

The ETS transcription factor ETV6 constrains the transcriptional activity of EWS–FLI to promote Ewing sarcoma

Received: 14 November 2021

Accepted: 24 November 2022

Published online: 19 January 2023

 Check for updates

Diana Y. Lu^{1,2,3}, Jana M. Ellegast^{2,3,10}, Kenneth N. Ross^{2,3,10}, Clare F. Malone^{2,3}, Shan Lin^{2,3}, Nathaniel W. Mabe ^{2,3}, Neekesh V. Dharia ^{2,3}, Ashleigh Meyer^{2,3}, Amy Conway^{2,3}, Angela H. Su^{2,3}, Julia Selich-Anderson⁴, Cenny Taslim⁴, Andrea K. Byrum⁴, Bo Kyung A. Seong^{2,3}, Biniam Adane^{2,3}, Nathanael S. Gray ^{5,6}, Miguel N. Rivera^{3,7,8}, Stephen L. Lessnick ^{4,9} & Kimberly Stegmaier ^{2,3} ✉

Transcription factors (TFs) are frequently mutated in cancer. Paediatric cancers exhibit few mutations genome-wide but frequently harbour sentinel mutations that affect TFs, which provides a context to precisely study the transcriptional circuits that support mutant TF-driven oncogenesis. A broadly relevant mechanism that has garnered intense focus involves the ability of mutant TFs to hijack wild-type lineage-specific TFs in self-reinforcing transcriptional circuits. However, it is not known whether this specific type of circuitry is equally crucial in all mutant TF-driven cancers. Here we describe an alternative yet central transcriptional mechanism that promotes Ewing sarcoma, wherein constraint, rather than reinforcement, of the activity of the fusion TF EWS–FLI supports cancer growth. We discover that ETV6 is a crucial TF dependency that is specific to this disease because it, counter-intuitively, represses the transcriptional output of EWS–FLI. This work discovers a previously undescribed transcriptional mechanism that promotes cancer.

As fundamental drivers of cell-type-specific identity and function, aberrant TFs represent an important class of genetic dependencies across distinct cancer types¹. Paediatric cancers exhibit few mutations genome-wide, but typically harbour sentinel mutations that alter TF proteins^{2–5}. Mutant TFs can hijack wild-type lineage-specific TFs into self-reinforcing, feed-forward core regulatory circuits (CRCs)^{6–14}. For example, MYCN in MYCN-amplified neuroblastoma and the PAX3–FOXO1 and PAX7–FOXO1 fusion proteins in alveolar rhabdomyosarcoma promote tumour growth by hijacking tumour-type-specific CRC

TFs^{15–19}. It is not known, however, to what extent distinct cancer types harbouring mutant TFs rely on this specific type of circuitry.

Ewing sarcoma, the second most common paediatric bone cancer, is defined by pathognomonic chromosomal translocations that fuse a gene member of the FET family of RNA-binding proteins with members of the ETS family of TFs^{20,21}. In 85–90% of cases, a translocation fuses the *EWSR1* and *FLI1* genes to encode the EWS–FLI fusion protein. EWS–FLI proteins exhibit the neomorphic ability to pioneer de novo enhancers at microsatellites that contain tandem ETS 5′-GGAA-3′ motif repeats^{22–31}

¹Harvard/MIT MD-PhD Program, Harvard Medical School, Boston, MA, USA. ²Department of Pediatric Oncology, Dana-Farber Boston Children's Cancer and Blood Disorders Center, Boston, MA, USA. ³The Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁴Center for Childhood Cancer and Blood Diseases, Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH, USA. ⁵Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA. ⁶Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. ⁷Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ⁸Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA. ⁹Division of Pediatric Hematology, Oncology and BMT, The Ohio State University College of Medicine, Columbus, OH, USA. ¹⁰These authors contributed equally: Jana M. Ellegast, Kenneth N. Ross. ✉e-mail: Kimberly_Stegmaier@dfci.harvard.edu

via multimerization and recruitment of chromatin-modifying complexes, which in turn lead to an altered gene expression programme^{20,32}.

Efforts to establish key dependencies in Ewing sarcoma have prioritized the identification of specific gene targets of EWS–FLI. Studies have described cell-type-specific TFs that are activated by, and cooperate with, EWS–FLI to reinforce oncogenic programmes^{23,24,32–39}, including in CRCs⁴⁰. Unbiased and systematic approaches are needed, however, to reveal crucial disease mechanisms specific to Ewing sarcoma.

Here we describe the results of a genome-scale CRISPR–Cas9 screen revealing that the wild-type ETS TF ETV6 (ETV6; also known as TEL) is a crucial Ewing-sarcoma-selective TF dependency. We validate this dependency in vitro and in vivo. In contrast to selective TF dependencies that reinforce the oncogenic programmes of mutant TFs in other cancer types, the repressive activity of ETV6 constrains EWS–FLI gene activation at 5′-GGAA-3′ repeat enhancers to promote Ewing sarcoma growth. We therefore discover a previously undescribed mechanism promoting cancer: competition on chromatin between an oncogenic fusion TF and a ‘restraining’ inhibitory TF.

Results

ETV6 is a selective TF dependency in Ewing sarcoma

We recently reported a genome-scale CRISPR–Cas9 loss-of-function screen performed in paediatric cancer cell lines (Paediatric Cancer DepMap) that identified transcriptional activators as a strongly enriched class of selective dependencies (genetic vulnerabilities unique to a specific cancer type) in several paediatric cancer subtypes⁴¹. Among the most highly selective dependencies were activating CRC TFs in neuroblastoma and rhabdomyosarcoma (Fig. 1a and Supplementary Tables 1–3). TFs involved in CRCs co-opted by MYCN in neuroblastoma (PHOX2B, HAND2, ISL1 and GATA3)^{15,16} and the PAX3–FOXO1 and PAX7–FOXO1 fusion proteins in alveolar rhabdomyosarcoma (MYOD1, MYOG, SOX8 and MYCN)^{17,18} constituted the strongest dependencies specific to each tumour type. By contrast, selective TF gene dependencies in Ewing sarcoma were transcriptional repressors, including the known dependencies *BCL11B* and *ZEB2* (refs. 36, 37), which are activated by EWS–FLI, and the previously uncharacterized dependency *ETV6* (Fig. 1a and Extended Data Fig. 1a). These TFs also scored in independent screens (Extended Data Fig. 1b). *ETV6* is not recurrently mutated in Ewing sarcoma^{42–44}, and *ETV6* dependency was not associated with a specific EWS–ETS fusion (that is, EWS–FLI or EWS–ERG) (Extended Data Fig. 1c). *ETV6* was not defined as a gene regulated by EWS–FLI in 18 of 19 gene sets (Supplementary Table 4) and did not exhibit a unique pattern of expression in cell lines (Fig. 1a) or in primary tumours⁴⁵ (Extended Data Fig. 1d and Supplementary Table 5) of Ewing sarcoma.

Fig. 1 | The repressive ETS TF ETV6 is a selective dependency in Ewing sarcoma cells.

a, Scatter plot depicting 18,333 genes interrogated in the DepMap CRISPR–Cas9 screen. $-\log_{10}(q \text{ value})$ of enrichment (x axis) measures the specificity of dependency for each tumour type⁴¹ (Ewing, $n = 14$; neuroblastoma, $n = 20$; rhabdomyosarcoma, $n = 11$). $-\log_{10}(q \text{ value})$ of enrichment (y axis) measures the specificity of gene expression for each tumour type⁷⁴ (Ewing, $n = 20$; neuroblastoma, $n = 28$; rhabdomyosarcoma, $n = 18$). TF genes⁸⁴ are red and labelled if $x > 8$ (except *ZEB2* = 3.28). Dashed lines show $-\log_{10}(0.05)$. **b**, Top: line graph depicting mean cell viability \pm s.e.m. in A673 Ewing sarcoma cells transduced with CRISPR–Cas9 constructs targeting *ETV6* (sgETV6-1 to sgETV6-4) or control single guide RNAs (sgChr2.2 cutting; sgLacZ non-cutting) ($n = 8$ biological replicates, two-way analysis of variance (ANOVA), Dunnett’s multiple comparisons, P adjusted < 0.0001). Represents two independent experiments. Bottom: western blot shows ETV6 with GAPDH loading control. **c**, Bar plot showing mean \pm s.e.m. number of A673 cell colonies in methylcellulose. ETV6 loss reduced the colony number (one-way ANOVA, $n = 3$ biological replicates, Sidak’s multiple comparisons, P adjusted < 0.0001). Represents two independent experiments. **d**, Schematic of the dTAG approach used to study ETV6. Ub, ubiquitin; VHL, von Hippel–Lindau. **e**, Western blot demonstrating ETV6–FKBP12^{F36V}–HA protein degradation and endogenous

By contrast, CRC TFs in neuroblastoma and rhabdomyosarcoma exhibited tumour-type-specific expression (Fig. 1a), as did *BCL11B* and *ZEB2* in Ewing sarcoma⁴⁵ (Extended Data Fig. 1e and Supplementary Table 5).

We validated an *ETV6* dependency in three cell lines of Ewing sarcoma, A673, EW8 and TC32, via CRISPR–Cas9 disruption. Loss of *ETV6* reduced cell growth in vitro (Fig. 1b and Extended Data Fig. 1f) and reduced anchorage-independent growth in methylcellulose (Fig. 1c and Extended Data Fig. 1g). We established a biochemical dTAG approach^{46,47} to perturb ETV6 abundance with precise temporal control and without eliciting acute DNA damage. FKBP12^{F36V}-tagged proteins can be acutely degraded following exposure to the dTAG small-molecule dTAG^V-1, which recruits the von Hippel–Lindau E3 ligase to ubiquitinate FKBP12^{F36V} (ref. 46). In the Ewing sarcoma cell lines A673 and EW8, we exogenously expressed ETV6 carboxy-terminally tagged with FKBP12^{F36V} and a human influenza haemagglutinin (HA) epitope (Fig. 1d). Simultaneously, we knocked out endogenous *ETV6* such that FKBP12^{F36V}-tagged ETV6 constituted the dominant form of ETV6 protein. ETV6–FKBP12^{F36V} degradation reduced anchorage-independent growth (Fig. 1e,f and Extended Data Fig. 1h). Degradation of ETV6 (Fig. 1g and Extended Data Fig. 2a) as well as CRISPR–Cas9-mediated knock-out of endogenous *ETV6* in parental A673 cells (Extended Data Fig. 2b) led to G1/G0 cell cycle arrest but did not induce apoptosis (Extended Data Fig. 2c and Supplementary Fig. 1).

In vivo, CRISPR–Cas9-mediated knockout of *ETV6* reduced the growth of subcutaneous TC32 tumours (Extended Data Fig. 2d). Using an orthotopic-like mouse model, in which A673 Ewing sarcoma cells implanted intramuscularly in the hindlimb are capable of metastasis⁴⁸, we observed that ETV6 loss reduced primary tumour growth (Fig. 1h). ETV6 loss reduced metastasis to liver tissues (Fig. 1i, left), and lung tissues displayed the same trend in one out of two *ETV6* knockout conditions (Fig. 1i, right).

Next we asked whether the DNA-binding domain (DBD) of ETV6 was crucial to its function. We knocked out endogenous *ETV6* and exogenously expressed wild-type *ETV6* or mutant *ETV6* bearing a C-terminal DBD deletion, which precluded ETV6 binding to chromatin and partially impeded its nuclear localization. This result is consistent with the report that the nuclear localization signal of ETV6 protein lies in its C terminus⁴⁹ (Extended Data Fig. 2e). Whereas wild-type *ETV6* expression rescued *ETV6* knockout, expression of the mutant *ETV6* did not (Extended Data Fig. 2f), which suggests that the specific activity of ETV6 on chromatin is crucial to its function in Ewing sarcoma.

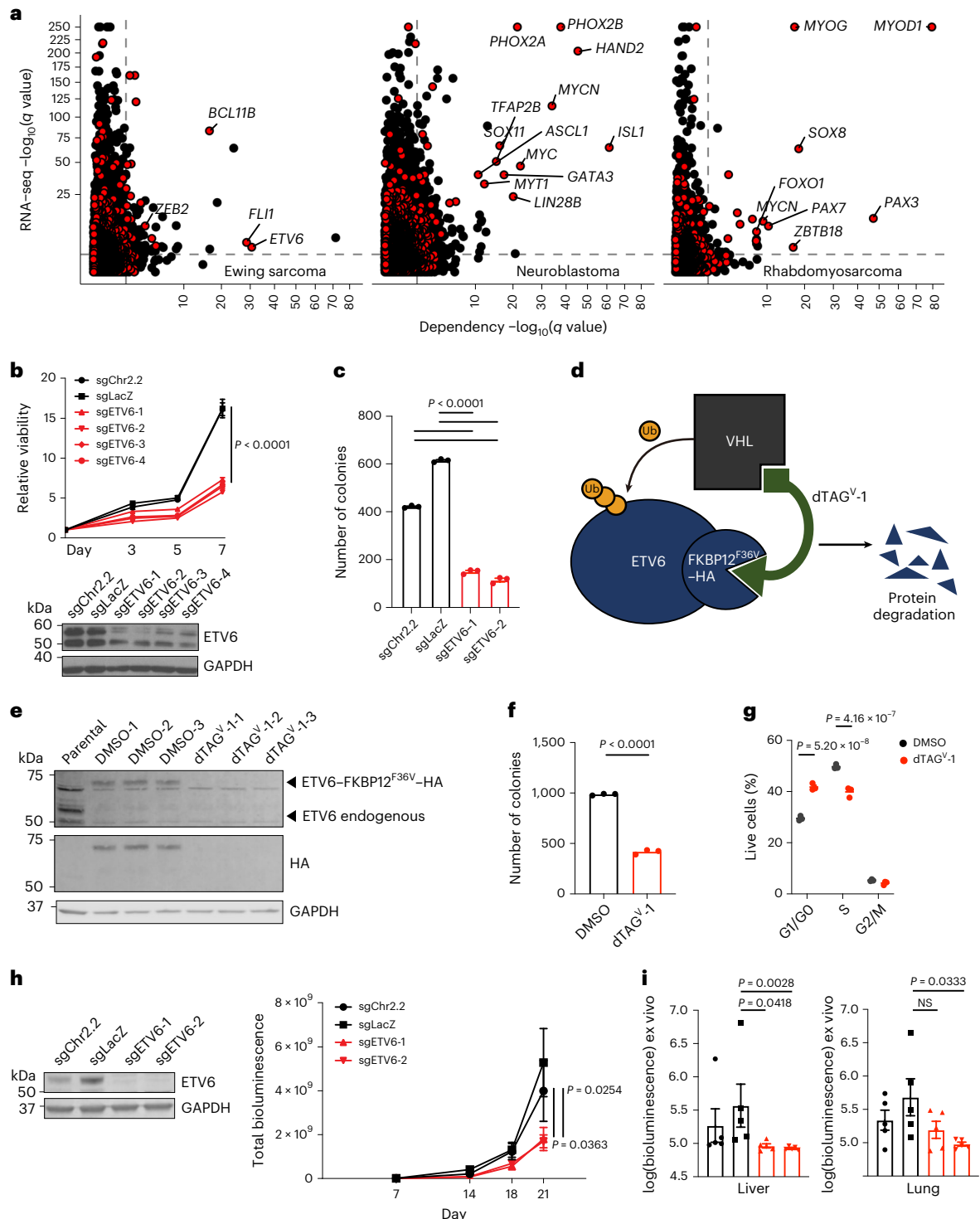
ETV6 and EWS–FLI co-occupy loci genome-wide

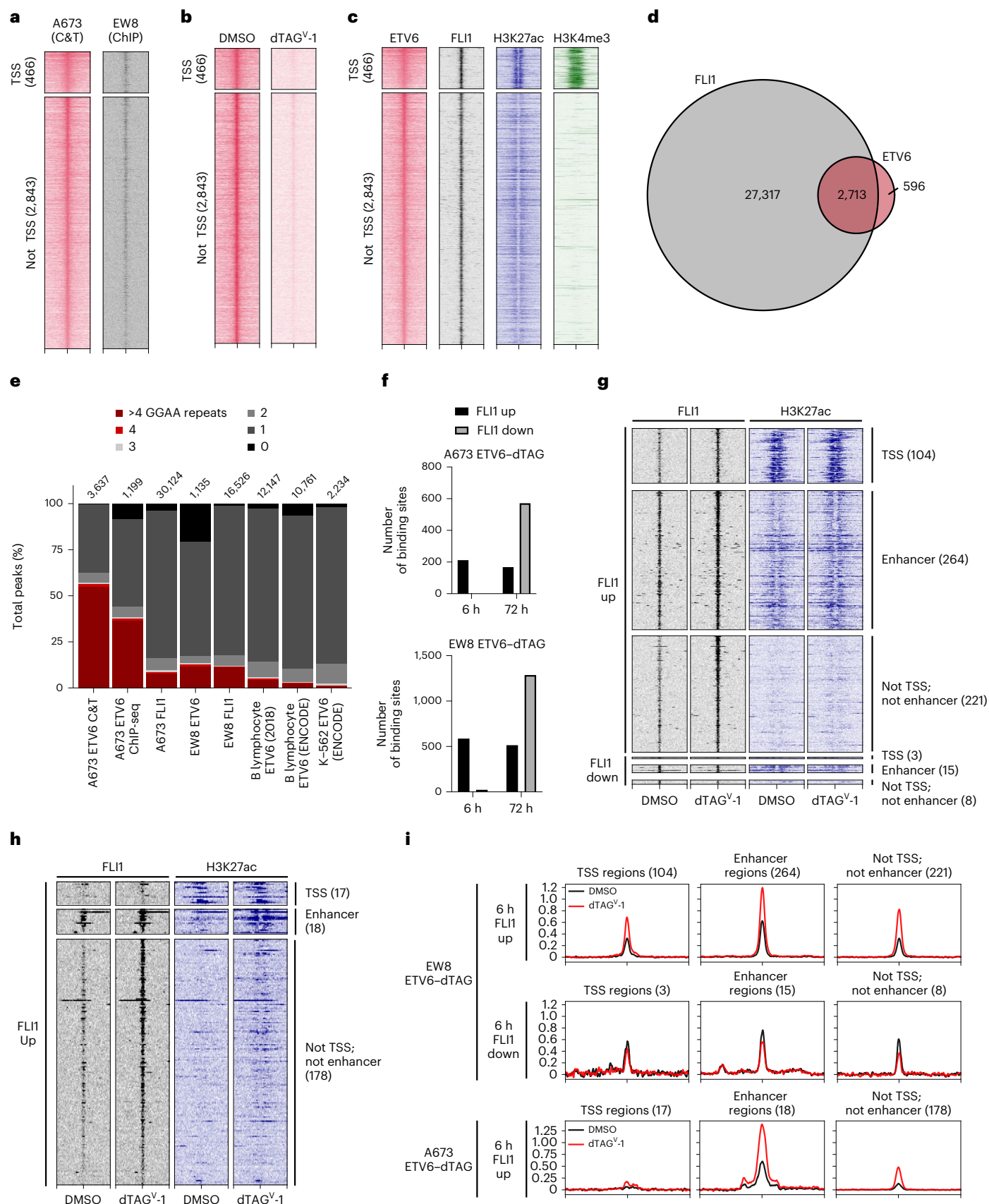
ETV6 and EWS–FLI harbour the ETS family DBD, which recognizes consensus 5′-GGA(A/T)-3′ motifs. We therefore asked whether they

ETV6 knockout in A673 ETV6–dTAG cells treated with dTAG^V-1 or DMSO for 6 h. Parental A673 lysates are shown on the left. Represents one experiment. **f**, Bar plot showing mean \pm s.e.m. number of A673 ETV6–dTAG cell colonies in methylcellulose ($n = 3$ biological replicates, two-tailed t -test, $P < 0.0001$). Represents two independent experiments. **g**, Cell cycle analysis of A673 ETV6–dTAG cells treated for 72 h with DMSO or dTAG^V-1 ($n = 3$ biological replicates, two-tailed t -test, Sidak’s multiple comparisons; G1/G0 phase, P adjusted = 5.20×10^{-8} ; S phase, P adjusted = 4.16×10^{-7}). Represents two independent experiments. **h**, Left: western blot of A673 cells implanted intramuscularly. Right: mean total body bioluminescence (\pm s.e.m.) ($n = 5$ mice per condition, biological replicates). ETV6 loss reduced tumour growth compared with sgChr2.2 (two-way ANOVA, Tukey’s multiple comparisons; sgETV6-1, P adjusted = 0.0363; sgETV6-2, P adjusted = 0.0254). **i**, Mean \pm s.e.m. log(bioluminescence) measurements of ex vivo resected liver and lung ($n = 5$). sgLacZ livers exhibited greater bioluminescence than sgETV6-1 (Kruskal–Wallis test, Dunn’s multiple comparisons; P adjusted = 0.0418) or sgETV6-2 livers (P adjusted = 0.0028). sgLacZ lungs exhibited greater bioluminescence than sgETV6-2 (P adjusted = 0.0333) but not sgETV6-1 lungs (not significant (NS), P adjusted = 0.2763). Key is the same as **h**.

co-localized on chromatin. We profiled endogenous ETV6 binding sites in parental A673 cells using cleavage under targets and tagmentation (CUT&Tag)⁵⁰ and profiled ETV6–FKBP12^{F36V}–HA binding sites in ETV6–dTAG cells using anti-HA chromatin immunoprecipitation with sequencing (ChIP-seq). These analyses defined a consensus list of ETV6-binding sites (Extended Data Fig. 3a and Fig. 2a). dTAG^{V-1} treatment reduced ETV6 abundance on chromatin in both dTAG models (Fig. 2b and Extended Data Fig. 3b). In parental Ewing sarcoma cells, we performed histone H3 lysine 27 acetylation (H3K27ac) ChIP-seq and analysed public histone H3 lysine 4 trimethylation (H3K4me3) ChIP-seq data²⁶ to annotate ETV6-binding sites. The results showed

that these sites occurred at active promoters and enhancers (Fig. 2c and Extended Data Fig. 3c). We performed ChIP-seq for EWS–FLI1 in A673 and EW8 parental cells by immunoprecipitating the C-terminal FLI1 domain. This is an accepted approach to identify EWS–FLI1-binding sites because wild-type *FLI1* typically is not expressed in Ewing sarcoma cells^{39,43}. EWS–FLI1 bound ubiquitously at ETV6-binding sites in both models (Fig. 2c and Extended Data Fig. 3c), although co-occupied binding sites constituted only a small proportion of total EWS–FLI1-binding sites (Fig. 2d and Extended Data Fig. 3d). EWS–FLI1 pioneers closed chromatin at GGAA repeat microsatellites²⁷, including at repeats of four or more²⁶. ETV6 localized at these longer consecutive GGAA repeats





at a higher frequency in Ewing sarcoma than in B lymphocytes or in K-562 leukaemia cells^{51,52}, which express *ETV6* ($P < 2.2 \times 10^{-16}$) (Fig. 2e and Extended Data Fig. 3e).

Loss of ETV6 increases EWS–FLI1 occupancy

We next asked whether loss of ETV6 alters EWS–FLI1 chromatin occupancy. We degraded ETV6 and profiled EWS–FLI1 binding by CHIP-seq

Fig. 2 | Acute loss of ETV6 leads to increased EWS–FLI binding. **a–c**, Heatmaps showing 3-kb windows centred at 3,309 consensus ETV6-binding sites, subplotted by overlap within 2.5 kb of transcription start sites (TSSs) and peaks ranked by maximum height. **a**, Left: CUT&Tag (C&T) of endogenous ETV6 in A673 parental cells. Right: anti-HA ChIP-seq of ETV6–FKBP12^{36V}–HA in EW8 ETV6–dTAG cells. **b**, Anti-HA CUT&Tag in A673 ETV6–dTAG cells treated with DMSO or dTAG^V-1 for 24 h. **c**, Left to right: endogenous ETV6 CUT&Tag, EWS–FLI ChIP-seq, H3K27ac ChIP-seq and published H3K4me3 ChIP-seq²⁶ in A673 cells. **d**, Venn diagram showing genomic locations of ETV6 consensus binding sites versus 30,030 EWS–FLI binding sites in A673 cells. **e**, Stacked column plot showing varying lengths of tandem 5′-GGAA-3′ motif repeats occurring at binding sites detected by (left to right): (1) endogenous ETV6 CUT&Tag in parental A673 cells; (2) ETV6–FKBP12^{36V}–HA ChIP-seq in A673 ETV6–dTAG cells; (3) EWS–FLI ChIP-seq in parental A673 cells; (4) ETV6–FKBP12^{36V}–HA ChIP-seq in EW8 ETV6–dTAG cells; (5) EWS–FLI ChIP-seq in parental EW8 cells; (6) and (7) endogenous ETV6 ChIP-seq in GM12878 B lymphocyte lymphoblastoid cells^{31,52}; and (8) endogenous

ETV6 ChIP-seq in K-562 chronic myelogenous leukaemia cells³². Number of binding sites in each dataset is shown. ETV6 bound to a higher percentage of >4 GGAA repeats in Ewing sarcoma compared to B lymphocyte ETV6 (2018) (A673 ETV6 C&T, $P < 1 \times 10^{-300}$; A673 ETV6 ChIP-seq, $P = 5.41 \times 10^{-214}$; and EW8 ETV6, $P = 1.34 \times 10^{-18}$; Fisher's exact tests). **f**, Bar plots showing the number of genomic regions exhibiting significantly altered EWS–FLI binding at 6 or 72 h following ETV6 degradation identified by CSAW (CSAW using the edgeR generalized linear model; $P < 0.05$)⁸⁹. FLII up sites exhibited increased EWS–FLI binding. FLII down sites exhibited decreased EWS–FLI binding. **g, h**, Heatmaps of EWS–FLI and H3K27ac ChIP-seq performed in EW8 ETV6–dTAG (**g**) and A673 ETV6–dTAG cells (**h**) at 6 h following DMSO or dTAG^V-1 treatment. Loci exhibiting significantly altered EWS–FLI binding are subplotted by direction of change (up or down) and overlap with TSS, enhancer or neither. Enhancer locations were defined using H3K27ac ChIP-seq in parental EW8 (**g**) and A673 (**h**) cells. **i**, Metaplots of FLII binding at regions shown in **g** and **h**.

at 6 and 72 h. At 6 h, in both dTAG models, significant alterations in EWS–FLI occupancy primarily constituted increases in binding (Fig. 2f). At 72 h, alterations were more dynamic, exhibiting both increases and decreases (Fig. 2f). We categorized loci by whether they gained or lost EWS–FLI binding at 6 h and whether they occurred at transcription start sites (TSSs) or at H3K27ac-defined enhancers (Fig. 2g, h). Regions that lost EWS–FLI binding did not change to as great a degree as regions that gained binding (Fig. 2i). Thus, the loss of ETV6 led acutely and predominantly to increased EWS–FLI binding, which provides support for the hypothesis that these TFs compete for binding. Additionally, ChIP-seq of H3K27ac at 6 h in both models (Fig. 2g, h) demonstrated a modest increase in H3K27ac abundance at enhancer regions that gained EWS–FLI binding (Extended Data Fig. 3f).

Differential EWS–FLI binding was highly dynamic at tandem 5′-GGAA-3′ repeats (Extended Data Fig. 3g). Notably, genomic regions that gained EWS–FLI binding were more likely to contain shorter tandem repeats of 2, 3 or 4 motifs compared with regions that lost EWS–FLI binding ($P < P = 6.974 \times 10^{-15}$). Consistent differences were not observed for single GGAA motifs or >4 GGAA repeats.

ETV6 is a transcriptional repressor in Ewing sarcoma

We next characterized genes regulated by ETV6, a reported transcriptional repressor^{53–57}. We performed RNA sequencing (RNA-seq) in both dTAG models at 6, 24 and 72 h following treatment with dimethylsulfoxide (DMSO) or dTAG^V-1 (Fig. 1e and Extended Data Fig. 4a). Globally, the expression profiles of each of the engineered dTAG cell lines approximated that of their corresponding parental cell lines (Extended Data Fig. 4b). At 6 h, the majority of differentially expressed genes were upregulated, which suggests that ETV6 acts predominantly as a transcriptional repressor in Ewing sarcoma (Fig. 3a). Strongly ETV6-repressed genes increased in expression over time following

ETV6 degradation (Fig. 3b). We observed concordance in regulated genes between dTAG models (Extended Data Fig. 4c) and identified a common set of 85 ETV6-repressed genes (Fig. 3c and Supplementary Table 6). We performed RNA-seq on parental A673 cells transduced with ETV6 CRISPR knockout (Fig. 3d and Extended Data Fig. 4d). The results showed that most of the 85 genes were also repressed by endogenous levels of wild-type ETV6 ($P = 2.66 \times 10^{-20}$). Consistent with the localization of ETV6 at active promoters and enhancers, ETV6-repressed genes were expressed and not completely silenced (Extended Data Fig. 4e). Additionally, ETV6-binding sites were enriched in ETV6-regulated genes (Fig. 3e and Extended Data Fig. 4f).

ETV6 is a master TF implicated in the normal development of neural and mesenchymal lineages^{58,59}. Developmental lineage-specific gene sets were enriched in ETV6-repressed genes (Fig. 3f and Supplementary Tables 6–11) and in ETV6-activated genes (Extended Data Fig. 4g and Supplementary Tables 12–17). ETV6-repressed genes, but not activated genes, were strongly enriched for genes regulated by histone deacetylases (HDACs), which may reflect the ability of ETV6 to recruit HDACs^{54,60–63}. We also observed strong enrichment of EWS–FLI-regulated genes in ETV6-regulated genes (Fig. 3f), consistent with their co-localization on chromatin.

Loss of ETV6 alters gene expression and the chromatin state

We next sought to associate locus-specific alterations in chromatin with differential gene expression after ETV6 degradation. Alterations in EWS–FLI binding at individual loci at 6 h were sustained at 72 h, and these alterations correlated between dTAG models (Fig. 4a). Consistently, the greatest degree of differential EWS–FLI binding was exhibited by loci that gained EWS–FLI binding; loci that lost EWS–FLI binding exhibited smaller changes in magnitude (Fig. 4a). This pattern was paralleled by alterations in H3K27ac abundance at EWS–FLI-binding

Fig. 3 | ETV6 is primarily a transcriptional repressor in Ewing sarcoma.

a, RNA-seq volcano plot in A673 ETV6–dTAG and EW8 ETV6–dTAG cells treated with DMSO or dTAG^V-1 for 6 and 72 h ($n = 3$ biological replicates). Red indicates genes upregulated in dTAG^V-1-treated cells (DESeq2 Wald test Benjamini–Hochberg P adjusted < 0.05 ; A673: $n = 423$ at 6 h and 2,554 at 72 h; EW8: $n = 123$ at 6 h and 1,614 at 72 h). Blue indicates genes downregulated in dTAG^V-1-treated cells (P adjusted < 0.05 ; A673: $n = 221$ at 6 h and 2,556 at 72 h; EW8: $n = 67$ at 6 h and 1,208 at 72 h). **b**, Row-normalized RNA-seq \log_2 (transcripts per million (TPM) + 1) heatmap of the 25 most differentially repressed genes in A673 ETV6–dTAG cells, identified from 6-h data, ranked by P value (DESeq2 P adjusted < 0.05 and \log_2 (fold change) > 1.5). **c**, Top left: Venn diagram of genes identified as ETV6-repressed in A673 ETV6–dTAG (423 genes) and EW8 ETV6–dTAG cells (123 genes) at 6 h (P adjusted < 0.05), identifying 85 common ETV6-repressed genes. Bottom: row-normalized RNA-seq \log_2 (TPM + 1) heatmap of 85 common ETV6-repressed genes in A673 ETV6–dTAG and EW8 ETV6–dTAG cells. **d**, Row-normalized \log_2 (TPM + 1) RNA-seq heatmap of 85 ETV6-repressed genes, ranked as shown in **c**, in parental A673 cells transduced with CRISPR–Cas9 vectors, identifying

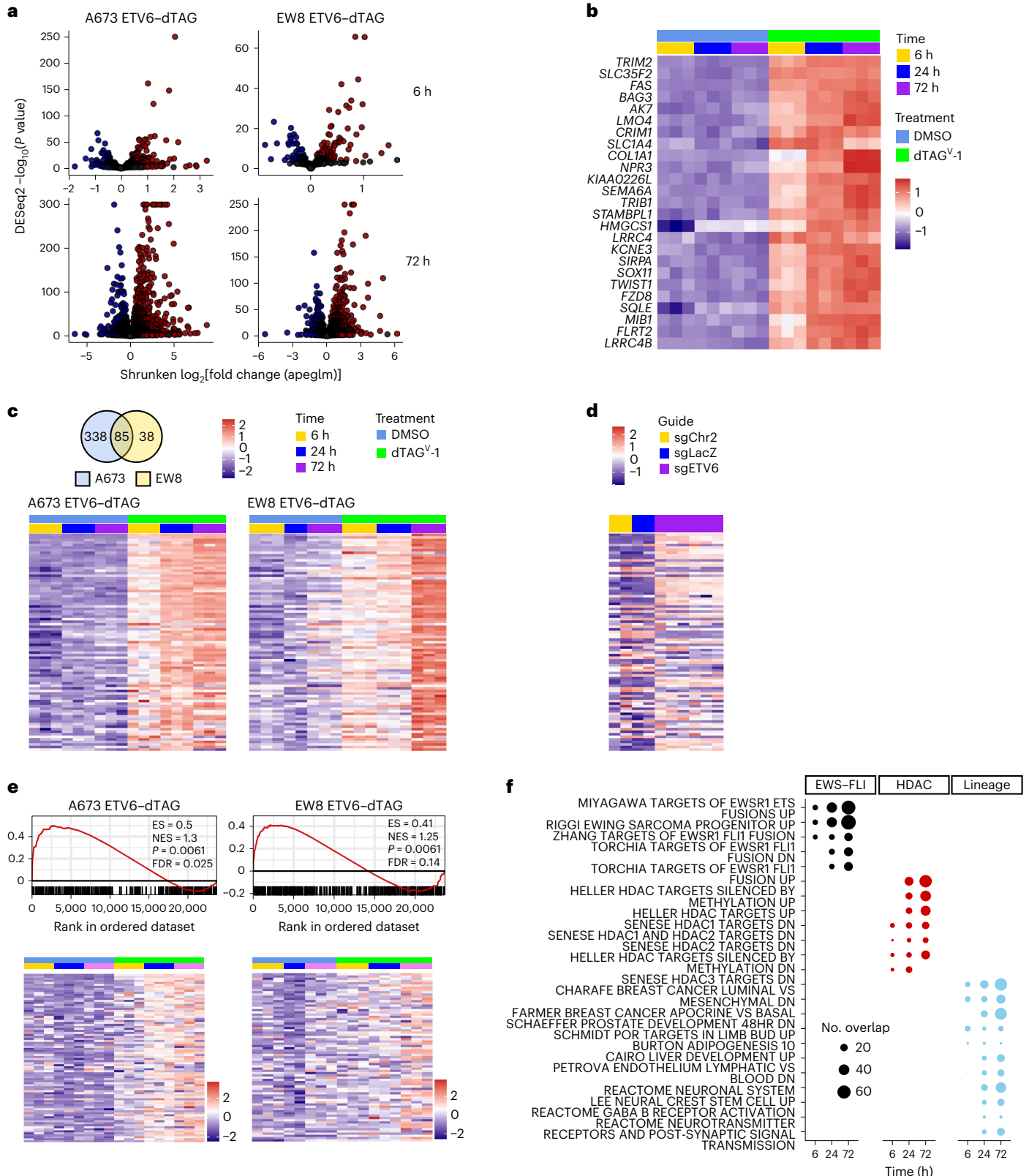
53 ETV6-repressed genes using this approach (one-sided hypergeometric test, $P = 2.66 \times 10^{-20}$). **e**, Top: gene set enrichment analysis (GSEA) plots of ETV6-bound genes enriched in ETV6-regulated genes in A673 ETV6–dTAG and EW8 ETV6–dTAG cells. ETV6-bound genes were defined by CUT&Tag and ChIP-seq in A673 cells and by ChIP-seq in EW8 cells. ETV6-regulated genes were defined by RNA-seq at 24 h. Bottom: RNA-seq heatmaps of ETV6-repressed core enrichment genes. ES, enrichment score; FDR, false discovery rate; NES, normalized enrichment score. Colour coding for time and treatment is the same as in **c**. **f**, Combined enrichment plot of MSigDB c2 pathways significantly enriched in ETV6-repressed genes defined by RNA-seq at 24 h common to both models (hypergeometric enrichment test, $P < 0.05$). Gene sets are ranked by significance. Dot size indicates the number of genes in the overlap between the gene set and common ETV6-repressed genes at 6, 24 and 72 h (85, 251 and 832 genes, respectively). Missing dots indicate non-significance. 'EWS–FLI', 'HDAC' and 'Lineage' gene sets characterize genes regulated by EWS–FLI, genes regulated by histone deacetylase enzymes, and genes underlying tissue-specific development or function, respectively.

sites (Fig. 4b and Extended Data Fig. 5a) and by alterations in chromatin accessibility (Fig. 4c). We assigned EWS–FLI-binding sites to nearby genes and examined their expression following ETV6 loss (Fig. 4d). Genes that gained or lost EWS–FLI binding exhibited significantly increased or decreased expression, respectively ($P < 1 \times 10^{-10}$), with genes in the former category exhibiting the greatest degree of change on average (Fig. 4d and Extended Data Fig. 5b). Thus, the most profound

consequences of ETV6 loss are increased EWS–FLI binding, opening of chromatin and increased gene expression.

Increased EWS–FLI occupancy upregulates gene expression

We knocked out *EWS–FLI* in A673 ETV6–dTAG cells to evaluate whether loss of EWS–FLI rescues gene expression changes with ETV6 loss. *FAS*, *ACTA2*, *TRIB1* and *SEMASB* were identified as ETV6-repressed genes that



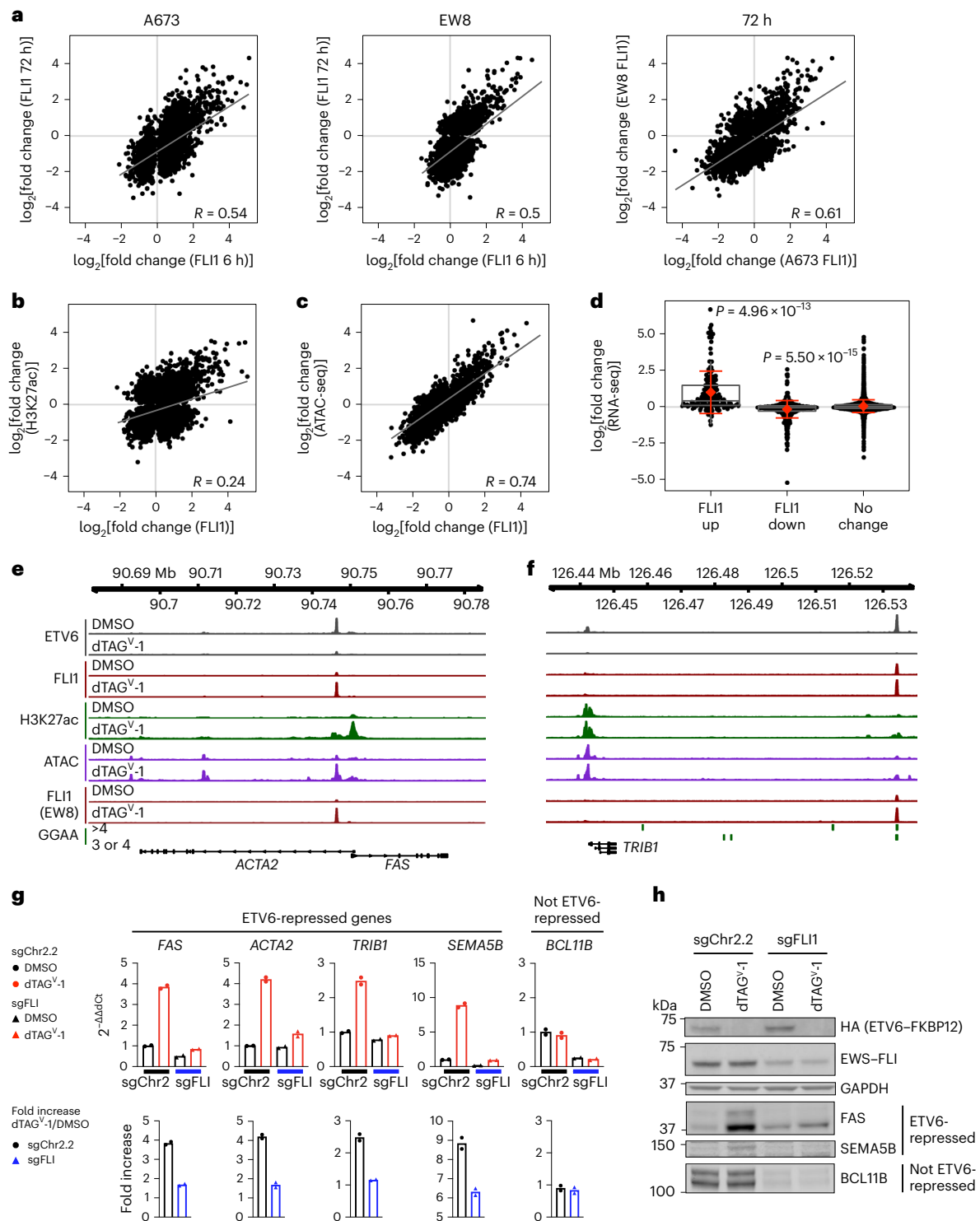


Fig. 4 | ETV6 constrains EWS-FLI1-induced gene expression. **a–c**, Lines indicate Pearson correlation; Pearson correlation value (R) is shown. **a**, Left and middle: scatter plots of \log_2 (fold change) in EWS-FLI1 binding at 6 and 72 h following DMSO or dTAG^{V-1} treatment in A673 and EW8 ETV6-dTAG cells ($n = 2$ biological replicates). Right: scatter plot comparing models. **b**, Scatter plot comparing \log_2 (fold change) in EWS-FLI1 binding to H3K27ac abundance detected by 6 h ChIP-seq in A673 ETV6-dTAG cells ($n = 2$ biological replicates). **c**, Scatter plot comparing \log_2 (fold change) in EWS-FLI1 binding in assay of transposase accessible chromatin sequencing (ATAC-seq) experiments ($n = 3$ biological replicates) at 72 h in A673 ETV6-dTAG cells. **d**, Plot comparing genes mapped from altered EWS-FLI1 binding sites (CSAW, $n = 2$ biological replicates) to \log_2 (fold change) in expression measured by RNA-seq in A673 ETV6-dTAG cells ($n = 3$ biological replicates) at 72 h. Grey boxes indicate median and first and third quartiles. Red diamond and error bars indicate mean expression \pm s.d.

(FLI1 up, $n = 148$, mean = 0.98; FLI1 down, $n = 542$, mean = -0.19; no change, $n = 4,585$, mean = 0.028). P values calculated using paired t -test, Benjamini-Hochberg corrections. **e, f**, Gviz-generated views of the *FAS-ACTA2* (**e**) and *TRIB1* (**f**) loci. ETV6 tracks show CUT&Tag of ETV6-FKBP12^{F36V}-HA in A673 ETV6-dTAG cells at 24 h. FLI1 tracks show ChIP-seq for EWS-FLI1 performed at 6 h, H3K27ac tracks show ChIP-seq for H3K27ac at 6 h and ATAC tracks show ATAC-seq at 72 h. FLI1 (EW8) tracks show ChIP-seq for EWS-FLI1 at 6 h in EW8 ETV6-dTAG cells. GGAA tracks indicate locations of tandem GGAA motif repeats. **g**, Top: bar plots showing qPCR in A673 ETV6-dTAG cells transduced with CRISPR-Cas9 constructs targeting control (sgChr2.2) or *EWS-FLI1* (sgFLI) and treated for 24 h with DMSO (black) or dTAG^{V-1} (red). Bars indicate mean $2^{-\Delta\Delta C_t}$ of $n = 2$ biological replicates, each representing the mean of technical triplicates. **h**, Western blot A673 ETV6-dTAG cells shown in **g** treated with DMSO or dTAG^{V-1} for 96 h.

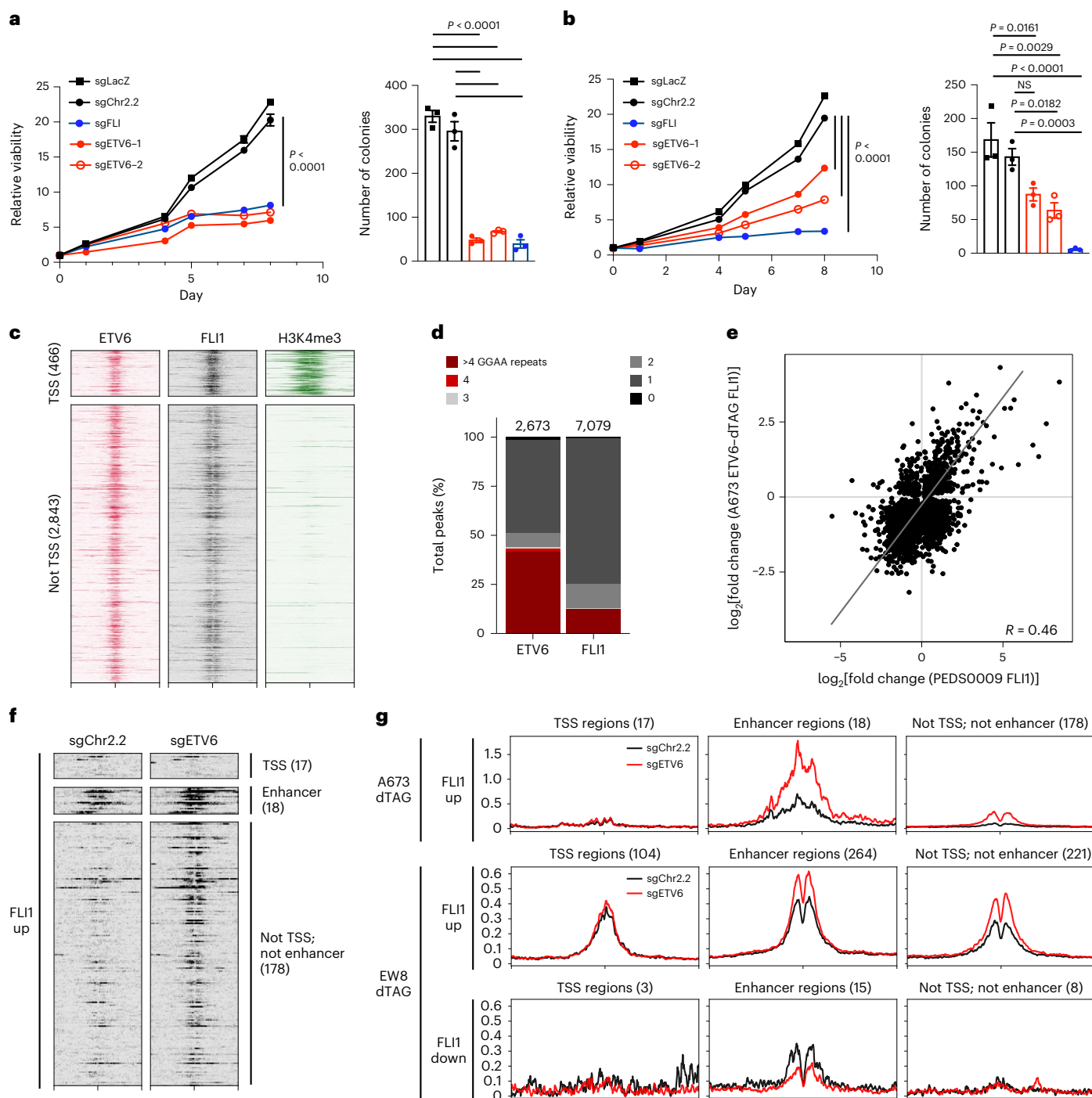


Fig. 5 | ETV6 competes with EWS-FLI1 for binding in clinically relevant Ewing sarcoma models. a, b, Cell growth in the newly derived Ewing sarcoma cell lines PEDS0009 (**a**) and PEDS0010 (**b**) following CRISPR-Cas9 knockout of *ETV6* (red) or *EWS-FLI1* (blue) compared to sgChr2.2 and sgLacZ controls (black). Line graphs show mean cell viability \pm s.e.m. ($n = 6$ biological replicates); knockout of *ETV6* and *EWS-FLI1* reduced viability in both lines compared to sgChr2.2 (two-way ANOVA, Tukey's multiple comparisons, P adjusted < 0.0001). Bar plots show mean cell colony number \pm s.e.m. ($n = 3$ biological replicates) in methylcellulose; *ETV6* and *EWS-FLI1* knockout reduced colony number in PEDS0009 cells (one-way ANOVA, Tukey's multiple comparisons, P adjusted < 0.0001 for all comparisons indicated) and PEDS0010 cells (sgLacZ versus sgETV6-1, P adjusted = 0.0161, sgETV6-2, P adjusted = 0.0029, sgFLI, $P < 0.0001$; sgChr2.2 versus sgETV6-1, NS P adjusted = 0.1120, sgETV6-2, P adjusted = 0.0182, sgFLI, P adjusted = 0.0003). **c,** Heatmaps showing 3-kb windows centred at 3,309 consensus *ETV6* binding sites, subplotted by overlap within 2.5 kb of TSSs. PEDS0009 cells were transduced with

sgChr2.2 control CRISPR-Cas9 constructs and profiled by CUT&Tag to detect endogenous *ETV6* (left) and CUT and release using nuclease (CUT&RUN) to detect *EWS-FLI1* (middle) and the histone mark H3K4me3 (right). **d,** Stacked column plot showing varying lengths of tandem 5'-GGAA-3' motif repeats occurring at *ETV6* (left) and *EWS-FLI1* (right) binding sites in PEDS0009 cells. **e,** Scatter plots of \log_2 (fold change) in *EWS-FLI1* binding in A673 *ETV6*-dTAG cells following 72 h of *ETV6* degradation (y axis) compared to CRISPR-Cas9-transduced PEDS0009 cells with knockout of *ETV6* (x axis). Line indicates Pearson correlation; Pearson correlation value (R) is shown. **f,** Heatmaps of *FLI1* CUT&RUN performed in control or *ETV6* knockout PEDS0009 cells. Loci shown are defined in Fig. 2h as regions that exhibited increased *EWS-FLI1* binding following *ETV6* loss in A673 *ETV6*-dTAG cells. **g,** Metaplots of *FLI1* binding in control or *ETV6* knockout PEDS0009 cells at genomic regions shown in **f** (top). Metaplots of *FLI1* binding at loci defined in Fig. 2g as regions that exhibited increased *EWS-FLI1* binding 72 h following *ETV6* degradation in EW8 *ETV6*-dTAG cells.

exhibit increased EWS–FLI binding, H3K27ac and chromatin accessibility at ETV6-vacated sites, some of which occurred at GGAA repeats (Fig. 4e,f and Extended Data Fig. 5c). We compared these genes to *BCL11B* because it is activated by EWS–FLI but is not repressed by ETV6 and does not exhibit altered EWS–FLI binding acutely following ETV6 loss (Extended Data Fig. 5d). Quantitative PCR (qPCR) demonstrated that degradation of ETV6 led to upregulation of ETV6-repressed genes but not *BCL11B* (Fig. 4g, top plots). *EWS–FLI* knockout significantly reduced the upregulation of ETV6-repressed genes (Fig. 4g, bottom plots). Immunoblotting validated that the attenuation of mRNA upregulation also affected protein levels (Fig. 4h). Thus, ETV6 and EWS–FLI antagonistically regulate *FAS*, *ACTA2*, *TRIB1* and *SEMASB* expression.

ETV6 functions similarly in clinically relevant Ewing sarcoma models

Well-established cancer cell lines may use distinct biological mechanisms to that of primary tumour cells. We therefore tested the relevance of our findings from cell lines in two newly derived Ewing sarcoma cell lines: CCLF_PEDS_0009_T (PEDS0009) and CCLF_PEDS_0010_T (PEDS0010)⁶⁴. *ETV6* knockout impaired cell growth in vitro and colony formation in methylcellulose (Fig. 5a,b and Extended Data Fig. 6a,b). Additionally, we tested cells from a minimally passaged Ewing sarcoma patient-derived xenograft (PDX): ES-PDX-001 (refs. 65, 66). Again, knockout of *ETV6* impaired cell growth in vitro (Extended Data Fig. 6c). In PEDS0009 cells, we observed ETV6 and EWS–FLI binding at previously defined EWS–FLI consensus binding sites (Fig. 5c). Concordant with our cell line data, ETV6 bound to GGAA microsatellites (Fig. 5d), and ETV6 loss resulted in increased EWS–FLI binding at the same loci that exhibited increased EWS–FLI occupancy in cell lines (Fig. 5e–g and Extended Data Fig. 6d,e). Genomic regions that gained EWS–FLI binding were more likely to contain shorter GGAA repeats of 2, 3 or 4 compared with regions that lost EWS–FLI (Extended Data Fig. 6f) ($P = 5.186 \times 10^{-11}$). These observations in minimally passaged cells were concordant with the data from well-established cell lines.

ETV6 and EWS–FLI antagonism at *SOX11* is functional

We next asked whether the antagonistic relationship between EWS–FLI and ETV6 is responsible for the dependency of Ewing sarcoma cells on ETV6. Almost half of the gene sets enriched in ETV6-repressed genes were related to developmental pathways (Extended Data Fig. 7a and Supplementary Table 18), and 46 of these included *SOX11* (Supplementary Table 19). *SOX11* expression exerts context-dependent effects on cancer cell survival, growth and metastasis^{67,68}. *SOX11* acts as an oncogene in mantle cell lymphoma⁶⁹ and promotes metastasis in breast cancer^{70,71}. Conversely, it also reduces proliferation and metastasis in prostate cancer⁷² and induces differentiation of glioma cells⁷³. In Ewing sarcoma cells, the exogenous expression of *SOX11* impaired cell growth,

whereas the expression of a DBD-deleted mutant did not (Extended Data Fig. 7b). These results provide support for a tumour-suppressive role for *SOX11* activity.

We observed differential EWS–FLI binding at a distal enhancer that mapped to *SOX11* as the nearest expressed gene (Fig. 6a, left). RNA-seq data from the Cancer Cell Line Encyclopedia⁷⁴ show that the neighbouring genes, *SILC1* and *LOC400940*, are not expressed in Ewing sarcoma. This enhancer occurred at tandem GGAA repeats and exhibited increased EWS–FLI binding, H3K27ac abundance and chromatin accessibility following ETV6 loss (Fig. 6a, right). RNA-seq confirmed that *SOX11* is repressed by ETV6-FKBP12^{F36V} in dTAG cells and by endogenous ETV6 in parental A673 cells (Extended Data Fig. 7c). EWS–FLI was required for *SOX11* upregulation after ETV6 loss (Fig. 6b,c). Knockout of *SOX11* in A673 ETV6–dTAG cells (Fig. 6d) rescued the effects of ETV6 degradation (Fig. 6e). Additionally, knockout of *SOX11* in A673 and TC32 cells (Fig. 6f and Extended Data Fig. 7d) rescued *ETV6* knockout (Fig. 6g and Extended Data Fig. 7e,f). In vivo, we observed rescue in TC32 cells grown as subcutaneous tumours in mice (Fig. 6h). These findings support the hypothesis that ETV6 dependency is specific to Ewing sarcoma cells because ETV6 constrains EWS–FLI activation of *SOX11* expression.

Finally, we asked whether co-regulation at *SOX11* by ETV6 and EWS–FLI could be recapitulated with ectopic expression of EWS–FLI. In rhabdomyosarcoma RD cells, we exogenously expressed wild-type EWS–FLI or the R340N DNA-binding mutant of EWS–FLI, which cannot bind to DNA⁷⁵. *SOX11* protein expression was induced by wild-type EWS–FLI but not the mutant (Extended Data Fig. 7g). Knockout of *ETV6* further upregulated *SOX11* abundance in the setting of wild-type EWS–FLI but not in the context of mutant EWS–FLI expression (Extended Data Fig. 7g). These findings demonstrate that the DBD of EWS–FLI is required for its activation of *SOX11* expression, an activity that is repressed by ETV6.

Discussion

In this study, we discovered an oncogenic mechanism underlying the paediatric cancer Ewing sarcoma. We demonstrated that the ETS TF ETV6 is a selective dependency in Ewing sarcoma because it antagonizes the transcriptional activity of EWS–FLI at ETS motifs. To our knowledge, this report constitutes the first description of transcriptional constraint of a fusion TF on chromatin as a crucial driver of tumour growth.

Although previous studies have described specific TFs as dependencies that reinforce the EWS–FLI transcriptional programme in Ewing sarcoma^{32–35,37}, including in CRCs⁴⁰, these targets were not identified in DepMap screening as selective gene dependencies. Instead, our discovery that ETV6 constrains EWS–FLI activity highlights a distinct, but equally central, epigenetic mechanism that drives tumour growth and reveals an unexpected contrast between Ewing sarcoma and other paediatric tumours in which CRCs are functionally dominant.

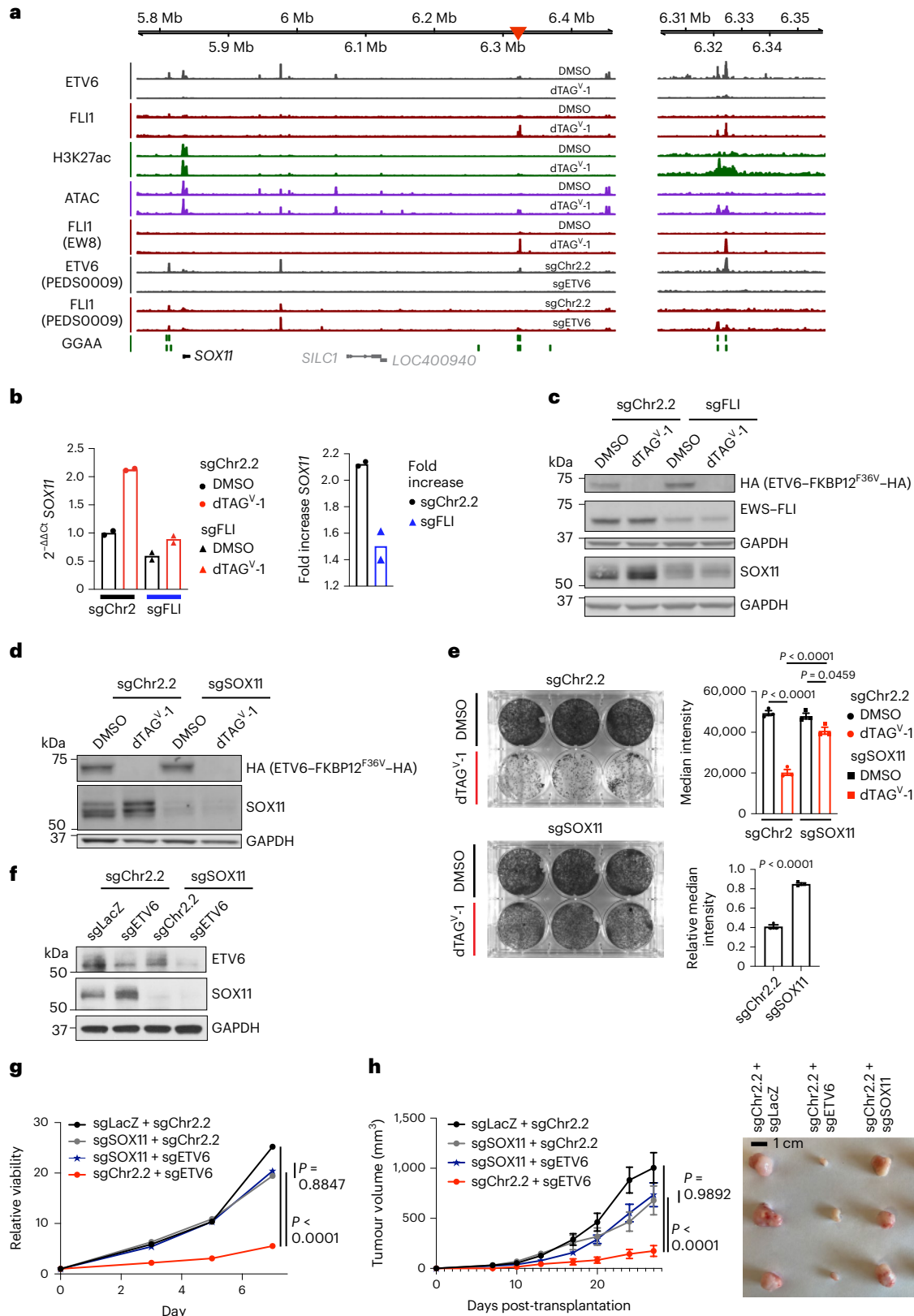
Fig. 6 | Knockout of the ETV6-repressed gene *SOX11* rescues the phenotype of ETV6 loss. **a**, Left: Gviz-generated view of the *SOX11* locus. Top four tracks show data generated in A673 ETV6–dTAG cells: ETV6, 24 h ETV6-FKBP12^{F36V}–HA CUT&Tag; FLII, 6 h EWS–FLI ChIP-seq; H3K27ac, 6 h H3K27ac ChIP-seq; ATAC, 72 h ATAC-seq. FLII (EWS) shows 6 h EWS–FLI ChIP-seq in EWS ETV6–dTAG cells. ETV6 (PEDS0009) and FLII (PEDS0009) show CUT&Tag for ETV6 and EWS–FLI, respectively, in PEDS0009 cells. GGAA shows tandem GGAA motif repeats. The red arrowhead indicates an enhancer region assigned to *SOX11*, the nearest expressed gene; *SILC1* and *LOC400940* are not expressed and labeled in grey⁷⁴. Right: magnified view of the enhancer. **b**, *SOX11* expression by qPCR, as described in Fig. 4g. **c**, Western blot of cells shown in **b** at 96 h. HA, EWS–FLI, GAPDH bands are also shown in Fig. 4h. **d**, Western blot of sgChr2.2-transduced or sgSOX11-transduced A673 ETV6–dTAG cells cultured in DMSO or dTAG^V-1. Represents two independent experiments. **e**, Left: cells in **d** stained with crystal violet. Right, top: bar plots showing mean \pm s.e.m. of median stain intensity per well (one-way ANOVA, $n = 3$ biological replicates, Sidak's multiple comparisons; DMSO versus dTAG^V-1 sgChr2.2, P adjusted < 0.0001, sgSOX11, P adjusted = 0.0459; sgChr2.2

versus sgSOX11 dTAG^V-1, P adjusted < 0.0001). Right, bottom: relative median intensity comparing dTAG^V-1-treated with DMSO-treated wells (two-tailed t -test, $n = 3$, $P < 0.0001$). Represents two independent experiments. **f**, Western blot of TC32 cells transduced with CRISPR–Cas9 constructs in combination. Represents one experiment. **g**, Line graph depicting mean viability in vitro ($n = 6$ biological replicates, s.e.m. bars too small to depict) of cells in **f**. *ETV6* knockout alone (red) reduced viability compared to control (black) (two-way ANOVA, Tukey's multiple comparisons, P adjusted < 0.0001). Simultaneous *ETV6* and *SOX11* knockout (blue star) did not reduce viability compared to *SOX11* knockout alone (grey) (NS, P adjusted = 0.8847) and exhibited greater viability than *ETV6* knockout alone (red) (P adjusted < 0.0001). **h**, Left: Line graph depicting mean subcutaneous tumour volume (mm³) \pm s.e.m. ($n = 6$ tumours, biological replicates) formed by cells shown in **f**. *ETV6* knockout alone reduced tumour volume (two-way ANOVA, Tukey's multiple comparisons, P adjusted < 0.0001). Simultaneous *ETV6* and *SOX11* knockout did not reduce tumour growth (NS, P adjusted = 0.9892) and exhibited greater growth than *ETV6* knockout alone (P adjusted < 0.0001). Right: representative tumours from each condition.

Cancer cells frequently co-opt mechanisms that underlie normal development⁷⁶. The competition between EWS–FLI and ETV6 in Ewing sarcoma bears resemblance to a mechanism of ETS TF competition governing cell-fate decisions in developing *Drosophila*. Pointed, the activating orthologue of human Ets-1, competes for binding at ETS motifs within specific enhancers with Yan, the repressive orthologue

of ETV6, to regulate the expression of key differentiation genes in distinct tissues^{77–81}. Here we described a similar mechanism that has been co-opted in cancer to regulate the transcriptional output of a fusion TF.

The epigenetic activity of ETS TFs other than EWS–FLI may contribute to the phenotype of ETV6 loss. Notably, *ETV7*, the homologue of *ETV6*, is not expressed in Ewing sarcoma cells (Supplementary Table 20),



and we did not observe strong changes in the expression of other ETS TFs with ETV6 loss. The maximum change exhibited by one gene was roughly threefold, and only five genes displayed a significant alteration in expression across the models evaluated. Moreover, none of the genes that displayed a change in expression were scored as dependencies or tumour suppressors in DepMap in Ewing sarcoma.

Although most human TF families contain paralogues that are co-expressed within distinct cell types^{82–84}, an understanding of their interactions at shared motifs is lacking. We began to unravel key *cis* regulatory principles that distinguish the specific functions of ETV6 and EWS–FLI, the antagonism of which on chromatin frequently occurred at shorter 5′-GGAA-3′ repeats. As the pathogenesis of EWS–FLI is typically associated with its activity at longer repeats or true microsatellites, we highlight a previously undescribed *cis* regulatory role for shorter GGAA repeats in this disease, which facilitates ETV6 fine-tuning of EWS–FLI. Even though the reconstitution of EWS–FLI for biochemical assays has been a challenge for the field, future work is needed to delineate the precise GGAA repeat code that determines the activities of each TF. Similarly, ETV6 and wild-type FLI proteins can engage in an inhibitory heterodimer⁸⁵, an interaction mediated by the amino-terminal Pointed (PNT) domain of ETV6, and further studies are needed to determine whether ETV6 and EWS–FLI engage in a protein–protein interaction. Notably, however, our experiments using an ETS DBD-deleted mutant of ETV6, with an intact PNT domain, demonstrated that the DNA-binding activity of ETV6 is crucial to its function in Ewing sarcoma.

ETV6 is a master TF in normal development and is recurrently mutated in cancer. ETV6 mutations include deletions and chromosomal translocations involving 30 distinct gene partners⁵³. Germline and somatic loss-of-function mutations frequently occur in pre-malignant disorders and leukaemias. For example, in B cell acute lymphoblastic leukaemia, ETV6 deletions frequently co-occur with ETV6–RUNX1 rearrangements, which result in biallelic loss of the ETV6 protein^{86,87}. Chromosomal translocations also fuse the N terminus of ETV6 with the tyrosine kinase domain from a number of receptor tyrosine kinases, which facilitate constitutive autophosphorylation and growth signalling⁵³. ETV6, however, has not been reported as recurrently mutated in Ewing sarcoma^{42–44}. Furthermore, ETV6 is not regulated by EWS–FLI (Supplementary Table 4) and does not exhibit a marked pattern of expression specific to this cancer type (Extended Data Fig. 1d). Nonetheless, we discovered its role as a crucial tumour-type-selective dependency in regulating EWS–FLI activity. As such, this report reaffirms the importance of performing unbiased functional screens at scale to reveal oncogenic mechanisms sustained by wild-type proteins.

Our findings suggest that a hallmark of Ewing sarcoma biology may involve the reliance on mechanisms constraining EWS–FLI activity to promote tumour growth. Indeed, we previously described mechanisms mediated by an E3 ligase (TRIM8) and cohesin that restrain EWS–FLI activity^{48,66,88}. Here we discovered a distinct mechanism in support of an EWS–FLI Goldilocks phenomenon⁶⁶ that is operative on chromatin. Future translational efforts could ultimately seek to modulate the activity of this pharmacologically challenging protein, either by decreasing or paradoxically increasing its activity.

In conclusion, we discovered the oncogenic role of TF competition on chromatin between a mutant TF and a wild-type paralogue. Our work contributes to an understanding of the dysregulated epigenetic mechanisms that can promote cancer, raising the possibility that similar mechanisms are relevant in other disease contexts.

Online content

Any methods, additional references, Nature Portfolio 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/s41556-022-01059-8>.

References

1. Tsherniak, A. et al. Defining a cancer dependency map. *Cell* **170**, 564–576.e16 (2017).
2. Filbin, M. & Monje, M. Developmental origins and emerging therapeutic opportunities for childhood cancer. *Nat. Med.* **25**, 367–376 (2019).
3. Panditharatna, E. & Filbin, M. G. The growing role of epigenetics in childhood cancers. *Curr. Opin. Pediatr.* **32**, 67–75 (2020).
4. Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
5. Gröbner, S. N. et al. The landscape of genomic alterations across childhood cancers. *Nature* **555**, 321–327 (2018).
6. Jahangiri, L. et al. Core regulatory circuitries in defining cancer cell identity across the malignant spectrum. *Open Biol.* **10**, 200121 (2020).
7. Saint-André, V. et al. Models of human core transcriptional regulatory circuitries. *Genome Res.* **26**, 385–396 (2016).
8. Ott, C. J. et al. Enhancer architecture and essential core regulatory circuitry of chronic lymphocytic leukemia. *Cancer Cell* **34**, 982–995.e7 (2018).
9. Riddick, G. et al. A core regulatory circuit in glioblastoma stem cells links MAPK activation to a transcriptional program of neural stem cell identity. *Sci. Rep.* **7**, 43605 (2017).
10. Sanda, T. et al. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer Cell* **22**, 209–221 (2012).
11. Kron, K. J. et al. TMPRSS2–ERG fusion co-opts master transcription factors and activates NOTCH signaling in primary prostate cancer. *Nat. Genet.* **49**, 1336–1345 (2017).
12. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional addiction in cancer. *Cell* **168**, 629–643 (2017).
13. Chapuy, B. et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* **24**, 777–790 (2013).
14. Sengupta, S. & George, R. E. Super-enhancer-driven transcriptional dependencies in cancer. *Trends Cancer* **3**, 269–281 (2017).
15. Boeva, V. et al. Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat. Genet.* **49**, 1408–1413 (2017).
16. Durbin, A. D. et al. Selective gene dependencies in MYCN-amplified neuroblastoma include the core transcriptional regulatory circuitry. *Nat. Genet.* **50**, 1240–1246 (2018).
17. Gryder, B. E. et al. Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma. *Nat. Genet.* **51**, 1714–1722 (2019).
18. Gryder, B. E. et al. PAX3–FOXO1 establishes myogenic super enhancers and confers BET bromodomain vulnerability. *Cancer Discov.* **7**, 884–899 (2017).
19. van Groningen, T. et al. Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nat. Genet.* **49**, 1261–1266 (2017).
20. Grünewald, T. G. P. et al. Ewing sarcoma. *Nat. Rev. Dis. Prim.* **4**, 5 (2018).
21. Delattre, O. et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162–165 (1992).
22. Toomey, E. C., Schiffman, J. D. & Lessnick, S. L. Recent advances in the molecular pathogenesis of Ewing’s sarcoma. *Oncogene* **29**, 4504–4516 (2010).
23. Gangwal, K., Close, D., Enriquez, C. A., Hill, C. P. & Lessnick, S. L. Emergent properties of EWS/FLI regulation via GGAA microsatellites in Ewing’s sarcoma. *Genes Cancer* **1**, 177–187 (2010).

24. Gangwal, K. et al. Microsatellites as EWS/FLI response elements in Ewing's sarcoma. *Proc. Natl Acad. Sci. USA* **105**, 10149–10154 (2008).
25. Guillon, N. et al. The oncogenic EWS–FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PLoS ONE* **4**, e4932 (2009).
26. Riggi, N. et al. EWS–FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* **26**, 668–681 (2014).
27. Boulay, G. et al. Cancer-specific retargeting of BAF complexes by a prion-like domain. *Cell* **171**, 163–178.e19 (2017).
28. Boulay, G. et al. Epigenome editing of microsatellite repeats defines tumor-specific enhancer functions and dependencies. *Genes Dev.* **32**, 1008–1019 (2018).
29. Riggi, N., Suva, M. L. & Stamenkovic, I. Ewing's sarcoma origin: from duel to duality. *Expert Rev. Anticancer Ther.* **9**, 1025–1030 (2009).
30. Tanaka, M. et al. Ewing's sarcoma precursors are highly enriched in embryonic osteochondrogenic progenitors. *J. Clin. Invest.* **124**, 3061–3074 (2014).
31. von Levetzow, C. et al. Modeling initiation of Ewing sarcoma in human neural crest cells. *PLoS ONE* **6**, e19305 (2011).
32. Cidre-Aranaz, F. & Alonso, J. EWS/FLI1 target genes and therapeutic opportunities in Ewing sarcoma. *Front. Oncol.* **5**, 162 (2015).
33. Fadul, J. et al. EWS/FLI1 utilizes NKX2-2 to repress mesenchymal features of Ewing sarcoma. *Genes Cancer* **6**, 129–143 (2015).
34. García-Aragoncillo, E. et al. DAX1, a direct target of EWS/FLI1 oncoprotein, is a principal regulator of cell-cycle progression in Ewing's tumor cells. *Oncogene* **27**, 6034–6043 (2008).
35. Kinsey, M., Smith, R., Iyer, A. K., McCabe, E. R. B. & Lessnick, S. L. EWS/FLI and its downstream target NROB1 interact directly to modulate transcription and oncogenesis in Ewing's sarcoma. *Cancer Res.* **69**, 9047–9055 (2009).
36. Wiles, E. T., Bell, R., Thomas, D., Beckerle, M. & Lessnick, S. L. ZEB2 represses the epithelial phenotype and facilitates metastasis in Ewing sarcoma. *Genes Cancer* **4**, 486–500 (2013).
37. Wiles, E. T., Lui-Sargent, B., Bell, R. & Lessnick, S. L. BCL11B is up-regulated by EWS/FLI and contributes to the transformed phenotype in Ewing sarcoma. *PLoS ONE* **8**, e59369 (2013).
38. Charville, G. W. et al. EWSR1 fusion proteins mediate PAX7 expression in Ewing sarcoma. *Mod. Pathol.* **30**, 1312–1320 (2017).
39. Smith, R. et al. Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. *Cancer Cell* **9**, 405–416 (2006).
40. Shi, X. et al. EWS–FLI1 regulates and cooperates with core regulatory circuitry in Ewing sarcoma. *Nucleic Acids Res.* **48**, 11434–11451 (2020).
41. Dharia, N. V. et al. A first-generation pediatric cancer dependency map. *Nat. Genet.* **53**, 529–538 (2021).
42. Brohl, A. S. et al. The genomic landscape of the Ewing sarcoma family of tumors reveals recurrent STAG2 mutation. *PLoS Genet.* **10**, e1004475 (2014).
43. Crompton, B. D. et al. The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov.* **4**, 1326–1341 (2014).
44. Tirode, F. et al. Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer Discov.* **4**, 1342–1353 (2014).
45. Vaske, O. M. Comparative tumor RNA sequencing analysis for difficult-to-treat pediatric and young adult patients with cancer. *JAMA Netw. Open* **2**, e1913968 (2019).
46. Nabet, B. et al. Rapid and direct control of target protein levels with VHL-recruiting dTAG molecules. *Nat. Commun.* **11**, 4687 (2020).
47. Nabet, B. et al. The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* **14**, 431–441 (2018).
48. Adane, B. et al. STAG2 loss rewires oncogenic and developmental programs to promote metastasis in Ewing sarcoma. *Cancer Cell* **39**, 827–844.e10 (2021).
49. Park, H., Seo, Y., Kim, J. I., Kim, W.-J. & Choe, S. Y. Identification of the nuclear localization motif in the ETV6 (TEL) protein. *Cancer Genet. Cytogenet.* **167**, 117–121 (2006).
50. Kaya-Okur, H. S. et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).
51. Venkataraman, A. et al. A toolbox of immunoprecipitation-grade monoclonal antibodies to human transcription factors. *Nat. Methods* **15**, 330–338 (2018).
52. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**, 636–640 (2004).
53. Rasighaemi, P. & Ward, A. C. ETV6 and ETV7: siblings in hematopoiesis and its disruption in disease. *Crit. Rev. Oncol. Hematol.* **116**, 106–115 (2017).
54. Chakrabarti, S. R. & Nucifora, G. The leukemia-associated gene TEL encodes a transcription repressor which associates with SMRT and mSin3A. *Biochem. Biophys. Res. Commun.* **264**, 871–877 (1999).
55. Irvin, B. J. et al. TEL, a putative tumor suppressor, induces apoptosis and represses transcription of Bcl-XL. *J. Biol. Chem.* **278**, 46378–46386 (2003).
56. Kim, C. A. et al. Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J.* **20**, 4173–4182 (2001).
57. Lopez, R. G. et al. TEL is a sequence-specific transcriptional repressor. *J. Biol. Chem.* **274**, 30132–30138 (1999).
58. Wang, L. C. et al. Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *EMBO J.* **16**, 4374–4383 (1997).
59. Wang, L. C. et al. The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. *Genes Dev.* **12**, 2392–2402 (1998).
60. Wang, L. & Hiebert, S. W. TEL contacts multiple co-repressors and specifically associates with histone deacetylase-3. *Oncogene* **20**, 3716–3725 (2001).
61. Guidez, F. et al. Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia-associated TEL-AML1 oncoprotein. *Blood* **96**, 2557–2561 (2000).
62. Fenrick, R. et al. TEL, a putative tumor suppressor, modulates cell growth and cell morphology of ras-transformed cells while repressing the transcription of stromelysin-1. *Mol. Cell. Biol.* **20**, 5828–5839 (2000).
63. Fisher, M. H. ETV6 germline mutations cause HDAC3/NCOR2 mislocalization and upregulation of interferon response genes. *JCI Insight* **5**, e140332 (2020).
64. Guenther, L. M. et al. A combination CDK4/6 and IGF1R inhibitor strategy for Ewing sarcoma. *Clin. Cancer Res.* **25**, 1343–1357 (2019).
65. García-Domínguez, D. J. et al. The combination of epigenetic drugs SAHA and HCI-2509 synergistically inhibits EWS-FLI1 and tumor growth in Ewing sarcoma. *Oncotarget* **9**, 31397–31410 (2018).
66. Seong, B. K. A. et al. TRIM8 modulates the EWS/FLI1 oncoprotein to promote survival in Ewing sarcoma. *Cancer Cell* **39**, 1262–1278.e7 (2021).
67. Tsang, S. M., Oliemuller, E. & Howard, B. A. Regulatory roles for SOX11 in development, stem cells and cancer. *Semin. Cancer Biol.* **67**, 3–11 (2020).
68. Yang, Z. et al. SOX11: friend or foe in tumor prevention and carcinogenesis? *Ther. Adv. Med. Oncol.* **11**, 1758835919853449 (2019).
69. Beekman, R., Amador, V. & Campo, E. SOX11, a key oncogenic factor in mantle cell lymphoma. *Curr. Opin. Hematol.* **25**, 299–306 (2018).

70. Shepherd, J. H. et al. The SOX11 transcription factor is a critical regulator of basal-like breast cancer growth, invasion, and basal-like gene expression. *Oncotarget* **7**, 13106–13121 (2016).
71. Zvelebil, M. et al. Embryonic mammary signature subsets are activated in *Brca1*^{-/-} and basal-like breast cancers. *Breast Cancer Res.* **15**, 25 (2013).
72. Yao, Z. et al. The role of tumor suppressor gene SOX11 in prostate cancer. *Tumour Biol.* **36**, 6133–6138 (2015).
73. Hide, T. et al. Sox11 prevents tumorigenesis of glioma-initiating cells by inducing neuronal differentiation. *Cancer Res.* **69**, 7953–7959 (2009).
74. Barretina, J. et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
75. Liang, H. et al. Solution structure of the ets domain of Fli-1 when bound to DNA. *Nat. Struct. Biol.* **1**, 871–876 (1994).
76. Boyer, L. A. et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
77. Boisclair Lachance, J. F., Webber, J. L., Hong, L., Dinner, A. R. & Rebay, I. Cooperative recruitment of Yan via a high-affinity ETS supersite organizes repression to confer specificity and robustness to cardiac cell fate specification. *Genes Dev.* **32**, 389–401 (2018).
78. Webber, J. L., Zhang, J., Massey, A., Sanchez-Luege, N. & Rebay, I. Collaborative repressive action of the antagonistic ETS transcription factors Pointed and Yan fine-tunes gene expression to confer robustness in *Drosophila*. *Development* <https://doi.org/10.1242/dev.165985> (2018).
79. Webber, J. L., Zhang, J., Mitchell-Dick, A. & Rebay, I. 3D chromatin interactions organize Yan chromatin occupancy and repression at the even-skipped locus. *Genes Dev.* **27**, 2293–2298 (2013).
80. Rebay, I. & Rubin, G. M. Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* **81**, 857–866 (1995).
81. O’Neill, E. M., Rebay, I., Tjian, R. & Rubin, G. M. The activities of two Ets-related transcription factors required for *Drosophila* eye development are modulated by the Ras/MAPK pathway. *Cell* **78**, 137–147 (1994).
82. Zhang, Y., Ho, T. D., Buchler, N. E. & Gordân, R. Competition for DNA binding between paralogous transcription factors determines their genomic occupancy and regulatory functions. *Genome Res.* **31**, 1216–1229 (2021).
83. Sizemore, G. M., Pitarresi, J. R., Balakrishnan, S. & Ostrowski, M. C. The ETS family of oncogenic transcription factors in solid tumours. *Nat. Rev. Cancer* **17**, 337–351 (2017).
84. Lambert, S. A. et al. The human transcription factors. *Cell* **172**, 650–665 (2018).
85. Kwiatkowski, B. A. et al. The ets family member Tel binds to the Fli-1 oncoprotein and inhibits its transcriptional activity. *J. Biol. Chem.* **273**, 17525–17530 (1998).
86. Raynaud, S. et al. The 12;21 translocation involving TEL and deletion of the other TEL allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood* **87**, 2891–2899 (1996).
87. Golub, T. R. et al. Fusion of the *TEL* gene on 12p13 to the *AML1* gene on 21q22 in acute lymphoblastic leukemia. *Proc. Natl Acad. Sci. USA* **92**, 4917–4921 (1995).
88. Surdez, D. et al. *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and *EWSR1-FLI1* activity in Ewing sarcoma. *Cancer Cell* **39**, 810–826. e9 (2021).
89. Lun, A. T. & Smyth, G. K. *csaw*: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.* **44**, e45 (2016).

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

Methods

Our research complied with all ethical guidelines determined by the Dana-Farber Cancer Institute Institutional Animal Care and Use Committee under Animal Welfare Assurance number D16-00010 (A3023-01). No human studies were performed. In Extended Data Fig. 6c, a minimally passaged cell line previously derived in our laboratory⁶⁶ from a previously characterized Ewing sarcoma PDX⁶⁵ (HSJD-ES-PDX-001) was studied. As such, this experiment was performed *in vitro* and did not involve the use of animals. As previously described⁶⁵, this PDX originated from a biopsy in a 21.7-year-old patient whose sex was not reported⁶⁵. It was collected with informed consent without compensation under an Institutional Review Board-approved protocol at Sant Joan de Déu Hospital (HSJD, Barcelona, Spain), protocol number HSJD135/11 (ref. 65).

CRISPR–Cas9 screen dependency analysis

All genome-scale dependency data are available at the DepMap portal website: <https://depmap.org>. DepMap AVANA 21Q1 dependency data were used (18,333 genes in 808 cell lines, https://figshare.com/articles/dataset/public_21q1/13681534). Twelve cell lines were not included in the analyses: four cell lines are classified as engineered lines; the origin of one cell line, CHLA57, is unknown as it is incorrectly identified as Ewing sarcoma; seven cell lines are listed as commonly misidentified cell lines in the ICLAC Register of Misidentified Cell Lines (<https://iclac.org/databases/cross-contaminations/>). Therefore, dependency data for 796 cell lines were examined. CERES gene effect scores were calculated as previously described^{41,90}. A lower CERES gene effect score indicates an increased likelihood that a specific gene is required for viability in that cell line. A CERES score of 0 indicates that gene deletion exhibited no effect on growth, whereas a score of –1 is comparable with the median of all commonly essential genes, that is, genes that were essential for growth in nearly every cell line across the entire screen. Tumour-type-enriched ‘selective’ dependencies were determined by performing a two-class comparison between the gene effect scores for cell lines of each tumour type (in-group) and the remainder of all other cell lines in the screen (out-group) for a specific gene as previously described⁴¹. In brief, effect size was calculated as the difference in the mean gene effect dependency score in the in-group compared with that in the out-group. In addition to two-sided *P* values, one-sided *P* values were generated to test whether the in-group exhibited, on average, greater or lesser dependency on a specific gene than the out-group. All *P* values were corrected for multiple hypothesis testing using the Benjamini–Hochberg correction and reported as *q* values. Tumour-type-enriched dependencies were identified in each tumour type as those with a *q* value of <0.05 and with a negative effect size (the mean of dependency gene effect score was more negative in the in-group than in the out-group). The same analyses were performed on the genome-scale CRISPR–Cas9 screens using the Broad Institute GeCKO library (18,478 genes in 43 cell lines, https://figshare.com/articles/dataset/DepMap_GeCKO_19Q1/7668407) as well as the Sanger library (17,799 genes in 318 cell lines, https://figshare.com/articles/dataset/Project_SCORE_processed_with_CERES/9116732/1).

Cancer cell line and primary tumour gene expression

RNA-seq gene expression data from the Cancer Cell Line Encyclopedia⁷⁴ were downloaded (19,177 genes in 1,376 cell lines) from the 21Q1 DepMap portal website (<https://depmap.org>). Tumour-type-enriched expression for each gene was calculated by performing a two-class comparison between the log₂(transcripts per million (TPM) + 1) gene expression for cell lines of each tumour type (in-group) and the remainder of all other cell lines profiled (out-group). All *P* values were corrected for multiple hypothesis testing using the Benjamini–Hochberg correction and reported as *q* values. RNA-seq gene expression data for primary tumours were downloaded from the Treehouse Childhood Cancer Initiative⁴⁵ (UCSC Genomics Institute, <https://treehousegenomics.soe.ucsc.edu/public-data>).

Cell samples and culture

All cell lines were genotyped by short tandem repeat analysis and tested for *Mycoplasma*. Whole-exome sequencing and RNA-seq were performed to validate cell line identity⁴³. The A673 cell line was purchased from the American Type Culture Collection (ATCC, CRL-1598). EW8 (originally derived by P. Houghton⁹¹) and TC32 (originally derived by T. Triche⁹²) cell lines were obtained from the Golub Lab. A673 and EW8 cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) (Thermo Fisher Scientific, MT10013CM), supplemented with 10% FBS (Sigma-Aldrich, F2442) and 1% penicillin–streptomycin (Life Technologies, 15140163). TC32 cells were grown in Roswell Park Memorial Institute (RPMI)-1640 medium (Thermo Fisher Scientific, MT10040CM), supplemented with 10% FBS and 1% penicillin–streptomycin. The PEDS0009 and PEDS0010 cell lines were obtained from the Cancer Cell Line Factory (Broad Institute) and were derived as previously described⁶⁴. They were cultured in RPMI-1640 supplemented with 10% FBS and 1% penicillin–streptomycin. The Ewing sarcoma PDX (HSJD-ES-PDX-001) was provided by J. Mora (HSJD)⁶⁵. To generate the minimally passaged cell line (ES-PDX-001), PDX tumours were processed as previously described⁶⁶. The RD cell line (ATCC, CRL-7731) was a gift from the DepMap group at the Broad Institute. RD cells were cultured in RPMI-1640 supplemented with 10% FBS.

CRISPR–Cas9 genome editing of Ewing sarcoma cell samples

The lentiCRISPRv2 plasmid backbone⁹³ (Addgene, 52961) encoding the Cas9 nuclease was digested with the restriction endonuclease Bsmbl (Esp3I) (Thermo Fisher Scientific, FERFD0454) and gel extracted (Qiagen, 28704). Synthetic oligonucleotides containing gene-targeting single guide RNA (sgRNA) sequences and adapter sequences (provided below) were ligated into the restriction digest site. Oligonucleotides were purchased from Integrated DNA Technologies (IDT), annealed and end-phosphorylated using T4 polynucleotide kinase (New England Biolabs, M0201S). Ligated vectors were transformed into One Shot Stbl3 *Escherichia coli* (Life Technologies, C737303), shaken at 37 °C for 1 h and grown overnight on 100 µg ml^{–1} ampicillin Luria broth plates (Teknova, L1004). Picked colonies were grown for 8–16 h in 100 µg ml^{–1} carbenicillin Luria broth starter cultures. Plasmids were DNA-extracted (Qiagen, 27104) and submitted for Sanger sequencing validation by Genewiz. Validated clones were cultured overnight, and plasmids were extracted (Qiagen, 12963).

CRISPR–Cas9 constructs were packaged into lentiviral particles. HEK293T cells were seeded at 70–80% confluence in 10 cm plates and co-transfected with 9 µg of lentiCRISPRv2-sgRNA construct plasmid, 0.9 µg pVSVg plasmid (Addgene 8454) and 9 µg pPAX2 plasmid (Addgene 19319) using Lipofectamine 3000 (Life Technologies, L3000015). Sixteen hours after transfection, medium was supplemented with 30% FBS. The following day, virus-containing medium was collected, 0.45 µm sterile-filtered and stored at –80 °C.

One million Ewing sarcoma cells were seeded per well of a 6-well plate and spin-infected using 2 ml of virus and 8 µg ml^{–1} polybrene (Santa Cruz Biotechnology, SC-134220) at 37 °C at 1,190 r.c.f. for 30 min. The following day, fresh medium containing 1 µg ml^{–1} puromycin (InvivoGen, ant-pr-1) was added. Cells were selected for at least 48 h. In experiments requiring knockout of two genes, the cells were co-transduced with constructs encoding two distinct sgRNAs, each conferring resistance to either puromycin or blasticidin. Cells were selected with 1 µg ml^{–1} puromycin and 5 µg ml^{–1} blasticidin (Life Technologies, A1113903) for at least 5 days. Separate samples of non-infected cells treated with drug were used to confirm cell death.

All sgRNA sequences used in the Broad Institute AVANA CRISPR–Cas9 screen are available for download at the DepMap Portal (<https://depmap.org>). The following sequences were used: sgETV6-1: 5′-GCAGCCAATTTACTGGAGCA-3′, sgETV6-2: 5′-GCAGGGATGACGTAGCCCAG-3′, sgETV6-3: 5′-GTGTGTGTATAGAGTTTCCA-3′, sgETV6-4: 5′-GTTATGTTGCACATTATCCA-3′, sgSOX11: 5′-CACCGA

AGATCCCGTTCATCCGGG-3', sgFLII: 5'-CACCGTGTCCGAGAGCAGCTCCAGG-3'. As previously described^{94,95}, sgChr2.2 served as a cutting control and targets a gene desert on chromosome 2, 5'-GGTGTGCGTATGAAGCAGTG-3'; sgLacZ served as a non-targeting control and targets a non-human gene, 5'-AACGGCGATTGACCGTAAT-3'. For ligation into the lentiCRISPRv2 plasmid, the additional bases 5'-CACCG-3' were added to the 5' end of the forward sequence. 5'-AAAC-3' and 5'-C-3' were added at the 5' and 3' ends of the reverse sequence, respectively.

Generation, culture and application of polyclonal dTAG cell lines

Four sgRNAs targeting exon–intron junctions of the endogenous *ETV6* locus were cloned, as described above, into the lentiCRISPRv2 plasmid backbone conferring blasticidin resistance (Addgene, 83480). The following sgRNA sequences were used: (1) sgETV6-endo-1, 5'-TCC TGCTCAGTGTAGCATT-3', (2) sgETV6-endo-2, 5'-GAACACTCACGCA GGTGCGC-3', (3) sgETV6-endo-3, 5'-TCCAGACTCTCACTGAATG-3', and (4) sgETV6-endo-4, 5'-AGTTCATAGAGCACATCACC-3'. A codon-optimized gBlock encoding coding sequences of *ETV6* was cloned into the pLEX_305 vector backbone (Addgene, 91798) to C-terminally tag the ETV6 protein with the FKBP12^{F36V} protein domain and a HA epitope tag. A673 and EW8 parental Ewing sarcoma cells were spin-infected and selected with puromycin and blasticidin as described above. Cells were cultured in DMEM supplemented with 15% FBS, 1% penicillin–streptomycin and 0.5 $\mu\text{g ml}^{-1}$ puromycin and 2.5 $\mu\text{g ml}^{-1}$ blasticidin to maintain selection. Cells were split at a ratio of 1:5 every other day.

The dTAG^V-1 molecule was provided by the Gray Laboratory (Dana-Farber Cancer Institute, Boston, MA) and used at a stock concentration of 10 mM suspended in DMSO. For 6, 24 and 72 h RNA-seq time points, A673 ETV6–dTAG cells were seeded at 1 million cells per 6 cm dish, 0.5 million cells per 6 cm dish and 0.5 million cells per 10 cm dish, respectively. EW8 ETV6–dTAG cells were seeded at 0.75 million cells per 6 cm dish, 0.5 million cells per 10 cm dish and 0.2 million cells per 10 cm dish, respectively. For each time point, three separate dishes were seeded and treated per DMSO or dTAG^V-1 condition. Cells were collected for total RNA extraction and western blot validation (described below). For 6 and 72 h ChIP-seq time points, A673 ETV6–dTAG cells were seeded at 5 million cells per 15 cm dish and 1 million cells per 15 cm dish, respectively. EW8 ETV6–dTAG cells were seeded at 7.7 million cells per 15 cm dish and 1 million cells per 15 cm dish, respectively. Twenty-four hours after seeding, existing medium was exchanged for DMSO or dTAG^V-1-containing medium. For all experiments, dTAG^V-1 was used at a final concentration of 1 μM . Equivalent volumes of DMSO were used as control.

A673 ETV6–dTAG cells were further perturbed to knockout *EW5-FLI* or *SOX11*. sgRNAs targeting *FLII* (5'-TGTCGGAGAGCAGCTCCAGG-3') or *SOX11* (5'-GAAGATCCCGTTCATCCGGG-3') from the Broad AVANA screen were cloned into lentiCRISPRv2 as described above. In total, 250,000 cells per well of a 6-well plate were spin-infected with 2 ml virus. Knockout was validated by western blotting.

Relative viability studies

Cells transduced with lentivirally packaged CRISPR–Cas9 constructs were seeded in 384-well plates at densities of 3,500 (TC32), 2,000 (A673), 250 (EW8) and 1,000 (PEDS0009, PEDS0010 and ES-PDX-001) cells per well suspended in 40–50 μl of medium per well containing 0.5–1 $\mu\text{g ml}^{-1}$ puromycin. Cells from each condition were grown separately in 6–8 wells per plate across 4 plates, which corresponded to day 0, 3, 5 and 7 time points. Wells at plate edges were filled with 50 μl of PBS to maintain humidity. To measure cell viability, 10 μl of CellTiter-Glo reagent (Promega, G7573) was added to each well, luminescing at an intensity proportional to ATP abundance, and plates were shaken at room temperature for 15 min. Luminescence was measured using a FLUOstar Omega microplate reader (BMG LabTech).

Relative viability was calculated by dividing the luminescence measurement of each well on day 7 by luminescence at day 0 using Microsoft Excel 16.50. In parallel, whole cell lysate was collected on day 7 for western blotting to confirm *ETV6* knockout. Statistics shown compare mean relative viability between conditions at day 7, analysed using GraphPad Prism 9.0.0.

Anchorage-independent growth

A 16-gauge blunt-end needle was used to transfer 12 ml of semi-solid methylcellulose-based medium (Stemcell Technologies, 03814) to a 50 ml conical tube and 3 ml of cell suspension containing 15,000 (A673 and /TC32), 5,000 (EW8) or 20,000 (PEDS0009, PEDS0010 and ES-PDX-001) cells. The mixture was vortexed and left at room temperature for 10–15 min until bubbles dissolved. A blunt-end needle was used to transfer 3 ml of the mixture to separate 6 cm dishes, which were placed inside a 15 cm plate containing a PBS-filled 6 cm dish used to maintain humidity. Colonies were stained 7 days later by adding 1 ml of a 1:1 mixture of PBS and MTT dye (Roche Diagnostics, 11465007001) per dish and incubating for 30–45 min at 37 °C. Colonies in each dish were imaged using an ImageQuant LAS 4000 imager (GE Healthcare) and quantified using ImageQuant TL 8.2 software (Cytiva). In parallel, whole cell lysate was collected from cultured cells for western blotting.

Flow cytometry and cell cycle analysis

Cell cycle analysis was performed using Click-iT Plus EdU Alexa Fluor 647 Flow Cytometry Assay kits (Life Technologies, C10424) per kit instructions with minor modifications. Cells were seeded and cultured separately before being pulsed with 10 μM of the modified nucleotide analogue 5-ethynyl-2'-deoxyuridine (EdU) for 90 min at 37 °C. Around 1–2 million cells per sample were trypsinized, washed, fixed, permeabilized and then treated with a reaction cocktail containing Alexa Fluor-647-conjugated picolyl azide to label incorporated EdU. Cells were stained with a RNase-containing propidium iodide solution (Cell Signaling, 4087S) for 45 min at 37 °C. Cells were analysed by flow cytometry at 5,000–10,000 cells per sample on a BD FACS Celesta instrument. Live cells were gated using FSC-A and SSC-A. The data were analysed using FlowJo v.10.6.1 software. Cells were collected from each sample for western blotting.

Mouse studies

All mouse studies were approved by the Dana-Farber Cancer Institute Institutional Animal Care and Use Committee (Animal Welfare Assurance number: D16-00010 (A3023-01)) and were performed in accordance with NIH guidelines for the humane care and use of animals. The intramuscular mouse xenograft experiment (Fig. 1h,i) studied immunodeficient NOD.Cg-Prkdcscid Il2rgtm1Wjl/Szj (NSG) mice ordered from Jackson Laboratory in a semi-orthotopic manner as previously described⁴⁸. A673 cells were lentivirally transduced to express luciferase and CRISPR–Cas9 constructs targeting *ETV6*. These cells were intramuscularly implanted in the hindlimbs of 7-week-old female mice. On the day of implantation, cells were suspended in a 1:1 mixture of PBS and Matrigel (Thermo Fisher Scientific, CB40230C) and injected directly into the hindlimb cranial thigh muscle, away from the sciatic nerve, at a concentration of 50,000 cells per mouse in 50 μl . Five mice per condition (sgChr2.2, sgLacZ, sgETV6-1 and sgETV6-2) were implanted. Disease progression was monitored by serial bioluminescence imaging of the whole body. Bioluminescence was measured 10 min following subcutaneous injection of luciferin using a PerkinElmer IVIS Spectrum (exposure time, 0.5–180 s; binning, 2–16; luminescent, 25,000) to determine the maximum bioluminescence exhibited by each mouse. Mice in each condition were imaged at the same time. Mice were euthanized at the end point. Lung and liver tissue samples were collected following euthanasia and placed in a 6-well dish for bioluminescence imaging. Subcutaneous mouse xenograft experiments were conducted in Jackson NSG mice (Extended Data Fig. 2d)

and CrTac:NCR-Foxn1^{nu} (nude) mice from Taconic Biosciences (Fig. 6h). In the former study, 12-week-old males were used. In the latter study, 6–8-week-old females were used. Cells were suspended in a 7:3 mixture of culture medium and Matrigel and injected bilaterally subcutaneously into sublethally irradiated mice at 3 million cells in 100 μ l. Three to four mice per condition received transplants. Tumours were measured with calipers serially twice weekly.

Animals were euthanized when tumours reached maximal 2 cm in at least one dimension or a humane end point such as ulceration or reduced mobility, in adherence to the NIH/NCI guidelines on limits of tumour size (equal to or less than 2.0 cm per tumour in any one dimension). This limit was not exceeded. Randomization was not appropriate in any study as drug treatments were not used. Mice from the same conditions were kept in different cages to minimize confounding environmental factors. Mice were housed with strictly controlled temperature and humidity and kept on 12-h light and dark cycles. No statistical methods were used to predetermine sample sizes, but sample sizes were similar to those reported in previous publications in which statistical significance was achieved^{48,66}. Data distribution was assumed to be normal, but this was not formally tested, with the exception of data shown in Fig. 1i, for which the data were not normal (Shapiro–Wilk $P < 0.05$) and thus log-transformed. Data collection and analysis were not performed blind to the conditions of the experiments. No animals or data points were excluded from the analyses.

Crystal violet staining and quantification

Cell samples were cultured separately and re-seeded at normalized cell densities (50,000 cells per well in 6-well plate) every 5 days with refreshed DMSO or 1 μ M dTAG^{V-1}. On day 20, each well was incubated with 1 ml of crystal violet stain, composed of 20% methanol and 1% w/v crystal violet powder (Sigma Aldrich, C6158) at room temperature for 20 min. Wells were washed with 3 ml of deionized H₂O five times and dried at room temperature. Plates were imaged using an ImageQuant LAS 4000 imager (GE Healthcare). The median intensity of stain in each well was quantified using ImageQuant TL 8.2 image analysis software (Cytiva).

Western blotting

Cells were lysed using cell lysis buffer (Cell Signaling Technology, 9803S), supplemented with protease inhibitor (Sigma Aldrich, 11836170001) and phosphatase inhibitor (Sigma Aldrich, 04906837001). Protein quantification of whole cell lysate was measured using a Bradford-based colorimetric assay (Bio-Rad, 5000006). Around 50–60 μ g of whole cell lysate was mixed with loading buffer (Life Technologies, NP0007), reducing buffer (Life Technologies, NP0009) and water and heated to 75 °C for 10 min. Samples were loaded onto 4–12% bis-tris 10-well gels (Life Technologies, NP0335BOX) and run at 100 V for 30 min followed by 150 V for 90 min using MOPS buffer (Life Technologies, NP0001). Gels were transferred to polyvinylidene difluoride membranes (Thermo Fisher Scientific, IPVH00010) at 100 V for 90 min using transfer buffer (Boston BioProducts, BP-190-1L) at 4 °C. Membranes were blocked in milk (Cell Signaling Technology, 9999S) for 60 min at room temperature. Membranes were rocked overnight at 4 °C in a solution of Tris-buffered saline and Tween-20 (TBST; Cell Signaling Technology, 9997S) containing 5% w/v BSA (Research Products International, A30075-1000.0), 0.02% sodium azide (Santa Cruz Biotechnology, SC-208393) and primary antibody. The following day, membranes were washed in TBST five times, 5 min per wash. For a subset of western blots, membranes were rocked for 1 h at room temperature in milk containing 1:5,000 dilution of HRP-conjugated secondary antibody against mouse (Cell Signaling Technology, 7076S) or rabbit (Cell Signaling Technology, 7074S). Membranes were then washed in TBST three times and immersed in a solution containing chemiluminescent substrate (Life Technologies, 34076), allowed to develop for 1 min, then imaged using film (Thermo Fisher Scientific, PI34091). Other western blots were imaged using a LI-COR system.

Membranes were rocked in a TBST solution containing a 1:10,000 dilution of secondary antibody against mouse (LI-COR Biosciences, 926-32210) and rabbit (LI-COR Biosciences, 926-68071) and 1:10,000 dilution of 10% SDS solution (Life Technologies, 15553027) for 1 h at room temperature. Membranes were washed in TBST three times and then briefly rinsed in PBS and imaged on an Odyssey CLx machine at medium resolution (ImageStudioLite 5.2.5).

The following primary antibodies were used at the following dilutions: anti-GAPDH at 1:10,000 (2118S, rabbit, monoclonal, Cell Signaling Technology); anti-ETV6 primary at 1:500 (WH0002120M1-100UG, mouse, monoclonal, Sigma Aldrich or SC-166835, mouse, monoclonal, Santa Cruz Biotechnology); anti-HA at 1:2,000 (3724S, rabbit, monoclonal, Cell Signaling Technology); anti-PARP at 1:1,000 (9542S, rabbit, polyclonal, Cell Signaling Technology); anti-cleaved caspase-3 at 1:1,000 (9664S, rabbit, monoclonal, Cell Signaling Technology); anti-FLI1 primary antibody at 1:1,000 (ab15289, rabbit, polyclonal, Abcam); anti-FAS at 1:500 (SC-8009, mouse, monoclonal, Santa Cruz Biotechnology); anti-SEMA5B at 1:500 (PA5113369, rabbit, polyclonal, Thermo Fisher Scientific); anti-BCL11B at 1:500 (12120S, rabbit, monoclonal, Cell Signaling Technology); and anti-SOX11 at 1:1,000 (58207S, rabbit, monoclonal, Cell Signaling Technology).

SOX11 overexpression

Complementary DNAs of wild-type *SOX11* and *SOX11* mutants harbouring a deletion of the DBD (H48-R119) were synthesized as gBlocks fragments (IDT), and then cloned into a pLX_TRC307 lentiviral expression vector co-expressing a puromycin resistance gene (obtained from the Genetic Perturbation Platform at the Broad Institute) using a Gibson Assembly Cloning kit (New England Biolabs E5510S). Constructs were lentivirally delivered to cells as described above.

Inducible *EWS-FLI* overexpression

The previously described vector pINDUCER20-EWS/FLI-HA⁶⁶, which encodes HA epitope-tagged EWS-FLI, was mutated to create the R340N mutation⁹⁶ by site-directed mutagenesis using a NEB Q5 Quick Change Site Directed Mutagenesis kit (E0554) using the following primers: forward 5'-CCGGGCCCTCAATTACTACTATGATAAAAAC-3'; reverse 5'-CTCAGCTTGTCGTAATTC-3'. The correct mutation was confirmed by Sanger sequencing with both a forward and reverse sequencing primer (forward: 5'-TCCCACACCGACCAGTCTCAC-3'; reverse: 5'-AGACTG CCTTGGGAAAAGCGCC-3'). pINDUCER20-GFP-HA vector was used as a control. RD cells stably expressing these three vectors were generated. For knockout experiments, RD cells stably expressing the inducible vectors were transduced with sgRNAs targeting Chr2.2 (cutting control) or *ETV6*. Expression of EWS-FLI was induced using 1 μ g ml⁻¹ doxycycline replenished every 48 h.

Rescue of *ETV6* knockout with wild-type and ETS-deleted *ETV6* overexpression

DNA fragments encoding codon-optimized *ETV6* wild-type (*ETV6-WT*) and mutant *ETV6* harbouring deletion of the ETS domain (*ETV6-ΔETS*) were purchased from gBlock (IDT) and cloned into pDONR-221 via BP gateway cloning. Constructs were further cloned into pINDUCER20 (Addgene, 44012) by LR cloning and lentivirally packaged as described above. A673 and EWS cells were transduced with lentivirus encoding either *ETV6-WT* or *ETV6-ΔETS* and incubated with 100 ng ml⁻¹ of doxycycline or vehicle for 24 h. Subcellular fractionation was performed according to the manufacturer's protocol (Thermo Fisher, PI78840). Western blotting and cell viability experiments were performed as described above.

qPCR

Total RNA was extracted from cells using an extraction kit with column-based genomic DNA removal (Qiagen, 74134). RNA was reverse transcribed to cDNA using an iScript kit (Bio-Rad Laboratories,

1708841) and diluted 1:7 with H₂O. For sgFlI rescue experiments, A673 ETV6–dTAG cells were transduced with sgChr2.2 or sgFlI CRISPR–Cas9 constructs and treated separately with DMSO or dTAG^V-1 in duplicate. All qPCR reactions were performed using a TaqMan system (Thermo Fisher Scientific) with technical triplicates. Probes were selected to span exon–exon junctions when possible. Specific probes were as follows: *GAPDH*: Hs02758991_g1; *FAS*: Hs00236330_m1; *ACTA2*: Hs00426835_g1; *TRIB1*: Hs00179769_m1; *SEMA5B*: Hs00400720_m1; *BCL11B*: Hs01102259_m1; and *SOX11*: Hs00846583_s1. In each qPCR reaction, the gene of interest was measured using FAM dye, whereas *GAPDH* control was measured using VIC dye. Samples were analysed in 384-well plate format using 5 µl TaqMan gene expression master mix (Thermo Fisher Scientific, 4369016), 0.5 µl of FAM-emitting probe, 0.5 µl of VIC-emitting *GAPDH* probe and 4 µl of diluted cDNA for a total of 10 µl per reaction. qPCR plates were analysed using a QuantStudio 6Flex Real-Time PCR machine and the accompanying QuantStudio Real-Time PCR software v.1.7 (Thermo Fisher Scientific). The delta-threshold cycle number (Δ Ct) was calculated as the difference in threshold cycle number (Ct) between the gene of interest and *GAPDH*. The $\Delta\Delta$ Ct was calculated as the difference between the Δ Ct of a particular sample and the average Δ Ct of the DMSO-treated, sgChr2.2 control samples. Fold increase in gene expression (after the loss of ETV6) was calculated as the ratio of $2^{-\Delta\Delta Ct}$ in dTAG^V-1-treated cells to the average $2^{-\Delta\Delta Ct}$ in DMSO-treated cells, in either the sgChr2.2 or the sgFlI conditions.

RNA-seq

All RNA-seq experiments were performed using total RNA extracted using a column-based kit (Qiagen, 74104) and treated with DNase digestion. The Life Technologies external RNA control consortium (ERCC) RNA spike-in samples were added to each sample for normalization per kit instructions (Thermo Fisher Scientific, 4456740). For all RNA-seq experiments, except the A673 sgETV6 CRISPR–Cas9 experiments, RNA-seq library preparation and sequencing were performed by Novogene (<https://en.novogene.com>) at a depth of roughly 20 million reads per sample. Per Novogene correspondence, the quality control for the RNA samples was performed using Qubit fluorometric quantitation (Thermo Fisher Scientific) and a Bioanalyzer instrument (Agilent). Libraries were then prepared using a New England Biolabs NEBNext Ultra II non-directional RNA Library Prep kit. Library quality and concentrations were assessed using Labchip (Perkin Elmer) and qPCR. Libraries were sequenced in 150-bp paired-end fashion on a Novaseq6000 instrument (Illumina). For the A673 sgETV6 CRISPR–Cas9 experiments, polyA-tailed mRNA was isolated from 1 µg total RNA using a magnetic bead-based kit per kit instructions (New England Biolabs, E7490S). RNA-seq library preparation was performed using a NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs, E7760S). Libraries were quantified using a Qubit dsDNA high sensitivity assay (Q32851). The distribution of DNA fragment sizes was measured using a High Sensitivity D1000 assay (Agilent, ScreenTape, 5067-5584; reagents, 5067-5585). The molarity of each library was calculated and normalized to 4 nM. Libraries were pooled and sequenced on a Nextseq 500 instrument (Illumina) (single-end; 75 cycles at a depth of roughly 40 million reads per sample) using a Nextseq 500 sequencing kit (Illumina, 20024906).

CUT&Tag

CUT&Tag was performed as previously described⁵⁰ with slight modifications by the Lessnick Laboratory (Nationwide Children's Hospital, Columbus, OH). About 250,000 cells per CUT&Tag condition were bound to BioMag Plus Concanavalin A-coated magnetic beads (Bangs Laboratories, BP531) and incubated with primary antibodies (ETV6 rabbit, Bethyl A303-674, 1:50; ETV6 mouse, Sigma WH0002120M1, 1:50; rabbit anti-mouse, Abcam, ab46540, 1:50) overnight at 4 °C, and secondary antibodies (guinea pig anti-rabbit IgG, Antibodies-Online ABIN101961, 1:100; rabbit anti-mouse, Abcam ab46540, 1:100) for 1 h at

room temperature. Adapter-loaded protein A–Tn5 fusion protein was added at a dilution of 1:250 and incubated for 1 h at room temperature. To activate Tn5, tagmentation buffer containing MgCl₂ was added and samples were incubated for 1 h at 37 °C. Reactions were stopped by addition of EDTA, and DNA was solubilized with SDS and proteinase K for 1 h at 50 °C. Total DNA was purified using phenol–chloroform extraction followed by ethanol precipitation. CUT&Tag libraries were prepared using NEBNext HiFi 2× PCR master mix (NEB, M0541S) and indexed primers⁹⁷ using a combined annealing–extension step at 63 °C for 10 s and 15 cycles followed by a 1.1× post-amplification AMPure XP (Beckman Coulter, A63880) bead clean-up. Libraries were pooled and sequenced (2 × 150 bp paired end) on an Illumina HiSeq4000 platform (Nationwide Children's Hospital Institute for Genomic Medicine). Two independent replicates of each CUT&Tag sample were prepared.

CUT&RUN

CUT and release using nuclease (CUT&RUN) was performed as previously described^{98–100} with slight modifications. In brief, 500,000 cells per condition were bound to activated ConA beads (EpiCypher 21-1401). Next, the ConA bead–cell mixture was resuspended in a cold antibody buffer and FLI-1-(ab133485; 1 µg per sample) antibody or 0.5 µg H3K4me3 (EpiCypher, 13-0041) as positive and 0.5 µg IgG (EpiCypher, 13-0042) as negative control were per sample added overnight. pAG-MNase (EpiCypher, 15-1016) was then added to each reaction to allow binding to the antibody-labelled chromatin. *E. coli* spike-in DNA (EpiCypher, 18-1401) was added following MNase activation. Subsequently, targeted chromatin was digested and released by the addition of CaCl₂. The fragmented chromatin was purified using a CUTANA DNA Purification kit (EpiCypher, 14-0050). Quantification, library preparation and sequencing were performed by the genomics core at Dana-Farber Cancer Institute.

ChIP-seq

Antibodies were conjugated to magnetic beads. For each immunoprecipitation (IP), 100 µl of protein A Dynabeads (Thermo Fisher Scientific, 10002D) were washed three times in 1 ml BSA blocking solution (0.5% w/v sterile-filtered BSA in H₂O) and resuspended in 250 µl. Beads were then rotated overnight at 4 °C with antibody, using 5 µg to target H3K27ac (Abcam, 4729) or 10 µg to target TFs (anti-HA, Abcam, ab9110; anti-FLII, Abcam, ab15289). For comparative studies (that is, comparing the relative binding of EWS–FLI), 2 µg of spike-in antibody recognizing a *Drosophila*-specific histone variant was added (Active Motif, 61686). The following morning, the antibody-conjugated beads were washed four times in 1 ml BSA blocking solution and then resuspended in 100 µl of the solution and stored at 4 °C.

To prepare sheared chromatin, Ewing sarcoma cells (20 million cells per ChIP reaction) were collected in a 1.5 ml tube and washed twice in 1 ml PBS. Cells were then crosslinked by resuspension in 1 ml PBS containing 1% methanol-free formaldehyde (Thermo Fisher Scientific, 28906) and rotated for 10 min at room temperature at 12 r.p.m. The reaction was quenched with 100 µl of 1.25 M glycine (Sigma Aldrich, G7126) and 100 µl 1 M Tris-HCl pH 8.0 (Thermo Fisher Scientific, 15568025). Cell pellets were washed twice with 1 ml PBS at room temperature and resuspended in 1 ml of SDS lysis buffer (0.5% SDS, 5 mM EDTA, 50 mM Tris-HCl pH 8.0) supplemented with protease inhibitor cocktail (Thermo Fisher Scientific, PI78429) and incubated at room temperature for 2 min with gentle agitation. Lysates were centrifuged at 15,000g for 10 min at 4 °C. The nuclear pellet was re-suspended in 950 µl of ChIP IP buffer (2 parts SDS lysis buffer and 1 part Triton dilution buffer, which was composed of 100 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM EDTA, 0.2% NaN₃ and 5% Triton X-100) supplemented with protease inhibitor and transferred to a milliTUBE (Covaris, 520130). Sonication was performed on an E220 Focus Ultra sonicator (Covaris) at 5% duty cycle, 140 W peak power, 200 cycles per burst, at 4 °C for 30 min per milliTUBE. Sheared chromatin was

transferred to a 1.5 ml tube and centrifuged at 15,000g for 10 min at 4 °C. The supernatant of sheared chromatin was transferred to a new reaction tube. To prepare the ChIP DNA input sample, 5 µl of sheared chromatin was transferred to a PCR strip-tube and mixed with 40 µl de-crosslinking buffer (100 mM NaHCO₃ and 1% SDS buffer), 1 µl RNase A (Thermo Fisher Scientific, 12091021) and 1 µl proteinase K (Thermo Fisher Scientific, AM2546). The tube was incubated for 2 h at 65 °C in a thermal cycler to de-crosslink DNA–protein covalent bonds. DNA was isolated using Agencourt AMPure XP bead-based purification at a 1.2× ratio (Beckman Coulter, A63881), eluted in 50 µl H₂O and stored at –20 °C. The remaining sheared chromatin was divided or pooled according to the target of interest; at least 5 million cells were used for IP of histone marks and 40 million cells for TFs. Each IP reaction was brought up to a total volume of at least 1 ml with ChIP IP buffer. Pooled reactions were conducted in 2 ml or 5 ml reaction tubes. 50 ng or 20 ng of *Drosophila* spike-in chromatin was added for each H3K27ac or TF ChIP reaction, respectively. The 100 µl conjugated bead–antibody solution was then added to the sheared chromatin. IP reactions were rotated overnight at 4 °C.

ChIP reactions were washed twice in 1 ml low-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, and 150 mM NaCl), high-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, and 500 mM NaCl), lithium chloride buffer (0.25 M LiCl, 1% IGEPAL-CH 630, 1% sodium deoxycholate, 10 mM Tris-HCl pH 8.0, and 1 mM EDTA) and then once in 700 µl Tris-EDTA buffer (Sigma Aldrich, 93283). Chromatin was eluted using 100 µl fresh ChIP elution buffer (1% SDS and 0.1 M NaHCO₃) and rotated at room temperature for 15 min. Eluate was transferred to PCR tubes and mixed with 8 µl 2.5 M NaCl, 1 µl RNase A and 1 µl proteinase K. Samples were de-crosslinked for 12–16 h at 65 °C on a thermal cycler. ChIP DNA was extracted from the de-crosslinked samples using AMPure XP beads at a 1.2× ratio and eluted in 20 µl of H₂O. DNA was quantified using a Qubit dsDNA high sensitivity assay (Q32851). DNA fragment sizes were measured with a TapeStation instrument using a High Sensitivity D1000 assay (Agilent, ScreenTape, 5067-5584; reagents, 5067-5585).

ChIP-seq libraries were prepared using a SMARTer ThruPLEX single-index DNA-Seq kit (Takara Bio, R400674, R400695). H3K27ac and TF samples were PCR-amplified 4 and 10 cycles, respectively. Libraries were prepared as described above and sequenced in 37-bp paired-end fashion for 75 cycles (Illumina, 20024906) at a depth of roughly 30 million reads per sample on the NextSeq 500.

ATAC-seq

A673 ETV6–dTAG cells were seeded and treated separately with DMSO or 1 µM dTAG^V-1 for 72 h. ATAC-seq was performed as previously described⁹⁷ on samples of 100,000 cells using a publicly available protocol (available at [https://www.med.upenn.edu/kaestnerlab/assets/user-content/documents/ATAC-seq-Protocol-\(Omni\)-Kaestner-Lab.pdf](https://www.med.upenn.edu/kaestnerlab/assets/user-content/documents/ATAC-seq-Protocol-(Omni)-Kaestner-Lab.pdf)) without modifications. The molarity of each library was calculated using a Qubit dsDNA Broad Range Assay kit (Thermo Fisher Scientific, Q32850) and a TapeStation D5000 dsDNA assay (Agilent; ScreenTape, 5067-5588; Ladder, 5067-5590; reagents, 5067-5589). Libraries were pooled and sequenced in 37-bp paired-end fashion for 75 cycles on an Illumina NextSeq 500 instrument.

ChIP-seq data analysis

The raw Illumina sequencer output was converted to fastq format using the program bcl2fastq (v.2.17). Sequencing read quality was examined using FastQC (<http://www.bioinformatics.babraham.ac.uk>) (v.0.11.9). Trimming of low-quality reads and clipping of sequencing adapters was done using the program Trimmomatic (v.0.36)¹⁰¹, and all reads shorter than 40 bp after trimming were discarded. Reads were aligned to the human genome (hg19) using Bowtie2 (v.2.3.5)^{102,103} using the ‘–very_sensitive’ preset collection of parameters. File conversion of .bam to .sam was done using SamTools (v.1.9q)¹⁰⁴, and duplicate reads

were removed using Picard-tools (v.2.19.0) (<http://picard.sourceforge.net>). ChIP-seq peaks were called using MACS2 (ref. 105) with a false discovery rate (FDR) $q < 0.01$ unless otherwise stated. The MACS2 algorithm utilizes a dynamic Poisson distribution to capture local biases in the genomic sequence, which allows for a sensitive and robust prediction of peaks. Unless otherwise noted, peaks were assigned to the closest gene within ±400 kb using the ChIPseeker package in R¹⁰⁶. Visualizations of the ChIP-seq data tracks were produced with the R Bioconductor Gviz package¹⁰⁷.

CUT&Tag data analysis

Quality control on raw sequencing reads were performed using FastQC (<http://www.bioinformatics.babraham.ac.uk>) (v.0.11.4). Adapter sequences and/or low-quality reads were trimmed using trim_galore (<http://www.bioinformatics.babraham.ac.uk>) (0.4.4_dev). Reads were aligned to human (hg19) and spike-in *E. coli* (Escherichia_coli_K_12_DH10B NCBI 2008-03-17) genomes using Bowtie2 (v.2.3.4.3)^{102,103} with the following options: --no-unal --no-mixed --no-discordant --dovetail --phred33 -q -110 -X 700. The option --very-sensitive was added when aligning to the spike-in genome. SamTools (v.1.9)¹⁰⁴ was used to convert .sam to .bam with the ‘-bq 10’ option. Counts of mapped reads were spike-in normalized by calculating a scale factor using the ‘median ratio method’ from DESeq2. Spike-in normalization in conjunction with the median ratio method provide a robust normalization method to appropriately account for global changes of ETV6 occupancies^{108,109}. Peaks in each biological replicate were called using MACS2 (v.2.2.7.1)¹⁰⁵ with the spike-in normalization. All duplicate reads were kept in the analysis. To ensure reproducibility and consistency of peaks called across multiple biological replicates, we calculated irreproducibility discovery rate¹⁰⁹ values and combined the replicates with rabbit anti-mouse as controls using DiffBind (v.2.14.0)¹¹⁰ and DESeq2 (v.1.26.0)¹⁰⁸. To ensure high-quality peaks that are most likely to represent biological signals, the final peak lists were generated with following thresholds: irreproducibility discovery rate < 0.005, FDR < 0.05, log₂(fold change) > 3 and mean normalized counts of ETV6 > 80 (Bethyl antibody) and >300 (Sigma antibody).

CUT&RUN data analysis

CUT&RUN FLII data for the PEDS0009 sample used a pipeline based on the bulk-level method outlined in CUT&RUNTools 2.0 (ref. 111) that is largely the same as the ChIP-seq pipeline. The changes to the ChIP-seq pipeline are an extra adapter trimming step run after Trimmomatic using kseq from CUT&RUN Tools and the addition of the ‘–dovetail’ flag to the Bowtie2 command. CUT&RUN samples also included *E. coli* spike-in for sample normalization and it was aligned to the *E. coli* (Escherichia_coli_K_12_DH10B NCBI 2008-03-17) genome.

Differential ChIP-seq binding

Differential binding analysis in ETV6–dTAG ChIP-seq samples was performed with the R Bioconductor package CSAW⁸⁹. CSAW uses a sliding window approach to count reads across the genome from sorted and indexed .bam files, for which each window is tested for significant differences between libraries using statistical methods from the edgeR package. Differential CSAW analysis was performed on A673 and EW8 ETV6–dTAG at 6 and 72 h in FLII and H3K27ac. The differential analysis performed here normalized samples based on *Drosophila* spike-in values, the reads of which were aligned to the dm6 version of the *Drosophila* genome. The differential ChIP-seq analysis procedure generally followed the approach outlined in the CSAW introductory usage tutorial as follows. The .bam files were read in allowing a maximum fragment length of 800, a minimum $q = 20$ and discarding any reads that fell in the hg19 or dm6 ENCODE blacklist files. A window size of 150 bases was used for analysis and tiled across the genome in 50 base steps. The ChIP-seq input control samples were used to help filter out regions containing just background reads by binning input control reads into

10,000 base blocks with a threshold of minimum prior counts of 2. The binned input reads were then compared with the ChIP-seq binding across all regions, and all ChIP-seq regions with a fold change of less than 3 over input were filtered out. After filtering, adjacent and overlapping 150 base regions were merged together to reduce the number of hypotheses tested (for example, A673 6 h ETV6–dTAG FLII had an average merged window width of 494 bases). *Drosophila* spike-in control reads were processed similarly to the human reads except, as there was no input control for the spike-in control, the spike-in reads were filtered using a global filtering method that required regions to be threefold above background. The counts for all enriched spike-in regions were used to calculate the normalization factors by applying the trimmed mean of M-values method on these counts via the function `normFactors`. Differential binding is tested for significance using the quasi-likelihood framework in the `edgeR` package, whereby `edgeR` models the counts using a negative binomial distribution that accounts for over-dispersion between biological replicates. To account for multiple hypothesis testing, CSAW converts per-window statistics into a *P* value for each region and then applies the Benjamini–Hochberg method to calculate the corrected FDR.

ChIP-seq heatmaps

ChIP-seq heatmaps were produced by functions in the following `deeptools` package (v.3.3.0)¹¹²: `computeMatrix`, `plotProfile` and `plotHeatmap`. All heatmaps were made using data in `.bigWig` files that have been generated by `deeptools bamCompare` that generates `.bigWig` files based on the comparison of a ChIP-seq sample `.bam` file to its corresponding input (from the same cell line and same batch) while being simultaneously normalized for sequencing depth. The function `computeMatrix` was then used to calculate scores for genome regions and to prepare an intermediate file that can be used with `plotHeatmap` and `plotProfiles`. Unless otherwise stated, the genome regions were regions defined by a BED file corresponding to ETV6 or FLII peaks. For Fig. 2a–c and Extended Data Fig. 3a–c, `computeMatrix` was used with multiple `.bigWig` score files and two BED region files, in which the ETV6 peaks are split into two groups depending on whether the ETV6 peak overlapped with a region defined by gene TSSs ± 2.5 kb according to UCSC hg19 refGene transcript definitions. Figure 2g,h used regions defined by differential FLII regions from $P < 0.05$ CSAW, whereby regions not intersecting with a TSS were further divided into two groups according to whether the region intersects with a H3K27ac ChIP-seq peak from MACS2 with $q < 0.01$ in the parental A673 or EW8 cell line.

GGAA repeat frequency at peak locations

Stacked bar plots were created in R using frequencies of overlap from the function `summarizePatternInPeaks` from R Bioconductor package `ChIPpeakAnno` (v.3.9)¹¹³. The function `summarizePatternInPeaks` was used to calculate the frequency of overlap of regions of the standard hg19 reference genome with GGAA repeats (from a single GGAA up to five consecutive GGAA sequences without any gaps) with peaks in FLII and ETV6 as called by MACS2. The ENCODE datasets analysed were from the Gene Expression Omnibus: GSE96274 (B lymphocyte) and GSE95877 (K-562).

Differential ATAC-seq regions

Processing of ATAC-seq data (that is, Fig. 4c) used the same pipeline as the ChIP-seq data, although an extra step was added after `Bowtie2` alignment that used `samtools` to remove mitochondrial reads (ChrM). CSAW was used for the differential analysis of ATAC-seq data in the same manner as CSAW was used with ChIP-seq data, except that there was no input control for filtering or spike-in control for sample normalization. In the absence of a matching input control, CSAW region filtering was performed by requiring regions to be threefold above the local background, whereby local background was defined by using wider windowing function of 2,000 bases and requiring regions to be

threefold above the neighbouring regions. Within CSAW, ATAC-seq samples were normalized to the background using 10,000 base windows to calculate compositional biases of samples.

RNA-seq data analysis

Gene expression values were derived from paired-end RNA-seq data, except for the A673 sgETV6 CRISPR–Cas9 RNA-seq experiment, which was sequenced in single-end fashion. The RNA-seq processing pipeline was roughly modelled on the GTEx pipeline (<https://github.com/broadinstitute/gtex-pipeline/>)¹¹⁴. FastQC was used to evaluate read quality on raw RNA-seq reads. Reads were aligned to the human genome (hg19) using STAR¹¹⁵. Transcript-level quantifications were calculated using RSEM (v.1.3.1)¹¹⁶. Gene counts from STAR were then used to quantify differentially expressed genes between the experimental and control conditions using the R Bioconductor package `DESeq2` (ref. 108) using the approximate posterior estimation for GLM coefficients (`apecglm`) method for effect size. Normalized expression values for individual samples were obtained from RSEM \log_2 (TPM) values with the RSEM \log_2 (TPM + 1) values used for GSEA and producing RNA-seq heatmap plots.

The RNA-seq samples included the ERCC spike-in control mix¹¹⁷. Sequences for the ERCC transcripts were added to the hg19 reference for the STAR transcript alignment, and the fold changes of ERCC probes were examined in the `DESeq2` output. Fold changes for ERCC probes were typically very small between the conditions in the ETV6–dTAG sample sets (for example, average fold change for 24 h A673 ETV6–dTAG of 0.995 between conditions). As such, ERCC spike-ins were not used to perform sample normalization.

Gene set pathway enrichment analysis

Gene set pathway enrichment analysis was performed with signatures from v.6.0 of the Broad Institute's molecular signature database (MSigDB) (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>) using the c2 curated gene sets from various sources such as online pathway databases, the biomedical literature and knowledge of domain experts. These pathway enrichment results are shown in Fig. 3f and Extended Data Fig. 4g. Pathway enrichment analysis was performed in R using the `clusterProfiler` package that provides the `enricher` function for a hypergeometric test for a test of over-representation of pathway genes in a set of user-defined genes. Figure 3f shows a combined enrichment plot of the top MSigDB c2 pathways enriched in the ETV6-repressed genes at 6, 24 and 72 h common to both A673 and EW8 (genes up in ETV6 dTAG^V-1 treatment RNA-seq). The plot shows a selected subset of the top enriched c2 gene sets, and the complete set of enriched sets is shown Supplementary Tables 7–11. The dot size corresponds to the number of genes in the gene set out of the total number of significantly ETV6-repressed genes at 6, 24, and 72 h (85, 251 and 832 genes, respectively). The colour corresponds to the gene set grouping. Missing points at times along the *x* axis represent times at which the enrichment was not significant with $P < 0.05$. The pathways are ordered first by the gene group and then by the average gene ratio (count of repressed genes in a pathway/number of repressed genes) across the three time points. Extended Data Fig. 4g shows a combined enrichment plot of the top MSigDB c2 pathways enriched in the ETV6-activated genes at 6, 24 and 72 h common to both A673 and EW8 (genes down in ETV6 dTAG^V-1 treatment RNA-seq). The plot shows a selected subset of the top enriched c2 gene sets, and the complete set of enriched sets is shown Supplementary Tables 12–17. The dot size corresponds to the number of genes in the gene set out of the total number of significantly ETV6-activated genes at 6, 24 and 72 h (33, 130 and 543 genes, respectively). The colour corresponds to the gene set grouping. Missing points at times along the *x* axis represent times at which the enrichment was not significant. The pathways are ordered first by the gene group and then by the average gene ratio (count of repressed genes in a pathway/number of repressed genes) across the three time

points. Extended Data Fig. 7a shows a pie chart of the top 100 enriched c5 gene sets, ranked by significance, in A673 ETV6-dTAG cells at 24 h. Each c5 gene signature was assigned to one of the categories listed; a complete list is shown in Supplementary Table 19.

GSEA

The GSEA algorithm^{118,119} was used to evaluate the association of gene sets with ETV6 regulation. GSEA was run with 2,500 permutations of the phenotype using signal-to-noise to rank genes. This GSEA algorithm was used for Fig. 3e to test enrichment and generate enrichment plots of ETV6-bound genes in ETV6-regulated genes. The A673 ETV6 peak locations are defined by the peaks that overlap in all three A673 ETV6 samples (two A673 ETV6 CUT&Tag samples from two ETV6 antibodies and one untreated A673 ETV6 dTAG HA sample) and the EW8 ETV6 peak locations are defined by peaks in the EW8 ETV6 HA sample. ETV6-bound genes were identified by mapping the peaks to their nearest genes using the R package ChIPseeker.

Statistics and reproducibility

Further information is available in the Nature Portfolio Reporting Summary linked to this article. Figure panels displaying data from experiments with $n = 1$ include Figs. 2a–e, 4h, 5c–g and 6c,f,h and Extended Data Figs. 2c,e, 3a–d, 6d–g and 7g. Figure panels displaying data from experiments with $n = 2$ include Figs. 2f–i, 4a–c,e–g and 6a,b and Extended Data Figs. 3f,g and 5a,c,d. All other figure panels display data from experiments with at least $n = 3$.

Reporting summary

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

Data availability

CRISPR–Cas9 screen data and the genomic characterization of cancer cell lines (RNA-seq) used in this study are publicly available at <https://depmap.org>. Gene expression data from the Treehouse Childhood Cancer Initiative characterizing primary tumours is publicly available at <https://treehousegenomics.soe.ucsc.edu/public-data/>. The Broad Institute's MSigDB is publicly available at <http://www.broadinstitute.org/gsea/msigdb/index.jsp>. Genomics data shown in this study have been deposited in the Gene Expression Omnibus under accession code GSE181554. Source data are provided with this paper.

Code availability

No custom code was used in this study. Analysis was done with publicly available pipelines using approaches described in the Methods.

References

- Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
- Khoogar, R. et al. Single-cell RNA profiling identifies diverse cellular responses to EWSR1/FLI1 downregulation in Ewing sarcoma cells. *Cell Oncol.* **45**, 19–40 (2022).
- Reynolds, C. P. et al. Biological classification of cell lines derived from human extra-cranial neural tumors. *Prog. Clin. Biol. Res.* **271**, 291–306 (1988).
- Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
- Price, C. et al. Genome-wide interrogation of human cancers identifies EGLN1 dependency in clear cell ovarian cancers. *Cancer Res.* **79**, 2564–2579 (2019).
- Malone, C. F. et al. Selective modulation of a pan-essential protein as a therapeutic strategy in cancer. *Cancer Discov.* **11**, 2282–2299 (2021).
- Welford, S. M., Hebert, S. P., Deneen, B., Arvand, A. & Denny, C. T. DNA binding domain-independent pathways are involved in EWS/FLI1-mediated oncogenesis. *J. Biol. Chem.* **276**, 41977–41984 (2001).
- Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21 (2015).
- Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).
- Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, e21856 (2017).
- Schmid, M., Durussel, T. & Laemmli, U. K. ChIC and ChEC; genomic mapping of chromatin proteins. *Mol. Cell* **16**, 147–157 (2004).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
- Langmead, B., Wilks, C., Antonescu, V. & Charles, R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35**, 421–432 (2019).
- Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- 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).
- Hahne, F. & Ivanek, R. Visualizing genomic data using Gviz and Bioconductor. *Methods Mol. Biol.* **1418**, 335–351 (2016).
- Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
- Chen, K. et al. The overlooked fact: fundamental need for spike-in control for virtually all genome-wide analyses. *Mol. Cell. Biol.* **36**, 662–667 (2015).
- Stark, R. & Brown, G. *DiffBind: Differential Binding Analysis of ChIP-Seq Peak Data* (Bioconductor, 2021).
- Yu, F., Sankaran, V. G. & Yuan, G.-C. CUT&RUNTools 2.0: a pipeline for single-cell and bulk-level CUT&RUN and CUT&Tag data analysis. *Bioinformatics* **38**, 252–254 (2021).
- Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
- Zhu, L. J. et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**, 237 (2010).
- Battle, A., Brown, C. D., Engelhardt, B. E. & Montgomery, S. B. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
- Jiang, L. et al. Synthetic spike-in standards for RNA-seq experiments. *Genome Res.* **21**, 1543–1551 (2011).
- Mootha, V. K. et al. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

Acknowledgements

D.L. was supported by a NIH NIGMS Medical Scientist Training Program Award (T32GM007753), a T32 Award (T32GM722643), a Harvard Landry Cancer Biology Consortium Fellowship, and a NIH F30 Award (F30CA246925); K.S. by the National Cancer Institute R35 CA210030, R01 CA204915 and a Cancer Moonshot U54 CA231637, a St. Baldrick's Foundation Robert J. Arceci Innovation Award, the St. Jude Children's Research Hospital Collaborative Research Program, and the Brian MacIsaac Sarcoma Foundation; J.M.E. by The Helen Gurley Brown Presidential Initiative (The Pussycat Foundation); C.F.M. by a Ruth L. Kirschstein National Research Service Award (F32CA243266); S.L. by NCI K99 CA263161 and a Leukemia & Lymphoma Society fellowship; N.V.D. by the Julia's Legacy of Hope St. Baldrick's Foundation Fellowship; B.K.A.S. by the Department of Defense PRCRP Horizon Award CA181249; N.S.G. by the Katherine L. and Steven C. Pinard Research Fund and the Hale Center for Pediatric Cancer; and S.L.L. by U54 CA231641. We thank J. Mora for the gift of PDX Ewing cells, J. Perry for administrative support, A. Krivtsov and J. Gadrey for experimental guidance, C. Hatton for bioinformatics assistance, and A. Balboni Iniguez and A. Rubin for mentorship and scientific discussions.

Author contributions

D.L. and K.S. conceived and designed the experiments, interpreted results and wrote the manuscript. K.N.R. performed computational analyses of the genomics data and interpreted results. J.M.E., C.F.M., S.L., N.W.M. and A.H.S. assisted with experimental design, execution and interpretation. N.V.D. performed analyses of CRISPR-Cas9 screen data. A.M. and A.C. assisted with animal experiments. B.K.A.S. assisted with cloning for the dTAG system. B.A. assisted with cell cycle experiments. J.S.-A., C.T. and A.K.B. performed CUT&Tag, computational analyses and interpretation. N.S.G. provided expertise and materials for establishment of the dTAG system. S.L.L. and M.N.R. provided reagents, resources and interpreted results.

Competing interests

K.S. receives grant funding from the DFCI/Novartis Drug Discovery Program and from KronosBio, is a member of the SAB and has stock options with Auron Therapeutics and has consulted for AstraZeneca. S.L.L. declares a competing interest as a member of the advisory board for Salaris Pharmaceuticals and is a listed the inventor on US patent no. US 7,939,253 B2, 'Methods and compositions for the diagnosis and treatment of Ewing's sarcoma' and US patent no. US 8,557,532, 'Diagnosis and treatment of drug-resistant Ewing's sarcoma.' N.V.D. is an employee of Genentech, a member of the Roche Group. N.S.G. is a founder, scientific advisory board member and equity holder in Syros, C4, Allorion, Jengu, B2S, Inception, EoCys, Larkspur (board member) and Soltego (board member). The remaining authors declare no competing interests.

Additional information

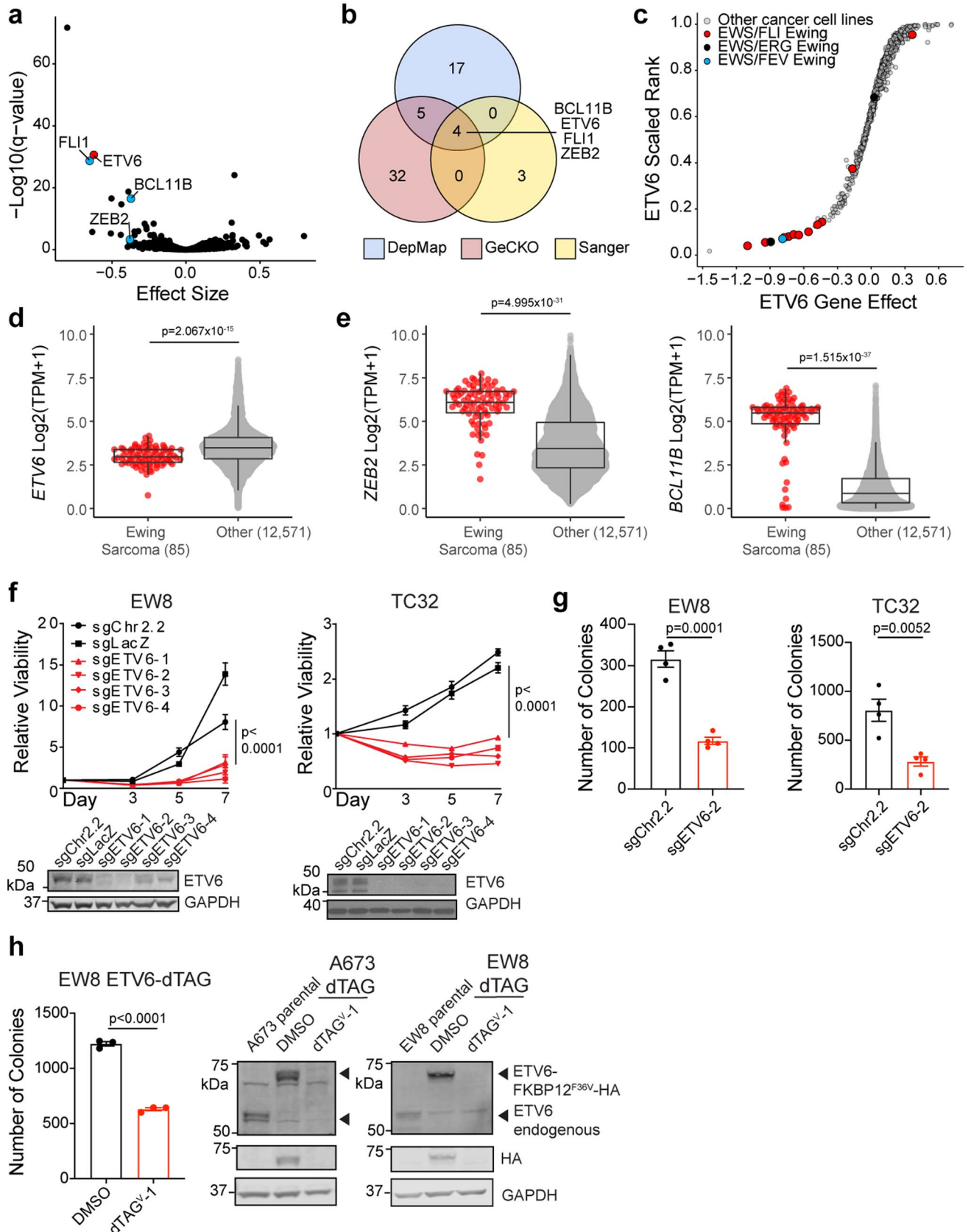
Extended data is available for this paper at <https://doi.org/10.1038/s41556-022-01059-8>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41556-022-01059-8>.

Correspondence and requests for materials should be addressed to Kimberly Stegmaier.

Peer review information *Nature Cell Biology* thanks Thomas Grunewald, Tiebang Kang, and Aykut Uren for their contribution to the peer review of this work.

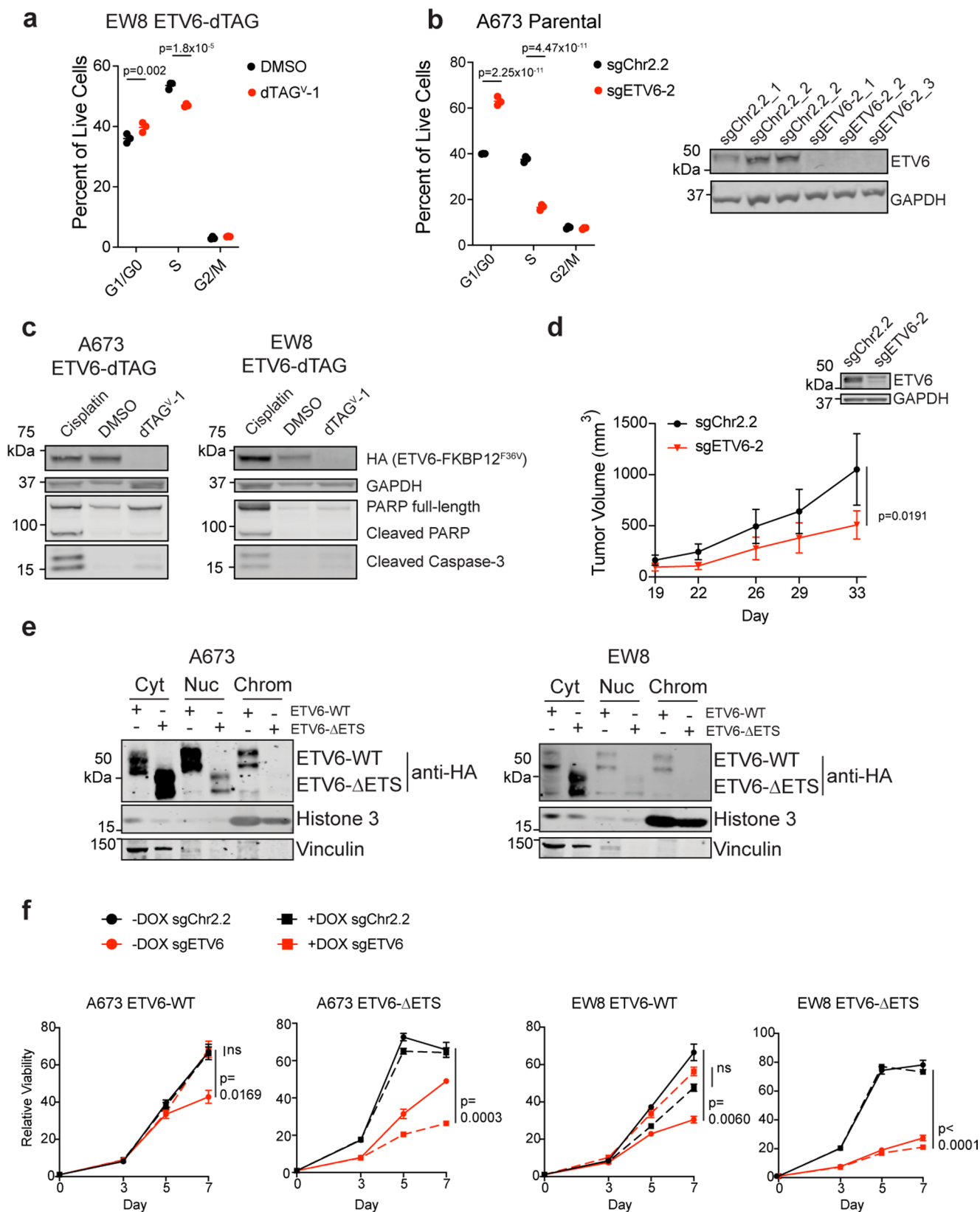
Reprints and permissions information is available at www.nature.com/reprints.



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

Extended Data Fig. 1 | The repressive ETS transcription factor, ETV6, is a selective dependency in Ewing sarcoma cells. a. Volcano plot of genes in the DepMap screen in Ewing sarcoma cell lines ($n = 14$) compared to all other cell lines ($n = 782$). Effect size (x-axis) indicates the impact of gene deletion on growth. $\text{Log}_{10}(\text{q-value})$ (y-axis) indicates specificity of dependency in Ewing sarcoma. Blue marks known selective TF dependencies. **b.** Venn diagram of Ewing sarcoma selective TF dependencies in the DepMap, GeCKO, and Sanger CRISPR/Cas9 screens. **c.** Scaled rank plot depicting all cell lines in DepMap. Gene effect (x-axis) measures *ETV6* deletion impact in each cell line. Ewing lines are enlarged and color-coded by specific EWS/ETS fusion (EWS/FLI $n = 11$, EWS/ERG $n = 2$, EWS/FEV $n = 1$). **d and e.** Expression ($\text{log}_2(\text{TPM} + 1)$; TPM, transcripts per million) of *ETV6* (**d**) and *BCL11B*, *ZEB2* (**e**) in primary tumors (Treehouse Childhood Cancer Initiative³⁷, Ewing sarcoma $n = 85$; other $n = 12,571$). Points show the full range between maxima and minima. Boxes show values for the 25th and 75th percentiles; middle line shows median (50th percentile). Whiskers extend no further than 1.5

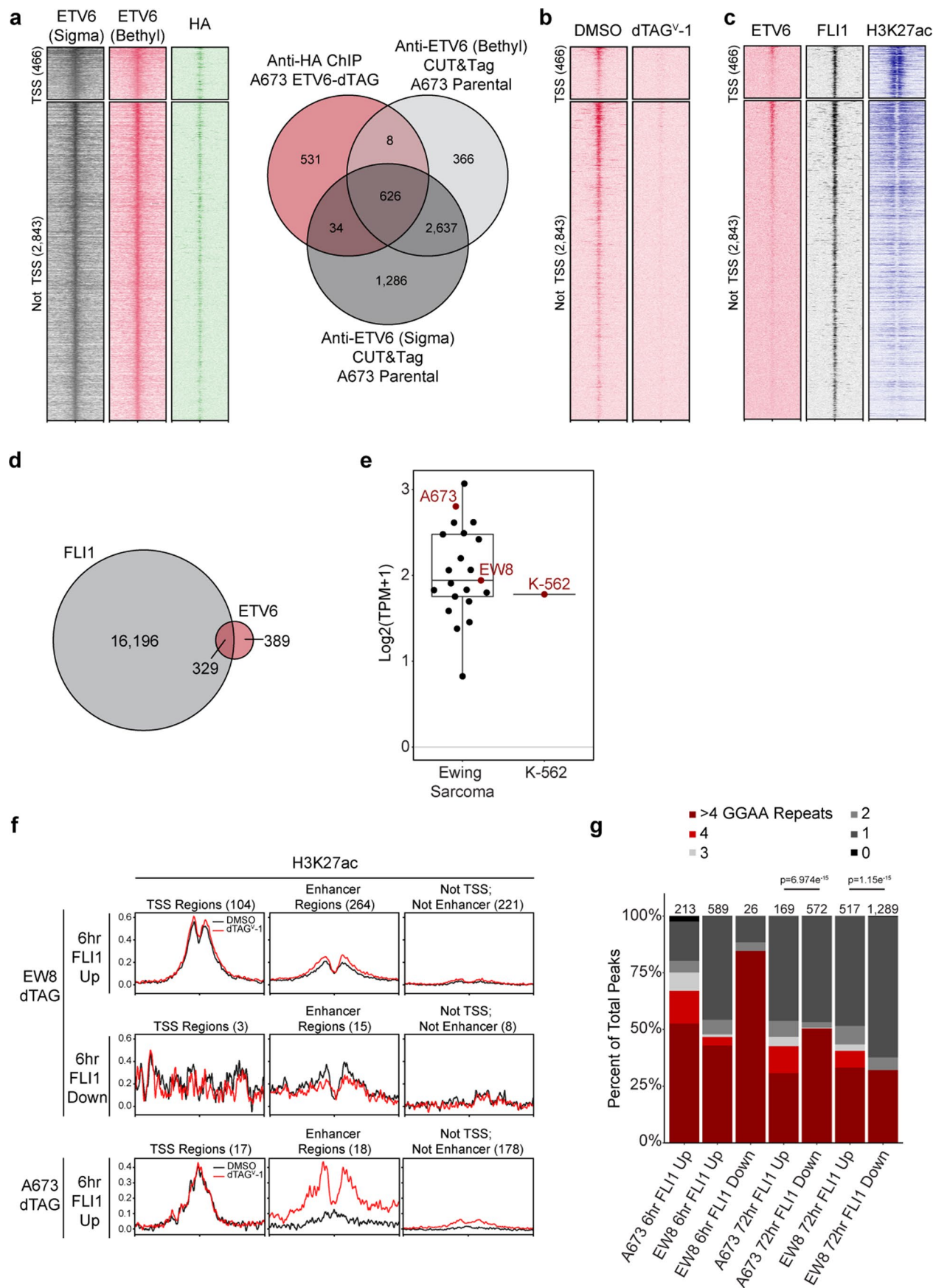
times the inter-quartile range. Gene expression in Ewing sarcoma was different from other tumor types (Welch 2-sample T-test, *ETV6* $p = 2.067 \times 10^{-15}$, *ZEB2* $p = 4.995 \times 10^{-31}$, *BCL11B* $p = 1.515 \times 10^{-37}$). **f.** Line graphs depicting mean cell viability \pm SEM in Ewing sarcoma cell lines ($n = 8$ biological replicates). *ETV6* knock-out cells exhibited lower viability than control (2-way ANOVA, Dunnett multiple comparisons EW8 $p\text{-adj} < 0.0001$; TC32 $p\text{-adj} < 0.0001$). Represents two independent experiments. Westerns show *ETV6* and GAPDH loading control (kDa, kiloDaltons). **g.** Bar plots showing mean \pm SEM number of methylcellulose cell colonies. *ETV6* knock-out samples formed fewer colonies (EW8 two-tailed t-test, $n = 4$ biological replicates, $p = 0.0001$; TC32 two-tailed t-test, $n = 4$ biological replicates, $p = 0.0052$). **h.** (Left) Bar plot showing mean \pm SEM number of EW8 *ETV6*-dTAG cell colonies. dTAG^V-1-treated cells formed fewer colonies than control ($n = 3$ biological replicates, two-tailed t-test, $p < 0.0001$). (Right) Western blot of dTAG cells shown here and in Fig. 1f.



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

Extended Data Fig. 2 | ETV6 promotes growth in Ewing sarcoma cells. **a.** Cell cycle analysis in EW8 ETV6-dTAG cells treated for 96 hours with DMSO or dTAG^V-1 (n = 3 biological replicates, two-tailed t-test, Sidak's multiple comparisons, G1/G0 p-adjust=0.002, S p-adjust=1.8 × 10⁻⁵). **b.** (Left) Cell cycle analysis in A673 cells transduced with CRISPR/Cas9 constructs (n = 3 biological replicates, two-tailed t-test, Sidak's multiple comparisons, G1/G0 p-adjust=2.25 × 10⁻¹¹, S p-adjust=4.47 × 10⁻¹¹). (Right) Western blot of ETV6. **c.** Western blot of A673 and EW8 ETV6-dTAG cells treated with 1 μM cisplatin (as a positive control) for 24 hours or with DMSO or dTAG^V-1 for 72 hours. Represents one independent experiment. **d.** (Left) Mean subcutaneous tumor volume in cubic millimeters ±SEM (n = 5, biological replicates) in mice implanted with CRISPR/Cas9-transduced TC32 cells (2-way ANOVA, p = 0.0191). (Right) Western blot showing ETV6. **e.** Western blots detecting exogenous expression of wild-type ETV6 protein (ETV6-WT) or mutant ETV6 harboring an ETS DNA binding domain deletion (ETV6-ΔETS) in

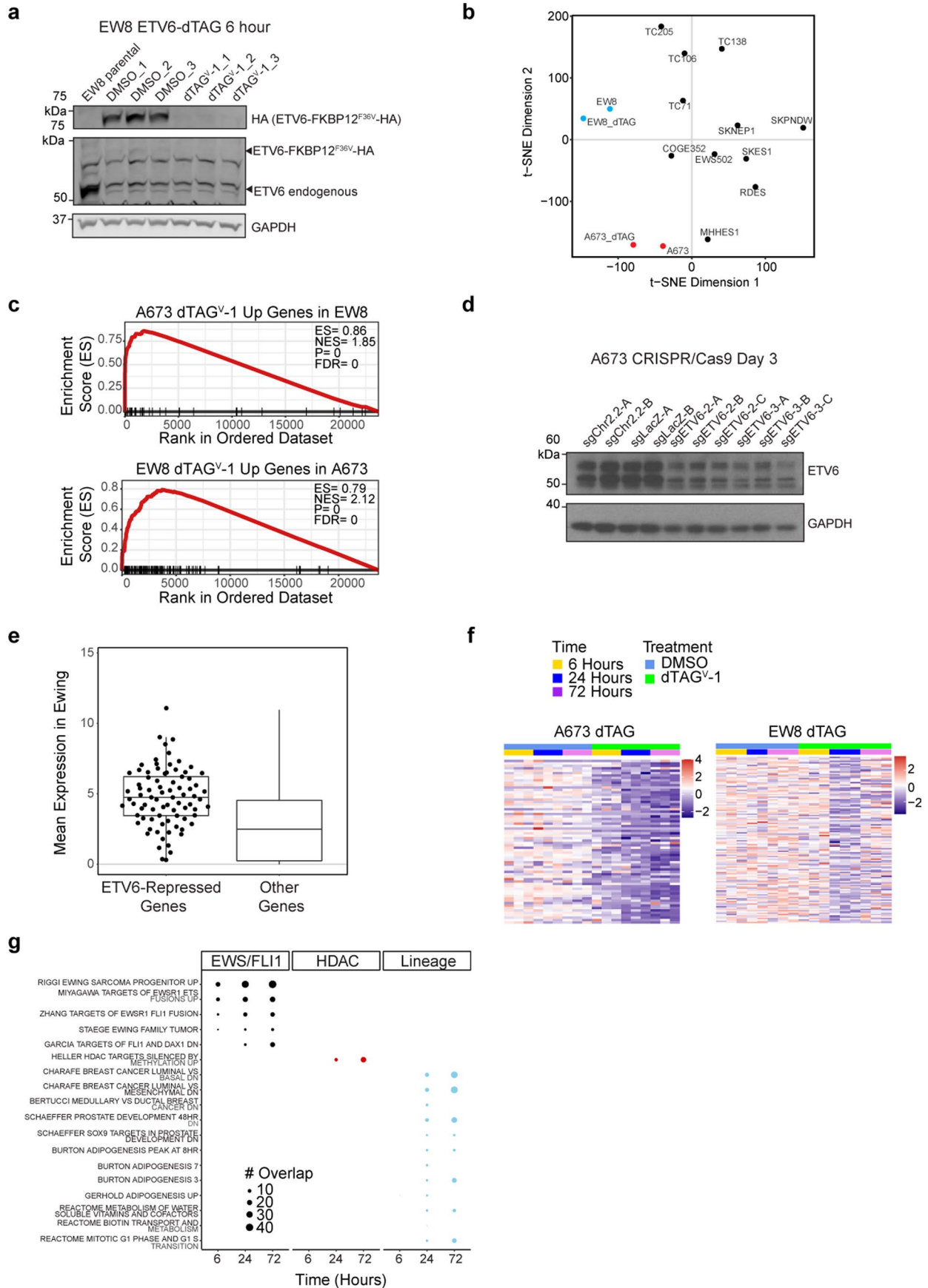
cytosolic (Cyt), nuclear (Nuc), and chromatin (Chrom) subcellular fractions in A673 (left) and EW8 (right) cell lines. Vinculin and Histone 3 demonstrate the quality of fractionation. Represents one independent experiment. **f.** Line graphs depicting mean cell viability ±SEM (n = 4 biological replicates) in A673 and EW8 cells transduced with control (sgChr2.2, black) or ETV6-targeting (sgETV6, red) CRISPR/Cas9 constructs and expressing doxycycline-induced wild-type ETV6 (ETV6-WT) or ETS DNA binding domain-deleted ETV6 (ETV6-ΔETS). Cells were treated with either vehicle (circles, solid lines) or doxycycline (squares, dashed lines). Knockout of ETV6 alone reduced cell viability (2-way ANOVA, Tukey's multiple comparisons, A673 p-adjust=0.0169; EW8 p-adjust=0.0060). ETV6 knockout in combination with wild-type ETV6 exogenous expression did not alter cell viability (ns, A673 p-adjust=0.9769; EW8 p-adjust=0.0972), but ETV6 knockout in combination with DBD-deleted ETV6 expression reduced cell viability (A673 p-adjust=0.0003; EW8 p-adjust<0.0001).



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Acute loss of ETV6 leads to increased EWS/FLI binding. **a.** (Left) Heatmaps of endogenous ETV6 CUT&Tag in parental A673 cells using two commercial antibodies and anti-ETV6-FKBP12^{F36V}-HA ChIP-seq in A673 ETV6-dTAG cells. Shown are ETV6 consensus binding sites sub-plotted by TSS overlap, ranked by height. (Right) Venn diagram showing that consensus binding sites were detected in at least 2 data sets. **b.** Heatmaps at consensus binding sites detected by anti-HA ChIP-seq in EW8 ETV6-dTAG cells at 24 hours DMSO or dTAG^V-1 treatment. **c.** (Left to right): Anti-HA ChIP-seq in EW8 ETV6-dTAG cells, anti-FLI1 ChIP-seq in parental EW8 cells, anti-H3K27ac ChIP-seq in parental EW8 cells. **d.** Venn diagram showing overlap between 718 ETV6-FKBP12^{F36V}-HA binding sites in EW8 ETV6-dTAG cells and 16,525 EWS/FLI binding sites in parental EW8 cells. **e.** Log₂(TPM + 1) expression of *ETV6* in Ewing sarcoma cell lines (n = 21)

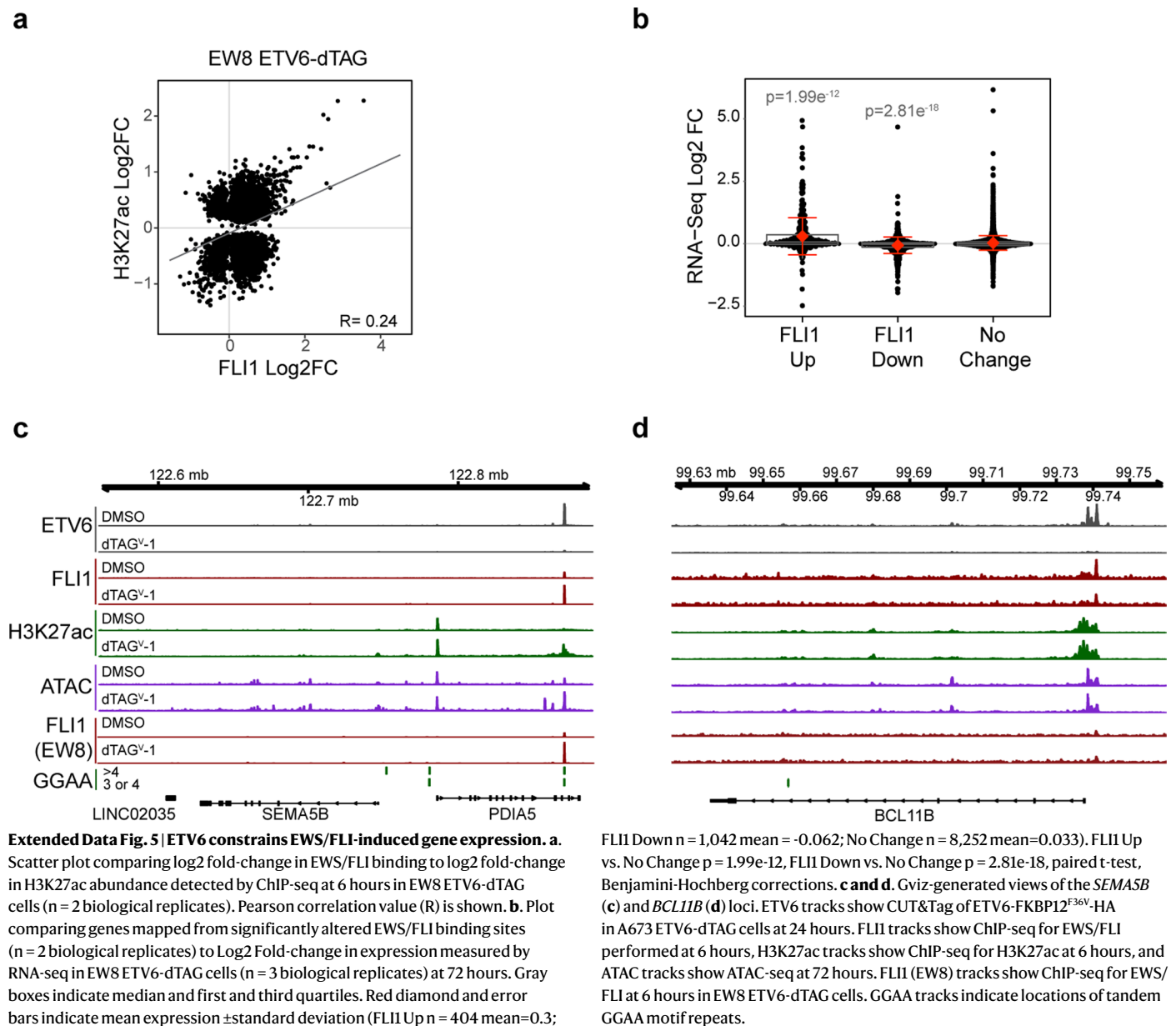
and the K-562 leukemia cell line (n = 1) from CCLE. Specific cell lines are in red. Points show the full range between maxima and minima. Boxes show values for the 25th and 75th percentiles; middle line shows median (50th percentile). Whiskers extend no further than 1.5 times the inter-quartile range. **f.** Metaplots of H3K27ac abundance at regions shown in Fig. 3g, h. **g.** Stacked column plot showing varying lengths of tandem 5'-GGAA-3' motif repeats occurring at genomic regions exhibiting significantly altered EWS/FLI binding (CSAW using the edgeR generalized linear model; p < 0.05) in A673 or EW8 dTAG cells at 72 hours. The number of peaks in each data set is listed. No genomic regions lost EWS/FLI binding in A673 dTAG at 6 hours. At 72 hours, regions that gained EWS/FLI binding were enriched for repeats of 2, 3, and 4 compared to regions that lost EWS/FLI binding (Fisher exact test, A673 p = 6.974 × 10⁻¹⁵, EW8 p = 1.15 × 10⁻¹⁵).

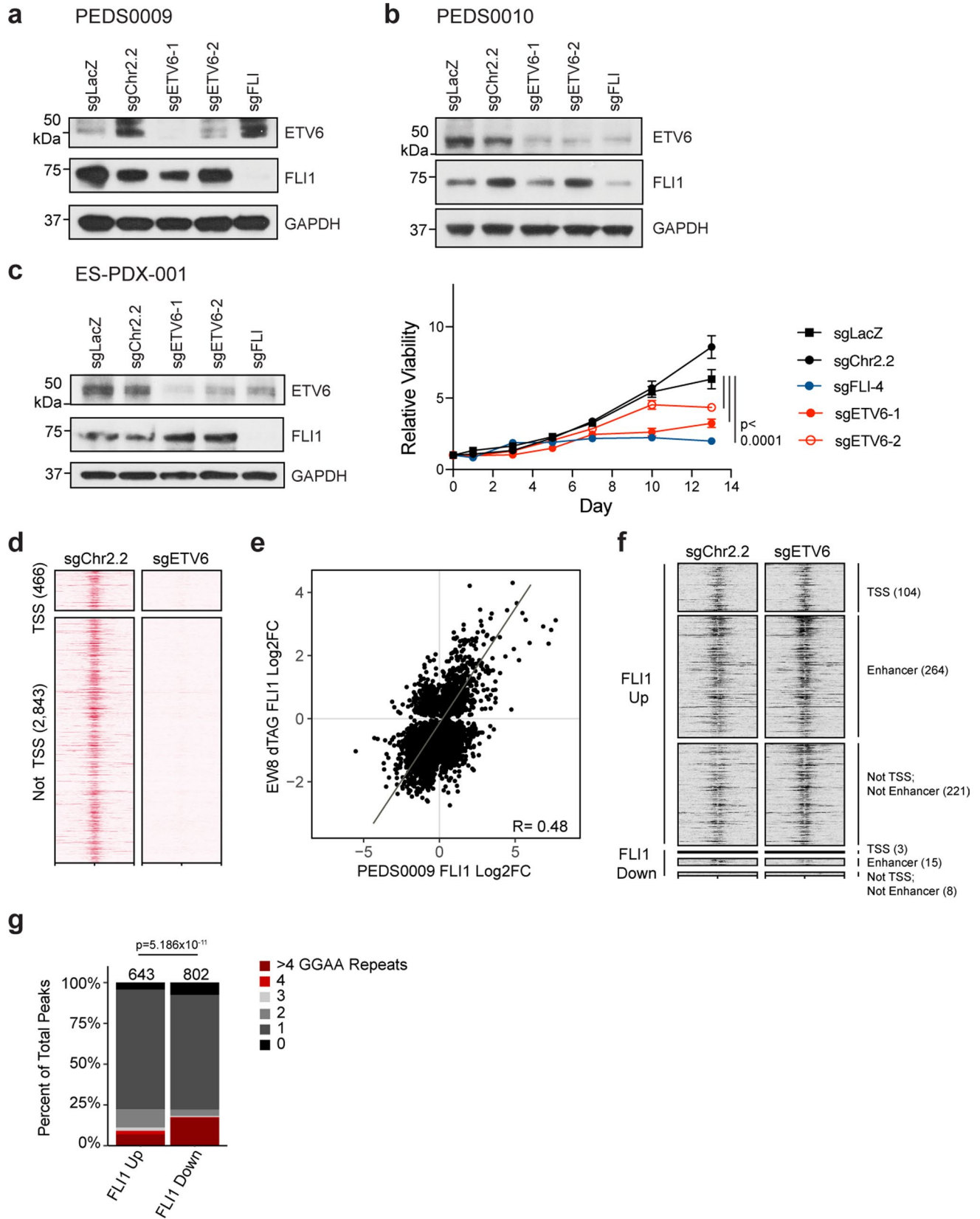


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | ETV6 is primarily a transcriptional repressor in Ewing sarcoma. a. Western blot of EW8 ETV6-dTAG cells treated for 6 hours with DMSO or dTAG^{V-1}. Represents one experiment. **b.** t-SNE of RNA-seq expression in parental Ewing sarcoma cell lines from the Cancer Cell Line Encyclopedia (CCLE)⁶⁶ and in A673 ETV6-dTAG and EW8 ETV6-dTAG cells treated for 6 hours with DMSO. EW8 parental and EW8 ETV6-dTAG cell samples are blue. A673 parental and A673 ETV6-dTAG cell samples are red. **c.** Gene set enrichment analysis (GSEA) of ETV6-repressed genes identified at 6 hours in A673 ETV6-dTAG compared to EW8 ETV6-dTAG RNA-seq data at 6 hours (top) and vice versa (bottom). **d.** Western blot of A673 parental cells transduced with sgChr2.2, sgLacZ, and sgETV6 CRISPR/Cas9 constructs. Represents one independent experiment. **e.** RNA-seq mean Log₂(TPM + 1) for 81 of 85 ETV6-repressed genes, compared to all other genes, across Ewing sarcoma cell lines in CCLE⁶⁶. Boxes show the values for the 25th and 75th percentiles; middle line shows the median

(50th percentile). Whiskers extend up from the 75th percentile and down from the 25th percentile, no further than 1.5*IQR (where IQR is the inter-quartile range, or distance between the 25th and 75th percentiles). Points show the full range between maxima and minima. **f.** RNA-seq heatmap of ETV6-activated core enrichment genes, identified by GSEA described in Fig. 3e, in A673 (68 genes) and EW8 dTAG cells (126 genes). **g.** Combined enrichment plot of MSigDB c2 pathways enriched in ETV6-activated genes in both dTAG models (24-hour RNA-seq, hypergeometric enrichment test; $p < 0.05$). Gene sets are ranked by significance; missing dots indicate insignificance. Dot size indicates the number of ETV6-activated genes at 6, 24, and 72 hours also in the gene set (33, 130, and 543 genes, respectively). 'EWS/FLI', 'HDAC', and 'Lineage' gene sets characterize EWS/FLI-regulated genes, histone deacetylase enzyme-regulated genes, and genes exhibiting tissue-specific expression, respectively.

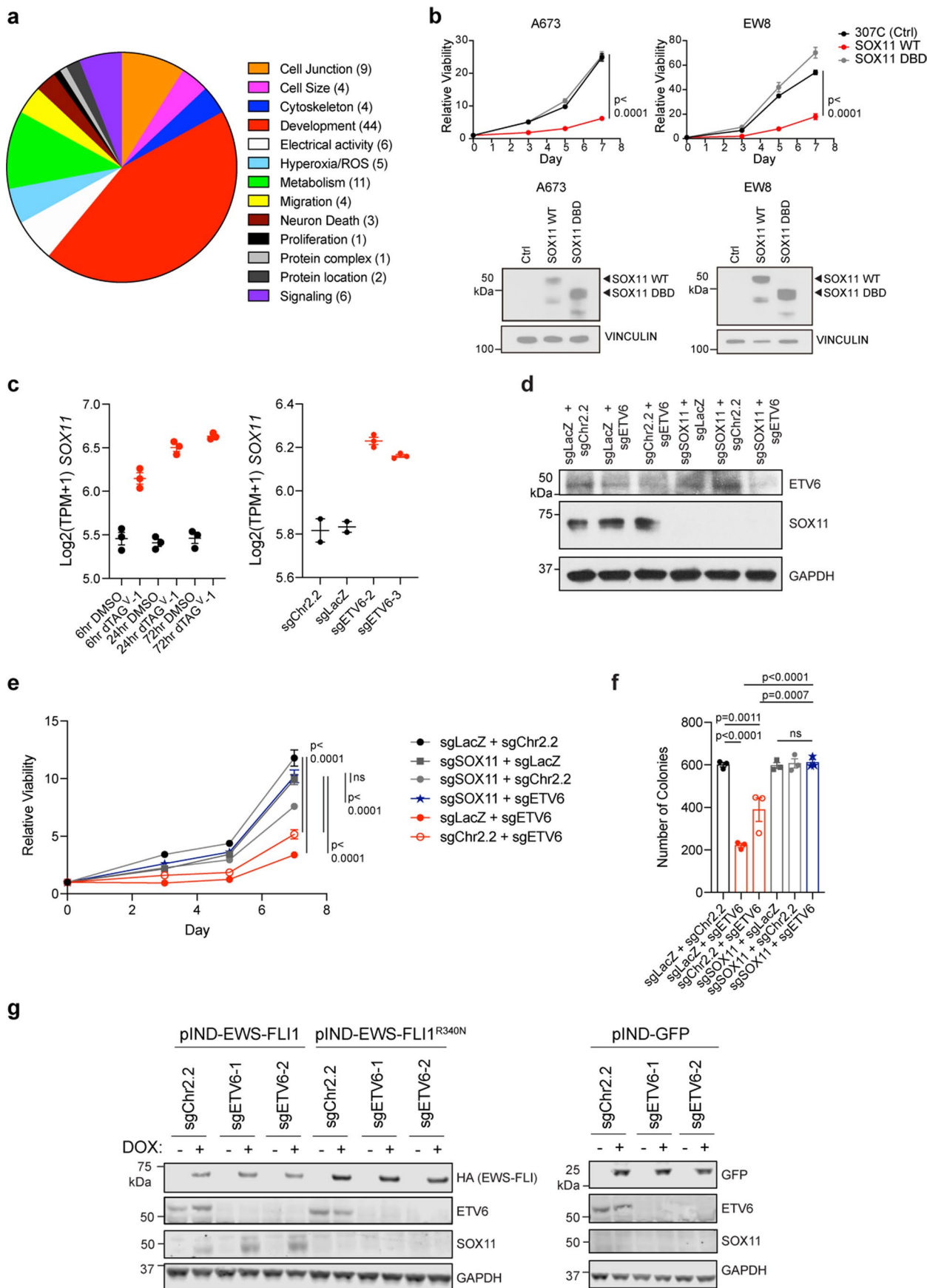




Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | ETV6 competes with EWS/FLI for binding in clinically relevant Ewing sarcoma models. a and b. Western blot of cells from the newly derived Ewing sarcoma cell lines, PEDS0009 (**a**) and PEDS0010 (**b**), transduced with CRISPR/Cas9 constructs. Represents one independent experiment. **c.** (Left) Western blot of cells from the Ewing sarcoma patient-derived xenograft, ES-PDX-001, transduced with CRISPR/Cas9 constructs. (Right) Line graphs depicting mean cell viability \pm SEM *in vitro* (n = 6 biological replicates). Knockout of *ETV6* and *EWS/FLI* reduced cell growth compared to sgChr2.2 control (2-way ANOVA, Tukey's multiple comparisons, p-adj = <0.0001). **d.** Heatmaps showing 3-kilobase (kb) windows centered at 3,309 consensus ETV6 binding sites, sub-plotted by overlap within 2.5 kb of transcription start sites (TSS). Shown are ETV6 peaks detected in sgChr2.2 control and *ETV6*-knockout PEDS0009 cells profiled

by CUT&RUN. **e.** Scatter plots of log₂ fold-change in EWS/FLI binding in EWS *ETV6*-dTAG cells following 72 hours of treatment with DMSO or dTAG^V-1 (y-axis) and in control and *ETV6*-knockout PEDS0009 cells profiled by FLII CUT&RUN (x-axis). Pearson correlation value (R) is shown. **f.** Heatmaps of FLII CUT&RUN performed in control and *ETV6*-knockout PEDS0009 cells. Loci shown were defined in Fig. 2g as regions that exhibited increased EWS/FLI binding upon *ETV6* loss in EWS *ETV6*-dTAG cells. **g.** Stacked column plot showing varying lengths of tandem 5'-GGAA-3' motif repeats occurring at genomic regions exhibiting significantly (CSAW; p < 0.05) increased (FLII Up) or decreased (FLII Down) EWS/FLI occupancy in PEDS0009 cells upon *ETV6* knockout. FLII Up regions were more likely than FLII Down regions to contain GGAA repeats of 2, 3, or 4 (Fisher Exact test, p = 5.186e-11).



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

Extended Data Fig. 7 | Knock-out of the *ETV6*-repressed gene, *SOX11*, rescues the phenotype of *ETV6* loss. **a. Top 100 significantly enriched MSigDB c5 Gene Ontology gene set categories in *ETV6*-repressed genes (24-hour RNA-seq, A673 *ETV6*-dTAG, parentheses show gene set number; full list in Supplementary Table 19). **b.** Line graphs depicting mean viability \pm SEM (n = 8 biological replicates) in A673 (left) and EW8 (right) cells exogenously expressing wild-type *SOX11* (*SOX11* WT, red), DBD-deleted mutant *SOX11* (*SOX11* DBD, gray), or empty pLX_TRC307 vector control (307 C, black). Wild-type *SOX11* expression reduced viability compared to control (2-way ANOVA, Dunnett's multiple comparisons, p -adj < 0.0001). Western blots show *SOX11* and Vinculin loading control. **c.** RNA-seq $\log_2(\text{TPM} + 1) \pm$ SEM of *SOX11* expression in A673 *ETV6*-dTAG cells (left) (n = 3 biological replicates) and in CRISPR/Cas9-perturbed parental A673 cells (right) (controls n = 2 biological replicates; *ETV6*-knockout n = 3 biological replicates). **d.** Western blot of CRISPR/Cas9-perturbed A673 cells. **e.** Mean viability \pm SEM (n = 6 biological replicates) of A673 cells shown in **d** *in vitro*. *ETV6* knockout (red open and closed circles) reduced viability compared to control (black circles)**

(2-way ANOVA, Tukey's multiple comparisons, p -adj < 0.0001). Simultaneous *ETV6* and *SOX11* knockout (blue star) did not reduce viability compared to *SOX11* knockout alone (gray square and circle) (sgSOX11 + sgETV6 vs. sgSOX11 + sgLacZ, not significant 'ns' p -adj = 0.7343) and exhibited greater viability than *ETV6* knockout alone (p -adj < 0.0001). **f.** Mean methylcellulose colony number \pm SEM (n = 3 biological replicates) formed by A673 cells shown in **d** and **e**. *ETV6* knockout reduced colony number compared to control (2-way ANOVA, Tukey's multiple comparisons, sgLacZ + sgETV6 vs. sgLacZ + sgChr2.2 p -adj < 0.0001, sgChr2.2 + sgETV6 vs. sgLacZ + sgChr2.2 p -adj = 0.0011). Simultaneous *ETV6* and *SOX11* knockout did not reduce colony number compared to *SOX11* knockout alone (sgSOX11 + sgLacZ vs. sgSOX11 + sgETV6, ns, p -adj = 0.9984) and increased colonies compared to *ETV6* knock-out alone (sgSOX11 + sgETV6 vs. sgLacZ + sgETV6 p -adj < 0.0001; vs. sgChr2.2 + sgETV6 p -adj = 0.0007). **g.** Western blot of rhabdomyosarcoma RD cells expressing doxycycline-inducible HA-tagged GFP, wild-type EWS/FLI, or DNA binding-incompetent R340N mutant EWS/FLI, in combination with CRISPR/Cas9 perturbation. One independent experiment.

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) 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

Data analysis

Microsoft Excel Version 16.50
 GraphPad Prism Version 9.0.0
 FlowJo_v10.6.1
 ImageQuant TL 8.2 (Cytiva)
 ImageStudioLite 5.2.5
 QuantStudio Real-Time PCR Software v1.7

ChIP-seq analysis: bcl2fastq (v2.17), FastQC (<http://www.bioinformatics.babraham.ac.uk>) (v0.11.9), Trimmomatic (v0.36), Bowtie2 (v2.3.5 and v2.3.4.3), SamTools (v1.9.0), Picard-tools (v2.19.0), MACS2 (version 2.1.4), deeptools (v3.3.0)

RNA-seq analysis: FastQC (<http://www.bioinformatics.babraham.ac.uk>) (v0.11.9), STAR (v2.7.2b), RSEM (v1.3.1). The RNA-seq processing pipeline was roughly modeled on the GTEx pipeline (<https://github.com/broadinstitute/gtex-pipeline/>).

RStudio Version 1.2.5033
 R version 4.0.4 (2021-02-15)
 R packages: annotate_1.68.0, AnnotationDbi_1.52.0, askpass_1.1, assertthat_0.2.1, beeswarm_0.3.1, Biobase_2.50.0, BiocFileCache_1.14.0, BiocGenerics_0.36.0, BiocParallel_1.24.1, biomaRt_2.46.3, Biostrings_2.58.0, bit_4.0.4, bit64_4.0.5, bitops_1.0-6, blob_1.2.1, cachem_1.0.4, caTools_1.18.1, colorspace_2.0-0, compiler_4.0.4, cowplot_1.1.1, crayon_1.4.1, curl_4.3, data.table_1.14.0, DBI_1.1.1, dbplyr_2.1.0, DelayedArray_0.16.2, DESeq2_1.30.1, digest_0.6.27, DO.db_2.9, DOSE_3.16.0, dplyr_1.0.5, ellipsis_0.3.1, fansi_0.4.2, farver_2.1.0, fastmap_1.1.0, fastmatch_1.1-0, fgsea_1.16.0, genefilter_1.72.1, geneplotter_1.68.0, generics_0.1.0, GenomeInfoDb_1.26.4, GenomeInfoDbData_1.2.4, GenomicAlignments_1.26.0, GenomicFeatures_1.42.2, GenomicRanges_1.42.0, ggbeeswarm_0.6.0, ggplot2_3.3.3,

ggrepel_0.9.1, glue_1.4.2, GO.db_3.12.1, GOSemSim_2.16.1, gplots_3.1.1, graph_1.68.0, grid_4.0.4, gridExtra_2.3, GSEABase_1.52.1, gtable_0.3.0, gtools_3.8.2, hms_1.0.0, httr_1.4.2, IRanges_2.24.1, KernSmooth_2.23-18, lattice_0.20-41, lifecycle_1.0.0, locfit_1.5-9.4, magrittr_2.0.1, Matrix_1.3-2, MatrixGenerics_1.2.1, matrixStats_0.58.0, memoise_2.0.0, munsell_0.5.0, openssl_1.4.3, pheatmap_1.0.12, pillar_1.5.1, pkgconfig_2.0.3, plyr_1.8.6, prettyunits_1.1.1, progress_1.2.2, purrr_0.3.4, qvalue_2.22.0, R6_2.5.0, rappdirs_0.3.3, RColorBrewer_1.1-2, Rcpp_1.0.6, RCurl_1.98-1.3, reshape2_1.4.4, rlang_0.4.10, Rsamtools_2.6.0, RSQLite_2.2.4, rstudioapi_0.13, rtracklayer_1.50.0, S4Vectors_0.28.1, scales_1.1.1, splines_4.0.4, stringi_1.5.3, stringr_1.4.0, SummarizedExperiment_1.20.0, survival_3.2-9, tibble_3.1.0, tidyselct_1.1.0, tools_4.0.4, utf8_1.2.1, vctrs_0.3.6, vipor_0.4.5, XML_3.99-0.6, xml2_1.3.2, xtable_1.8-4, XVector_0.30.0, zlibbioc_1.36.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

CRISPR/Cas9 screen data and the genomic characterization of cancer cell lines (RNA-seq) used in this study are publicly available at [depmap.org](#). Gene expression data from the Treehouse Childhood Cancer Initiative characterizing primary tumors is publicly available at <https://treehousegenomics.soe.ucsc.edu/public-data/>. Gene set pathway enrichment analysis was performed with signatures from version 6.0 of the Broad Institute's molecular signature database (MSigDB) (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>). All genomics data is uploaded to Gene Expression Omnibus (GSE181554).

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	For CRISPR-Cas9 screening of cancer cell lines, no sample size was predetermined as the goal was to screen all possible human cancer cell lines. At this time, data for 796 human cancer cell lines were used. For other experiments, no statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications. For all low-throughput validation experiments, studies were performed with a minimum of triplicate replicates. For the animal experiments, three to five animals were analyzed per condition based on lab experience. qPCR analysis was performed on two separate biological replicate cell samples, each analyzed in technical triplicate. Sample sizes are indicated in figure legends.
Data exclusions	From the DepMap analysis, 12 cell lines were not included in the subsequent analyses for the following reasons. Four cell lines are classified as engineered lines. The origin of one cell line, CHLA57, is unknown as it is incorrectly identified as an Ewing sarcoma line. Seven cell lines, including the Ewing sarcoma cell line, SKNMC, were removed because they were listed as commonly misidentified cell lines in the ICLAC Register of Misidentified Cell Lines (https://iclac.org/databases/cross-contaminations/). Therefore, dependency data for 796 cell lines were examined.
Replication	For CRISPR-Cas9 screening of cancer cell lines, replicate information is available at depmap.org . Figure legends indicate the nature of replicates shown as well as independently repeated experiments.
Randomization	Not applicable; animal studies did not involve drug treatment.
Blinding	Not applicable; animal studies did not involve drug treatment.

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

Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Anti-GAPDH (#2118S, rabbit, monoclonal, clone 14C10, Cell Signaling Technology)
 Anti-ETV6 (#WH0002120M1-100UG, mouse, monoclonal, clone 3B10, Sigma Aldrich; #SC-166835, mouse, monoclonal, clone E-1, Santa Cruz Biotechnology; #A303-674, rabbit, polyclonal, Bethyl)
 Anti-H3K27ac (#ab4729, rabbit, polyclonal, Abcam)
 Anti-HA (#3724S, rabbit, monoclonal, clone C29F4, Cell Signaling Technology)
 Anti-PARP (#9542S, rabbit, polyclonal, Cell Signaling Technology)
 Anti-cleaved Caspase-3 (#9664S, rabbit, monoclonal, clone 5A1E, Cell Signaling Technology)
 Anti-FLI1 (#ab15289, rabbit, polyclonal, Abcam)
 Anti-FAS (#SC-8009, mouse, monoclonal, clone B-10, Santa Cruz Biotechnology)
 Anti-SEMA5B (#PA5113369, rabbit, polyclonal, Thermo Fisher Scientific)
 Anti-BCL11B (#12120S, rabbit, monoclonal, clone D6F1, Cell Signaling Technology)
 Anti-SOX11 (# 58207S, rabbit, monoclonal, clone E7G9N, Cell Signaling Technology)
 Horseradish peroxidase anti-mouse (Cell Signaling Technology, #7076S)
 Horseradish peroxidase anti-rabbit (Cell Signaling Technology, #7074S)
 IRDye anti-mouse (LICOR Biosciences, #926-32210)
 IRDye anti-rabbit (LICOR Biosciences, #926-68071)

Validation

All antibodies were reportedly validated by the manufacturer. Anti-PARP and anti-cleaved Caspase-3 were validated by the manufacturer by treating cells with DNA-damaging agents. Similarly, in this study, cisplatin-treated cells were used as a positive control validating these antibodies (ED Fig 2c). Antibodies validated by target knock-out and western blot with data shown in this paper include anti-FLI1 (Fig 4h; ED Fig 6a/b/c), anti-SOX11 (Fig 6d; ED Fig 7d). Anti-ETV6 was validated in the same manner (Fig 1b/e/h; ED Fig 1f; ED Fig 2b; ED Fig 4a/d; ED Fig 6a/b/c; ED Fig 7g) as well as by CUT&Run (ED Fig 6d). Anti-HA was validated by western blot and ChIP-seq comparing DMSO vs. dTAG-treated cells in ETV6-dTAG cells (eg, Fig 1e, ED Fig 1h, ED Fig 3b). Additionally, we validated antibodies for the following proteins by knock-out and western blot, though this data is not shown in this paper: BCL11B, SEMA5B, FAS.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

All cell lines used for the genome-scale CRISPR-Cas9 screen are detailed in Meyers RM et al., Nature Genetics 2017, and on depmap.org. A673 cells were purchased from ATCC (CRL-1598). EW8 and TC32 were a kind gift from the Golub Lab (Broad Institute). EW8 cells were originally established in the Houghton Lab (Greehey Children's Cancer Research Institute, San Antonio, TX). TC-32 cells were originally obtained from Children's Oncology Group (COG) Childhood Cancer Repository. The PEDS0009 and PEDS0010 cell lines were obtained from the Cancer Cell Line Factory (Broad Institute, Cambridge, MA). NO human studies were performed. In Extended Data Fig. 6c, a minimally passaged Ewing sarcoma cell line previously derived in our laboratory (Seong et al., 2021) from a previously characterized patient-derived xenograft (PDX) (Garcia-Dominguez et al., 2018) (HSJD-ES-PDX-001) was studied. (As such, this experiment was performed in vitro and did not involve the use of animals.) As previously described (Garcia-Dominguez et al., 2018), this PDX originated from a scapular biopsy in a 21.7 year-old patient whose sex was not reported. It was collected with informed consent without compensation under an Institutional Review Board-approved protocol at Sant Joan de Déu Hospital (HSJD, Barcelona, Spain), animal protocol number HSJD 135/1165, and provided to our laboratory by Dr. Jaume Mora. The Rhabdomyosarcoma RD cell line was a gift from the DepMap group at the Broad Institute and was originally obtained from ATCC (catalog no. CRL-7731).

Authentication

All cell lines used were STR tested for identity at either the Dana-Farber Cancer Institute molecular diagnostics core facility or The Broad Institute of MIT and Harvard.

Mycoplasma contamination

All cells were tested for mycoplasma and confirmed negative.

Commonly misidentified lines
(See [ICLAC](#) register)

All ICLAC lines were removed from the analysis of the DepMap CRISPR/Cas9 dependency data, including the Ewing sarcoma cell line, SKNMC.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	The study shown in Fig. h/i used 7 week-old female NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice ordered from Jackson Laboratory. The study shown in Extended Data Fig. 2d used 12 week-old male NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice ordered from Jackson Laboratory. The study shown in Fig. 6h used 6-8 week-old female CrTac:NCR-Foxn1<nu> (nude) mice from Taconic Biosciences. Mice were housed at Dana-Farber Cancer Institute with regulated temperature and humidity and standard 12 hour light/dark cycles.
Wild animals	No wild animals were used.
Field-collected samples	No field-collected samples were used.
Ethics oversight	All of our mouse studies were approved by the DFCI IACUC committee and were performed in accordance with NIH guidelines for the humane care and use of animals. Animal Welfare Assurance Number: D16-00010 (A3023-01).

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.

We have deposited all new genomics data to GEO (GSE181554) for publication.

Files in database submission

GSM5505948
 GSM5505949
 GSM5505950
 GSM5505951
 GSM5505952
 GSM5505953
 GSM5505954
 GSM5505955
 GSM5505956
 GSM5505957
 GSM5505958
 GSM5505959
 GSM5505960
 GSM5505961
 GSM5505962
 GSM5505963
 GSM5505964
 GSM5505965
 GSM5505966
 GSM5505967
 GSM5505968
 GSM5505969
 GSM5505970
 GSM5505971
 GSM5505972
 GSM5505973
 GSM5505974
 GSM5505975
 GSM5505976
 GSM5505977
 GSM5505978
 GSM5505979
 GSM5505980
 GSM5505981
 GSM5505982
 GSM5505983
 GSM5505984
 GSM5505985
 GSM5505986
 GSM5505987
 GSM5505988
 GSM5505989
 GSM5505990
 GSM5505991

GSM5505992
GSM5505993
GSM5505994
GSM5505995
GSM5505996
GSM5505997
GSM5505998
GSM5505999
GSM5506000
GSM5506001
GSM5506002
GSM5506003
GSM5506004
GSM5506005
GSM5506006
GSM5506007
GSM5506008
GSM5506009
GSM5506010
GSM5506011
GSM5506012
GSM5506013
GSM5506014
GSM5506015
GSM5506016
GSM5506017
GSM5506018
GSM5506019
GSM5506020
GSM5506021
GSM5506022
GSM5506023
GSM5506024
GSM5506025
GSM5506026
GSM5506027
GSM5506028
GSM5506029
GSM5506030
GSM5506031
GSM5506032
GSM5506033
GSM5506034
GSM5506035
GSM5506036
GSM5506037
GSM5506038
GSM5506039
GSM5506040
GSM5506041
GSM5506042
GSM5506043
GSM5506044
GSM5506045
GSM5775771
GSM5775772
GSM5775773
GSM5775774
GSM5775775
GSM5775776
GSM5775777
GSM5775778
GSM5775779
GSM5775780
GSM5775781
GSM5775782
GSM5775783
GSM5775784
GSM5775785
GSM5775786
GSM5775787
GSM5775788
GSM6503406
GSM6503407
GSM6503408
GSM6503409

GSM6503410
 GSM6503411
 GSM6503412
 GSM6503413
 GSM6503414
 GSM6503415
 GSM6503416
 GSM6503417
 GSM6503418
 GSM6503419

Genome browser session
 (e.g. [UCSC](#))

Not applicable.

Methodology

Replicates	At least 2 replicates were used for all ChIP-seq experiments.
Sequencing depth	All sequencing was done to a minimum depth of 20 million reads per sample.
Antibodies	ChIP-seq primary antibodies: Anti-H3K27ac (Abcam, # 4729) Anti-HA (Abcam, # ab9110) Anti-FLI1 (Abcam, # ab15289) Spike-in Anti-Drosophila histone-variant (Active Motif, # 61686)
Peak calling parameters	ChIP-seq peaks were called using MACS2 with the False Discovery Rates (FDR) $q < 0.01$ unless otherwise stated.
Data quality	All ChIP-seq experiments were performed in duplicate to ensure rigor. The raw Illumina sequencer output was converted to FASTQ format using the program bcl2fastq (v2.17). Sequencing read quality was examined using FastQC (http://www.bioinformatics.babraham.ac.uk) (v0.11.9). Trimming of low-quality reads and clipping of sequencing adapters was done using the program Trimmomatic (v0.36)77 and all reads shorter than 40bp after trimming were dropped. Reads were aligned to the human genome (hg19) using Bowtie2 (v2.3.5)78,79 using the '—very_sensitive' preset collection of parameters. Bam to Sam file conversion was done with SamTools (v1.9q)80 and duplicate reads were removed using Picard-tools (v2.19.0) (http://picard.sourceforge.net). ChIP-seq peaks were called using MACS281 with the False Discovery Rates (FDR) $q < 0.01$ unless otherwise stated. The MACS2 algorithm utilizes a dynamic Poisson distribution to capture local biases in the genomic sequence, which allows for a sensitive and robust prediction of peaks. Unless otherwise noted, peaks were assigned to the closest gene within +/- 400kb using the ChIPseeker package in R82. Visualizations of the ChIP-seq data tracks were produced with the R Bioconductor Gviz package83.
Software	All ChIP-seq experiments were performed in duplicate to ensure rigor. The raw Illumina sequencer output was converted to FASTQ format using the program bcl2fastq (v2.17). Sequencing read quality was examined using FastQC (http://www.bioinformatics.babraham.ac.uk) (v0.11.9). Trimming of low-quality reads and clipping of sequencing adapters was done using the program Trimmomatic (v0.36)77 and all reads shorter than 40bp after trimming were dropped. Reads were aligned to the human genome (hg19) using Bowtie2 (v2.3.5)78,79 using the '—very_sensitive' preset collection of parameters. Bam to Sam file conversion was done with SamTools (v1.9q)80 and duplicate reads were removed using Picard-tools (v2.19.0) (http://picard.sourceforge.net). ChIP-seq peaks were called using MACS281 with the False Discovery Rates (FDR) $q < 0.01$ unless otherwise stated. The MACS2 algorithm utilizes a dynamic Poisson distribution to capture local biases in the genomic sequence, which allows for a sensitive and robust prediction of peaks. Unless otherwise noted, peaks were assigned to the closest gene within +/- 400kb using the ChIPseeker package in R82. Visualizations of the ChIP-seq data tracks were produced with the R Bioconductor Gviz package83.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cell cycle analysis was performed using the Click-iT™ Plus EdU Alexa Fluor™ 647 Flow Cytometry Assay Kit (Life Technologies, #C10424). Cells were analyzed per kit instructions with minor modifications. Cells were pulsed with 10 μ M of the modified nucleotide analogue, EdU (5-ethynyl-2'-deoxyuridine), by spiking the molecule provided at stock concentration (10 mM) into the existing media for 90 minutes in a tissue culture incubator. According to kit instructions, 1-2 million cells per sample were trypsinized, washed, fixed, permeabilized, and then treated with a reaction cocktail containing Alexa Fluor-647-conjugated picolyl azide to label incorporated EdU. Following this step, cells were washed and then stained with an RNase-containing propidium iodide solution (Cell Signaling, #4087S) for 45 minutes in a tissue culture incubator. Lastly, cells were analyzed by flow cytometry at 5,000-10,000 cells per sample. Live cells were gated using FSC-A and SSC-A. dTAG cell samples were seeded in triplicate and treated separately with DMSO or dTAGV-1 prior to EdU pulsing. A673 ETV6 knock-out cells were
--------------------	--

	seeded and cultured separately prior to EdU pulsing. In all experiments, additional cells were collected from each sample for western blot validation.
Instrument	BD FacsCelesta
Software	FlowJo_v10.6.1
Cell population abundance	This experiment did not involve sorting.
Gating strategy	A figure exemplifying gating strategy has been provided separately and demonstrates the following: Live cells were gated using FSC-A and SSC-A. G1/G0 cells were gated as cells exhibiting relatively less propidium iodide and less Alexa Fluor-647 staining. G2/M cells were gated as cells exhibiting relatively high propidium iodide and less AF-647 staining. S cells were gated as cells exhibiting relatively high AF-647 and intermediate propidium iodide staining.

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