An enhancer RNA recruits MLL1 to regulate transcription of *Myb*

Juhyun Kim[1,6*], Luis F. Diaz[1,2,6], Matthew J. Miller[1,3,6], Benjamin Leadem[1,4], Ivan Krivega[1,5] and Ann Dean[1*]

[1]Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

[2]Oregon Health and Sciences University, Portland, OR 97239

[3]University of Iowa Medical School, Iowa City, IA 52242

[4]GeneDx, Gaithersburg, MD 20877

[5]Sonothera, South San Francisco, CA 94080

[6] Joint authors

[*]Address for correspondence: ann.dean@nih.gov; Tel: 301-496-6068.

juhyun.kim@nih.gov; Tel: 301-402-5704

ABSTRACT

The *Myb* proto-oncogene encodes the transcription factor c-MYB, which is critical for the proliferation and differentiation of hematopoietic stem and progenitor cells. Distant enhancers of *Myb* expression form a hub of interactions with the *Myb* promoter but the regulation of *Myb* during hematopoiesis is still incompletely understood. Here we identified a novel nuclear *Myb* long non-coding enhancer RNA (*Myrlin*) that originates from the -81 kb murine *Myb* enhancer within the *Myb* ─ *Hbs1l* intergenic region. *Myrlin* and *Myb* are coordinately regulated in a developmental stage-specific fashion during maturation of erythroid progenitors and upon differentiation of MEL cells. CRISPR/Cas9 genome editing of the *Myrlin* transcription start site at the -81kb enhancer reduced both *Myrlin* and *Myb* expression. Deletion of the *Myrlin* TSS reduced occupancy by the looping protein LDB1, compromising long-range hub contacts between the *Myb* promoter and its enhancer and reducing RNA Pol II occupancy across the *Myb* locus. In contrast, silencing of *Myrlin* using CRISPRi, while similarly reducing both *Myrlin* and *Myb* expression, left the *Myb* enhancer hub undisturbed, revealing that chromatin looping and transcription activation of *Myb* can be decoupled. Investigating a role for *Myrlin*, *per se*, we found that *Myrlin* interacts with MLL1 complex, a transcriptional coactivator that plays an essential role in regulating gene expression during hematopoiesis. *Myrlin* CRISPRi compromised MLL1 occupancy in the *Myb* locus and decreased CDK9 and RNA Pol II binding. *Myrlin* CRISPRi further resulted in pausing of RNA Pol II in the *Myb* first exon/intron. These data document that Myrlin directly participates in activating *Myb* transcription by recruiting MLL1.

INTRODUCTION

While the vast majority of the mammalian genome is transcribed, only a small fraction of these transcripts encodes proteins. The functional relevance of most non-coding transcription remains largely unknown. RNA polymerase II (Pol II) non-coding transcripts that are greater than 200 nt in length and lack coding potential are known as long non-coding RNAs (lncRNAs). LncRNAs, much like their protein coding counterparts, can be spliced and polyadenylated, but they are biased towards two-exon transcripts that remain localized to the nucleus (Cabili et al., 2011; Derrien et al., 2012; Djebali et al., 2012). Nuclear-localized lncRNAs can be involved in gene regulation via interactions with chromatin remodelers, histone modifying complexes and transcription factors (Gil and Ulitsky, 2020; Herman et al., 2022; Hu et al., 2012; Kim and Shiekhattar, 2016; Statello et al., 2021). LncRNAs derived from enhancer regulatory regions, known as enhancer RNAs or eRNAs, have the potential ability to function together with the enhancers from which they are derived to regulate expression of target genes (Arnold et al., 2019; Hou and Kraus, 2021; Sartorelli and Lauberth, 2020).

Enhancers increase the transcriptional output of target genes in a cell-type specific fashion, often bridging substantial genomic distances (Furlong and Levine, 2018; Kim and Dean, 2021; Kim and Wysocka, 2023; Levine et al., 2014; Panigrahi and O'Malley, 2021; Plank and Dean, 2014). Both active enhancers and genes are marked by histone H3 lysine 4 (H3lys4) methylation. This modification is deposited by the MLL/COMPASS family of histone methyltransferases including SET1A/B and four mixed lineage leukemia complexes (MLL1-4) (Cenik and Shilatifard, 2021; Dou et al., 2006). MLL1 forms a large macromolecular complex with WDR4, Menin, RbBP5, ASH2L and DPY-30 and interacts with the basic transcription machinery, including RNA Pol II (Milne et al., 2005). MLL1 selectively binds non-methyl CpG DNA through its CXXC domain (10). Additionally, MLL1 adaptor subunit WDR5 promotes the recruitment of MLL1 to genomic targets at a subset of genes (11). In HOXA loci, MLL1 complex is recruited by the lncRNA HOTTIP through direct binding with WDR5, establishing H3K4me3 deposition and driving gene transcription (Wang et al., 2011).

The *Myb* proto-oncogene encodes c-MYB (hereafter MYB), a critical hematopoietic regulator of cell proliferation and differentiation (Wang et al., 2018). MYB is a repressor of human fetal hemoglobin production (Jiang et al., 2006). Given that elevation of fetal hemoglobin in adults moderates the symptoms of Sickle Cell Disease and β-Thalassemia, MYB is a potential target of therapeutic manipulation. *Myb* is regulated by microRNAs and by a series of enhancers distributed over more than 100 kb between *Myb* and the adjacent upstream *Hbs1l* gene in mouse and human (Sankaran et al., 2011; Stadhouders et al., 2014; Stadhouders et al., 2012; Wahlberg et al., 2009). In the mouse, five enhancers, -36, -61, -68, -81 and -109 kb, with respect to the *Myb* transcription start site (TSS), establish proximity with the *Myb* promoter in an active chromatin hub (Stadhouders *et al.*, 2012). Repression of *Myb*, which is required for terminal differentiation of erythroid cells, is accompanied by loss of these contacts. The enhancers are occupied by the LDB1 transcription factor complex, which mediates enhancer looping, and reduction of LDB1 using an shRNA compromises formation of the *Myb* enhancer hub (Song et al., 2007; Stadhouders et al., 2015). However, how chromatin looping and transcription activation at the *Myb* locus are regulated remains unknown.

Long non-coding RNAs have been linked to erythropoiesis and the regulation of numerous erythroid genes, including the adult β-globin and fetal γ-globin genes (Alvarez-Dominguez et al., 2014; Gurumurthy et al., 2021; Hu et al., 2011; Ivaldi et al., 2018; Morrison et al., 2018; Paralkar et al., 2014). We identified a novel lncRNA derived from the murine -81 kb *Myb* enhancer termed *Myrlin* for *My*b enhance<u>r</u> <u>l</u>ong <u>in</u>tergenic non-coding RNA. The *Myrlin* transcript was not required for formation of the *Myb* enhancer hub. However, *Myrlin* loss reduced MLL1/WDR5 recruitment in the *Myb* locus and decreased CDK9 and RNA Pol II occupancy. Furthermore, *Myrlin* loss resulted in pausing of RNA Pol II within the *Myb* first exon/intron. These results tie the *Myb* locus lncRNA *Myrlin* to the detailed mechanism of *Myb* transcription activation and suggest novel avenues that could become therapeutic targets for increasing HbF in β-globin hemoglobinopathies.

METHODS

Cell culture and animals

Control and CRISPR-edited mouse erythroid leukemia (MEL) cell lines were cultured in a 5% CO2 humidified incubator at 37 °C in DMEM with L-Glutamine, 10% fetal bovine serum, 1% Penicillin and Streptomycin and 1mM sodium pyruvate. MEL differentiation was induced with 2% DMSO for 4 days. Mouse embryonic and fetal RNA prepared from E8.5 yolk sac and E14.5 fetal liver was a generous gift from Dr. Xiang Guo.

CRISPR-Cas9 genome editing of MEL cells

CRISPR gRNAs were designed using the http://crispr.mit.edu/ website (see Supplemental Table S1 for gRNA sequences). gRNAs targeting the *Myrlin* transcription start site were cloned into the CRISPR-Cas9 and gRNA expression vector pSpCas9(BB)-2A-GFP (PX458) (gift from Feng Zhang, Addgene plasmid #48138) as described (Ran et al., 2013). MEL cells were transfected with Escort IV lipid transfection reagent (Sigma-Aldrich L3287) according to the manufacturer's instructions. Fluorescent cells were sorted 48 hours later and plated at limiting dilution to isolate clones. Clonal lines were genotyped by PCR using EmeraldAmp GT PCR Master Mix (Takara) and target specific primers flanking the *Myrlin* transcription start site. Deletions were validated by sequencing.

Stable MEL cell clones expressing dCAS9-KRAB were generated using Lenti-dCAS9-KRAB-blast (Gift from Dr. Gary Hon, Addgene plasmid #89567). MEL cells (2 x 106 cells) were suspended in 100ul of High-Performance Electroporation Solution (BTXpress) with 5-10ug of plasmid DNA and electroporated with the Gene Pulser Xcell System (Bio-Rad) using 2 pulses at 200V for 5ms. Cells were diluted with 100ul of pre-warmed media and transferred to 2 ml of media in a 12-well culture dish. Cells were selected in 10µg/ml Blasticydin for one week and plated at limiting dilution to isolate clones. Clonal lines were checked for production of S. pyogenes dCAS9 by RT-qPCR. dCAS9-KRAB MEL cells were electroporated as above with gRNAs targeting the *Myrlin* transcription start site cloned into LentiGuide-puro (gift from Feng Zhang, Addgene plasmid #52963). Cells were selected in 10µg/ml Blasticydin and 1µg/ml

puromycin for one week and plated at limiting dilution to isolate clones. Transfected cells were checked for expression of *Myrlin* by RT-qPCR.

5' and 3' rapid amplification of cDNA ends (RACE)

RACE was performed using the FirstChoice RLM-RACE kit (ThermoFisher Scientific) following the manufacturer's protocol. Total RNA from MEL cells was extracted and reverse transcribed using the 3' RACE adapter and the sequence of interest was amplified by nested PCR (3' RACE). Alternatively, a sample of the same RNA was treated with Calf Intestine Alkaline Phosphatase, then Tobacco Acid Pyrophosphatase and finally ligated to the 5' RACE adapter. De-capped adapter-ligated RNA was then reverse transcribed, and the sequence of interest was amplified by nested PCR (5' RACE). For RACE primers, see Supplemental Table S1.

Reverse-transcription qPCR (RT-qPCR)

RNA was isolated from 1x106 MEL cells with the RNeasy kit (Qiagen). One µg of RNA was treated with DNase I (ThermoFisher Scientific) for 15 min at 25°C. RNA was reverse-transcribed using the Superscript III First-Strand Synthesis System (ThermoFisher Scientific) following manufacturer's instructions. RT-qPCR was performed using the SYBR Green Supermix (Bio-Rad) with the ABI 7900HT (Applied Biosystems). Data was normalized to ActB. For RT-qPCR primers see Supplemental Table S1 and (Stadhouders *et al.*, 2012).

Chromatin immunoprecipitation (ChIP)

ChIP was performed as described (Krivega and Dean, 2017). MEL cells were cross-linked with 1% formaldehyde, nuclei were prepared, and chromatin was sheared by sonication to an average fragment size of 200–1000 bp. Chromatin was precleared, incubated with antibodies overnight at 4°C and subsequently immunoprecipitated using Protein A/G PLUS Agarose (Santa Cruz sc-2003). Immunoprecipitated samples were washed under stringent conditions, eluted, de-crosslinked and DNA was purified using the Genomic DNA Clean & Concentrator Kit according to the manufacturer's instructions (Zymo Research). Real-time qPCR was carried out with published primers (Stadhouders *et al.*, 2012) and see Supplemental Table S1), using SYBR

Green Supermix (Bio-Rad) on the ABI 7900HT (Applied Biosystems). Enrichment was calculated against input. For antibodies see Supplemental Table S2.

Chromatin conformation capture assay (3C)

3C was performed as described (Hagege et al., 2007) with HindIII restriction endonuclease. After reversal of crosslinks, ligation products were purified using the Genomic DNA Clean & Concentrator Kit as instructed by manufacturer (Zymo Research). Relative cross-linking between the *Myb* promoter and fragments of interest was analyzed by real-time qPCR with published TaqMan probes and primers (Stadhouders *et al.*, 2012). Ligation products of HindIII digested BAC DNA containing the mouse *Myb-Hbs1l* intergenic region were used to determine primer efficiency. Results were normalized to ligation frequency for a primer pair in ERCC3. Additional primers are listed in Supplemental Table S1.

***Myrlin* RNA cellular localization**

MEL cells were washed with ice cold PBS and lysed in hypotonic buffer as described (Ivaldi *et al.*, 2018). After centrifugation, RNA from the supernatant cytoplasmic material and the nuclear pellet were purified using the RNeasy Mini Kit (Qiagen 74104) as described by the manufacturer. RNA was reverse transcribed, and cDNA was measured by real-time qPCR. Expression was normalized to Tub1a. For RT-qPCR primers see Supplemental Table S1.

Chromatin Isolation by RNA purification (ChIRP)

The *Myrlin* probes were designed using the Stellaris Probe Designer version 4.2 and synthesized with 3' Bio-TEG modification by IDT. LacZ probes were procured from Millipore (CS216572). ChIRP-seq was performed as described (Chu C, Quinn J, Chang HY. Chromatin Isolation by RNA Purification (ChIRP). Jove-Journal of Visualized Experiments. 2012(61)). Briefly, MEL cells were fixed with 1% glutaraldehyde (Sigma-Aldrich, G5882) for 10mins at room temperature, then subjected to sonication with Bioruptor (Diagenode), and then aliquoted into input, odd, and even samples. 100pM probes were introduced and incubated for 4hrs at 37°C with 600rpm shaking. RNA was extracted with the miRNAeasy Mini Kit (Qiagen, 217004), and

SuperScript III First-Strand Synthesis System (Invitrogen, 18080051) was used to reverse transcribe RNA to cDNA. DNA products were then isolated using phenol: chloroform. qPCR was performed using SYBR green chemistry (SYBR Green Supermix) with the ABI 7900HT. For RT-qPCR primers see Supplemental Table S1.

RNA pulldown

Full-length *Myrlin* was transcribed conducted using the MEGAscript T7 Kit (Invitrogen, AM1333) according to the manufacturer's protocol. RNA pulldown was performed using the Pierce™ Magnetic RNA-Protein Pull-Down Kit (ThermoFisher, 2064) according to the manufacturer's protocol. Briefly, 1µg of biotinylated *Myrlin* RNA was incubated with 1mg of precleared protein extracted from MEL cells for 4 hrs at 4°C. Following this, Streptavidin magnetic beads was added and incubated for an additional 2hrs. Finally, proteins were eluted and subjected to western blotting. For antibodies see Supplemental Table S2.

RNA-ChIP

The RNA ChIP-IT Kit (Active Motif, 53024) was used according to the proprietary methods. Immunoprecipitated RNA (three biological replicates) was reverse transcribed using the SuperScript III First-Strand Synthesis SuperMix (ThermoFisher Scientific) and analyzed using SYBR green chemistry (SYBR Green Supermix) with the ABI 7900HT. For RT-qPCR primers see Supplemental Table S1.

CUT&Tag library preparation and data processing

CUT&Tag libraries (two biological replicates) were prepared using the CUT&Tag-IT™ Assay Kit (Active Motif, 53160) following to the manufacturer's protocol. 10^6 of Mel cells were collected for each biological replicate and two replicates were prepared. The Mel cells were bound to Concanavalin A Beads and Incubated with 1:50 rabbit polyclonal Phospho-Rpb1 CTD Ser5 and Ser2 antibody (Cellsignaling 13523 and 13499). Guinea pig a-rabbit antibody was used at 1:100 dilution as secondary antibody. Tagmentation was performed using pA-Tn5 Transposomes at 37°C for 60 mins. DNA was purified by DNA Purification Column, then

universal i5 primers and uniquely barcoded i7 primers were added to the DNA with 14cycles of PCR. Individual libraries were purified with SPRI beads and eluted with 20 µL DNA Purification Buffer. The libraries were sequenced on a MiSeq. For data analysis, see Supplementary Materials.

Quantification and statistical analysis

Statistical analysis was carried out using Graphpad Prism (version 9). Three biological replicates were performed for all experiments. All data were expressed as mean ± SD. Two-tailed Student's *t* test was used for significance test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. For complete data analysis, see Supplementary Methods.

RESULTS

The -81 kb murine *Myb* enhancer contains the TSS for a spliced, long intergenic non-coding RNA

The murine *Myb-Hbs1l* intergenic region contains previously characterized regulatory enhancers (Stadhouders *et al.*, 2012). In MEL cells and in primary erythroid cells, the five enhancers, located -36, -61, -68, -81 and -109 kb upstream of the *Myb* promoter, are occupied by the LDB1 complex that includes DNA binding transcription factors GATA1 and TAL1, bridging protein LMO2 and looping protein LDB1 (Figure 1A) (Stadhouders *et al.*, 2012). ChIP-seq and RNA-seq data for uninduced MEL cells indicates that RNA Pol II occupies several of the enhancer sites, but active RNA transcription is only observed at the -81 kb enhancer (Figure 1A) (Stadhouders *et al.*, 2012).

Rapid amplification of cDNA ends (5' and 3' RACE) revealed an unannotated 2-exon transcript at the *Myb* -81 kb enhancer that exists as two spliced isoforms with a single intron spanning more than 20 kb (Figure 1B, C). The primary transcript is 809 nt with a minor 357 nt shorter form attributable to early termination in exon 2. Stranded RNA-seq from induced and uninduced MEL cells publicly sourced from ENCODE indicates that *Myb* and the -81 kb transcript are divergently transcribed from opposite DNA strands (Figure S1). The transcript has very low coding potential according to the Coding Potential Assessment Tool (CPAT) (Wang et al., 2013) (Figure 1D). Thus, the transcript qualifies as a long non-coding RNA and its low abundance (about 10-fold lower than *Myb*) is consistent with that of an enhancer RNA (Li et al., 2016).

We named the transcript *Myrlin* (*Myb* enhancer long intergenic non-coding RNA). In uninduced MEL cells *Myrlin* is primarily nuclear localized, consistent with a potential role in gene regulation (Figure 1E). Expression of *Myb* decreases upon erythroid cell maturation, which is mirrored by decreases in both *Myb* and *Myrlin* upon differentiation of MEL cells by DMSO (Figure 1F). Like *Myb*, *Myrlin* is expressed at very low levels in E8.5 yolk sac primitive erythroid cells and then more robustly in E14.5 fetal liver definitive erythroid cells (Tober et al., 2008) (Figure 1G). These findings show that expression of *Myb* and *Myrlin* is coordinately

regulated in a developmental stage-specific fashion in murine erythroid cells and raises the possibility that *Myrlin* may play a role in regulating *Myb* expression.

Deletion of the *Myrlin* TSS reduces *Myb* expression

JASPAR motif analysis (http://jaspar.binf.ku.dk/) identified a TATA box located -25 nucleotides upstream of the 5' end of *Myrlin* as determined by RACE, and a GATA1 binding motif, site of LDB1 complex occupancy, located -56 nucleotides upstream (Figure 2A, Figure S2). This organization is consistent with the finding that most mouse erythroid-expressed non-coding RNAs are transcribed from conventional promoters regulated by known transcription factors and are regulated by similar Pol II release mechanisms (Henriques et al., 2018; Paralkar *et al.*, 2014).To investigate a role for *Myrlin* in *Myb* expression, we generated several MEL cell lines with small deletions that were designed to target the transcription start site (TSS) of *Myrlin* (Figure 2A, Figure S2). Three different mutations were obtained all of which reduced *Myrlin* and *Myb* transcription to varying degrees in uninduced MEL cells (Figure 2B). We chose for further study the 17 base pair deletion (Δ17, hereafter ΔTSS), which removes most of the sequence between the TATA box and the initiator element, and results in the strongest reduction of *Myrlin* and *Myb*.

To begin to characterize the impairment of *Myb* transcription ΔTSS cells, we performed ChIP to detect the occupancy of RNA Pol II. Compared to a control MEL cell line generated with a plasmid lacking an sgRNA, occupancy of Pol II in ΔTSS cells was reduced at the *Myb* promoter and at the -81 kb enhancer/*Myrlin* TSS as well as at the other enhancer sites (Figure 2C). TBP ChIP revealed that only the -81 kb enhancer was occupied by this member of the pre-initiation complex and in ΔTSS cells there was a marked reduction (Figure 2D). In accordance with results showing that fetal γ-globin transcription increases when *Myb* regulatory microRNAs are reduced in human cells (Bauer and Orkin, 2011; Bianchi et al., 2010; Mucenski et al., 1991; Sankaran *et al.*, 2011), after differentiation of ΔTSS MEL cells, there was a several fold increase in murine embryonic βh1 globin transcription (Figure 2E), consistent with *Myb* reduction. These results

support the idea that *Myrlin*, transcribed from the -81 kb *Myb* enhancer locus, is a positive regulator of *Myb* transcription.

Long-range *Myb* promoter and enhancer contacts are reduced after *Myrlin* TSS deletion

Myb expression is regulated by long-range interactions between the *Myb* promoter and LDB1-bound enhancers within the *Myb-Hbs1l* intergenic region (Stadhouders *et al.*, 2012). A CTCF site 30 kb downstream of the *Myb* promoter also participates in the *Myb* enhancer hub, likely through direct interaction between CTCF and LDB1 (Lee et al., 2017). Moreover, the disruption of this enhancer hub underlies *Myb* downregulation during erythroid differentiation (Stadhouders *et al.*, 2012). We used chromatin conformation capture (3C) to determine whether ΔTSS influenced contacts between *Myb* and its intergenic enhancers. Compared to control cells, ΔTSS cells displayed reduced interaction frequency between the *Myb* promoter and enhancers, which was particularly evident at the -36 kb and -81 kb enhancer sites and the -30kb CTCF site (Figure 3A).

The reduction of interaction frequency between *Myb* and its multiple enhancers by ΔTSS deletion closely resembles the reduced interactions observed upon LDB1 knock down using an shRNA in MEL cells (Stadhouders *et al.*, 2012). To investigate further, we carried out ChIP experiments to determine the occupancy of the LDB1 complex after reduction of *Myrlin* in ΔTSS cells. We observed that diminished long-range interactions in ΔTSS cells correlated with reduced LDB1 and TAL1 across the enhancers but reduced GATA1 occupancy was only notable at the -81 and -36 enhancers (Figure 3B-D). The H3K27ac mark, which indicates active enhancers, was not significantly affected in ΔTSS cells (Figure 3E). Together, our results show that the decrease in *Myb* expression upon deletion of the *Myrlin* TSS is accompanied by reduced LDB1 complex occupancy across the *Myb* enhancers and *Myb* enhancer hub disruption, although the enhancers remain in a potentially active state, retaining the H3K27ac mark.

CRISPRi for *Myrlin* affects *Myb* transcription but not enhancer interactions

We next employed CRISPRi as an alternative approach to reduce the *Myrlin* transcript without altering the sequence context at the *Myrlin* -81 kb enhancer. Catalytically dead CAS9 (dCAS9) retains the ability to target DNA and can be an adaptable block to transcription elongation when fused to a KRAB repression domain. Using two different gRNAs to target dCAS9-KRAB to *Myrlin* exons 1 or 2, we observed a 50% to 60% reduction in *Myrlin* transcription leading to a similar drop in *Myb* transcription as observed in the *Myrlin* ΔTSS deletion (Figure 4A).

ChIP localization of histone modifications at *Myb* enhancers revealed H3K9me3 at the -81kb enhancer in ΔTSS cells, a signature heterochromatin mark of KRAB-mediated repression (Figure 4B). There was no change in H3K27ac at any enhancer after *Myrlin* CRISPRi, (Figure 4C), similar to what we observed after the *Myrlin* ΔTSS deletion. However, in contrast to broad reduction of LDB1 at enhancer sites after Myrlin ΔTSS deletion, LDB1 enhancer occupancy was not significantly reduced by CRISPRi (Figure 4D). In addition, 3C experiments revealed relatively little change in interaction frequency between *Myb* and its enhancers after CRISPRi compared to a control clone generated with a non-targeted dCAS9-KRAB vector (Figure 4E). Thus, downregulation of *Myb* after CRISPRi silencing of *Myrlin* does not involve loss of the *Myb* enhancer hub, separating looping and transcription activation at this locus. We conclude that *Myb* downregulation after *Myrlin* CRISPRi silencing does not require loss of the *Myb* enhancer hub, raising the possibility that *Myrlin* has a direct role in *Myb* transcription activation.

Myrlin maintains H3K4me3 in the *Myb* locus through MLL1-WDR5

To further explore the role of *Myrlin* in *Myb* transcription activation, we focused on the *Myb* promoter and first exon/intron, which contain a CpG island that is highly enriched for H3K4me3 when *Myb* is active (Figure S3). ChIP-qPCR revealed strong H3K4me3 enrichment across these sequences, which was reduced following *Myrlin* transcriptional repression by ΔTSS CRISPRi (Figure 5A). The Mixed Lineage Leukemia 1 (MLL1) complex has been shown to deposit H3K4me3 modifications at the *Myb* locus and activate transcription (Nguyen et al., 2022). Therefore, we carried out ChIP for MLL1 complex components MLL1 and WDR5. We found a

strong decrease in both the MLL1 and WDR5 enrichment at the *Myb* locus after *Myrlin* CRISPRi (Figure 5B, C). Interestingly, HOTTIP long non-coding RNA targets the MLL1/WDR5 complex to promoters of HOXA genes to facilitate gene expression (Wang *et al.*, 2011), but currently no known eRNAs have been associated with MLL/WDR5 complex in the context of gene regulation.

The results so far suggest that *Myrlin* may directly participate in activation of *Myb* transcription through recruiting MLL1. To further investigate, we performed *Myrlin* RNA-ChIP and found that the *Myrlin* transcript interacts with the MLL1 component WDR5 (Figure 5D). RNA-ChIP also detected interaction that were not statistically significant between *Myrlin* and MLL1 and with Menin, another component of the MLL1 complex. Therefore, we performed RNA pull down using biotinylated *Myrlin* and blotted against these components of the MLL1 complex. The results confirmed interaction of *Myrlin* with WDR5 and further revealed that *Myrlin* could pulldown complex components MLL1 and Menin. Biotinylated *Myrlin* did not pull down SET1a or SET1b, which are members of different COMPASS-like complexes, nor was there any interaction with Tubulin (TUB), which served as a negative control (Figure 5E). These results strongly support that *Myrlin* plays a role in MLL1 complex recruitment to promote H3K4me3 deposition and subsequent transcription of *Myb* and that *Myrlin* loss compromises this series of events and *Myb* transcription activation.

RNA Pol II pausing and CDK9 occupancy are affected by CRISPRi of *Myrlin*

MLL1 complexes can maintain target gene expression through regulating both transcriptional initiation and elongation. MLL1 loss results in loss of the CDK9, the protein kinase subunit of pTEFb that confers phosphorylation on Pol II ser2 to drive transcription elongation in hematopoietic cells (Fujinaga et al., 2023; Mishra et al., 2014). Previous studies have shown that inhibition of CDK9 strongly reduced transcription elongation through the *Myb* gene body and resulted in pausing of Pol II within the *Myb* first intron (Mitra et al., 2012; Stadhouders *et al.*, 2012). Using ChIP-qPCR, we found that CDK9 recruitment across the *Myb* enhancers and in *Myb* exon 1 and intron 1 was severely diminished after *Myrlin* CRISPRi (Figure 6A).

Subsequent ChIP-qPCR analysis of RNA Pol II localization in the *Myb* locus revealed decreased occupancy at several *Myb* enhancers but no significant difference at the *Myb* promoter (Figure 6B). Interestingly, after *Myrlin* CRISPRi, RNA Pol II accumulated within the *Myb* gene body, particularly at *Myb* exon1/intron 1, which is consistent with reduction of *Myb* transcripts.

To further investigate this result, we performed CUT & TAG for the Ser5P and Ser2P phosphorylated forms of RNA Pol II, which represent the initiating and elongating forms of the enzyme, respectively. The genome browser view in Figure 6C illustrates the reduction of both phosphorylated forms of RNA Pol II across the *Myb* locus after *Myrlin* CRISPRi. Quantitation of these data shows about a 30% reduction of Pol II Ser5P and Ser2P within the *Myb* gene body and reduced occupancy at each of the enhancers (Figure 6D, E). To obtain increased resolution of this result, we analyzed the signals of the shorter (<120 bp) and longer (>270 bp) Pol II Ser5P CUT & TAG fragments, which distinguishes Pol II in the pre-initiation state at the TSS and paused Pol II at peaks up- and downstream of the TSS, respectively (Miura and Chen, 2020) (Figure 6F). *Myrlin* CRISPRi reduced the signal for the shorter fragments, suggesting some decrease in Pol II recruitment upon *Myrlin* loss. At the same time, the signal associated with longer fragments, indicative of paused Pol II, strongly increased across exon/intron 1.

We calculated the pausing index of Pol II Ser5P from the longer fragment data based on the Ser5P ratio at the TSS (exon1) and in the gene body and found about a 3-fold higher pausing index in *Myb* after *Myrlin* CRISPRi compared to controls (Figure 6G). Similar results were obtained for Pol II pausing in *Myb* using RNAseq peaks without separation by fragment size (Figure S4A, B). These results connect *Myrlin* to the mechanism by which transcription of *Myb* is regulated and suggest that *Myrlin*, through MLL1, may help to recruit or stabilize CDK9 at *Myb,* which is necessary for efficient RNA Pol II elongation through the gene.

KLF1 interacts with *Myrlin* transcripts within the *Myb* active chromatin hub

What could be the basis of *Myrlin* local function in *Myb* transcription? *Myrlin* interaction with MLL1 at the actively transcribed *Myb* promoter raised the question of whether *Myrlin* is retained within the enhancer hub by tethering to the -81 kb enhancer, adjacent to its transcription start site. We performed ChIRP experiments to determine *Myrlin* localization in the *Myb* locus. Two sets of *Myrlin* probes (odd and even probes) can specifically capture the *Myrlin* transcript (Figure 7A). We found prominent peaks for *Myrlin* at its two exons and at the -81 kb enhancer site but not at other sites across the locus (Figure 7B). Since the -81 kb enhancer is close to *Myrlin* exon 1, the signal at that site could have been contributed to by cross-linking of *Myrlin* bound to exon1. Therefore, we asked whether *Myrlin* interacted with -81 kb-bound transcription factors, specifically KLF1, which uniquely binds to the -81 site among the enhancers. KLF1 is required for full *Myb* transcription activation (Stadhouders *et al.*, 2012). Indeed, RNA ChIP for KLF1 revealed interaction with *Myrlin*, which was confirmed by biotinylated RNA pull down (Figure 7C, D). ChIP-qPCR confirmed occupancy of KLF1 at the -81 kb *Myb* enhancer but not at -68 kb, as expected (Figure 7E). KLF1 occupancy was not affected by *Myrlin* CRISPRi. We conclude that KLF1 may contribute to the localization of *Myrlin* within the *Myb* enhancer hub.

In summary, structural components contributing to *Myb*-enhancer interactions within the hub, including the LDB1 complex and KLF1, are present and the hub forms normally when *Myrlin* is lost through CRISPRi (Figure 7F). However, in the absence of *Myrlin*, recruitment of MLL1/WDR5 is reduced, resulting in poor binding of CDK9 and RNA Pol II in the *Myb* locus. Moreover, RNA Pol II pausing within the *Myb* first exon/intron and failure to complete *Myb* transcripts are observed upon *Myrlin* loss. We conclude that the *Myrlin* eRNA has important recruitment functions for MLL1 complex that are required for *Myb* transcription activation beyond formation of the *Myb* enhancer hub.

DISCUSSION

Hundreds of long non-coding RNAs are expressed during erythropoiesis but evidence for their function remains anecdotal (Alvarez-Dominguez *et al.*, 2014; Hu *et al.*, 2011; Ivaldi *et al.*, 2018; Paralkar *et al.*, 2014). We identified *Myrlin*, an eRNA, transcribed from the -81 kb murine *Myb* enhancer as a positive regulator of *Myb*. A rigorous paradigm to deduce the function of lncRNA and eRNA loci involves dissection of the functional roles of the RNA transcript versus the act of enhancer transcription, both of which can influence target gene expression and/or chromatin organization (Bassett et al., 2014). To this end, we analyzed the role of the RNA molecule itself by using *Myrlin* RNA ChIP and biotinylated *Myrlin* pull down assays to document the importance of the transcript, *per se*. We found that *Myrlin* interacts with MLL1/WDR5 to promote *Myb* transcription. In the absence of *Myrlin*, MLL1 recruitment to *Myb* and *Myb* promoter H3K4me3 modification is reduced. In addition, CDK9 is reduced at the *Myb* promoter and RNA Pol II accumulates excessively within the *Myb* first exon/intron. Thus, *Myrlin* is required for Pol II pause release to promote *Myb* transcription.

In our initial experiments to reduce *Myrlin* transcription, we deleted 17 bp of the *Myrlin* TSS. This resulted in loss of the LDB1 looping complex from all the *Myb* enhancer sites, disruption of the *Myb* enhancer hub and reduction of *Myb* transcription. The results suggest interdependence among multiple *Myb* enhancers for formation of the hub. Hubs are understood from single allele interaction experiments to represent close association of all hub sequences within which multiple enhancers show preferential aggregation with each other and with the genes they regulate (Oudelaar et al., 2018). Enhancer interdependence has been observed at the *IgK* super enhancer cluster where deletion of one enhancer affects interactions among the others (Jiang et al., 2016; Proudhon et al., 2016). This model comports well with the idea that extrusion of chromatin loops through the action of cohesin complexes and CTCF will bring many points within an extended locus into close proximity where their interactions may be stabilized by specific transcription factors (Fudenberg et al., 2016; Sanborn et al., 2015). Such a hub of interactions and transcription factor density may facilitate accumulation of transcriptional components such as RNA Pol II, possibly involving liquid-liquid phase separation, favoring transcription (Furlong and Levine, 2018).

The precise role of *Myrlin* in *Myb* enhancer hub formation, if any, remains to be clarified, since the LDB1 complex was lost from the enhancer sites in the ΔTSS cells, which is well documented to compromise LDB1 enhancer looping in this locus (Stadhouders *et al.*, 2012). Intriguingly, LDB1 binding, and the hub formation were not affected when *Myrlin* transcription was repressed by CAS9-KRAB, which did not affect LDB1 complex occupancy. It remains unclear why LDB1 is lost from the enhancer sites in ΔTSS cells but not after *Myrlin* CRISPRi. One possibility is that the 17 bp deletion of the TSS, near the LDB1 complex motif at the -81 kb enhancer, distorts the local chromatin architecture sufficiently to reduce LDB1 binding, which then destabilizes binding at the other hub enhancers, and looping. Regardless, losing *Myrlin* under the circumstances of KRAB repression did not have this effect on LDB1 binding yet still resulted in *Myb* downregulation, separating looping from transcription activation. Looping in the absence of transcription has been observed at the β-globin locus in erythroid cells and, more generally, in leukemia cells after drug treatment to inhibit BET proteins (Crump et al., 2021; Krivega et al., 2014; Palstra et al., 2008). These data support the idea that enhancer looping mechanistically precedes transcription activation.

To explore the difference between looped and transcriptionally active *Myb* loci versus inactive looped loci, we localized RNA Pol II and found a decrease in Pol II recruitment to the *Myb* promoter after *Myrlin* CRISPRi and significant pausing in intron/exon 1. Pol II pausing is correlated with transcriptional repression during terminal erythroid repression (Murphy et al., 2021). Previous work had suggested a Pol II pause region in *Myb* intron 1 that was proposed to correspond to either a stem-loop forming sequence followed by a poly (dT) tract 1.7 Kb downstream of the *Myb* TSS or to a CTCF site about 2 Kb downstream (Mitra *et al.*, 2012; Stadhouders *et al.*, 2012). These groups reported that transcription elongation through *Myb* was inhibited in erythroid cells or in breast cancer cells by CDK9 inhibitors. Our high-resolution CUT & TAG results for Pol II Ser5 and Pol II Ser3 showed localization of paused transcripts across *Myb* exon 1 and into intron 1, likely encompassing the sites previously suggested as pause sites. We were able to show that *Myrlin* functions to recruit or stabilize CDK9 in the *Myb* exon/intron1 Pol II pause region. It has been suggested that interactions of an enhancer with its

target promoter can stimulate Pol II pause release (Ghavi-helm 2014, Rahl 2010) and that this property may be related to CDK9 activity (Ghavi-Helm et al., 2014; Gressel et al., 2017; Rahl et al., 2010). *Myrlin* provides a link between an eRNA and CDK9 at a target gene.

We observed that *Myrlin* interacts with several subunits of the MLL1 complex, including MLL1, WDR5 and Menin, and that *Myrlin* is important for MLL1 occupancy in the *Myb* locus and for H3K4me3 modification at the *Myb* promoter CpG island. Reduced MLL1 may underlie loss of CDK9 at the *Myb* locus, as CDK9 reduction was reported on a global scale after deletion of MLL1 in hematopoietic stem and progenitor cells (Mishra *et al.*, 2014). In this scenario, the *Myrlin* eRNA becomes important for *Myb* transcription after the enhancer hub is formed. We propose that *Myrlin* interacts with MLL1 and the CDK9 component of pTEFb to recruit RNA Pol II to the locus and assure efficient elongation through the *Myb* exon/intron 1 pause region. After *Myrlin* CRISPR, MLL1 and CDK9 are reduced, and Pol II pausing reduces *Myb* transcripts. The above putative functions of *Myrlin* would require its localized presence in the *Myb* locus/enhancer hub. We explored the potential role of KLF1 in such localization since it binds uniquely to the -81 enhancer where the *Myrlin* eRNA is transcribed. Indeed, we documented interaction of *Myrlin* and KLF1 by biotinylated RNA pull down.

One of the most promising strategies for treating Sickle Cell Disease and β-Thalassemia is reactivation of fetal hemoglobin production in erythroid cells of adult patients. GWAS revealed an association between single nucleotide polymorphisms (SNPs) in the *Myb* ─ *Hbs1l* intergenic region, encompassing the several *Myb* enhancers, that reduce *Myb* gene expression and elevate fetal hemoglobin in human adult erythroid cells (Jiang *et al.*, 2006; Stadhouders *et al.*, 2014; Wahlberg *et al.*, 2009). Two of these SNPs, i.e., rs66650371 and rs77755698, are located within the LDB1 complex GATA1/TAL1 binding peak at the human -84 kb *Myb* enhancer, which is the homologue of the murine -81 kb enhancer (Farrell et al., 2011; Stadhouders *et al.*, 2014). These SNPs reduce LDB1 binding to the -84 kb enhancer and decrease interaction frequency with the *MYB* promoter (Stadhouders *et al.*, 2014). The effect of the SNPs on interactions of the other enhancers with MYB was not tested in this work but our results suggest that overall formation of the hub is likely affected.

Recently, a 1,283 bp non-coding RNA, HMI-LNCRNA, was reported to arise from the -84 kb human *MYB* enhancer (Morrison *et al.*, 2018). Thousands of human lncRNAs have homologues in other species with similar expression patterns but low sequence conservation (Hezroni et al., 2015; Paralkar *et al.*, 2014). However, BLATN sequence comparison of HMI-LNCRNA to the mouse genome revealed a conserved 'patch' of 378 nt (85.8% homology) shared with the 5' end of *Myrlin*, something commonly observed for these poorly conserved lncRNAs (Hezroni *et al.*, 2015) (Figure S5). Within this homology patch lie the GATA1/LDB1 complex binding site that mediates looping to the *Myb* promoter in human and mouse erythroid cells, the -81 kb KLF1 binding site and the *Myrlin* TATA box and first exon. Thus, it seems likely that the pause release function of *Myrlin* at the *Myb* gene may be conserved between species. While any *Myb* regulatory function ascribed to *Myrlin* in mouse cells remains to be established for the related transcript in human cells, the *Myb* transcription mechanisms participated in by *Myrlin* suggest the possibility of their utility as targets to increase HbF production to ameliorate the severity of hemoglobinopathies.

ACKNOWLEDGEMENTS

We acknowledge computational support from Dr. Lecong Zhou, the NIDDK Genomics Core for sequencing support and the NIH HPC (Biowulf) for computational support. This work was funded by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH (DK 075033 to AD).

AUTHOR CONTRIBUTION

AD and IK conceived the project; IK, JK and BL supervised experiments. LFD, JK and MJM performed experiments. LFD, JK and AD wrote the paper; all authors edited the paper.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

REFERENCES

Alvarez-Dominguez, J.R., Hu, W., Yuan, B., Shi, J., Park, S.S., Gromatzky, A.A., van, O.A., and Lodish, H.F. (2014). Global discovery of erythroid long noncoding RNAs reveals novel regulators of red cell maturation. Blood *123*, 570-581. blood-2013-10-530683 [pii];10.1182/blood-2013-10-530683 [doi].

Arnold, P.R., Wells, A.D., and Li, X.C. (2019). Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. Front Cell Dev Biol *7*, 377. 10.3389/fcell.2019.00377.

Bassett, A.R., Akhtar, A., Barlow, D.P., Bird, A.P., Brockdorff, N., Duboule, D., Ephrussi, A., Ferguson-Smith, A.C., Gingeras, T.R., Haerty, W., et al. (2014). Considerations when investigating lncRNA function in vivo. Elife *3*, e03058. 10.7554/eLife.03058 [doi].

Bauer, D.E., and Orkin, S.H. (2011). Update on fetal hemoglobin gene regulation in hemoglobinopathies. Curr. Opin. Pediatr *23*, 1-8. 10.1097/MOP.0b013e3283420fd0 [doi].

Bianchi, E., Zini, R., Salati, S., Tenedini, E., Norfo, R., Tagliafico, E., Manfredini, R., and Ferrari, S. (2010). c-myb supports erythropoiesis through the transactivation of KLF1 and LMO2 expression. Blood *116*, e99-110. 10.1182/blood-2009-08-238311.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev *25*, 1915-1927. gad.17446611 [pii];10.1101/gad.17446611 [doi].

Cenik, B.K., and Shilatifard, A. (2021). COMPASS and SWI/SNF complexes in development and disease. Nat Rev Genet *22*, 38-58. 10.1038/s41576-020-0278-0.

Crump, N.T., Ballabio, E., Godfrey, L., Thorne, R., Repapi, E., Kerry, J., Tapia, M., Hua, P., Lagerholm, C., Filippakopoulos, P., et al. (2021). BET inhibition disrupts transcription but retains enhancer-promoter contact. Nat Commun *12*, 223. 10.1038/s41467-020-20400-z.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome. Res *22*, 1775-1789. 22/9/1775 [pii];10.1101/gr.132159.111 [doi].

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature *489*, 101-108. nature11233 [pii];10.1038/nature11233 [doi].

Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L., Allis, C.D., and Roeder, R.G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol *13*, 713-719. 10.1038/nsmb1128.

Farrell, J.J., Sherva, R.M., Chen, Z.Y., Luo, H.Y., Chu, B.F., Ha, S.Y., Li, C.K., Lee, A.C., Li, R.C., Li, C.K., et al. (2011). A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. Blood *117*, 4935-4945. 10.1182/blood-2010-11-317081.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Rep *15*, 2038-2049. S2211-1247(16)30530-7 [pii];10.1016/j.celrep.2016.04.085 [doi].

Fujinaga, K., Huang, F., and Peterlin, B.M. (2023). P-TEFb: The master regulator of transcription elongation. Mol Cell *83*, 393-403. 10.1016/j.molcel.2022.12.006.

Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. Science *361*, 1341-1345. 10.1126/science.aau0320.

Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., and Furlong, E.E. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. Nature *512*, 96-100. nature13417 [pii];10.1038/nature13417 [doi].

Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. Nat Rev Genet *21*, 102-117. 10.1038/s41576-019-0184-5.

Gressel, S., Schwalb, B., Decker, T.M., Qin, W., Leonhardt, H., Eick, D., and Cramer, P. (2017). CDK9-dependent RNA polymerase II pausing controls transcription initiation. Elife *6*. 10.7554/eLife.29736.

Gurumurthy, A., Yu, D.T., Stees, J.R., Chamales, P., Gavrilova, E., Wassel, P., Li, L., Stribling, D., Chen, J., Brackett, M., et al. (2021). Super-enhancer mediated regulation of adult beta-globin gene expression: the role of eRNA and Integrator. Nucleic Acids Res *49*, 1383-1396. 10.1093/nar/gkab002.

Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de, L.W., and Forne, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc *2*, 1722-1733.

Henriques, T., Scruggs, B.S., Inouye, M.O., Muse, G.W., Williams, L.H., Burkholder, A.B., Lavender, C.A., Fargo, D.C., and Adelman, K. (2018). Widespread transcriptional pausing and elongation control at enhancers. Genes Dev *32*, 26-41. 10.1101/gad.309351.117.

Herman, A.B., Tsitsipatis, D., and Gorospe, M. (2022). Integrated lncRNA function upon genomic and epigenomic regulation. Mol Cell *82*, 2252-2266. 10.1016/j.molcel.2022.05.027.

Hezroni, H., Koppstein, D., Schwartz, M.G., Avrutin, A., Bartel, D.P., and Ulitsky, I. (2015). Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. Cell Rep *11*, 1110-1122. 10.1016/j.celrep.2015.04.023.

Hou, T.Y., and Kraus, W.L. (2021). Spirits in the Material World: Enhancer RNAs in Transcriptional Regulation. Trends Biochem Sci *46*, 138-153. 10.1016/j.tibs.2020.08.007.

Hu, W., Alvarez-Dominguez, J.R., and Lodish, H.F. (2012). Regulation of mammalian cell differentiation by long non-coding RNAs. EMBO Rep *13*, 971-983. embor2012145 [pii];10.1038/embor.2012.145 [doi].

Hu, W., Yuan, B., Flygare, J., and Lodish, H.F. (2011). Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. Genes Dev *25*, 2573-2578. gad.178780.111 [pii];10.1101/gad.178780.111 [doi].

Ivaldi, M.S., Diaz, L.F., Chakalova, L., Lee, J., Krivega, I., and Dean, A. (2018). Fetal gamma-globin genes are regulated by the BGLT3 long noncoding RNA locus. Blood *132*, 1963-1973. blood-2018-07-862003 [pii];10.1182/blood-2018-07-862003 [doi].

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector, T.D., and Thein, S.L. (2006). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood *108*, 1077-1083. 108/3/1077 [pii];10.1182/blood-2006-01-008912 [doi].

Jiang, T., Raviram, R., Snetkova, V., Rocha, P.P., Proudhon, C., Badri, S., Bonneau, R., Skok, J.A., and Kluger, Y. (2016). Identification of multi-loci hubs from 4C-seq demonstrates the functional importance of simultaneous interactions. Nucleic Acids Res *44*, 8714-8725. 10.1093/nar/gkw568.

Kim, J., and Dean, A. (2021). Enhancers navigate the three-dimensional genome to direct cell fate decisions. Curr Opin Struct Biol *71*, 101-109. 10.1016/j.sbi.2021.06.005.

Kim, S., and Wysocka, J. (2023). Deciphering the multi-scale, quantitative cis-regulatory code. Mol Cell *83*, 373-392. 10.1016/j.molcel.2022.12.032.

Kim, T.K., and Shiekhattar, R. (2016). Diverse regulatory interactions of long noncoding RNAs. Curr. Opin. Genet. Dev *36*, 73-82. S0959-437X(16)30023-5 [pii];10.1016/j.gde.2016.03.014 [doi].

Krivega, I., Dale, R.K., and Dean, A. (2014). Role of LDB1 in the transition from chromatin looping to transcription activation. Genes Dev *28*, 1278-1290. gad.239749.114 [pii];10.1101/gad.239749.114 [doi].

Krivega, I., and Dean, A. (2017). LDB1-mediated enhancer looping can be established independent of mediator and cohesin. Nucleic Acids Res *45*, 8255-8268. 3831094 [pii];10.1093/nar/gkx433 [doi].

Lee, J., Krivega, I., Dale, R.K., and Dean, A. (2017). The LDB1 Complex Co-opts CTCF for Erythroid Lineage-Specific Long-Range Enhancer Interactions. Cell Rep *19*, 2490-2502. S2211-1247(17)30745-3 [pii];10.1016/j.celrep.2017.05.072 [doi].

Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping back to leap forward: transcription enters a new era. Cell *157*, 13-25. S0092-8674(14)00201-3 [pii];10.1016/j.cell.2014.02.009 [doi].

Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat. Rev. Genet *17*, 207-223. nrg.2016.4 [pii];10.1038/nrg.2016.4 [doi].

Milne, T.A., Dou, Y., Martin, M.E., Brock, H.W., Roeder, R.G., and Hess, J.L. (2005). MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci U S A *102*, 14765-14770. 10.1073/pnas.0503630102.

Mishra, B.P., Zaffuto, K.M., Artinger, E.L., Org, T., Mikkola, H.K., Cheng, C., Djabali, M., and Ernst, P. (2014). The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. Cell Rep *7*, 1239-1247. 10.1016/j.celrep.2014.04.015.

Mitra, P., Pereira, L.A., Drabsch, Y., Ramsay, R.G., and Gonda, T.J. (2012). Estrogen receptor-alpha recruits P-TEFb to overcome transcriptional pausing in intron 1 of the MYB gene. Nucleic Acids Res *40*, 5988-6000. 10.1093/nar/gks286.

Miura, M., and Chen, H. (2020). CUT&RUN detects distinct DNA footprints of RNA polymerase II near the transcription start sites. Chromosome Res *28*, 381-393. 10.1007/s10577-020-09643-0.

Morrison, T.A., Wilcox, I., Luo, H.Y., Farrell, J.J., Kurita, R., Nakamura, Y., Murphy, G.J., Cui, S., Steinberg, M.H., and Chui, D.H.K. (2018). A long noncoding RNA from the HBS1L-MYB intergenic region on chr6q23 regulates human fetal hemoglobin expression. Blood Cells Mol. Dis *69*, 1-9. S1079-9796(17)30440-0 [pii];10.1016/j.bcmd.2017.11.003 [doi].

Mucenski, M.L., McLain, K., Kier, A.B., Swerdlow, S.H., Schreiner, C.M., Miller, T.A., Pietryga, D.W., Scott, W.J., Jr., and Potter, S.S. (1991). A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. Cell *65*, 677-689. 10.1016/0092-8674(91)90099-k.

Murphy, Z.C., Murphy, K., Myers, J., Getman, M., Couch, T., Schulz, V.P., Lezon-Geyda, K., Palumbo, C., Yan, H., Mohandas, N., et al. (2021). Regulation of RNA polymerase II activity is essential for terminal erythroid maturation. Blood *138*, 1740-1756. 10.1182/blood.2020009903.

Nguyen, N., Gudmundsson, K.O., Soltis, A.R., Oakley, K., Roy, K.R., Han, Y., Gurnari, C., Maciejewski, J.P., Crouch, G., Ernst, P., et al. (2022). Recruitment of MLL1 complex is essential for SETBP1 to induce myeloid transformation. iScience *25*, 103679. 10.1016/j.isci.2021.103679.

Oudelaar, A.M., Davies, J.O.J., Hanssen, L.L.P., Telenius, J.M., Schwessinger, R., Liu, Y., Brown, J.M., Downes, D.J., Chiariello, A.M., Bianco, S., et al. (2018). Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. Nat Genet *50*, 1744-1751. 10.1038/s41588-018-0253-2.

Palstra, R.J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., and de, L.W. (2008). Maintenance of Long-Range DNA Interactions after Inhibition of Ongoing RNA Polymerase II Transcription. PLoS. ONE *3*, e1661.

Panigrahi, A., and O'Malley, B.W. (2021). Mechanisms of enhancer action: the known and the unknown. Genome Biol *22*, 108. 10.1186/s13059-021-02322-1.

Paralkar, V.R., Mishra, T., Luan, J., Yao, Y., Kossenkov, A.V., Anderson, S.M., Dunagin, M., Pimkin, M., Gore, M., Sun, D., et al. (2014). Lineage and species-specific long noncoding RNAs during erythro-megakaryocytic development. Blood *123*, 1927-1937. blood-2013-12-544494 [pii];10.1182/blood-2013-12-544494 [doi].

Plank, J.L., and Dean, A. (2014). Enhancer Function: Mechanistic and Genome-Wide Insights Come Together. Mol. Cell *55*, 5-14. S1097-2765(14)00523-1 [pii];10.1016/j.molcel.2014.06.015 [doi].

Proudhon, C., Snetkova, V., Raviram, R., Lobry, C., Badri, S., Jiang, T., Hao, B., Trimarchi, T., Kluger, Y., Aifantis, I., et al. (2016). Active and Inactive Enhancers Cooperate to Exert Localized and Long-Range Control of Gene Regulation. Cell Rep *15*, 2159-2169. 10.1016/j.celrep.2016.04.087.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432-445. 10.1016/j.cell.2010.03.030.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc *8*, 2281-2308. nprot.2013.143 [pii];10.1038/nprot.2013.143 [doi].

Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. U. S. A *112*, E6456-E6465. 1518552112 [pii];10.1073/pnas.1518552112 [doi].

Sankaran, V.G., Menne, T.F., Scepanovic, D., Vergilio, J.A., Ji, P., Kim, J., Thiru, P., Orkin, S.H., Lander, E.S., and Lodish, H.F. (2011). MicroRNA-15a and -16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. Proc Natl Acad Sci U. S. A *108*, 1519-1524. 1018384108 pii];10.1073/pnas.1018384108 [doi].

Sartorelli, V., and Lauberth, S.M. (2020). Enhancer RNAs are an important regulatory layer of the epigenome. Nat Struct Mol Biol *27*, 521-528. 10.1038/s41594-020-0446-0.

Song, S.-H., Hou, C., and Dean, A. (2007). A positive role for NLI/Ldb1 in long-range β-globin locus control region function. Mol. Cell *28*, 810-822.

Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., van, I.W., Lenhard, B., Rooks, H., Best, S., Menzel, S., et al. (2014). HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. J. Clin. Invest *124*, 1699-1710. 71520 [pii];10.1172/JCI71520 [doi].

Stadhouders, R., Cico, A., Stephen, T., Thongjuea, S., Kolovos, P., Baymaz, H.I., Yu, X., Demmers, J., Bezstarosti, K., Maas, A., et al. (2015). Control of developmentally primed erythroid genes by combinatorial co-repressor actions. Nat Commun *6*, 8893. 10.1038/ncomms9893.

Stadhouders, R., Thongjuea, S., Andrieu-Soler, C., Palstra, R.J., Bryne, J.C., van den Heuvel, A., Stevens, M., de, B.E., Kockx, C., van der Sloot, A., et al. (2012). Dynamic long-range chromatin interactions control Myb proto-oncogene transcription during erythroid development. EMBO J *31*, 986-999. emboj2011450 [pii];10.1038/emboj.2011.450 [doi].

Statello, L., Guo, C.J., Chen, L.L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol *22*, 96-118. 10.1038/s41580-020-00315-9.

Tober, J., McGrath, K.E., and Palis, J. (2008). Primitive erythropoiesis and megakaryopoiesis in the yolk sac are independent of c-myb. Blood *111*, 2636-2639. blood-2007-11-124685 [pii];10.1182/blood-2007-11-124685 [doi].

Wahlberg, K., Jiang, J., Rooks, H., Jawaid, K., Matsuda, F., Yamaguchi, M., Lathrop, M., Thein, S.L., and Best, S. (2009). The HBS1L-MYB intergenic interval associated with elevated HbF levels shows characteristics of a distal regulatory region in erythroid cells. Blood *114*, 1254-1262. blood-2009-03-210146 [pii];10.1182/blood-2009-03-210146 [doi].

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124. 10.1038/nature09819.

Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res *41*, e74. 10.1093/nar/gkt006.

Wang, X., Angelis, N., and Thein, S.L. (2018). MYB - A regulatory factor in hematopoiesis. Gene *665*, 6-17. S0378-1119(18)30439-6 [pii];10.1016/j.gene.2018.04.065 [doi].

FIGURE LEGENDS

Figure 1. The *Myb* -81 kb enhancer is the transcription start site for the noncoding *Myrlin* RNA. (A) ENCODE ChIP-Seq of LDB1 complex components (GATA1, LDB1 and TAL1), CTCF, and PolII in the *Myb-Hbs1l* locus in MEL cells. PolyA RNA-seq is shown in MEL cells. Intergenic LDB1 complex binding sites are highlighted (grey vertical bars). (B) Nested PCR products from 5' and 3' RACE in MEL cells. (C) Poly-A RNA-seq and Ldb1 ChIP-seq tracks from ENCODE for MEL at the -81 kb enhancer and downstream *Myrlin* exon 2. (D) Prediction of coding potential for *Myrlin* and select other transcripts as determined by CPAT (see text). (E) Relative expression of *Myrlin* in nuclear and cytoplasmic fractions of MEL cells determined by RT-PCR. *MALAT1* lncRNA and *ActB* provided nuclear and cytoplasmic controls, respectively. (F) Total RNA of uninduced and induced MEL cells was used to determine relative expression of *Myrlin* and *Myb* by RT-PCR. Expression was normalized to *ActB*. (G) Total RNA of E8.5 yolk sac cells (YS) and E14.5 fetal liver cells (FL) was used to determine relative expression of *Myrlin*, *Myb* and *Hbb-b1* by RT-PCR. Expression was normalized to *ActB*. Error bars indicate SEM of 3 independent biological experiments. (*) $P < 0.05$, (**) $P<0.01$, (***) $P<0.001$ by Student's t-test.

Figure 2. CRISPR deletions at the TSS for *Myrlin* affect *Myb* expression. (A) Schematic diagram of the CRISPR/Cas9-mediated 17bp deletion upstream of the *Myrlin* TSS. Smaller deletions were within the 17 bp deletion or extended upstream. PCR (below) shows the relative sizes for the 7bp, 8bp and 17bp deletion PCR products using WT or mutant gDNA. (B) Relative expression of *Myrlin* (exon 1, exon1/2 junction, exon 2) and *Myb* (exon 2) in MEL CRISPR/Cas9 control cells (no gRNA) and MEL 7bp, 8bp and 17bp CRISPR/Cas9 deletion mutants. Expression was normalized to *ActB*. (C) RNA Pol II-ChIP in MEL CRISPR/Cas9 control cells and MEL 17bp CRISPR/Cas9 deletion mutant (ΔTSS) at the *Myb* gene promoter and enhancer sites (-36, -61, -68, -71 and -109 kb). (D) ChIP for TBP at the *Myb* enhancer sites as in panel C for MEL CRISPR/Cas9 control cells and MEL ΔTSS CRISPR/Cas9 deletion mutant. (E) Relative expression of Hbb-bh1, Hbb-b1 and Hbb-y in induced MEL CRISPR/Cas9 control cells and MEL ΔTSS CRISPR/Cas9 deletion mutant. Error bars indicate SEM of 3 independent biological experiments. (*) $P < 0.05$, (**) $P<0.01$, (***) $P<0.001$ by Student's t-test.

Figure 3. Chromatin organization of the *Myb* locus and transcription factor occupancy is affected by reduction of *Myrlin* in ΔTSS cells. (A) Chromatin conformation capture (3C) interaction frequency between the enhancers found within the *Myb-Hbs1l* intergenic region using the *Myb* promoter as the anchor (black bar) observed for ΔTSS and control cell uninduced MEL cells. (B) GATA1 occupancy at *Myb* enhancers in control and ΔTSS uninduced MEL cell lines determined by ChIP-qPCR. (C) TAL1 occupancy at *Myb* enhancers in control and ΔTSS uninduced MEL cell lines determined by ChIP-qPCR. (D) LDB1 occupancy at *Myb* enhancers in control and ΔTSS uninduced MEL cell lines determined by ChIP-qPCR. (E) H3K27ac normalized to H3 occupancy at *Myb* enhancers in control and ΔTSS uninduced MEL cell lines determined by ChIP-qPCR. Error bars indicate SEM of 3 independent biological experiments. (*) $P < 0.05$ and (**) $P<0.01$ by Student's t-test.

Figure 4. CRISPRi targeting of *Myrlin* compromises *Myb* transcription but not hub formation. (A) Expression of *Myb* and *Myrlin* monitored by RT-qPCR in CRISPRi uninduced MEL cells targeted with dCas9 sgRNA1, sgRNA3 or without an sgRNA (control). (B) ChIP-qPCR for H3K9me3 across the *Myb* locus before and after *Myrlin* CRISPi in uninduced MEL cells. (C) ChIP-qPCR for H3K27ac across the *Myb* locus before and after *Myrlin* CRISPi. (D) ChIP-qPCR for LDB1 across the *Myb* locus before and after *Myrlin* CRISPi. (E) Chromatin conformation capture (3C) interaction frequency between the *Myb* enhancers using the *Myb* promoter as the anchor (black bar) observed after *Myrlin* CRISPRi in uninduced MEL cells targeted with dCas9 sgRNA or without an sgRNA (control). Error bars indicate SEM of 3 independent biological experiments. (*) $P < 0.05$ and (**) $P<0.01$ by Student's t-test.

Figure 5. *Myrlin* interacts with MLL1 complex. (A) ChIP-qPCR for H3K4me3 across *Myb* sequences. (B) ChIP-qPCR for MLL1 at the *Myb* gene. (C) ChIP-qPCR for MLL1 complex component WDR5 at the *Myb* gene. (D) Myrlin RNA ChIP. (E) RNA pull down using biotinylated *Myrlin* and blotting using antibodies to MLL1 components. Tubulin served as a

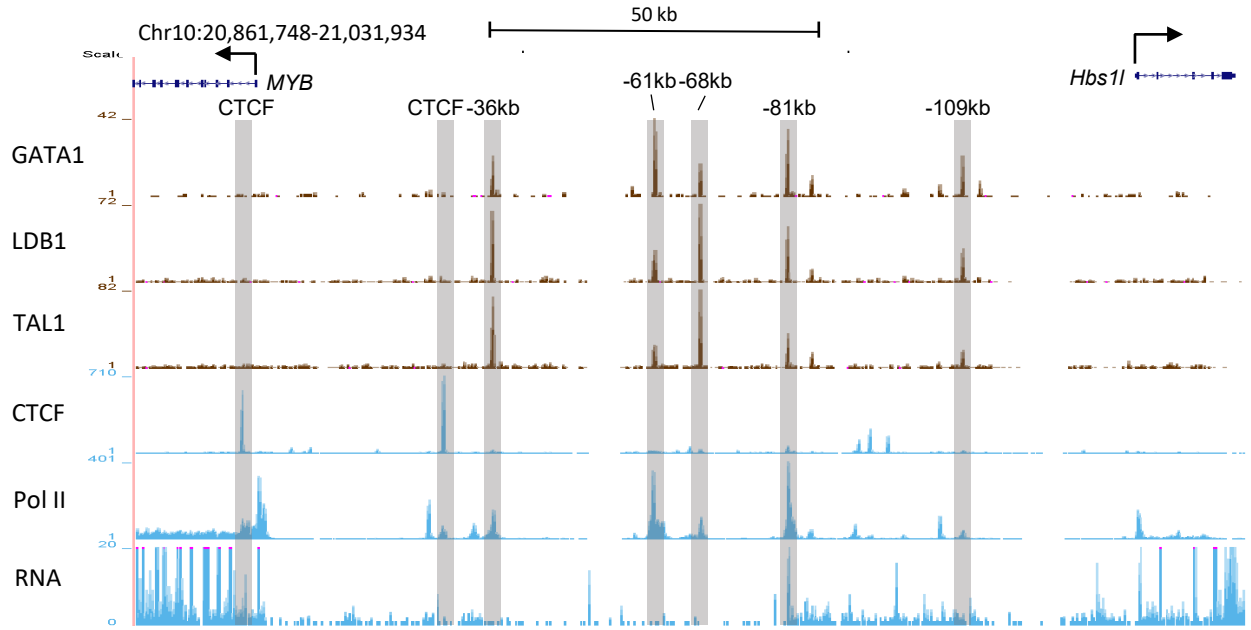
negative control. Error bars indicate SEM of 3 independent biological experiments. (*) $P < 0.05$ and (**) $P<0.01$ by Student's t-test.

Figure 6. Pol II and CDK9 occupancy within the *Myb* enhancer hub is affected by CRISPRi targeting of *Myrlin*. (A) ChIP-qPCR for CDK9 in *Myrlin* CRISPRi uninduced MEL cells targeted with dCas9/KRAB or without an sgRNA (control). (B) RNA Pol II occupancy in the *Myb* promoter/exon 1 region after *Myrlin* CRISPRi and in control cells. (C) CUT&Tag for RNA Pol II Ser5 and Ser2 phosphorylated forms. (D, E) Quantitation of CUT&TAG data showing Pol II Ser5P and Ser2P occupancy in the *Myb* gene body and at each of the enhancers. (F) Separate analysis of the shorter (<120 bp) and longer (>270 bp) Pol II Ser5P CUT & TAG fragments displaying Pol II Ser5 occupancy (Miura and Chen, 2020). (G) Pausing index calculated for Pol II Ser5 across *Myb*. Error bars indicate SEM of 3 independent biological experiments. (*) $P < 0.05$ and (**) $P<0.01$ by Student's t-test.

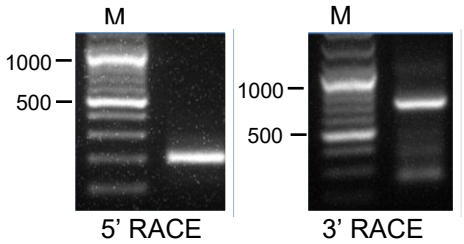
Figure 7. **KLF1 interaction with *Myrlin* contributes to localization within the *Myb* enhancer hub**. (A) *Myrlin* RNA pull down was conducted to determine efficiency of probes for ChIRP. (B) ChIRP DNA pull down by *Myrlin* across *Myb* and the *Myb* enhancers. (C) RNA ChIP for KLF1. (D) Biotinylated *Myrlin* pull-down and blotting with KLF1 antibodies. (F) Model of the *Myb* locus enhancer hub. Transcription of *Myb* is depicted with and without the -81 kb enhancer *Myrlin* eRNA after *Myrlin* CRISPRi. Large, shaded circle represents Pol II and LDB1 transcription factor density within the *Myb* enhancer hub which is diminished when *Myrlin* transcription is reduced. Error bars indicate SEM of 3 independent biological experiments. (*) $P < 0.05$ and (**) $P<0.01$ by Student's t-test.

Figure 1

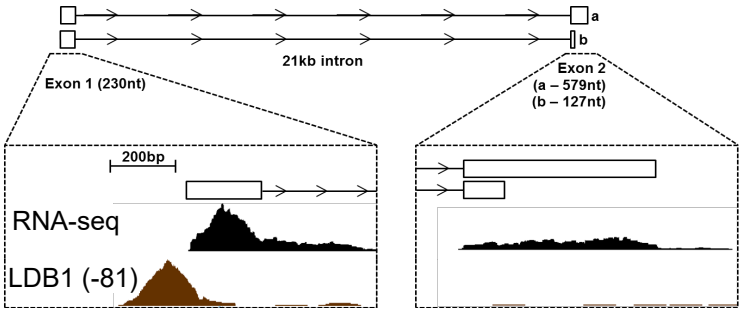
A



B



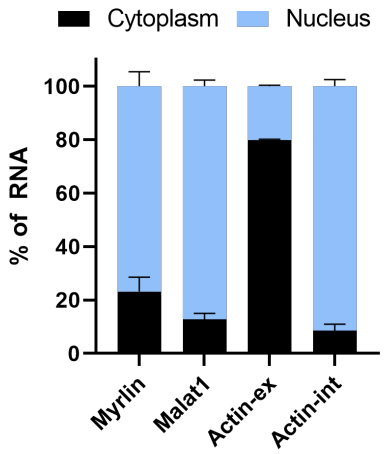
C



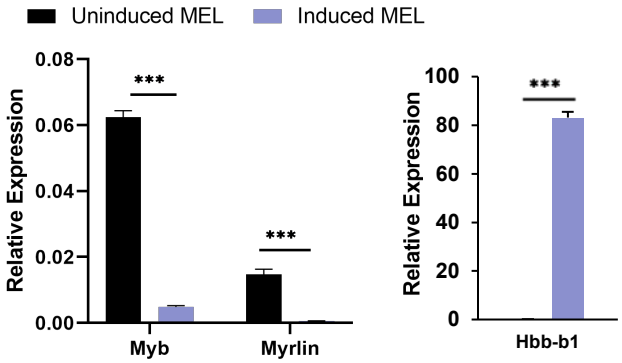
D

Sequence name	Coding probability	Coding label
c-Myb	1.000	yes
Hbs1l	1.000	yes
Ahi1	1.000	yes
-81kb	0.017	no
Lockd	0.108	no
Malat1	0.056	no

E



F



G

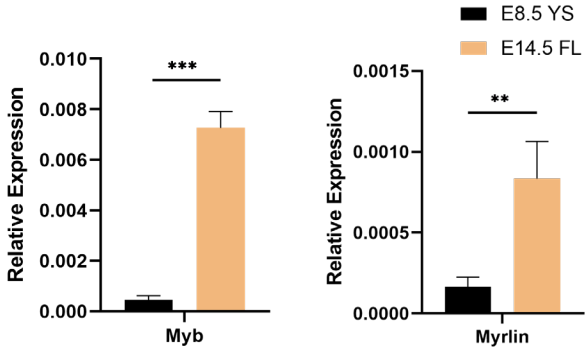


Figure 2

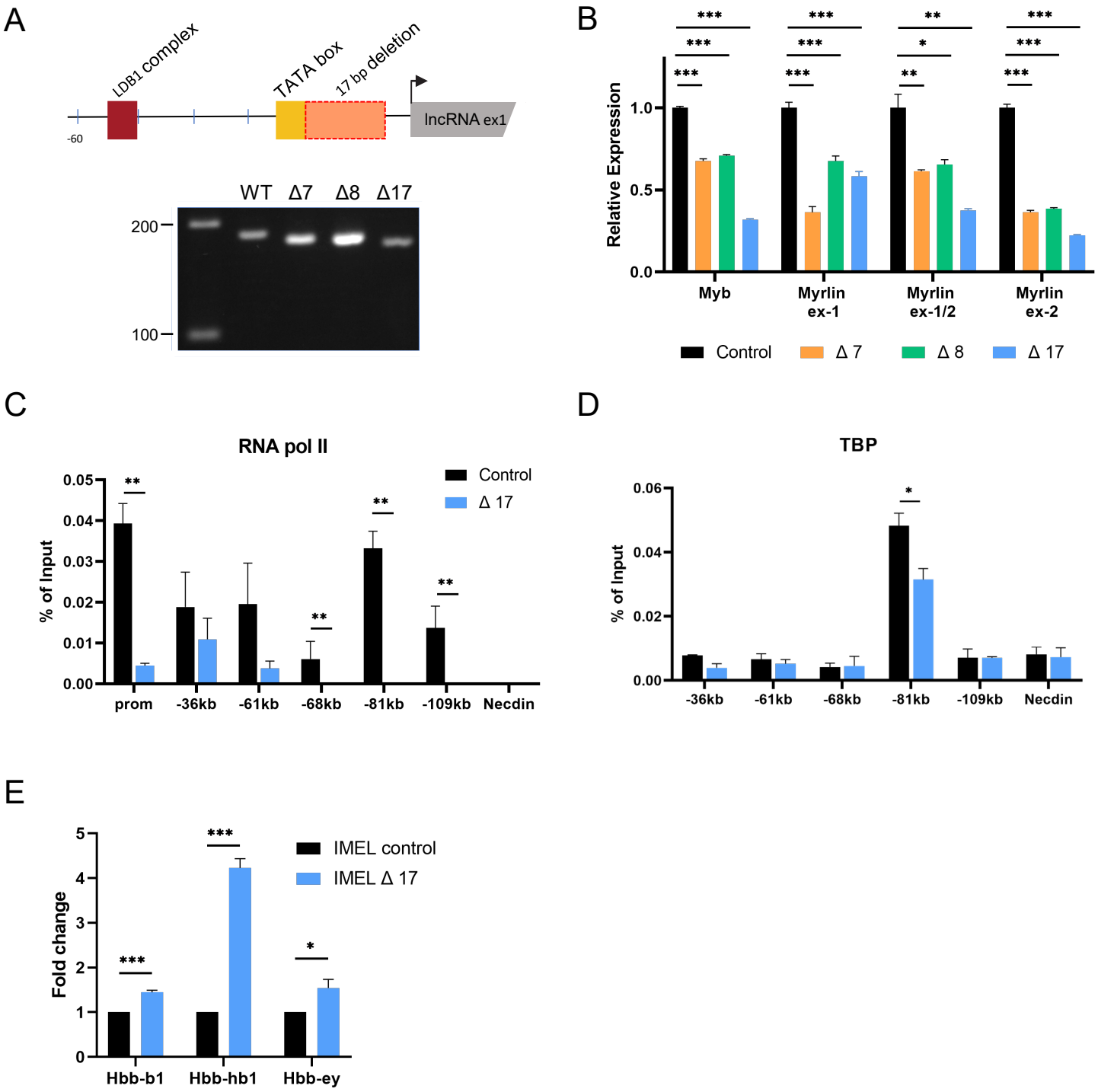

A
LDB1 complex
TATA box
17 bp deletion
lncRNA ex1
-60
WT Δ7 Δ8 Δ17
200
100
B
Relative Expression
1.0
0.5
0.0
Myb
Myrlin ex-1
Myrlin ex-1/2
Myrlin ex-2
Control
Δ 7
Δ 8
Δ 17
C
RNA pol II
% of Input
Control
Δ 17
prom
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
D
TBP
% of Input
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
E
Fold change
IMEL control
IMEL Δ 17
Hbb-b1
Hbb-hb1
Hbb-ey

Figure 3

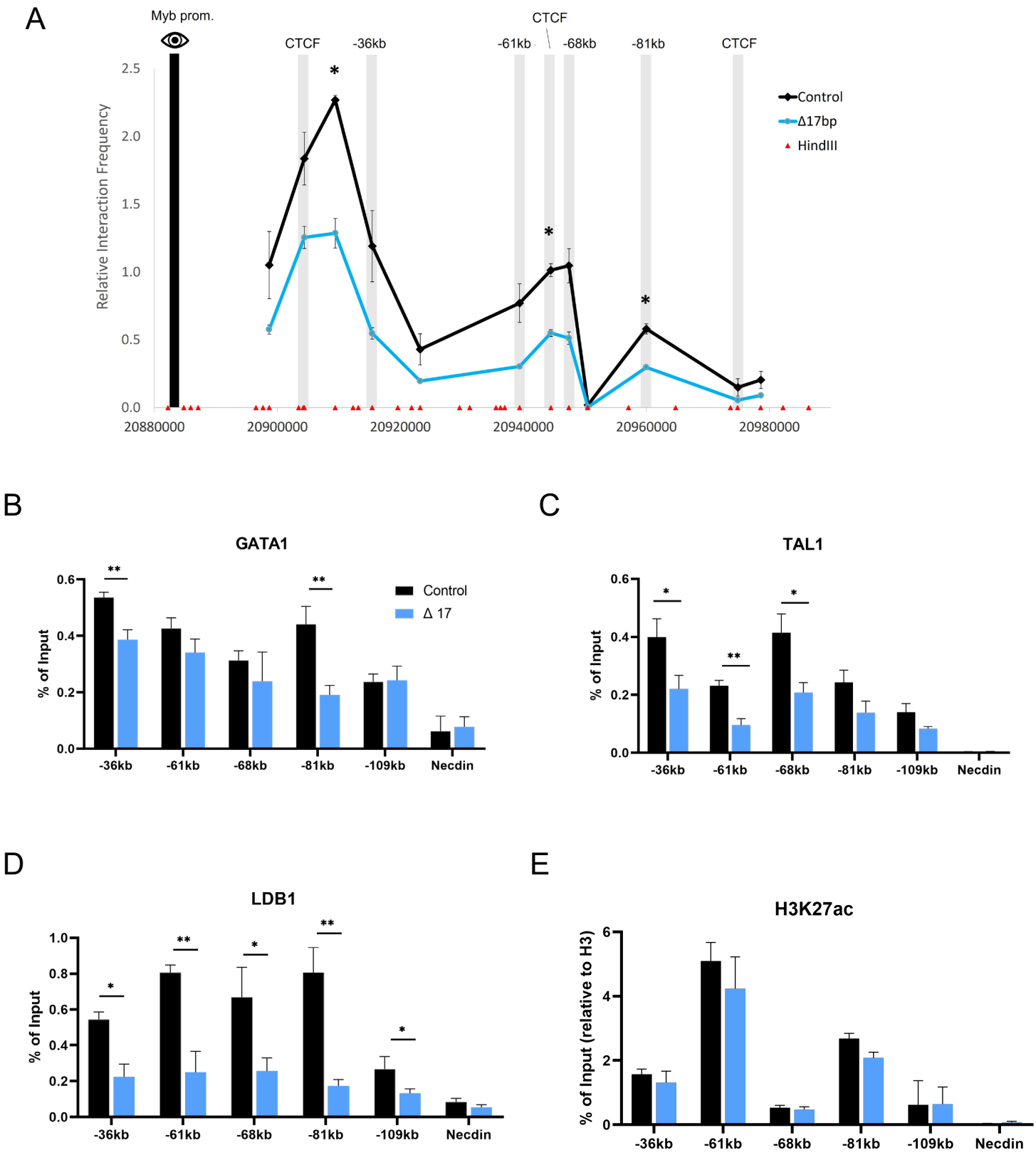

A
Myb prom.
CTCF
-36kb
-61kb
CTCF
-68kb
-81kb
CTCF
Relative Interaction Frequency
2.5
2.0
1.5
1.0
0.5
0.0
20880000
20900000
20920000
20940000
20960000
20980000
Control
Δ17bp
HindIII
B
GATA1
% of Input
Control
Δ 17
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
C
TAL1
% of Input
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
D
LDB1
% of Input
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
E
H3K27ac
% of Input (relative to H3)
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin

Figure 4

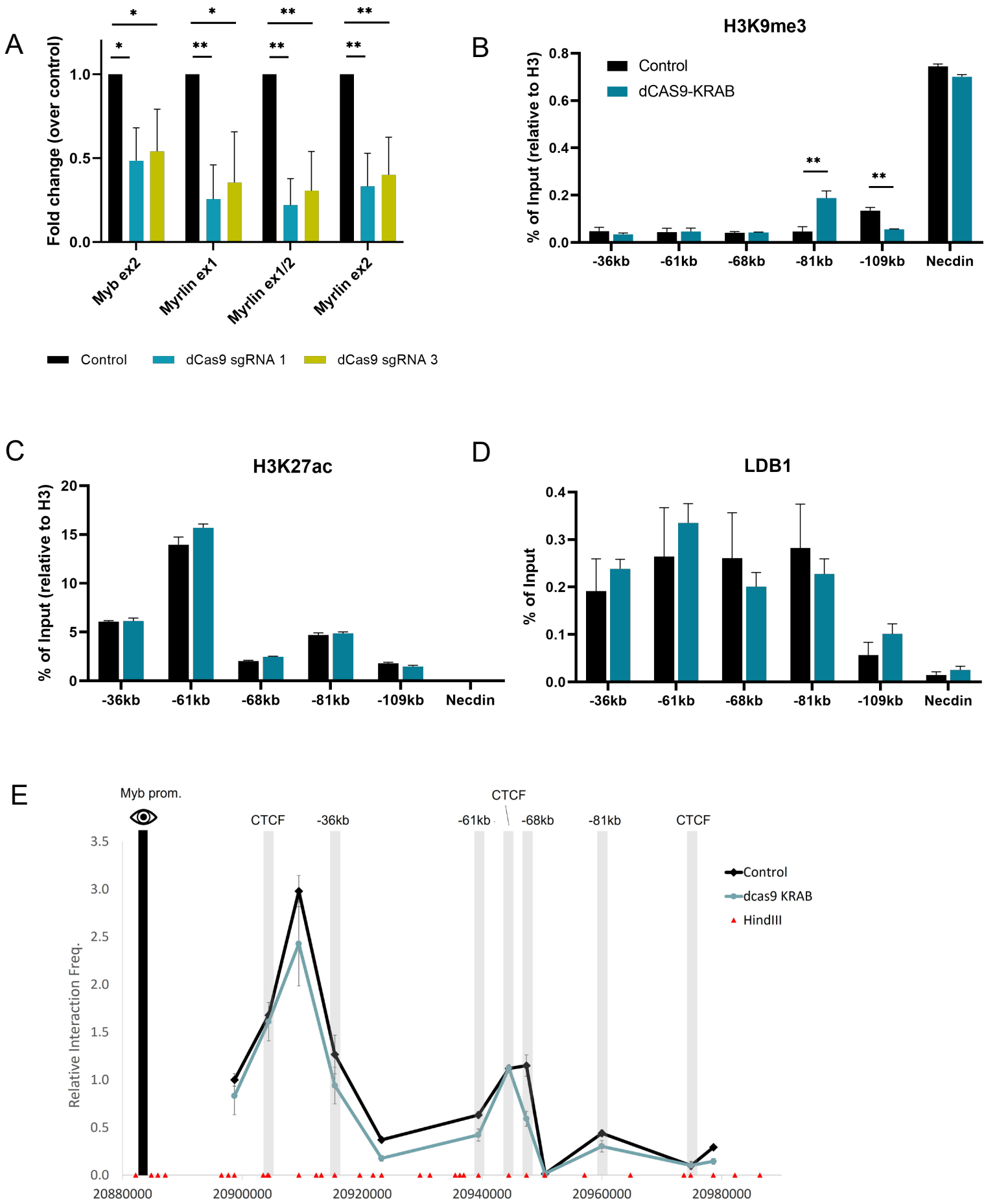
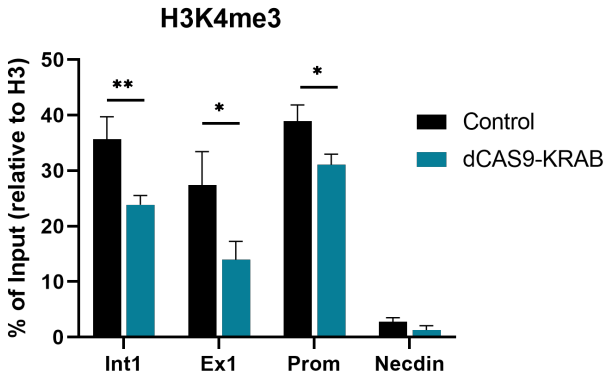
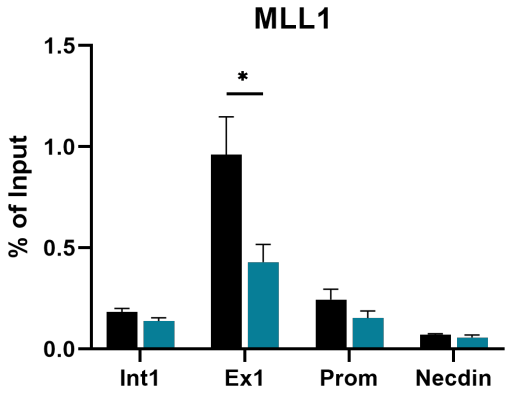

A
Fold change (over control)
1.0
0.5
0.0
*
*
*
**
**
**
**
**
Myb ex2
Myrlin ex1
Myrlin ex1/2
Myrlin ex2
Control
dCas9 sgRNA 1
dCas9 sgRNA 3
B
H3K9me3
% of Input (relative to H3)
0.8
0.6
0.4
0.2
0.0
Control
dCAS9-KRAB
**
**
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
C
H3K27ac
% of Input (relative to H3)
20
15
10
5
0
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
D
LDB1
% of Input
0.4
0.3
0.2
0.1
0.0
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
E
Myb prom.
CTCF
-36kb
CTCF
-61kb
-68kb
-81kb
CTCF
Relative Interaction Freq.
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Control
dcas9 KRAB
HindIII
20880000
20900000
20920000
20940000
20960000
20980000

Figure 5

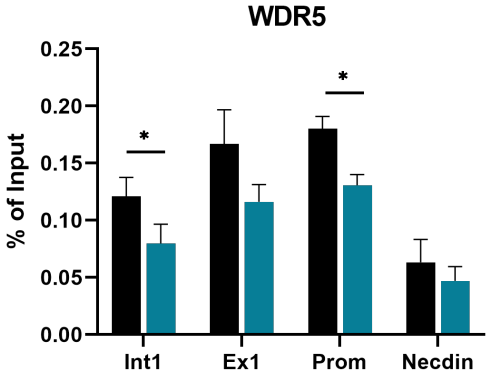
A



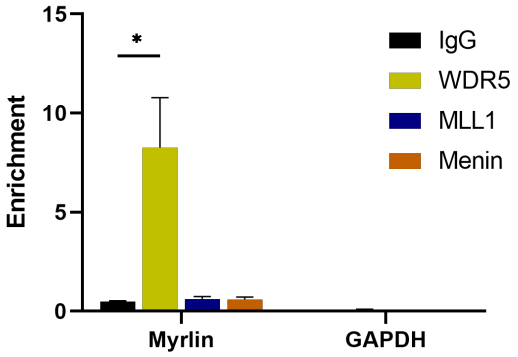
B



C



D



E

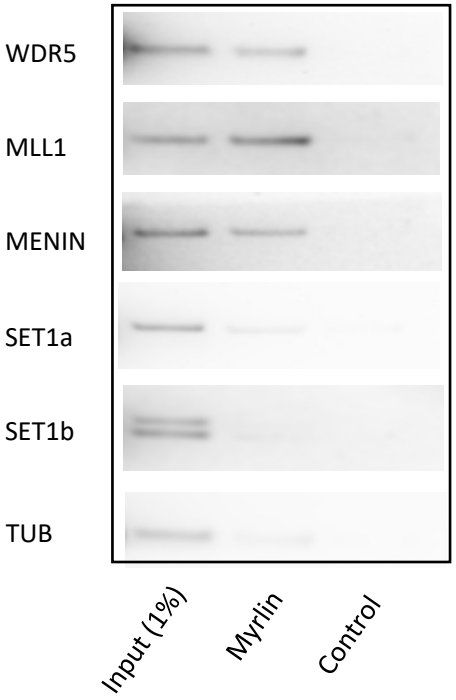


Figure 6

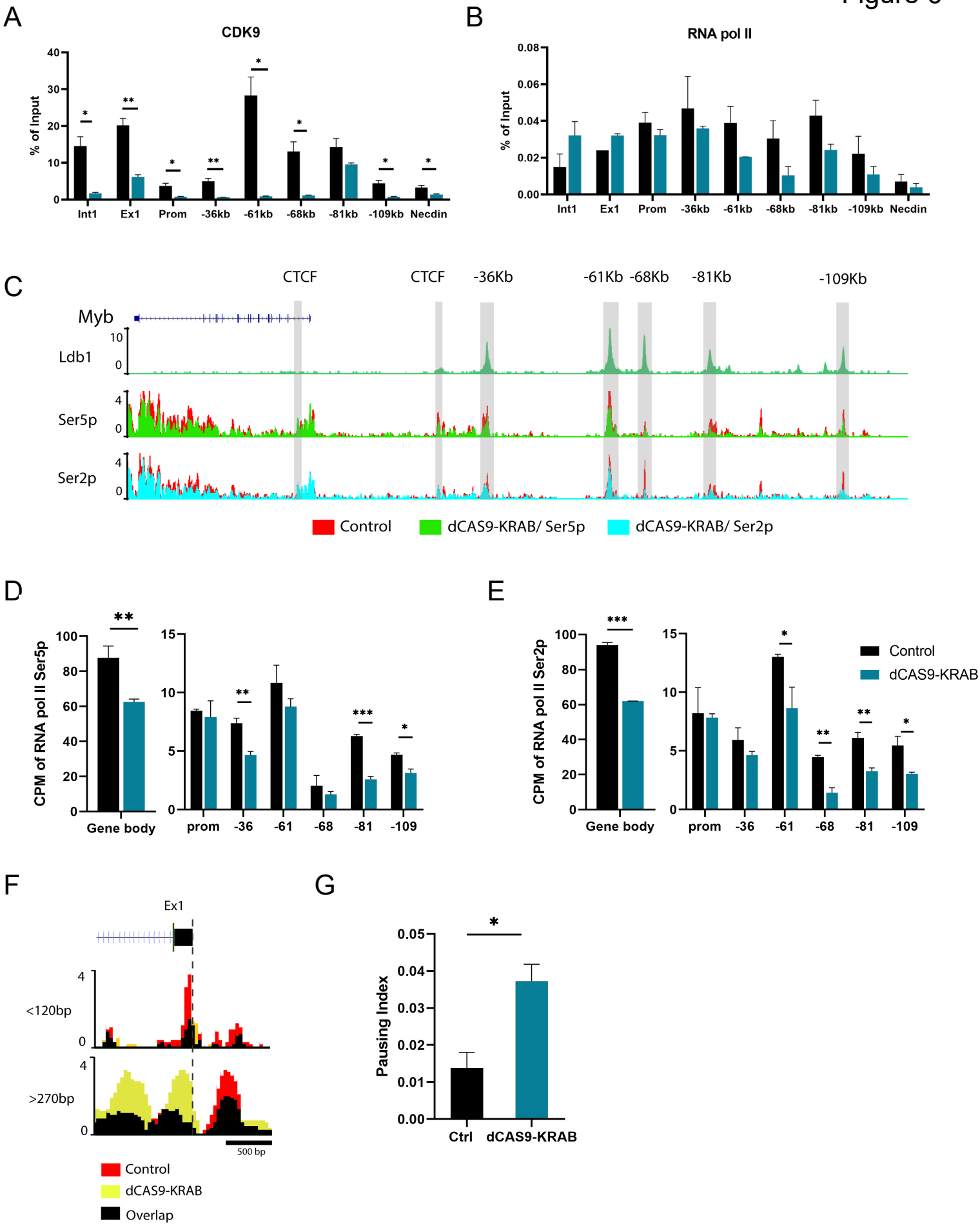
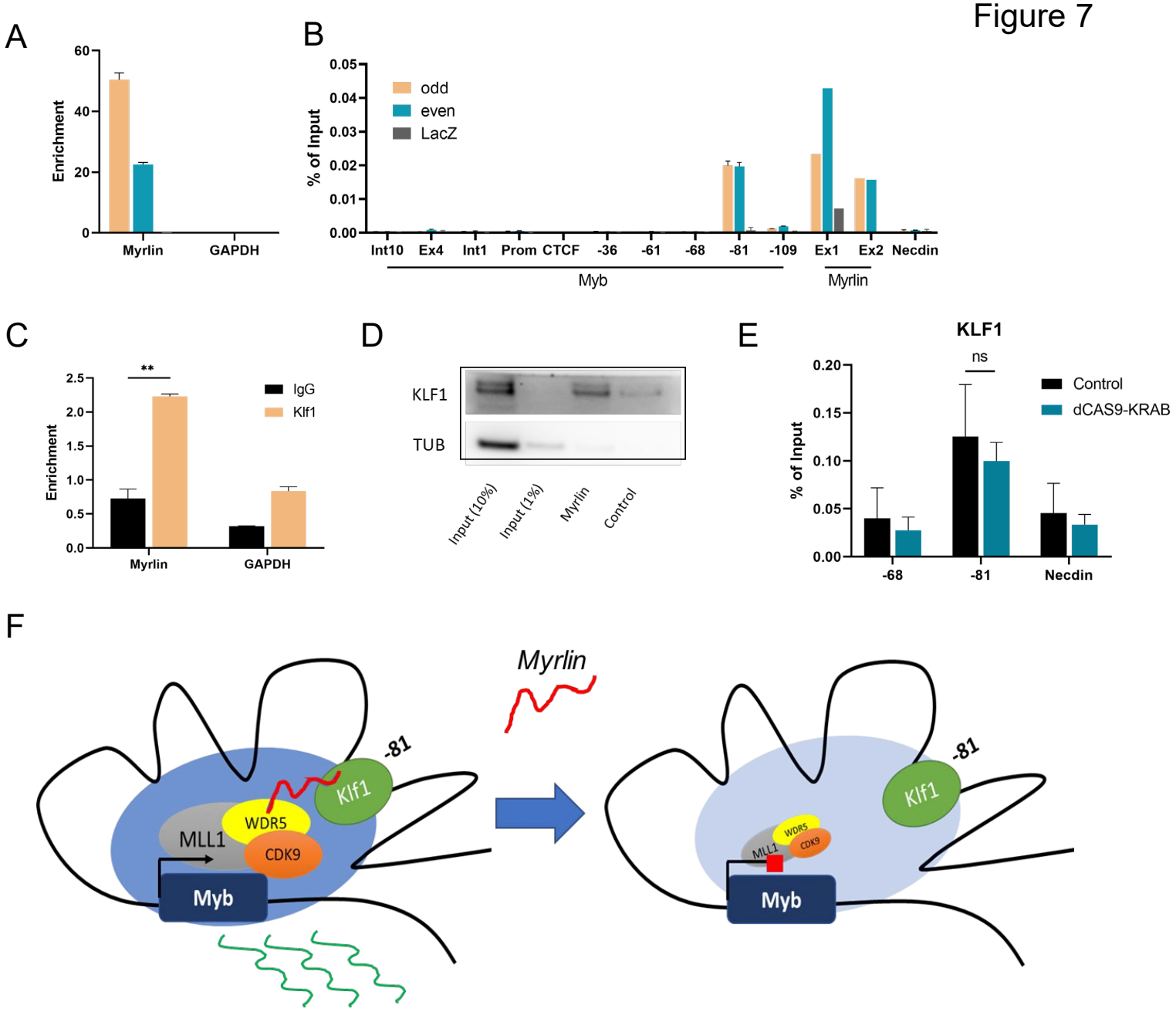

A
CDK9
% of Input
Int1
Ex1
Prom
-36kb
-61kb
-68kb
-81kb
-109kb
Necdin
B
RNA pol II
% of Input
C
CTCF
CTCF
-36Kb
-61Kb
-68Kb
-81Kb
-109Kb
Myb
Ldb1
Ser5p
Ser2p
Control
dCAS9-KRAB/ Ser5p
dCAS9-KRAB/ Ser2p
D
CPM of RNA pol II Ser5p
Gene body
prom
-36
-61
-68
-81
-109
E
CPM of RNA pol II Ser2p
Control
dCAS9-KRAB
F
Ex1
<120bp
>270bp
500 bp
Control
dCAS9-KRAB
Overlap
G
Pausing Index
Ctrl
dCAS9-KRAB

Figure 7


A
Enrichment
60
40
20
0
Myrlin
GAPDH
B
% of Input
0.05
0.04
0.03
0.02
0.01
0.00
odd
even
LacZ
Int10
Ex4
Int1
Prom
CTCF
-36
-61
-68
-81
-109
Ex1
Ex2
Necdin
Myb
Myrlin
C
Enrichment
2.5
2.0
1.5
1.0
0.5
0.0
**
IgG
Klf1
Myrlin
GAPDH
D
KLF1
TUB
Input (10%)
Input (1%)
Myrlin
Control
E
KLF1
% of Input
0.20
0.15
0.10
0.05
0.00
ns
Control
dCAS9-KRAB
-68
-81
Necdin
F
Myrlin
-81
Klf1
WDR5
MLL1
CDK9
Myb
-81
Klf1
WDR5
MLL1
CDK9
Myb