Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma

Berkley E. Gryder[1,*], Silvia Pomella[1,2], Carly Sayers[3], Xiaoli S. Wu[4,5], Young Song[1], Anna M. Chiarella[6], Sukriti Bagchi[3], Hsien-Chao Chou[1], Ranu S. Sinniah[1], Ashley Walton[1], Xinyu Wen[1,7], Rossella Rota[2], Nathaniel A. Hathaway[6], Keji Zhao[8], Jiji Chen[9], Christopher R. Vakoc[4], Jack F. Shern[3], Benjamin Z. Stanton[1,10], Javed Khan[1,*]

[1]Genetics Branch, NCI, NIH, Bethesda, MD, USA

[2]Department of Oncohematology, Ospedale Pediatrico Bambino Gesu' Research Institute, IRCCS, Rome, Italy

[3]Pediatric Oncology Branch, CCR, NCI, NIH, Bethesda, MD, USA

[4]Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

[5]Genetics Program, Stony Brook University, Stony Brook, New York 11794, USA

[6]Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, Chapel Hill, NC 27599, USA

[7]Advanced Biomedical Computational Science, Frederick National Laboratory for Cancer Research, Frederick, MD, USA

[8]Systems Biology Center, NHLBI, NIH, Bethesda, MD, USA

[9]Advanced Imaging and Microscopy Resource, NIBIB, NIH, Bethesda, MD, USA

[10]The Research Institute at Nationwide, Nationwide Children's Hospital, Columbus, OH, USA

Abstract

Core regulatory transcription factors (CR TFs) orchestrate the placement of super-enhancers (SEs) to activate transcription of cell-identity specifying gene networks, and are critical in promoting cancer. Here, we define the core regulatory circuitry of rhabdomyosarcoma (RMS) and identify critical CR TF dependencies. These CR TFs build SEs that have the largest levels of histone acetylation, yet paradoxically SEs also harbor the highest amounts of histone deacetylases

*Corresponding authors: Berkley E. Gryder, Ph.D. berkley.gryder@nih.gov, Javed Khan, M.D. khanjav@mail.nih.gov.

(HDACs). We find that hyperacetylation selectively halts CR TF transcription. To investigate the architectural determinants of this phenotype, we developed Absolute Quantification of Architecture (AQuA) HiChIP, revealing erosion of native SE contacts, and aberrant spreading of contacts involving histone acetylation. Hyperacetylation removes RNA Pol2 from core regulatory genetic elements, and eliminates RNA-Pol2 but not BRD4 phase condensates. This study identifies a SE-specific requirement for balancing histone modification states to maintain SE architecture and CR TF transcription.

Introduction

There are more than 1,500 transcription factors (TFs) encoded in the human genome[1]. Some TFs are used across all human cell types (such as the General Transcription Factors[2]), while many TFs are restricted to a particular time and place in development[3,4]. In a given cell type, a few core regulatory (CR) TFs, expressed at the highest levels, tend to dominate and determine the placement of large histone acetylation deposits, termed super enhancers (SEs)[5], which form around a mosaic array of CR TF binding sites and drive cell-type specific gene expression[6]. CR TFs are themselves driven by a subset of the SEs they form, and can be co-opted as essential dependencies in cancer[7,8].

CR TFs function by recruiting acetylation writers (CBP/p300), readers (BRD4) and erasers (histone deacetylases, HDACs), among many other co-activators, to create SEs[9]. The entire axis of histone acetylation is essential for CR TF transcription[10]. While the need to chemically add or recognize acetylation for enhancer-driven RNA Pol2 transcription is well documented[11–14], why CR TFs recruit HDAC-containing complexes to SEs is not understood.

Here, we determine and dissect the essential regulatory networks underlying childhood rhabdomyosarcoma in primary tumors and cell lines, and use this disease context to mechanistically interrogate the consequences of hyperacetylation at the chromatin template. Using a combination of RNA-seq, single-cell RNA-seq and nascent ChRO-seq, we find CR TFs have a rapid and high sensitivity to histone deacetylase inhibition. Spike-in normalized ChIP-Rx and AQuA-HiChIP shows that hyperacetylated histones spread and disrupt the three-dimensional (3D) organization of SEs, which destabilizes CR TF and RNA Pol2 binding at SEs and dissolves RNA Pol2 but not BRD4 condensate assembly. Thus, while histone acetylation is considered an "active" chromatin modification, its deposition must be tempered and controlled to facilitate SE-driven core regulatory transcription.

Results

RMS Core Regulatory Nodes Include SOX8 and are Selectively Required for Growth

To understand the epigenetic networks driving RMS, we sought to identify its regulatory circuitry. We performed analysis of SE-associated TFs across 21 RMS samples, both primary tumors and cell lines. Because RMS shares reliance on myogenic TFs, we cross-analyzed 7 samples from the muscle lineage. SEs were defined with H3K27ac ChIP-seq experiments, for which we incorporated sample-matched RNA-seq data. For a given SE-

associated TF_a (expressed at least 4 TPM in RNA-seq), the circuitry input (*in degree*) was calculated as the number of all TFs ($motif_a + motif_b + motif_c$ and so on) with a motif present in nucleosome-depleted valleys of TF_a's SE. The output for TF_a (Fig. 1a) was calculated as the total number of SEs (with putative TF target genes; SE1, SE3, SE4, and so on) which had TF_a's motif (*out degree*). Total connectivity (*in + out degree,* normalized to 1 = maximum connectivity in the sample) predicted the TFs with high connectivity, the "core" of the regulatory circuitry (Fig. 1a). In RMS samples, CR TFs formed 4 modules: (1) a pan-RMS module including MYOD and MYOG, (2) a FP-RMS only module including MYCN and FOXO1 (the SE regulating *PAX3/7-FOXO1*)[15,16], (3) a FN-RMS module unified by consistent presence of PAX7 and AP1-family TFs, and (4) a normal muscle-specific module of CR TFs including Nur77 (encoded by *NR4A1*) and MEF2D (Supplementary Fig. 1a).

CR TF prediction identified SOX8 as consistently high-scoring across all PAX3-FOXO1 samples (Fig. 1a). SOX8 was validated by ChIP-seq, which revealed that it co-localizes with the other CR TFs in FP-RMS (Fig. 1b). ATAC-seq peaks in SEs which contain SOX8 (n = 839) were more strongly bound by all other CR TFs and have the largest H3K27ac signal (Fig. 1b). SOX8 binds to 623 of 776 SEs in RH4 cells (Fig. 1c). Among SOX family members, *SOX8* was most highly expressed (Supplementary Fig. 1b) and overexpressed compared to normal tissue (Supplementary Fig. 1c). Histone acetylation network modeling placed SOX8 as a central hub (Supplementary Fig. 1d). Western blot analysis showed SOX8 present at the protein level in two primary FP-RMS tumors (Supplementary Fig. 1e). These data support the inclusion of SOX8 as a previously unrecognized component of the CRC in RMS.

Analysis of Project Achilles CRISPR data demonstrated *SOX8* and other CR TFs were essential to the growth of FP-RMS (Supplementary Fig. 1f), and uniquely so compared to 389 other cell lines across both pediatric and adult malignancies[17]. To validate this, we targeted 6 sgRNAs per DNA-binding domain of all TFs in a pooled fashion (Fig. 1d) in RH4 cells. This cell line harbors 50 CR TFs, with an average of 1.8 SEs nearby each CR TF gene locus (Supplementary Fig. 2a), which are the most connected among 113 TFs that are SE associated and not pan-essential (Supplementary Fig. 2b). We found that the CR TFs in FP-RMS, including SOX8, were critical to maintenance of cancer cell proliferation (Fig. 1e). Those CR TFs with a depletion score less than −1 (on a $\log_2$ scale) were defined as the "top" CR TFs (n = 13) in RH4 (Supplementary Fig. 2c–d). The dependence on these TFs agrees with recent evidence for myogenic factors as oncogenic in RMS[18].

Functional dissection of Core Regulatory circuitry reveals anti-myogenic attribute of SOX8 in RMS

The hallmark t(2;13)(q35;q14) translocation of FP-RMS creates PAX3-FOXO1. As the primary oncogenic driver, we considered the possibility that other CR TFs (SOX8, MYOD1, MYOG) were supporting growth of FP-RMS by contributing to *PAX3-FOXO1*'s transcription. ChIP-seq showed binding of these factors to one-another's enhancer constituents (Fig. 2a, with the exception of PAX3-FOXO1 showing no peaks near *MYOG*). To assay the 3D connectivity among these elements and their putative target genes, were performed HiChIP of H3K27ac, which identified direct interactions between SEs, nearby

typical enhancers, and each TF gene (Fig. 2a). The SE on chromosome 13 (near *FOXO1*) is connected to many additional enhancers and the *PAX3* promoter on chromosome 2. Yet, individual dissection of CR TF function (using the 2 most potent sgRNAs per CR TF from our pooled screen) showed PAX3-FOXO1's expression was not reduced by disruption of MYOD, MYOG or SOX8, while all factors were negatively regulated by disruption of PAX3-FOXO1 itself (Fig. 2b). Disruption of either MYOD or MYOG resulted in strong transcriptional depletion of *MYOD1*, *MYOG*, *SOX8*, and other less prominent CR TFs (Fig. 2b–c).

SOX8 disruption caused upregulation of *MYOD1* and *MYOG* that persisted, suggesting a unique negative-regulatory role in FP-RMS. Consistent with this notion, it has been reported that SOX8 is a marker for muscle satellite cells that inhibits myogenesis[19]. We performed gene set enrichment analysis (GSEA) for genes upregulated during differentiation from myoblasts to myotubes. This program is strongly downregulated by CRISPR-cas9 disruption of MYOD or MYOG, while perturbation of PAX3-FOXO1 or SOX8 activates these myogenic genes (Fig. 2c–d). Additionally, disruption of PAX3, MYOD or MYOG (but not SOX8) caused downregulation of CR TFs in RH4 cells (Fig. 2d). The distinct characteristics of these CR TFs was orthogonally confirmed by shRNA knock-down (Supplementary Fig. 2e–g). As PAX3-FOXO1 directly upregulates *MYOD1*, yet FP-RMS is unable to complete the MYOD-driven myogenic differentiation program, the discovery that PAX3-FOXO1 also drives *SOX8* supports a core regulatory model with both positive and negative regulation (Fig. 2e) able to lock RMS in its de-differentiated state.

HDACs are employed by CR TFs

Gene activation is strongly positively correlated with histone acetylation. The most abundant sites of H3K27ac in the RMS epigenome are bound by PAX3-FOXO1, especially SEs[15]. Yet, PAX3-FOXO1 co-immunoprecipitates with the enzymes that erase this mark, nuclear histone deacetylases (HDACs)[20]. SEs are associated with high levels of HDAC1/2 binding in mouse ES cells[5,9].

To see if PAX3-FOXO1 co-binds with HDACs on the epigenome, we performed ChIP-seq on all three nuclear HDACs (1, 2 and 3). We found that each of these HDACs co-associated with PAX3-FOXO1 and other CR TFs at SEs (Figure 3a). Increased HDAC genomic binding predicted greater expression (independent of CR TF binding), and CR TF targets were much more often found among genes with the highest levels of HDACs at their promoters (Fig. 3b). We next sub-classified genes by transcriptional output, finding more HDAC binding to the epigenome as expression increased, with the strongest association at CR TF targets (Supplementary Fig. 3a). We found that HDAC binding was strongest at the approximately 9,000 sites co-bound by all three HDACs, and that 91% of these triply bound sites were co-occupant with CR TFs (Supplementary Fig. 3b). Our data highlighted that HDAC1/2/3 binding is positively correlated with CT TF networks at the chromatin level. This raised the question: is the removal of excess acetylation essential for transcription at core regulatory domains?

Dynamics of CR TF Transcription upon HDAC inhibition

The high-levels of HDACs at loci encoding CR TF genes could translate into a selective transcriptional vulnerability. RNA-seq after treatment with an inhibitor selective for nuclear HDACs, Entinostat, revealed a marked decrease in expression of *SOX8*, *MYOD1*, *MYOG* and *MYCN* (Supplementary Fig. 3c). All core regulatory genes in RH4 cells, especially the top CR TFs, were selectively downregulated by inhibition of HDACs 1/2/3, while housekeeping genes, other TFs, or even SE genes generally were not downregulated (Fig. 3c). Interestingly, the first hour of HDAC inhibition resulted in increased expression of CR TFs, before a sharp decrease at 6 hours and continued suppression at 24 hours (Fig. 3d, Supplementary Fig. 3d). Genetic depletion of individual HDACs revealed incomplete transcriptional hinderance of CR TFs (Supplementary Fig. 3e). To examine if selective changes in CR TFs was owing to differences in nascent transcription at these loci (as opposed to differences only in RNA transcript stability), we performed chromatin run-on and sequencing (ChRO-seq)[21] in RH4 cells treated with Entinostat for 10 minutes, 1 hour and 6 hours. We found a rapid initial increase and eventual decrease in nascent transcription at CR TF loci, exemplified at *MYOD1* and *MYOG* (Fig. 3e) and seen across all CR TFs (Fig. 3f). Furthermore, we calculated relative mRNA stability using combined analysis of ChRO-seq and total RNA-seq[22], and find indeed that CR TF transcripts are much less stable (Supplementary Fig. 3f–h), allowing relatively tight kinetics between transcriptional changes and net mRNA reduction, especially for the least stable transcripts such as *MYCN* (as reported for *MYC* in previous studies[23]).

As enhancers control transcriptional bursts[24], we hypothesized the initial increase at 10 minutes and 1 hour did not represent an increased number of transcripts in each cell, but rather an increase in the proportion of cells actively engaged. Using single-cell RNA-seq (scRNA-seq) we tested and confirmed that the proportion of single cells expressing any given CR TF increased (Fig. 3g) while the expression levels of any given TF did not increase. It appears that co-expression of any two specific factors (considering *SOX8* and *MYOD1* as an example, Figure 3g) is uncommon, but we interpret this as an artifact of sparsity and limited detection in scRNA-seq. The increase of apoptotic genes, unlike the increase of CR TFs, was the result of an increased quantity of reads per cell, not only an increased proportion (Supplementary Fig. 3i); also, while CR TFs decreased at 6 hours, apoptotic genes continued to increase. Therefore, the scRNA experiments indicate that acute hyperacetylation does not necessarily increase the maximum output of SE controlled CR TFs, but activates transcriptionally resting cells, before the eventual halting of CR TF transcription at later time points.

Enhancer Spreading and Loss of Architectural Integrity at CR TFs

We hypothesized that increased histone acetylation at CR TFs could explain the initial burst and subsequent crash of CR transcription. In previous studies of H3K27ac changes with HDAC inhibitors, ChIP-seq showed drug-induced spreading of acetylation, and yet a decrease in acetylation at enhancer peaks[25,26]. We reasoned that, because these previous studies normalized by sequencing depth, a global increase in H3K27ac could not have been detected even if it is was occurring, preventing correct interpretation of the data. To evaluate this, we incorporated reference exogenous *Drosophila* chromatin (ChIP-Rx)[27]. Indeed,

analysis and normalization without spike-in gave an apparent decrease in H3K27ac at SEs, whereas reference normalization revealed a strong increase (Supplementary Fig. 4a). The boundaries of SEs were stable within the first hour of HDAC inhibition with focal increases in acetylation. In contrast, by 6 hours, we observed spreading (or, outward diffusion) of acetylated chromatin beyond SE boundaries (Fig. 4a).

CR TFs require looping machinery, such as cohesin and YY1, to bring TF-bound SEs into close spatial proximity to their gene body, to insulated neighborhoods by CTCF[28,29]. Drug-induced hyperacetylation at SEs, constituting new chemical composition of the chromatin template, may alter looping. To interrogate this, we began with ChIP-Rx of YY1 and BRD4, in RH4 cells treated with DMSO or Entinostat for 6 hours. YY1 and BRD4 bound more both at CTCF sites and the surrounding chromatin upon treatment with Entinostat for 6 hours (Supplementary Fig. 4b–c). YY1 primarily occupied SE constituents and target promoters whereas RAD21 (cohesin) preferentially bound to CTCF-occupied sites (insulators, Supplementary Fig. 4d). Upon HDAC inhibition, RAD21 increased subtly, whereas YY1 loading increased markedly (Supplementary Fig. 4d). Similarly, BRD4 showed increased binding, while acetylation writer p300 and erasers HDAC2 and HDAC3 were mildly reduced (Supplementary Fig. 4d). HDACi caused accessibility (ATAC-seq) increases at nucleosome depleted valleys in SEs (Supplementary Fig. 4d) flanked by regions of H3K27ac spread highlighting the connection between genomic deposition of H3-ac and accessibility.

Molecules that inhibit HDAC isoforms non-selectively have been used in prior studies that found histone acetylation spreading.[25,26] Entinostat also achieves spreading, but inhibits only the Class I nuclear HDAC isoforms 1/2/3. HDACs 1/2/3 co-occupy SEs with near identical patterning (Supplementary Fig. 5a). In an attempt to further narrow-down the isoforms of HDAC required for CR TF transcription, we employed inhibitors selective for HDAC1/2 (Merck60) or HDAC3 (LW3). Strong downregulation of *MYOD1* (Supplementary Fig. 5a) required pharmacological inhibition of all three Class I HDACS, and either HDAC1/2i of HDAC3i alone were unable to increase histone acetylation to the same extent as HDAC1/2/3i (Fig. 4b, Supplementary Fig. 5b). We confirmed that 3 days of CRISPR-mediated individual disruption of HDAC1, HDAC2 or HDAC3 caused only a slight increase in H3K27ac, and no spreading (Supplementary Fig. 5c). HDAC1/2/3 triple inhibition with Entinostat increases histone acetylation not only at H3K27, but also H2BK5 and H4K16, and to a lesser extent at H3K36 (Fig. 4b).

To test if 2-dimensional SE spreading was associated with aberrant engagement of SEs to loci encoding CR network promoters or other genomic targets, we devised a modified HiChIP[30,31] to capture global changes in contact frequencies that may be masked without external normalization. Termed Absolute Quantification of Architecture (AQuA)-HiChIP, the method incorporated the addition of an identical amount of fixed mouse cells to human cells immediately prior to *in situ* Hi-C contact generation. After enrichment (in this case H3K27ac) with an antibody possessing cross-species reactivity, contacts were captured by biotin and library preparation was performed on streptavidin beads. The ratio of mouse and human contact pairs was quantified, and AQuA contact frequency (reference normalized contacts per million) allowed us to define absolute changes in 3D SE dynamics upon

HDACi. Indeed, the hyperacetylation structures in normal HiChIP were masked, but uncovered by AQuA-HiChIP (Supplementary Fig. 6a).

AQuA-HiChIP in RH4 cells revealed that upon HDACi, new interactions invade unmodified chromatin, such as the 40-kb gap between *MYOD1* SEs (Fig. 4c). These new interactions were not observed in classical HiChIP but were uncovered by the AQuA-HiChIP methodology. While many aberrant new interactions grow, the prominent SE-to-SE interaction at *MYOD1* is diminished (Fig. 4d). The apparent "spreading" of AQuA-HiChIP is greater than that seen in ChIP-Rx (Supplementary Fig. 5c), but each 3D contact pair can be comprised of both a direct and DNA-ligation-mediated association with an acetylated histone. The spreading phenomenon was observed at other CR TFs regulated by one or more SEs, such as *SOX8* (Supplementary Fig. 6b), *PAX3-FOXO1*, *FOXM1*, *JUN*, *MYCN*, *MYOG*, *RARA* and *SIX2* (Supplementary Fig. 6c). New hyperacetylation induced interactions are seen at SE-pairs genome-wide but are confined by CTCF boundaries (Fig. 4e).

The loss of SE-contacts, and the gain of excessive aberrant contacts, could help explain why transcription of CR TFs is most sensitive to HDAC inhibition. But if the role of HDACs at SEs is to prevent spreading, why are they bound to the epicenters (with p300) rather than the boundaries? We reasoned that the shape of binding indicates nucleosome movement and turnover[32]: p300 and HDACs do not directly overlap H3K27ac, but are immediately adjacent, and the signal of H3K27ac at SEs tapers off over distances (> 2,000 bp) much greater than a single nucleosome (147 bp). Thus, interfering with the catalytic balance of acetylation writers by inhibiting HDACs at the epicenter could then affect spreading at distal boundaries in a tapered fashion by increasing the local concentration of acetylated histones. This agrees with the shape and breadth of spreading induced by Entinostat. To directly recapitulate transcriptional downregulation by increased in histone acetylation past the boundary regions at a single SE, we chose to navigate p300 to specific locations using dCas9 with chemical induced proximity binding[33]. Thus, we engineered RH4 cells to stably express dCas9 and MS2-FKPB to enable sgRNA directed recruitment of endogenous p300 via a bifunctional small molecule CEM114 (FK506 linked to a binder of the bromodomain of p300)[34]. Targeting the SE epicenters had little impact on transcription (possibly because p300 abundance was already high at these locations), but recruitment of p300 past the SE boundaries caused downregulation of *MYOD1* (and, only upon co-administration of 50 nM CEM114, Fig. 4f).

HDAC inhibition decommissions binding of CR TFs and RNA Pol2 at SEs

To further interrogate the directness of the effects of HDAC1/2/3 inhibition, we next performed ChIP-Rx of PAX3-FOXO1, MYOD, SOX8, MYOG after 6 hours with Entinostat. We then ranked each CR TF binding site by the total amount of change, finding PAX3-FOXO1, SOX8 and MYOD to be reduced at the SEs regulating these important CR TFs (Fig. 5a–b), while MYOG showed very subtle changes in binding (data not shown). At the same time, hyperacetylation caused RNA Pol2 to be largely removed, exemplified at *MYOD1* (Fig. 5c), which loses H3K27ac in the gene body and undergoes H3K27ac spreading into the TES region. The loss of RNA Pol2 at CR TFs coincides with increased

H3K27ac in the transcriptional end site region (TESR); this effect was not seen at typical TFs (Supplementary Fig. 7a). The ratio of genic H3K27ac spread was greatest for shorter TFs (including *SOX8*, *MYOG*, *MYOD1* and *MYCN*), compared to all Pol2 bound genes, regular TFs or even other SE genes (Supplementary Fig. 7b–c). Reduction of Pol2 in the gene body and at the TES could occur by inhibition of pause release, but HDACi-induced pausing is ruled out at *MYOD1* as the decrease did not coincide with an increase of Pol2 at the TSS-proximal pause site. Rather, unloading of Pol II was seen at the promoter, TSS region and the gene body. CR TFs exhibit more elongation in the ground state than other genes (Supplementary Fig. 7d, left), but while HDAC inhibition induces pausing genome-wide (Supplementary Fig. 7d) in agreement with previous work[25], this does not occur at CR TFs (Supplementary Fig. 7d, right).

Disruption of RNA-Polymerase clusters

While many genomic sites lose RNA-Pol2 upon HDACi, Pol2 clusters found at SEs are most disrupted (Supplementary Fig. 7e, *t* test, $P < 2.2 \times 10^{-16}$). We hypothesized that such clusters would be visibly dissipated by HDACi, reflecting the ChIP-Rx observations. Recent high-resolution imaging of RNA-Pol2 in live cells revealed a clustering of up to 80 molecules of RNA-Pol2[35], and these clusters reside in liquid-liquid phase separated droplets at SEs[36]. Using RH4 cells stably transduced with RNA polymerase II subunit RPB3 tagged with GFP, we identified RNA Pol2 clusters in single cells that were stable over time (Fig. 5d). These RNA-Pol2-GFP clusters were rapidly dissipated upon HDAC inhibition (Fig. 5d). The size and brightness of RNA-Pol2 puncta is asymmetrically distributed, similar to SEs, and is reduced by HDACi (Supplementary Fig. 7f).

PAX3-FOXO1 and the other CR TFs both recruit and rely on BRD4 to mediate their output at SEs. To evaluate if BRD4 forms condensates in RMS, we knocked-in mEGFP to the endogenous human *BRD4* gene in RH4 cells, finding it forms the predicted structures. In contrast to rapid disappearance of RNA Pol2 puncta, BRD4 puncta are resilient to hyperacetylation. These visual changes were reflective of ChIP-Rx binding changes: while RNA Pol2 is lost at CR TF gene loci, BRD4 is gained (Fig. 5f). Considering many relevant factors by ChIP-Rx, Entinostat reduces binding of 3 of the 4 essential CR TFs measured, reduces RNA Pol2, while the competing acetylation enzymes (p300, HDACs) and are stable, with increases in YY1, RAD21 and BRD4 (Fig. 5g). These divergent responses to hyperacetylation may be undergirded by rules of inclusion or exclusion from phase condensates at SEs that future work can elucidate.

The loss of SE architectural integrity by hyperacetylation of histones and increased accessibility, with concomitant dissipation of RNA-Pol2 clusters, provides a plausible mechanism by which a drug with a non-gene-specific target (HDAC) could yield a focused impact on only certain genes which are particularly dependent on co-localization with large enhancer elements (Fig. 5h). Thus, 3D genome architecture can be targeted as a vulnerability in RMS via modulation of HDAC activity.

Discussion

Core regulatory TFs not only govern epigenetic status in FP-RMS, they also represent disease-critical targets[5,7]. We discovered and confirmed developmental gene *SOX8* is critical among CR TFs. SOX8 regulates early neural crest development[37]. While CR TFs MYOD and MYOG are pro-myogenic differentiation, in RMS as in satellite muscle cells, SOX8 counteracts the ability of these factors to complete the muscle lineage. As SOX8 has SEs bound by PAX3-FOXO1, it provides a mechanistic link to the anti-differentiation activity of the fusion. SOX8 is reminiscent of SOX2 among CR TFs in embryonic cell circuitry[4], as both factors promote the de-differentiated state, yet these are distinct in that SOX2 is a positive regulator of embryonic CR TFs, and SOX8 is a negative regulator of RMS CR TFs. This opens an important avenue of future research in rhabdomyosarcoma biology.

Phase condensates form compartments of concentrated components needed for biochemical reactions[38], and SEs perform this function at the chromatin level[39]. Furthermore, oncogenic TFs can position phase condensates[40]. The choice of components (which proteins, nucleic acid polymers and small molecules exist in a phase condensate) is driven by biophysical properties, and post-translational acetylation is a means of decreasing a fundamental property, charge. Our data show that SEs do not fit the classic paradigm that more acetylation and decreased HDAC activity are associated with more transcription. We instead suggest a paradigm whereby at critical CR TF loci, transcription cannot be maintained without the counterbalance to acetyltransferases, HDACs. This model is supported by HDACi-induced erosion of SE boundaries, the invasion of acetylated histones into surrounding unmarked chromatin, the dissipation of 3D enhancer loops, the unloading of RNA-Pol2 at CR TFs and dissolving of RNA-Pol2 clusters upon hyperacetylation. When considered together, these experiments begin to shed light on why HDACs are essential to CR network circuitry. Furthermore, this offers a new mechanism of action for reinterpreting historical studies of HDACi in cancer therapeutics, perhaps especially for TF-driven cancers. Future experiments will help determine how overall charge balance and distribution of modifications on the disordered histone tail influences the formation of larger condensates and modulates transcription. These properties will shape an important new framework for interpreting chemical epigenomics.

Methods

Cell Lines and Primary Tumors

Cell lines were tested for mycoplasma within one or two passages of each experiment and found to be negative, and cell line identities were ensured by RNA-seq and genotyping. RH4, RH3, RH5 and RH41 were kind gifts from Dr. Peter Houghton, SCMC from Dr. Janet Shipley, RD, CTR and Birch from Dr. Lee Helman. NIH3T3 and HEK293T cells were purchased from ATCC. Validation was performed by DNA fingerprinting AmpFlSTR Identifiler PCR Amplification Kit (Catalog Number 4322288) by Life Technologies. Cell lines were grown at 5% CO_2 and 37 °C in DMEM supplemented with 10% fetal bovine serum (FBS) and pen/strep. Primary tumors were acquired via the NCI coordinated ClinOmics protocol as previously described[41].

ChIP-seq and ChIP-Rx

ChIP-seq was performed as previously described[42,43]. Briefly, cells or tumor tissue was formaldehyde fixed (1% final concentration) for 14 minutes, dounce homogenized, pelleted and resuspended in ChIP buffer with protease inhibitors (Active Motif, Cat# 53040). Then, samples were sheared for 27 cycles (1 cycle = 30 seconds of sonication, 30 seconds resting) with the Active Motif EpiShear Probe Sonicator. Sheared chromatin samples were immunoprecipitated overnight at 4 ˚C with antibodies (listed in the Supplementary Table 6) and purified by Agarose beads (Active Motif ChIP-IT Sigh Sensitivity Kit, Cat# 53040). For ChIP-seq with reference exogenous spike-in chromatin, we added *Drosophila* chromatin (Active Motif, Cat# 53083) with a *Drosophila*-specific antibody (Active Motif, Cat# 61686) to each condition, careful to keep both the amounts of human chromatin and *Drosophila* chromatin constant across drug or vehicle treated samples. In each experiment, validation of enrichment was assessed using qPCR with the ChIP-IT qPCR Analysis Kit (Active Motif, Cat# 53029) and the primers listed in Supplementary Table 1. ChIP-seq or Chem-seq DNA libraries were made with Illumina TruSeq ChIP Library Prep Kit, DNA was size selected with SPRIselect reagent kit (250–300 bp insert fragment size). Libraries were multiplexed and sequenced using NextSeq500 High Output Kit v2 (75 cycles), cat. # FC-404–2005 on an Illumina NextSeq500 machine. 25,000,000–35,000,000 reads were generated for each ChIP-seq and Chem-seq sample.

ATAC-seq library preparation

Assays of transposase-accessible chromatin (ATAC) were performed as previously described[44,45]. Briefly, 50,000 cells were isolated, and nuclei were generated by incubating on ice with 500 μl lysis buffer (RSB with 0.1% Tween-20) for 10 min. The resulting nuclei were centrifuged at $500 \times g$ for 10 min, and resuspended in 1× Tagment DNA buffer (Illumina) with 2.5 μl Tagment DNA Enzyme (Illumina) and incubated at 37 °C for 30 minutes. For each transposition reaction, the volume was 50 μl. The transposition mixtures were quenched with 500 μl PB buffer (Qiagen) and purified by standard protocol with MinElute PCR purification kit. Each ATAC library was amplified with Nextera primers for 16 PCR cycles and purified with Agencourt AMPure XP (Beckman Coulter) to remove excess primers. The resulting ATAC libraries were sequenced with NextSeq500 with paired-end reads.

Analysis of ChIP-seq and ATAC-seq data

ChIP-seq and ATAC-seq reads were aligned to the human genome version hg19 using BWA version 0.7.17, and was visualized in IGV after extended to the average fragment size and binning to 25 bp using IGVtools -count function. Samples with *Drosophila* spike-in were additionally normalized to reads per million mapped dm3 reads[27]. Peaks were called using MACS2 (version 2.1.1.20160309, https://github.com/taoliu/MACS) using "narrow" mode for all targets reported in this paper, as they form sharp genomic peaks (rather than broad swaths, as is seen for H3K27me3). Parameters for MACS2 usage: [--format BAM --control input.bam --keep-dup all --pvalue 0.0000001]. Regions called as peaks which are known to be spurious mapping artifacts were removed before any further analysis (reference locations for sites black-listed by the ENCODE consortium, https://sites.google.com/site/

anshulkundaje/projects/blacklists). Motif analysis was performed on peaks called from MACS2, using findMotifsGenome.pl from HOMER version 4.9.1 (http://homer.ucsd.edu/homer/index.html). Super enhancers (SEs) were identified using the ROSE2 package (https://github.com/linlabbcm/rose2) employing stitching parameter of 12,500 bp. In analysis steps considering signal in "Core Regulatory Domains", we defined the domain of each CR TF by stitching together its gene body with all associated SEs within 200 kb.

Core regulatory transcription factor network analysis

The degree of inward and outward binding (ie, TF network prediction for core regulatory transcription factors) was analyzed from SE calls in RMS cell lines and tumors for which H3K27ac ChIP-seq data and paired RNA-seq data were available. The scripts used are available here (https://pypi.org/project/coltron/) and require ROSE2 super-enhancer calls for input. Briefly, the following computational steps for each sample are performed: a list of active TFs was generated RNA-seq (cutoff = 4 TPM), then each super-enhancer from ROSE2 is associated with the nearest expressed TF (distance cutoff = 0.5 Mb), then a list of coordinates for each candidate TF is generated, then valleys within SEs are identified using bamliquidator (https://github.com/BradnerLab/pipeline/wiki/bamliquidator), then FASTA sequences for each valley are generated, then TF motifs are identified and network connection are mapped, from which each TFs "in degree" and "out degree" are tabulated[8].

RNA-seq sample preparation and data analysis

Briefly, RNA was extracted from RMS cell lines, or tumors, as well as cell lines treated with drugs, using the RNeasy mini kit (Qiagen). Poly-A selected RNA libraries were prepared and sequenced on an Illumina HiSeq2000. RNA-seq reads were aligned to hg19 using STAR version 2.5.3a. Gene expression values were calculated as Transcription Per Million (TPM) or Fragments Per Kilobase Million (FPKM) using RSEM version 1.3.0 and the UCSC hg19 reference at the gene level. For gene set enrichment analysis (GSEA)[46] the TPM values for each protein coding gene were compared to DMSO by $\log_2$ fold-change comparison, followed by rank-ordering. Bubble-plots of enrichment output from GSEA analysis of custom and public gene sets were created in R using custom scripts (https://github.com/GryderArt/VisualizeRNAseq).

ChRO-seq

Chromatin Run On followed by massively parallel sequencing (ChRO-seq) was performed on formaldehyde-fixed (same as ChIP-seq) RH4 cells treated with DMSO (6 hours), or Entinostat (10 minutes, 1 hour or 6 hours). All buffers were made in RNase-free water, and extreme care was taken to maintain a complete RNase-free environment, including frequent use of RNaseZap (Ambion, cat # AM9780) on all surfaces, containers and gloves. Drug treated, fixed and frozen cell pellets were thawed and resuspended in 1 ml of freshly prepared, ice-cold NUN buffer (0.3 M NaCl, 1 M Urea, 1% NP-40, 0.2 mM EDTA, 7.5 mM $MgCl_2$, 20 mM HEPES, 1 mM DTT, 1× PIC, 20 units/ml SUPERase In RNase Inhibitor), vortexed vigorously for 30 seconds, then 400 μl additional NUN buffer was added and vortexed for 30 seconds. Samples were incubated for 30 minutes at 4 °C and 1,500 RPM in a Thermomixer, centrifuged for 30 minutes at 4 °C and 12,500 × g to pellet chromatin. NUN supernatant was removed, and pellet was washed 3 times with 50 mM Tris-HCl, centrifuged

at $10,000 \times g$ (5 minutes, 4 °C). Samples were then resuspended in 50 μl cold Storage Buffer (50 mM Tris-HCl, 25% glycerol, 5 mM MgAc$_2$, 0.1 mM EDTA, 5 mM DTT, 40 units/ml SUPERase) and sonicated (30 second cycles, 30% amplitude, 5 cycles, kept cold). Run-on was performed at 37 °C for 5 minutes after adding 50 μl of Run-On master mix (containing 60 μM biotin-11-CTP, 60 μM biotin-11-UTP, 400 μM ATP, 400 μM GTP, 10 mM Tris-HCl, 5 mM Mg, 300 mM KCl, 0.8 units/μl SUPERase, 1% Sarkosyl and 1 mM DTT), and stopped with 500 μl Trizol LS and placed on ice. Adapters with a random hexamer was used as a unique molecule index (UMI). Base hydrolysis, streptavidin bead binding, RNA adapter ligation and reverse transcription steps were performed as previously described[21]. Sequencing was performed in paired-end mode to a depth of ~24 million reads per condition. Reads clean-up and processing was performed using the Proseq 2.0 pipeline (https://github.com/Danko-Lab/proseq2.0) and visualized in IGV. ChRO-seq sample statistics are available in Supplementary Table 5.

Estimation of mRNA stability

To calculate an estimate of mRNA stability and half-life, we compared RNA (TPM values calculated as above from total RNA-seq, exons and UTRs only) and nascent RNA-seq from ChRO-seq experiments (nascent TPM, with reads counted for the whole transcript including exons, introns, UTRs and transcriptional end site regions). We adapted the same assumptions as reported in a recent preprint from the Danko and Siepel labs[22] (except that we used ChRO-seq data in place of PRO-seq data), namely that the relative decay rates of transcripts is proportional to the ratio of RNA production (nascent RNA, in TPM) divided by the steady state transcript levels (mRNA). We then applied a first order half-life equation $T_{1/2}{}^{CR} = \ln(2)/(\text{mRNA decay rate})$, where "$CR$" denotes estimation from ChRO-seq/RNA-seq data. We used matched datasets from RH4 cells treated with DMSO for 6 hours to define relative mRNA half-life values for protein coding genes in FP-RMS.

Pooled CRISPR screening

The design of human transcription factor (TF) sgRNA pooled libraries was described in a previous study[47]. Briefly, annotations of TF DNA binding domain were retrieved from NCBI Conserved Domains Database. In average, six different sgRNAs were designed for targeting each individual domain of 1,427 TFs in human genome. All the sgRNAs with high predicted off-target effect were then excluded. Specific domain targeting and positive/negative control oligonucleotides were synthesized in duplicate using an array platform (Twist Bioscience) and then PCR cloned into BsmbI-digested LRG2.1 vector using Gibson Assembly kit (NEB). Deep-sequencing analysis were then performed on a MiSeq instrument (Illumina) to verify that 100% of the designed sgRNAs were cloned in the LRG2.1 backbone and the abundance of >95% of the sgRNA was within 5-fold of the mean.

CRISPR-based negative selection genetic screenings were performed in non-clonal rhabdomyosarcoma cell lines with stable Cas9 expression (LentiV-Cas9-Puro vector). Lentiviral of pooled sgRNA library for TF library was produced in HEK293T. Briefly, HEK293T cells were transfected with pooled TF library sgRNAs and helper packaging plasmid (pVSVG and psPAX2) with polyethylenimine (PEI 25000) transfection reagent to produce lentivirus. HEK293T cells were plated one day before transfection with 70%−80%

confluency in 10 cm tissue culture dish. For one 10 cm dish of HEK293T cells, 10 μg of plasmid DNA, 5 μg of VSVG, 7.5 μg of psPAX2, and 32 μl of 1 mg/ml PEI were mixed, incubated and added to the cells. Fresh media were changed 8 hours post-transfection, and lentivirus-containing supernatant was collected at 24, 48, 56 and 72 hours post-transfection and pooled together.

For lentiviral infection of TF library, target rhabdomyosarcoma cells were mixed with the virus and 4 μg/ml polybrene, and then centrifuged at 1,700 rpm for 20 minutes at room temperatue. Fresh media was changed at 24 hr post-infection. Virus titer was measured by infectiong cells with serially diluted virus. To ensure a single copy sgRNA transduction per cell, multiplicity of infection (MOI) was set to 0.3~0.4. To maintain the representation of sgRNAs during screen, the number of sgRNA-positive cells was kept at least 1,000 times the number of sgRNAs in the library. Cells were harvest at initial (day 3 post-infection) and final (around 12 doubling times after the initial passage) time points. Genomic DNA was extracted using QIAamp DNA mini kit (QIAGEN) according to the manufacturer's instructions.

Sequencing libraries were constructed as described previously[48]. Briefly, sgRNA cassette (~200 bp) was amplified by PCR from genomic DNA, followed by end-repair with T4 DNA polymerase (NEB B02025), DNA Polymerase I, Large (Klenow) Fragment (NEB M0210L), and T4 polynucleotide kinase (NEB M0201L), and addition of 3' A-overhang with Klenow Fragment (3'−5' exo-) (NEB). The DNA fragment then were ligated to diversity-increased custom barcodes with Quick ligation kit (NEB M2200L), and attached with Illumina paired-end sequencing adaptors with Phusion master mix (Thermo Fisher Scientific F548L). The final libraries were pooled together in equal molar ratio and sequenced by MiSeq (Illumina) with MiSeq Reagent Kit v3 (Illumina).

The sequencing data were de-multiplexed and trimmed to contain only the sgRNA sequence cassettes. The read counts of each individual sgRNA was calculated, allowing no mismatches to the reference sgRNA sequence as previously described[48]. The total read counts were normalized to each sample. Individual sgRNAs with the read count lower than 50 in the initial time point were discarded. The average $\log_2$ fold-change of abundance of all sgRNA against a specific domain/gene was calculated. Data for TF dependencies, paired with RNA-seq expression of TFs in RMS cells, along with Achilles data used to categorize "pan-essential TFs" as those TFs depleted across all cell lines with a median CRISPR depletion score of less than −0.4, is available in Supplemental Table 4. Follow-up CRISPR experiments used the top 2 sgRNA sequences from the pooled screen, per target, and these lentiviral constructs are available upon request.

shRNA lentivirus production and cell infection

HEK293T were transfected with either pLKO.1-puro shRNA Scramble or specific shRNA (MYOD, MYOG, MYCN, SOX8) and helper packaging plasmid (pVSVG and psPAX2) with Fugene 6 (Sigma Aldrich). Transfection medium was replaced 24 hours later and 48 hours after transfection the lentiviral containing medium was collected, spun to remove cell debris, and the supernatant filtered. For lentiviral infection, RH4 were transduced at a 5–10

multiplicity of infection (MOI) for 24 hours with polybrene (5 μg/ml; Sigma) and 10% FCS. After a further 24 hours cells were selected with 0.75 μg/ml puromycin for 2 days.

Small molecule compounds

All molecules were dissolved in DMSO to a final concentration of 10 mM, and diluted to a final DMSO concentration of <0.03% by volume in DMEM for cell culture experiments. Entinostat was generously supplied by Developmental Therapeutics Program (NCI, NIH).

Single-cell RNA-sequencing

RH4 cells were treated with either DMSO for 1 h or with 1 μM Entinostat for 1 h or 6 h. Drop-seq was subsequently performed as previously communicated[49]. Briefly, the treated cells were trypsinized, washed, filtered through a 40-μm filter, and resuspended in PBS with 0.01% BSA at a concentration of 100 cells/μl. Barcoded beads (ChemGenes Cat# Macosko-2011–10) were resuspended in complete lysis buffer at a concentration of 120 cells/μl. Cells, beads and oil (Bio-Rad #186–4006) were loaded into syringes and then pumped through a microfluidic chip (FlowJEM) at a rate of 4 ml/h (beads and cells) or 12 ml/h (oil) to create nanoliter-sized droplets containing a single bead and/or cell. The outflow of droplets was collected. Droplets were disrupted with perfluorooctanol (Sigma Cat# 370533–25G) to release the beads, which were then washed several times. Reverse transcription of the mRNAs bound to the capture beads was performed using Maxima H-RTase (Thermo Scientific Cat# EP0753) and Template Switch Oligos (Supplementary Table 3). Single-stranded nucleotides were removed by exonuclease I treatment (Thermo Scientific Cat# EN0581). Beads were counted, and twenty PCR reactions each containing 8,000 beads were prepared and performed using 2× Kapa HiFi Hotstart Readymix (Kapa Biosystems Cat# KK2605) and a custom SMART PCR primer. The resulting cDNA was purified and concentrated using AMPure XP beads (Beckman Coulter Cat# A63880). Quality of the purified cDNA was assessed using an Agilent TapeStation, and then a DNA library was prepared on 600 pg of the cDNA using a Nextera XT kit (Illumina Cat# FC-131–1024) with the custom New-P5-SMART PCR hybrid oligo. The DNA library was purified using AMPure XP beads and sequenced using a NextSeq 500/550 High Output v2 kit (75 cycles) (Illumina Cat# FC-404–2005) using a custom Read 1 primer. For each experimental condition, mRNA captured from around 8,000 cells was sequenced on an Illumina NextSeq500.

Single-cell bioinformatic analysis

Sequencing data were processed following the Drop-seq Computational Protocol and using the Drop-seq software tools v.1.2 from the McCarroll lab (http://mccarrolllab.com/wp-content/uploads/2016/03/Drop-seqAlignmentCookbookv1.2Jan2016.pdf) using the default parameters. Barcodes were extracted and reads were aligned to the hg19 genome. Cells containing more than 500 unique genes were kept for downstream processing. The Seurat version 2.3.0 R package was used for downstream data analysis. The percent mitochondrial reads were calculated and cells having more than 15% mitochondrial reads were discarded from further analysis. The gene expression measurement per cell was normalized by the total expression and multiplied by the scaling factor of 10,000 and log transformed. A canonical correlation analysis was run to combine results from all three single-cell runs. Core

regulatory TF list used in this analysis was defined as the complete candidate list of SE regulated TFs in RH4 cells without any thresh-holds for inward or outward binding.

AQuA-HiChIP

For a detailed protocol of the AQuA HiChIP protocol, please see the associated protocol exchange (https://protocolexchange.researchsquare.com/article/nprot-7121/v1). Briefly, NIH3T3 (mouse) cells were grown and fixed in parallel with human cancer cells (RH4) treated with DMSO or Entinostat for 6 hours. 2 million fixed mouse cells were mixed with 7 million fixed RH4 cells, for each condition. Gentle lysis was used to release the nuclei that were then permeabilized in 0.5% SDS (10 minutes at 62 °C) and then quenched with 10% Triton X-100, followed by digestion with Mbol restriction enzyme (2 hours at 37 °C) and heat inactivated (20 minutes at 62 °C). Biotin was incorporated (biotin-14-dATP, Thermo Scientific, Cat# 19524–016) with DNA Polymerase I, Large (Klenow) Fragment (NEB, Cat# M0210) for 1 hour at 37 °C, followed by *in situ* ligation with T4 DNA ligase (4 hours, room temperature). Nuclei were then pelleted, sonicated (11 minutes shearing 'on' time with 30 seconds 'on' 30 seconds off, using *Active Motif* Epi-shear probe sonicator at 30% power), and immunoprecipitated with cross-species reactive antibody against H3K27ac (Active Motif, Cat# 39133). After overnight incubation (rotating, 4 °C), the ChIP reactions were buffer exchanged into Dynabeads Protein A (ThermoFischer Scientific, Cat# 10002D) and incubated (2 hours, 4 °C). ChIP reaction was then washed, eluted, treated with Proteinase K (45 minutes at 55 °C), and crosslinks reversed by increasing temperature to 67 °C for 2 hours, and purified. Biotin capture and washing was performed with Dynabeads M-280 Streptavidin (ThermoFischer Scientific, Cat# 11205D), followed by on-bead library preparation (end repair, A-tailing, adapter ligation and library amplification)[50]. Amplified library DNA was purified with AMPure beads, quantified and multiplexed, and paired-end sequenced with Illumina NextSeq High Output v2 kit (150 cycles, Cat# FC-404–2002) to generate 100–200 million read pairs per condition.

AQuA-HiChIP bioinformatic analysis

A step-by-step bioinformatic protocol for AQuA-HiChIP, with associated code, is available (https://github.com/GryderArt/AQuA-HiChIP). Briefly, paired end reads were mapped once to the human genome (hg19), and in parallel also mapped to the mouse genome (mm10), using bowtie2 within the HiC-pro pipeline[51]. Mapped read pairs were then filtered first for junction validity (containing two fragments from restriction digestion), then duplicates were removed. Contact maps were converted to Juicebox compatible *.hic file format using the hicpro2juicebox.sh tool (https://github.com/nservant/HiC-Pro) provided with HiC-pro, and matrices were extracted using the Juicer[52] dump tool (https://github.com/theaidenlab/juicer/wiki/Data-Extraction). AQuA factor incorporation was done by calculating the Contacts Per Million (CPM = Raw contact frequency matrix × 1,000,000/total valid contacts), then converting to Reference normalized Contacts Per Million (RCPM = CPM × hg19 valid contacts/mm10 valid contacts). Virtual 4C analysis was performed by isolating a single row by coordinate of interest (the virtual viewpoint anchor) from a dense RCPM normalized matrix within a region of interest. For Aggregate Peak Analysis (APA) plots, Juicer was used (juicer_tools apa -r 10000,5000 -k NONE -n 0) to generate raw APA data matrices (sum of contact frequencies for all pairs of SE to SE peaks, both inter-domain and intra-domain, in

both DMSO and HDAC treated cells) were normalized to RCPM, then plotted in R using pheatmap. For a detailed bioinformatic protocol, see ref [53].

CEM activation

RH4 cells were transduced with dCas9-HA (with blasticidin resistance) and MS2-FKBPx2 (wit hygromycin resistance) and selected with 20 μg/ml blasticidin and 200 μg/ml hygromycin, and expanded for 3–4 passages. Once established, we used RH4-dCas9-MS2-FKBPx2 cells for CEM (chemical epigenetic modifier) activation using (1) guide RNAs with MS2 loops to bind MS2-FKPBx2 and (2) a heterobifunctional small molecule CEM114 (FK506 linked to a p300 bromodomain binder) to recruit endogenous p300. Guide RNAs were transduced overnight prior to activation with 50 nM CEM114 for various durations. Constructs for guide RNAs included tandem insertion of each set (4 sgRNAs for sgBoundary around the *MYOD1* super enhancer boundaries, 6 sgRNAs for sgEpiCenters targeting *MYOD1* super enhancer centers) into a lentiviral backbone (plasmids available upon request).

Super resolution imaging of RNA-Pol2-GFP and BRD4-mEGFP

RH4 cells were stably transduced by lentivirus with RBP3 (RNA-Pol2 subunit) tagged with GFP (construct a gift of Dr. Dan Larson, NIH) and purified by FACS sorting. Endogenous CRISPR-cas9-mediated knock-in of mEGFP to *BRD4* in RH4 cells was down with homology directed repair (constructs a gift from Dr. Richard Young, MIT), followed by sorting for positive cells, as described previously[39].

Super-resolution structured-illumination microscopy was performed with a home-built instant structured illumination microscope (iSIM)[54]. The iSIM was equipped with a 60× Oil-immersion objective lens (Olympus, N.A. = 1.42), a stage top incubator (Okolab), xyz automated stage (ASI, USA). A 488 nm laser (Genesis MX-Series, Coherent) was used as excitation source and the emission was collected by an sCMOS camera (pco.edge 4.2, PCO) after passing through 488 notch filter (NFD01–488). The microscope stage, lasers, AOTF and the camera were controlled through custom written Python scripts. Imaging volumes were acquired every 10 minutes, using a 0.25 μm z spacing. Imaging was performed at 37 °C.

Widefield imaging was conducted on an Olympus IX 81 microscope equipped with a 60× Silicone Oil-immersion objective lens (Olympus, N.A. 1.3), a live cell environmental control chamber (INUBG2A-PI, Tokai Hit), xy automated stage (ASI, USA) and piezo z-stage (Mad City Labs, USA). An LED light source (Xcite 110LED, Excelitas Technologies) with single band filter (FF02–472/30, Semrock, USA) provided excitation and the emission was collected by the same objective passing through a bandpass filter (520/35, Semrock, USA) in front of the electron-multiplying charge-coupled device (EMCCD) (iXon DU-888, Andor Technologies). The microscope stage, LED and the camera were controlled through μ-manager (2.0 Beta)[55]. Imaging volumes with 0.5 μm z spacing were collected every 10 minutes for time-lapse imaging, at multiple xy stage positions. Imaging was performed at 37° C.

Western Blot

Whole-cell lysates were prepared with RIPA (supplemented with protease and phosphatase inhibitors). Cells were lysed by sonication, incubated for 30 minutes at 4° C, and then lysates were centrifuged for 15 minutes. Protein concentration of the resulting supernatant was estimated by BCA assay. 40 μg of sample were separated on 4–12% Bis-Tris gels and transferred to PVDF membrane, blocked 1 hour in 5% non-fat milk in Tris buffered saline and Tween-20 (TBS-T). Membranes were incubated at 4° C overnight with 1:1,000 diluted SOX8 (Abcam, Cat # ab104245) primary antibody, washed three times in TBS-T, then incubated in 1:10,000 diluted HRP-labeled secondary antibody (Santa Cruz Biotechnology, Cat #sc-2004) at room temperature for 1 hour, washed an additional three times with TBS-T, and then developed with chemiluminescent reagents (Thermo Fischer Scientific, SuperSignal West Femto).

Statistics and Reproducibility

Statistical tests were either performed in GraphPad Prism (version 7 from GraphPad Software), R (version 3.5.1 from the Foundation for Statistical Computing), or GSEA (version 4.0 from the Broad Institute). For distribution violin-boxplots and statistical calculations of significance in Figure 3c, the size (n) of the gene sets were: RH4 TE (typical enhancer) Genes (n = 7,446), RH4 SE (super enhancer) Genes (n = 1,591), House Keeping Genes (n = 389), All TF Genes (n = 1,427), All RH4 CR (core regulatory) TFs (n = 50), Top RH4 CR TFs (n = 13). All gene sets are available in Supplementary Table 2. Gene set enrichment P values, NES values, and False Discovery Rate (FDR) values reported throughout are calculated with 1,000 permutations in the GSEA software, run in pre-ranked mode. CRISPR for human TFs in RMS cells was performed and sequenced as a single experiment per cell line, and was performed across 4 RMS cell lines (data available in Supplementary Table 4). ChRO-seq experiments were performed at 4 different time points to increase confidence in the data trends, and were performed once each (Fig. 3e). ChIP-seq experiments were often performed across multiple FP-RMS cell lines or primary tumors rather than performing "biological" replicates in the same cell line; in some instances we increased confidence in the accuracy of results by performing an orthogonal experimental condition or related ChIP target (performing ChIP-seq for HDAC1, HDAC2 and HDAC3, rather than HDAC1 in triplicate). Related to Figure 3, for HDAC1 and HDAC3 we performed ChIP-seq once each, and even though these were from different RH4 cell passages these single experiments showed high-concordance with HDAC2 ChIP-seq that was repeated twice in RH4 and once in each of 3 other FP-RMS samples (RH5, SCMC and a PDX tumor, data not shown) all giving very similar ChIP-seq profiles. ChIP-seq of CR TFs MYOG, MYOD, and SOX8 were independently repeated twice with very similar results; MYCN, YY1 and MED1 were performed once and were bound to expected regions; BRD4 results are representative of 6 independent ChIP-seq experiment in RH4 cells; PAX3-FOXO1 ChIP-seq data were representative of 4 experiments across different RH4 cell passages with similar results. For histone acetylation after DMSO or Entinostat treatment (related to Figure 4), we saw similar profiles and changes (spreading) among diverse histone lysine modifications (H3K27ac, H3K16ac, H2BK5ac), and these experiments were performed once each from different RH4 cell preparations; orthogonally the H3K27ac HiChIP experiments were qualitatively similar (when flattened) to H3K27ac ChIP-seq

results (see Supplementary Fig. 5c). AQuA-HiChIP data are representative of two separate biotin-captures, library preparations and sequencing runs that gave similar results. H3K27ac ChIP-seq profiles in FP-RMS were similar in 38 different independent experiments across 14 different cell lines, tumors and PDX models (15 of these ChIP-seq experiments of H3K27ac were performed from RH4 cells, as independently repeated control experiments for various chemical or genetic perturbations, and all giving very similar profiles).

Reporting Summary

Additional information regarding sample statistics, antibody information including lot numbers and dilutions used, ChIP-seq peak calling outputs, software used herein, read depth for experiments, and experimental designs, is available in the Life Sciences Reporting Summary linked to this article.

Data and Software Availability

The data reported herein are made publicly available through the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/). The GEO accession number for all ChIP-seq, ChIP-Rx, AQuA-HiChIP and RNA-seq data is GSE116344. Code is available at https://github.com/GryderArt. A list of related resources and reagents used in our study can be found in Supplementary Table 6, which includes a list of all software and data used herein and their sources.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Lambert SA et al. The human transcription factors. Cell 172, 650–665 (2018). [PubMed: 29425488]

2. Orphanides G, Lagrange T & Reinberg D The general transcription factors of RNA polymerase II. Genes & development 10, 2657–2683 (1996). [PubMed: 8946909]

3. Lee Tong I. & Young Richard A. Transcriptional Regulation and Its Misregulation in Disease. Cell 152, 1237–1251 (2013). [PubMed: 23498934]

4. Boyer LA et al. Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell 122, 947–956 (2005). [PubMed: 16153702]

5. Whyte WA et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319 (2013). [PubMed: 23582322]

6. Saint-André V et al. Models of human core transcriptional regulatory circuitries. Genome Research (2016).

7. Durbin AD et al. Selective gene dependencies in MYCN-amplified neuroblastoma include the core transcriptional regulatory circuitry. Nature Genetics 50, 1240–1246 (2018). [PubMed: 30127528]

8. Lin CY et al. Active medulloblastoma enhancers reveal subgroup-specific cellular origins. Nature 530, 57–62 (2016). [PubMed: 26814967]

9. Hnisz D et al. Super-Enhancers in the Control of Cell Identity and Disease. Cell 155, 934–947 (2013). [PubMed: 24119843]

10. Gryder BE et al. Chemical genomics reveals histone deacetylases are required for core regulatory transcription. Nature Communications 10, 3004 (2019).

11. Allfrey V, Faulkner R & Mirsky A Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proceedings of the National Academy of Sciences 51, 786–794 (1964).

12. Pogo B, Allfrey V & Mirsky A RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. Proceedings of the National Academy of Sciences 55, 805–812 (1966).

13. Kanno T et al. BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. Nat Struct Mol Biol 21, 1047–57 (2014). [PubMed: 25383670]

14. Moriniere J et al. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. Nature 461, 664–U112 (2009). [PubMed: 19794495]

15. Gryder BE et al. PAX3-FOXO1 Establishes Myogenic Super Enhancers and Confers BET Bromodomain Vulnerability. Cancer Discovery (2017).

16. Stewart E et al. Identification of Therapeutic Targets in Rhabdomyosarcoma through Integrated Genomic, Epigenomic, and Proteomic Analyses. Cancer Cell 34, 411–426.e19 (2018). [PubMed: 30146332]

17. Meyers RM et al. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. Nature genetics 49, 1779 (2017). [PubMed: 29083409]

18. Tenente IM et al. Myogenic regulatory transcription factors regulate growth in rhabdomyosarcoma. eLife 6, e19214 (2017). [PubMed: 28080960]

19. Schmidt K, Glaser G, Wernig A, Wegner M & Rosorius O Sox8 Is a Specific Marker for Muscle Satellite Cells and Inhibits Myogenesis. Journal of Biological Chemistry 278, 29769–29775 (2003). [PubMed: 12782625]

20. Böhm M et al. Helicase CHD4 is an epigenetic coregulator of PAX3-FOXO1 in alveolar rhabdomyosarcoma. The Journal of Clinical Investigation 126, 4237–4249 (2016). [PubMed: 27760049]

21. Chu T et al. Chromatin run-on and sequencing maps the transcriptional regulatory landscape of glioblastoma multiforme. Nature Genetics (2018).

22. Blumberg A et al. Characterizing RNA stability genome-wide through combined analysis of PRO-seq and RNA-seq data. BioRxiv, 690644 (2019).

23. Sharova LV et al. Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. DNA research 16, 45–58 (2008). [PubMed: 19001483]

24. Fukaya T, Lim B & Levine M Enhancer Control of Transcriptional Bursting. Cell 166, 358–368 (2016). [PubMed: 27293191]

25. Greer Celeste B. et al. Histone Deacetylases Positively Regulate Transcription through the Elongation Machinery. Cell Reports 13, 1444–1455 (2015). [PubMed: 26549458]

26. Sanchez GJ et al. Genome-wide dose-dependent inhibition of histone deacetylases studies reveal their roles in enhancer remodeling and suppression of oncogenic super-enhancers. Nucleic acids research 46, 1756–1776 (2017).

27. Orlando David A. et al. Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. Cell Reports 9, 1163–1170 (2014). [PubMed: 25437568]

28. Rao Suhas S.P. et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665–1680 (2014). [PubMed: 25497547]

29. Dowen Jill M. et al. Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes. Cell 159, 374–387 (2014). [PubMed: 25303531]

30. Mumbach MR et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Meth 13, 919–922 (2016).

31. Mumbach MR et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. Nature Genetics 49, 1602 (2017). [PubMed: 28945252]

32. Chory EJ et al. Nucleosome Turnover Regulates Histone Methylation Patterns over the Genome. Molecular Cell 73, 61–72.e3 (2019). [PubMed: 30472189]

33. Stanton BZ, Chory EJ & Crabtree GR Chemically induced proximity in biology and medicine. Science 359, eaao5902 (2018). [PubMed: 29590011]

34. Chiarella AM et al. Dose-dependent activation of gene expression is achieved using CRISPR and small molecules that recruit endogenous chromatin machinery. Nature Biotechnology (2019).

35. Cho W-K et al. RNA Polymerase II cluster dynamics predict mRNA output in living cells. eLife 5, e13617 (2016). [PubMed: 27138339]

36. Cho W-K et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science (2018).

37. Weider M & Wegner M SoxE factors: Transcriptional regulators of neural differentiation and nervous system development. Seminars in Cell & Developmental Biology 63, 35–42 (2017). [PubMed: 27552919]

38. Banani SF, Lee HO, Hyman AA & Rosen MK Biomolecular condensates: organizers of cellular biochemistry. Nature Reviews Molecular Cell Biology 18, 285 (2017). [PubMed: 28225081]

39. Sabari BR et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science (2018).

40. Boulay G et al. Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. Cell 171, 163–178.e19 (2017). [PubMed: 28844694]

Methods Only References

41. Chang W et al. Multi-Dimensional ClinOmics for Precision Therapy of Children and Adolescent Young Adults with Relapsed and Refractory Cancer: A report from the Center for Cancer Research. Clinical Cancer Research (2016).

42. Barski A et al. High-Resolution Profiling of Histone Methylations in the Human Genome. Cell 129, 823–837 (2007). [PubMed: 17512414]

43. Yohe ME et al. MEK inhibition induces MYOG and remodels super-enhancers in RAS-driven rhabdomyosarcoma. Science Translational Medicine (2018).

44. Buenrostro JD, Giresi PG, Zaba LC, Chang HY & Greenleaf WJ Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Meth 10, 1213–1218 (2013).

45. Hodges HC et al. Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. Nature Structural & Molecular Biology (2017).

46. Subramanian A et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences 102, 15545–15550 (2005).

47. Tarumoto Y et al. LKB1, Salt-Inducible Kinases, and MEF2C Are Linked Dependencies in Acute Myeloid Leukemia. Molecular Cell 69, 1017–1027.e6 (2018). [PubMed: 29526696]

48. Shi J et al. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nature Biotechnology 33, 661 (2015).

49. Macosko Evan Z. et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214 (2015). [PubMed: 26000488]

50. Kidder BL & Zhao K Efficient library preparation for next-generation sequencing analysis of genome-wide epigenetic and transcriptional landscapes in embryonic stem cells in Stem Cell Transcriptional Networks 3–20 (Springer, 2014).

51. Servant N et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biology 16, 259 (2015). [PubMed: 26619908]

52. Durand NC et al. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Systems 3, 95–98 (2016). [PubMed: 27467249]

53. Gryder B, Khan J & Stanton B Absolute Quantification of Architecture (AQuA-HiChIP) Enables Measurement of Differential Chromatin Interactions. Nature Protocols (2019).

54. York AG et al. Instant super-resolution imaging in live cells and embryos via analog image processing. Nature Methods 10, 1122 (2013). [PubMed: 24097271]

55. Edelstein AD et al. Advanced methods of microscope control using μManager software. Journal of biological methods 1(2014).
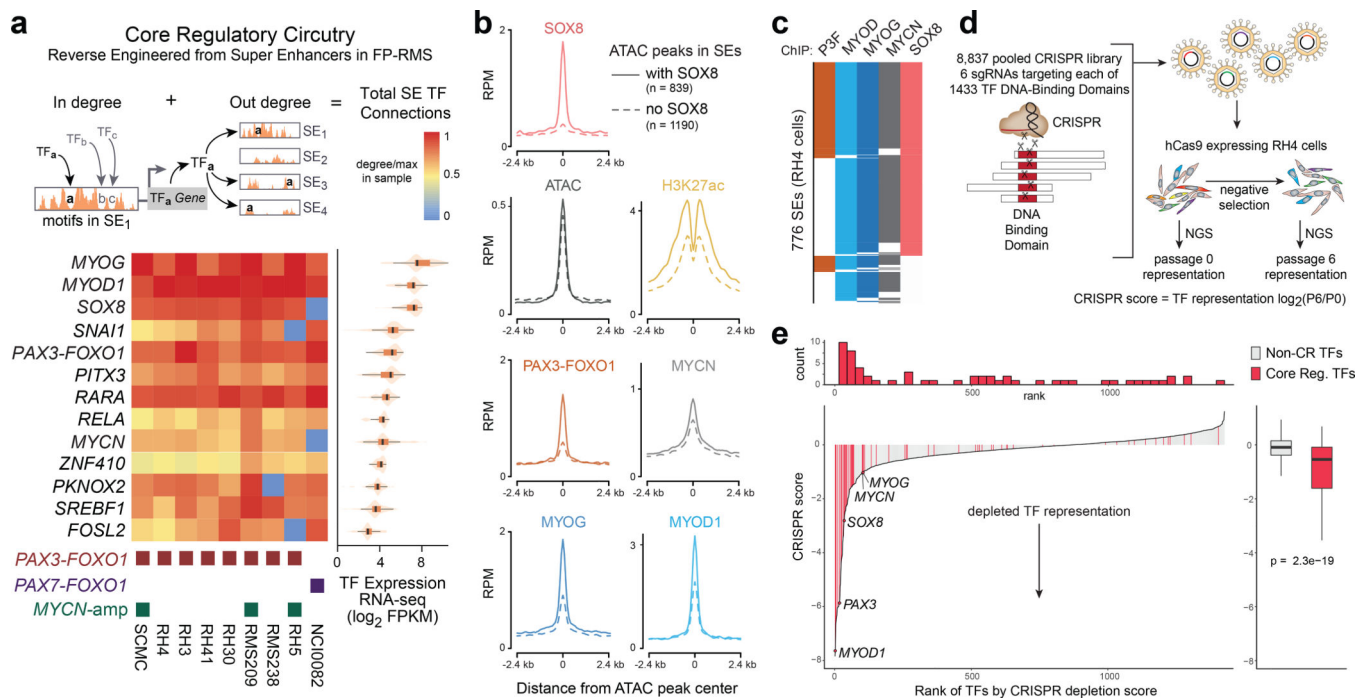
Figure 1. Core regulatory circuitry includes SOX8 and is critical for FP-RMS

a, Core regulatory circuitry identified by analysis of motif-networks in super-enhancer (SE) associated transcription factors (TFs). Heatmap shows predicted CR TFs found in FP-RMS cell lines and tumors (n = 9 independent samples), clustered and colored by degree of connectivity (scaled to 1 = maximum connectivity in the sample). Expression of CR TFs is plotted on the right (box plots show quartiles, with whiskers showing the $1.5 \times$ inter-quartile range, with data distribution as violin plots).

b, Core regulatory TF validation ChIP-seq in RH4 ChIP-seq (a FP-RMS cell line). Metagene plots at ATAC-seq peaks in SEs were divided into SOX8-bound (solid lines, n = 839 peaks) and SOX-unbound (dotted lines, n = 1,190 peaks).

c, Prevalence and co-occurrence of CR TFs in SEs in RH4.

d, Functional genetic screening to evaluate TF dependencies in FP-RMS cells, using pooled CRISPR-cas9 methodology. Each sgRNA (n = 8,837) is designed to recruit Cas9 to DNA sequence coding for the DNA-binding domain(s) of each human TF (n = 1433).

e, Depletion of sgRNAs targeting CR TFs, compared to all non-CR TFs, over 6 passages of negative selection. Top panel shows a histogram of the number of CR TFs (red bar). Lower panel shows a CRISPR score for each TF as an average $\log_2$ fold change of sgRNAs (n = 6 distinct sequences per TF) targeting DNA-binding domains. Right panel shows summary statistics in box (quartiles) and whisker ($1.5 \times$ inter-quartile range) plot, with *P* value comparing extent of depletion between CR TFs and all other TFs calculated with an unpaired, two-sided student's *t* test with Welch's correction.
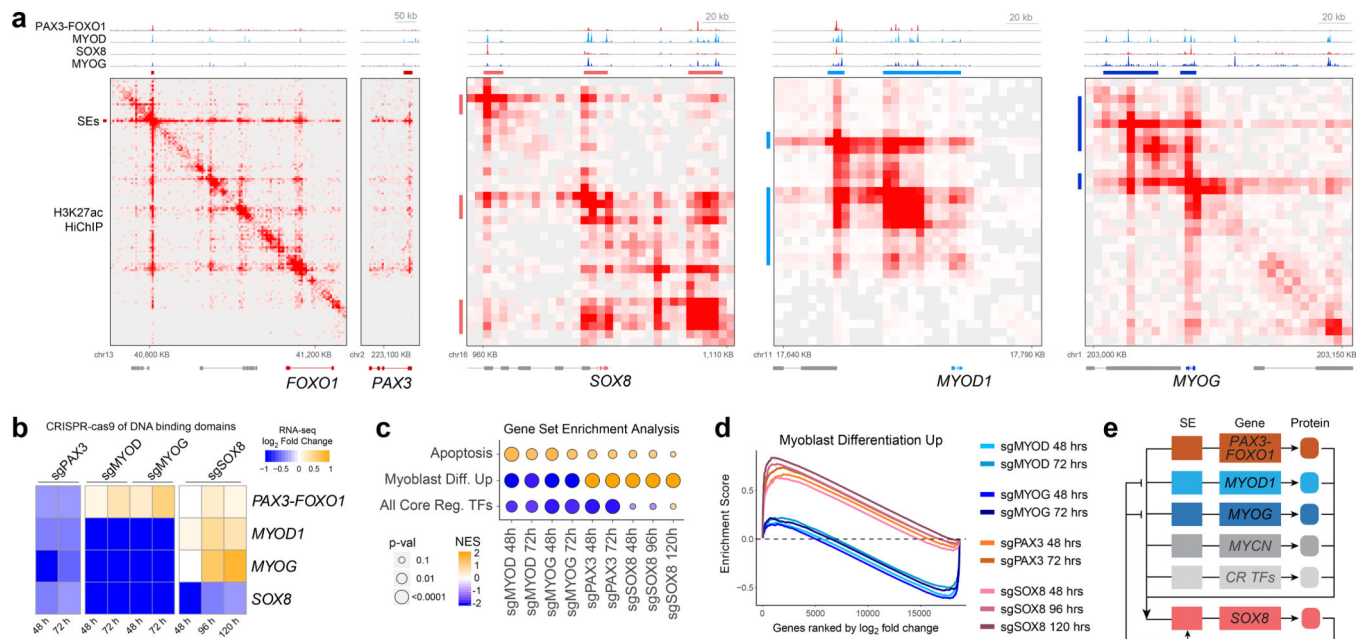
Figure 2. Genetic dissection of core regulatory network reveals a SOX8-mediated myogenic blockade in RMS

(a) Connectivity of CR TFs shown by ChIP-seq of PAX3-FOXO1, MYOD, MYOG and SOX8 at super-enhancers (SEs) that interact with the promoters of *PAX3*, *MYOD1*, *MYOG* and *SOX8*. Heatmaps depict interaction frequency between genomic elements, assayed by H3K27ac HiChIP (one representative of two experiments in RH4 cells).

(b) Co-regulation of CR TFs evaualted in RH4 cells with CRISPR-cas9 targeted to the DNA binding domains of PAX3, MYOD, MYOG or SOX8. Experiments were repeated orthogonally with shRNA with similar results (Supplementary Fig. 2e).

(c) Gene set enrichment analysis shows divergent transcriptional impact of individual CR TF disruption in RH4 cells (gene sets available in Supplemental Table 2). Similar results were obtained using shRNA (Supplementary Fig. 2f). *P* value is generated using the GSEA algorithm of the enrichment score relative to the null distribution calculated with 1,000 permutations. Each bubble represents one gene set analyzed against an individual RNA-seq experiment (sgRNA target v. non-targeting sgRNA) at the indicated time points.

(d) Activation and enrichment of myoblast differentiation genes upon sgRNA mediated disruption of PAX3-FOXO1 or SOX8, in constrast to depletion of the myogenic program upon MYOD or MYOG knockout, and was also seen upon shRNA knockdown (Supplementary Fig. 2g).

(e) Model of auto-regulatory feed-forward (PAX3-FOXO1, MYOD1, MYOG, MYCN and other CR TFs) and negative feedback (from SOX8) circuitry in FP-RMS.
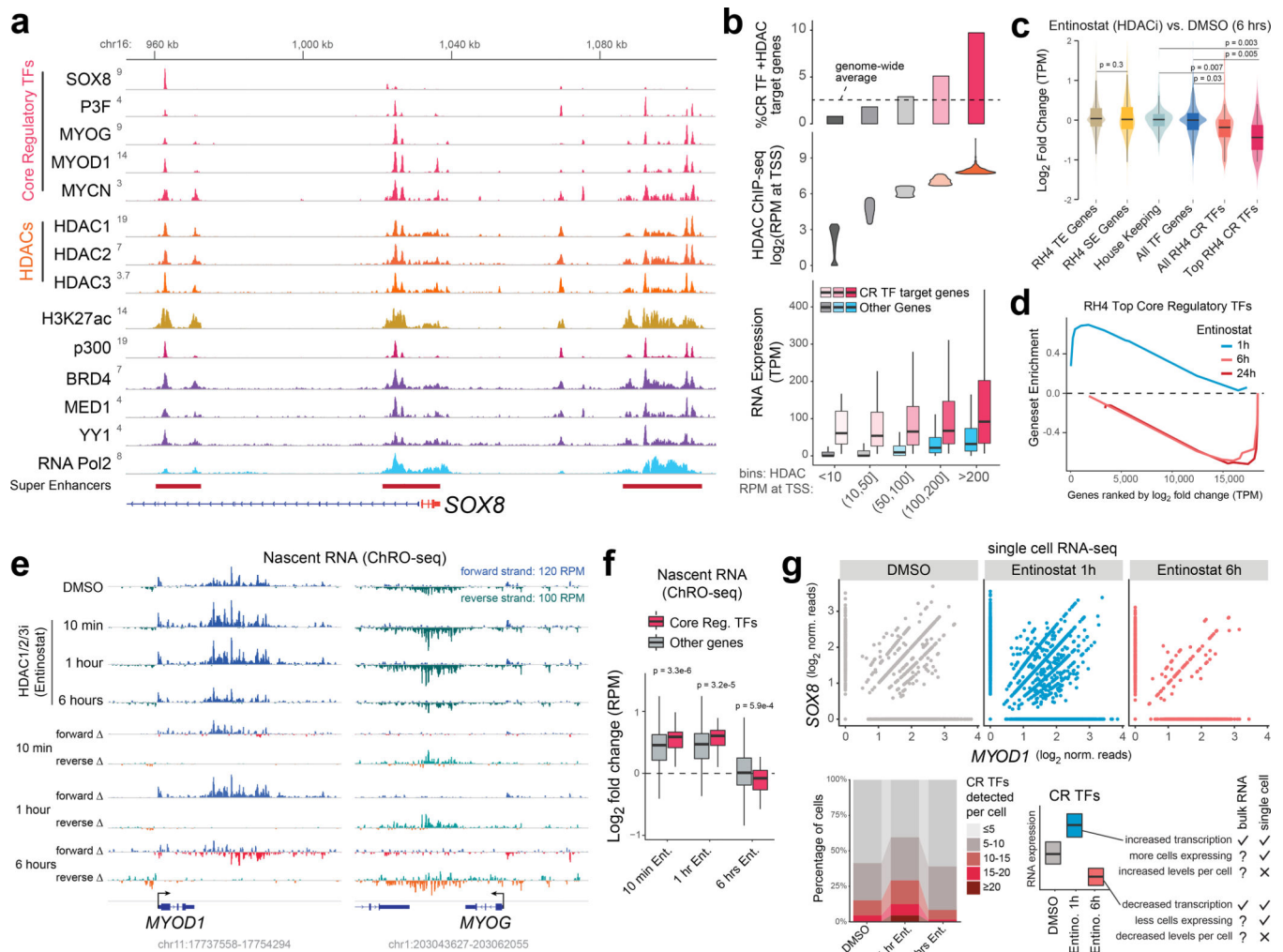
Figure 3. Core regulatory TFs require HDAC1/2/3 for their transcription

a, Genome browser view at the *SOX8* locus, including 3 super-enhancers (red bars) shown by ChIP-seq to be heavily decorated by CR TFs, HDACs (nuclear isoforms 1, 2 and 3), histone acetylation, other co-activators and RNA Pol2. Each track is one representative of two independent ChIP-seq experiments, except for HDAC3, MYCN, YY1 and MED1 which were performed once, and PAX3-FOXO1 and H3K27ac ChIP-seq data which were performed more than 4 times across different RH4 cell passages.

b, Enrichment of CR TFs in SEs associated with genes that have high amounts of HDAC bound to their promoters (top, dotted line indicate genome-wide average), genes divided into 5 categories of HDAC ChIP-seq density (middle, violin plots), which correlated positively with expression (bottom, boxplots with median and quartiles, whiskers showing the 1.5 × inter-quartile range). Bins were defined using a single representative HDAC2 ChIP-seq, and similar binding was seen with HDAC1 and HDAC3 (see Supplementary Fig. 3a–b).

c, Changes in gene expression upon HDAC1/2/3 inhibition with Entinostat for 6 hours in RH4 cells. Violin plots are overlaid with box plots of median and quartiles, whiskers showing the 1.5 × inter-quartile range. *P* values were calculated with a two-tailed unpaired *t* test with Welch's correction.

d, Time-course RNA-seq followed by geneset enrichment reveals that Entinostat causes an initial increase in core regulatory transcription (1 hr), followed by a selective downregulation at 6 and 24 hours in RH4 cells. FDR was calculated by GSEA with 1,000 permutations for each timepoint, and were for 1 hour, q = 0.006; 6 hours, q = 0; 24 hours, q = 0.

e, Chromatin run-on and sequencing (ChRO-seq) to measure nascent RNA shows transcriptional increase at CR TFs *MYOD1* and *MYOG* in 10 minutes, sustained for the first hour, and decreased at 6 hours of Entinostat treatment (1 μM in RH4 cells).

f, Nascent RNA changes over time course treatment with Entinostat, as meansured by ChRO-seq and summarized by box plots (median with quartiles, whiskers representing 1.5 × interquartile range in the data). A two-sided, unpaired *t* test with Welch's correction was used to calculate *P* values.

g, Single-cell RNA-seq analysis of cells treated with DMSO (n = 2,925 cells) or Entinostat for 1 hr (n = 3,805 cells) or 6 hrs (n = 3,240 cells). Top panel, shows increased number of cells expressing the CR TFs *SOX8* and *MYOD1* at 1 hour of Entinostat and a decreased number at 6 hours. Bottom left panel, single-cell RNA-seq data of CR TFs abundance per cell in RH4 cells. These experiments were performed once. Bottom right shows heuristic boxplots and describes some relevant advantages of single cell RNA-seq over bulk RNA-seq.
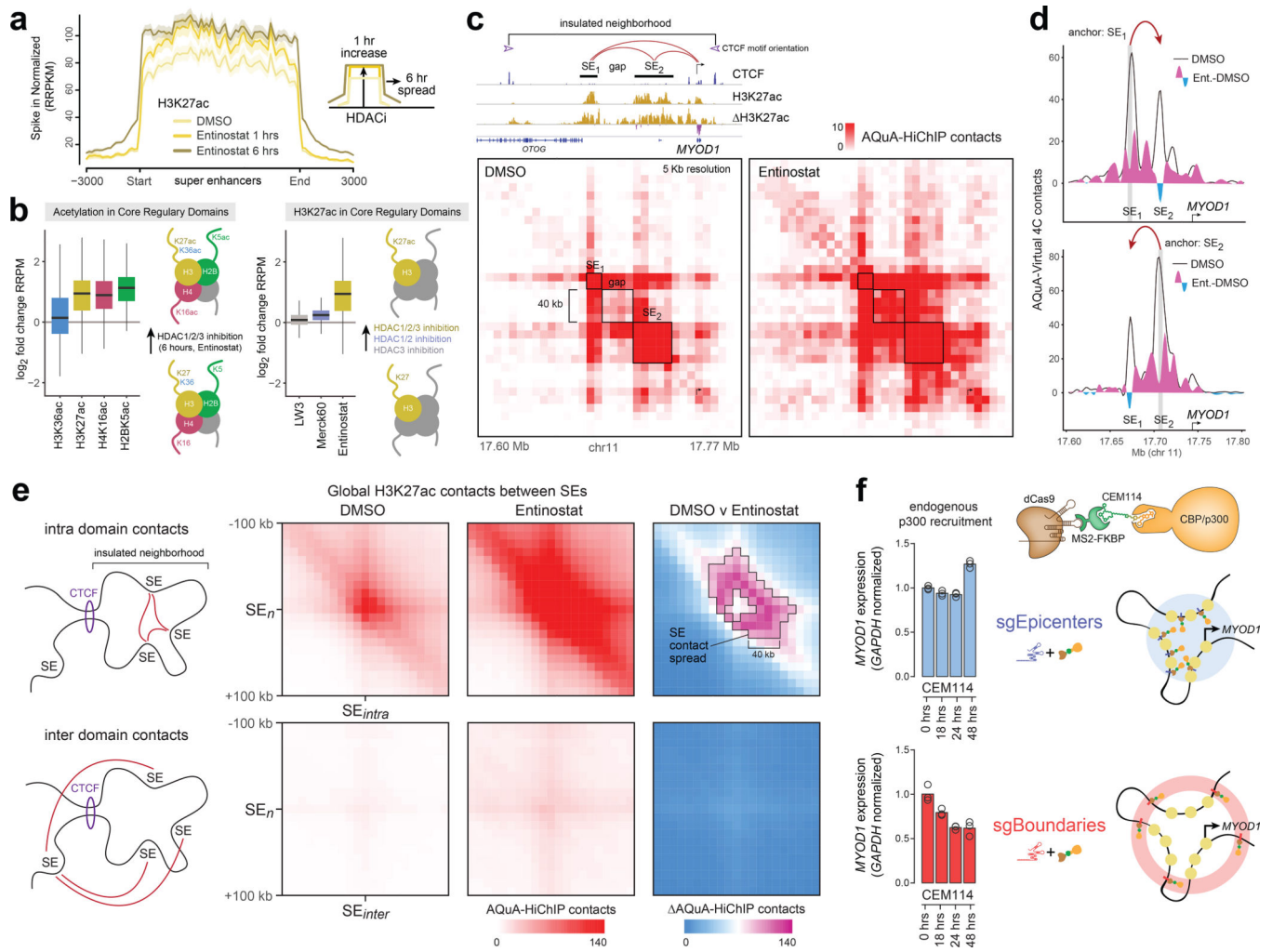
Figure 4. AQuA-HiChIP shows disruption of super-enhancer architecture by hyperacetylation

a, Super-enhancer dynamics upon HDAC inhibition, revealing an acute increase in H3K27ac after 1 hour, followed by a spread beyond the endogenous SE boundary at 6 hours. ChIP-Rx with exogenous spike in is reported as Reference-normalized Reads Per Kilobase per Million mapped reads (RRPKM). Shading shows the SEM.

b, Acetylation changes quantified on diverse lysine residues on histone subunits H3 (K36, K27), H2B (K5) and H4 (K16) by ChIP-Rx. Reference-normalized Reads Per Million mapped reads (RRPM) differences for diverse acetylation sites on histones are shown for Core Regulatory Domains after treatment with either DMSO or Entinostat for 6 hours (left), or for H3K27ac upon treatment with HDAC1/2 inhibitor Merck60, HDAC3 inhibitor LW3 or HDAC1/2/3 inhibitor Entinostat (right). All experiments were in RH4 cells treated for 6 hours with 1 μM of the indicated inhibitors. Box plots show the median and quartiles, with whiskers showing the 1.5 × inter-quartile range.

c, Absolute Quantification of Architecture HiChIP (AQuA-HiChIP) identifies spreading of SE-mediated contacts within the insulated neighborhood of CR TF *MYOD1*. Contact map is shown at 5-kb resolution (5-kb by 5-kb contact squares), and is scaled to AQuA normalized contacts per million.

d, Gained aberrant contacts and lost SE-to-SE contacts are visualized by AQuA-Virtual 4C at *MYOD1* from viewpoint anchor SE_1 (top) and SE_2 (bottom). Virtual 4C is representative of two replicate biotin captures and library preparations, both with similar results, and agrees with non-virtual 4C experiments at *MYOD1* SEs under the same treatment conditions.

e, SE contact spreading as seen by AQuA-HiChIP Aggregate Peak Analysis (APA) plots of all SE-to-SE contacts within insulated neighborhoods (SE_n to SE_{intra}, top) or SEs nearby but outside insulated neighborhood CTCF boundaries (SE_n to SE_{inter}, bottom). Resolution is shown at 10-kb by 10-kb squares.

f, Endogenous p300 recruitment to *MYOD1* SE elements. Chemical Epigenetic Modifier-114 (CEM-114, bi-functional FKBP-binder and p300 bromodomain binder) enables dCas9-guided recruitment of p300 to MYOD1 SE epicenters (sgEpicenters) or SE boundaries (sgBoundaries) in RH4 cells. Triplicate values are one representative of two independent cell treatments each with two RT-qPCR replicates that gave similar results.
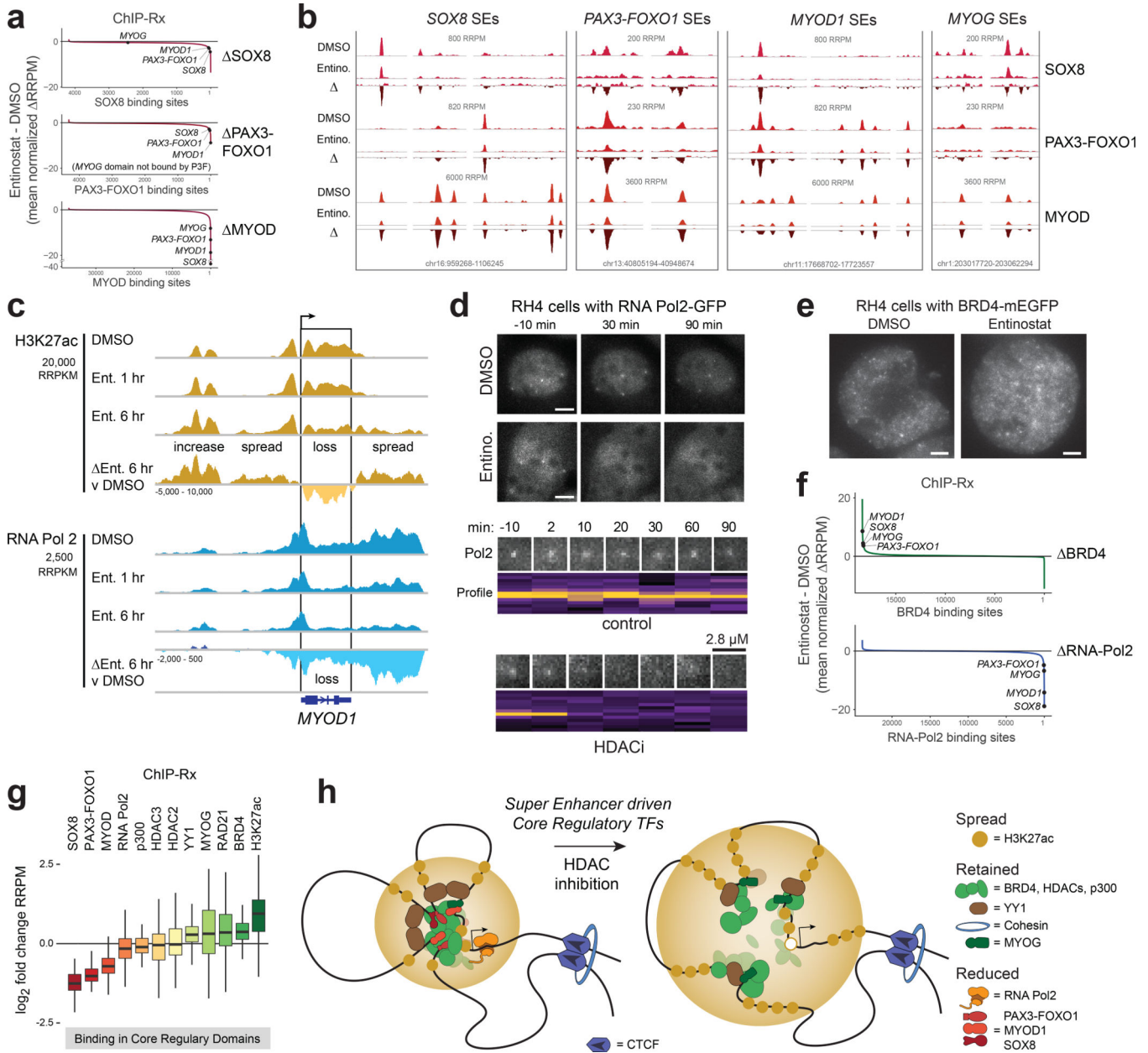
Figure 5. SE clusters and phase condensates are disrupted by hyperacetylation

a, ChIP-Rx binding sites for CR TFs, ranked by change in binding upon 6 hours of Entinostat treatment in RH4 cells.

b, Entinostat-induced changes in binding of CR TFs shown at *cis*-regulatory elements for *SOX8, PAX3-FOXO1*, *MYOD1*, and *MYOG* in RH4 cells.

c, RNA Pol2 unloading along all genic positions of *MYOD1* at 1 and 6 hours of Entinostat treatment, and associated changes in H3K27ac, as measured by ChIP with spike-in normalization (ChIP-Rx).

d, Clusters of RNA Pol2 tagged with GFP (in live FP-RMS cells, RH4), imaged in a time course with or without addition of HDAC inhibitor Entinostat. Scale bar equals 5 μM. Representative single puncta of RNA Pol2 clusters over time are shown (2.8 μM squares)

from control or HDAC inhibitor treated RH4 cells. Cross-section profiles of pixel brightness for corresponding time points are shown below image frames. Images are from wide-field (see Online Methods) and are representative of 50 images across 2 independent experiments with similar results. Alternative imaging modality iSIM gave similar results to widefield and is shown in Supplementary Figure 7f.

e, Live cell imaging of BRD4 (endogenously tagged with mEGFP in RH4 cells) treated for 6 hours with DMSO (left) or Entinostat (right). Scale bar represents 5 μM. Images are representative of 20 images across 2 independent experiments with similar results.

f, Quantification of RNA Pol2 (top) and BRD4 (bottom) binding changes upon HDAC1/2/3 inhibition with Entinostat (1 μM, RH4 cells) by ChIP-Rx signal.

g, HDAC inhibition induced changes in binding of key factors within Core Regulatory Domains, measured by ChIP-Rx in RH4 cells, summarized by box plots representing the median and quartiles, whiskers showing $1.5 \times$ interquartile range.

h, A model whereby histone hyperacetylation disrupts normal looping interactions, reduces binding of certain CR TFs, reduces binding of RNA Pol2 and causes loss of transcription at super-enhancer-dependent CR TF genes.