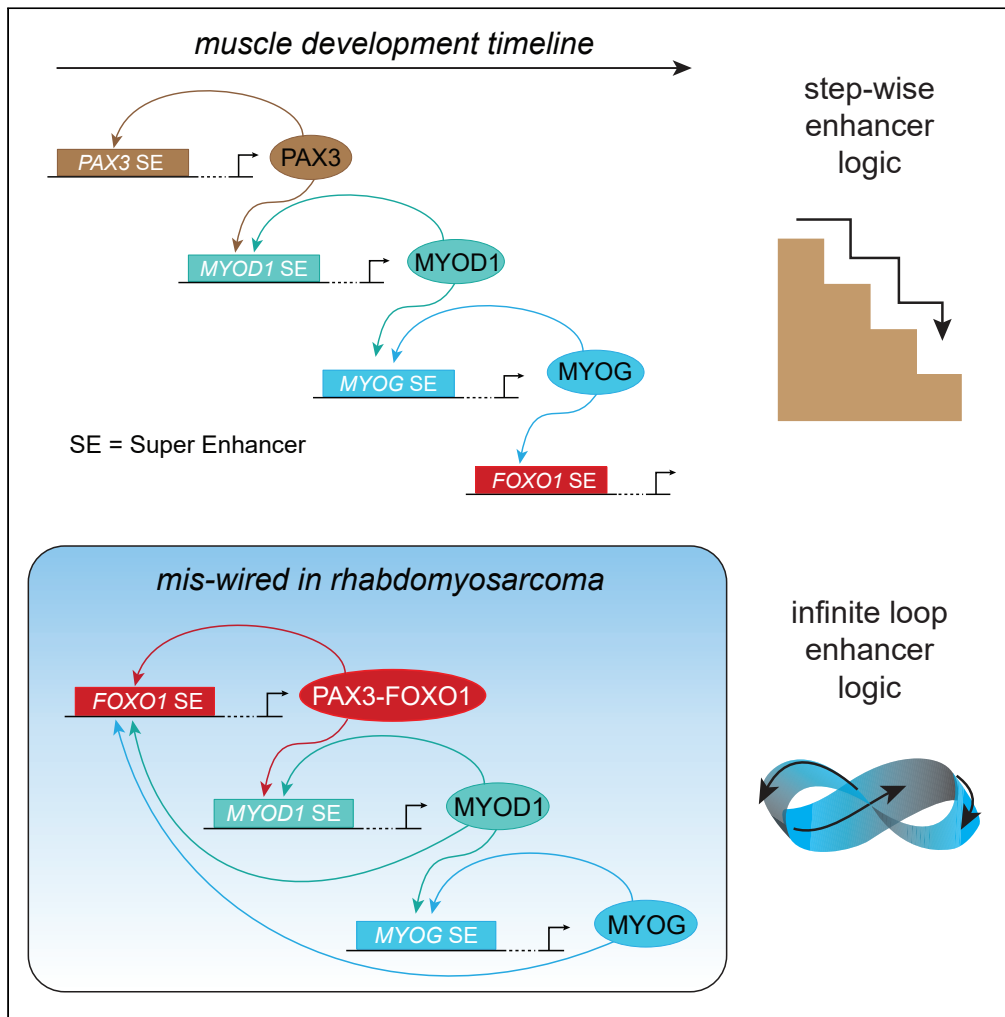


Article

# Miswired Enhancer Logic Drives a Cancer of the Muscle Lineage



Berkley E. Gryder,  
Marco Wachtel,  
Kenneth  
Chang, ..., Beat  
Schäfer,  
Christopher R.  
Vakoc, Javed Khan

berkley.gryder@nih.gov  
(B.E.G.)  
khanjav@mail.nih.gov (J.K.)

**HIGHLIGHTS**

Muscle lineage enhancer logic is miswired in PAX3-FOXO1-driven rhabdomyosarcoma

PAX7-FOXO1, PAX3-NCOA1, and PAX3-INO80D tumors show evidence of miswired enhancers

FOXO1 super enhancer is myogenically activated and sustains PAX3-FOXO1 expression

Oncogenic translocations select for miswired enhancers around disordered proteins



## Article

Miswired Enhancer Logic Drives  
a Cancer of the Muscle Lineage

Berkley E. Gryder,<sup>1,\*</sup> Marco Wachtel,<sup>2</sup> Kenneth Chang,<sup>3</sup> Osama El Demerdash,<sup>3</sup> Nicholas G. Aborenden,<sup>1</sup> Wardah Mohammed,<sup>1</sup> Winston Ewert,<sup>4</sup> Silvia Pomella,<sup>5</sup> Rossella Rota,<sup>5</sup> Jun S. Wei,<sup>1</sup> Young Song,<sup>1</sup> Benjamin Z. Stanton,<sup>6</sup> Beat Schäfer,<sup>2</sup> Christopher R. Vakoc,<sup>3</sup> and Javed Khan<sup>1,7,\*</sup>

## SUMMARY

**Core regulatory transcription factors (CR TFs) establish enhancers with logical ordering during embryogenesis and development. Here we report that in fusion-positive rhabdomyosarcoma, a cancer of the muscle lineage, the chief oncogene *PAX3-FOXO1* is driven by a translocated *FOXO1* super enhancer (SE) restricted to a late stage of myogenesis. Using chromatin conformation capture techniques, we demonstrate that the extensive *FOXO1* cis-regulatory domain interacts with *PAX3*. Furthermore, RNA sequencing and chromatin immunoprecipitation sequencing data in tumors bearing rare *PAX3* translocations implicate enhancer miswiring across all fusion-positive tumors. HiChIP of H3K27ac showed connectivity between the *FOXO1* SE, additional intra-domain enhancers, and the *PAX3* promoter. We show that *PAX3-FOXO1* transcription is diminished when this network of enhancers is ablated by CRISPR. Our data reveal a hijacked enhancer network that disrupts the stepwise CR TF logic of normal skeletal muscle development (*PAX3* to *MYOD* to *MYOG*), replacing it with an “infinite loop” enhancer logic that locks rhabdomyosarcoma in an undifferentiated stage.**

## INTRODUCTION

Control of the expression of the core regulatory transcription factors (CR TFs) that guide developmental decision making are directed by logical enhancer elements (Boyer et al., 2005; Chen et al., 2008; Lee and Young, 2013; Vermunt et al., 2019). These genomic elements, when heavily activated, become super enhancers (SEs) with unusually large deposits of active histone marks, chromatin regulators, and transcriptional coactivators (Hnisz et al., 2013). Chromosomal rearrangements allowing SEs to drive oncogene expression is an emerging mechanism in tumor biology (Bandopadhyay et al., 2016; Northcott et al., 2014; Xia and Wei, 2019; Zimmerman et al., 2018). Alveolar (fusion-positive) rhabdomyosarcoma (FP-RMS), an aggressive skeletal muscle cancer of childhood, often possesses chromosomal translocations, involving commonly *PAX3* and *FOXO1* genes, rarely *PAX7-FOXO1*, and in exceptional cases *PAX3-INO80D* and *PAX3-NCOA1* fusions (Shern et al., 2014). Disruption of CR TF transcription is effectual as FP-RMS treatment (Gryder et al., 2017, 2019a, 2019b). During normal skeletal muscle development, *PAX3* initiates specification of the muscle lineage and is shut off during myogenic differentiation. Consequently, master regulators *MYOD* and finally *MYOG* promote muscle progenitor cells to exit cell division and complete muscle differentiation (Hettmer and Wagers, 2010). However, although FP-RMS cells express these master regulators needed to trigger muscle differentiation program, they are halted in an early myoblastic and thus more proliferative state and are not able to complete cell differentiation. Fusion gene products are thought to be responsible for the inability of FP-RMS to differentiate. However, the mechanism of how the oncogenic fusions lock FP-RMS cells in their myoblast state has not been fully understood. In this study, we test the hypothesis that the chromosomal translocation event resulted in novel enhancer/promoter interactions to maintain robust expression of the oncogenic fusion protein in FP-RMS.

Previously, we uncovered a strong dependence on general SE function for tumor survival, with *PAX3-FOXO1* being a chief determinant of SE formation in collaboration with *MYOD* and *MYOG* (Gryder et al., 2017). Using chromatin conformation capture (3C, 4C-seq, HiChIP) and chromatin immunoprecipitation (ChIP) (ChIP sequencing [ChIP-seq], ChIP-Rx)-based assays, we here study a key SE 300 kb distal to *FOXO1*, and its interconnected smaller enhancer elements, and examine its function in FP-RMS. We

<sup>1</sup>Genetics Branch, National Cancer Institute, NIH, Bethesda, MD 20892, USA

<sup>2</sup>University Children's Hospital, Zurich, Switzerland

<sup>3</sup>Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

<sup>4</sup>Biologic Institute, Redmond, WA, USA

<sup>5</sup>Department of Oncohematology, Ospedale Pediatrico Bambino Gesù' Research Institute, IRCCS, Rome, Italy

<sup>6</sup>Center for Childhood Cancer & Blood Diseases, The Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH, USA

<sup>7</sup>Lead Contact

\*Correspondence: [berkley.gryder@nih.gov](mailto:berkley.gryder@nih.gov) (B.E.G.), [khanjav@mail.nih.gov](mailto:khanjav@mail.nih.gov) (J.K.)  
<https://doi.org/10.1016/j.isci.2020.101103>



propose that hijacking SEs bound by myogenic CR TFs allows for continued expression of oncogenic PAX fusions, thus circumventing normal myogenic enhancer logic.

## RESULTS

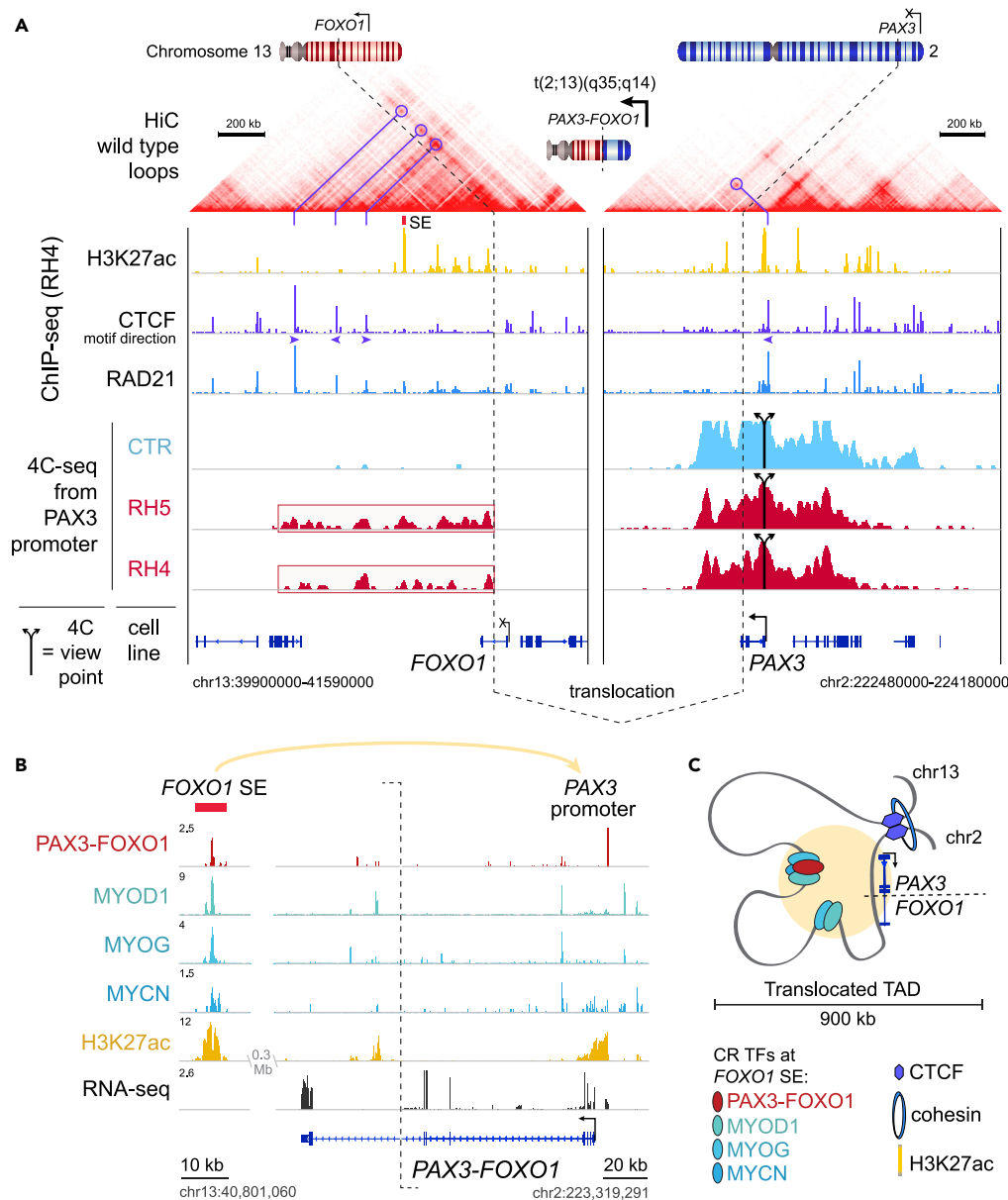
### Chromosomal Translocation Imports the FOXO1 Super Enhancer to the PAX3 Promoter

Precisely how PAX3-FOXO1 locks the cells into a myoblastic state unable to differentiate is unknown. Proper enhancer-promoter interactions are enabled by constraints in 3D chromatin folding, determined by CTCF and cohesin-formed loops at topologically associated domain (TAD) boundaries (Barrington et al., 2019; Dixon et al., 2012; Downen et al., 2014; Nora et al., 2017). PAX3 is normally silenced during progression past the myoblast stage of muscle differentiation. PAX3 expression during embryogenesis is tightly controlled, and structural variation that disrupts the PAX3 TAD causes limb malformation (Lupiáñez et al., 2015). We hypothesized that the fusion event results in novel enhancer/promoter looping events to maintain fusion protein expression independent of normal lineage control. Hi-C data (Rao et al., 2014) indicated three candidate topological loops containing wild-type FOXO1 that exist in normal cells. We found by ChIP-seq that all of these were occupied by RAD21 (of the cohesin complex) and CTCF in FP-RMS RH4 cells (Figure 1A). CTCF-binding events that form loops most often have binding motif sequences that are antiparallel (and point inward toward each other) (Rao et al., 2014). The CTCF motif orientation at the first and third of these sites near FOXO1 were found to be antiparallel with the CTCF motif near the PAX3 promoter, permissive of chromatin loop formation via extrusion after the translocation.

To identify interacting domains *cis* to the PAX3 promoter after the translocation, we used circularized chromatin conformation capture followed by sequencing (4C-seq) from viewpoint anchors around the PAX3 promoter and FOXO1 genes on chromosomes 2 and 13. Remarkably, looping was detected between the PAX3 promoter and multiple candidate enhancers downstream of FOXO1 and was restricted between the intronic fusion breakpoint in FOXO1 and the predicted topological boundary (Figure 1A). The outermost TAD-boundary looping interaction was confirmed by Sanger sequencing of the 3C PCR product (Figures S1A–S1C). Notably, each of the 3 CTCF sites 3' of FOXO1 formed looping interactions with PAX3 only in translocation-positive RH4, but not in the translocation-negative RMS cell line CTR (Figure S1D). A previous study of 4C-seq in FP-RMS cell lines has shown similar interactions consistent with our results (Vicente-García et al., 2017), but here we provide the first ChIP-seq-informed functional interpretation of these interactions as CTCF-bounded enhancers that contain critical CR TFs. Next, we analyzed this region with genomewide Hi-C contact map data in RH30 cells overlaid ChIP-seq for H3K27ac, confirming interactions between the FOXO1 *cis*-regulome and PAX3 (Figure S1E) in an independent FP-RMS cell line. We hypothesize that these newly juxtaposed enhancer elements keep the PAX3 promoter on the translocated allele active. These enhancers could sustain the continual PAX3-FOXO1 expression and the oncogenic process because they harbor strong binding sites for MYOD1, MYCN, and MYOG (Figures 1B and 1C).

### Rare PAX Fusions Implicate Enhancer Miswiring

Besides the PAX3-FOXO1 translocation, there are several other PAX translocation variations in FP-RMS including PAX7-FOXO1, PAX3-NCOA1, and PAX3-INO80D (Figure 2A). In each of these cases the 5' end (N terminal) of PAX3 or PAX7 is fused with the 3' end (C terminal) of FOXO1 (exons 1–7 [out of 9] of PAX3 with exons 2–3 of FOXO1), NCOA1 (type 1; exons 1–6 of PAX3 with exons 13–22 of NCOA1 and type 2; exons 1–7 of PAX3 with exons 12–22 of NCOA1), and INO80D (exons 1–7 of PAX3 fused with exons 9–11 of INO80D). There is protein homology between PAX3 and PAX7 (with similar DNA-binding domains), whereas NCOA1 or INO80D do not share any protein homology with the FOXO1 transactivation domain. However, RNA sequencing (RNA-seq) reveals remarkably similar transcriptome profiles from tumors harboring these diverse oncogenic fusions (Shern et al., 2014). In addition, SEs (found in RH4 cells to be bound by FP-RMS-specific CR TFs) exist near all known translocation partner genes (Figure 2B). Therefore, we hypothesized that enhancer miswiring as a result of translocations may be the common theme among all PAX fusion tumors. However, because translocations of NCOA1 and INO80D are rare and cell lines do not exist for them that may allow full epigenomic profiling we reasoned that we could infer epigenetic states (epistates) of the fusion and reciprocal alleles by measuring the exon-level gene expression of fusion partners. We expect that gene fusions involving one active and one inactive gene would result in an exon imbalance, from which we could infer an allele-specific epistate (Figure S2A). In the instance where both alleles have active epistates, we would expect less exonic imbalance (Figure S2B). Thus, from RNA-seq we could infer allele-specific epistates and inform a view of enhancer miswiring in rare forms of FP-RMS.



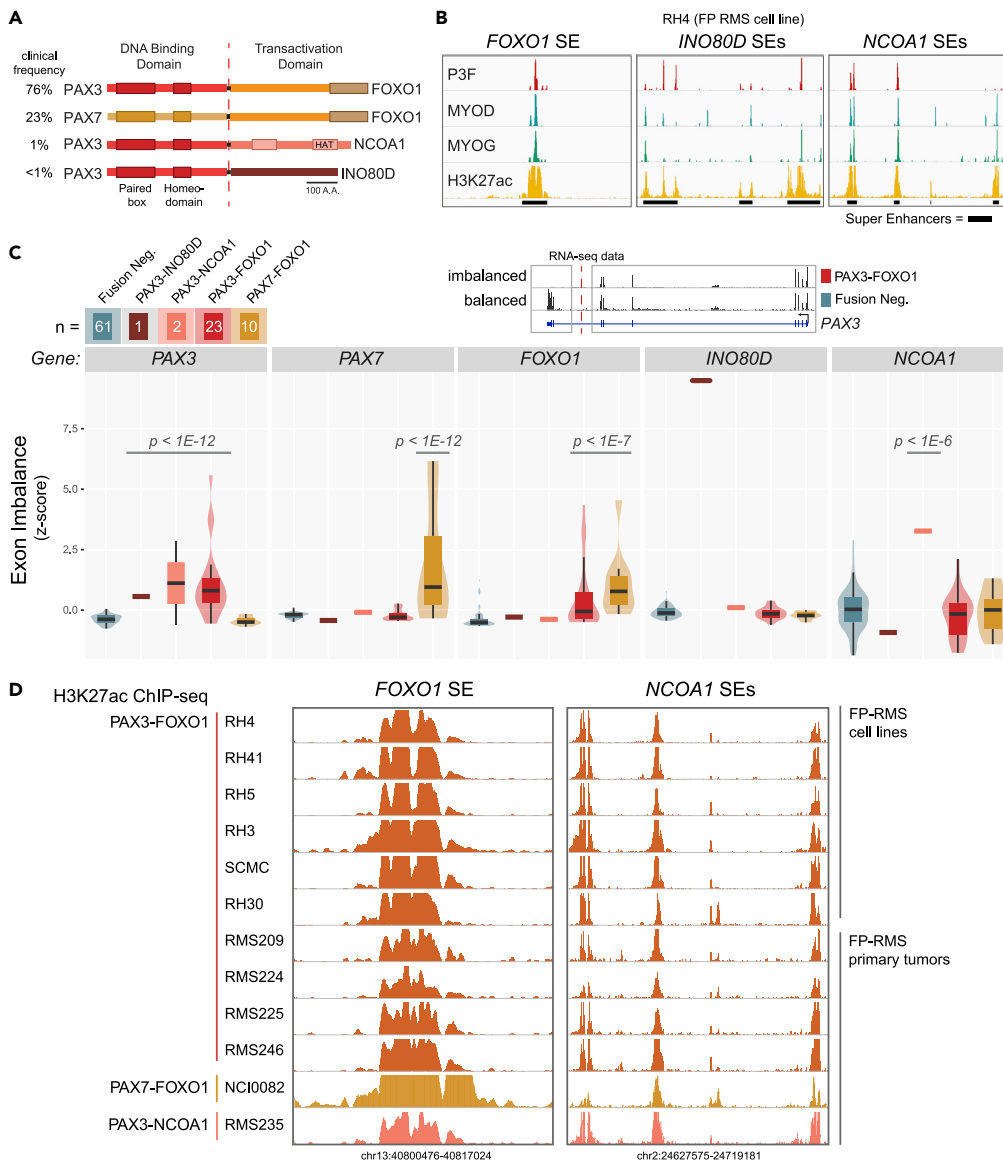
**Figure 1. Translocation Structures an Insulated Neighborhood Surrounding PAX3-FOXO1**

(A) Wild-type loops indicated by Hi-C profile from human GM12878 cells. ChIP-seq demonstrates binding locations of H3K27ac, CTCF, and RAD21 in RH4 cells. 4C-seq reveals looping between viewpoints at CTCF sites bounding FOXO1 enhancers, and the PAX3 promoter, in translocation-negative (CTR) and translocation-positive (RH5, RH4) cells. Viewpoints are indicated by split arrows, and translocation breakpoints are indicated by dotted lines.

(B) ChIP-seq signal for master transcription factors and H3K27ac, and RNA-seq signal, in reads per million (RPM), at the FOXO1 super enhancer (SE) and PAX3-FOXO1 fusion gene, in RH4 cells.

(C) Schematic of the translocation creating a new topologically associated domain (TAD) bringing the PAX3 promoter (chr2) under the control of FOXO1 SE and other smaller enhancers (chr13).

In FP-RMS with a PAX3-FOXO1 translocation, if only the promoter of PAX3 determines the expression of the PAX3 gene on the wild-type allele and PAX3-FOXO1 fusion gene on the translocated allele, the expression of all PAX3 exons will show low variance (as measured by Z scoring). The expression of PAX3 exons will show a high Z score if PAX3-FOXO1 is regulated by the abnormal juxtaposition of the FOXO1 SE, because the last exons of PAX3 are not influenced by the FOXO1 SE (both from the remaining wild-type PAX3 and the reciprocal FOXO1-PAX3 translocated allele). Therefore, we examined exon-level expression of the



**Figure 2. SEs and Allele-Specific Expression at Rare PAX3 Translocation Partners**

(A) PAX fusions and their clinical frequency in tumors of patients with FP-RMS. (B) SEs in RH4 (PAX3-FOXO1-bearing cells) include not only those near *FOXO1* but also *INO80D* and *NCOA1* in RH4 cells. (C) Exonic imbalance measure in RNA-seq data from primary tumors and cell lines of FN-RMS (n = 61), FP-RMS with PAX3-INO80D fusion (n = 1), PAX3-NCOA1 fusion (n = 2), PAX3-FOXO1 (n = 23), and PAX7-FOXO1 (n = 10). Box plots show the median, quartiles with whiskers at the 1.5 × interquartile range, and distributions plotted as violins. p values were calculated using a two-tailed t test with Welch's correction. Example genome browser of RNA-seq data in FP-RMS is from RH4 and FN-RMS is from SMS-CTR. (D) H3K27ac ChIP-seq in FP-RMS cell lines (n = 6) and primary tumors (n = 6) at the *FOXO1* SE and the *NCOA1* SEs.

genes involved in translocation using RNA-seq data from tumors of patients with FP-RMS (see [Transparent Methods](#)). The RNA-seq data showed that exons before the translocation (3' or N terminal) are always expressed significantly higher than those beyond the translocation breakpoint (5' or C terminal) as demonstrated by high Z score of exonic expression of *PAX3* or *PAX7* (Figure 2C). This is demonstrated by the high expression of *PAX3* exons 1–7 and low expression of exons 8–9 in the RH4 cell line with *PAX3-FOXO1*, indicating that only the exons involved in the fusion event are expressed due to activation by the *FOXO1* SE. Conversely, exon usage of a gene is balanced in patients lacking the translocation, e.g., *PAX3* in fusion-negative RMS (FN-RMS, SMS-CTR cells, Figure 2C). Importantly, inferring allele selective

expression via RNA-seq allows interrogation of extremely rare PAX fusions, *PAX3-INO80D* and *PAX3-NCOA1*. All fusion gene partners showed exonic imbalance (high Z scores) resulting from favored expression of the translocated exons (Figures 2C and S2C).

ChIP-seq data in FP-RMS cell lines and patients allowed us to discover recurrent SEs surrounding not only *FOXO1* but also rare partner *NCOA1* (Figure 2D). Although SEs represent only ~4% of enhancers, their unique and consistent presence argues that these diverse fusions may uniformly rewire SEs, which can be active in the epigenomic state of all patients with FP-RMS. Here we report the first epigenomic data generated for a *PAX3-NCOA1* patient, and we found that this rare fusion epigenetically phenocopies tumors with *PAX3-FOXO1* or *PAX7-FOXO1* (Figure 2D). When miswired, these SEs are key elements driving the expression of fusion genes.

### CRISPR Reveals Essentiality of *cis*-regulatory Elements Regulating *PAX3-FOXO1*

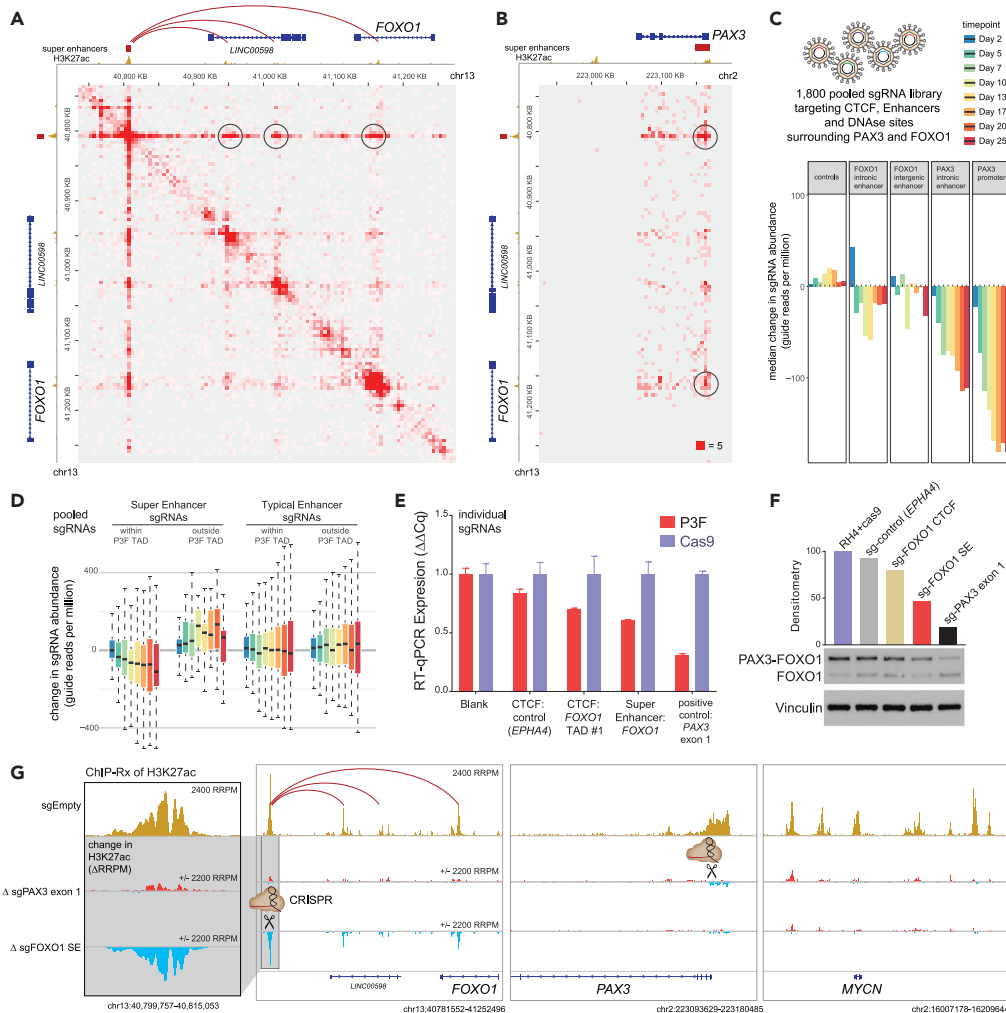
To build on the evidence from 4C, we set out to gain a more complete dataset confirming the interactions between the enhancer network controlling *PAX3-FOXO1*. Thus, we used HiChIP against H3K27ac to capture protein-directed interaction frequency between acetylated chromatin sites at enhancers and promoters (Gryder et al., 2020; Mumbach et al., 2016). The results identified that the *FOXO1* SE is connected not only to the *PAX3* promoter but also to three smaller intergenic and intronic enhancer elements near or within *FOXO1* (Figures 3A and 3B).

We next measured the contribution of these enhancer elements to the overall survival of FP-RMS cells. We designed a library of single guide RNAs (sgRNAs) against each enhancer or promoter constituent, each DNase hypersensitive site, and each CTCF peak as defined by genome-wide profiles in RH4 cells. We introduced them in a pooled fashion by viral infection into RH4 cells expressing Cas9. The abundance of each sgRNA in the population was then quantified over time using next-generation sequencing (at days 2, 5, 7, 10, 13, 17, 20, and 25). sgRNAs that target the *PAX3* promoter had the strongest impact on RH4 cell viability, as inferred from the largest reduction in guide representation over time (Figure 3C). Among CTCF sites, two candidate anchor sites (*FOXO1*-distal sites #2 and #3) had no negative influence, whereas the outermost CTCF sites (*FOXO1*-TAD boundary site #1 and *PAX3*-TAD boundary) were both reduced by negative selection (Figure S3A). Among TF-bound and H3K27ac-decorated enhancers, sgRNAs targeting SEs within the TAD (the translocation-induced insulated neighborhood containing *PAX3-FOXO1*) were more effective than sgRNAs targeting SEs outside this neighborhood, and also more than typical enhancers (Figure 3D).

Individual sgRNAs were next used to study the impact on *PAX3-FOXO1* transcription. We confirmed our hypothesis that disruption of the *FOXO1* SE reduced *PAX3-FOXO1* at the transcript level and protein level at 24 h after sgRNA infection (Figures 3E and 3F). This resulted in cell growth impairment over time (Figure S3B). To attribute the effect of this sgRNA to direct impairment of the enhancer, we assayed H3K27ac changes by ChIP-Rx (spike in reference normalized ChIP-seq) (Orlando et al., 2014). The results revealed that the *FOXO1* SE was depleted of H3K27ac and that the associated enhancer interaction network (as HiChIP identified) was also drastically reduced of acetylation at the sites interacting with *FOXO1* SE (Figure 3G, see Transparent Methods). Conversely, this enhancer network was not impaired by an sgRNA targeting the first exon of *PAX3*, except for slight reduction in acetylation levels at the *PAX3* promoter (Figure 3G). These data demonstrated that the *FOXO1* SE is essential to maintain the expression of *PAX3-FOXO1* oncogene in RMS.

### *FOXO1* SE Is Activated during a Key Step in Myogenesis

To examine if the activity of the *FOXO1* SE was coordinated with myogenic steps, we utilized ENCODE data mapping H3K27ac in various stages of the muscle lineage. We found that the *FOXO1* enhancer is transiently transformed into an SE during myogenesis at the same time *MYOG* acquires an SE (Figures 4A and 4B). *FOXO1*, *MYOG*, and *MYOD1* have more highly ranked SEs in FP-RMS when compared with FN-RMS (Figure 4C), in agreement with the notion that FP status is more advanced toward myotubes and FN status is more similar in the earlier myoblast state. *MYOG* activation is commonly prevented by mutant RAS signaling through MEK/ERK in FN-RMS tumors, which can be rapidly released via small molecule inhibitors of MEK/ERK (Yohe et al., 2018). Using this system, we asked if the *FOXO1* SE and concomitant *FOXO1* expression was induced alongside *MYOG* activation. Indeed, we not only found *FOXO1* to be up-regulated but also observed *MYOG* invasion on the same SE that is recruited during the *PAX3-FOXO1* translocation event in FP-RMS (Figure S4A).



**Figure 3. CRISPR Functional Mapping of Non-coding Elements Controlling PAX3-FOXO1**

(A) H3K27ac HiChIP reveals structure of FOXO1 SE interactions with smaller intra-TAD enhancer elements (data generated in RH4 cells previously; Gyder et al., 2019a).

(B) Interaction by H3K27ac HiChIP between PAX3 promoter and FOXO1 SE and intronic enhancer element.

(C) Pooled sgRNA CRISPR screening tiling against cis-regulatory genomic elements surrounding PAX3 and FOXO1 defined their degree of essentiality. RH4 cells expressing Cas9 were sampled by counting sgRNA abundance using sequencing at the indicated time intervals.

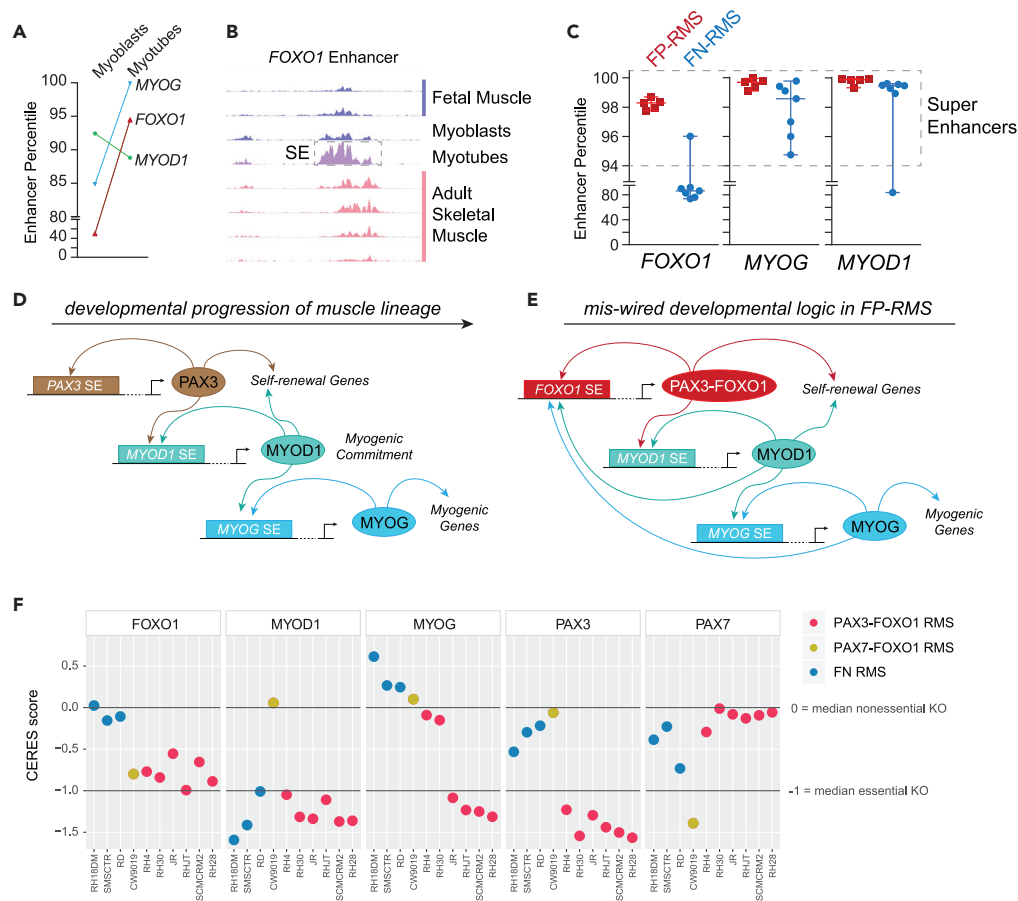
(D) Change in sgRNA abundance of pooled CRISPR shows that intra-TAD super enhancers are more critical for RH4 cell survival than typical enhancers or SEs outside TAD boundaries. Data are shown as box plots of the median and first and third quartiles, with whiskers showing 1.5 × interquartile range.

(E) Individual sgRNA impact on PAX3-FOXO1 gene expression after 1 day of transduction in FP-RMS RH4 cells expressing cas9. Bars show median, and error bars represent the SD of technical triplicates.

(F) PAX3-FOXO1 protein levels are reduced by individual sgRNAs targeting key cis-regulatory elements, especially those targeting the FOXO1 SE and the first exon of PAX3.

(G) ChIP-seq with reference exogenous spike-in (ChIP-Rx) for H3K27ac was employed to interrogate the chromatin impact of the sgRNA targeting the FOXO1 SE. The top track is control ChIP-Rx (sgEmpty), the second track is showing the change (delta RRPm) upon CRISPR of the first exon of PAX3, and the third track shows the change in H3K27ac upon CRISPR of the FOXO1 SE at 24 h post sgRNA transduction. All experiments were performed in RH4 cancer cells.

Together, our data suggested that the myogenic lineage timeline (Figure 4D) is miswired in rhabdomyosarcoma to maintain early TFs (granting the self-renewal capacity afforded by the de-differentiated cell state) despite the presence of late terminal differentiation factors (prominently MYOD and MYOG, Figure 4E). This implies that the infinite loop logic could be broken by removal of MYOD and MYOG from RMS cells, and although these TFs are not directly druggable as yet, DepMap CRISPR data strongly



**Figure 4. Miswired Super Enhancer Logic Able to Maintain An Oncogenic Cell State**

(A) Increased FOXO1 and MYOG enhancer rank (by percentile) in the transition from myoblasts to myotubes.  
 (B) H3K27ac signal at the FOXO1 enhancer in samples from different positions along the muscle lineage development timeline (right).  
 (C) Enhancer percentile for enhancers surrounding FOXO1, MYOG, and MYOD1 in RMS, as measured by rank of H3K27ac ChIP-seq signal. Each point shows the enhancer percentile in a different FP- or FN-RMS cell line or primary tumor.  
 (D) Model illustrating normal development of the muscle lineage.  
 (E) Miswiring of myogenic circuitry due to translocation of FOXO1 SE to PAX3 promoter, allowing MYOD and MYOG to activate PAX3-FOXO1.  
 (F) CRISPR dependency data of FOXO1, MYOG, MYOD1, PAX3, and PAX7 in RMS cell lines from the Achilles DepMap project, using CERES scoring scaled to 0 for the median nonessential gene knockout (KO) and  $-1$  for the median essential gene KO.

suggests that these (and the infinite loop they support) are essential for PAX3-FOXO1-driven RMS growth (Figure 4F).

## DISCUSSION

Two factors can be selected for in rearrangement-driven cancers: oncogenic biochemical function in the case of a resulting fusion protein (Kadoch and Crabtree, 2013) and aberrant expression levels of a proto-oncogene (such as MYC or GFI1) via enhancer hijacking (Northcott et al., 2014). In FP-RMS, PAX fusions are selected for both an oncogenic fusion protein product and miswired enhancer logic, effectively reprogramming core regulatory TF networks. Lacking a translocation, the FN-RMS subtype has aberrations in signaling pathways that enable circumvention of normal myogenic logic (Slemmons et al., 2017; Yohe et al., 2018).

PAX3 activates MYOD by binding and activating MYOD1 SEs, but then shuts off presumably because MYOD does not work backward to upregulate PAX3 (MYOD ChIP-seq shows no binding in the PAX3 promoter, unlike



the MYOG promoter, [Figure S4B](#)). Lacking enhancers responsive to MYOD/MYOG, the remaining wild-type alleles of PAX3/7 in FP-RMS tumors are silent. Our data suggest that newly juxtaposed enhancer elements initiate and continually drive PAX3-FOXO1 expression, implicating that enhancer miswiring is at the heart of the oncogenic process in FP-RMS. When the FOXO1 SE is translocated to regulate PAX3, late myogenic factors (MYOG/MYOD) work through this SE to drive an early myogenic factor (PAX3), changing a “progressive” enhancer logic into an “infinite loop” enhancer logic. Similarly, CR TF logic circuits are self-reinforcing in self-renewal and proliferative states such as embryonic stem cells ([Whyte et al., 2013](#)).

Analysis of RNA-seq for patients with non-FOXO1 partners (INO80D, NCOA1) of PAX3 reveals a remarkably similar transcriptome ([Shern et al., 2014](#)), despite not being TFs themselves, and having no sequence homology to the activation domain (AD) of FOXO1. It has been shown that TFs can perform their function even when their ADs are swapped with those of other TFs ([Hope and Struhl, 1986](#)). This tolerance to diverse AD sequence may be explained by the fact they share the common property of being intrinsically disordered, a feature needed to support phase separation capacity of TFs ([Bojja et al., 2018](#)). Indeed, the portions of FOXO1, INO80D, and NCOA1 involved in PAX3 fusion oncoproteins are predicted to be heavily disordered ([Figures S5A–S5C](#)). Remarkably, although transcription factors as a class are heavily disordered ([Figure S5D](#)), PAX fusion partners are particularly disordered (similar to FET family fusions like EWSR1). A related partner MAML3 is almost entirely disordered ([Figure S5E](#)), and PAX3-MAML3 fusions occur in biphenotypic sinonasal sarcoma (SNS) but not in RMS ([Wang et al., 2014](#)). MAML3 lacks an SE in myogenesis or RMS, and our model would therefore predict the absence of PAX3-MAML3 translocations in RMS. SNS may arise from a cell of origin whose epigenome has a lineage-restricted enhancer at MAML3, which gets recruited in SNS tumorigenesis. We believe that even these non-TF fusion partners are actually acting as TFs by inducing liquid-liquid phase condensate formation and likely serving as a scaffold for the intrinsically disordered domains of proteins (i.e., Mediator, BRD4, and RNA Pol2) that are essential for PAX3 fusions to drive transcription from distal enhancers.

Many disordered proteins in the genome are not involved in PAX3 fusions. A parallel criterion for a successful tumorigenic fusion could be the presence of an active SE in the same lineage step as PAX3 (such as those SEs proximal to FOXO1, INO80D, and NCOA1 in the RMS-specific epigenetic state). SE-containing loci may be enriched in translocations for two reasons. First, active enhancers are transcriptionally active and early replicating ([Siefert et al., 2017](#)), and thus likely more susceptible to double-strand breakage ([Barlow et al., 2013](#)). Second, among translocations which form, those resulting in overexpression of an oncogene are selected for, and SEs can enable such continued overexpression. The appearance of certain SEs is transient and logically restricted to certain points in development and thus may be restricting the potential miswiring events that could give rise to an “infinite loop” in CR TF logic. We propose that this can explain, at least in part, the selection of translocation partners in FP-RMS tumors and provides a paradigm likely relevant to other translocation-driven cancers.

### Limitations of the Study

RNA-seq and ChIP-seq analyses used for inferring enhancer miswiring in rare PAX3/7 fusion tumors is limited as we have only validated it functionally in cell lines bearing the most common fusion PAX3-FOXO1. Furthermore, the bioinformatic analysis and predictions we have made regarding the intrinsically disordered regions of FOXO1, INO80D, and NCOA1, have not been validated with experiments to show liquid droplet formation. Last, the enhancer data for myogenesis are limited to only a few steps along this lineage, according to what is publicly available. These limitations leave important work to be pursued in future work by the field.

### Resource Availability

#### Lead Contact

Javed Khan ([khanjav@nih.gov](mailto:khanjav@nih.gov)).

#### Materials Availability

Plasmids, cell lines are available upon request. Primary tumor samples are unavailable for public distribution but please contact us if interested in collaborative exploration with these rare materials.

### Data and Code Availability

Newly generated ChIP-seq data from primary RMS tumors has been made publicly available through the Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>). The GEO accession number is GSE136799. Code is available at <https://github.com/GryderArt>.

### METHODS

All methods can be found in the accompanying [Transparent Methods supplemental file](#).

### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101103>.

### ACKNOWLEDGMENTS

We thank Michael Kuehl and Katherine Masih for thoughtful discussions of the manuscript. We thank the CCR Genomics Core at the National Cancer Institute, NIH, Bethesda, MD, for Sanger sequencing. This work was supported by the NCI, NIH. S.P. is a recipient of a Fondazione Veronesi fellowship.

Neither does the content of this publication necessarily reflect the views or policies of the Department of Health and Human Services nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

### AUTHOR CONTRIBUTIONS

B.E.G. and J.K. conceived the project and designed experiments. B.E.G. performed 3C and 4C-seq experiments. B.E.G. and Y.S. performed ChIP-seq, ChIP-Rx, and RNA-seq experiments and generated data. B.E.G. and B.Z.S. performed HiChIP experiments. B.E.G. and W.E. wrote scripts and performed bioinformatic analysis. N.G.A., W.M., and M.W. performed individual CRISPR-Cas9-based experiments under supervision from B.E.G., J.K., and B.S., and results were interpreted by all authors. Western blot experiments were performed by S.P. under the supervision of J.K. and R.R. K.C., O.E.D., and C.R.V. designed the pooled CRISPR screening of regulatory elements and assisted with bioinformatic analysis of the resulting data. B.E.G. wrote the manuscript, and all authors contributed to editing the manuscript. The final draft was approved by J.K.

### DECLARATION OF INTERESTS

The authors declare no conflicts of interest.

Received: February 18, 2020

Revised: March 31, 2020

Accepted: April 23, 2020

Published: May 22, 2020

### REFERENCES

- Bandopadhyay, P., Ramkissoon, L.A., Jain, P., Bergthold, G., Wala, J., Zeid, R., Schumacher, S.E., Urbanski, L., O'Rourke, R., and Gibson, W.J. (2016). MYB-QKI rearrangements in angiocentric glioma drive tumorigenicity through a tripartite mechanism. *Nat. Genet.* 48, 273.
- Barlow, J.H., Faryabi, R.B., Callén, E., Wong, N., Malhowski, A., Chen, H.T., Gutierrez-Cruz, G., Sun, H.-W., McKinnon, P., and Wright, G. (2013). Identification of early replicating fragile sites that contribute to genome instability. *Cell* 152, 620–632.
- Barrington, C., Georgopoulou, D., Pezic, D., Varsally, W., Herrero, J., and Hadjuri, S. (2019). Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. *Nat. Commun.* 10, 1–14.
- Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* 175, 1842–1855.e16.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947–956.
- Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.
- Downen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K., et al. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159, 374–387.
- Gryder, B.E., Khan, J., and Stanton, B.Z. (2020). Measurement of differential chromatin

interactions with absolute quantification of architecture (AQuA-HiChIP). *Nat. Protoc.* 15 (3), 1209–1236.

Gryder, B.E., Pomella, S., Sayers, C., Wu, X.S., Song, Y., Chiarella, A.M., Bagchi, S., Chou, H.-C., Sinniah, R.S., Walton, A., et al. (2019a). Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma. *Nat. Genet.* 51, 1714–1722.

Gryder, B.E., Wu, L., Woldemichael, G.M., Pomella, S., Quinn, T.R., Park, P.M.C., Cleveland, A., Stanton, B.Z., Song, Y., Rota, R., et al. (2019b). Chemical genomics reveals histone deacetylases are required for core regulatory transcription. *Nat. Commun.* 10, 3004.

Gryder, B.E., Yohe, M.E., Chou, H.-C., Zhang, X., Marques, J., Wachtel, M., Schaefer, B., Sen, N., Song, Y.K., Gualtieri, A., et al. (2017). PAX3-FOXO1 establishes myogenic super enhancers and confers BET bromodomain vulnerability. *Cancer Discov.* 7, 884–899.

Hettmer, S., and Wagers, A.J. (2010). Muscling in: uncovering the origins of rhabdomyosarcoma. *Nat. Med.* 16, 171–173.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947.

Hope, I.A., and Struhl, K. (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* 46, 885–894.

Kadoch, C., and Crabtree, G.R. (2013). Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in Synovial sarcoma. *Cell* 153, 71–85.

Lee, T.I., and Young, R.A. (2013). Transcriptional regulation and its misregulation in disease. *Cell* 152, 1237–1251.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause

pathogenic rewiring of gene-enhancer interactions. *Cell* 161, 1012–1025.

Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* 13, 919–922.

Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* 169, 930–944.e2.

Northcott, P.A., Lee, C., Zichner, T., Stütz, A.M., Erkek, S., Kawachi, D., Shih, D.J., Hovestadt, V., Zapatka, M., and Sturm, D. (2014). Enhancer hijacking activates GF11 family oncogenes in medulloblastoma. *Nature* 511, 428–434.

Orlando, D.A., Chen, M.W., Brown, V.E., Solanki, S., Choi, Y.J., Olson, E.R., Fritz, C.C., Bradner, J.E., and Guenther, M.G. (2014). Quantitative ChIP-seq normalization reveals global modulation of the epigenome. *Cell Rep.* 9, 1163–1170.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680.

Shern, J.F., Chen, L., Chmielecki, J., Wei, J.S., Patidar, R., Rosenberg, M., Ambrogio, L., Auclair, D., Wang, J., Song, Y.K., et al. (2014). Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic Axis in fusion-positive and fusion-negative tumors. *Cancer Discov.* 4, 216–231.

Siefert, J.C., Georgescu, C., Wren, J.D., Koren, A., and Sansam, C.L. (2017). DNA replication timing during development anticipates transcriptional programs and parallels enhancer activation. *Genome Res.* 27, 1406–1416.

Slemmons, K.K., Crose, L.E., Riedel, S., Sushnitha, M., Belyea, B., and Linardic, C.M. (2017). A novel notch-YAP circuit drives stemness and tumorigenesis in embryonal rhabdomyosarcoma. *Mol. Cancer Res.* 15, 1777–1791.

Vermunt, M.W., Zhang, D., and Blobel, G.A. (2019). The interdependence of gene-regulatory elements and the 3D genome. *J. Cell Biol.* 218, 12–26.

Vicente-García, C., Villarejo-Balcells, B., Irastorza-Azcárate, I., Naranjo, S., Acemel, R.D., Tena, J.J., Rigby, P.W.J., Devos, D.P., Gómez-Skarmeta, J.L., and Carvajal, J.J. (2017). Regulatory landscape fusion in rhabdomyosarcoma through interactions between the PAX3 promoter and FOXO1 regulatory elements. *Genome Biol.* 18, 106.

Wang, X., Bledsoe, K.L., Graham, R.P., Asmann, Y.W., Viswanatha, D.S., Lewis, J.E., Lewis, J.T., Chou, M.M., Yaszemski, M.J., Jen, J., et al. (2014). Recurrent PAX3-MAML3 fusion in biphenotypic sinonasal sarcoma. *Nat. Genet.* 46, 666–668.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319.

Xia, J.-H., and Wei, G.-H. (2019). Enhancer dysfunction in 3D genome and disease. *Cells* 8, 1281.

Yohe, M.E., Gryder, B.E., Shern, J.F., Song, Y.K., Chou, H.-C., Sindiri, S., Mendoza, A., Patidar, R., Zhang, X., and Guha, R. (2018). MEK inhibition induces MYOG and remodels super-enhancers in RAS-driven rhabdomyosarcoma. *Sci. Transl. Med.* 10, eaan4470.

Zimmerman, M.W., Liu, Y., He, S., Durbin, A.D., Abraham, B.J., Easton, J., Shao, Y., Xu, B., Zhu, S., and Zhang, X. (2018). MYC drives a subset of high-risk pediatric neuroblastomas and is activated through mechanisms including enhancer hijacking and focal enhancer amplification. *Cancer Discov.* 8, 320–335.

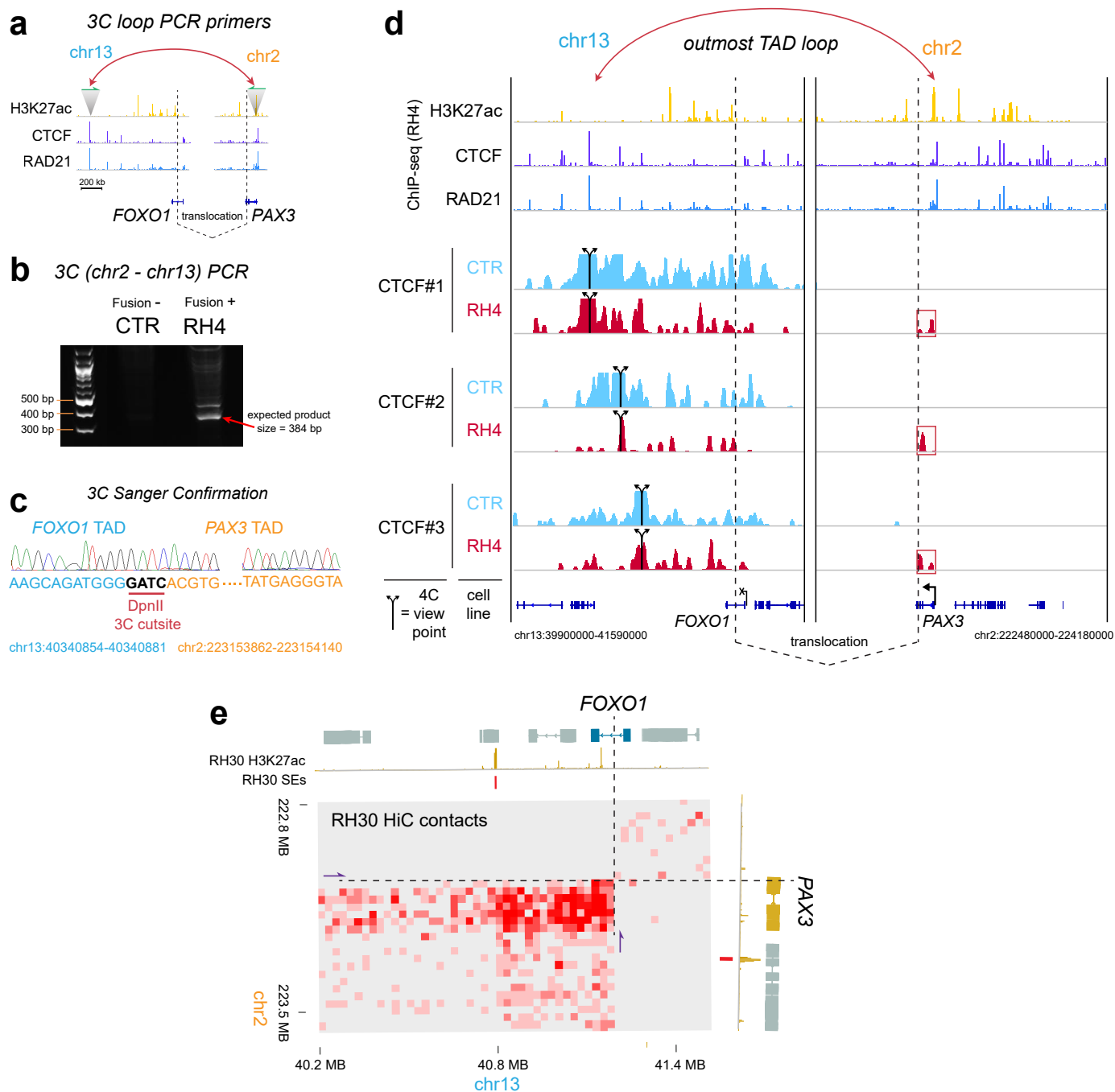
iScience, Volume 23

## **Supplemental Information**

### **Miswired Enhancer Logic Drives**

#### **a Cancer of the Muscle Lineage**

**Berkley E. Gryder, Marco Wachtel, Kenneth Chang, Osama El Demerdash, Nicholas G. Aboreden, Wardah Mohammed, Winston Ewert, Silvia Pomella, Rossella Rota, Jun S. Wei, Young Song, Benjamin Z. Stanton, Beat Schäfer, Christopher R. Vakoc, and Javed Khan**



**Figure S1, related to Figure 1.**

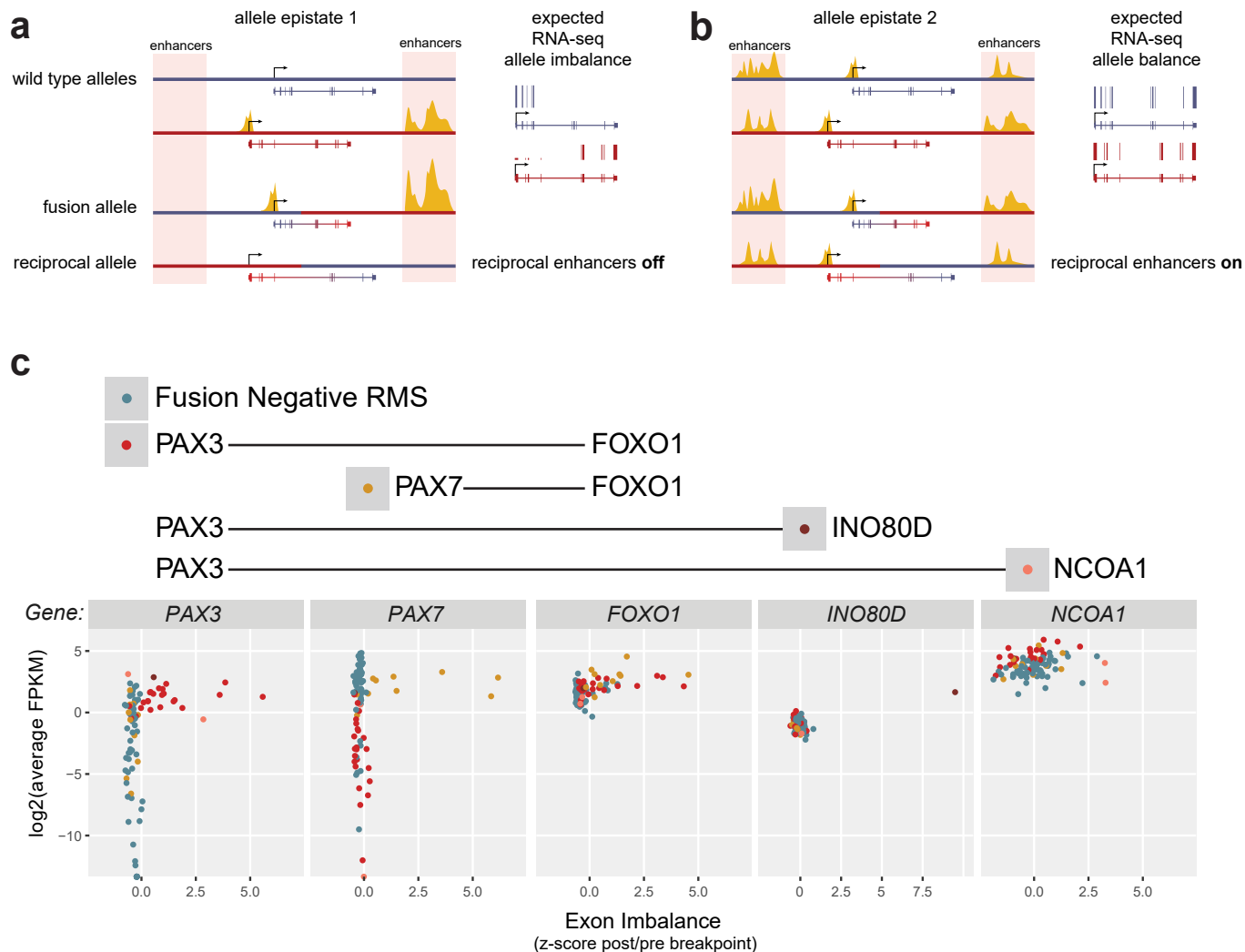
**a.** Diagram of 3C confirmation PCR primer placement

**b.** 3C confirmation PCR in fusion negative (CTR) and fusion positive (RH4) cells.

**c.** Sanger validation of 3C PCR product.

**d.** 4C-seq of CTCF sites: 3 viewpoint sites distal to *FOXO1* on chromosome 13, which show interaction with chromatin proximal to the *PAX3* promoter and CTCF anchor.

**e.** *PAX3-FOXO1* translocation induced HiC contacts between the *FOXO1* cis-regulome and the *PAX3* gene body/promoter in RH30. H3K27ac ChIP-seq from RH30 also shown, with SEs for the same cell line.

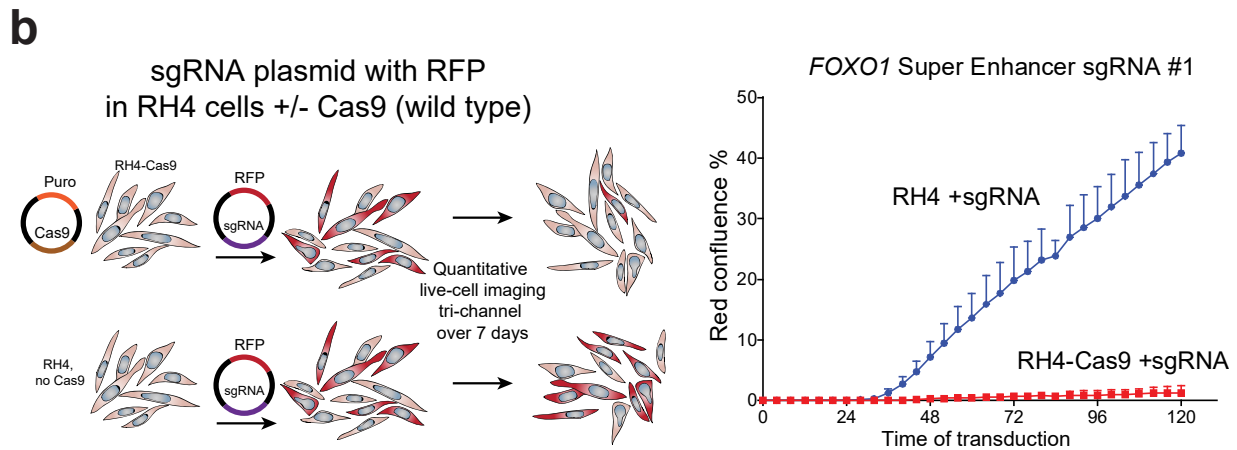
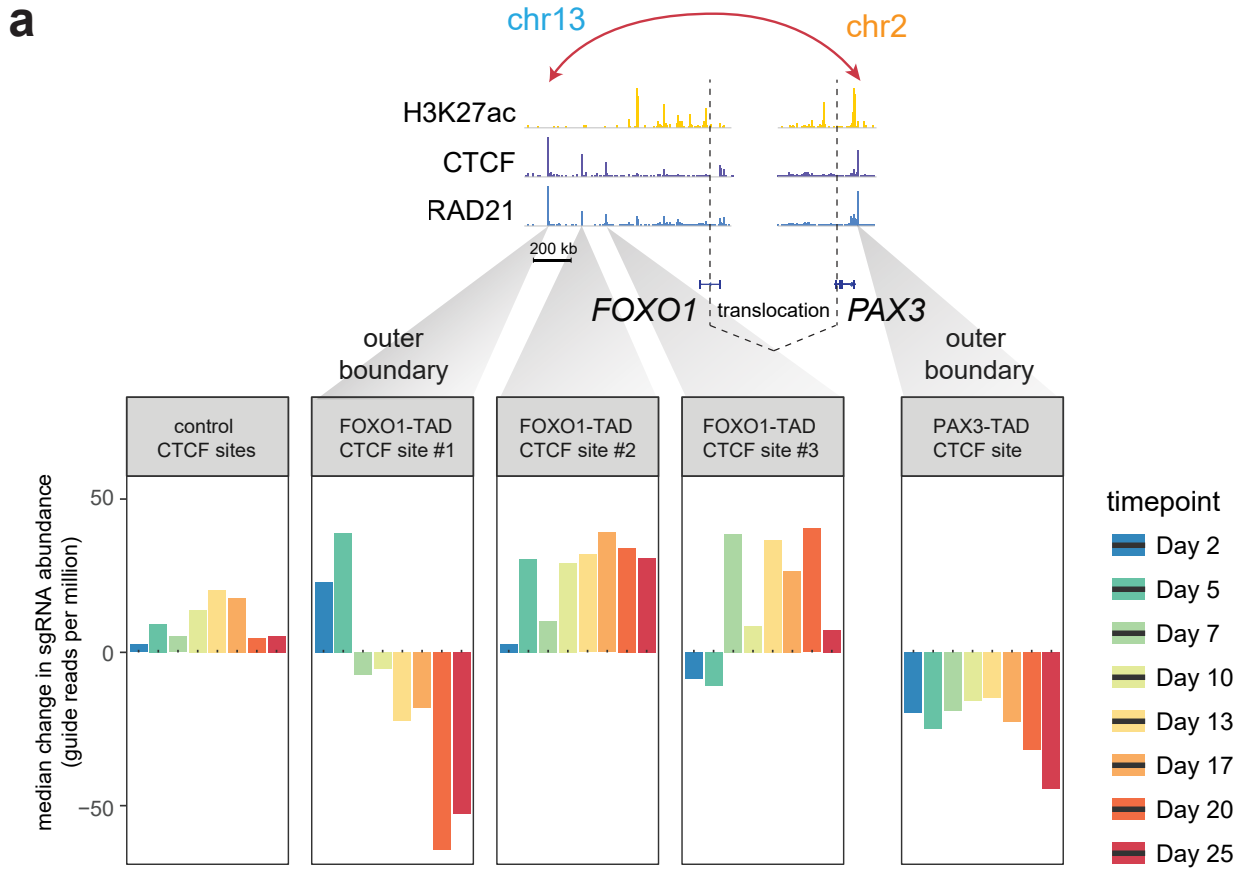


**Figure S2, related to Figure 2.**

**a.** Allele specific epigenetic state (epistate) is expected to create an exonic imbalance by activating fused allele and not reciprocal allele lacking active enhancers.

**b.** If both partners of a fusion gene are surrounded by active enhancers then the wild type, fusion and reciprocal alleles being active would result in a more balanced exon level expression than in (a).

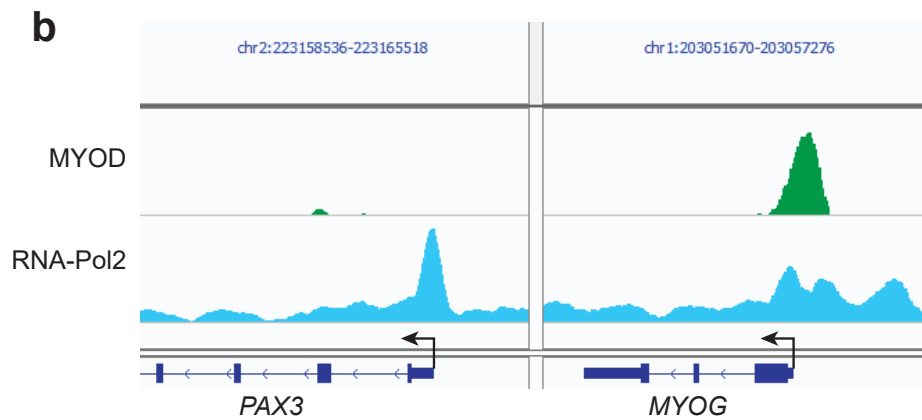
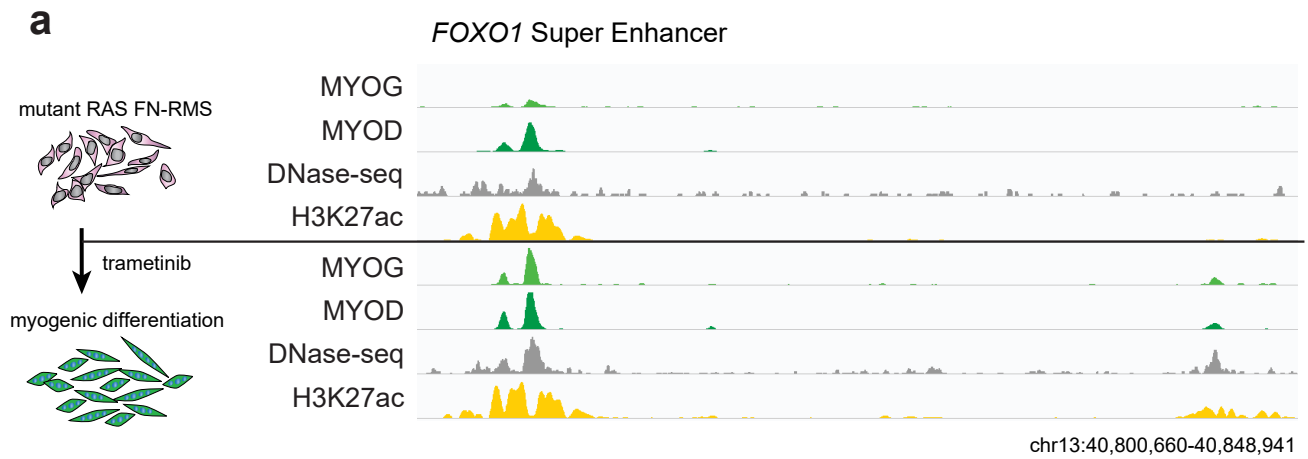
**c.** Exon imbalance for PAX fusion partner genes, plotted against RNA expression levels.



**Figure S3, related to Figure 3.**

**a.** Pooled sgRNA CRISPR interrogation of CTCF sites surrounding *PAX3* and *FOXO1* reveals importance of the two most distal sites, but not two intervening CTCF bound locations distal to *FOXO1*.

**b.** Time course Cas9-negative selection test with sgRNA targeting *FOXO1* SE and Red Fluorescent Protein (RFP) expression, with and without Cas9 in RH4 cells.

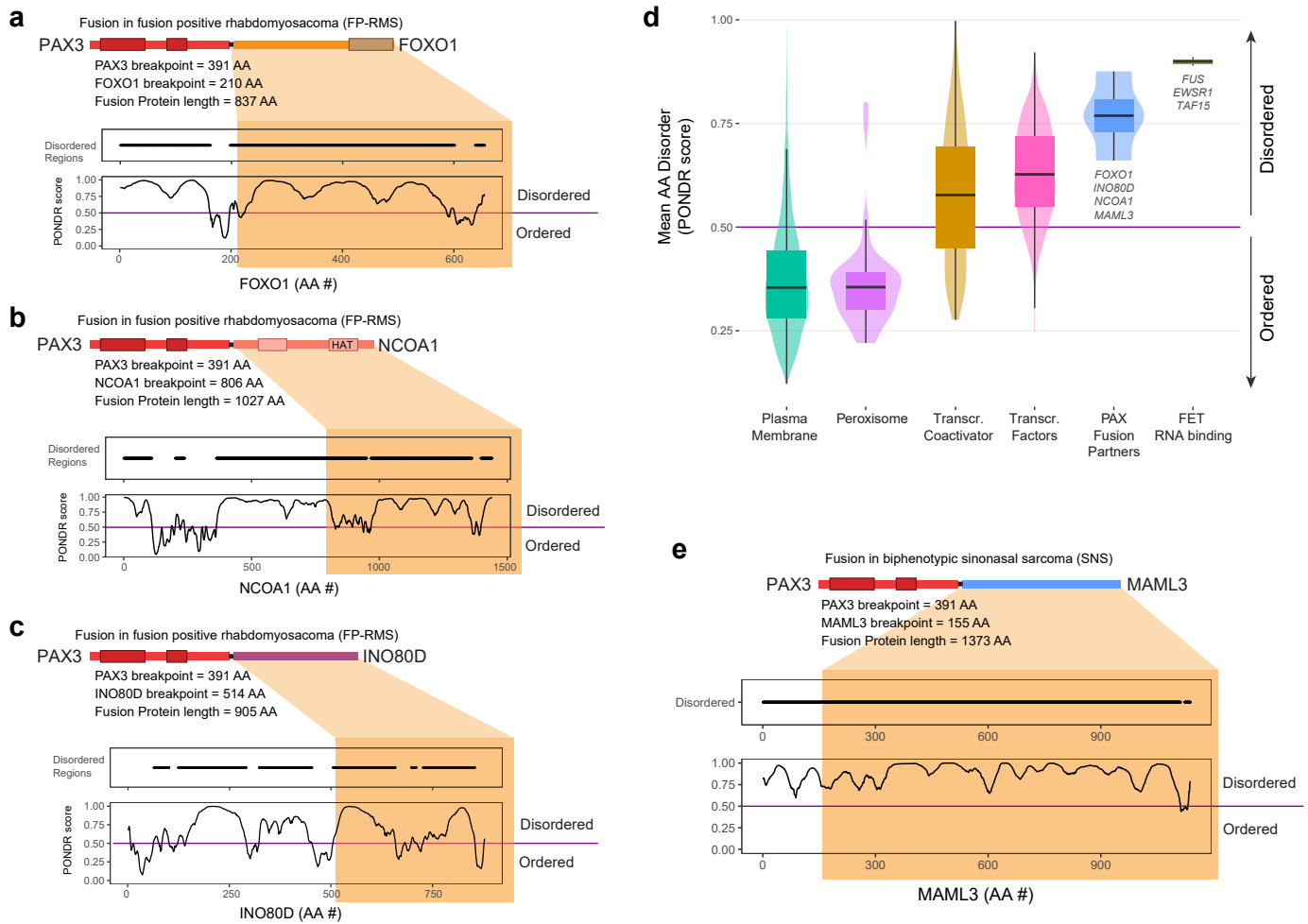


**Figure S4, related to Figure 4.**

a. MYOG invades the *FOXO1* super enhancer during trametinib-induced myogenesis in SMS-CTR cells, a PAX-fusion lacking cell line driven by mutant HRAS (Q61K).

b. MYOD ChIP-seq alongside RNA Pol2 ChIP-seq at the *PAX3* and *MYOG* promoters in RH4 cells (*PAX3-FOXO1* translocated FP-RMS cancer cells)





**Figure S5, related to Figure 4.**

- PAX3-FOXO1 fusion gene, with PONDNR (Predictor of Natural Disordered Regions) score (<http://www.pondr.com/>) and disordered regions mapped for FOXO1 amino acids.
- PAX3-NCOA1 fusion gene, with PONDNR score for NCOA1.
- PAX3-INO80D fusion gene, with PONDNR score for INO80D
- Distribution of average disorder PONR scores among protein families, including transcriptional coactivators and transcription factors, alongside PAX3 fusions and FET family fusion partners.
- PAX3-MAML3 fusion gene, with PONDNR score for MAML3.

## Transparent Methods

### Circularized chromatin conformation capture (4C-seq)

4C-seq was performed on RH4, RH5 and CTR cells as previously described (Gryder et al., 2017). Briefly, cells were grown in DMEM at 37 °C, then chemically crosslinked with 1% formaldehyde for 12 minutes. Then, *in-situ* digestion and *in-situ* re-ligation of 3D contacts (keeping nuclei intact) was performed using DpnII as the 4-bp DNA cutter. Following re-ligation, Csp6I was used to reduce template size, followed by relegation to circularize for inverse PCR at viewpoints of interest surrounding *PAX3* and *FOXO1*. 4C samples of RMS cell lines were amplified using bait region primers for *PAX3*-promoter (F: CAAGGAGTCCTGGTGCCAAA, R: CACTGGCCGGTGAGAAGG). We also studied three CTCF sites on the *FOXO1* side with the following primers: FOXO1-CTCF.1 (F: GCTCCCACAGAAGAAGCAGA, R: GGTGGAGACAGAGGCAGTAC), FOXO1-CTCF.2 (F: CACACACAAGCAAGCACAGA, R: AGCCTCATTACCACTTTTGAACA), and FOXO1-CTCF.3 (F: TCAGGAAGGTTCAAACCTTACTTTCC, R: CACGCACGCATAAAAGAGCA). Illumina TruSeq ChIP Library Prep Kit was used on purified inverse PCR products, and sequenced (75 bp single end) with an Illumina NextSeq500 (and, are typically multiplexed with ChIP-seq experiments to increase sequence diversity, which is needed because most reads from a single 4C viewpoint are identical duplicated products from self-ligation).

### 4C-seq data analysis

Reads from 4C experiments were first filtered to keep only reads containing the bait primer sequence (from the viewpoint), tolerating 1 mismatched base pair. Then, barcodes and viewpoint sequence were trimmed, followed by mapping to hg19 with bwa. Reads surrounding the viewpoint (within 4 kb) were removed to aid visualization of informative distal contacts. Smoothing was performed by averaging over a sliding window of 5 kb and visualized in IGV.

### 3C confirmation PCR and sequencing

Primers were designed to validate interaction between the outermost CTCF boundaries surrounding *PAX3* and *FOXO1* (Extended Data Fig. 1). The left primer overlapped a digestion site detected to form a long-range ligation at high frequency in 4C data, with the sequence: AAGCAGATGGGG**GATCACGTG** (DpnII cutsite in bold). The right primer was designed close to the *PAX3* CTCF site, but before any additional cutsites: GCAGCCAGTGAGATAAGACATTA. PCR was performed using Phusion High Fidelity Master Mix (NEB) with 30 pmol of each primer and 70 ng of 3C DNA template (same template from which 4C was prepared, prior to the second Csp6I cut), with 25 cycles (94 °C for 15 seconds, 55 °C for 30 seconds, 70 °C for 30 seconds) with completed with 1 minute at 70 °C. PCR product was purified with PCR QiaQuick kit (Qiagen), then Sanger sequencing was performed at the Center for Cancer Research Sequencing Facility (CCR-SF).

### Exonic imbalance analysis of RMS RNA-seq

RNA-seq read counts were generating using cufflinks at the exon level for each sample (RMS primary tumor data) (Shern et al., 2014). Then, for each RMS tumor (n = 97), the exonic balance before and after the known break points were calculated (average FPKM of pre-breakpoint exons/average FPKM of post-breakpoint exons) for all known translocation partners (*PAX3*, *PAX7*, *FOXO1*, *NCOA1*,

*INO80D*). Isoforms used for analysis were *PAX3*: NM\_181458; *PAX7*: NM\_001135254; *FOXO1*: NM\_002015; *NCOA1*: NM\_147223; *INO80D*: NM\_017759

Normal exonic bias was corrected for by taking a z-score across all samples (including FN-RMS samples). Plots of exon imbalance were made in R using ggplot2.

### HiC and HiChIP data analysis

HiC data for high-resolution contact frequencies in GM12878 cells (Rao et al., 2014) was downloaded and visualized through the Juicebox desktop application (Durand et al., 2016). Candidate native loop contacts (purple circles, Fig. 1) were annotated manually from visual inspection, and CTCF motif orientation was derived by taking the intersect of CTCF peaks in RH4 cells called with MACS2 (previously published, GEO accession number GSM2214099) and CTCF recognition sequences annotated in the HOMER known motifs collection (<http://homer.ucsd.edu/homer/motif/genomeWideMotifScan.html>). HiC data in RH30 cells was obtained from ENCODE3.

H3K27ac HiChIP data analysis of data generated in RH4 cells (Gryder et al., 2019) was performed using the Hi-Pro pipeline (Gryder et al., 2020; Servant et al., 2015). We kept only high-quality paired-end tags (PETs), using the HiC-Pro pipeline to filter out reads lacking the restriction site (DpnII), filter out duplicates, and filter out poorly mapped reads. Valid contact pairs remaining were converted to .hic format using the hicpro2juicebox.sh script (<https://github.com/nservant/HiC-Pro/tree/master/bin/utils>), to enable compatibility with the Juicebox visualization tool.

### Lentivirus production

A lentivirus plasmid mix was made by combining Lenti-Rev, Lenti-PM2, Lenti-Tat and Lenti-Vsv-G plasmids at a ratio of 1:1:1:2. HEK293T cells were seeded at 50-60% confluency (2 million cells) in a 10 cm dish in 10 mL media (DMEM, 10% FBS) and incubated at 37°C, 5% CO<sub>2</sub>. The next day, 5 µg of transfer plasmid, 10 µg of lentivirus plasmid mix and 36 µL of X-tremeGENE HP DNA transfection reagent (Sigma) were mixed into 1.2 mL serum-free OptiMEM (Gibco), vortexed and incubated for 15 minutes at RT and added to the cells. Media was changed 24 hours after transfection. Viral supernatant was harvested the following three days and filtered through a 0.45 µm syringe filter. Each collection was diluted with 5X PEG-it Virus Precipitation Solution (System Biosciences) and stored at 4°C until the final collection. After the final collection, viral supernatants were centrifuged at 1500g for 30 minutes at 4°C. Viral pellets were resuspended in PBS, aliquoted and stored at -80°C. All transductions were conducted with 8 µg/mL polybrene.

### Pooled CRISPR Library Design and Construction

Sequences of enhancer regions (~100 sites) surrounding *PAX3-FOXO1*, CTCF, and SE peaks were derived such that all possible sgRNA guides for *S.pyogenes* Cas9 can be designed within those regions. A total of 1830 sgRNAs were produced. Guides were selected based on having an off-target score of 1.0 denoting the maximum possible score for guides not having any off-target.

Oligonucleotides were synthesized as a pool commercially (Twist Biosciences) and then PCR cloned into BsmBI-cut sgRNA expression vector, LRG2.1T, by using gibbon assembly. Deep sequencing analysis was performed on Illumina platform and verified that all of sgRNA designs were cloned (data not shown).

### **Pooled CRISPR-Cas9 screening**

The lentiviral sgRNA library was produced as described above. RH4.Cas cells were first transduced with varying concentrations of the pooled lentiviral library to determine the amount needed for a MOI of ~0.3. RH4.Cas cells were then seeded in three 15 cm dishes at 8 million cells per dish. 24 hours later, cells were transduced with the lentiviral library at a low MOI, leading to ~27% positive cells (as determined by FACS analysis of GFP expression) and ensuring over 1000x representation of the library. 48 hours after transduction, virus-containing media was removed and replaced with fresh media. 40 million cells were collected and pelleted at days 2, 5, 7, 10, 13, 17, 20 and 25 post-transduction and 12 million cells were plated at each timepoint for subsequent cell sampling. Coverage at cell level was kept above 1000x throughout the screen.

Sequencing libraries were constructed essentially as previously described (Shi et al., 2015). Genomic DNA (gDNA) from cell pellets was extracted using the AllPrep kit (Qiagen). 32 parallel PCR reactions were performed to amplify sgRNA sequences from gDNA harvested at each timepoint. The total volume of each PCR reaction was 50  $\mu$ L containing 400 ng of gDNA template, 0.2  $\mu$ M forward (TCTTGTGGAAAGGACGAAACACCG) and reverse (TCTACTATTCTTCCCTGCACTGT) primers, and 25  $\mu$ L Platinum PCR SuperMix (Thermo Fisher). PCR cycles were: 1x (98°C, 2min) 35x (98°C, 8s, 67°C, 12s 72°C, 10s), 1x (72°C, 5min). Parallel PCR products were pooled together and purified with the QiaQuick kit (Qiagen). After production of sgRNA amplicons from gDNAs by PCR, ends of fragments were repaired using T4 DNA polymerase (NEB), DNA polymerase I large fragment (Klenow) (NEB), and T4 polynucleotide kinase (NEB), followed by addition of 3' A-overhang with Klenow (3'-5'-exo-) (NEB). Amplicons from each sample were ligated with a unique barcode adaptor (pool) and purified using AMPure magnetic beads (Beckman Coulter). Samples were deep sequenced on Illumina platform. Read counts for each sample were ascertained by mapping raw reads to the sgRNA library sequences. Read counts of each sample were normalized to facilitate data analysis.

### **Construction of individual sgRNA expression plasmids**

sgRNA sequences were designed using the CRISPR design tool from MIT (crispr.mit.edu). Pairs of DNA oligonucleotides encoding the protospacer sequences were annealed together to create double-stranded DNA fragments with 4-bp overhangs. These fragments were ligated into BsmBI digested Shuttle\_sg\_RFP657 plasmid. Plasmid constructs were confirmed via sanger sequencing using the LKO1\_5 primer (GACTATCATATGCTTACCGT) on the U6 promoter.

### **DepMap Achilles Data Analysis**

CRISPR-cas9 screening for essential genes was performed by the Broad Institute's DepMap and Achilles team (Meyers et al., 2017), and the data was downloaded from the 2020 Q1 release (<https://depmap.org/portal/download/all/>) and was processed using custom R scripts (<https://github.com/GryderArt/CRISPRtoolkit/>).

## Supplementary References

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Systems* 3, 99-101.

Gryder, B.E., Khan, J., and Stanton, B.Z. (2020). Measurement of differential chromatin interactions with absolute quantification of architecture (AQuA-HiChIP). *Nature Protocols*.

Gryder, B.E., Pomella, S., Sayers, C., Wu, X.S., Song, Y., Chiarella, A.M., Bagchi, S., Chou, H.-C., Sinniah, R.S., Walton, A., *et al.* (2019). Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma. *Nature Genetics* 51, 1714-1722.

Gryder, B.E., Yohe, M.E., Chou, H.-C., Zhang, X., Marques, J., Wachtel, M., Schaefer, B., Sen, N., Song, Y.K., Gualtieri, A., *et al.* (2017). PAX3-FOXO1 Establishes Myogenic Super Enhancers and Confers BET Bromodomain Vulnerability. *Cancer Discovery*.

Meyers, R.M., Bryan, J.G., McFarland, J.M., Weir, B.A., Sizemore, A.E., Xu, H., Dharia, N.V., Montgomery, P.G., Cowley, G.S., and Pantel, S. (2017). Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature genetics* 49, 1779.

Rao, Suhas S.P., Huntley, Miriam H., Durand, Neva C., Stamenova, Elena K., Bochkov, Ivan D., Robinson, James T., Sanborn, Adrian L., Machol, I., Omer, Arina D., Lander, Eric S., *et al.* (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159, 1665-1680.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology* 16, 259.

Shern, J.F., Chen, L., Chmielecki, J., Wei, J.S., Patidar, R., Rosenberg, M., Ambrogio, L., Auclair, D., Wang, J., Song, Y.K., *et al.* (2014). Comprehensive Genomic Analysis of Rhabdomyosarcoma Reveals a Landscape of Alterations Affecting a Common Genetic Axis in Fusion-Positive and Fusion-Negative Tumors. *Cancer Discovery*.