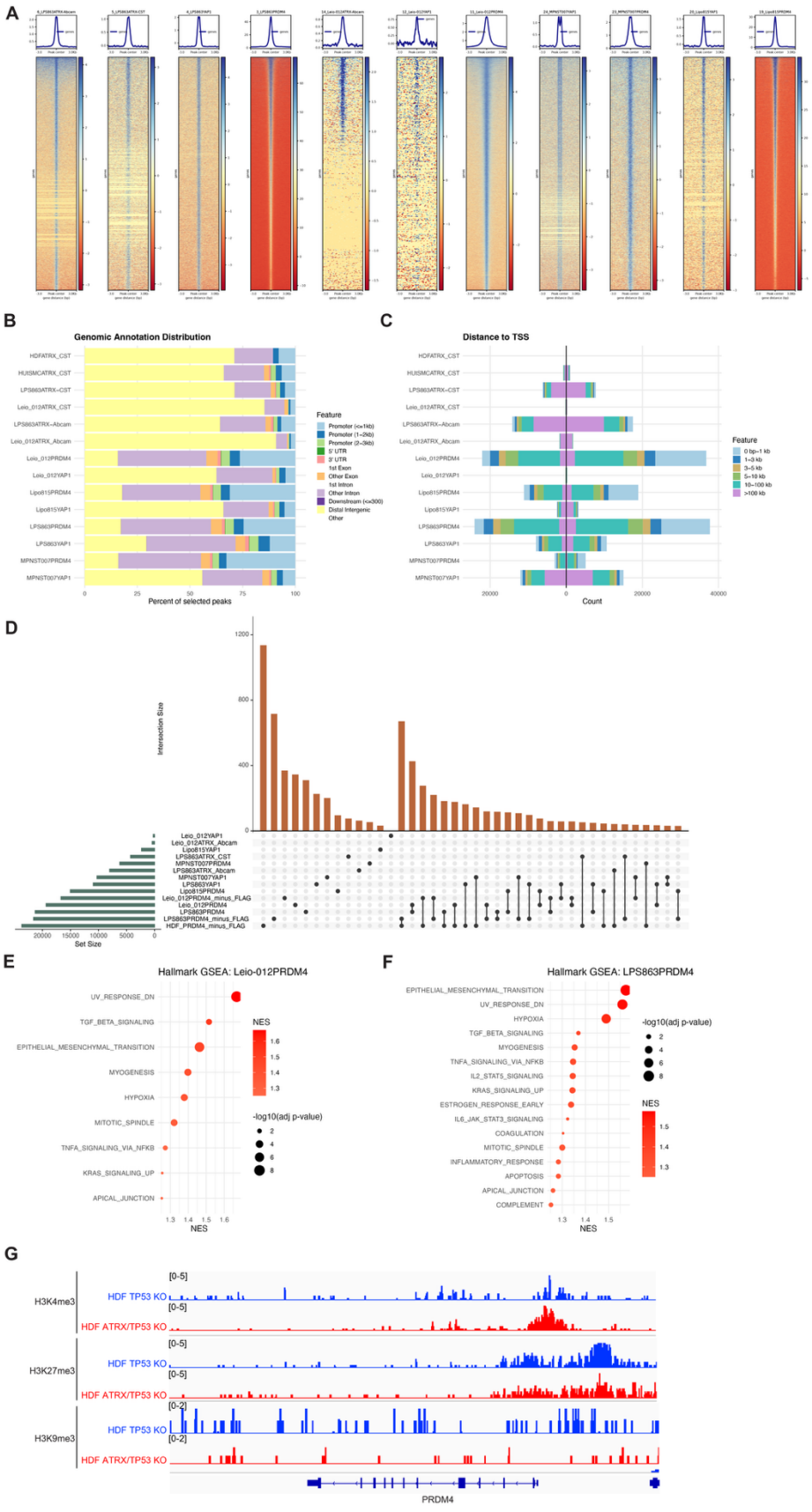
Supplemental Figure S18. Upregulation of YAP-TEAD signaling in ATRX-deficient sarcomas.

(A) Heatmap of DEG expression between ATRX functional (ATRX-WT TERT-high) and inactive (ATRX-mutated TERT-low) tumor samples from TCGA SARC UPS dataset. (B) Expression of YAP signaling pathway genes (from GSEA pathway terms GOBP_HIPPO_SIGNALING and REACTOME_SIGNALING_BY_HIPPO) which are increased in ATRX inactive UPS tumors.



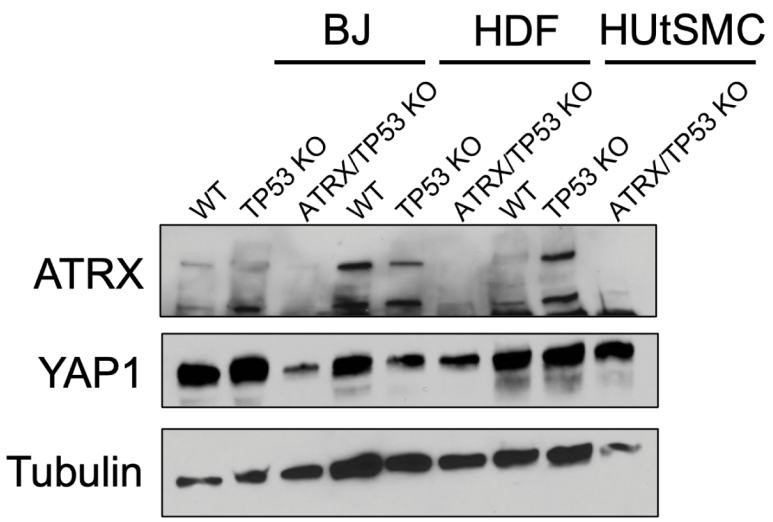
Supplemental Figure S19. CUT&RUN analysis of ATRX, PRDM4, and YAP1.

(A) Heatmap analysis of CUT&RUN peaks. (B) Genomic annotation of CUT&RUN peaks. (C-D) Gene set enrichment analysis using Hallmark collections for PRDM4-regulated peaks in Leio-012 (C) and LPS863 cells (D). (E-F) Pathway analysis using hallmark gene set enrichment analysis for the genes identified by PRDM4 CUT&RUN in Leio-012 (E) and LPS863 (F) cells. (G) Track plots from CUT&RUN at the PRDM4 locus for histone marks associated with activation (H3K4me3) or repression (H3K27me3, H3K9me3).


A
B
Genomic Annotation Distribution
HDFATRX_CST
HUISMCATRX_CST
LPS863ATRX-CST
Leio_012ATRX_CST
LPS863ATRX-Abcam
Leio_012ATRX_Abcam
Leio_012PRDM4
Leio_012YAP1
Lipo815PRDM4
Lipo815YAP1
LPS863PRDM4
LPS863YAP1
MPNST007PRDM4
MPNST007YAP1
Feature
Promoter (<=1kb)
Promoter (1~2kb)
Promoter (2~3kb)
5' UTR
3' UTR
1st Exon
Other Exon
1st Intron
Other Intron
Downstream (<=300)
Distal Intergenic
Other
Percent of selected peaks
C
Distance to TSS
Feature
0 bp~1 kb
1~3 kb
3~5 kb
5~10 kb
10~100 kb
>100 kb
Count
D
Intersection Size
Set Size
Leio_012YAP1
Leio_012ATRX_Abcam
Lipo815YAP1
LPS863ATRX_CST
MPNST007PRDM4
LPS863ATRX_Abcam
MPNST007YAP1
LPS863YAP1
Lipo815PRDM4
Leio_012PRDM4_minus_FLAG
Leio_012PRDM4
LPS863PRDM4
LPS863PRDM4_minus_FLAG
HDF_PRDM4_minus_FLAG
E
Hallmark GSEA: Leio-012PRDM4
UV_RESPONSE_DN
TGF_BETA_SIGNALING
EPITHELIAL_MESENCHYMAL_TRANSITION
MYOGENESIS
HYPOXIA
MITOTIC_SPINDLE
TNFA_SIGNALING_VIA_NFKB
KRAS_SIGNALING_UP
APICAL_JUNCTION
NES
-log10(adj p-value)
F
Hallmark GSEA: LPS863PRDM4
EPITHELIAL_MESENCHYMAL_TRANSITION
UV_RESPONSE_DN
HYPOXIA
TGF_BETA_SIGNALING
MYOGENESIS
TNFA_SIGNALING_VIA_NFKB
IL2_STAT5_SIGNALING
KRAS_SIGNALING_UP
ESTROGEN_RESPONSE_EARLY
IL6_JAK_STAT3_SIGNALING
COAGULATION
MITOTIC_SPINDLE
INFLAMMATORY_RESPONSE
APOPTOSIS
APICAL_JUNCTION
COMPLEMENT
NES
-log10(adj p-value)
G
H3K4me3
H3K27me3
H3K9me3
HDF TP53 KO
HDF ATRX/TP53 KO
[0-5]
[0-2]
PRDM4

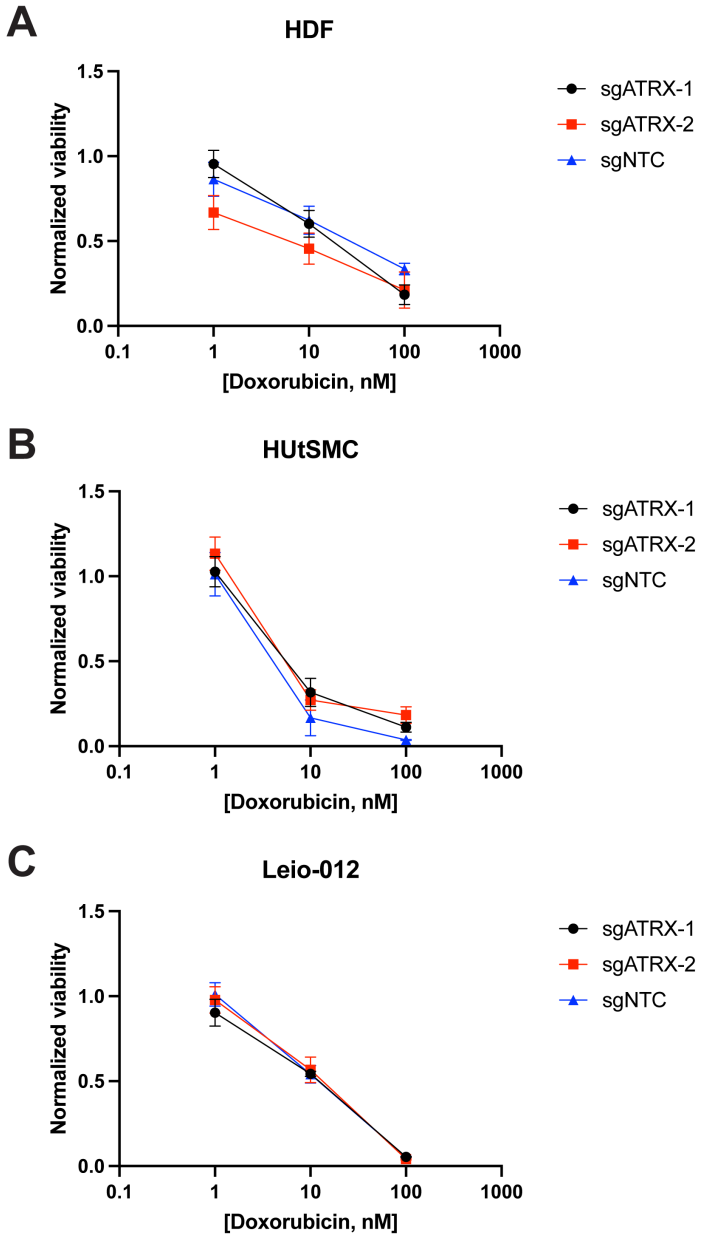
Supplemental Figure S20. YAP1 levels do not change significantly with ATRX loss.

YAP1 western blotting in BJ, HDF, and HUtSMC cells.



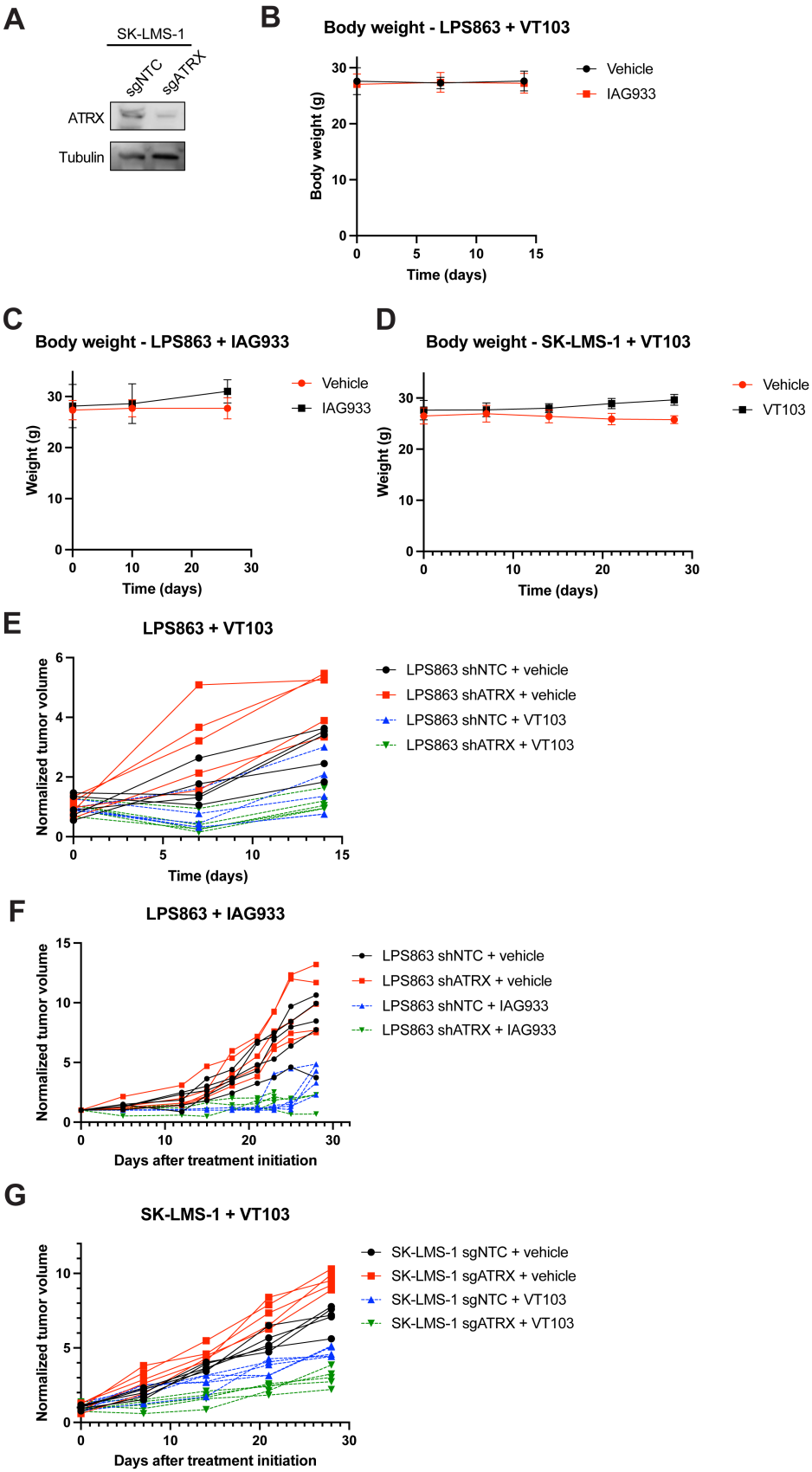
Supplemental Figure S21. ATRX loss does not sensitize to other stressors.

HDF (A), HutSMC (B), and Leio-012 (C) expressing Cas9 and either gRNA targeting ATRX (2 different guides were used) or a non-targeting control gRNA (sgNTC) were treated with the indicated doses of doxorubicin. MTT assay was used to assess viability after 3 days of treatment. Viability was normalized to untreated controls within each cell line and condition.



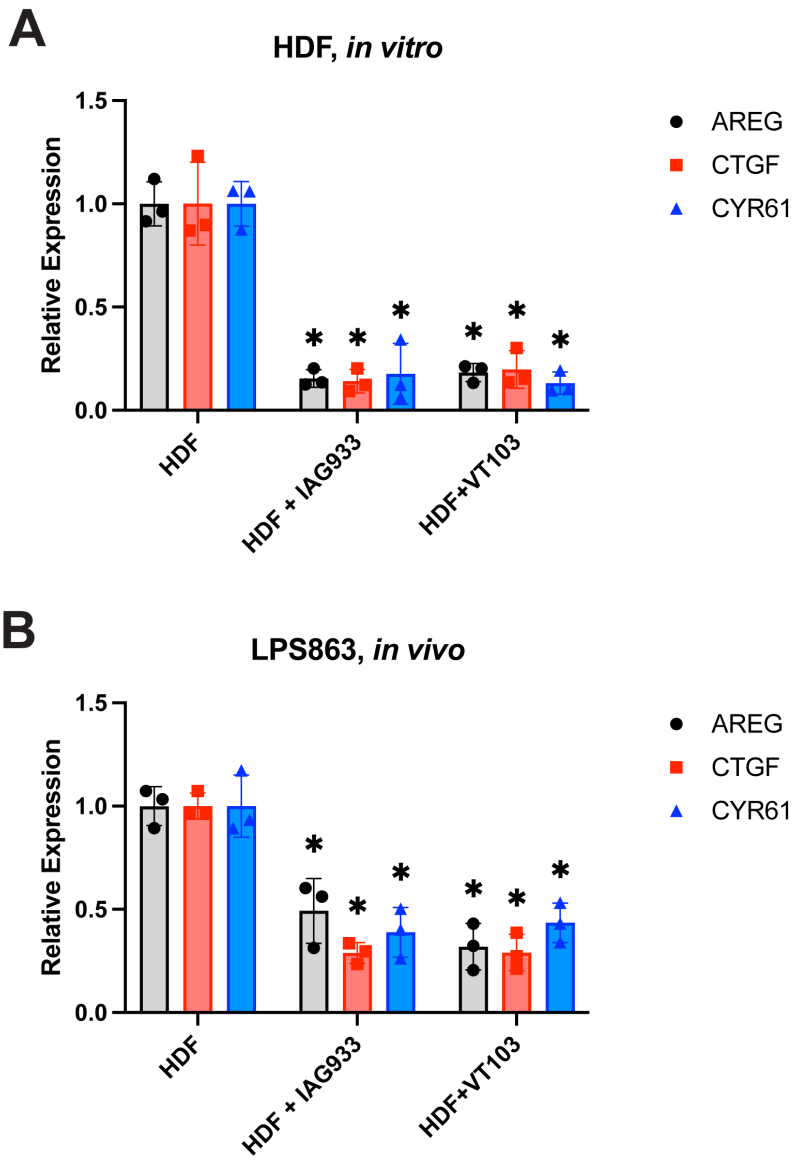
Supplemental Figure S22. Treatment of sarcoma xenograft models with TEAD inhibitors .

(A) Western blot showing depletion of *ATRX* using CRISPR in SK-LMS-1. (B-D) Body weight of mice over time for each of the following experiments: LPS863 xenografts treated with VT103 (B), LPS863 xenografts treated with IAG933 (C), and SK-LMS-1 xenografts treated with VT103 (D). (E-G) Individual growth curves for: LPS863 xenografts treated with VT103 (E), LPS863 xenografts treated with IAG933 (F), and SK-LMS-1 xenografts treated with VT103 (G).


A
SK-LMS-1
sgNTC
sgATRX
ATRX
Tubulin
B
Body weight - LPS863 + VT103
Body weight (g)
Time (days)
Vehicle
IAG933
C
Body weight - LPS863 + IAG933
Weight (g)
Time (days)
Vehicle
IAG933
D
Body weight - SK-LMS-1 + VT103
Weight (g)
Time (days)
Vehicle
VT103
E
LPS863 + VT103
Normalized tumor volume
Time (days)
LPS863 shNTC + vehicle
LPS863 shATRX + vehicle
LPS863 shNTC + VT103
LPS863 shATRX + VT103
F
LPS863 + IAG933
Normalized tumor volume
Days after treatment initiation
LPS863 shNTC + vehicle
LPS863 shATRX + vehicle
LPS863 shNTC + IAG933
LPS863 shATRX + IAG933
G
SK-LMS-1 + VT103
Normalized tumor volume
Days after treatment initiation
SK-LMS-1 sgNTC + vehicle
SK-LMS-1 sgATRX + vehicle
SK-LMS-1 sgNTC + VT103
SK-LMS-1 sgATRX + VT103

Supplemental Figure S23. Treatment with TEAD inhibitors reduces TEAD target genes.

(A) HDF cells treated *in vitro* with the TEAD inhibitors IAG933 and VT103 were assessed for TEAD target gene expression by qRT-PCR. (B) LPS863 xenografted tumors treated *in vivo* were explanted and utilized for qRT-PCR to assess TEAD target gene expression.



Supplemental Tables

Table S1. Patient characteristics summary.

Table S2. ATRX status by sarcoma subtype.

Table S3. Multivariate analysis.

Table S4. ATRX co-dependencies.

Table S5. ATAC-seq data quality.

Table S6. Transcription factor motif analysis.

Table S7. PRDM4 expression correlation with YAP1-TEAD target gene expression.

Table S8. Antibodies.

Table S9. qRT-PCR primer sequences.

Supplemental Methods

ATAC-seq

ATAC was performed as previously described (80) using Illumina Tagment DNA Enzyme and Buffer kit (20034197). Sequencing was performed on an Illumina NextSeq500 at the MDACC Advanced Technology Genomics Core. Analysis and figures from ATAC-seq data were generated using Pluto (https://pluto.bio). In addition, we also performed analysis as follows: first, we estimated the raw paired-end sequences quality, then the paired-end sequencing reads were processed using cutadapt to remove sequencing adapters. Adapter sequences were trimmed from both read pairs, and reads shorter than 20 bp after trimming were discarded. All ambiguous nucleotides were also discarded. If at least one read passed filtering criteria, the sequences were included for analysis. FastQC was used to assess read quality, then sequences were aligned to the human reference genome (Hg38) using Bowtie2. Aligned BAM files were sorted and indexed using SAMtools. Reads were filtered to retain only properly paired fragments with mapping quality ≥ 30. We removed unmapped, mate unmapped, secondary, QC-fail, and supplementary sequences. Peaks were called using MACS3 with default parameters. Genome-wide signal tracks were generated in bigwig format using deepTools for visualization in genome browsers. Signal tracks were then visualized using Integrative Genomics Viewer to assess chromatin accessibility. Peak quality was assessed by fragment size distribution, TSS enrichment, and fraction of reads in peaks (FRiP) (**Table S5**). High-quality libraries exhibited characteristic nucleosome-free and mono-/di- nucleosome fragment patterns and strong enrichment at TSS regions. High-confidence peaks were retained for downstream analyses. Read counts within peaks were quantified to generate a peak-by-sample matrix for subsequent differential accessibility analysis. Differential chromatin accessibility analysis was performed using DESeq2. Overlaps among peak sets across

samples and transcription factors were computed using BEDTools, and the resulting intersections were visualized through heatmaps and signal profile plots. Peaks were further annotated to nearby genes using ChIPseeker.

CUT&RUN

Cleavage under targets and release using nuclease (CUT&RUN) was performed following EpiCypher's CUTANA protocol and reagents. Briefly, cells were harvested with 0.05% trypsin, and $5x10^5$ cells were used per reaction. Cells were bound to concanavalin-conjugated paramagnetic beads. Antibodies (**Table S8**) were added (0.5µg per reaction) and allowed to bind overnight. Stop buffer was added with E. coli DNA spike-in. pAG-MNase was added and activated with $CaCl_2$. DNA was purified using SPRIselect, and concentration checked by Qubit fluorometer per manufacturer's instructions. Library prep was performed using 5 ng input DNA and CUTANA CUT&RUN Library Prep Kit. Libraries were pooled (each sample diluted to 4nM) and sequenced at Genewiz using an Illumina NovaSeq.

Sequencing reads were assessed for quality using FastQC and MultiQC. Adapter trimming was performed, and reads with low-quality bases and adapters were trimmed by cutadapt. Base quality trimming was applied to both ends of each read using a Phred score cutoff of 20. Reads shorter than 20 base pairs after trimming were discarded. Ambiguous nucleotides (Ns) at read ends were removed. Only the resulting high-quality paired reads were written to compressed FASTQ files, and detailed trimming statistics were recorded for each sample in log files. Trimmed sequences were again assessed for quality using FastQC and MultiQC and aligned to Hg38 human reference genome using Bowtie2 with default settings to generate BAM files. In parallel, reads were also aligned to the E.coli K12 genome to quantify spike-in normalization factors. We kept

only aligned reads with high-confidence fragments, and those mapping to mitochondrial DNA or with a mapping quality score below 30 were excluded. We compared each aligned sequence to corresponding IgG control sequence to get Spike-in factors. The aligned BAM files were downsampled according to calculated Spike-in factor. BigWig files were generated using deepTools bamCoverage for visualization of normalized signal tracks across the genome. Track visualizations for each sample were examined using Integrative Genomics Viewer. Peak calling was performed using MACS3 and SEACR, using IgG as the background control. HOMER was used to identify enriched motifs. Peak overlaps between samples and different transcription factors were calculated using Bedtools to intersect and plot heatmaps or profiles to visualize. Peaks were annotated to nearby genes using ChIPseeker or ChIPpeakAnno R packages.

RNA-sequencing

Total RNA was extracted using Qiagen RNeasy Mini kit. RNA quality and integrity was assessed using an Agilent bioanalyzer, and samples with RNA integrity number (RIN) ≥7 were used. RNA-seq libraries were prepared at sequenced at Genewiz. The quality of raw sequencing reads was assessed using FastQC. Summary reports across all samples were generated with MultiQC to provide an overview of sequencing quality metrics. Adapter sequences and low-quality bases were removed using trimming tools such as Cutadapt or Trimmomatic, with reads shorter than 20 base pairs discarded to ensure reliable downstream alignment. Post-trimming quality was reassessed using FastQC to confirm improvement in read quality and successful adapter removal. The trimmed reads are aligned to reference genome (Hg38) or transcriptome using aligners such as Hisat2. After alignment, we marked duplicates, performed quality control on the alignments, and quantified gene expression at both gene and transcript levels. Quality of

alignments were performed using tools like Qualimap to assess metrics such as mapping rate, coverage, and gene body coverage. Gene-level expression quantification was performed using featureCounts from the Subread package, generating count matrices for downstream analyses. Differential expression analysis was using DESeq2 to identify differentially expressed genes between sample groups. Comparisons were performed between *ATRX*-deficient and control samples within each cell line and combined. Genes with an adjusted p-value < 0.05 and absolute log2 fold change ≥ 1 were considered significantly differentially expressed. Gene set enrichment analysis was performed using the GSEA software (Broad Institute) to identify pathways associated with *ATRX* deficiency. Enrichment was assessed against curated gene sets from the Molecular Signatures Database (MSigDB), including Hallmark, GO, and KEGG collections. Pathways with a false discovery rate (FDR) q-value < 0.25 were considered significantly enriched.