The Texas Medical Center Library
DigitalCommons@TMC

Dissertations and Theses (Open Access)

MD Anderson UTHealth Houston Graduate School

8-2023

Acute Manipulation Of Erna Level For Dissecting Its Roles In Transcriptional Regulation

Lanxin Bei

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations

 Part of the Cell Biology Commons, and the Molecular Biology Commons

Recommended Citation

Bei, Lanxin, "Acute Manipulation Of Erna Level For Dissecting Its Roles In Transcriptional Regulation" (2023). *Dissertations and Theses (Open Access)*. 1293.
https://digitalcommons.library.tmc.edu/utgsbs_dissertations/1293

This Thesis (MS) is brought to you for free and open access by the MD Anderson UTHealth Houston Graduate School at DigitalCommons@TMC. It has been accepted for inclusion in Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digcommons@library.tmc.edu.

Acute manipulation of eRNA level for dissecting its roles in transcriptional regulation

By:

Lanxin Bei

APPROVED:

wenbo li

Wenbo Li, Ph.D.

Advisory Professor

Xiaodong Cheng, Ph.D.

Blaine Bartholomew, Ph.D.

Nidhi Sahni, Ph.D.

Ambro van Hoof, Ph.D.

APPROVED:

Dean, The University of Texas

MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

Acute manipulation of eRNA level for dissecting its roles in transcriptional regulation

A

Thesis

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth Houston

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

by

Lanxin Bei

Houston, Texas

August 2023

Acknowledgements

I would like to express my deep gratitude to Dr. Wenbo Li, who has played an exceptional role as my mentor throughout my research career. His guidance and encouragement have been invaluable in shaping my understanding of the molecular processes of transcription and the role of RNA as a transcriptional regulator. Without Dr. Li's unwavering support, I would not have been able to successfully complete my thesis project.

I would also like to extend my thanks to all the members of my lab who have provided their support during my time there. Dexin has been instrumental in assisting me with NGS data analysis and sharing essential bioinformatic skills. Phouc has been a great teacher, helping me navigate the basics of laboratory work when I first started my research journey in the United States. Joo has provided invaluable insights into basic epigenetics and shared his research experiences with me. Ruoyu has offered me tremendous help in both scientific matters and personal life, providing valuable perspectives. Xiaoyu has been a reliable partner in troubleshooting perplexing experimental results. Lana and Yi-ting have always been there to uplift my spirits and provide constant encouragement.

I would also like to express my sincere gratitude to all my committee members for their insightful advice on my projects and their meticulous review of my thesis.

Furthermore, I would like to extend my appreciation to my friends both in the United States and back home. As an international student venturing into a foreign city and facing a completely different culture, starting a new life was a mixture of excitement and uncertainty. I consider myself fortunate to have had a group of friends who supported each other, whether in science or in life, during my time at GSBS. I am also grateful to my friends and former lab members in China for their encouragement and spiritual support. It means a lot to me that we have stayed in touch, continuously exchanging new thoughts about life, science, self-reflection, and interesting facts.

I am deeply thankful to everyone who has played a part in shaping my research journey and personal growth.

Abstract

Acute manipulation of eRNA level for dissecting its roles in transcriptional regulation

Lanxin Bei

Advisory Professor: Wenbo Li, Ph.D.

Enhancers are the central genetic elements controlling cell-type and state specific transcription programs to dictate cell fates during development. Mechanistic understanding of enhancer action is important for both biology and disease research. In the human genome, more than 60k human enhancers were found to produce non-coding transcripts named enhancer RNAs (eRNAs). These created a new challenge to understand enhancer functions, which now are not only DNA elements that promote transcription but also RNA-producing transcription units themselves. Importantly, deregulation of eRNAs was associated with various diseases such as cancer, immune disorders, and neurodegeneration. However, the direct role of eRNAs in transcriptional regulation and enhancer-promoter looping remains debatable, and in cases that eRNAs may bear functions the mechanisms are incompletely understood. One important problem is that currently used perturbation methods of enhancer/eRNAs are not acute and cannot distinguish direct from indirect effects.

To reveal the dynamics of eRNA-regulated target gene transcription and epigenetic features, I attempted to establish acute and controllable systems to manipulate eRNA levels with high temporal precision. Three perturbation methods were tested on a testbed eRNA, including dCas9 based CRISPR activation or inhibition system and antisense oligonucleotide (ASO) based RNA inhibition system. I found that these CRISPR based systems show variable effects and sub-optimal acute perturbation of target eRNA, while ASOs are more acute for target perturbation. My study benchmarked several methods and established an acute perturbation system for controlling eRNAs as well as other types of RNAs, which could potentially overcome the current confounding issues of potential secondary effects. These results provided a foundation for further studies of fine-scale dynamics of eRNA-mediated transcriptional regulation and chromatin organization.

Table of Contents

List of Illustrations

List of Tables

List of Abbreviations

eRNA: enhancer RNA

KRAB: Krüppel associated box

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

gRNA: guide RNA, mostly stands for single guide RNA

PAM: protospacer adjacent motif

NHEJ: non-homologous end joining

DSB: double strand break

dCas9: endonuclease deficient Cas9

TF: transcription factor

CRISPR-SAM: CRISPR-based synergistic activation mediator

RISC complex: RNA-induced silencing complex

TSS: transcription start site

tetO: tetracycline operator

TetR: tetracycline repressor, a transcription repressor that bind to tetO

tTA: tetracycline-controlled trans-activator, consists of TetR + VP16

rtTA: reverse tetracycline-controlled trans-activator, consists of rTetR + VP16

m6A: N6 methyl adenosine

YTHDC1: YTH domain-containing protein 1

FBS: Fetal Bovine Serum

PBS: Phosphate Buffered Saline

DMEM: Dulbecco's Modified Eagle Medium

SDS-PAGE: Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis

TRE: tet-responsible element

RT-qPCR: Reverse transcription-quantitative polymerase chain reaction

ASO: antisense oligonucleotide

LNA: locked nucleic acid

ChIP: chromatin immunoprecipitation

TT-seq: transient transcriptome sequencing

ATAC-seq: assay for transposase-accessible chromatin with sequencing

H3K27ac: acetylation of the lysine residue at position 27 in the N-terminal region of the histone H3 protein

H3K4me1: mono-methylation at the 4th lysine residue of the histone H3 protein

RNA Pol II/Pol II: RNA polymerase II

PFA: paraformaldehyde

CHAPTER 1: Introduction

1.1 The role of enhancer in transcriptional regulation and other biology process

Enhancers are DNA regulatory element that attenuate the transcription of their distal target genes[1]. In the commonly agreed model, enhancers regulate their target genes by recruiting transcription factors (TFs)/co-factors and looping to the promoters in an orientation and distance independent manner[1–3]. Enhancers are marked by several features including high accessibility, specific DNA sequences for transcription factor binding, enrichment of H3K27ac and H3K4me1 histone modifications, and binding of transcription co-factors[4–6]. Enhancer-mediated transcriptional regulation determines development and cell-fate decisions in metazoan, cellular responses to stimuli, and its dysregulation contributes to various diseases[7]. For example, during the evolution of snake, a 17bp DNA deletion occurred in the enhancer regulating the Shh gene, which determines the development of limbs, and the deletion of this enhancer led to the evolutionary regression of limbs in snakes[8]. For another example, in humans, a variant in the intronic enhancer of RET proto-oncogene leads to significantly higher risk for Hirschsprung's disease[9]. Enhancers are also responsible for organizing synchronized and acute response to key physiological signaling stimulation including Notch[10], retinoic acid (RA)[11], and estrogen[12] signaling.

1.2 Existing knowledge of enhancer derived RNA (eRNA)

The activation of an enhancer involves a coordinated process involving DNA sequences, transcription factors, co-factors, histone modifications, and enhancer-derived RNAs (eRNAs)[13–15]. Enhancer activation is considered to be initiated by the binding of pioneer transcription factors (like FOXA, Gal4, and GATA1) during lineage determination or early development[16,17]. Pioneer factors then recruit lineage specific TFs and co-factors like histone methyltransferases, histone acetyltransferases like CBP/P300[17]. These co-factors help deposit active histone markers like H3K4me1/2, H3K27ac, that further recruit epigenetic "readers" or activation co-factors like BRD4[13]. These steps concomitantly or subsequently prepare the enhancer for further recruitment of RNA pol2 and the bi-directional transcription of eRNAs. This

cis-activation complex, likely as an entirety, plays a key role to regulate target genes transcription by interacting with the gene promoter through looping[3,18] or by the formation of transcriptional condensate[19,20].

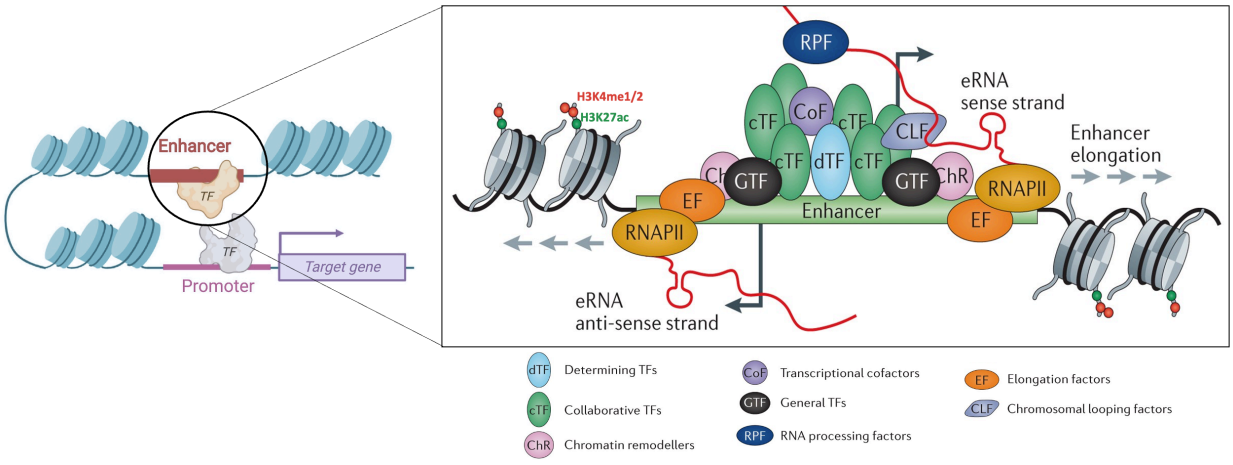

Figure 1.1: Features of an active enhancer including transcription of enhancer RNA (eRNA). Illustration partially is generated by BioRender, and partially cited from [7,13].

Ever since the first genome-wide identification of eRNAs in 2010[14,21], debates exist as to whether eRNA transcripts are functional transcripts, or simply byproducts of the transcriptional machinery. Over the past decade, evidence supported that at least a subset of eRNAs can be functional either by the action of RNA polymerase transcription process on enhancers or by the eRNA transcripts[13]. However, due to the abundant number of eRNAs genome-widely, it remains not clear what portion of eRNAs have biological functions, how they function, and if any functions are directly or indirectly achieved. Globally, the transcription level of eRNA is correlated with enhancer activity[14,22]. Mechanistically, eRNAs are reported to regulate several stages of transcription, including the recruitment of TFs and coregulators[23,24], RNA Pol II pause-release and elongation[25–27], and enhancer-promoter looping[13,18,28,29]. For example, in TLR4 signaling, the deposition of H3K4me1/2 at signal-induced enhancers was largely dependent of the transcription process of eRNA[30]. At the enhancers of some immediate early genes upon signal induction in neurons, the eRNA transcripts could act as decoy for negative elongation factor (NELF) and facilitate the elongation of gene transcription[25,26]. The

binding of nascent eRNA transcripts to histone acetyltransferase CBP could stimulate its enzymatic activity and transcription of target genes[31]. eRNAs are globally activated by various cellular signaling[14,18,32,33], supporting their roles in synchronizing rapid transcriptional changes upon stimulation. Importantly, deregulation of eRNAs was associated with diseases including cancer[34], immune disorders, and neurodegeneration[35,36]. As an example, the multivalent proto-oncogene MYC is regulated by a long-range interacting eRNA named CCAT1 (colon cancer-associated transcript 1), whose deregulation leads to various cancer types[37–39].

1.3 limitations of current efforts/methods in eRNA studies

Despite progress in the past few years, the direct role of eRNAs still remains debatable. One major problem is that currently used perturbation methods of enhancer/eRNAs are not acute and cannot distinguish direct or indirect effects upon perturbation. On one hand, identification of direct target genes is hard to achieve by long-term perturbation. Many eRNA directly regulate critical transcription factors like MYC[37], Sox2, and KLF4[40]. In such cases, a long-term perturbation may result in waves of regulation event and alter the transcription program of hundreds of genes, confounding the identification of direct targets. On the other hand, even between eRNA and its direct target gene, the regulation evolves a series of biological events including eRNA transcription, target gene transcription, the binding of transfections factors/co-factors, deposition of histone marks, and the alteration of enhancer-promoter looping[41,42]. Long-term perturbation can't resolve the dynamics and causal relationship of these biological processes. It is still under debate whether H3K4me1/3, or enhanced enhancer-promoter looping is required or a result of target gene transcription. Understanding the immediate and causal step of eRNA mediated regulation is critical for understanding the fine-scale transcription process and may potentially facilitate therapeutics.

Current CRISPR/dCas9-based perturbation can achieve highly efficient activation and inhibition by targeting the endonuclease deactivated Cas9 (dCas9) fused with activation[43] or inhibition[44–46] domain to transcription start sites. But the stable cell line with constitutively expressed dCas9 and gRNA could only confer long-term perturbation. For overexpression,

vector-based transgene expression could achieve high induction level, but the ectopically-expressed transgene might not fully recapitulate the authentic function of eRNAs, which mostly functions within the territory of their original transcribed sites. For inhibition, siRNA-based RNA knockdown is efficient for most genes and could be done by transient transfection, but the activity of RISC complex in the nucleus was unclear. The knockdown efficiency of eRNA in a short time was seldom reach 90% in previous publications[47,48].

An acute and controllable system to manipulate eRNAs is critical for clarifying the direct roles of eRNAs in transcriptional regulation. In this study, I aim to establish acute perturbation systems for upregulating and inhibiting eRNA levels. With this system, I will further investigate how acute eRNA manipulation impacts target gene transcription and epigenetic features, as well as the dynamics and causal-consequence relationships of these features. Establishing a previously unavailable acute perturbation system for controlling eRNAs will provide a useful tool to study eRNA functions by overcoming the current confounding issues of potential secondary effects. This system could also be applied to a broad range of other RNA transcripts to facilitate their functional studies.

CHAPTER 2: Exploring CRISPR perturbation for rapid eRNA manipulation

2.1 Background

2.1.1 Overview of the CRISPR system

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) has become the most used genetic editing tool since its first application in mammalian cells. It is originally the component of a bacterial immune system defending against virus DNA invasion. When the invading virus sequence enters the bacteria genome, CRISPR RNA (crRNA) that is transcribed from the “spacer” sequence obtained from previous invaders matches with the inserting sequence. The crRNA assembles with Cas9 enzyme produced by the CRISPR system and can therefore bind and recognize DNA sequence of the invading virus[49]. The Cas9 enzyme further induce double strand break by its HNH and RuvC enzymic domain. Since early 2012[50–52], the CRISPR system has become a powerful tool for genetic engineering and has been

continuously improved to edit DNA sequence as well as chemically modify DNA and histone (i.e., the epigenome)[53–56]. The CRISPR tool for genetic editing composes of a single-guide RNA (sgRNA) that could binds to the targeting DNA region by base-pairing, and with the help of gRNA, the Cas9 protein can cut specific DNA adjacent to a PAM trinucleotide. This will include double strand DNA break (DSB) to the targeted DNA, which can then be repaired by NHEJ (non-homologous end joining) or by homologous recombination (if a repair donor can be provided). This feature has been taken advantage to edit genomic DNA to cause point mutations, frame shift mutations or to create specific genetic changes needed for specific scientific purposes.

2.1.2 CRISPR/dCas9-based epigenome editing

Modified from the CRISPR/Cas9 system, the CRISPR/dCas9 system has been newly developed for epigenome editing[44–46,57]. dCas9 stands for enzyme defective Cas9 that bears two point-mutations at the HNH and RuvC domains, which abolish its endonuclease activity. This dCas9 can bind specific DNA regions of interests with high precision and specificity with the help of gRNA without causing DNA DSB. For epigenome editing, dCas9 is normally fused with effector domains like transcriptional activators or inhibitors, including VP64[43,57], P300[58], and KRAB[46]. Which can be “piggybacked” by the dCas9/gRNA to the genetic elements of interests to modify the epigenetic state.

Based on different effector domain and co-factor applied, various versions of CRISPR/dCas9 system have been developed. For CRISPRa, the SAM system overall exhibits best performance[43,59]. The SAM system consists of three components: a dCas9 fused with VP64 activation domain modified from virus VP16 activator; a gRNA scaffold contains 2 virus stem-loop sequences that specifically binds to the MS2 coat protein; and the co-activators HSF1 and P65 fused with MS2 protein that could assembly with the gRNA scaffold. The SAM complex can be directed to the transcription start site (TSS) of a target locus, and VP64 and MS2-HSF1-P65 can then activate target gene transcription by recruiting transcription factors and co-factors including TFIIA, TFIIB, TFIID, P300, Mediator, SWI/SNF[60]. For CRISPRi, the

dCas9-KRAB based inhibition has been widely applied in single locus studies[61] and large-scale screening[62]. The KRAB domain is a family of transcriptional inhibition domains in many zinc finger-based transcription factors and most of the current CRISPRi system uses the KRAB domain of Kox1[63]. The KRAB domain itself does not have enzyme activity but functions through recruiting other suppressors, prominently TRIM28[64]. TRIM28 inhibits transcription through recruiting inhibitory proteins, including the chromosome remodeler component CHD3 and H3K9 methyltransferase SETDB1, which induce H3K9me2/3 that induces heterochromatinization[64].

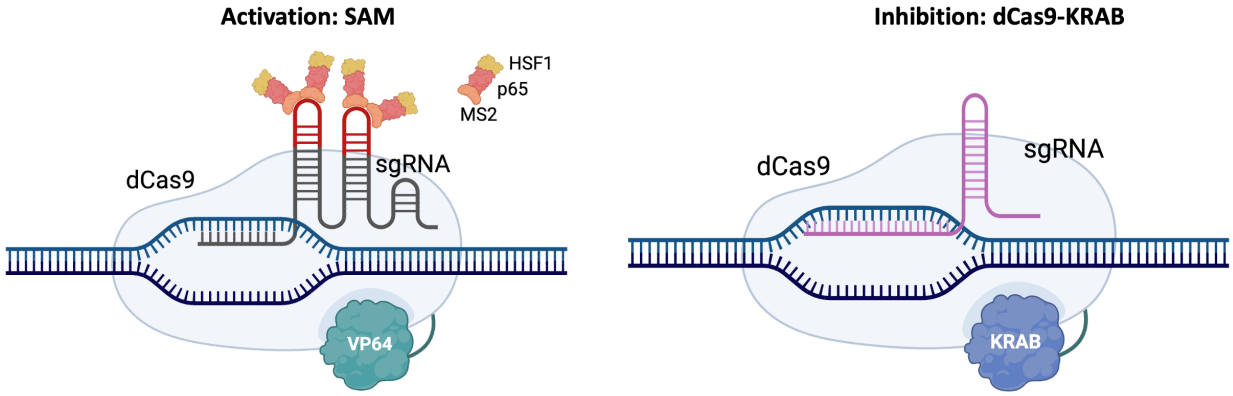

Figure 2.1: Model of CRISPR-SAM system for transcription activation and CRISPR/dCas9-KRAB system for transcriptional inhibition. Illustration is generated by BioRender.

2.1.3 Tet-on system for inducible dCas9 expression

For long-term perturbation over 10 days, the current CRISPR-based systems exhibit competent activation and inhibition effects for functional study[65]. However, acute perturbation by CRISPR/dCas9 system has not yet been widely reported. To achieve acute perturbation, one possible method is to control the expression of the dCas9 protein. For controlling the on-and-off of protein expression, the most applied system is Tet-on system modified from bacteria[66]. In the rtTA-based Tet-on system[67], the protein of interest is driven by a minimal CMV promoter whose upstream enhancer was removed to achieve low basal expression. Tet-responsible element (TRE) is placed upstream of the promoter. Another essential component is rtTA protein consisting of a rTetR domain that binds to the TRE only in the presence of tetracycline or doxycycline, and three tandem repeats of viral protein VP16 that activate the

transcription of the target when the rtTA bind to TRE. Therefore, doxycycline us a transcription inducer. This Tet-on system has been applied for controllable expression of various genes[66] and could be potentially used for expressing dCas9 in CRISPRa/i to study the direct function of enhancer RNA.

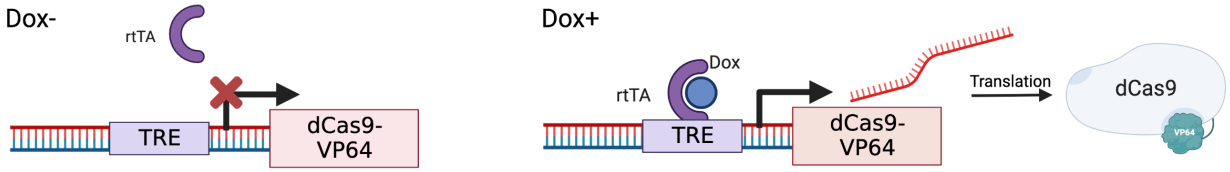

Figure 2.2: Model of Dox-inducible dCas9 expression through the rtTA-based Tet-on system. Illustration is generated by BioRender.

2.2 Materials and methods

2.2.1 cell culture

The MCF-7 cells were grown in DMEM High Glucose medium (Corning, 10-013-CM) supplemented with 10% FBS (GenDepot cat. F0900-050) in a cell culture incubator set at 37 degrees Celsius with 5% CO2.

Similarly, the HEK293T cells used for virus packaging were cultured in DMEM High Glucose medium supplemented with 10% FBS (GenDepot, cat. F0900-050) under the same conditions.

The cell line containing doxycycline-inducible dCas9-VP64/KRAB elements was cultured in DMEM supplemented with 10% Tet-free FBS (GenDepot, cat. F0500-050). To induce expression, doxycycline (MilliporeSigma, D9891-5G) was added at a concentration of 2 ng/μL.

2.2.2 Lentivirus packaging and infection

To produce lentivirus, HEK293T cells were transfected with three plasmids. Plates were coated with 1 mL of 10 ng/mL Poly-D-lysine (Gibco, A3890401) and incubated for 30 minutes, followed by a wash with PBS. Cells were detached using Trypsin (Sigma, T4049) and seeded at 30% confluency with 800 μL of medium per well. Solution 1 (1 μg psPAX2, 0.33 μg pMD2.G,

1.33 µg target plasmid) was mixed with 100 µL of Opti-MEM (Gibco, 31985070). Solution 2 (5 µL lipofectamine 2000 reagent, 100 µL Opti-MEM) was prepared. Both solutions were incubated for 5 minutes at room temperature, then combined and incubated for 20 minutes. Next, 200 µL of the mixture was added to the cells, followed by incubation at 37°C, 5% CO2 for 12-16 hours. After 15 hours, the medium was replaced with 10% FBS DMEM and cells were incubated for 24 hours. Lentiviral particles were harvested from the medium, filtered through a 0.45 µm filter, and immediately used or stored at -80°C.

To perform lentivirus infection, MCF-7 cells were detached using Trypsin and seeded in 6-well plates at 30% confluency with 500 µL of medium per well. Then, 1000 µL of virus containing medium was added to the suspended cells, and polybrene (Sigma, #H9268) was included at a final concentration of 8 µg/mL. After incubating for 24 hours, the media was replaced with a DMEM with 10% FBS containing antibiotics for selection.

2.2.3 Establishment of cell lines

Addgene #61425, #61426[57] plasmids were used for parental cell line of constitutive SAM system; Addgene #50916[68], #61426 plasmids were used for parental cell line of dox-inducible SAM system; Addgene #50917[68], were used for parental cell line of dox-inducible CRISPR/dCas9-KRAB system; Addgene #61427[57] were used for gRNA expression in all cell lines.

To isolate individual clones of the SAM parental cell lines, cells were seeded in 15 cm round dishes at a total cell number of 5000. The cells were cultured for 10 days, and single clones were carefully selected using 200 µL tips under a microscope. These single clones were then cultured in 12-well plates to assess the growth and expression level of dCas9-VP64 through Western Blot analysis.

Next, different gRNA-encoding viruses (using plasmid Addgene #61427) were introduced into the parental cell lines. For the inducible and constitutive SAM systems, single clones of the parental cell lines were used, whereas the bulk parental cell line was employed for the inducible dCas9-KRAB system. RNA samples were collected 14 days after Zeocin

selection, and RT-qPCR was performed to evaluate the constitutive or doxycycline-induced perturbation effect of each gRNA.

Regarding antibiotic selection, puromycin (Gibco, A1113802) was used for selection over a period of 1 week at a concentration of 1 μg/mL. Zeocin (InvivoGen, ant-zn-1p), Blasticidin (InvivoGen, ant-bl-10p), Hygromycin (Gold Biotechnology, H-270-1), and Geneticin (Gold Biotechnology, G-418-1) selections were performed for a minimum of 2 weeks, with concentrations of 200 μg/mL, 10 μg/mL, 400 μg/mL, and 1000 μg/mL, respectively.

2.2.4 Western Blot

Cells were washed with cold PBS and lysed using 1x Laemmli Buffer (Bio-rad, Cat#1610737). The lysate was heated at 95°C for 10 minutes to denature the proteins. Next, the protein samples were loaded onto an SDS-PAGE gel and electrophoresed in SDS tris-glycine running buffer (Bio-rad, 1610772EDU). The proteins were transferred to a 0.45μm PVDF membrane (Bio-rad, Cat#11620177) using wet transfer. The membrane was blocked with 5% milk for 1 hour. Then, it was incubated overnight at 4°C with primary antibodies: Cas9 (Cas9 ThermoFisher, 1:1000, cat. 8C1-F10, mouse) and GAPDH (Proteintech, 1:2000, cat. 60001-1-lg, mouse). After washing with TBST three times for 10 minutes each, the membrane was incubated with a secondary antibody (Jackson ImmunoRearch, 115-035-166) for 30 minutes at room temperature. Following three additional TBST washes, the membrane was exposed to an ECL Substrate (Bio-Rad, 170-5060) for chemiluminescent detection. The signal was captured and analyzed using an Imaging System (Bio-Rad).

2.2.5 gRNA design for CRISPRa and CRISPRi

gRNA were designed using the CRISPOR webportal (http://crispor.tefor.net/). For CRISPR activation, target sequence was -500bp-0bp to the transcription start site (TSS) of target eRNA. For CRISPR inhibition, target sequence was 0bp-400bp to the TSS[45]. The genome of Homo sapiens were selected for filtering the specificity of gRNA. NGG were selected as the Protospacer Adjacent Motif (PAM). lentiGuide-Puro (Zhang lab) were selected as Addgene plasmid for cloning.

gRNAs used in this study:

gRNA name	gRNA sequence
TFF1e gRNA_1_Fw	CACCGCCAGTGATCTGGCTCTGCGT
TFF1e gRNA_1_Rv	AAACACGCAGAGCCAGATCACTGGC
TFF1e gRNA_2_Fw	CACCGCCAGATCACTGGGTAAACAC
TFF1e gRNA_2_Rv	AAACGTGTTTACCCAGTGATCTGGC
TFF1e gRNA_3_Fw	CACCGCTGGCAACGACCTGTCCCAA
TFF1e gRNA_3_Rv	AAACTTGGGACAGGTCGTTGCCAGC
TFF1e gRNA 4_Fw	CACCGCTGGCAACGACCTGTCCCAA
TFF1e gRNA 4_Rv	AAACTTGGGACAGGTCGTTGCCAGC
TFF1e gRNA 5_Fw	CACCGCTCTGCGCGGGCTACCTGAC
TFF1e gRNA 5_Rv	AAACGTCAGGTAGCCCGCGCAGAGC
TFF1e gRNA 6_Fw	CACCGAACCACAGGGACGTGTACGG
TFF1e gRNA 6_Rv	AAACCCGTACACGTCCCTGTGGTTC
*TFF1e*_gRNA7_Fw	CACCGAAATCAAAGGGACGGCCGCG
*TFF1e*_gRNA7_Rv	AAACCGCGGCCGTCCCTTTGATTTC

Table 1: Primer sequences for gRNA targeting *TFF1e* for activation and inhibition.

2.2.6 cloning of gRNAs

5 µg of the Addgene 61427 plasmid, expressing MS2-gRNA, was linearized using BsmBI (NEB, R0134S) following the manufacturer's instructions. Enzyme cutting product were run on 1% agarose gel for purification. Bright positive band on the gel were cut and purified by gel extraction kit (Qiagen, 28706X4). The gRNA primers from Sigma were annealed and phosphorylated by incubating with PNK (NEB, M0201S) using the following program: 37°C for 30 minutes, 95°C for 5 minutes, and then gradually cooled to 25°C at a rate of 5°C per minute (0.1°C/s). For ligation, 50 ng of the linearized plasmid and the annealed oligo, with a 10x amount of plasmid, were mixed with 1 µL of 10x T4 buffer (NEB, B0202S), 0.5 µL of T4 DNA ligase (NEB, M0202S), and topped up to 10 µL with ddH2O. The ligation reaction was carried

out by incubating for 20 minutes at 16°C. Next, 10 µL of the ligation product was transformed into DH5-alpha competent *E. coli*. (NEB, C2987P). Bacterial clones were picked, and Sanger sequencing was performed using a specific primer (5'-tacaaaatacgtgacgtag-3') targeting the U6 promoter to verify the inserted sequence.

2.2.7 RNA Isolation, cDNA synthesis, and quantitative PCR (qPCR)

RNA isolation was performed using the Quick-RNA Miniprep kits (Zymo Research, Cat#R1055). For cDNA synthesis, the SuperScript IV First-Strand Synthesis kit (Invitrogen, Cat#18091050) was used, utilizing 500ng total RNA as the template. The generated cDNA was quantified by qPCR using the SYBR Green SSo Advanced reagent (Bio-rad, Cat# 1725274).

qPCR primers:

primer name	primer sequence
GAPDH_F	CCTGTTCGACAGTCAGCCGCATC
GAPDH_R	GGTGACCAGGCGCCCAATACG
*TFF1e*_F	ATCTGCTGCTGCTTCCACTT
*TFF1e*_R	ACATGACCCTGCAGACCTTC
*TFF1*_intron_F	TATCTGGATGGGCCTTGGGA
*TFF1*_intron_R	AGACCTTCGAGAAGTGCGAC
*TFF1*_mRNA_F	CACCATGGAGAACAAGGTGA
*TFF1*_mRNA_R	TGACACCAGGAAAACCACAA

Table 2: Primer sequences for RT-qPCR.

2.2.8 Statistical analysis

All bar graphs were presented as mean ± SD using R version 4.1.2.

2.3 Results

2.3.1 Model cell line and testbed eRNA

MCF-7 cell line was the model cell line in this study because of the accessibility of abundant transcription and epigenetic data.

As proof of concept, an enhancer RNA on chromosome 21 upstream of the *TFF1* protein coding gene was used as testbed RNA in this study and was named *TFF1e*. *TFF1e* is specifically expressed in a subset of breast cancer cell lines and is marked by typical enhancer features including H3K27ac, H3K4me1 histone modifications, the binding of transcription co-factors like BRD4 and P300[15], and the binding of cell-type specific transcription factor ER-alpha[18]. 4C data[35] has support a high contact frequency between the promoter region of *TFF1* gene and the *TFF1e*, indicating *TFF1e* directly regulates *TFF1* gene.

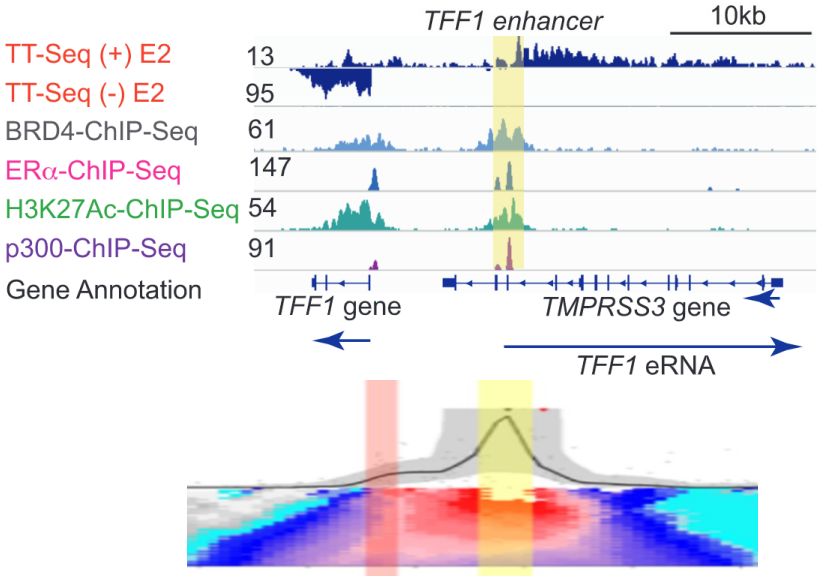

Figure 2.3: Genomic context of *TFF1e* and *TFF1* gene locus. Upper: IGV snapshot of TT-seq for profiling transient transcription[69] and ChIP-seq for profiling the binding of TFs and co-TFs, from[70]; lower: 4C data showing contact frequency of *TFF1* promoter and *TFF1e*, from[35]. Numbers on the track represent RPKM.

2.3.2 Establishment of parental cell lines for CRISPR perturbation

For all CRISPRa/i system in this study, stable parental cell lines were established first. Parental cell lines represent CRISPRi/a cell lines with all other components including the dCas9 protein and co-activators except the gRNA sequence. Once the parental cell line is established, perturbation of multiple transcripts could be done by only infecting the cell with different gRNA-encoding viruses instead of performing the infection of all CRISPR components every time.

SAM parental line with constitutively expressed dCas9-VP64 and MS2-P65-HSF1 was made for quick screening of effective gRNA. Parental cell line for dox-inducible dCas9-VP64/KRAB expression was also successfully established and the dox induction effect was validated. Dox treatment for 48hrs was sufficient for dCas9-VP64 protein to reach to a comparable level as constitutively expressed dCas9-VP64 (figure 2.4). Either longer treatment or higher dox concentration did not further increase the protein level (figure 2.4). Dox wash-off for 3 days was sufficient to tune down the protein level (figure 2.4), showing the capability of the tet-on system to turn on and shut down protein expression. Protein level of dCas9-KRAB upon dox treatment was not as high as dCas9-VP64, potentially due to a difference in protein stability or lower binding affinity of Cas9 antibody to dCas9-VP64. mRNA level of dCas9-KRAB upon dox induction was around half of that of dCas9-VP64 (figure 2.5) while the variance in their protein level by western-blot was more than 10-fold, indicating the transcriptional variance may not be the major cause of variance in protein level.

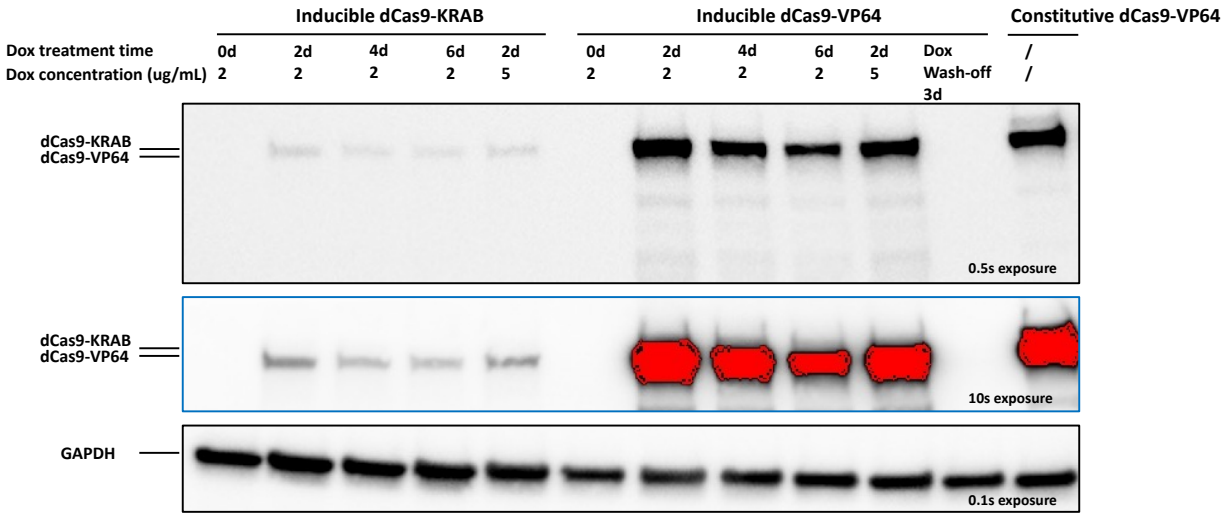

Figure 2.4: Western-blot showing the induced expression of dCas9-KRAB and dCas9-VP64, and constitutive expression of dCas9-VP64 in parental cell lines.

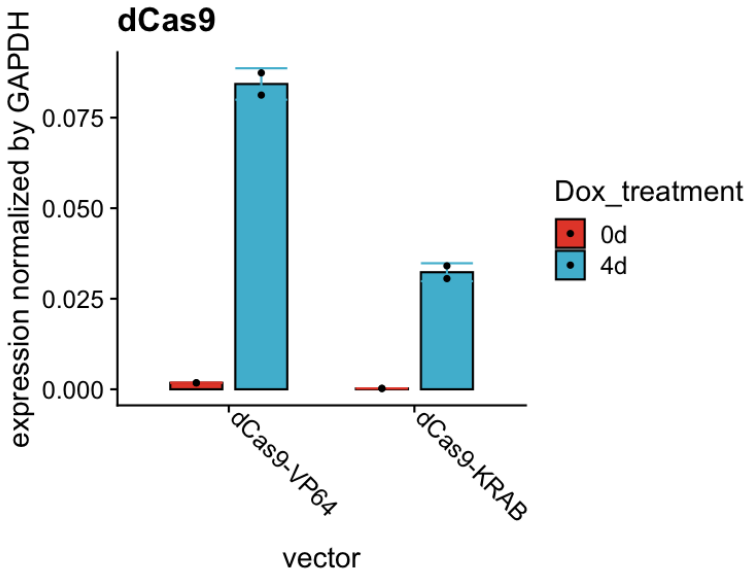

Figure 2.5: RT-qPCR analyzing the mRNA level of dCas9-VP64 and dCas9-KRAB upon 4 days of dox treatment. Each dot represents one technical replicate.

Parental cell lines post drug selection had the component for CRISPR perturbation integrated in their genome, but they are still bulk groups of heterogeneous cells with variations of dCas9 copy number, genome integration sites, so as dCas9 protein expression level. To get a more homogeneous system and potentially more repeatable results, I picked single clones for constitutive and inducible SAM parental cell lines. Clone 1 of constitutive SAM parental cell line exhibit higher basal dCas9-VP64 level, and clone 2 of inducible SAM parental line exhibit higher dCas9-VP64 level upon dox treatment. These two clones were used for the following experiments.

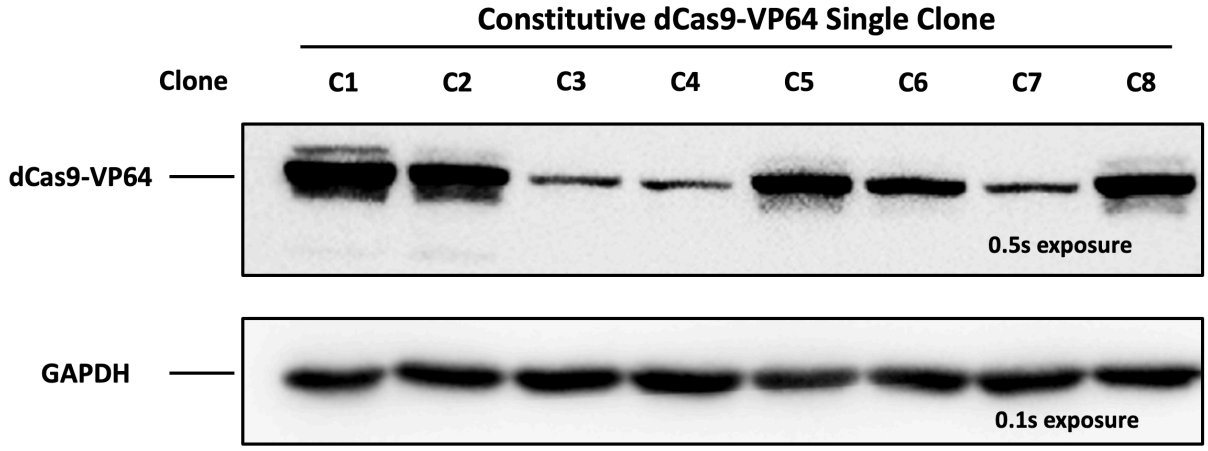

Figure 2.6: Western Blot showing the protein level of dCas9-VP64 in 8 clones of constitutive SAM parental line.

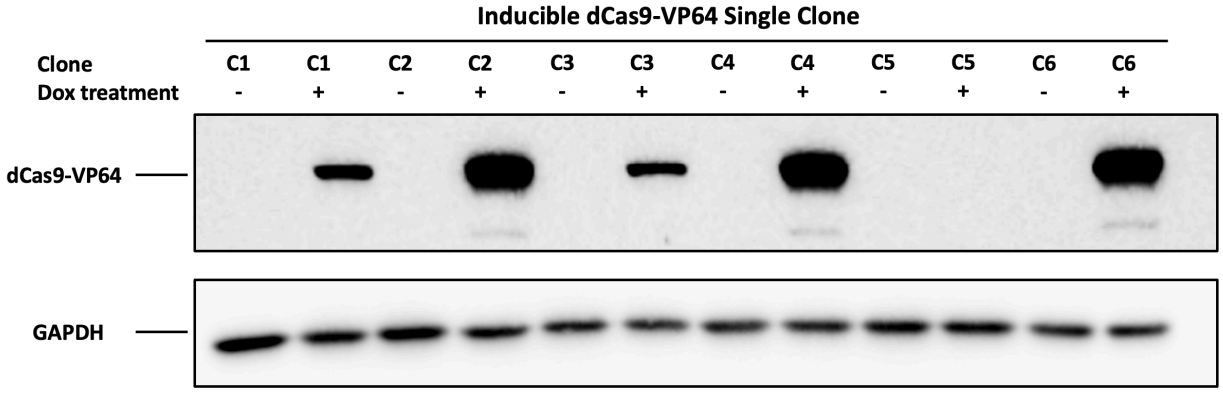

Figure 2.7: Western-blot showing the protein level of dCas9-VP64 in 6 clones of inducible SAM parental line.

2.3.3 design and cloning of gRNAs

7 gRNAs were designed targeting either upstream or downstream the TSS of *TFF1e*, which are proposed to be preferred gRNA position for activation and inhibition, respectively (figure 2.8). The 7 gRNAs were individually cloned into Addgene 61427 plasmid and confirmed by Sanger sequencing (figure 2.9).

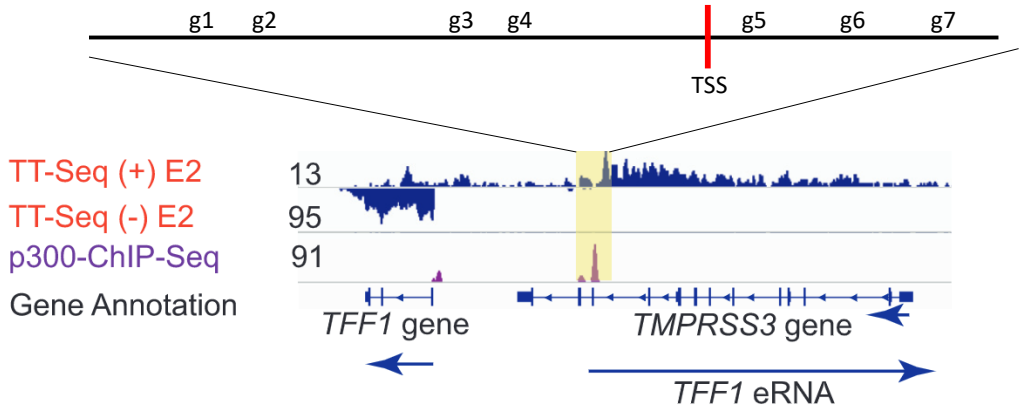

Figure 2.8: Position of 7 gRNAs near the TSS of *TFF1e*.

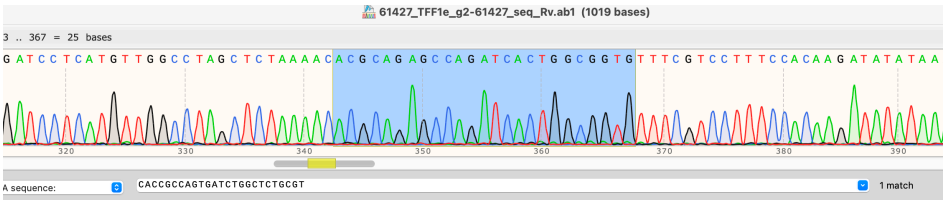

Figure 2.9: Representative snapshot of Sanger sequencing result showing successful insertion of *TFF1e* gRNA2 into the 61427 plasmids.

2.3.4 Activation and inhibition of *TFF1e* by CRISPR perturbation

Viruses encoding gRNA-1, -2, -3, and -7 were infected into constitutive SAM parental line for testing their activation effect. gRNA 2 and 3 that are located upstream of TSS exhibited >8-fold *TFF1e* induction, showing the high potential of SAM system in activating eRNA. gRNA2 that confers the best activation were used for dox-inducible SAM system for acute activation of *TFF1e*. To our disappointment, 48h of dox treatment only brought the *TFF1e* to 2.5-fold of its basal expression level, an induction rate insufficient for functional study.

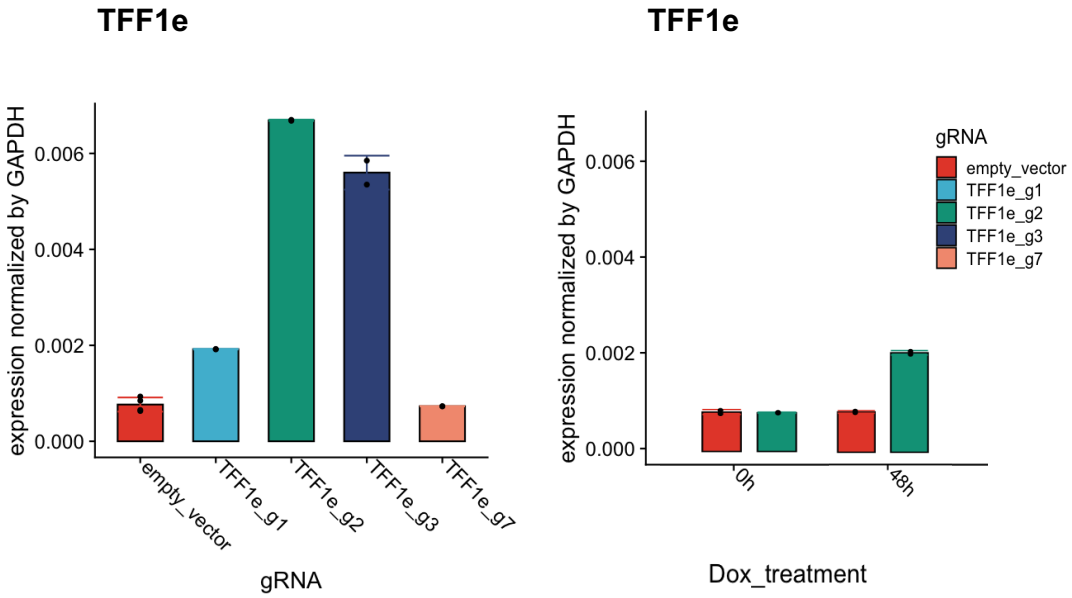

Figure 2.10: RT-qPCR analyzing activation effect of four gRNAs by the constitutive SAM system and activation effect of gRNA2 by the inducible SAM system. Parental cell line infected with virus encoding Addgene #61427 plasmid backbone served as negative control. Each dot represents one technical replicate.

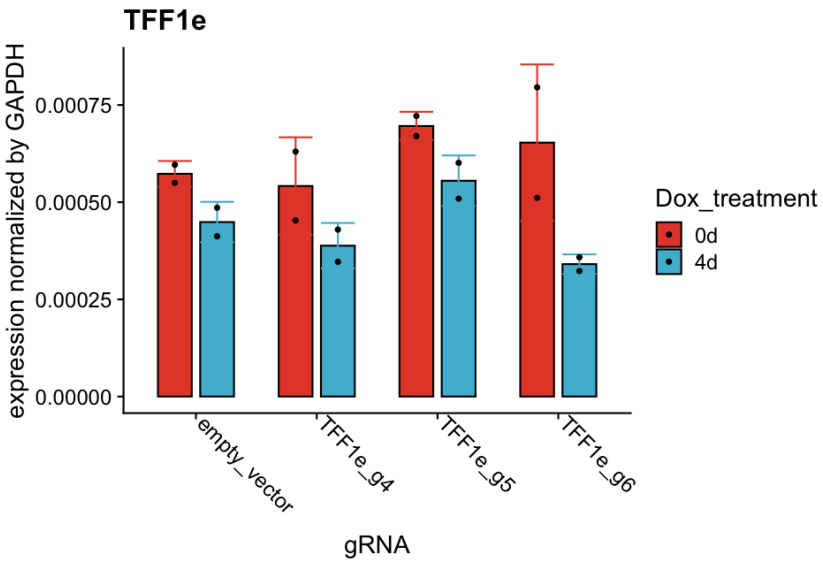

Figure 2.11: RT-qPCR analyzing inhibition effect of three gRNAs by the inducible dCas9-KRAB system. Parental cell line infected with Addgene 61427 plasmid backbone served as negative control. Each dot represents one technical replicate.

gRNA-4, -5, and -6 were infected into dox-inducible dCas9-KRAB parental line for testing their inhibition effect. All three gRNAs inhibit the transcription of *TFF1e* upon 72h of dox treatment, but even the most effective gRNA-5 only reduced the transcription to ~50%. Although dCas9-KRAB induction for longer time may achieve heterochromatinization and transcriptional inhibition to a greater extend, transcriptional inhibition for over 72 hours can certainly lead to waves of secondary effect. Therefore, CRISPR-inhibition was not regarded as an optimal choice for acute inhibition.

2.4. Discussion

CRISPR/dCas9-based perturbations are proved as powerful tools for controlling transcription activities by their applications in various single locus studies and high-content screenings[65,71]. Most of the studies achieved efficient target gene activation or inhibition by long-term perturbation of more than 10 days[65,71], a time frame that is too long to meet the demand in our study. My data have shown that 48h of activation and even 96h of inhibition did not achieve substantial perturbation at least for *TFF1e*, indicating CRISPR/dCas9-based perturbation might not be the ideal tool for studying the direct effect of eRNAs.

Nevertheless, the dox-inducible SAM activation system has achieved robust activation of another eRNA target within 24 hours by me, demonstrating its potential for single-locus study. The activation effect of inducible SAM has been shown to be largely dependent on genomic region, with some eRNA easier to activate than others. This limits the universal application of inducible CRISPRa to eRNAs. Delayed activation effect in the inducible SAM system might be caused by 1) prolonged dCas9-VP64 translation and translocation to the nucleus; 2) time-consuming steps for the establishment of active transcription state including recruitment of transcription factors and co-factors by VP64, and the deposition of activate histone marks.

One caveat of CRISPR perturbation that hinders their application in our study is potential off-target perturbation of spatially proximal genes. In a fine scale, the DNA is spatially organized into enhancer-promoter loops and other structural loops. The SAM complex or dCas9-KRAB being targeted to the eRNA TSS might be directly brought to the promoter region of the target gene by enhancer-promoter looping. The potential of dCas9 to directly perturb the target gene poses a challenge in distinguishing whether the altered level of enhancer RNA (eRNA) or the dCas9 complex itself exerts a direct influence on the transcription of the target gene. The dCas9 complex might also be spatially brought close to other genomic regions that loop to the enhancer and cause unwanted transcriptional changes to these regions.

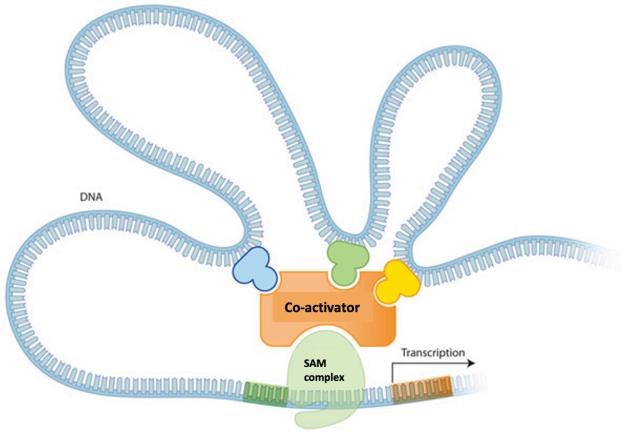

Figure 2.12: Model showing potential direct off-target transcriptional activation of spatially proximal genes. Modified from: Scitable by Nature Education, topic page: gene expression (https://www.nature.com/scitable/topicpage/gene-expression-14121669/).

Some alternative methods leveraging the CRISPR/dCas9 could be further tested. For reducing the response time of activation and inhibition, an inducible gRNA may be quicker because it wouldn't require time associated with dCas9 translation and nuclear import. For minimizing the effects on proximal genes caused by spatial interaction, fusing alternative effector protein to dCas9, like P300[58] and HDAC[72] that directly rewrite histone acetylation, may exert more specific perturbation than VP64 and KRAB that recruits a large group of transcription regulators. For CRISPR inhibition, specifically, targeting only the dCas9 downstream of the TSS could also sterically block the elongation of RNA pol II[46]. However, it has been demonstrated that this inhibitory effect is not as pronounced as that achieved with dCas9-KRAB[46]. Recently, CRISPR-based methods directly targeting the RNA transcripts has been published, utilizing CRISPR/Cas13[73–75] and CRISPR-Csm[76]. Directly targeting RNA exhibit higher potential of reaching acute eRNA depletion, but the potential collateral effect needs to be carefully addressed[73,77].

CHAPTER 3: Acute eRNA depletion by ASO mediated knockdown

3.1 Background

3.1.1 Mechanism of ASO action

Anti-sense oligonucleotide (ASOs) applied in biomedical research are synthetic single strand DNAs typically 14-24 nucleotides long[78]. ASOs can regulate RNA and protein level though several mechanisms. For transcriptional regulation, the most well studied mechanism leverages the property of RNaseH that recognizes DNA-RNA hybrid. Once an ASO binds with its target RNA by complementary base pairing, the DNA-RNA hybrid becomes a substrate for RNaseH that specifically cleaves the RNA strand. The cleavage products are then degraded by the RNA surveillance machinery in the nucleus and cytoplasm[79]. ASOs can also modulate RNA splicing by binding to splice sites to block certain splicing events[78]. In the protein translation

stage, ASO could be designed to target the AUG translation start site for blocking the binding of ribosome and suppressing protein translation[78].

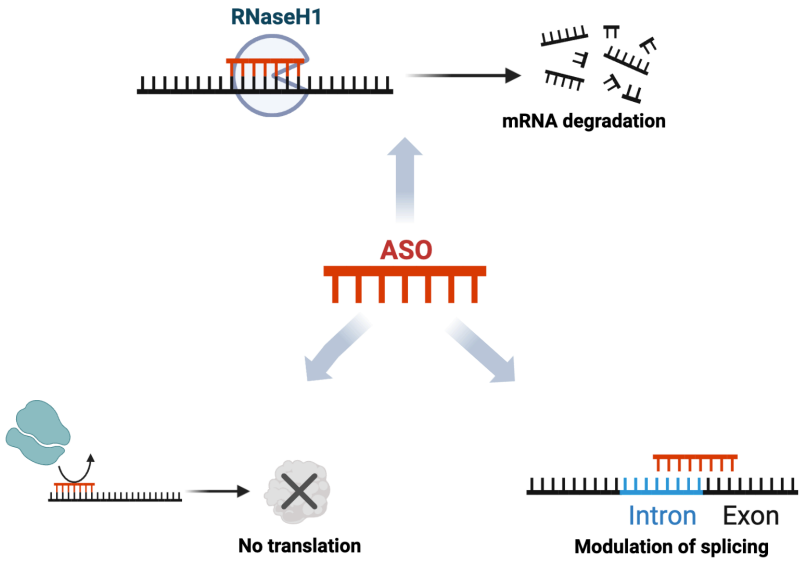

Figure 3.1: Three ways of how ASOs regulate RNA and protein levels. Illustration is generated by BioRender.

The native DNA ASO showed limited application potential due to rapid degradation of phosphoribose backbone in vivo and pervasive off-targeting effect[78]. Chemical modifications developed in the last few decades has greatly improved ASO properties. The application of phosphorothioate (PS) backbone with a sulfur atom replacing the non-bridging oxygen increase its nuclease resistance. The chimeric design of adding modified RNA, including 2′-O-methyl (2′-OMe), 2′-O-methoxyethyl (2′-MOE)[78], and locked nucleic acid, at both ends of DNA oligos has further enhanced stability and target affinity of the ASO. Locked nucleic acids (LNAs) are widely utilized modified RNA monomers that are commonly integrated into ASOs. “Locked” in the name comes from the methylene bridge between the 2′ oxygen and the 4′ carbon of the RNA ribose ring, which locks the ribose ring in a conformation that favors Watson-Crick binding, resulting in a high thermal stability when hybridized to target RNA. The incorporation of each LNA monomer increases the melting temperature (Tm) of the ASO-RNA duplex by 2-8C[80], making it possible to design shorter ASOs with a relative high Tm. As a result, LNA-

incorporated ASOs possess high target affinity and unprecedented mismatch discrimination ability, making it a currently optimal choice for targeted RNA degradation.

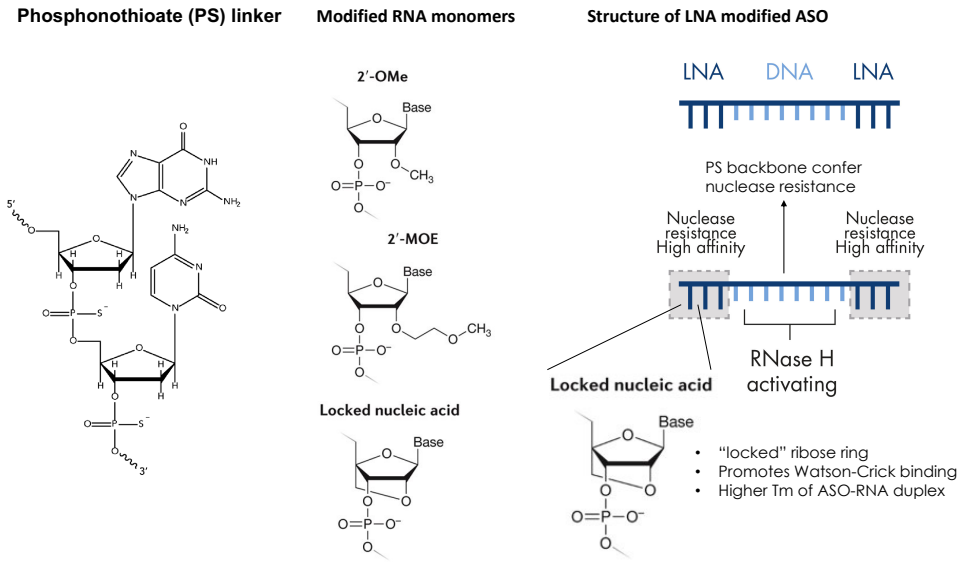

Figure 3.2: Chemical modifications to ASO and the detailed structure of LNA modified ASO. Cited and modified from Qiagen Antisense LNA GapmeRs Handbook and[78].

3.1.2 Some potential caveats of ASO mediated knockdown

For mechanistic study and especially acute perturbation of enhancer RNA, some potential caveat of ASO should be carefully considered. Firstly, cytotoxicity should be closely monitored because the PS modified backbones show extensive non-specific protein binding that might cause toxicity, especially at high concentrations[81]. Secondly, potential off targeting might still knockdown unspecific transcript. A 16nt ASO with 3 mismatched nucleotides still has the potential of causing a 50% inhibition of RNA (lab unpublished data by me). Therefore, ASO specificity should be checked genome-widely by BLAST. Thirdly, ASO close to the 5' end of a transcript was reported to potentially trigger pre-mature transcription termination[82], which might confound the result if one wants to distinguish the inhibition of RNA transcript or the process of transcription. In such cases, pre-mature termination needs to be monitored.

3.2 Materials and methods

3.2.1 ASO design and synthesis

The sequence of the first 2803nt of *TFF1e* transcript were input into web server of Stellaris FISH probe designer (https://www.biosearchtech.com/stellaris-designer) for ASO design. Parameters were set as follow: organism: Human; masking level: 5; max. number of probes: 96, oligo length (18-22nt): 18; min. spacing length (nt): 0-50. Minimal spacing length were set with different value to generate more available probes. 16nt is the preferred length for LNA-modified ASO but the minimum oligo length option on the web server was 18. Therefore, three 16-nt sequences were derived from the 18-nt sequence by picking the left, middle, and right 16-nt. BLAST was performed for all the 16nt probes with the hg38 reference transcriptome as reference. Sequences with reverse-complementary match to untargeted transcripts equal to or over 13 nt were regarded with high off-targeting potential and were discarded. ASOs were synthesized by IDT (Integrated DNA Technologies) with full phosphorothioate modification in the backbone and LNA modification at three nucleotides on both the 3' end and 5' end.

LNA name	LNA sequence
SCR_LNA	AACACGTCTATACGC
*TFF1e*_LNA1	GCAAAAGGTCCGACGA
*TFF1e*_LNA2	CACCAATCTACGCTCC
*TFF1e*_LNA3	GCTAACCCGGATGCTT
*TFF1e*_LNA4	AACGCTGGGAGATCTT

Table 3: Sequence of LNA-modified ASOs targeting *TFF1e*.

3.2.2 ASO transfection

To prepare mixture 1, 2 µL or 4 µL of 20 uM ASO was combined with 100 µL of Opti-MEM and incubated at room temperature for 5 minutes. For mixture 2, 3 µL of RNAiMax reagent (Thermo Fisher Scientific, Cat#13778075) was added to 100 µL of Opti-MEM and incubated at room temperature for 5 minutes as well. After the incubation, the two mixtures were combined and further incubated at room temperature for 20 minutes. Meanwhile, 0.3 million MCF7 cells were seeded onto 6-well plates. Following the 20-minute incubation period, the combined mixture was added to the cells, bringing the total volume to 1mL. The cells were

then incubated in a CO2 incubator for either 12 or 24 hours before fixation (for imaging) or RNA collection (for qPCR).

3.3.3 Cell fixation for confocal imaging

Cells were seeded on Poly-D lysine coated slides in a 6-well plate during the fluorescent-ASO transfection. After 24 hours, the cells were rinsed twice with ice-cold PBS and fixed with 4% PFA (Fisher Scientific, AAJ19943K2) for 10 minutes. Following the fixation, the cells were washed twice with PBS and blocked with 200 µL of an immunofluorescence (IF) blocking solution (Cell Signaling, 12411) for 30 minutes at room temperature. Two PBS washes were performed, and then the cells were stained with 300nM DAPI (Invitrogen, D1306) for 2 minutes. The slides were rinsed with methanol and mounted with a mounting medium (Invitrogen, P36930). Finally, the slides were imaged using a confocal microscope following standard instructions.

3.3.4 RNA Isolation, cDNA synthesis, and quantitative PCR (qPCR)

RNA isolation, cDNA synthesis, and quantitative PCR was performed as described in 2.2.7.

3.3.5 Statistical analysis

All bar graphs were presented as mean ± SD using R version 4.1.2. Statistical analysis involved a two-tailed Student's t-test for comparation of means between groups as specified. Results were considered significant at $p < 0.05$, and the significance level is indicated by asterisks in each figure panel: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

3.3 Results

3.3.1 Localization of LNA-modified ASO after transfection

To visualize the transfection efficiency and LNA localization, we transfected MCF-7 with fluorescently labeled LNA that has no specific target genome-widely. 24 hours post transfection, >80% of the cells were positively transfected. The fluorescent ASO was detectable both in the cytoplasm and nucleus.

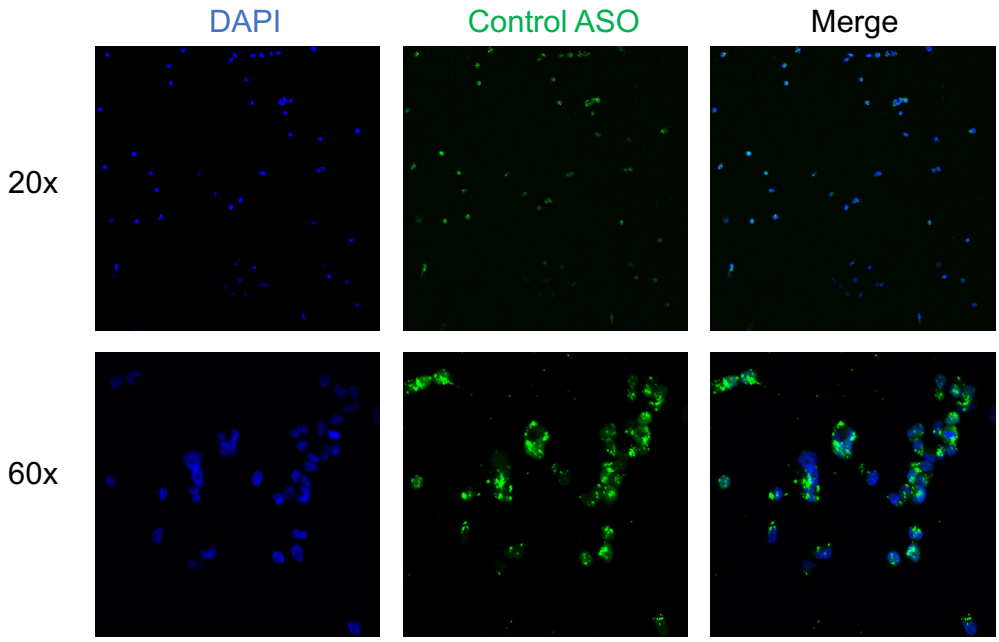

Figure 3.3: Confocal image showing positive transfection rate and localization of florescent-labeled LNA 24 hours post-transfection. LNA concentration in the medium was 80nM.

3.3.2 Acute and effective *TFF1e* knockdown by LNA modified ASOs

The inhibition efficiency of four LASOs targeting the first 2kb of *TFF1e* has been tested by transfecting MCF-7 cells at a concentration of 40nM for 24 hours. *TFF1e* LNA1 and *TFF1e* LNA4 both significantly depleted *TFF1e* RNA (figure 3.4 B). *TFF1e* LNA1 showed an depletion efficiency of >90% and was used for all following experiments. Time course knockdown was performed using *TFF1e* LNA1 at a concentration of 80nM to test whether *TFF1e* can be knocked down in shorter time scale. The result showed that even 6 hours is sufficient for reaching an effective knockdown of over 80% (figure 3.4 C), demonstrating the high potential of LNA in studying the direct role of eRNAs. After 48h of transfection, there was no significant cell death observed both in the control group and *TFF1e* KD group, demonstrating 40nM and 80nM is within the safety range (figure 3.4 D). The viability and morphology of cells was comparable between the control group and *TFF1e* KD group, demonstrating *TFF1e* might not directly regulate essential genes for cell survival.

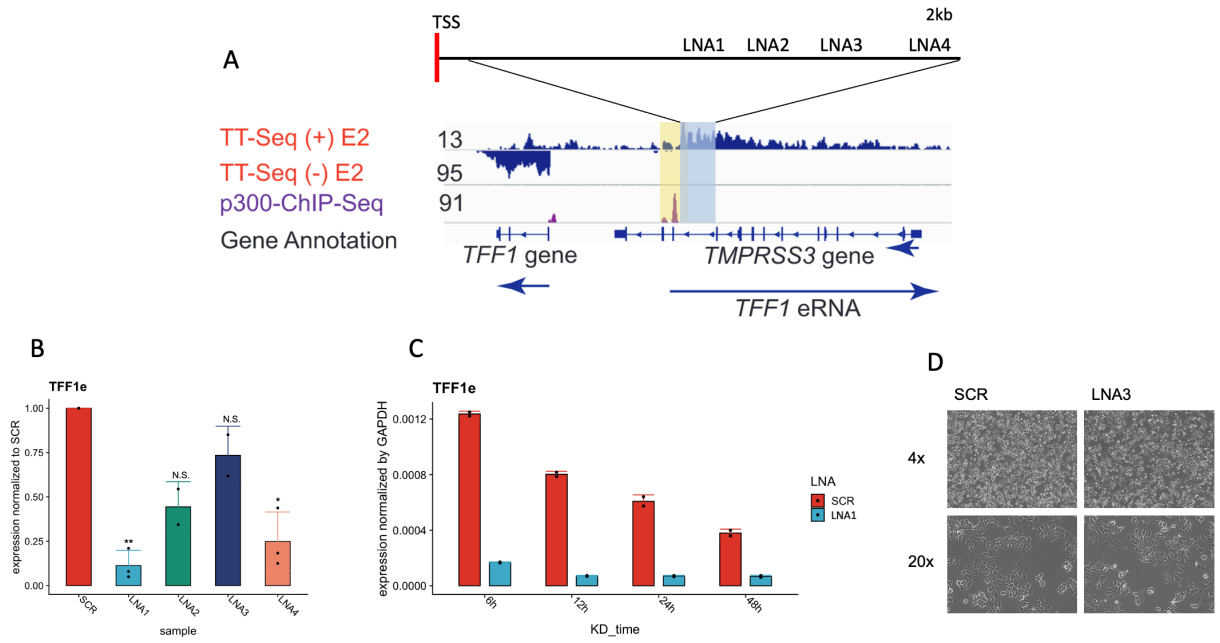

Figure 3.4: Acute and robust *TFF1e* knockdown by LNA. (A) Position of four LNAs targeting the first 2kb of *TFF1e* transcript; (B) RT-qPCR analyzing the knockdown effect of four LNAs to *TFF1e* after 24 hours of transfection at a concentration of 40nM, each dot represents one biological replicate; (C) RT-qPCR analyzing the knockdown effect of *TFF1e* LNA1 to *TFF1e* after 6, 12, 24, 48 hours of transfection at a concentration of 80nM, each dot represents one technical replicate; (D) Microscope image showing cell viability after transfection of SCR LNA and *TFF1e* LNA1 for 48 hours at a concentration of 40nM.

3.3. Discussion

We have demonstrated ASO mediated knockdown can achieve acute and effective inhibition for eRNA within 6 hours of transfection. Therefore, ASO knockdown has been selected as the acute inhibition system for investigating the direct role of eRNA in this study. Compared with CRISPRi (KRAB), which inhibits the process of transcription and in turn reduce the level of RNA transcripts, ASO mediated knockdown directly target the RNA without directly perturbing the transcription process. This property could facilitate study on the function of the RNA transcript itself free from the transcription process.

The knockdown effect of 4 LNAs targeting the first 2kb of *TFF1e* did not show a correlation between LNA position and inhibition efficiency. The knockdown effect might be

more relevant to LNA binding affinity and RNA secondary structure. Nevertheless, LNAs proximal to the TSS are still preferred for knocking down non-coding RNA without well-defined transcription termination site, whose transcription might terminate upstream the annotated site. LNAs that target downstream of the termination site are incapable of knocking down those "short transcripts", which might still be functional through their 5' motifs.

Off-targeting effects remains a concern for LNA mediated knockdown, as with all methods involving reverse-complementary base pairing. Although we discarded all the sequence with >=13nt match in the human transcriptome, the possibility of off targeting still exists. 13nt is an arbitrary threshold that we used because no sequence was left if we lower the threshold to 12. There is no absolute data supporting the exact minimum match to trigger a RNaseH cleavage, but the existing studies has shown mismatches significantly dampen the knockdown efficiency[83]. Moreover, performing BLAST with the hg38 reference transcriptome may fail to detect potential mismatches to some un-annotated transcripts. Off-targeting effects can be mitigated by individually performing knockdown using two different ASOs for one target and only looking for consistently altered genes. Nevertheless, even if the *TFF1e* ASO is off-targeting to some protein-coding gene, the corresponding cellular protein level might not be substantially affected after 6 or 12 hours because most of proteins have their half-lives over 12 hours[84]. The ability to achieve efficient knockdown within a span of 6 hours has reduced the sensitivity of antisense oligonucleotides (ASOs) to off-target effects. This is particularly advantageous when studying the direct function of eRNA.

CHAPTER 4: Dissecting eRNA's role in regulating target gene transcription and epigenetic features

4.1 Background

4.1.1 epigenetic features at enhancers and promoters

Histone modifications has been reported closely associated with active or inactive transcription stages[2]. However, whether the association comes from a direct causal role of histone modification on transcription is still under years of debate. Whether the transcription

could also affect the deposition of histone marks has also been discussed[85,86]. With recent studies of enhancer RNA, transcription of eRNA and deposition of histone modifications seems to mutually affect each other[30,31]. Understanding the interplay between eRNA transcription and histone modification could facilitate our understanding to transcription process.

Histone acetylation has been reported as both the cause and consequence of transcription, as described below. The acetyl group was first recognized to mediate transcriptional activation by neutralizing positive charge of Lysine residual in the histone tail and reduce its affinity to DNA, making the locus more accessible and transcriptionally active[86,87]. Histone acetylation is now recognized to mediate transcriptional activation by recruiting specific reader proteins. Bromo-domain containing protein like BRD4 bind to H3K27ac and recruits the positive transcription elongation factor b (P-TEFb) to promote transcription elongation[88,89]. A recent study on enhaner RNA has found the IDR (intrinsic disorder domain) of BRD4 can interact with the IDR or RNA m6A reader YTHDC1 and form transcriptional condensates[70]. YEATS domain of YEATS family proteins[90,91] binds to H3K9ac, H3K27ac, and H3K18ac and mediate various molecular process including transcriptional regulation, transcriptional elongation, and chromatin remodeling[90,91]. Recent study on eRNAs has revealed nascent eRNA transcripts stimulate the HAT activity of CBP, increase H3K27ac and H3K18ac at the enhancer and target promoter, thereby promote gene expression[31]. This observation supports at least in the enhancer region histone acetylation could be a consequence of eRNA transcription. The fact that histone acetylation acts as both cause and consequence of transcription shows the complexity of transcriptional regulation with feed-back and enhancing effect. For a specific locus, an acute perturbation system could be powerful at identifying whether a histone modification is more of a cause or consequence of eRNA expression.

Histone methylation is associated with both transcriptional inhibition and activation, depending on the position of the methylated lysine. H3K4me, H3K36me, and H3K79me are associated with gene activation, while H3K9me, H3K27me, and H3K20me are associated with

gene inhibition[92]. Transcriptional activation or inhibition associated with different histone methylation are largely dependent on reader proteins. For example, the chromodomain of HP1 proteins recognize methylated H3K9 and mediate heterochromatin spreading and transcriptional silencing[93,94]. H3K4me1 preferentially marks active enhancers, while H3K4me3 preferentially marks active promoters, but normally both are observed at promoters and enhancers[15]. Methylated H3K4 is recognized by PHD domain (plant homeodomain) of BPTF (Bromodomain PHD finger transcription factor) that opens condensed chromatin to increase DNA accessibility and induce transcriptional activation[95]. A recent study has argued a distinct role of H3K4me3 in facilitating promoter proximal Pol II pause-release by recruiting integrator complex subunit 11 (INTS11) rather than initiation[96]. H3K4 methylation has also been identified as consequence of transcription as it was reported that deposition of H3K4me1/2 follows eRNA induction upon TLR signaling and was significantly reduced by inhibiting Pol II elongation[30].

4.1.2 ChIP assay for profiling transcription related features

ChIP (Chromatin Immuno-Precipitation)[97] is the commonly used method for profiling genomic binding sites of specific proteins, including transcription factors, histones, RNA polymerase, and other proteins interacting with DNA. The basic workflow of ChIP is first perform chemical crosslinking to preserve the protein-DNA interaction, sonicate the sample to break DNA into small segments of 200-300bp, add an antibody targeting a specific protein to bind with the protein-DNA complex, then pulldown the complex with magnetic beads that have high affinity to immunoglobin. DNA can be purified for profiling by either qPCR (ChIP-qPCR) or DNA-sequencing (ChIP-seq). The distribution of histone modifications on certain gene promoter or enhancer can be profiled by using antibodies targeting the modified histone, like H3K4me1/2/3, H3K27ac, etc. Transcription activity at certain regions can be profiled by using antibody targeting unphosphorylated Rpb1 subunit of RNA pol2, or Rpb1 with phosphorylated Ser5/2 specifically for transcription initiation and elongation.

4.2 Materials and methods

4.2.1 RNA Isolation, cDNA synthesis, and quantitative PCR (qPCR)

RNA isolation, cDNA synthesis, and quantitative PCR was performed as described in 2.2.7. Sequences for primers targeting the *TFF1* intron and *TFF1* mRNA could be found in 2.2.7.

4.2.2 ChIP

ChIP was performed as previously described[70] with minor modifications. MCF-7 cells were cross-linked with 1% formaldehyde (Millipore Sigma, 252549) for 10 minutes in a 10cm dish. Each immunoprecipitation (IP) group used 1-2 million cells. Sonication were performed with a 20% amplitude for 9 minutes. 2mg of H3K27ac antibody (Abcam 4729, batch: GR3251520-1) and 6 µL of Rpb1 antibody (Cell Signaling 14958s, batch: 5) were incubated overnight with each IP group.

4.4.3 Sequencing Library Preparation and Sequencing

ChIP-Seq libraries were prepared using the NEB Ultra II DNA Library Kit (NEB, E7645). Sequencing was performed on the NextSeq 550 Sequencing system according to the manufacturer's instructions using a paired-end 40bp mode.

4.2.4 Next generation sequencing data processing

To process the reads, Trim Galore (https://github.com/FelixKrueger/TrimGalore) was utilized to remove adapters. The resulting clean reads were then aligned to the human genome hg38 using BWA-MEM 0.7 (Li and Durbin, 2009). Duplicate reads caused by PCR were removed using Picard (http://broadinstitute.github.io/picard/). Only uniquely mapped reads were kept and converted into bigwig files using deepTools' bamCoverage function. The conversion process involved normalization using RPKM (Ramírez et al., 2014). Finally, the resulting bigwig files were visualized using IGV 2.16.0 (Robinson et al., 2011).

4.2.5 Statistical analysis

Bar graphs for qPCR were presented as mean ± SD using R version 4.1.2. Bar graphs for ChIP-seq quantification were presented by RPM using R version 4.1.2.

4.3 Results

4.3.1 Acute *TFF1e* depletion directly inhibit the transcription of target gene

We performed knock down of *TFF1e* with *TFF1e* LNA1 at a concentration of 40nM for 24hrs. qPCR primer targeting the intron of *TFF1* gene was applied for detecting the transcription activity of the gene. Primers that cross an exon-exon junction are more suitable for detecting the mature mRNA level and may not show significant changes even if the transcription activity was inhibited for short term because mRNA usually has longer half-life and the existing mRNA has much higher level than newly synthesized pre-mRNA. Therefore, primers targeting an intron is more suitable for detecting the transcription activity. Upon the *TFF1e* knockdown, qPCR primers targeting an intron saw reduction of ~50%, while primers targeting the *TFF1* mRNA didn't see reduction. The result shows acute *TFF1e* inhibition directly affect the target gene transcription before mRNA level is substantially affected. This observation also emphasis the necessity of using qPCR primers targeting an intron to detect transient transcriptional changes.

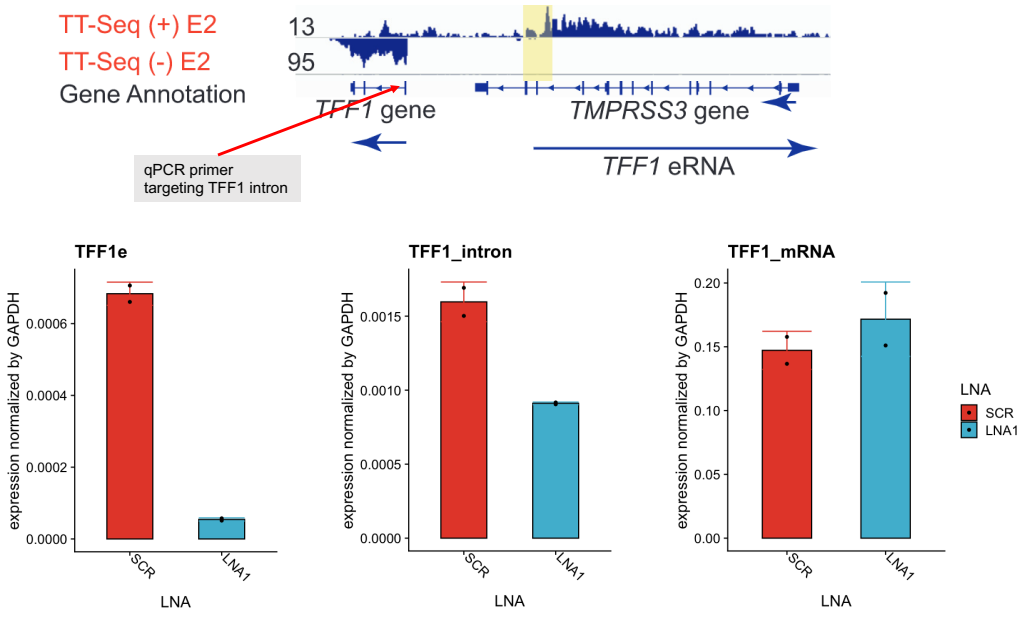

Figure 4.1: RT-qPCR analyzing the changes of intronic RNA and mRNA of *TFF1* gene upon 24 hours of *TFF1e* knockdown. LNA concentration in the medium was 40nM. Each dot represents one technical replicate.

4.3.2 Acute *TFF1e* depletion directly inhibit the RNA Pol II activity at target gene

For identifying the alteration of RNA Pol II activity and histone modification, ChIP was performed after acute depletion of *TFF1e* for 12hrs using *TFF1e* LNA4 and *TFF1e* LNA1 independently at a concentration of 80nM. Antibodies targeting Rpb1 (CTD) or H3K27ac were used for pulling down the associated genomic DNA. Rpb1 is the largest subunit of RNA polymerase II and the binding of which directly reveals the RNA pol II occupancy at that region, partially implying the transcription activity. H3K27ac marks active enhancers. Upon 12hrs of *TFF1e* depletion by both LNA1 and LNA4, Rpb1 signal at the *TFF1* gene and *TFF1e* was reduced, further validating the conclusion by intron qPCR that *TFF1e* inhibition directly affect the target gene transcription. Unexpectedly, the level of H3K27ac was only slightly reduced after *TFF1e* knockdown, indicating transcription changes at *TFF1* gene might precede the alteration of H3K27ac. Therefore, reduced H3K27ac might not be the immediate effect of eRNA depletion and might not be the cause of inhibited *TFF1* transcription.

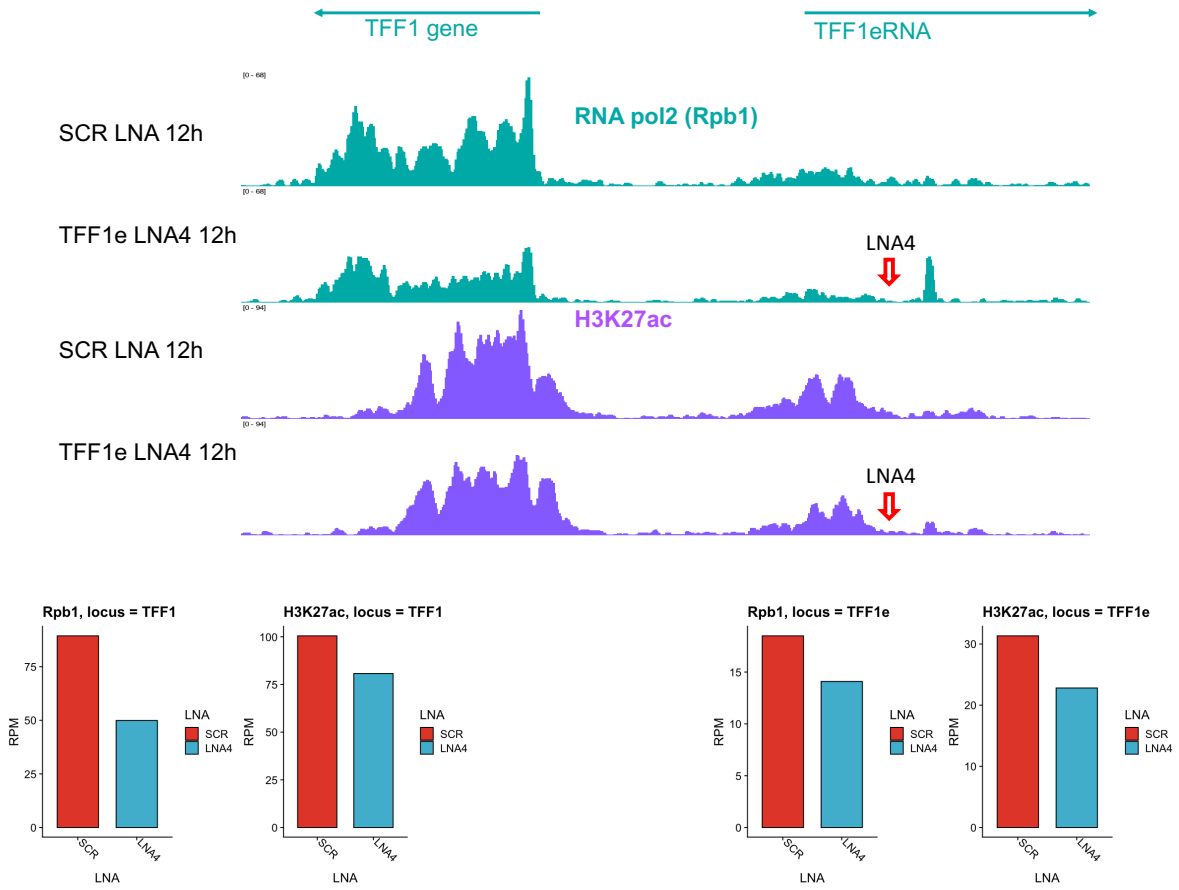

Figure 4.2: Top: IGV snapshot of ChIP-seq profiling Rpb1 binding and H3K27ac level at *TFF1e* and *TFF1* gene region upon *TFF1e* knockdown for 12 hours by LNA4. Numbers on the track represent RPKM. LNA concentration in the medium was 80nM. Bottom: Quantification of ChIP-seq signal using RPM of the peak region.

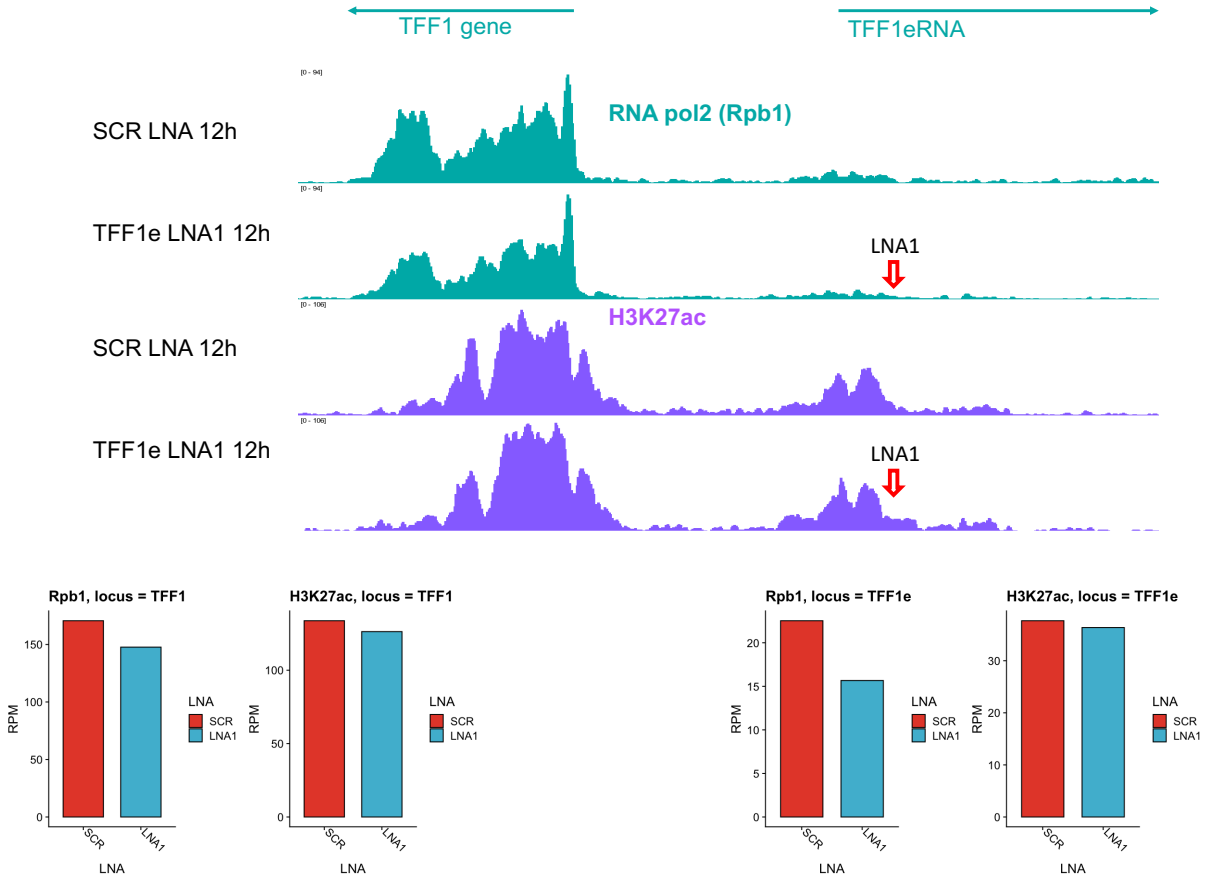

Figure 4.3: Top: IGV snapshot of ChIP-seq profiling Rpb1 binding and H3K27ac level at *TFF1e* and *TFF1* gene region upon *TFF1e* knockdown for 12 hours by LNA1. Numbers on the track represent RPKM. LNA concentration in the medium was 80nM. Bottom: Quantification of ChIP-seq signal using RPM of the peak region.

4.4 Discussion

Revealed by my intron RT-qPCR and ChIP-seq data, LNA-mediated knockdown of enhancer RNA was able to induce transcription changes of the target gene upon short-term perturbation and could be a powerful tool for mechanistic study of eRNA. Additionally, I found robustness of a perturbation method is essential for the result interpretation. Even using *TFF1e*

LNA1, there were times when the knockdown efficiency didn't reach >80% (as shown in one replicate in figure 3.4 B). In such cases the reduction in *TFF1* intron might be minor or undetectable, and only when a certain level of perturbation was achieved can alterations in the target genes or the epigenome be detected. Thinking of CRISPR inhibition or activation screening followed by single-cell sequencing, the effect of some perturbed transcripts might be underestimated due to insufficient perturbation level. Although LNA mediated knockdown shows potent robustness in this single locus study, their application in large scale screening was not achieved in the current stage because large-scale LNA synthesis is still under development.

Our data of unchanged H3K23ac upon 12 hours of eRNA knockdown was inconsistent with previously published work by Bose et al. proposing eRNA's role in activating the enzymic activity of CPB[31] and facilitating the deposition of H3K27ac. ASO mediated knockdown of eRNAs was also performed in their study and reductions in H3K27ac was observed. The inconsistency might be due to 1) locus specific effect; 2) knockdown in their study was performed for over 36 hours, while the knockdown in my study was performed for 12 hours. Nevertheless, a slight reduction of H3K27ac at both the promoter and enhancer was still observed upon 12 hours of *TFF1e* knockdown. I cannot exclude that slight reduction of H3K27ac is capable of reducing BRD4 binding and p-TEFb recruitment, and in turn tune down the transcription. For dissecting whether H3K27ac plays a causal role, time course experiment should be performed to identify the time point from which H3K27ac starts to change.

Time course knockdown should also be performed to study the dynamic changes of additional epigenomic features in eRNA transcription including H3K4me1/3 deposition, BRD4 binding, MED1 binding, CTCF binding, cohesin binding, and enhancer-promoter looping as they are all closely involved in transcription process. It has not been fully elucidated that which step is the immediate response to eRNA depletion due to the lack of acute perturbation methods. For example, evidence has shown eRNA inhibition is associated with both target gene downregulation and compromised enhancer-promoter looping[18], but it's still under debate

whether eRNA depletion directly inhibits looping and in turn reduces the target gene transcription, or the alteration in looping follows the reduction in target gene transcription. Acute eRNA depletion could potentially resolve this debate at least in several single loci.

ChIP-seq has come up with one unprecedented result that the signal of Rpb1 binding at *TFF1e* was also reduced upon the knockdown. The result could be interpreted as *TFF1e* knockdown directly inhibits not only the eRNA transcript but also the transcription process. Two potential explanations are either the local eRNA level/density directly affects eRNA transcription through a feed-back loop, or the LNA1 caused pre-mature transcription termination. The proposed model for LNA-mediated premature termination involves RNA cleavage at the targeted LNA position, resulting in the generation of phosphorylated end without 5' cap. This end could be recognized by the exoribonuclease XRN2, which degrades the newly synthesized RNA until it reaches the transcribing Pol II, similar to the torpedo model observed in canonical transcriptional termination[98]. If pre-mature termination occurs at *TFF1e* region, the Pol II binding near the transcription start site should remain unchanged and the binding downstream the LNA targeting sites should be reduced. However, reduction of Pol II binding was observed near the TSS and throughout the whole *TFF1e* region, rejecting the hypothesis of pre-mature termination triggered by *TFF1e* LNA1. Therefore, it's likely that *TFF1e* transcripts also promote the transcription of itself, and the depletion of which directly inhibited eRNA transcription through a feed-back loop.

Biologically, I have confirmed *TFF1* gene as a direct target of *TFF1* enhancer RNA by acute eRNA depletion, and the H3K27ac might not be the causal event driving *TFF1* downregulation. The *TFF1e* locus is just a testbed eRNA in the current stage and I am looking forward to investigating multiple eRNAs with similar method for summarizing a global pattern. Methodologically, I first explored CRISPR-based methods for eRNA perturbation and have proposed several potential caveats. As a major achievement, I proved LNA-mediated knockdown as a potent method for acute perturbation of eRNA and established an efficient workflow from LNA design to effect testing. This method could be further applied to other

eRNAs or all other types of RNAs, to perform time course perturbation, and to confirm more potential functions of eRNA including recruitment of multiple co-factors and regulating targets on other chromosome (trans-regulation)[28].

Bibliography

1. Shlyueva, D., Stampfel, G. & Stark, A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15,** 272–286 (2014).

2. Calo, E. & Wysocka, J. Modification of Enhancer Chromatin: What, How, and Why? *Mol Cell* **49,** 825–837 (2013).

3. Schoenfelder, S. & Fraser, P. Long-range enhancer–promoter contacts in gene expression control. *Nat Rev Genet* **20,** 437–455 (2019).

4. Levine, M. Transcriptional Enhancers in Animal Development and Evolution. *Curr. Biol.* **20,** R754–R763 (2010).

5. Blackwood, E. M. & Kadonaga, J. T. Going the Distance: A Current View of Enhancer Action. *Science* **281,** 60–63 (1998).

6. Bulger, M. & Groudine, M. Functional and Mechanistic Diversity of Distal Transcription Enhancers. *Cell* **144,** 327–339 (2011).

7. Claringbould, A. & Zaugg, J. B. Enhancers in disease: molecular basis and emerging treatment strategies. *Trends Mol Med* **27,** 1060–1073 (2021).

8. Kvon, E. Z., Kamneva, O. K., Melo, U. S., Barozzi, I., Osterwalder, M., Mannion, B. J., Tissières, V., Pickle, C. S., Plajzer-Frick, I., Lee, E. A., Kato, M., Garvin, T. H., Akiyama, J. A., Afzal, V., Lopez-Rios, J., Rubin, E. M., Dickel, D. E., Pennacchio, L. A. & Visel, A. Progressive Loss of Function in a Limb Enhancer during Snake Evolution. *Cell* **167,** 633-642.e11 (2016).

9. Emison, E. S., McCallion, A. S., Kashuk, C. S., Bush, R. T., Grice, E., Lin, S., Portnoy, M. E., Cutler, D. J., Green, E. D. & Chakravarti, A. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature* **434,** 857–863 (2005).

10. Falo-Sanjuan, J., Lammers, N. C., Garcia, H. G. & Bray, S. J. Enhancer Priming Enables Fast and Sustained Transcriptional Responses to Notch Signaling. *Dev. Cell* **50,** 411-425.e8 (2019).

11. Su, G., Wang, W., Zhao, X., Chen, J., Zheng, J., Liu, M., Bi, J., Guo, D., Chen, B., Zhao, Z., Shi, J., Zhang, L. & Lu, W. Enhancer architecture-dependent multilayered transcriptional regulation orchestrates RA signaling-induced early lineage differentiation of ESCs. *Nucleic Acids Res.* **49,** 11575–11595 (2021).

12. Hah, N., Murakami, S., Nagari, A., Danko, C. G. & Kraus, W. L. Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res.* **23,** 1210–1223 (2013).

13. Li, W., Notani, D. & Rosenfeld, M. G. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat Rev Genet* **17,** 207–223 (2016).

14. Kim, T.-K., Hemberg, M., Gray, J. M., Costa, A. M., Bear, D. M., Wu, J., Harmin, D. A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., Markenscoff-Papadimitriou, E., Kuhl, D., Bito, H., Worley, P. F., Kreiman, G. & Greenberg, M. E. Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465,** 182–187 (2010).

15. Kim, T.-K. & Shiekhattar, R. Architectural and Functional Commonalities between Enhancers and Promoters. *Cell* **162,** 948–959 (2015).

16. Mayran, A. & Drouin, J. Pioneer transcription factors shape the epigenetic landscape. *J. Biol. Chem.* **293,** 13795–13804 (2018).

17. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13,** 613–626 (2012).

18. Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A. Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., Oh, S., Kim, H.-S., Glass, C. K. & Rosenfeld, M. G. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498,** 516–520 (2013).

19. Sabari, B. R., Dall'Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., Abraham, B. J., Hannett, N. M., Zamudio, A. V., Manteiga, J. C., Li, C. H., Guo, Y. E., Day, D. S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T. I., Cisse, I. I., Roeder, R. G., Sharp, P. A., Chakraborty, A. K. & Young, R. A. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361,** (2018).

20. Shrinivas, K., Sabari, B. R., Coffey, E. L., Klein, I. A., Boija, A., Zamudio, A. V., Schuijers, J., Hannett, N. M., Sharp, P. A., Young, R. A. & Chakraborty, A. K. Enhancer Features that Drive Formation of Transcriptional Condensates. *Mol. Cell* **75,** 549-561.e7 (2019).

21. Santa, F. D., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B. K., Muller, H., Ragoussis, J., Wei, C.-L. & Natoli, G. A Large Fraction of Extragenic RNA Pol II Transcription Sites Overlap Enhancers. *PLoS Biol.* **8,** e1000384 (2010).

22. Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I. E., Males, M., Viales, R. R. & Furlong, E. E. M. The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Gene Dev* **32,** 42–57 (2018).

23. Lai, F., Orom, U. A., Cesaroni, M., Beringer, M., Taatjes, D. J., Blobel, G. A. & Shiekhattar, R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **494,** 497–501 (2013).

24. Sigova, A. A., Abraham, B. J., Ji, X., Molinie, B., Hannett, N. M., Guo, Y. E., Jangi, M., Giallourakis, C. C., Sharp, P. A. & Young, R. A. Transcription factor trapping by RNA in gene regulatory elements. *Science* **350,** 978–981 (2015).

25. Schaukowitch, K., Joo, J.-Y., Liu, X., Watts, J. K., Martinez, C. & Kim, T.-K. Enhancer RNA Facilitates NELF Release from Immediate Early Genes. *Mol Cell* **56,** 29–42 (2014).

26. Gorbovytska, V., Kim, S.-K., Kuybu, F., Götze, M., Um, D., Kang, K., Pittroff, A., Brennecke, T., Schneider, L.-M., Leitner, A., Kim, T.-K. & Kuhn, C.-D. Enhancer RNAs stimulate Pol II pause release by harnessing multivalent interactions to NELF. *Nat Commun* **13,** 2429 (2022).

27. Zhao, Y., Wang, L., Ren, S., Wang, L., Blackburn, P. R., McNulty, M. S., Gao, X., Qiao, M., Vessella, R. L., Kohli, M., Zhang, J., Karnes, R. J., Tindall, D. J., Kim, Y., MacLeod, R., Ekker, S. C., Kang, T., Sun, Y. & Huang, H. Activation of P-TEFb by Androgen Receptor-Regulated Enhancer RNAs in Castration-Resistant Prostate Cancer. *Cell Reports* **15,** 599–610 (2016).

28. Tsai, P.-F., Dell'Orso, S., Rodriguez, J., Vivanco, K. O., Ko, K.-D., Jiang, K., Juan, A. H., Sarshad, A. A., Vian, L., Tran, M., Wangsa, D., Wang, A. H., Perovanovic, J., Anastasakis, D., Ralston, E., Ried, T., Sun, H.-W., Hafner, M., Larson, D. R. & Sartorelli, V. A Muscle-Specific Enhancer RNA Mediates Cohesin Recruitment and Regulates Transcription In trans. *Mol Cell* **71,** 129-141.e8 (2018).

29. Hsieh, C.-L., Fei, T., Chen, Y., Li, T., Gao, Y., Wang, X., Sun, T., Sweeney, C. J., Lee, G.-S. M., Chen, S., Balk, S. P., Liu, X. S., Brown, M. & Kantoff, P. W. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. *Proc. National Acad. Sci.* **111,** 7319–7324 (2014).

30. Kaikkonen, M. U., Spann, N. J., Heinz, S., Romanoski, C. E., Allison, K. A., Stender, J. D., Chun, H. B., Tough, D. F., Prinjha, R. K., Benner, C. & Glass, C. K. Remodeling of the

Enhancer Landscape during Macrophage Activation Is Coupled to Enhancer Transcription. *Mol Cell* **51,** 310–325 (2013).

31. Bose, D. A., Donahue, G., Reinberg, D., Shiekhattar, R., Bonasio, R. & Berger, S. L. RNA Binding to CBP Stimulates Histone Acetylation and Transcription. *Cell* **168,** 135-149.e22 (2017).

32. Tyssowski, K. M., DeStefino, N. R., Cho, J.-H., Dunn, C. J., Poston, R. G., Carty, C. E., Jones, R. D., Chang, S. M., Romeo, P., Wurzelmann, M. K., Ward, J. M., Andermann, M. L., Saha, R. N., Dudek, S. M. & Gray, J. M. Different Neuronal Activity Patterns Induce Different Gene Expression Programs. *Neuron* **98,** 530-546.e11 (2018).

33. Hou, T. Y. & Kraus, W. L. Analysis of estrogen-regulated enhancer RNAs identifies a functional motif required for enhancer assembly and gene expression. *Cell Reports* **39,** 110944 (2022).

34. Zhang, Z., Lee, J.-H., Ruan, H., Ye, Y., Krakowiak, J., Hu, Q., Xiang, Y., Gong, J., Zhou, B., Wang, L., Lin, C., Diao, L., Mills, G. B., Li, W. & Han, L. Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. *Nat Commun* **10,** 4562 (2019).

35. Oh, S., Shao, J., Mitra, J., Xiong, F., D'Antonio, M., Wang, R., Garcia-Bassets, I., Ma, Q., Zhu, X., Lee, J.-H., Nair, S. J., Yang, F., Ohgi, K., Frazer, K. A., Zhang, Z. D., Li, W. & Rosenfeld, M. G. Enhancer release and retargeting activates disease-susceptibility genes. *Nature* **595,** 735–740 (2021).

36. Wang, Y., Zhang, C., Wang, Y., Liu, X. & Zhang, Z. Enhancer RNA (eRNA) in Human Diseases. *Int. J. Mol. Sci.* **23,** 11582 (2022).

37. Xiang, J.-F., Yin, Q.-F., Chen, T., Zhang, Y., Zhang, X.-O., Wu, Z., Zhang, S., Wang, H.-B., Ge, J., Lu, X., Yang, L. & Chen, L.-L. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res* **24,** 513–531 (2014).

38. Pomerantz, M. M., Ahmadiyeh, N., Jia, L., Herman, P., Verzi, M. P., Doddapaneni, H., Beckwith, C. A., Chan, J. A., Hills, A., Davis, M., Yao, K., Kehoe, S. M., Lenz, H.-J., Haiman, C. A., Yan, C., Henderson, B. E., Frenkel, B., Barretina, J., Bass, A., Tabernero, J., Baselga, J., Regan, M. M., Manak, J. R., Shivdasani, R., Coetzee, G. A. & Freedman, M. L. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat. Genet.* **41,** 882–884 (2009).

39. You, Z., Liu, C., Wang, C., Ling, Z., Wang, Y., Wang, Y., Zhang, M., Chen, S., Xu, B., Guan, H. & Chen, M. LncRNA CCAT1 Promotes Prostate Cancer Cell Proliferation by Interacting with DDX5 and MIR-28-5P. *Mol. Cancer Ther.* **18,** 2469–2479 (2019).

40. Wang, X., Cairns, M. J. & Yan, J. Super-enhancers in transcriptional regulation and genome organization. *Nucleic Acids Res* **47,** 11481–11496 (2019).

41. Cramer, P. Organization and regulation of gene transcription. *Nature* **573,** 45–54 (2019).

42. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Bio* **19,** 621–637 (2018).

43. Liao, Z., Lee, J.-H., Krakowiak, J., Wang, R. & Li, W. Enforced Activation of Enhancer RNAs In Situ through the dCas9 Synergistic Activation Mediator System. *J Vis Exp* (2020). doi:10.3791/61460

44. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. & Lim, W. A. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152,** 1173–1183 (2013).

45. Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., Guimaraes, C., Panning, B., Ploegh, H. L., Bassik, M. C., Qi, L. S., Kampmann, M. & Weissman, J. S. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159,** 647–661 (2014).

46. Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S. & Qi, L. S. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154,** 442–451 (2013).

47. Mousavi, K., Zare, H., Dell'Orso, S., Grontved, L., Gutierrez-Cruz, G., Derfoul, A., Hager, G. L. & Sartorelli, V. eRNAs Promote Transcription by Establishing Chromatin Accessibility at Defined Genomic Loci. *Mol. Cell* **51,** 606–617 (2013).

48. Melo, C. A., Drost, J., Wijchers, P. J., van de Werken, H., de Wit, E., Vrielink, J. A. F. O., Elkon, R., Melo, S. A., Léveillé, N., Kalluri, R., de Laat, W. & Agami, R. eRNAs Are Required for p53-Dependent Enhancer Activity and Gene Transcription. *Mol. Cell* **49,** 524–535 (2013).

49. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513,** 569–573 (2014).

50. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337,** 816–821 (2012).

51. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339,** 819–823 (2013).

52. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. & Church, G. M. RNA-Guided Human Genome Engineering via Cas9. *Science* **339,** 823–826 (2013).

53. Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I. & Liu, D. R. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551,** 464–471 (2017).

54. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533,** 420–424 (2016).

55. Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Bio* **17,** 5–15 (2016).

56. Shechner, D. M., Hacisüleyman, E., Younger, S. T. & Rinn, J. L. CRISPR Display: A modular method for locus-specific targeting of long noncoding RNAs and synthetic RNA devices in vivo. *Nat Methods* **12,** 664–670 (2015).

57. Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., Hsu, P. D., Habib, N., Gootenberg, J. S., Nishimasu, H., Nureki, O. & Zhang, F. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517,** 583–588 (2015).

58. Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E. & Gersbach, C. A. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33,** 510–517 (2015).

59. Wang, K., Escobar, M., Li, J., Mahata, B., Goell, J., Shah, S., Cluck, M. & Hilton, I. B. Systematic comparison of CRISPR-based transcriptional activators uncovers gene-regulatory features of enhancer–promoter interactions. *Nucleic Acids Res.* **50,** 7842–7855 (2022).

60. Hirai, H., Tani, T. & Kikyo, N. Structure and functions of powerful transactivators: VP16, MyoD and FoxA. *Int. J. Dev. Biology* **54,** 1589–1596 (2011).

61. Thakore, P. I., D'Ippolito, A. M., Song, L., Safi, A., Shivakumar, N. K., Kabadi, A. M., Reddy, T. E., Crawford, G. E. & Gersbach, C. A. Highly Specific Epigenome Editing by CRISPR/Cas9 Repressors for Silencing of Distal Regulatory Elements. *Nat Methods* **12,** 1143–1149 (2015).

62. Schraivogel, D., Gschwind, A. R., Milbank, J. H., Leonce, D. R., Jakob, P., Mathur, L., Korbel, J. O., Merten, C., Velten, L. & Steinmetz, L. M. Targeted Perturb-seq enables genome-scale genetic screens in single cells. *Nat Methods* **17,** 629–635 (2020).

63. Alerasool, N., Segal, D., Lee, H. & Taipale, M. An efficient KRAB domain for CRISPRi applications in human cells. *Nat Methods* **17,** 1093–1096 (2020).

64. Schultz, D. C., Friedman, J. R. & Rauscher, F. J. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2α subunit of NuRD. *Genes Dev.* **15,** 428–443 (2001).

65. Gasperini, M., Hill, A. J., McFaline-Figueroa, J. L., Martin, B., Kim, S., Zhang, M. D., Jackson, D., Leith, A., Schreiber, J., Noble, W. S., Trapnell, C., Ahituv, N. & Shendure, J. A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. *Cell* **176,** 377-390.e19 (2019).

66. Das, A. T., Tenenbaum, L. & Berkhout, B. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr. Gene Ther.* **16,** 156–167 (2016).

67. Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W. & Bujard, H. Transcriptional Activation by Tetracyclines in Mammalian Cells. *Science* **268,** 1766–1769 (1995).

68. Kearns, N. A., Genga, R. M. J., Enuameh, M. S., Garber, M., Wolfe, S. A. & Maehr, R. Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development* **141,** 219–223 (2013).

69. Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., Gagneur, J. & Cramer, P. TT-seq maps the human transient transcriptome. *Science* **352,** 1225–1228 (2016).

70. Lee, J.-H., Wang, R., Xiong, F., Krakowiak, J., Liao, Z., Nguyen, P. T., Moroz-Omori, E. V., Shao, J., Zhu, X., Bolt, M. J., Wu, H., Singh, P. K., Bi, M., Shi, C. J., Jamal, N., Li, G., Mistry, R., Jung, S. Y., Tsai, K.-L., Ferreon, J. C., Stossi, F., Caflisch, A., Liu, Z., Mancini, M. A. & Li, W. Enhancer RNA m6A methylation facilitates transcriptional condensate formation and gene activation. *Mol Cell* **81,** 3368-3385.e9 (2021).

71. Schmidt, R., Steinhart, Z., Layeghi, M., Freimer, J. W., Bueno, R., Nguyen, V. Q., Blaeschke, F., Ye, C. J. & Marson, A. CRISPR activation and interference screens decode stimulation responses in primary human T cells. *Science* **375,** eabj4008 (2022).

72. Kwon, D. Y., Zhao, Y.-T., Lamonica, J. M. & Zhou, Z. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. *Nat. Commun.* **8,** 15315 (2017).

73. Ai, Y., Liang, D. & Wilusz, J. E. CRISPR/Cas13 effectors have differing extents of off-target effects that limit their utility in eukaryotic cells. *Nucleic Acids Res* (2022). doi:10.1093/nar/gkac159

74. Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A., Lander, E. S., Voytas, D. F., Ting, A. Y. & Zhang, F. RNA targeting with CRISPR–Cas13. *Nature* **550,** 280–284 (2017).

75. Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., Shmakov, S., Makarova, K. S., Semenova, E., Minakhin, L., Severinov, K., Regev, A.,

Lander, E. S., Koonin, E. V. & Zhang, F. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353,** aaf5573 (2016).

76. Colognori, D., Trinidad, M. & Doudna, J. A. Precise transcript targeting by CRISPR-Csm complexes. *Nat. Biotechnol.* 1–9 (2023). doi:10.1038/s41587-022-01649-9

77. Tong, H., Huang, J., Xiao, Q., He, B., Dong, X., Liu, Y., Yang, X., Han, D., Wang, Z., Wang, X., Ying, W., Zhang, R., Wei, Y., Xu, C., Zhou, Y., Li, Y., Cai, M., Wang, Q., Xue, M., Li, G., Fang, K., Zhang, H. & Yang, H. High-fidelity Cas13 variants for targeted RNA degradation with minimal collateral effects. *Nat Biotechnol* **41,** 108–119 (2023).

78. Rinaldi, C. & Wood, M. J. A. Antisense oligonucleotides: the next frontier for treatment of neurological disorders. *Nat Rev Neurol* **14,** 9–21 (2018).

79. Lima, W. F., Hoyos, C. L. D., Liang, X. & Crooke, S. T. RNA cleavage products generated by antisense oligonucleotides and siRNAs are processed by the RNA surveillance machinery. *Nucleic Acids Res.* **44,** 3351–3363 (2016).

80. Lundin, K. E., Højland, T., Hansen, B. R., Persson, R., Bramsen, J. B., Kjems, J., Koch, T., Wengel, J. & Smith, C. I. E. Chapter Two Biological Activity and Biotechnological Aspects of Locked Nucleic Acids. *Adv. Genet.* **82,** 47–107 (2013).

81. Shen, W., Hoyos, C. L. D., Migawa, M. T., Vickers, T. A., Sun, H., Low, A., Bell, T. A., Rahdar, M., Mukhopadhyay, S., Hart, C. E., Bell, M., Riney, S., Murray, S. F., Greenlee, S., Crooke, R. M., Liang, X., Seth, P. P. & Crooke, S. T. Chemical modification of PS-ASO therapeutics reduces cellular protein-binding and improves the therapeutic index. *Nat Biotechnol* **37,** 640–650 (2019).

82. Lee, J.-S. & Mendell, J. T. Antisense-Mediated Transcript Knockdown Triggers Premature Transcription Termination. *Mol Cell* **77,** 1044-1054.e3 (2020).

83. Rasmussen, S. & Roberts, P. Functional studies of microRNA based on knockdown using Locked Nucleic Acid probes. *Nat Methods* **4,** iii–iv (2007).

84. Rolfs, Z., Frey, B. L., Shi, X., Kawai, Y., Smith, L. M. & Welham, N. V. An atlas of protein turnover rates in mouse tissues. *Nat. Commun.* **12,** 6778 (2021).

85. Wang, Z., Chivu, A. G., Choate, L. A., Rice, E. J., Miller, D. C., Chu, T., Chou, S.-P., Kingsley, N. B., Petersen, J. L., Finno, C. J., Bellone, R. R., Antczak, D. F., Lis, J. T. & Danko, C. G. Prediction of histone post-translational modification patterns based on nascent transcription data. *Nat Genet* **54,** 295–305 (2022).

86. Chen, Y.-J. C., Koutelou, E. & Dent, S. Y. R. Now open: Evolving insights to the roles of lysine acetylation in chromatin organization and function. *Mol. Cell* **82,** 716–727 (2022).

87. Eberharter, A. & Becker, P. B. Histone acetylation: a switch between repressive and permissive chromatin. *EMBO Rep.* **3,** 224–229 (2002).

88. Yang, Z., Yik, J. H. N., Chen, R., He, N., Jang, M. K., Ozato, K. & Zhou, Q. Recruitment of P-TEFb for Stimulation of Transcriptional Elongation by the Bromodomain Protein Brd4. *Mol. Cell* **19,** 535–545 (2005).

89. Jang, M. K., Mochizuki, K., Zhou, M., Jeong, H.-S., Brady, J. N. & Ozato, K. The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. *Mol. Cell* **19,** 523–534 (2005).

90. Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., Ren, Y., Jin, Q., Dent, S. Y. R., Li, W., Li, H. & Shi, X. AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation. *Cell* **159,** 558–571 (2014).

91. Zhao, D., Li, Y., Xiong, X., Chen, Z. & Li, H. YEATS Domain—A Histone Acylation Reader in Health and Disease. *J. Mol. Biol.* **429,** 1994–2002 (2017).

92. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13,** 343–357 (2012).

93. Danzer, J. R. & Wallrath, L. L. Mechanisms of HP1-mediated gene silencing in Drosophila. *Development* **131,** 3571–3580 (2004).

94. Fischer, T., Cui, B., Dhakshnamoorthy, J., Zhou, M., Rubin, C., Zofall, M., Veenstra, T. D. & Grewal, S. I. S. Diverse roles of HP1 proteins in heterochromatin assembly and functions in fission yeast. *Proc. Natl. Acad. Sci.* **106,** 8998–9003 (2009).

95. Li, H., Ilin, S., Wang, W., Duncan, E. M., Wysocka, J., Allis, C. D. & Patel, D. J. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442,** 91–95 (2006).

96. Wang, H., Fan, Z., Shliaha, P. V., Miele, M., Hendrickson, R. C., Jiang, X. & Helin, K. H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature* **615,** 339–348 (2023).

97. Park, P. J. ChIP–seq: advantages and challenges of a maturing technology. *Nat Rev Genet* **10,** 669–680 (2009).

98. Proudfoot, N. J. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* **352,** aad9926 (2016).

Vita

Lanxin Bei was born in China, Guilin. She graduated from Shanghai Jiao Tong University in China with Bachelor of Animal Science. She then matriculated into the master's program at The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences in August 2021.