The role of human SMARCA5 in chromatin structure and function

By

Monica Lorraine Bomber

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

May 12, 2023

Nashville, Tennessee

Approved:

Bruce Carter, Ph.D., Chair

William Tansey, Ph.D.

Emily Hodges, Ph.D.

Manuel Ascano, Jr., Ph.D.

Scott Hiebert, Ph.D., Advisor

Dedication

To my parents, Kristen and Phil Bomber, as well as my siblings, Helena, Max, and Conrad. Thank you for all your support.

Acknowledgements

The research in this dissertation was made possible by the Biochemical and Chemical Training for Cancer Research training grant, T32CA009582.

Firstly, I would like to thank my advisor and committee. Thank you to my mentor, Dr. Scott Hiebert, for the guidance and support throughout my years in his lab. Thank you to my committee members, Dr. Bruce Carter, Dr. Emily Hodges, Dr. Manual Ascano, and Dr. William Tansey, for all their support, advice, and guidance over the years.

I would also like to thank present and past lab members. Dr. Kristy Stengel has been a great mentor and sounding board in the lab, leading by example with her great scientific discoveries. Dr. Shilpa Sampathi, as a post-doctoral researcher in the lab was always a great teacher, helping me understand different molecular mechanisms. Another post-doctoral researcher, Dr. Pankaj Archarya, taught me how to use different computational technologies, inspiring me to learn more about dry lab techniques. The first graduate student Dr. Gretchen Wemhoener-Johnston, was a great example of a graduate student, helping me with my class work throughout the years and always made great desserts. Susu Zhang, the second graduate student in the lab provided great feedback and great suggestions of places to eat. Hillary Layden, the third graduate student in the lab, provided great feedback on papers, computational problems, and presentations (and a place to live).

I would especially like to help those lab members working more behind the scenes. Our first lab manager Steven Pierce, was a great help, aiding me in experiments and helping me locate any resource I needed. The lab technician, Brenda Stark, provided a lot of morale support, great desserts, and all the clean pipettes and lab glassware we

could ever need. Our newest lab technician, Jacob Ellis, offered great morale support and always offered to help. The final behind-the-scenes member I would like to thank is Leslie, a member of the maintenance staff who was there to chat and keep me company during the late nights in the lab, his positive encouragement was always appreciated.

I cannot forget about each of the Vanderbilt Cores that helped me along the way. Thank you to the Flow Cytometry Core, including Brittany Matlock and David Flaherty, for all the assistance and laughs. Thank you to VANTAGE, especially Angela Jones, for helping me troubleshoot all my sequencing experiments.

A special thank you goes out to my family and friends. The support of my parents, siblings, extended family, and friends has been crucial to me as I went through this long and stressful process called a “Ph.D.”. I would like to provide a special thanks to the Secret Bagel Society Girls (Victoria, Katie, Jenny, and Bethany) for listening during hard times and providing feedback when needed.

The final and last thank you goes to Dr. Michael Stadler from Swiss Institute of Bioinformatics, for all his help with the loess modeling and code that he provided for the nucleosome repeat length calculations.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
7-AAD	7-amino-actinomycin D
ACF	ATP-Utilizing Chromatin Assembly and Remodeling Factor
ACF1	ATP-Utilizing Chromatin Assembly and Remodeling Factor 1
AID	Auxin-inducible degradation
AML	Acute myeloid leukemia
AML1-ETO	AML1-ETO Fusion Protein
APC	Allophycocyanin
ATAC-seq	Assay for transposase-accessible chromatin with high-throughput sequencing
ATCC	American Type Culture Collection
ATM	ATM Serine/Threonine Kinase
ATP	Adenosine triphosphate
BET	Bromodomain and extra-terminal
BFP	Blue Fluorescent Protein
BORIS	Brother of the Regulator of Imprinted Sites
BRD	Bromodomain-containing protein
BRD4	Bromodomain-containing protein 4
Brg1	Protein Brahma Homolog 1
BSA	Bovine Serum Albumin
CAF1	CCR4-Associated Factor 1
CBX7	Chromobox 7

CD11b	Cell Surface Glycoprotein MAC-1 Subunit Alpha
CD34	Hematopoietic progenitor cell antigen CD34
CD82	CD82 molecule
CDK	Cyclin-Dependent Kinase
CDK7	Cyclin-Dependent Kinase 7
CDK9	Cyclin-Dependent Kinase 9
CENP	Centromere Protein
ChIP-seq	Chromatin immunoprecipitation followed by sequencing
CHRAC	Chromatin Accessibility Complex
CHRAC1	Chromatin Accessibility Complex Subunit 1
CML	Chronic myelogenous leukemia
CO_2	Carbon dioxide
CPSF73	Cleavage and Polyadenylation-Specific Factor 3
CPM	Counts per million
CRBN	Cereblon
CRIPSR	Clustered regularly interspaced short palindromic repeats
CTCF	CCCTC-Binding Factor
CTD	C-Terminal Domain
CUT&RUN	Cleavage under targets & release using nuclease
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase-seq	DNase I-hypersensitive site sequencing

DNMT	DNA methyltransferase
DNMT1	DNA (cytosine-5)-methyltransferase 1
DNMT3	DNA (cytosine-5)-methyltransferase 3
DOT1L	DOT1-Like Histone Lysine Methyltransferase
DSIF	DRB Sensitivity-Inducing Factor
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ENL	MLLT1 Super Elongation Complex Subunit
ETO	Eight-twenty-one
ETOZnf	Eight-twenty-one zinc finger
FACT	Facilitates Chromatin Transcription Complex
FKBP12	FKBP Prolyl Isomerase 1A
FPKM	fragments per kilobase of exon per million mapped fragments
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
GB	Gene body
GTF	General transcription factor
HAT	Histone Acetyl Transferase
HBD	Histone-binding domain
HCl	Hydrochloric acid
HDAC	Histone Deacetylase
HDAC1	Histone Deacetylase 1
Hi-C	Genome wide chromatin conformation capture

HDR	Homology-directed repair
HP1	Heterochromatin Protein 1
HSF	Heat Shock Transcription Factor 1
HSPC	Hematopoietic stem and progenitor cell
HSS	HAND-SANT-SLIDE domain
IKZF1	IKAROS Family Zinc Finger 1
IKZF3	IKAROS Family Zinc Finger 3
IMDM	Iscove's modified Dulbecco's medium
INO80	Inositol-requiring protein 80
IP	Immunoprecipitation
ISWI	Imitation switch
KCL	Potassium chloride
KDM	Lysine Demethylase
KO	Knockout
LSD1	Lysine-Specific Demethylase 1
MACS2	Model-based analysis of ChIP-seq
MCM2	Minichromosome Maintenance Complex Component 2
MDM2	MDM2 Proto-Oncogene, E3 Ubiquitin Protein Ligase
mESCs	Mouse embryonic stem cells
MetAP-2	Methionine Aminopeptidase-2
MM	Multiple myeloma
Mnase-seq	Micrococcal nuclease digestion with deep sequencing

mRNA	Messenger ribonucleic acid
MYC	MYC Proto-Oncogene, BHLH Transcription Factor
NEB	New England Biolabs
NELF	Negative elongation factor
NFR	Nucleosome-free region
NGS	Next-generation sequencing
NIR	Near-infrared
NM1	Nuclear myosin 1
NMR	Nuclear magnetic resonance
NoRC	Nucleolar remodeling complex
NP40	Nonyl phenoxypolyethoxylethanol
NRL	Nucleosome repeat length
NURF	Nucleosome Remodeling Factor
ORC	Origin recognition complex
PAF	Platelet-activating factor
PAS	Polyadenylation signal
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with Tween
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase chain reaction
PDVF	Polyvinylidene difluoride
PI	Proprium iodide

PIC	Preinitiation complex
PMA	Phorbol Myristate Acetate
POI	Protein of interest
PP	Promoter Proximal
PP2A	Protein phosphatase 2
PRC1	Polycomb repressive complex 1
PRO-seq	Precision nuclear run-on followed by sequencing
PROTAC	Proteolysis targeting chimera
p-TEFb	Positive transcription elongation factor
PTM	Post-translational modification
PU.1	Refer to SPI1
RAD21	Human homolog of radiation sensitive mutant 21
REST	RE1 Silencing Transcription Factor
RNA	Ribonucleic acid
RNAPI	RNA Polymerase I
RNAPII	RNA Polymerase II
RNAPIII	RNA Polymerase III
RNP	RNA
RPMI	Roswell Park Memorial Institute Medium
rRNA	Ribosomal RNA
RSF	Remodeling and spacing factor
RTF1	RNA Polymerase-Associated Protein RTF1 Homolog

RUNX1	RUNX Family Transcription Factor 1
SCF	Skp1-Culline-F box containing Hrt
SDS	Sodium dodecyl sulfate
SEC	Super elongation complex
SMARCA5	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin A5
SPI1	Spi-1 Proto-Oncogene
SPT6	Transcription elongation factor SPT6
SRCAP	Snf2-related CREBBP activator protein
SWR1	SWI2/SNF2-Related 1 Chromatin Remodeling Complex
TAD	Topologically associated domains
TBP	TATA Box-Binding Protein
TF	Transcription factor
TFIIA	General Transcription Factor IIA Subunit 1
TFIIB	General Transcription Factor IIB
TFIID	Transcription factor II D
TFIIE	Transcription factor II E
TFIIF	Transcription factor II F
TFIIH	Transcription factor II H
THZ1	Covalent cyclin-dependent kinase 7 (CDK7) inhibitor
TIP5	Transcription termination factor 1 interacting protein 5
TSS	Transcription start site
WB	Western blot

WICH	WSTF-SNF2h complex
WSTF	Williams Syndrome Transcription Factor
WT	Wild type
XRN2	5'-3' Exoribonuclease 2
YY1	YY1 Transcription Factor

1. CHAPTER 1: Introduction

1.1 Chromatin structure

The central dogma of gene expression states that deoxyribonucleic acid (DNA) is transcribed into ribonucleic acid (RNA) by RNA polymerases, which is then translated into protein by the ribosome (Fig 1.1A) (Brenner et al., 1961; Crick, 1970, 1958). To understand this process, we must determine the basic biochemistry of DNA structure. DNA is a double helix consisting of pairs of nucleotides, also known as bases. Four bases (guanine, adenine, thymidine, and cytosine) encode DNA and are later transcribed into RNA. The human genome contains ~2 meters of DNA, which is equivalent to a total of 3 billion base pairs. The DNA double helix is negatively charged and wraps around positively charged histone octamers, called nucleosomes. Histone octamers consist of two copies of four positively charged core histone proteins, H2A, H2B, H3, and H4 (Kornberg, 1977; Luger et al., 1997). The negatively charged DNA helix can wrap 1.7 times around each nucleosome equaling ~147 bp, thereby creating an organized array of nucleosomal-wrapped DNA known as “beads on a string” (Fig 1.1B) (Kornberg, 1977; Luger et al., 1997). A linker histone, H1, binds to the histone octamer and DNA between nucleosomes. The H1 histone acts as the glue to promote the folding and structure of the 147-bp nucleosome and the 20–50 bp linker DNA to create higher-order chromatin (Thoma et al., 1979). The spacing between nucleosomes can vary within a cell, changing the fluidity of chromatin compaction and gene regulation through nucleosome mobilization (Sandro Baldi et al., 2020).

The 3 billion base pairs of DNA in a human cell are divided into 23 pairs of compact chromatin structures, known as chromosomes. The 3 billion base pairs within the 23

chromosome pairs encode about 20,000 genes. Regulation of these 20,000 genes within chromosomes requires the balanced coordination of *histone*-modifying enzymes, chromatin-remodeling enzymes, and transcription factors (TFs) along the DNA and histone octamers. DNA-templated processes regulate these genes during transcription, replication, and repair. In this dissertation, I will discuss a chromatin remodeling ATPase, SWI/SNF-related matrix-associated actin-dependent regulator of chromatin A5 (SMARCA5), and how its role in chromatin structure affects DNA-templated processes such as TF binding, transcription, and DNA replication.

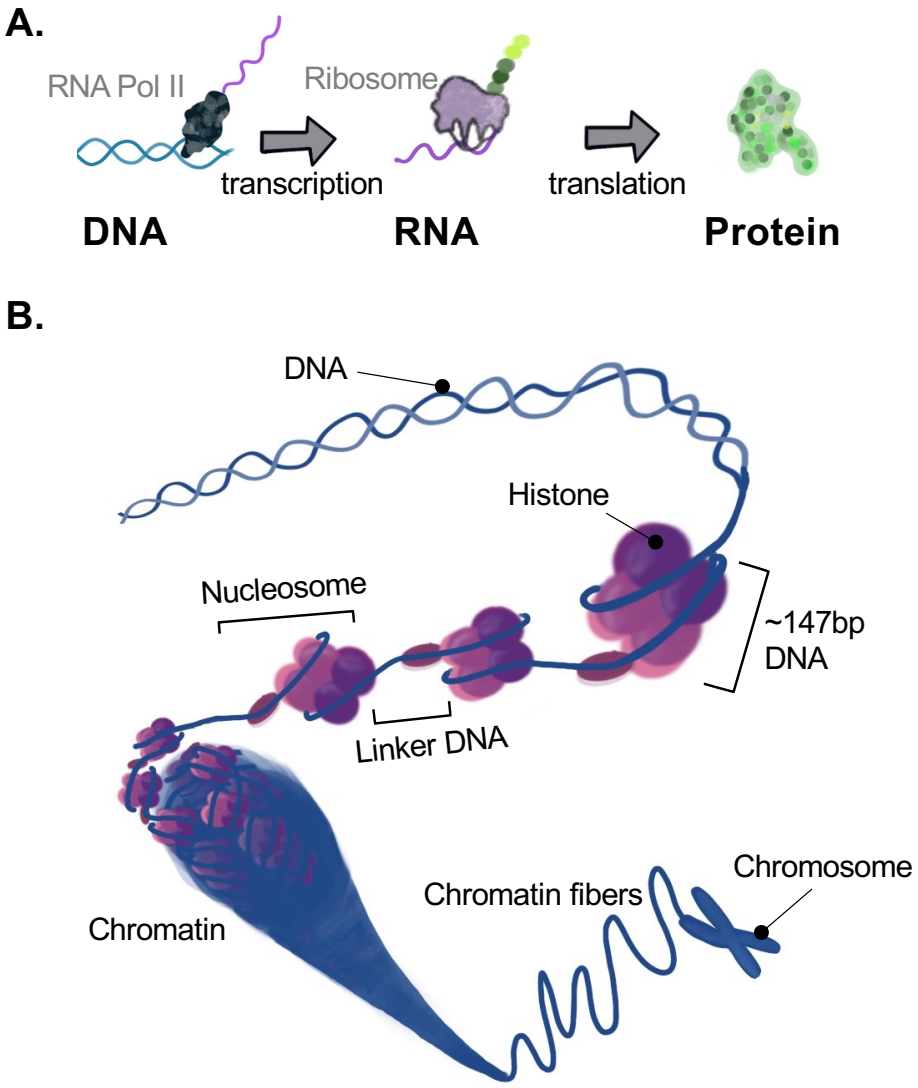

Figure 1.1: The DNA components of the cell.

(A) The central dogma. DNA (blue) with RNA Polymerase II transcribing mRNA (purple). Then, the ribosome translates mRNA into amino acids (green). The amino acids then fold into proteins. (B) The DNA double helix is wrapped around a histone octamer, which is known as the nucleosome core. Nucleosomes also contain a linker Histone H1 on the piece of linker DNA between the nucleosome cores. The nucleosome structures are compacted into chromatin, which then forms the chromatin fibers of the chromosomes.

1.2 Histone modifications

It is essential to understand the building blocks of the nucleosome, histones, and their role in chromatin dynamics. The histones that make up the nucleosome were first discovered to be modified by methylation or acetylation marks in the 1960s (Allfrey et al., 1964; Murray, 1964; Phillips and Simson,1962). Studies in the 1970s determined that these marks were located on the histone 'tails' (Lilley et al., 1976). Subsequent research established that histone-modifying enzymes add or remove post-translational modifications (PTMs) to histone tails, and these enzymes are called writers and erasers (Fig 1.2). The writers, such as histone acetyltransferases (HATs), add chemical modifications to the histone tails in the form of acetylation. Other marks that can be added to histone tails by different histone-modifying enzymes are methylation, phosphorylation, and ubiquitination. Erasers remove these chemical marks. An example eraser includes histone deacetylases (HDACs), which remove the acetylation mark catalyzed by HATs (Fig 1.2).

Readers are another class of these histone-modifying enzymes that identify and interpret chemical modifications (Fig 1.2). Histone tail acetylation is associated with active transcription or open chromatin; an example of an activating factor recruited to acetylated tails is the reader BRD4 (Hargreaves et al., 2009; Moon et al., 2005). Histone tail methylation is associated with both activation and repression of transcription. The activation mark H3K4me3 at promoters is read by the PHD finger domain of TAF3 within the TFIID complex, subsequently leading to more efficient formation of the preinitiation complex (PIC) for transcription activation (Lauberth et al., 2013; Vermeulen et al., 2007). The repressive mark H3K27me3 is read by EED, which recruits the rest of the PRC2 to

add more H3K27me3 to neighboring nucleosomes for repression of developmental genes (Margueron et al., 2009). We can measure global changes in the histone code using methods such as chromatin immunoprecipitation followed by sequencing (ChIP-seq) and cleavage under targets and release using nuclease (CUT&RUN) (Mardis, 2007; Skene et al., 2018). The histone code combined with histone-modifying enzymes and chromatin remodeling complexes alters the chromatin landscape and the mechanism of DNA-templated processes (Li et al., 2007; Shi et al., 2022).

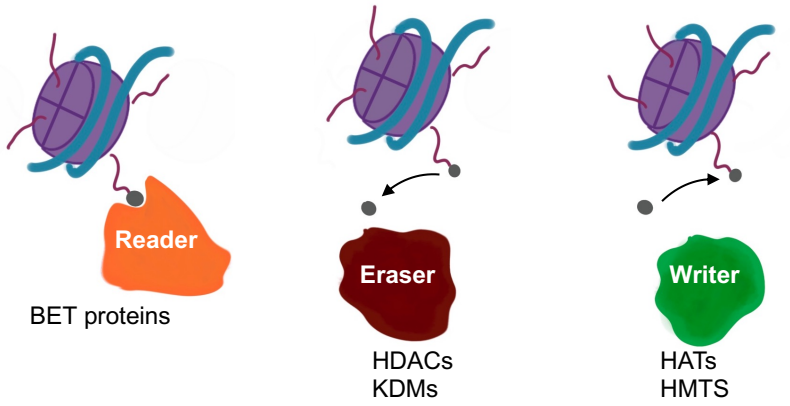

Figure 1.2: Histone-modifying enzymes.

Readers (e.g., BET proteins) read the marks (acetylation) on histone tails. Erasers (e.g., HDACS and KDM) remove the acetylation and methylation marks on histone tails. Writer enzymes add different marks (e.g., acetylation and methylation) to histone tails.

1.3 Histone Variants

Nucleosomes consist of a DNA segment wrapped around eight positively charges histones, known as a histone octamer. As previously mentioned, the histone octamer consists of two copies of H2A, H2B, H3, and H4. Each histone family is encoded by multiple genes, some of which vary greatly from the consensus sequence. Incorporation of these histone variants into nucleosomes can be dependent on cell cycle, and chromatin state and have described roles in gene regulation and replication (Long et al., 2019; Strobino et al., 2020). While there are multiple histone variants, in this study, I will focus on the H2A core histone variant, H2A.Z.

H2A.Z is one of four replication-independent H2A variants. The difference between the H2A core histone and H2A.Z variant is in the C-terminal tail, while the histone fold domain has extensive homology between the two variants (Suto et al., 2000). SWI2/SNF2-Related 1 Chromatin Remodeling Complex (SWR1) places H2A.Z while Inositol-requiring protein 80 (INO80) evicts the H2A.Z histone variant (Luk et al., 2010; Papamichos-Chronakis et al., 2011). Vertebrates have two homologs of SWR1, p400, and Snf2-related CREBBP activator protein (SRCAP) which place H2A.Z.

H2A.Z is associated with gene activation and gene silencing, and the SWR1 complex places H2A.Z at regulatory elements and gene promoters (Giaimo et al., 2019). H2A.Z is enriched around regulatory elements containing CTCF sites (Fu et al., 2008). Loss of H2A.Z leads to improper heterochromatin formation, disorganization of centromere structure, and chromosome segregation defects (Greaves et al., 2007; Hou et al., 2010). Additionally, H2A.Z stimulates the ATPase activity of ISWI remodelers significantly more than the H2A core histone (Goldman et al., 2010). Understanding how

H2A.Z influences DNA-templated processes such as transcription, replication, and chromatin remodeling could provide essential insights into chromatin structure and function.

1.4 Nucleosome repeat length

Chromatin structure can be altered by PTM-mediated changes in histone tails and chromatin remodeling enzymes driven by ATP-dependent reactions. Chromatin remodeling enzymes can install, slide, exchange, and evict nucleosomes within the DNA to alter chromatin structure. The chromatin structure is called open or closed depending on nucleosome spacing along the DNA. Euchromatin also called open chromatin, has equally spaced nucleosomes and/or nucleosome-free regions (NFR) that allow the binding of different factors for active transcription or replication. Heterochromatin (closed chromatin) has unevenly spaced nucleosomes that are tightly compacted into higher-order chromatin structures. Other chromatin structures enable the binding of different factors to DNA, thereby achieving different transcription and replication outcomes and different cell states (Dixon et al., 2015).

Chromatin is highly dynamic and can be analyzed with different experimental techniques to measure changes after specific cell alterations. Techniques that identify chromatin structure measure nucleosome positioning to understand how chromatin structure affects DNA-templated processes such as TF binding, transcription, replication, and repair (Fig 1.3 A). Genome-wide assays can measure the arrangement of nucleosomes within chromatin by performing next-generation sequencing (NGS) experiments such as assay for transposase-accessible chromatin using sequencing (ATAC-seq) and Micrococcal nuclease followed by sequencing (MNase-seq) (Fig 1.3 B, C). ATAC-seq uses a transposase with adapters to insert itself into NFRs, subsequently providing a range of information from NFRs to mono-, di-, and tri-nucleosomal fragments

(Fig 1.3 B). MNase-seq analysis can be used to determine the mononucleosomal pattern within a cell population (Fig 1.3 C). MNase-seq can provide information on the position of mononucleosomes and use that information to measure the length between nucleosome positions or the nucleosome repeat length (NRL) (Fig 1.3 A). The NRL is not constant but changes depending on the genome area, cell type, developmental stage, or transcriptional state of the cells (Sandro Baldi et al., 2020).

Understanding the chromatin structure around a gene or protein binding site can provide information about the state of transcription or replication. The chromatin structure around the transcription start site (TSS) of a repressed gene contains poorly positioned/disorganized nucleosomes; notably, nucleosomes over the TSS block the binding of activating TFs (Fig 1.4 A). The structure of an active gene consists of NFRs around the TSS and well-positioned nucleosomes throughout the gene body (Fig 1.4 B). The NFRs at the TSSs are surrounded by well-positioned nucleosomes, especially +1 and –1 nucleosome (Fig 1.4 B). Knowledge of the chromatin structures associated with specific transcriptional states allows us to better understand the cascades caused by chromatin remodeling enzymes and histone-modifying factors that can alter the chromatin landscape.

Changes in chromatin structure affect transcriptional control and other DNA-templated processes, such as DNA replication and TF binding at intergenic and intronic regions. To replicate DNA, the origin of replication must be available for the replisome to read and initiate DNA replication (Fig 1.4 C). The cell may stop dividing when the origins of replication are closed and chromatin remodelers such as INO80 do not function, leading to cell death or cell cycle arrest. Cells also can exhibit cell death or cell cycle

arrest due to the deregulation of transcription from the loss of TF binding events. The imitation switch (ISWI) chromatin remodeling protein equally spaces the nucleosomes around CTCF motifs and other TF binding sites, thereby enabling processes such as DNA looping (Fig 1.4 D) (refs throughout this section). Disorganization of these equally spaced nucleosome arrays can block the binding of DNA replication proteins and TFs, leading to the loss of essential cellular processes. The knowledge we gain by examining chromatin structure, and chromatin remodeling enzymes can inform us when, where, and how changes in DNA-templated processes can lead to alterations in the cell cycle and cell development.

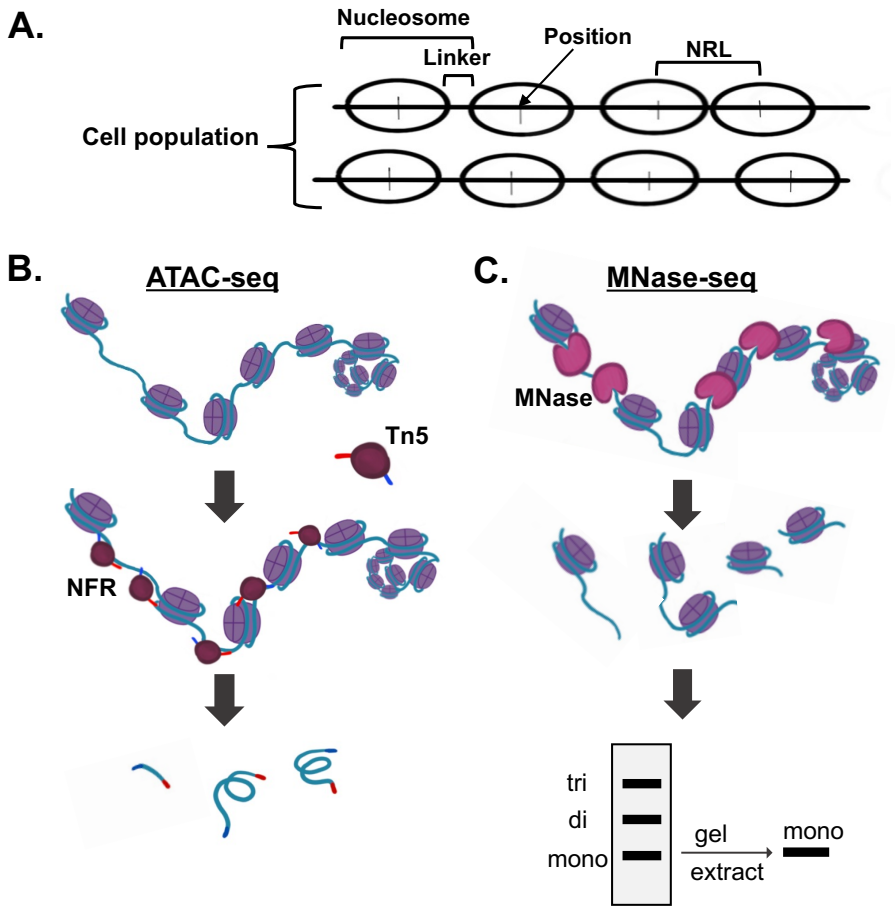

Figure 1.3: Techniques used to analyze nucleosome positioning.

(A) Diagram of the nucleosome structure within a cell population, where the nucleosome contains the histone octamer and the linker DNA. The nucleosome is positioned at the center of the nucleosome, which allows the measurement of NRL. (B) ATAC-seq is a method for measuring NFRs in a cell using a Tn5 transposase incubated with adapters that can insert into open regions of chromatin. This provides small fragments of NFRs and mono-, di-, and tri-nucleosome fragments. (C) MNase-seq is a method for measuring mononucleosomes in a cell using micrococcal nuclease digestion to cut either side of the nucleosome. This provides mono-, di-, and tri-nucleosome fragments; therefore, the mononucleosome fragments must be gel-extracted for library preparation.

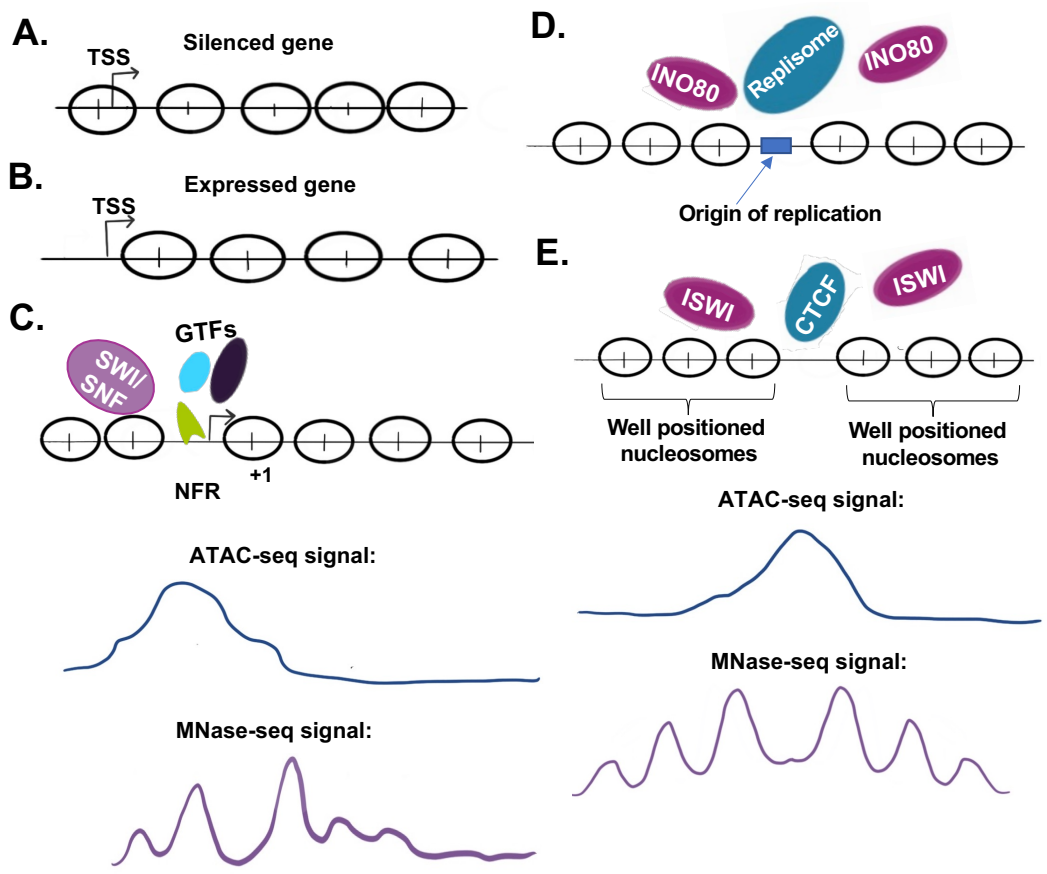

Figure 1.4: Nucleosome structure throughout the genome.

(A) Nucleosome structure around the TSS of a silenced gene. (B) Nucleosome structure of the TSS of an actively transcribed gene. (C) Nucleosome structure around the TSS of a gene with the SWI/SNF chromatin remodeling enzymes and GTFs. The resulting ATAC-seq signal is blue, and the MNase-seq signal is purple. (D) Nucleosome structure around an active replication origin with the replisome in blue and the INO80 chromatin remodeling enzymes in purple. (E) Nucleosome structure around the motif binding site of the CTCF transcription factor (blue) with chromatin remodeling enzymes. The resulting ATAC-seq signal is blue, and the MNase-seq signal is purple.

1.5 Transcriptional regulation

The transcription cycle for the regulation of gene expression is essential for cell development, growth, and survival. The steps involved in transcription involve chromatin remodeling enzymes, DNA-binding TFs, RNA Polymerase II (RNAPII), and general transcription factors (GTFs), which coordinately achieve RNA synthesis and gene expression. This process involves the initiation, promotor proximal pausing, pause release, elongation, and termination of RNAPII. Any alterations in these steps can lead to dysregulation of gene expression and cell death or disease. Dissecting each step during the transcriptional cycle will allow us to better understand how alterations could lead to disease initiation and progression.

1.5.1 Initiation

Transcriptional initiation starts with the formation of the preinitiation complex (PIC) at promoters with open regions of chromatin or NFRs (Fig 1.5 A). PIC formation begins with the binding of the TATA box binding protein (TBP), a component of TFIID, along with other TFIID elements that bind downstream of the TSS (Shopland et al., 1995). The GTFs, TFIIA, and TFIIB then bind to TBP to facilitate recruitment of RNAPII to DNA (Fig 1.5 B). Mediator then binds to TFIID to stabilize the PIC and create long-range interactions between enhancers and promoters (Chen et al., 2021). Next, RNAPII binds to the promoter with TFIIE, TFIIF, and TFIIH to expose the DNA for transcription initiation. The TFIIH complex containing the CDK7 kinase first phosphorylates the fifth residue in the C-terminal domain (CTD) of RNAPII, Ser5 (Fig 1.5 C) (Glover-Cutter et al., 2009). Phosphorylation of Ser5 on the CTD of RNAPII results in the eviction of the Mediator complex and activation of RNA capping enzymes, and RNAPII starts to transcribe 20 to

60 nucleotides of RNA before pausing due to binding of Negative elongation factor (NELF) (Allen and Taatjes, 2015; Li et al., 2013).

1.5.2 Pausing and pause release

RNAPII pausing between the promoter and the first (+1) nucleosome is stabilized by the binding of NELF and DRB sensitivity-inducing factor (DSIF) to RNAPII on DNA (Vos et al., 2018b; Yamaguchi et al., 1999). NELF and DSIF stabilize paused RNAPII along the DNA, thereby creating a kinetic barrier for RNAPII. The first (+1) nucleosome also creates a physical barrier that impedes RNAPII elongation, ultimately creating the promoter-proximal pause (Fig 1.5 D). To be released from this pause, RNAPII can have premature termination from the promoter-proximal pause, or Positive transcription elongation factor b (p-TEFb) can mediate pause release. P-TEFb contains the CDK9 and Cyclin T kinase that is recruited to paused RNAPII by TFs such as c-MYC, BRD4 bound to acetylated histone tails, or the Super elongation complex (SEC) to activate pause release (Eberhardy and Farnham, 2001; Luo et al., 2012; Moon et al., 2005). Once p-TEFb binds, CDK9 phosphorylates NELF, DSIF, and Ser2 on the RNAPII CTD, resulting in pause release and elongation; that is if the PP2A-Integrator complex does not block Ser2 phosphorylation (Fig 1.5 E) (Peterlin and Price, 2006; Vervoort et al., 2021).

1.5.3 Elongation

To continue productive elongation through the gene body, RNAPII must overcome different barriers, usually nucleosomes, by SWI/SNF chromatin remodeling enzymes. Polymerase-associated factor complex (PAF) is necessary for elongation and allows RNAPII to reach optimal speed and processivity by the binding of SPT6 and RTF1. SPT6 facilitates transcription through nucleosomes and optimizes elongation velocity at the

beginning of gene bodies, whereas RTF1 is necessary to maintain elongation velocity throughout the gene body (Fig 1.5 F) (Chen et al., 2009; Vos et al., 2018a; Žumer et al., 2021). Facilitates chromatin transcription (FACT) destabilizes nucleosome barriers during transcription by disassembling the H2A-H2B dimer from nucleosomes to allow RNAPII to proceed through the gene body, then redeposits the H2A-H2B dimer after RNAPII transcription (Formosa, 2012).

1.5.4 Termination

Transcription can be terminated by the following processes: mRNA polyadenylation, RNAPII eviction from DNA, RNAPII slowing, and/or mRNA release from RNAPII for nuclear export. The main process to terminate the transcriptional cycle is through the polyadenylation signal (PAS), which is then recognized and cleaved by the XRN2 exonuclease and the CPSF73 endonuclease (Eaton et al., 2018). RNAPII can then be recycled and recruited to the next active gene.

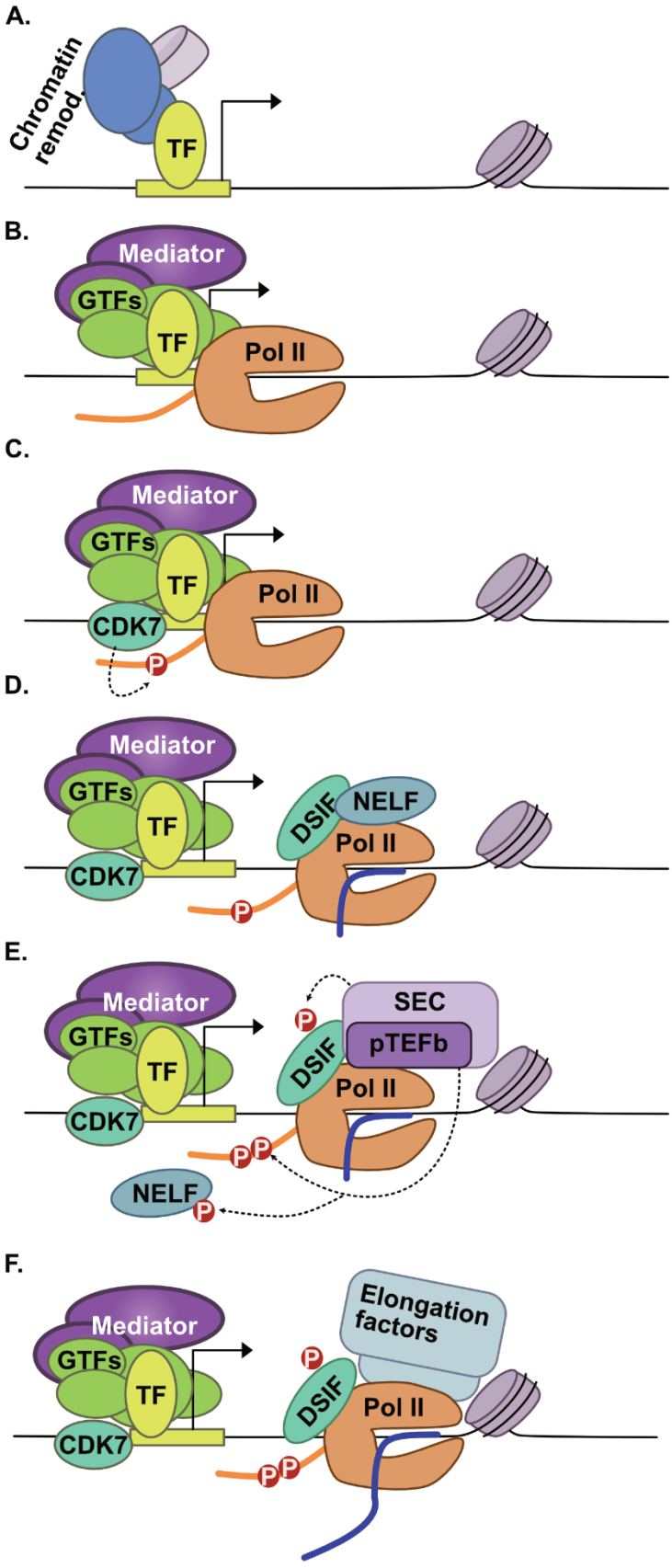
A.
Chromatin remod
TF
B.
Mediator
GTFs
TF
Pol II
C.
Mediator
GTFs
TF
Pol II
CDK7
P
D.
Mediator
GTFs
TF
DSIF
NELF
Pol II
CDK7
P
E.
Mediator
GTFs
TF
SEC
pTEFb
DSIF
Pol II
CDK7
P
P
NELF
P
F.
Mediator
GTFs
TF
Elongation factors
P
DSIF
Pol II
CDK7
P
P

Figure 1.5: RNA polymerase II transcription.

(A) Chromatin remodeling enzymes are recruited by transcription factors to move nucleosomes around the TSS. (B) Mediator and other GTFs bind the TSS to recruit RNAPII. C. CDK7 phosphorylates the serine 5 on the RNAPII C terminal tail. D. The phosphorylation of serine 5 leads to release of Mediator as RNAPII transcribes ~20bp and is paused by binding of NELF and DSIF. E. The SEC binds and p-TEFb phosphorylates NELF, DSIF, and serine 2 on the C-terminal tail of RNAPII leading to pause release and elongation. F. Elongation factors binds RNAPII leading to positive elongation.

1.6 Nucleosome-dependent replication

Human cells contain ~2 meters of DNA that must be replicated every time a cell divides, approximately every 24 hours. This takes DNA-templated process takes many different factors, such as chromatin remodeling enzymes, histone chaperones, helicases, and DNA polymerases to bind open chromatin. Human chromatin creates a barrier to replisome assembly (Azmi et al., 2017). Chromatin remodeling enzymes, such as INO80 or ISWI, must move or slide nucleosomes away from the origin of replication to allow access to DNA and replisome assembly (Längst and Becker, 2001; Papamichos-Chronakis and Peterson, 2008). The ACF-ISWI complex has been shown to be important for replication through heterochromatin and replication fork velocity after knockdown (Bhaskara et al., 2013; Collins et al., 2002). INO80 is also necessary for replication fork velocity, as well as stabilized by BAP1 at replication forks, and necessary for recovery from stalled replication forks (Lee et al., 2014; Shimada et al., 2008). Together this data shows how important chromatin remodeling enzymes are at replication forks.

The replisome also requires the binding of a helicase, histone chaperones, and nucleosome disassembly factor to an open region of chromatin (DNA without nucleosomes) to assemble with the DNA clamp, primase, and polymerase (Berbenetz et al., 2010; Lipford and Bell, 2001; Yeeles et al., 2015). In eukaryotic cells, the minichromosome maintenance (MCM) helicase binds to open regions of double-stranded DNA to transform it into single-stranded DNA as FACT disassembles the nucleosomes and histones in front of the replication fork (Evrin et al., 2009; Formosa, 2012; Labib et al., 2000). The polα primase is then loaded onto the DNA to create RNA primers for the polδ and polε polymerases (Stillman, 2008). PCNA, the polymerase loading clamp, loads

leading strand polymerase, polδ, and lagging strand polymerase, polε, for DNA replication (Navas et al., 1995; Nick McElhinny et al., 2008; Pursell et al., 2007). As DNA is replicated, CAF1 mediates the assembly of new nucleosomes and histones onto the new sets of daughter DNA (Smith and Stillman, 1989). The chromatin structure has a critical role in DNA replication as cells cannot replicate without access to DNA.

1.7 SMARCA5

1.7.1 Discovery of ISWI homologs

The first descriptions of SWI/SNF-related matrix-associated actin-dependent regulator of chromatin A5 (SMARCA5) occurred in studies of yeast homologs. The first yeast ATPase (SNF2/SWI2) was described in 1986 as required for de-repression of the glucose-repressible gene *Suc* (Abrams et al., 1986). The yeast SNF2/SWI2 was discovered as a possible helicase in 1992 (Okabe et al., 1992). Two years after the discovery of the SNF2/SWI function in yeast, the *Drosophila* SWI2-related genes, *brm* and *ISWI*, were cloned, characterized, and identified (Elfring et al., 1994). In 1998, the Human Genome Project identified the human homolog of *Drosophila* ISWI from *SMARCA5* cDNA, with 73% amino acid homology to ISWI (Aihara et al., 1998). It was theorized that SMARCA5 could contribute to the onset of hepatocellular carcinomas as there was a loss of heterozygosity in the chromosomal location of *SMARCA5*. These studies laid the groundwork for understanding the phenotypic and molecular roles of SMARCA5.

1.7.2 SMARCA5 as a member of the ISWI subfamily

SMARCA5 is a member of the ISWI subfamily of SWI/SNF complexes. SMARCA5 is an ATP-dependent helicase and chromatin remodeling enzyme necessary for sliding nucleosomes along DNA to regulate processes such as replication, repair, and transcription (Fig 1.6 B). The general structure and function of the ISWI chromatin remodeling ATPase homolog, SMARCA5, is conserved from yeast to mammals. Members of the ISWI subfamily contain RecA-like lobes/domains that constitute the

ATPase, thereby allowing DNA translocation from the nucleosome (Fig 1.6 A) (Iestyn et al., 2003).

Two auto-inhibitory domains (AutoN and NegC) surround the translocation domain (Fig 1.6 A) (Clapier and Cairns, 2012). The AutoN domain inhibits the translocation domain; however, AutoN is antagonized when bound to the histone H4 tail, allowing increased ATPase and translocation activity (Hwang et al., 2014) (Fig 1.6 C Step 1). The NegC domain also inhibits translocation activity but is antagonized by the last domain, the HAND-SANT-SLIDE (HSS) domain, which binds to linker DNA between nucleosomes (Fig 1.6 C Step 2) (Grüne et al., 2003). As the HSS domain pulls along the linker DNA to equally space the nucleosomes, it antagonizes NegC and allows the translocation domain to remove DNA from the nucleosome, subsequently sliding the nucleosome along (Fig 1.6 C Step 3). The HSS domain is released when it reaches the end of the linker DNA, allowing NegC to inhibit the translocation domain (Fig 1.6 C Step 4). When the HSS domain runs out of linker DNA and runs into the next nucleosome, the chromatin remodeling enzyme sets the internucleosomal distance and is then released from the nucleosome, ultimately leading to equally spaced nucleosome arrays.

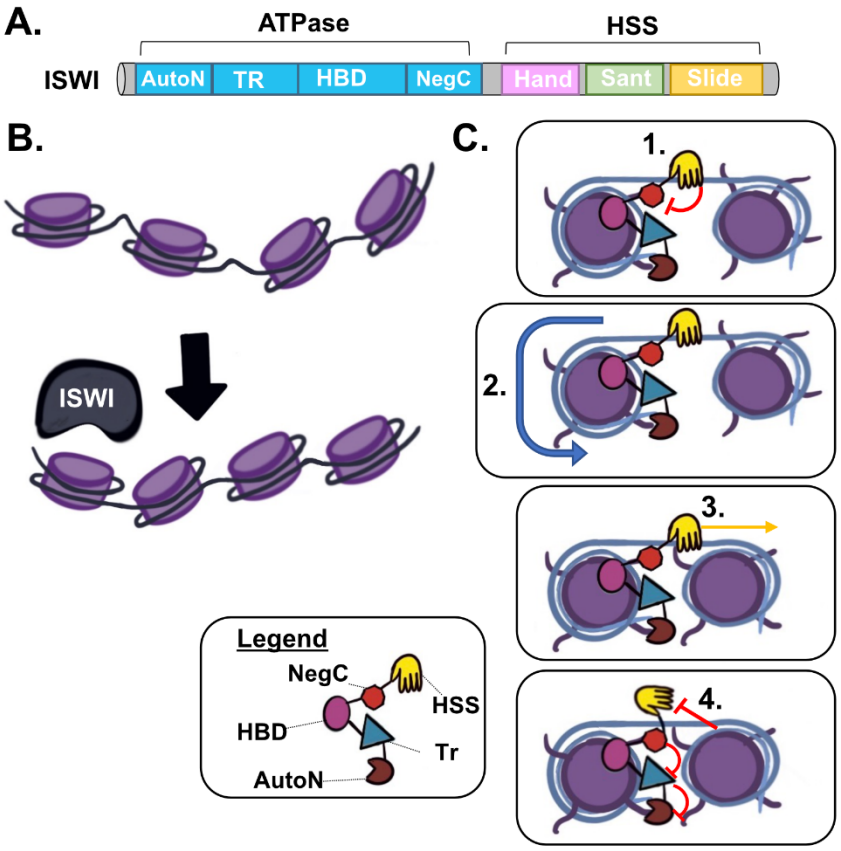

Figure 1.6: SMARCA5 structure and function.

(A) SMARCA5 has an ATPase domain that contains two auto-inhibitory domains (AutoN and NegC), a histone binding domain (HBD), and the translocase domain. The SMARCA5 C-terminal domain is a HAND-SANT-SLIDE domain (HSS). (B) ISWI slides nucleosomes to create equally spaced arrays of nucleosomes. (C) (1) The HSS domain binds the linker DNA between nucleosomes and blocks the NegC domain from inhibiting the translocation domain (Tr). HBD binds to the H4 histone tail, inhibiting the AutoN domain and allowing the Tr domain full activity (2) The HSS domain antagonizes NegC as it pulls along the linker DNA to equally space the nucleosomes, allowing the translocation domain to remove the DNA from the nucleosome and (3) slide it along. (4) Once the HSS domain reaches the end of the linker DNA it is released, allowing NegC to inhibit the Tr domain. The chromatin remodeling enzyme is then released from the nucleosome after setting the precise internucleosomal distance, leading to equally spaced nucleosome arrays.

1.7.3 SMARCA5 complexes

SMARCA5 constitutes the catalytic subunit of five complexes: ATP-utilizing chromatin assembly and remodeling factor (ACF), Chromatin accessibility complex (CHRAC), Remodeling and spacing factor (RSF), Nucleolar remodeling complex (NoRC), and WSTF-imitation switch (B-WICH/WICH) (Table 1.1). The ACF complex was first discovered in 1997 in *Drosophila*; it contained ACF1 and ISWI, which express ATP-dependent chromatin remodeling activity (Ito et al., 1997). The human ACF complex was not purified until 2000 when it was shown to function *in vitro* in nucleosome sliding along with SMARCA5 (LeRoy et al., 2000). A study in 2002 reported that ACF1 and SMARCA5 were necessary for S phase replication after knockdown, specifically within heterochromatin (Collins et al., 2002). Since then, the ACF complex has been shown to be important for DNA double-strand break repair, nucleosome sliding, and G_2/M checkpoint function (Sánchez-Molina et al., 2011).

Varga-Weisz et al. (1997) identified two of the five members of the *Drosophila* CHRAC complex, ISWI and Topoisomerase II, and determined that the CHRAC complex had different remodeling activity than that of the NURF complex (Varga-Weisz et al., 1997). Several years later, the entire human CHRAC complex was identified and reported to consist of SMARCA5, ACF1, p15, and p17, but not Topoisomerase II (Poot et al., 2000). Human CHRAC was determined to have nucleosome spacing activity similar to that of the *Drosophila* CHRAC complex *in vitro*.

The RSF complex consists of RSF1 and SMARCA5, and has a role in facilitating activator-dependent transcription initiation on chromatin template RNAPII *in vitro* (LeRoy et al., 1998). The RSF complex also is required for the assembly of centromeric chromatin

and double-stranded DNA break repair (Helfricht et al., 2013; Loyola et al., 2003; Min et al., 2014; Perpelescu et al., 2009).

The NoRC complex was discovered by identifying the novel Transcription termination factor 1 interacting protein 5 gene (*TIP5*) and its interacting partner SMARCA5 (Strohner et al., 2001). The NoRC complex is recruited to rDNA promoters to mediate gene silencing through heterochromatin formation. NoRC silences rDNA by recruiting different interacting factors to hypoacetylated Histone H4 and methylate H3K9 around the rDNA promoters (Guetg et al., 2010; Santoro et al., 2002; Zhou et al., 2002).

The WICH complex, which contains SMARCA5 and Williams syndrome transcription factor (WSTF), functions in late S phase to control replication through heterochromatin (Bozhenok et al., 2002). In 2006, a complex containing the WICH components SMARCA5 and WSTF, along with Nuclear myosin (NM1), was identified as the B-WICH complex (Cavellán et al., 2006). The B-WICH complex is required to regulate chromatin structure by activating rDNA with RNAPI and RNAPIII (Sadeghifar et al., 2015). Together, these complexes function to specifically alter the chromatin remodeling activity of SMARCA5 in DNA-templated processes such as transcription, replication, and repair (He et al., 2008).

Complex	Components	Functions
ACF	SMARCA5, ACF1	DNA replication and repair
CHRAC	SMARCA5, ACF1, p15, p17	DNA replication
RSF	SMARCA5, RSF1	dsDNA break repair, centromeric chromatin assembly
NoRC	SMARCA5, TIP5	RNAPI transcription
WICH	SMARCA5, WSTF	Replication through heterochromatin
B-WICH	SMARCA5, WSTF, NM1	RNAPI and RNAPIII transcription

Table 1.1: SMARCA5 complexes and their functions.

1.7.4 SMARCA5 in hematopoiesis

In the early 2000s, Stopka and Skoultchi studied the function of SMARCA5 in mammalian development and discovered that it was essential for early murine embryogenesis (Stopka and Skoultchi, 2003). The lethality of *Smarca5* knockout (KO) in early murine development was due to a degeneration of *Smarca5*-null blastocyst at day 2 of embryogenesis (Stopka and Skoultchi, 2003). Not only is Smarca5 important for healthy development of mice, it is also is overexpressed in $CD34^{+}$ AML (Stopka et al., 2000); therefore, these investigators studied the role of SMARCA5 in adult erythropoiesis using K562 human erythroleukemia cells (Stopka and Skoultchi, 2003). Stopka and Skoultchi discovered that *Smarca5* knockdown caused cytokine-induced erythropoiesis blockage *in vitro*. This led to studies of the specific function of *Smarca5* in the hematopoietic system in mouse models. *Smarca5* knockout using Vav1-Cre and Epor-Cre mice resulted in embryonic lethality due to anemia (Kokavec et al., 2017). These results were supported by more recent studies in zebrafish that demonstrated the importance of *Smarca5* in fetal hematopoiesis in zebrafish (Ding et al., 2021b). Bone marrow samples from patients with AML have significantly higher SMARCA5 expression at the time of diagnosis than after complete remission (Zikmund et al., 2020). Together, these data indicate that SMARCA5 has a key role in hematopoietic development.

1.7.5 SMARCA5 functions in transcription and chromatin structure

Numerous studies have investigated SMARCA5 function in chromatin structure and transcription *in vitro*. SMARCA5 slides the nucleosome from 1-2bps in ~20 seconds, and slides nucleosomes in a non-directional manner in isolation *in vitro* (Gamarra et al.,

2018; Oppikofer et al., 2017). Combining ACF1 or WSTF1 with SMARCA5 *in vitro* remodels the nucleosome to the center of an 80 bp DNA fragment, suggesting that these subunits are important for the directionality of SMARCA5 remodeling (Oppikofer et al., 2017).

Smarca5 deletion is embryonic lethal in mice and impairs blastocyst-derived stem cell proliferation, but Barisic *et al.* (2019) used CRISPR Cas9 to generate SMARCA5 knockout (KO) clones in mouse embryonic stem cells (mESCs) with a homozygous frameshift mutation in exon 6 of *Smarca5*. SMARCA5 KO mESCs have normal morphology and a similar cell cycle as WT but lost their differentiation potential and grow about 13% slower than WT. The inability to differentiate suggests that there were specific differences between WT and *Smarca5* KO; therefore, the Schübeler laboratory performed RNA-seq to identify differences in the steady-state mRNA levels. They detected 2000 changes in steady-state mRNA levels that showed enrichment for differentiation and proliferative function.

SMARCA5 was previously described as a “nucleosome ruler” *in vitro*. Barisic *et al*. (2019) tested this using the SMARCA5 KO model. Analyzing nucleosome repeat length (NRL) in combination with MNase-seq data showed an increase in NRL in *Smarca5* KO mESCs. As a control, they analyzed NRL in *Brg1* KO cells, which had no changes in NRL compared to WT. Dr. Michael Stadler created a Loess model to calculate NRL within ATAC-seq data sets, removing NFR fragments and utilizing mono-, di-, and tri-nucleosomal fragments. The ATAC-seq of *Smarca5* KO exhibited an increase in NRL compared to that of WT, with an increase of up to 10 bp around genes and intergenic regions.

Smarca5 knockdown in HELA cells disrupts nucleosome organization around CTCF motifs and reduces human homolog of radiation sensitive mutant 21, RAD21, localization around CTCF motifs (Dluhosova et al., 2014). An association between CTCF and SMARCA5 was found that led to increased PU.1 expression (Dluhosova et al., 2014). SMARCA5 is also associated with RAD21 (Hakimi et al., 2002). *Smarca5* KO in mESCs led to a global decrease in nucleosomes specifically around CTCF motifs but not around REST motifs, which was measured by tracking mononucleosome spacing around motif centers using MNase-seq and by differential motif analysis of NFRs using ATAC-seq. These global changes in nucleosome structure around CTCF motifs led to a loss of CTCF binding in *Smarca5* KO, which was rescued by the addition of WT *Smarca5*. In a different outcome, the mESC KO of *Brg1,* another chromatin remodeling ATPase and member of the SWI/SNF family, led to loss phasing around REST sites and REST binding, but no changes CTCF binding or phasing around CTCF sites (Barisic et al., 2019). *Smarca5* KO genome-wide chromatin capture conformation, Hi-C, data indicated that there was a global loss in topologically associated domains (TADs), similar to what was observed in *Ctcf* KO (Barisic et al., 2019). These combined data provide evidence for the indirect and probably secondary effects of SMARCA5 loss due to the use of knockdown and knockout models.

1.8 Fast biology and proteolysis targeting chimeras (PROTACs)

Genome dynamics within a cell display rapid changes in response to internal or external changes, such as drug treatments, gene knockdown/knockout, or environmental effects. These changes can lead to a chromatin remodeling enzyme sliding a nucleosome along DNA within 20 seconds, leaving spaces of open chromatin and nucleosome-free

regions (NFR) or the alternative, closing regions of chromatin (Fig 1.7) (Gamarra et al., 2018). A TF can find binding sites along the NFRs of DNA in 0.5 to 12 seconds, leading to transcriptional changes (Fig 1.7) (Mazza et al., 2012; Morisaki et al., 2014). These rapid changes, called Fast biology, can alter the genome structure and transcriptional content. The transcriptional changes are accompanied by changes in gene networks and protein expression, leading to activation of secondary targets. This cascade of events can lead to rapid phenotypic changes, such as cell differentiation or apoptosis within one day (Fig 1.7).

To resolve each of these stages of molecular and cellular changes in chromatin structure, experimental techniques have evolved from gene knockdown or knockout (which requires days to establish) to chemical inhibitors. Many of these inhibitors are nonspecific and affect other protein family members. Many other proteins are not considered “druggable” because they function as scaffolds and/or do not contain an ATPase domain. This barrier of undruggable targets is overcome by using proteolysis targeting chimeras (PROTACs), which are heterobifunctional molecules that bind to the protein of interest (POI) and utilize the cell proteasome system [e.g., Cereblon (CRBN) E3 ubiquitin ligase] to ubiquitinate the POI and send it to the degradation pathway. PROTACs enable protein degradation within minutes.

The ability to target E3 ubiquitin ligases for direct degradation of a POI was first described 20 years ago. They designed Protac-1 to recruit methionine aminopeptidase-2 (MetAP-2) to Skp1-Culline-F box–containing Hrt (SCF) for ubiquitination followed by degradation (Sakamoto et al., 2001). This led to increasing interest in creating PROTAC therapies for different diseases. One of the early targets was the androgen receptor to

the E3 ubiquitin ligase MDM2 (Schneekloth et al., 2008). This was followed by the discovery of thalidomide and lenalidomide as heterobifunctional molecules that bind to the CRBN and Ikaros family members IKZF1 and IZKF3 (Ito et al., 2010). IKZF3 is highly upregulated in chronic lymphocytic (CML) and multiple myeloma (MM) patient samples. Treating cells with lenalidomide or thalidomide resulted in anti-cancer effects and led to many clinical trials (Avet-Loiseau et al., 2009; Krönke et al., 2014).

Drug development for customized PROTACs can take years and requires a molecular binding site for the heterobifunctional molecule and CRBN. Therefore, *in vivo* studies have begun by modifying the N-terminal or C-terminal portion of the POI with a degron tag. The dTAG system was developed by utilizing the $FKBP12^{F36V}$ mutant that serves as the tag for the POI and the pocket created by the mutation allows for binding of the degron molecules dTAG-7 and dTAG-13 (Clackson et al., 1998; Erb et al., 2017). This enables incorporation into the endogenous locus of the POI or overexpression of the tagged POI by lentiviral infection followed by knockout of the endogenous copies. Examples of early studies with $FKBP12^{F36V}$ are the lentiviral fusion ENL by James Bradner's group using dTAG-13, followed by endogenous tagging of YY1 by Richard Young's laboratory with the use and creation of dTAG-47 (Fig 1.8) (Erb et al., 2017; Weintraub et al., 2017). These studies have revolutionized biological, biochemical, and genomic analyses of chromatic structure, allowing quick removal of POIs from the system within minutes to hours, versus days to weeks from knockout and knockdown experiments.

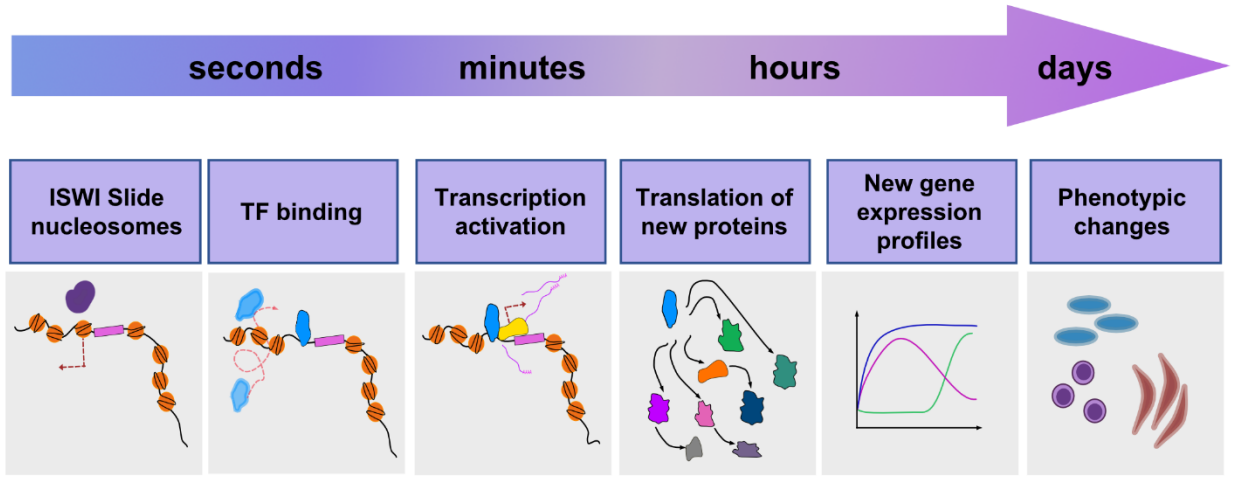

Figure 1.7: Fast biology.

ISWI remodelers and other chromatin remodeling enzymes can slides nucleosomes a few base pairs every 20 seconds. Transcription factors can bind and activate transcription of genes within 20-30 seconds. Translation of protein can occur with a few minutes, all leading to new gene expression profiles and possible phenotypic changes in cells. (Swift and Coruzzi, 2017)

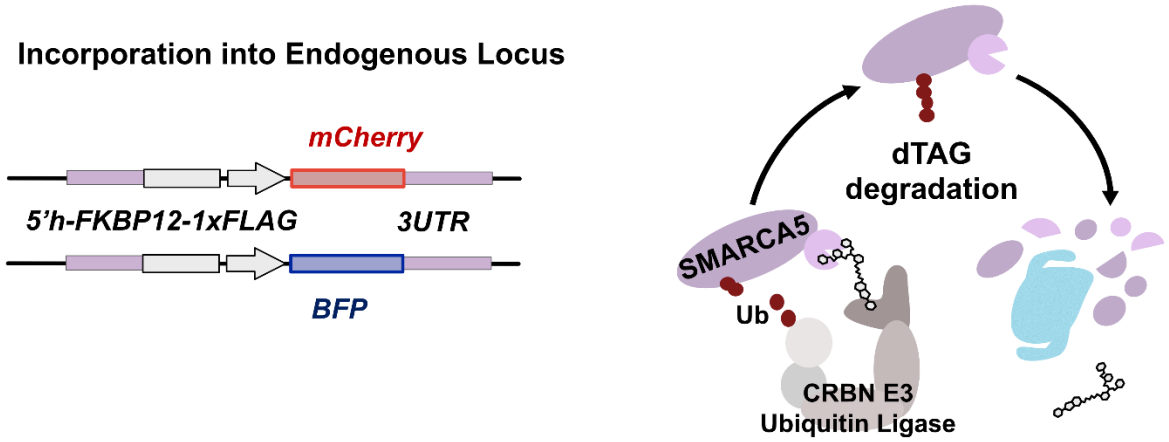

Figure 1.8: Utilizing the dTAG system.

Schematic of homology-directed repair plasmids (HDR) containing the FKBP12^{F36V} FLAG tag that is incorporated into the endogenous alleles of the protein of interest (POI) and bulk sorted based on mCherry and BFP fluorescence. The endogenously tagged POI can be degraded with a heterobifunctional proteolysis targeting chimera (PROTAC) molecule known as dTAG that binds both FKBP12 ^{F36V} and Cereblon E3 ubiquitin ligase, ubiquitinating the POI, and sending it to the proteasome for degradation.

1.9 Scope of dissertation

The subsequent work represents a chapter-by-chapter summation of the dissertation work on Human SMARCA5. Determining the direct role of chromatin remodeling enzymes such as SMARCA5 can allow for a better understanding of how DNA-templated processes occur. In this dissertation I utilize engineered cell lines with degron tags in the endogenous locus of *SMARCA5*, to determine which DNA-templated processes SMARCA5 is directly required for and those that are secondary effects of SMARCA5 knockout.

Chapter 1 presents background on chromatin structure, transcription, replication, SMARCA5 history, and recent development of innovative tools in the field. Chapter 2 presents the Material and Methods used in the studies described in this dissertation. Chapter 3 presents detailed descriptions of which DNA-templated processes were determined to be directly regulated by human SMARCA5, which were published in *Molecular Cell* (Bomber et al., 2022). The combination of short inducible degradation time courses and deep sequencing techniques gave us the temporal resolution required to identify the direct role of SMARCA5 in transcription, protein binding, and overall nucleosome positioning. Chapter 4 presents a discussion of the conclusions and limitations of this studies and future directions for this research moving forward.

2. CHAPTER 2: Materials and Methods

2.1 Antibodies

Antibody	Dilution	Species	Company	Catalog	Use
7AAD	5uL		BD	51-2359KC	Flow
AnnexinV (FITC)	5uL		BD	556419	Flow
BrdU (FITC)	15uL		BD	556025	Flow
CD11b	15uL		BD		Flow
CD82	5uL		BD		Flow
CTCF	1:100	Rabbit	Millipore	07-729	C&R
ETOZnf	1:100	Rabbit	House		C&R
Flag	1:200	Mouse	Sigma	F1804	C&R
H2AZ	1:100	Rabbit	abcam	ab4174	C&R
H3K27me3	1:100		CST	C36B11	C&R
HA	1:800	Rabbit	CST	C29F4	C&R
GAPDH	1:1000	Mouse	SantaCruz		WB
Propium Iodide			BD	51-66211E	Flow
RUNX1	1:100	Mouse	SantaCruz	sc-365644	C&R

Secondary Mouse	1:200		abcam	ab46540	C&R
Secondary Mouse	1:5000		LICOR		WB
Secondary Rabbit	1:200		Invitrogen	31238	C&R
Secondary Rabbit	1:5000		LICOR		WB
SMARCA5	1:1000	Rabbit	Abcam	ab72499	WB
Zombie NIR	1:2000		Biolegend	423106	Flow

Table 2.1: Antibodies

2.2 ATAC-seq

Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells at 5 x10^5/mL were treated with 500nM dTAG-47 at the 2hr, 6hr, and 24hr in duplicate. The nuclei were isolated and incubated with Tn5 following the protocol as previously described (Barnett et al., 2020), adding in 1% *Drosophila* S2 nuclei. The sequencing libraries were created as previously described (Buenrostro et al., 2013). Samples were sequenced by the VANTAGE Sequencing Core on the NovaSeq6000 instrument. Kasumi-1, HEL, and OCI-LY1 *SMARCA5*^{FKBP12F36V}-expressing cells at 5 x10^5/mL were treated with 500nM dTAG-47 at the indicated times in duplicate. The nuclei were isolated and incubated with Tn5 following the Active motif protocol (Active Motif, 53150) adding in 1% *Drosophila* S2 nuclei in duplicate. The sequencing libraries were created and SPRI clean-up was performed with the Active Motif kit per protocol. Samples were sequenced by the VANTAGE Sequencing Core on the NovaSeq6000 instrument. Data were trimmed and analyzed as described in github (https://github.com/monnieb92/SMARCA5_paper), using trimmomatic, bowtie2, and samtools (Bolger et al., 2014; Langmead and Salzberg, 2012; Li et al., 2009). Data were further analyzed as described in github (https://github.com/monnieb92/SMARCA5_paper), using Genrich (v0.6.1, available at https://github.com/jsh58/Genrich, parameters: -q 0.05 -j -r -E) for peak calling, Diffbind for creating a counts table, and DESeq2 with or without batch effect to normalize counts tables and call differential peaks (Love et al., 2014; Ross-Innes et al., 2012; Stark and Brown, 2011; Zhang et al., 2008). ATAC-seq data were visualized using deepTools (Ramírez et al., 2016). HOMER was used to annotate peaks (Heinz et al., 2010).

2.3 Cell Culture

The Kasumi-1 cell line was purchased from ATCC and grown in RPMI supplemented with 20% heat inactivated fetal bovine serum, 1% L-glutamine, and 1% penicillin and streptomycin at 37˚C with 5% CO_2. The HEL cell line was grown in RPMI supplemented with 10% fetal plex, 1% L-glutamine, and 1% penicillin and streptomycin at 37˚C with 5% CO_2. The OCI-LY1 cell line was grown in IMDM supplemented with 10% fetal plex, 1% L-glutamine, and 1% penicillin and streptomycin at 37˚C with 5% CO_2. *Drosophila* S2 cells were gown at 28˚C in Schnieder media supplemented with 10% heat inactivated fetal bovine serum and 0.5% penicillin and streptomycin.

2.4 Cell Cycle Analysis

The Kasumi-1 Parental and Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells were seeded at 3 x 10^5/mL and treated with 500nM dTAG-47 for 24h and 48h time points. BrdU (20mM) was diluted 1:1000 into cell cultures for 1.5 hours at 37˚C 5% CO_2, the cells were washed with PBS, and then fixed overnight with 5 mL of 70% EtOH. Cells were collected by centrifuging at 1500x rpm, then resuspended in 1mL 2N HCl supplemented with 0.5mg/mL Pepsin (Sigma, P7012- 1G). After 30 minutes in 37˚C water bath, the solution was neutralized with 3mL of 0.1M Sodium Tetraborate pH 8.5. The cells were pelleted and washed in 0.5% BSA in PBS, then permeabilized with 1mL of 0.5% BSA/0.5% Tween-20 in PBS. Cells were stained with 15uL of FITC-anti-BrdU (BD, 556025) in 85uL of 0.5% BSA PBS and incubated for 45 minutes at 20˚C in the dark. Cells were washed with 1mL of 0.5% BSA/0.5% Tween-20 in PBS, collected using centrifugation and resuspended in 400uL of PBS with Proprium Iodide (PI) (BD, 51-

66211E) supplemented with RNAse A (Sigma, CAS 9001-99-4) before flow cytometry in which cells were acquired at 400 events per second.

2.5 Cell Proliferation

The Kasumi-1 Parental and Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells were seeded at 3 x 10^5/mL, then treated with 500nM dTAG-47 or DMSO. Cell viability was assessed using Trypan Blue (Gibco, 15250-061) exclusion and counted on a hemocytometer every day for three days. Alamar Blue was also used to assess proliferation by first seeding cells at 3 x 10^5/mL and treating with 500nM dTAG-47 or DMSO. 100uL aliquots of dTAG-47 and DMSO treated cells were put in a 96 well plate in triplicate, then mixed with 10uL of alamar Blue (Invitrogen, DAL1100). The 96 well plate of cells mixed with alamar Blue was then incubated at 37°C and 5% CO_2 for 4 hours, and absorbance measured at 590nm emission wavelength.

2.6 CRISPR Knockin and Electroporation

The homology directed DNA repair (HDR) plasmids were created using Gibson cloning (NEB, E5510S) to assemble the 5' and 3' SMARCA5 homology arms (amplified by PCR from genomic DNA) and the FKBP12 plasmid containing either mCherry or blue fluorescent protein (pAW62.YY1.FKBP.knock-in.mCherry/BFP; #104370/104371 AddGene). Cells were seeded at 6 x 10^5/ml the day before electroporation. The gRNAs complexes were mixed at a ratio of 1.5uL 100uM crRNA_C (IDT, Alt-R® CRISPR-Cas9 crRNA, 2 nmol, Table 2.2), 1.5uL 100uM crRNA_D (IDT, Alt-R® CRISPR-Cas9 crRNA, 2 nmol, Table 2.2), 1.5uL 200uM tracrRNA (IDT, 1072532), 1.5uL R buffer (Neon Kit) and heated at 95°C for 5 minutes and then allowed to cool to 20°C. Ribonucleoprotein (RNP) complexes were formed with 10ug of gRNA complex and 15 ug of Cas9D10A-Nickase

(IDT, 1081062) by incubation at 20˚C for 25 minutes. The Neon System was used to electroporate using a 100uL tip (ThermoScientific, MPK100025) with the following conditions: 1350V, 35ms, 1 pulse (Kasumi-1); 1300V, 20ms, 2 pulses (HEL); 1700V, 20ms, 1 pulse (OCI-LY1) . 1 x 10^6 cells were resuspended in 100uL of R buffer (ThermoScientific, MPK100025) and added to the RNP reaction with 4 ug of mCherry and 4ug of BFP HDR plasmids prior to electroporation. Electroporated cells were placed in antibiotic free media overnight at 37˚C 5% CO_2 and the media replaced 24 hours after electroporation.

gRNA type	Target	Sequence
crRNA_C	SMARCA5	5' GATGGCGCACCTGATGGTCG 3'
crRNA_D	SMARCA5	5' GGTGAAGACTGAAAGGGACAA 3'
tracrRNA	Universal	5’ AGCAUAGCAAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGUGGCACCG AGUCGGUGCUUU 3’

Table 2.2: gRNA sequences

2.7 CUT&RUN and Analysis

2.7.1 CUT&RUN

Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells at 5 x10^5/mL were treated at indicated time points with 500nM dTAG-47. Concanavalin A (Con A) Conjugated Paramagnetic beads were activated by washing 11uL/sample on a magnet with 100uL/sample of Bead Activation Buffer (20mM HEPES pH7.9, 10mM KCL, 1mM $CaCl_2$, 1mM $MnCl_2$) for a total of two washes. Con A Beads were then resuspended in 11uL/sample of Bead Activation Buffer and aliquot 10uL/sample in PCR tubes. Harvested 5x10^5 cells through centrifugation at 500xg for 3 minutes at 20^0C then washed with 100uL/sample Wash Buffer (20mM HEPES pH7.5, 150mM NaCl, 0.5mM Spermidine, 1x Roche complete mini tablet EDTA-free (Sigma, 4693159001)) for a total of two washes. The 100uL of cells was incubated with 10uL of activated ConA beads for 10 minutes at 20˚C, resuspended in 50uL Antibody buffer (Wash Buffer + 0.02% Digitonin, 2mM EDTA) with antibodies at proper concentrations: anti-Flag (1:200, Sigma, F1804), anti-CTCF (1:100, Millipore, 07-729), anti-RUNX1 (1:100, SantaCruz, sc-365644), or anti-ETOZnf (1:100, made in house) at 4˚C overnight on nutator. The next day cells were washed two times with Digitonin Buffer (Wash Buffer + 0.02% Digitonin), while on the magnet. Then the Con A Beads with cells were incubated with 50uL of Antibody buffer containing anti-mouse secondary (1:200, abcam, ab46540) or anti-rabbit secondary (1:200, Invitrogen, 31238) for 1.5 hours at 4˚C as described. Beads were washed two times with 250uL Digitonin Buffer and resuspended in 50uL Digitonin Buffer with 2.5uL of pA/G-MNAse (Epicypher, 15-1116) for 10 min at 20˚C. The Con A Beads with cells were washed two

times with 250uL Digitonin Buffer before being resuspended in 50uL of Digitonin Buffer and the 1mM CaCl2 for 2 hours at 4^0C to activate the MNase. Finally 33uL of Stop Buffer was added and the solutions was incubated at 37°C. The DNA fragments were isolated by removing the solution from the beads and purifying with the NEB Monarch DNA Cleanup Kit (NEB, T1030S) per the manufacture's protocol, eluting with 10uL. Libraries were created with the NEB Next Ultra II DNA Library Prep Kit (NEB, E7645S) per the manufacturer's protocol, skipping the initial size selection, and amplifying for 14 cycles. The samples were sequenced by the VANTAGE Sequencing Core on the NovaSeq 6000 instrument.

2.7.2 Bioinformatics Analysis

Data were trimmed and aligned as described in github: https://github.com/monnieb92/SMARCA5_paper), using trimmomatic, bowtie2, and samtools (Bolger et al., 2014; Langmead and Salzberg, 2012; Li et al., 2009). Bam files were analyzed using MACS2 for peak calling, and Diffbind was used to create counts tables, for the default normalization DESeq2 was used (Love et al., 2014; Stark and Brown, 2011; Zhang et al., 2008). Due to global changes in SMARCA5 and CTCF peaks, these analyses could not use DESeq2 for normalization, so normalization was done using total read counts. CUT&RUN data was visualized using deepTools (Ramírez et al., 2016) and HOMER was used to annotate peaks and perform motif analysis (Heinz et al., 2010).

2.8 Flow Cytometry Analysis

2.8.1 AnnexinV

Kasumi-1, HEL, and OCI-LY1 *SMARCA*^{5FKBP12F36V}-expressing cells were treated for the indicated times with 500nM dTAG-47 and $1x10^6$ cells were spun down at 1100xg

for 5 minutes, washed with PBS, then pelleted again. Cells were resuspend in 1x binding buffer (BD 556454) at $1x10^6$/mL and $1x10^5$/mL stained with 5uL FITC-AnnexinV (BD 556419), Zombie NIR (BioLegend, 423105) 1:2000 (substituting Zombie NIR for propidium iodide), and 10,000 events were acquired on the Flow cytometer.

2.8.2 CD11b

Kasumi-1-*SMARCA*^{5FKBP12F36V}-expressing cells were treated for the indicated times with 500nM dTAG-47 and $1x10^6$ cells were spun down at 1100xg for 5 minutes 4˚C, washed with PBS, then pelleted again. Cells were resuspended in 85uL of PBS with 15uL APC-CD11b (BD 550019) and incubated at 4˚C for 15 minutes. Cells were pelleted at 1100xg for 5 minutes 4˚C then resuspended in 400uL 0.5% BSA in PBS. 10,000 events were acquired on the Flow cytometer.

2.9 Gibson Cloning

The homology directed DNA repair (HDR) plasmids were created using Gibson cloning (NEB, E5510S) to assemble the 5' and 3' SMARCA5 homology arms (amplified by PCR from genomic DNA) and the FKBP12 plasmid containing either mCherry or blue fluorescent protein (pAW62.YY1.FKBP.knock-in.mCherry/BFP; #104370/104371 AddGene). The homology arms were amplified with Phusion (NEB, M0530S), the primers from Table 2.3, and either genomic DNA or plasmids. The PCR reactions were performed per NEB manufacture protocol and PCR products were gel purified (Promega, A9282) before Gibson Assembly.

Fwd/Rev	Sequence (5'-3')	Fragment
Fwd	tgggttgatgtctggtttttccttgaCTGTTAGTCTCTAAGCTTTGATTATTTCAG	5'h gDNA
Rev	gatggtttccacctgcactcctccggatccTAGTTTCAGCTTCTTTTTTCTTCCTCGACC	5'h gDNA
Fwd	ccaccggcggcatggacgagctgtacaagtaaATATGTTTTTGTTTCATAATCACTAACTTTAAACCAG	3UTR gDNA BFP
Fwd	accggcggcatggacgagctgtacaagtaaATATGTTTTTGTTTCATAATCACTAACTTTAAACC	3UTR gDNA mcherry
Rev	tcgagctcggtacccggggatccTGGAATTATAGTCAGTCTTGCTAACAC	3UTR gDNA
Rev	tcaaggaaaaaccagacatcaacccacttg	PlasmidFrag1
Fwd	GTCTCATACCTATTACTTTGTAGTAGTTGT	PlasmidFrag1
Fwd	GGTCGAGGAAGAAAAAAGAAGCTGAAACTAggatccggaggagtgcaggtggaaaccatc	PlasmidFrag2
Rev	AGTTAGTGATTATGAAACAAAAACATATttacttgtacagctcgtccatgccgccggtgg	PlasmidFrag2-mcherry
Rev	AAAGTTAGTGATTATGAAACAAAAACATATttaattaagcttgtgccccagtttgctagg	PlasmidFrag2-BFP

Table 2.3: SMARCA5-FKBP12-1xFlag Gibson Cloning Primers

2.10 MNase-seq

Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells at 5 x10^5/mL were treated with 500nM dTAG-47 at the indicated times in duplicate. 10 million cells were lysed with cell lysis buffer (10mM Tris- HCl pH 7.4, 300mM sucrose, 3mM $CaCl_2$, 2M $MgCl_2$, 1% Triton-X100, 100mM DTT), spun down at 600xg for 3 minutes, placed in reaction buffer (50mM Tris HCl pH8, 5mM $CaCl_2$) and 2 million cell equivalents were incubated with 0.1U/uL MNase (Sigma N3755) for 4 minutes at 20˚C (Chereji et al., 2019). Stop buffer (250mM EDTA pH 8.0, 5% SDS) was added along with 1uL RNAse A (10ug/uL) and 10uL 10% SDS then incubated at 37˚C for 30 minutes. The sample was treated with 1.uL Proteinase K (20mg/mL) and the cells were incubated at 55˚C for 1 hour. The DNA was extracted with phenol::chloroform isoamyl alcohol extractions using Qiagen MaXtract High Density tubes (Qiagen 129046) followed by a chloroform extraction, and repeated for a total of two phenol::chloroform isoamyl alcohol and chloroform extractions. Then the DNA was ethanol precipitated with 1/10th the volume of 3M NaOAc pH 5.5, 1uL glycogen 20mg/mL, and 3x volume of 100% Ethanol. The sequencing libraries were created from gel extracted mono-nucleosomes using the NEB Ultra II DNA Prep Kit (NEB E7645L) with 7 cycles of amplification. Samples were sequenced by the VANTAGE Sequencing Core on the NovaSeq6000 instrument. Data was trimmed and analyzed as described in github (https://github.com/monnieb92/SMARCA5_paper), using trimmomatic, “bowtie2 --very-sensitive”, and samtools (Bolger et al., 2014; Langmead and Salzberg, 2012; Li et al., 2009). MNase-seq data were further analyzed using “deepTools –alignmentseive” to isolate fragment sizes between 140 and 200 bp as well as remove “blacklist peaks” with ENCODE hg19 blacklist peaks, and the data visualized via metaplots using “deepTools -

-MNase" (Ramírez et al., 2016). A previously published loess model algorithm was used to determine nucleosome repeat length (Barisic et al., 2019; Valouev et al., 2011).

2.11 PRO-seq

Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells were seeded at 5 x10^5/mL for a total of 30 million cells per sample. The cells were treated with 500nM dTAG-47 and at the indicated times. Nuclei was isolated by first washing the cells twice with ice-cold PBS. The washed cell pellets were resuspended in 10 mL of swelling buffer (10mM Tris-HCl pH 7.4, 300mM Sucrose, 3mM $CaCl_2$, 2mM $MgCl_2$, and Protease inhibitors) and incubated on ice for 5 minutes. The cells were spun down for 5 minutes at 1000rpm and then resuspended in 1mL of cell lysis buffer without detergent (10mM Tris-HCl pH 7.4, 300mM Sucrose, 3mM $CaCl_2$, 2mM $MgCl_2$, 10% glycerol, and Protease inhibitors) followed by 1mL of cell lysis buffer with detergent (10mM Tris-HCl pH 7.4, 300mM Sucrose, 3mM $CaCl_2$, 2mM $MgCl_2$, 10% glycerol,0.2% Triton X-100 and Protease inhibitors). The cells were incubated on ice for 5 minutes and then dounce homogenized 40 times before moving to a new 15mL conical and adding 5mL of 1x detergent cell lysis buffer (Triton X-100 final concentration 0.1%). Cells were spun down at 1500 x rpm for 5 minutes and washed with 5mL 1x detergent cell lysis buffer. The nuclei were resuspended in 1mL glycerol storage buffer (50mM Tris-HCl pH 8.3, 40% glycerol, 0.1 mM EDTA, 5mM $MgCl_2$, and protease inhibitors) and counted. Nuclei are put at a final concentration of 100uL of glycerol storage buffer per 20 million nuclei and stored at -80˚C.

The following day, the nuclei are thawed on ice to prepare for the nuclear run on, base hydrolysis, bead binding, and 3' adapter ligation. The 100uL of 2x run on reaction (10mM Tris-HCl pH 8, 5mM MgCl, 1mM DTT, 300mM KCl, 375uM ATP, 375uM GTP,

375uM UTP, 375uM Biotin CTP (Perkin-Elmer, cat# NEL542001) , 0.8units/uL SuperseIN RNAse inhibitor (Thermo, cat# AM2694), 1% Sarkosyl) is equilibrated at 30°C for 5 minutes. The 100uL of 20 million nuclei and 0.2 million S2 cells are resuspended into the 100uL 2x run on reaction buffer and incubated for 3 minutes at 30°C. The RNA is then isolated from the nuclear run-ons by resuspending in 600uL Trizol LS (Ambion, cat# 10296010), rotating for 5 minutes at 25°C, adding 160uL of chloroform and centrifuging at 14,000 rpm 4°C for 5 minutes. The aqueous phased is moved to a new tube for ethanol precipitation with 1.5uL of 20mg/mL glycogen and 3x volume of ice cold 100% EtOH placed at -20°C for 10 minutes. The samples were centrifuged for at 4°C and max speed for 30 minutes then washed with 500uL 75% EtOH. The pellets were resuspended in 20uL DEPC H_2O, incubated at 25°C for 5 minutes, then heat denatured for 30 seconds at 65°C preparing for the base hydrolysis. The RNA is treated with 5uL 1N NaOH, incubated on ice for 10 minutes, neutralized with 25uL Tris-HCl pH 6.8, and passed through a P-30 column (BIO-RAD, cat# 7326231) for size selection and clean up. Strepavidin Dynabeads (30uL per sample) are washed once in 1mL of activation buffer (0.1N NaOH, 50mM NaCl), twice 1mL of 100mM NaCl, and resuspended in 50uL/sample of binding buffer (10mM Tris-HCl pH 7.5, 300mM NaCl, 0.1% Triton X-100). Biotinylated RNA is isolated with 50uL of activated and washed Strepavidin M-280 Dynabeads (Invitrogen, cat# 11205D) incubating on rotator for 20 minutes at 25°C. The Strepavidin Dynabeads bound to RNA is then washed in 500uL High salt wash (50mM Tris-HCl pH 7.5, 2M NaCl, 0.5% Triton X-100), 500uL of binding buffer, 500uL of low salt wash (5mM Tris-HCl pH 7.5, 0.1% Triton X-100), followed by two 500uL Trizol (Ambion, cat#

15596018) extractions of just the beads and an ethanol precipitation. The 3' adaptor is ligated by first resuspending the dried RNA pellet in 5uL of 2uM Rev 3 RNA adaptor (5'-GAUCGUCGGACUGUAGAACUCUGAAC-3') and incubating for 30 seconds at 65°C. Then add 5uL of 2x T4 RNA Ligation mix (1uL 10x T4 ligase buffer, 1uL DEPC H_2O, 1uL SuperaseIN, 1uL T4 RNA Ligase I (NEB, cat# M0204S), 1uL10mM ATP) to the RNA, transfer to PCR tubes, and incubate at 20°C for 6 hours then 4°C until the next step or next day.

The second day required a 2nd bead binding and enrichment, 5' de-capping, 5' phosphorylation, and 5' adaptor ligation. Added 40uL of DEPC H_2O to the RNA 3' adaptor mixture, and repeated a second bead binding, Trizol extraction, and ethanol precipitation. Dissolved the RNA pellet in 5uL of DEPC H_2O and incubated for 30 seconds at 65°C to denature to prepare for de-capping. Added 5uL of CAP CLIP Mix (3uL DEPC H_2O, 1uL 10x CAP CLIP buffer, 1uL SuperaseIN, 0.5uL CAP CLIP (CellScript, cat# C-CC15011H)) to the RNA and incubated at 37°C for 2 hours. Performed a Trizol extraction and ethanol precipitation. Redissolved the RNA pellet in 5uL of DEPC H_2O and incubated for 30 seconds at 65°C to denature to prepare for 5' hydroxyl repair. Added 5uL of PNK mix (2uL DEPC H_2O, 1uL 10x PNK buffer, 1uL 10mM ATP,0.5uL SuperaseIN, 0.5uL T4 PNK (NEB, cat# M0201S)) to the 5uL of RNA and incubated at 37°C for 1 hour. Performed another Trizol extraction and ethanol precipitation. Started the 5' adaptor ligation by dissolving the modified RNA in 5uL of 2uM VRA5 (5'-CCUUGGCACCCGAGAAUUCCA-3'), incubating for 30 seconds at 65°C, adding 5uL of T4 RNA ligase mixture (1uL 10x T4 ligase buffer, 1uL DEPC H_2O, 1uL SuperaseIN, 1uL T4 RNA Ligase I (NEB # M0204S),

1uL10mM ATP), and incubating overnight for 6 hours at 20°C then 4°C for the rest of the night.

The third day required a 3rd bead binding and enrichment, reverse transcription, library amplification and page purification. Brought the samples to 50uL by adding 40uL DEPC H_2O then completed the bead binding, Trizol extraction, and ethanol precipitation. Diluted the RNA pellet in 5uL of 5uM RP1 primer (5'- AATGATACGGCGACCAC CGAGATCTACACGTTCAGAGTTCTACAGTCCGA-3), dissolved for 5 minutes at 25°C, and denatured for 30 seconds at 65°C. Next added 4uL of the reverse transcription reaction (2uL 5x first strand synthesis buffer, 1uL 100mM DTT, 0.5uL dNTP (12.5mM each), 0.5uL SuperaseIN) to each sample, incubated in thermocycler at 48°C for 3 minutes before returning to ice and adding 1uL of Superscript II Reverse Transcriptase (Invitrogen, cat # 18080044) for the reverse transcription reaction (44°C for 20 minute, 52°C for 45 minutes, and then 4°C forever). For the full scale library amplification, 6uL DEPC H_2O was added to the reverse transcription reaction and 33uL PCR library mixture (10.5uL DEPC H_2O, 10.5uL 5x HF Buffer, 10uL 5M betaine, 1uL 12.5uM RP1 primer, 1uL dNTP (12.5mM each), 1uL Phusion polymerase (NEB, cat# M0530S)) was added to each sample along with 1uL 12.5uM RPIx of an unique Illumina library barcode for each sample. The thermocycler was ran as follows: 95°C for 2 min; 21 cycles of 95°C for 30 sec, 56°C for 30 sec, 72°C for 30 sec; 72°C for 3 min. The samples were brought to a final volume of 200uL for an ethanol precipitation. The cDNA library pellets were resuspend in 10uL of DEPC H_2O then ran on 8% native polyacrylamide gels, with two samples per gel and stained with SYBER Gold (Invitrogen, cat# S11494). The gels were

imaged and the cDNA library fragments between ~150bp and ~350bp were extracted and placed in eppendorf tubes. The cDNA was gel extracted by crushing the gel with a pestle, adding 1mL soaking buffer (TE (38.8 mL),150 mM NaCl (1.2 mL 5M NaCl), 0.02% Tween-20 (8 uL)), and incubating overnight with agitation at 37°C.

On the fourth and final day samples were spun on a tabletop centrifuge at top speed for 5 minutes at 25°C, the supernatant was kept in a new tube, and the gel was resuspended in 400uL of soaking buffer. The gel samples were incubated again with agitation at 37°C for 4 hours then spun at top speed for 5 minutes at 25°C. The samples were passed through Costar Spin-X centrifuge tube filters (Corning, cat# 8162) to remove debris, spinning at 4000 rpm on a tabletop centrifuge for 5 minutes. Finally the samples were extracted with phenol:chloroform:isoamyl alcohol, ethanol precipitation, and resuspended in 10uL 1mM EDTA. The libraries were sent to VANTAGE for sequencing on the Illumina NextSeq500.

2.12 RNA-seq

Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells at 5 x10^5/mL were treated with 500nM dTAG-47 at the indicated times in duplicate. RNA was extracted from 5 x 106 Kasumi-1 cells using 0.75mL Trizol (Invitrogen, 15596026) and incubated at 20˚C for 10 minutes. The RNA was then further extracted by adding 0.15mL chloroform, vortexed on high for 1 minute, and centrifuged at 16,000 x G for 15 minutes to separate the aqueous phase. The RNA was precipitated using isopropanol. DNA was removed using DNAse I (Invitrogen, 18068-015) degradation and the RNA collected using ethanol precipitation. The RNA-seq libraries were created and sequenced by the VANTAGE Sequencing Core (Vanderbilt) on the NovaSeq6000 instrument.

Analysis was performed as previously described (Trapnell et al., 2012). The heat maps were clustered using hierarchical ward.D2 method (Murtagh and Legendre, 2014).

2.13 Synchronized Cell Cycle

Parental Kasumi-1 and Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells were seeded at 4 x 10^5/mL then treated for 18hrs at 37˚C 5% CO_2 with thymidine (Sigma, T9250-10G) at a final concentration of 2mM. The cells were washed with pre-warmed PBS then placed in RPMI media supplemented with 20% heat inactivated fetal bovine serum, 1% L-glutamine, and 1% penicillin and streptomycin for 9hrs at 37˚C 5% CO_2. The cells were treated with a second thymidine block for 18hrs at 37˚C 5% CO_2 for a final concentration of 2mM thymidine. Before releasing the cells from the double thymidine block, they were or treated for 2hrs with DMSO or 500nM dTAG-47. Cells were released from G_1/G_0 by washing with pre-warmed PBS and incubating with fresh media as previously describes. Cells were then incubated with 1:1000 BrdU (20mM) , collected, fixed, stained, and analyzed at the indicated time points (0h, 2h, 4h, 6h, 8h, 12h, 24h) following the Cell Cycle Analysis methods section mentioned above.

2.14 Western Blot

Cells were washed with cold phosphate buffered saline (PBS), lysed with a buffer containing 50mM Tris-HCl pH 7.4, 150mM NaCl, 0.1% SDS, 0.5% DOC, 1% NP40, sonicated briefly on ice, and the lysates were clarified by centrifugation. The protein concentrations were measured with the Lowry assay Kit (BioRad, 5000113, 5000114, 5000115), measuring a 96 well at 750nM. The 100ug of protein lysates were mixed with 4X SDS running buffer, boiled, and loaded on an 8% polyacrylamide gel. The gel was run at 90 volts for 2-3 hours and transferred overnight at 4^0C onto PDVF activated with

methanol. Immunoblotting was performed with anti-SMARCA5 (1:1000 abcam, 72499), anti-GAPDH (1:1000 SantaCruz, sc-365062), anti-FLAG (1:500 Sigma M2Flag, F1804) in blocking buffer (5% BSA in PBST(PBS with 0.1% Tween-20)). The membrane was visualize using fluorescently tagged secondary antibodies from LICOR Biosciences and visualized with the Odyssey scanner.

2.15 Synchronized Cell Cycle ATAC-seq

Kasumi-1-*SMARCA5*^{FKBP12F36V}-expressing cells were synchronized with a double thymidine block as described above. Before releasing the cells from the double thymidine block, they were treated for 6hrs with DMSO or 500nM dTAG-47. Then performed ATAC-seq using the Active Motif Kit as previously described in 2.2.

3. CHAPTER 3: Results

3.1 Introduction

The basic unit of chromatin is the nucleosome, which consists of approximately 147 bp of DNA wrapped around a histone octamer core(Kornberg, 1977; Luger et al., 1997). This allows for both the efficient packaging of all the genetic material within the nucleus and mediates the accessibility of the DNA to other regulatory proteins, including those controlling transcription, DNA replication and DNA repair. Neighboring nucleosomes are typically separated by 20-50 bp of linker DNA, which is more easily accessible to regulatory proteins and can also be bound by structural proteins, such as histone H1, to promote the formation of higher order chromatin structures(Bakayev and Georgiev, 1976; S. Baldi et al., 2020; Thoma et al., 1979). Chromatin remodeling enzymes hydrolyze ATP and slide or evict nucleosomes to modify chromatin structure and accessibility, as well as cooperate with histone chaperones to facilitate histone variant exchange(Hirschhorn et al., 1992; Mizuguchi et al., 2004; Schwartz and Ahmad, 2005). Therefore, through complex mechanisms governing chromatin dynamics, these critical enzyme complexes have the ability to regulate all DNA-templated cellular processes.

Smarca5 (SNF2H) along with *Smarca1* (SNF2L) encode the two ATPase components of the imitation switch (ISWI) chromatin remodeling complex and participate in multiple distinct ISWI subcomplexes. SNF2L associates with BPTF and RbAp46/48 in the NURF complex, while SNF2H is a component of CERC2-containing remodeling complex (CERF), ATP-utilizing chromatin assembly factor (ACF), chromatin accessibility complex (CHRAC), nucleolar remodeling complex (NoRC), Williams Syndrome

Transcription Factor (WSTF)-ISWI chromatin remodeling complex (WICH), and remodeling and spacing factor (RSF)(Barak et al., 2003; Cavellán et al., 2006; Hamiche et al., 1999; Ito et al., 1997; Kukimoto et al., 2004; Loyola et al., 2003; Strohner et al., 2001; Yang et al., 2006; Zhou et al., 2002). ISWI complexes have been implicated in the control of multiple processes including DNA replication, DNA repair, and transcription. RNAi mediated knockdown of SMARCA5 impaired DNA replication and S-phase progression, and the ISWI complexes in particular were important for replication through heterochromatin(Bozhenok et al., 2002; Collins et al., 2002; Guetg et al., 2010; Santoro et al., 2002). Similarly, the WICH-ISWI complex associated with PCNA and localized to sites of DNA replication (Poot et al., 2004). In addition, multiple ISWI complexes have been implicated in the DNA damage response and double strand break repair(Aydin et al., 2014; Min et al., 2014; Poot et al., 2004; Varga-Weisz et al., 1997). Finally, the RSF-ISWI complex is required for the initiation of transcription from chromatinized DNA templates *in vitro*(LeRoy et al., 1998), suggesting that SMARCA5-containing complexes may directly regulate gene expression. In fact, defects in transcriptional control were proposed to underlie a number of phenotypes observed in *Smarca5*-deficient mouse and zebrafish models (Ding et al., 2021a, 2021c).

Unlike the related SWI/SNF chromatin remodeling complexes, the ISWI complex cannot evict nucleosomes, rather it slides them to maintain appropriately spaced nucleosome arrays (Aydin et al., 2014; Dao et al., 2020; Deindl et al., 2013; Gamarra et al., 2018; Grüne et al., 2003; Levendosky and Bowman, 2019). Deletion of *Smarca5* from mouse embryonic stem cells revealed that it was critical for the maintenance of appropriate nucleosome repeat length (NRL) (Barisic et al., 2019). As the length of the

core-associated DNA is static (147 bp), changes in nucleosome repeat length reflect a lengthening or shortening of the linker DNA. Interestingly, NRL correlates with specific genomic features. For example, the gene body of highly expressed genes tends to have a relatively short nucleosome repeat length, which may prevent nucleosomal packing, while heterochromatin is associated with a longer nucleosome repeat length, which may allow nucleosomal stacking and compaction[3]. In addition, nucleosome positioning can be variable at some genomic features, but well defined at others. For instance, chromatin insulators bound by CTCF and transcription start sites (TSS) of active genes exhibited more rigid nucleosome phasing patterns (S. Baldi et al., 2020; Barisic et al., 2019; Clarkson et al., 2019). It is likely that these unique nucleosomal patterns have both structural and regulatory consequences.

In addition to changes in nucleosome repeat length and consistent with previous studies, *Smarca5* deletion in mouse embryonic stem cells resulted in a loss of CTCF binding and disruption of topology associated domain (TAD) structure, suggesting that Smarca5 is required for CTCF binding and function (Barisic et al., 2019). SMARCA5 also associates with CTCF and RAD21 in co-immunoprecipitation assays, further connecting SMARCA5 with CTCF function (Dluhosova et al., 2014; Hakimi et al., 2002; Song et al., 2022). However, due to the static nature of the knockout studies, the exact mechanism by which Smarca5 altered CTCF binding could not be addressed. These results suggested that CTCF could direct SMARCA5 to slide nucleosomes to organize chromatin at its DNA binding sites. In contrast, a degron tagged version of the only ISWI member in *Drosophila* S2 cells caused changes in nucleosome positioning only around “housekeeping” genes whose expression was only modestly affected in the first 6hr after

degradation (Hendy et al., 2022). These results may reflect fundamental differences between the *Drosophila* cell line and mammalian cells. For example, there is a profound difference in the requirement of CTCF in *Drosophila* versus vertebrates, as deletion of CTCF in *Drosophila* had only modest effects on genome architecture(Kaushal et al., 2021).

While germline deletion of *Smarca5* caused embryonic lethality, tissue-specific deletion also suggested a profound role in normal tissue development and differentiation, which was associated with large changes in gene expression, as well as proliferation defects (Stopka et al., 2000; Stopka and Skoultchi, 2003; Zikmund et al., 2019). These genetic deletion models helped define the role of Smarca5 in biology and development, yet they cannot effectively distinguish direct and indirect chromatin effects, and thus have not allowed a mechanistic examination of Smarca5 function in cells. In contrast, *in vitro* biochemical reconstitution assays have provided details on how these chromatin remodeling enzymes bind to DNA and hydrolyze ATP to move nucleosomes (Gamarra et al., 2018; Sabantsev et al., 2019). In fact, single-molecule analysis of SMARCA5 nucleosomal sliding indicated that a 1-2 base pair step (or sub-step) occurs within 10-30 seconds, highlighting the dynamic nature of nucleosomal positioning (Gamarra et al., 2018; Sabantsev et al., 2019). These *in vivo* versus *in vitro* analyses highlight the challenge of defining the contribution of nucleosome remodeling complexes to gene expression, DNA replication and cell cycle progression.

Here, we used CRISPR-Cas9^{D10A} and homology-directed DNA repair to modify the endogenous allele of *SMARCA5* to make it sensitive to a small molecule proteolysis targeting chimera (PROTAC), dTAG-47, to begin to close this gap (Erb et al., 2017;

Layden et al., 2021). We modified the endogenous *SMARCA5* loci in acute myeloid leukemia cells (Kasumi-1), erythroleukemia cells (HEL), and a diffuse large B-cell lymphoma cell line (OCI-LY1). By using a small bi-functional molecule to rapidly degrade SMARCA5 we were able to show that SMARCA5 had only modest effects on nascent transcription or cytoplasmic pools of mRNA within the first 6hr after degradation. Conversely, ATAC-seq and MNase-seq identified rapid changes in nucleosome repeat length throughout the genome. Moreover, these effects were observed in synchronized cell cultures at the G_1/S phase and in the mid to late S phase of the cell cycle. This was particularly acute near CTCF DNA binding sites, where CTCF was rapidly lost from the genome. These results suggest that nucleosome positioning is dynamic throughout the cell cycle and might suggest that SMARCA5 plays a role as a counterweight to other nucleosome sliding events in the cell.

3.2 Results

3.2.1 Engineering an endogenous *SMARCA5-FKBP12^{F36V}* allele for inducible degradation

Chromatin-remodeling enzymes are the subject of intense drug discovery efforts, as they could be therapeutic targets in cancer, including acute leukemia (Stopka et al., 2000; Zikmund et al., 2020), and other chronic disease syndromes (Erdel and Rippe, 2011; Roberts and Orkin, 2004). Small molecule inhibitors of chromatin remodeling enzymes are also invaluable tools for the discovery of the mechanism of action of these enzymes, as they allow rapid enzyme inhibition. Because a selective small molecule inhibitor was not readily available for SMARCA5, we used CRISPR-Cas9^{D10A} nickase to induce two ssDNA breaks at the end of the coding region of *SMARCA5* and homology-directed repair to insert *FKBP12^{F36V}-FLAG* to generate a C-terminal fusion protein in the t(8;21)-containing acute myeloid leukemia (AML) cell line Kasumi-1, the human erythroleukemia cell line (HEL), and a diffuse large B-cell lymphoma cell line (OCI-LY1) (Fig 3.1A) (Weintraub et al., 2017). This approach yielded an endogenous SMARCA5 protein that was rapidly degraded upon treatment with the PROTAC, dTAG-47 (Fig 3.1A) (Erb et al., 2017; Weintraub et al., 2017). We performed RNA-seq analysis on the parental and CRISPR-edited Kasumi-1 cells to ensure that the addition of the degron tag alone did not dramatically alter gene expression. We observed few changes in gene expression between parental and SMARCA5-FKBP12^{F36V}-FLAG cell lines with a Pearson correlation coefficient of the $\log_2$ (FPKM) between these cells of 0.9965 (Fig 3.2A).

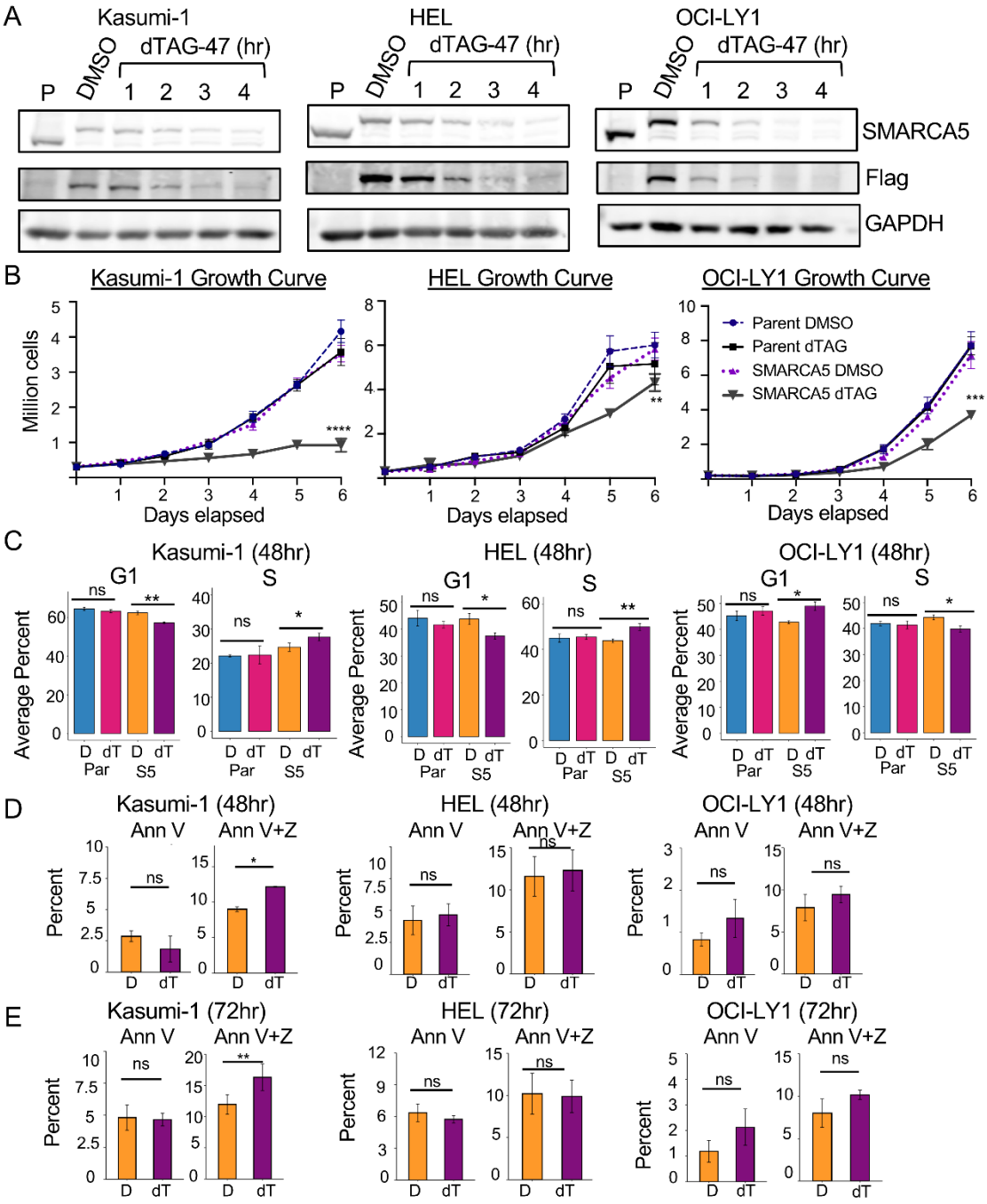
A
Kasumi-1
HEL
OCI-LY1
DMSO
dTAG-47 (hr)
P
1 2 3 4
SMARCA5
Flag
GAPDH
B
Kasumi-1 Growth Curve
HEL Growth Curve
OCI-LY1 Growth Curve
Million cells
Days elapsed
Parent DMSO
Parent dTAG
SMARCA5 DMSO
SMARCA5 dTAG
C
Kasumi-1 (48hr)
HEL (48hr)
OCI-LY1 (48hr)
G1
S
Average Percent
D dT D dT
Par S5
D
Kasumi-1 (48hr)
HEL (48hr)
OCI-LY1 (48hr)
Ann V
Ann V+Z
Percent
E
Kasumi-1 (72hr)
HEL (72hr)
OCI-LY1 (72hr)

Figure 3.1: Selective degradation of endogenously tagged SMARCA5 leads to decreased cell growth.

Degradation time course showing rapid loss of SMARCA5 in Kasumi-1, HEL, and OCI-LY1 cells (Parental and SMARCA5$^{FKBP12F36V\text{-}Flag}$) after 500nM dTAG-47 treatment. Western blot of SMARCA5, Flag, and loading control, GAPDH. (B) Cell growth curve of Kasumi-1, HEL, and OCI-LY1 Parental versus SMARCA5-FKBP12^{F36V}-FLAG cells over a 6-day period after dTAG-47 treatment. Cell viability was determined using Trypan Blue exclusion (n=3). (C) Bar graphs of flow cytometry analysis of BrdU incorporation versus propidium iodide (PI) staining of Kasumi-1, HEL, and OCI-LY1 Parental versus SMARCA5-FKBP12^{F36V}-FLAG-expressing cells after 48h dTAG-47 treatment (n=3). Two-way repeat measure ANOVA: **** = $p<0.0001$; *** = $p<0.0007$; ** = $p<0.0014$. (D, E) Bar graphs of flow cytometry analysis of Annexin V staining vs Zombie-NIR staining of Kasumi-1 (n=5), HEL (n=5), and OCI-LY1 (n=3) SMARCA5-FKBP12^{F36V}-FLAG-expressing cells after 48h (D) or 72h (E) dTAG-47 treatment. D, DMSO; dT, dTAG-47; P-values (Welch two-tailed *t* test) are indicated by ns = $p > 0.05$, * = $p <0.05$, ** $p< 0.001$. Error bars show standard deviation.

Both germline and *Vav-Cre*-mediated *Smarca5* deletion were embryonic lethal in mice (Kokavec et al., 2017; Stopka and Skoultchi, 2003). Therefore, we first tested the effect of SMARCA5 degradation on cell growth and viability (Fig 3.1B). In Kasumi-1 cells, cell growth slowed 2-3 days after the addition of dTAG-47, while SMARCA5 degradation in HEL and OCI-LY1 cell lines did not cause a decrease in growth until day 5 (Fig 3.1B). The decreased cell growth in Kasumi-1 cells was associated with an increase in CD11b expression, which is indicative of myeloid differentiation (Fig 3.2B). There were also subtle changes in BrdU incorporation, suggesting that SMARCA5 may have contributed to efficient transit through the S phase in the Kasumi-1 and HEL cell lines (Fig 3.1C and 3.2C), which was previously observed after *SMARCA5* knockdown (Bhaskara et al., 2013; Poot et al., 2000). The OCI-LY1 cell line showed a small decrease in the percentage of cells in the S phase with an accumulation in G_1. Degradation of SMARCA5 in Kasumi-1 cells led to a statistically significant increase in Zombie NIR/AnnexinV double positive cells at 48hr and 72hr following dTAG-47 treatment, while HEL and OCI-LY1 cells did not display increased cell death upon SMARCA5 degradation (Fig 3.1D, 3.1E and 3.2D), consistent with the more dramatic effect on cell growth in Kasumi-1 cells (Fig 3.1B).

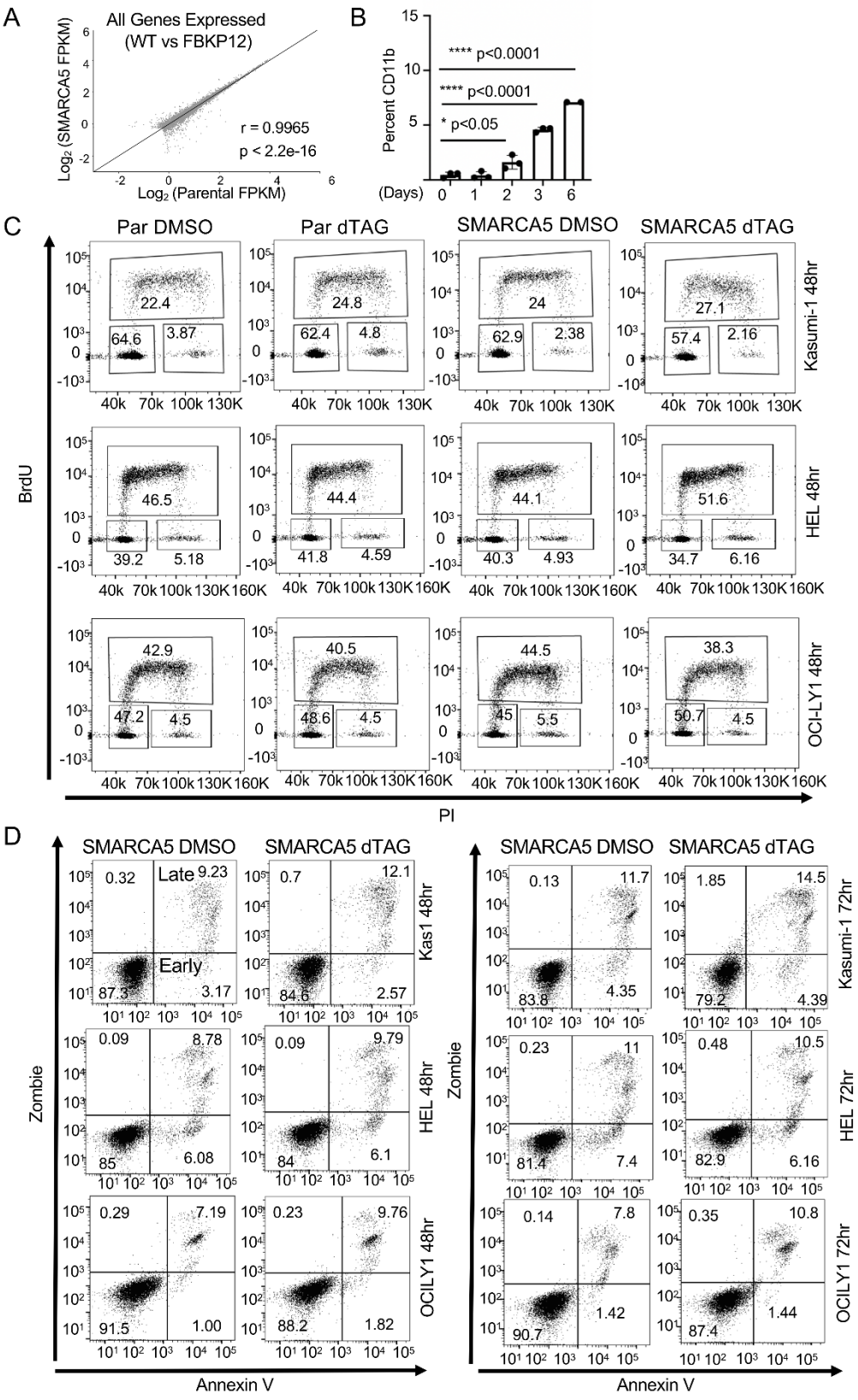
A
All Genes Expressed
(WT vs FBKP12)
Log2 (SMARCA5 FPKM)
Log2 (Parental FPKM)
r = 0.9965
p < 2.2e-16
B
Percent CD11b
**** p<0.0001
**** p<0.0001
* p<0.05
(Days) 0 1 2 3 6
C
Par DMSO
Par dTAG
SMARCA5 DMSO
SMARCA5 dTAG
Kasumi-1 48hr
HEL 48hr
OCI-LY1 48hr
BrdU
PI
D
SMARCA5 DMSO
SMARCA5 dTAG
Late
Early
Kas1 48hr
HEL 48hr
OCILY1 48hr
Kasumi-1 72hr
HEL 72hr
OCILY1 72hr
Zombie
Annexin V

Figure 3.2: Characterization of SMARCA5-FKBP12^{F36V}–Flag tagged cell lines.

(A) Scatter plot of log_2 normalized counts from RNA-seq comparing the log_2 transformation of FPKM of all genes expressed in Kasumi-1 Parental versus Kasumi-1 SMARCA5-FKBP12^{F36V}-FLAG- expressing cells. (B) Quantification of CD11b expression by flow cytometry as a marker of myeloid differentiation of parental Kasumi-1 or SMARCA5-FKBP12^{F36V}-FLAG–expressing cells at 1, 2, 3, and 6 days after SMARCA5 degradation (n=3). (C) Flow cytometry analysis of BrdU incorporation versus propidium iodide (PI) staining of Kasumi-1, HEL, and OCI-LY1 Parental versus SMARCA5-FKBP12^{F36V}- FLAG-expressing cells at 48hr following SMARCA5 degradation. (D) Flow cytometry analysis of Annexin V-FITC staining and Zombie-NIR staining of Kasumi-1, HEL, and OCI-LY1 SMARCA5- FKBP12^{F36V}-FLAG cells 48hr and 72hr after degradation of SMARCA5 compared to DMSO control samples.

3.2.2 Genomic localization of SMARCA5

When engineering the *SMARCA5-FKBP12*^{F36V} allele, we incorporated a FLAG epitope tag to simplify the analysis of endogenous SMARCA5. We used anti-FLAG in cleavage under targets and release using nuclease (CUT&RUN) analysis in the presence or absence of dTAG-47-mediated degradation to localize SMARCA5 throughout the genome (Skene et al., 2018). We identified 9,650 SMARCA5 peaks in Kasumi-1, 12,048 in HEL, and 8,546 in OCI-LY1 that were reduced by at least 1.5-fold following dTAG-47 treatment (Fig 3.3A). These changed peaks were enriched in promoters, intronic, and intergenic regions (Fig 3.4), and many of these sites overlapped between cell lines (Figure 3B). For example, the region around *SERTAD1* shows multiple conserved SMARCA5 (FLAG) peaks (Fig 3.3C). The consensus binding site for CTCF and the related brother of the regulator of imprinted sites, BORIS were commonly enriched under the changed peaks within each cell line (Figure 3.3D). These results are consistent with ChIP-seq data from mESCs demonstrating frequent colocalization of CTCF and SMARCA5 (Song et al., 2022).

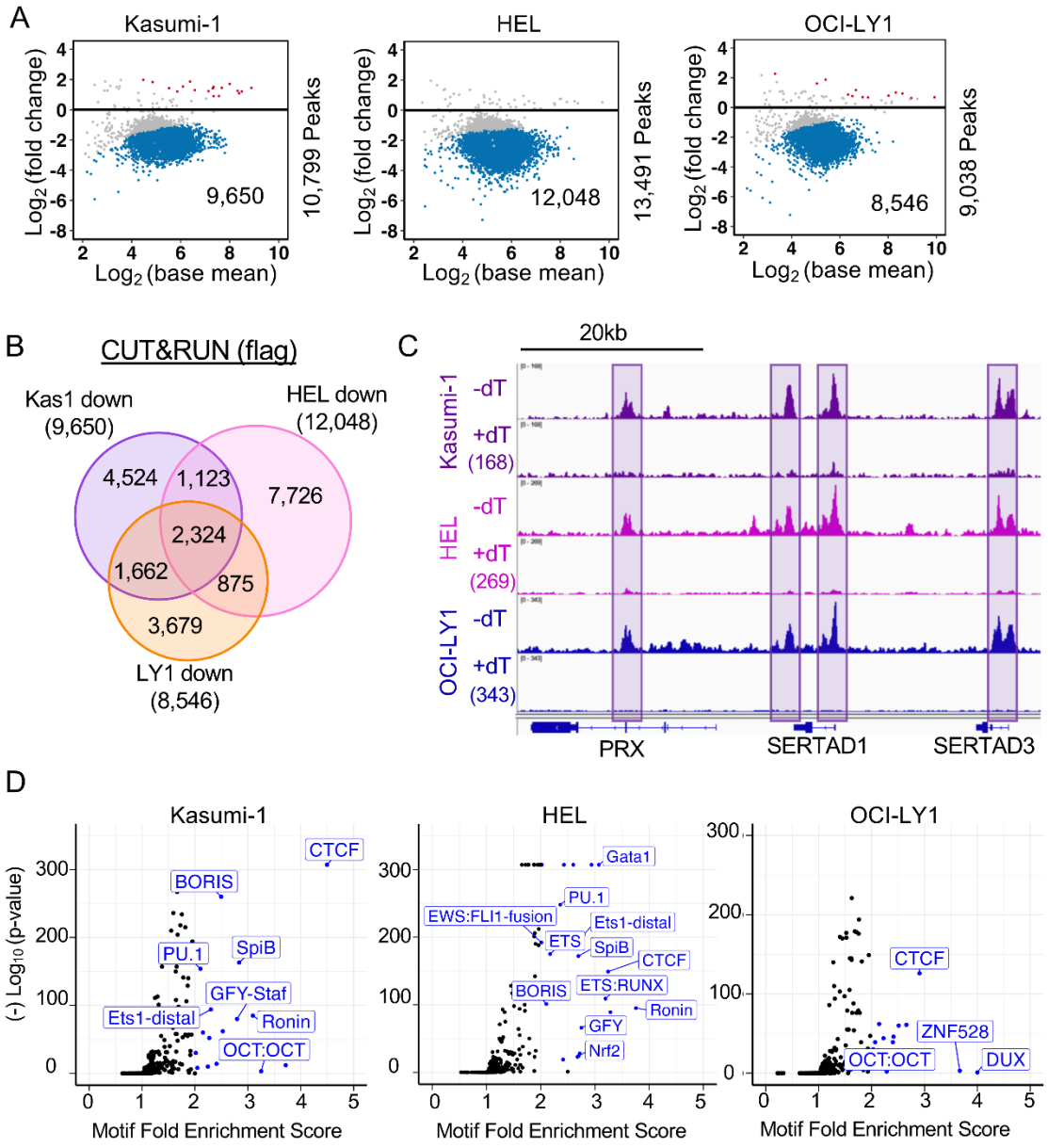

Figure 3.3 . Genomic localization of endogenously tagged SMARCA5-FKBP12F36V-FLAG.

(A) MA plots of the changes in peaks upon degradation of SMARCA5-FKBP12^{F36V}-FLAG. The blue dots represent peaks down-regulated at least 1.5-fold. (B) Venn diagram showing the overlap between the SMARCA5 CUT&RUN decreased peaks from each of the three cell lines. (C) IGV browser tracks showing an example of peaks conserved in all three cell lines. (D) Volcano plots of motif enrichment scores (percent enrichment/percent background) of the significantly down-regulated peaks versus the Log_{10} (p-value). Blue dots represent motifs enriched by 2-fold.

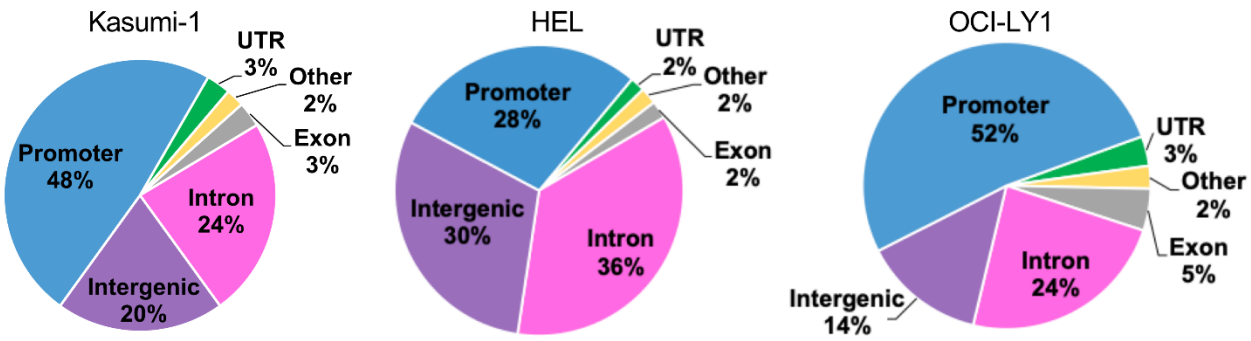

Figure 3.4: Genomic annotations of the SMARCA5 peaks using nearest neighbor analysis.

3.2.3 SMARCA5 loss affects CTCF localization

Smarca5 deletion in mouse ES cells resulted in a loss of CTCF binding (Barisic et al., 2019). Therefore, we performed CUT&RUN to assess the ability of CTCF to bind DNA during a time course of SMARCA5 degradation (Fig 3.5A-C). Remarkably, the intensity of all the CTCF binding was reduced following SMARCA5 degradation, with the observed decrease beginning at 2hr. While the number of peaks significantly reduced within 2hr of dTAG-47 treatment varied among the three cell lines, all the peaks were trending lower and by 6hr there was a dramatic and significant loss of CTCF DNA binding (Fig 3.5A-C). Although these cell types represent different hematopoietic lineages, there was a great deal of conservation among the CTCF sites that were lost at 6hr (Fig 3.5D).

Given that it appeared that all CTCF binding sites were changing over the time course of SMARCA5 degradation, we overlapped all CTCF peaks with the SMARCA5 peaks, and found a 70-80% overlap of CTCF with SMARCA5 peaks (Fig 3.5E-G). For example, in the region around *MMP28,* there were both SMARCA5 and CTCF peaks that were lost upon degradation of SMARCA5 in all three cell lines (Fig 3.5H). Thus, it appears that degradation of SMARCA5 caused a rapid loss of CTCF DNA binding.

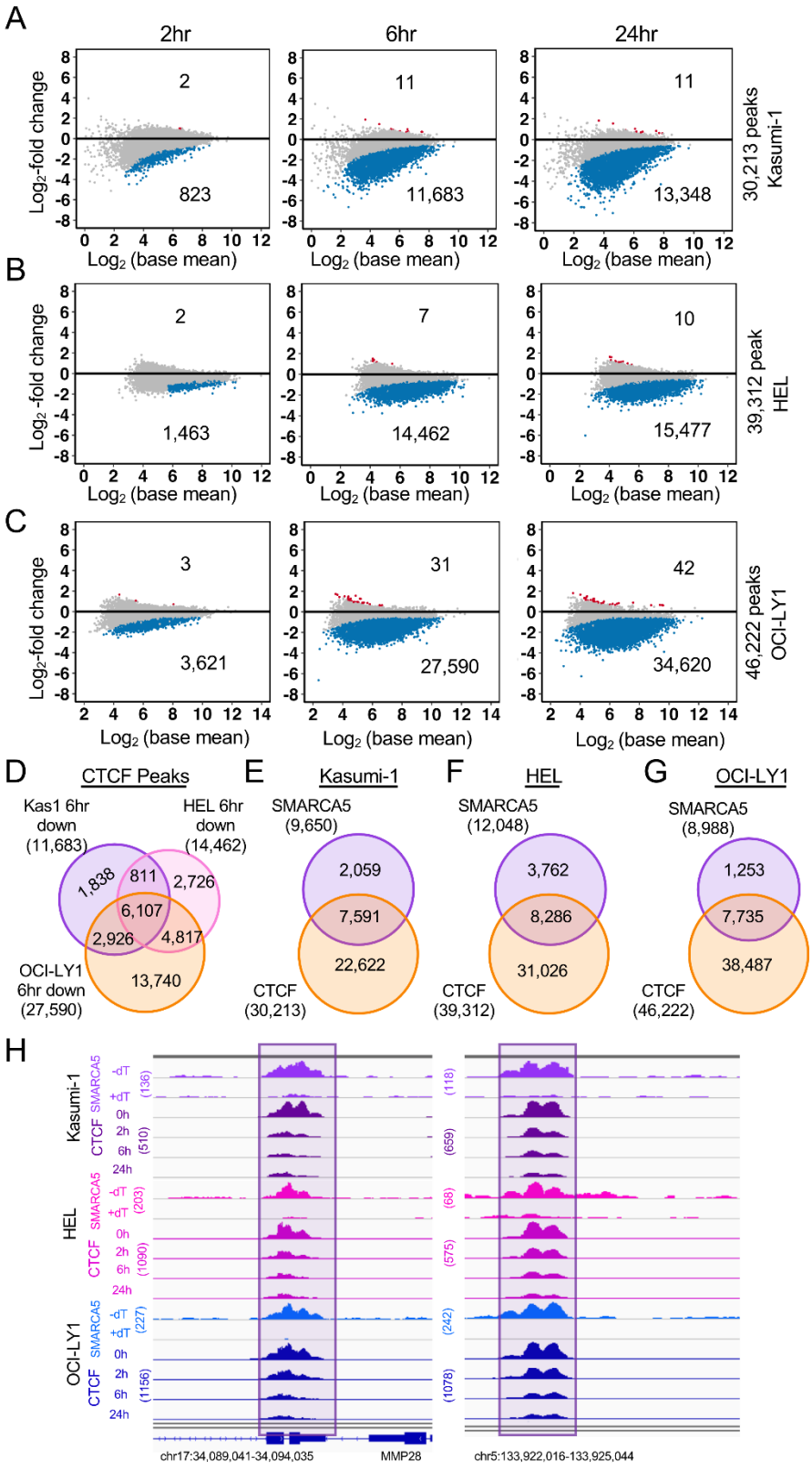
A
2hr
6hr
24hr
Log2-fold change
Log2 (base mean)
2
823
11
11,683
11
13,348
30,213 peaks
Kasumi-1
B
2
1,463
7
14,462
10
15,477
39,312 peak
HEL
C
3
3,621
31
27,590
42
34,620
46,222 peaks
OCI-LY1
D
CTCF Peaks
Kas1 6hr down (11,683)
HEL 6hr down (14,462)
OCI-LY1 6hr down (27,590)
1,838
811
2,726
6,107
2,926
4,817
13,740
E
Kasumi-1
SMARCA5 (9,650)
2,059
7,591
22,622
CTCF (30,213)
F
HEL
SMARCA5 (12,048)
3,762
8,286
31,026
CTCF (39,312)
G
OCI-LY1
SMARCA5 (8,988)
1,253
7,735
38,487
CTCF (46,222)
H
Kasumi-1
HEL
OCI-LY1
SMARCA5
CTCF
-dT
+dT
0h
2h
6h
24h
chr17:34,089,041-34,094,035
MMP28
chr5:133,922,016-133,925,044

Figure 3.5: SMARCA5 is required for CTCF DNA binding.

(A-C) MA plots of differential analysis of CTCF CUT&RUN peaks at the indicated time points in (A) Kasumi-1 (n=2), (B) HEL (n=2), (C) OCI-LY1 (n=2) cells expressing SMARCA5-FKBP12^{F36V}-FLAG. The blue dots represent the down-regulated CTCF peaks. The red dots represent the up-regulated CTCF peaks. (D) Venn diagram showing the CTCF sites that were lost 6hr after degradation of SMARCA5-FKBP12^{F36V}-FLAG from Kasumi-1, HEL, and OCI-LY1 cells. (E-G) Venn diagrams of the SMARCA5 CUT&RUN peaks and CTCF CUT&RUN peaks in (E) Kasumi-1, (F) HEL, (G) OCI-LY1 cells expressing SMARCA5-FKBP12^{F36V}-FLAG. (H) IGV browser tracks of SMARCA5 CUT&RUN and CTCF CUT&RUN at the indicated time after dTAG-47 treatment in Kasumi-1 (purple), HEL (pink), and OCI-LY1 (blue) SMARCA5-FKBP12^{F36V}-FLAG cell lines.

While the loss of CTCF DNA binding was closely associated with the degradation of SMARCA5, to ensure that other DNA binding proteins were not affected, we performed CUT&RUN assays to detect RUNX1 (Fig 3.6A) and the t(8;21) fusion protein, AML1-ETO (Fig 3.6B). Only 5 RUNX1 sites changed within 24hr of SMARCA5 degradation (Fig 3.6A), while 7,031 AML1-ETO sites were significantly down-regulated (Fig 3.6B). The larger number of AML1-ETO binding sites that were lost was intriguing, so we used motif analysis and found that these sites were not enriched for a RUNX1-binding motif, but contained a CTCF motif (Fig 3.6C). Not surprisingly, these down-regulated AML1-ETO sites did not overlap with the previously published AML1-ETO regulated enhancer peaks (Fig 3.6D) (Stengel et al., 2021). Indeed, many of these AML1-ETO binding sites were bound by CTCF (Fig 3.6E), suggesting that the loss of these weak AML1-ETO binding sites could be secondary to the loss of CTCF.

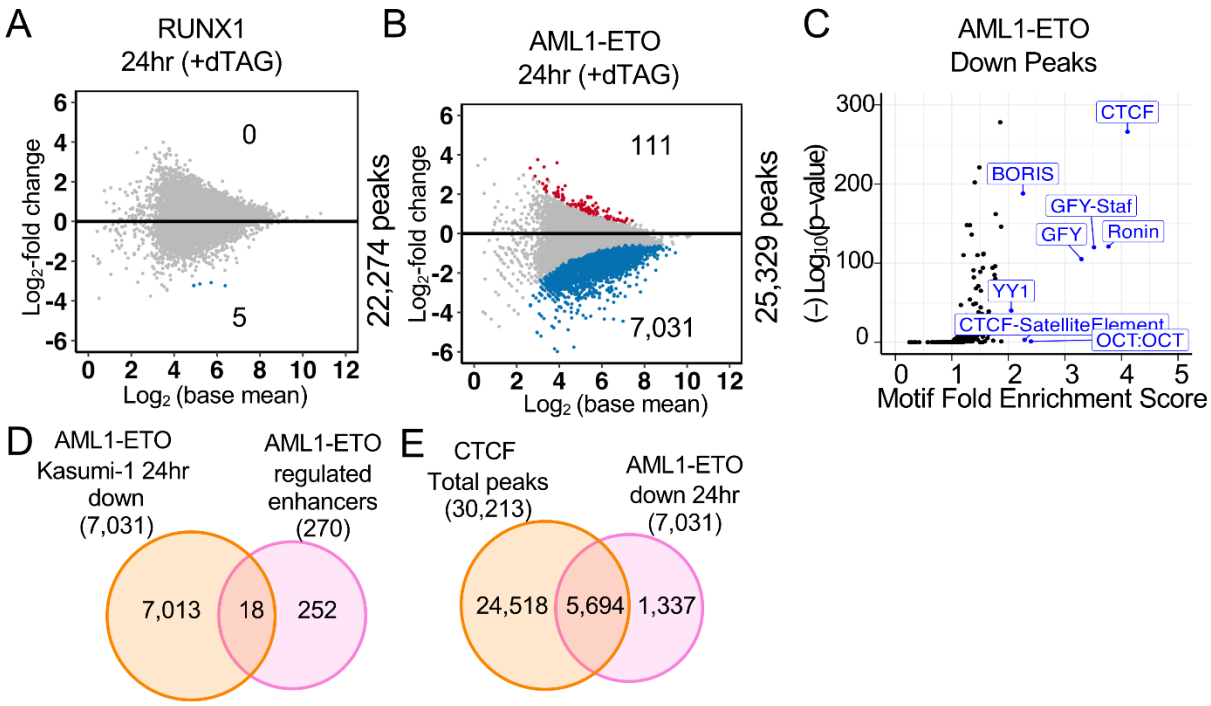

Figure 3.6: SMARCA5 is not required for RUNX1 localization.

(A and B) MA plots of differential analysis of (A) RUNX1 CUT&RUN or (B) AML1-ETO CUT&RUN peaks 24hr after degrading SMARCA5-FKBP12^{F36V}-FLAG in Kasumi- 1 cells (n=2). (C) Volcano plots of the negative Log_{10} (p-value) versus motif enrichment scores from the down-regulated AML1-ETO peaks in B. (D) Venn diagram showing the overlap between the AML1-ETO peaks down-regulated upon degradation of SMARCA5- FKBP12^{F36V}-FLAG with the previously identified AML1-ETO regulated enhancer peaks (Stengel et al., 2021). (E) Venn diagrams comparing the AML1-ETO down-regulated peaks with total CTCF CUT&RUN peaks.

3.2.4 SMARCA5 maintains accessible chromatin around CTCF motifs

Given the ability of SMARCA5 to slide nucleosomes along the DNA (Aalfs et al., 2001; He et al., 2008), we used ATAC-seq to identify areas of the genome that were affected by degradation of SMARCA5 (Buenrostro et al., 2013). We performed a time course analysis at 0, 2, 6, and 24hr after the addition of dTAG-47 to the Kasumi-1 cell line to assess changes in transposase accessibility over time. This analysis identified few areas of the genome that were affected within the first two hours, but there were substantial changes within 6hr with 8,513 peaks lost and 6,347 gained, and by 24hr there were 12,669 peaks lost and 10,766 gained (Fig 3.7A, 3.8A, and 3.8B). We also performed ATAC-seq after 6hr of dTAG-47 treatment in the HEL and OCI-LY1 cell lines containing the SMARCA5-FKBP12^{F36V} fusion and also found large changes in accessibility (Fig 3.7B and 3.7C). The down-regulated peaks in each of the three cell lines showed significant overlap with 888 peaks found in all three cell lines and roughly 50% overlap when comparing just two of the cell lines at a time (Fig 3.7D). On the other hand, the up-regulated peaks showed only 24 overlapping peaks between all three cell lines and 5% overlap when comparing just two cell lines (Fig 3.7E).

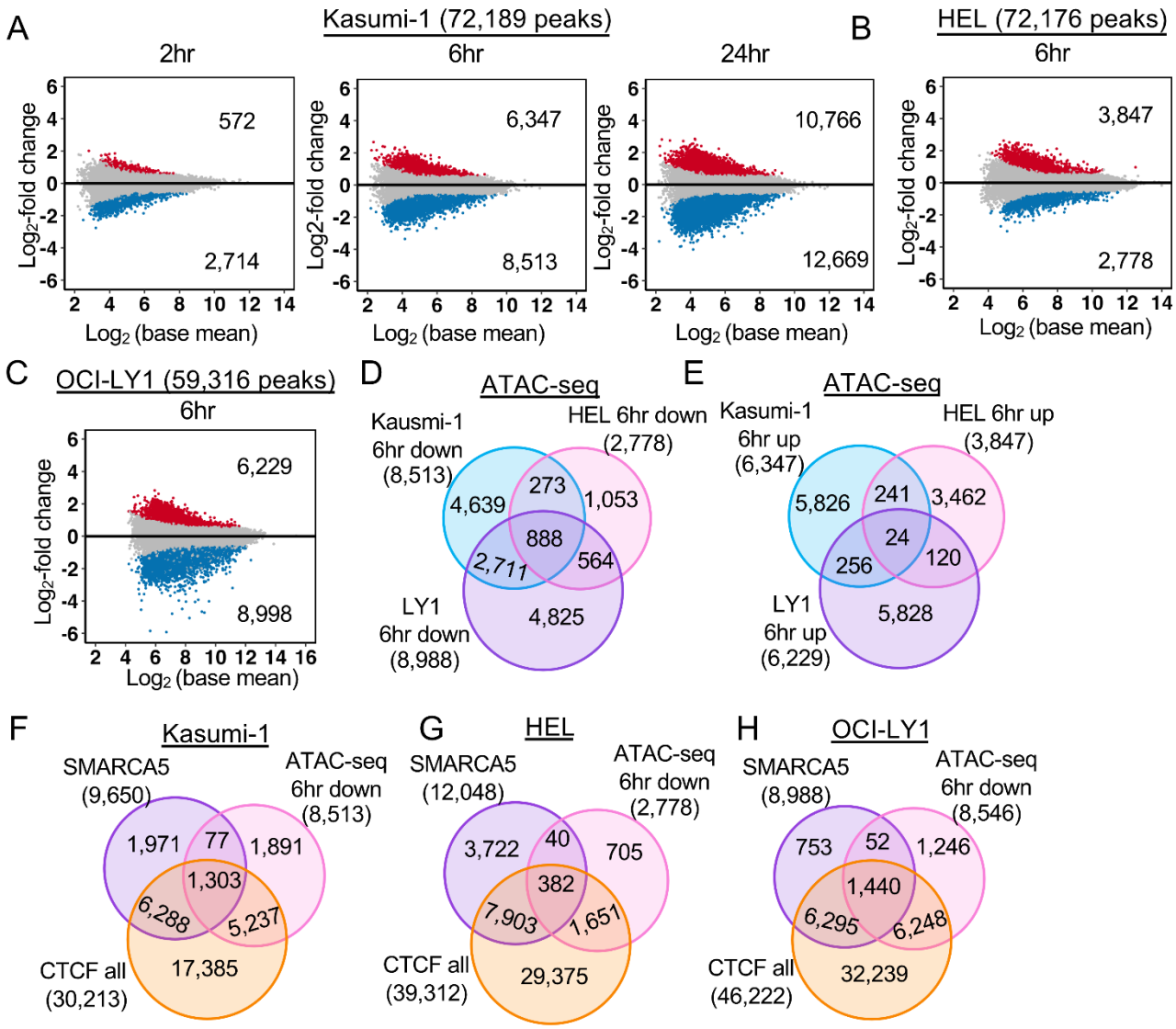

Figure 3.7: Loss of SMARCA5 alters chromatin accessibility.

(A-C) MA plots of differential analysis of ATAC-seq peaks at the indicated time points of dTAG-47-mediated degradation of SMARCA5-FKBP12^{F36V}-FLAG in (A) Kasumi-1 (n=4), (B) HEL (n=2), (C) OCI-LY1 (n=2). The blue dots represent peaks down-regulated and red dots up-regulated by at least 1.5-fold. (D-E) Venn diagrams showing that some down-regulated (D) and up-regulated (E) ATAC-seq peaks from A, B, and C are conserved in Kasumi-1, HEL, and OCI-LY1 cells. (F-H) Venn diagrams showing the overlap of the SMARCA5 CUT&RUN peaks, CTCF CUT&RUN peaks, and ATAC-seq 6hr after degradation of SMARCA5-FKBP12^{F36V}-FLAG in Kasumi-1 (F), HEL (G), and OCI-LY1 (H) cells.

Next, we asked if the accessibility changes were associated with the genomic binding sites of SMARCA5 or with the CTCF binding sites. The ATAC-seq peaks that were lost showed a significant overlap with CTCF binding sites in all three cell lines (Fig 3.7F-H). Interestingly, there was a better overlap between SMARCA5 DNA association and loss of accessibility in the Kasumi-1 and OCI-LY1 cells than in HEL cells (Fig 3.7F-H). Nevertheless, motif enrichment analysis of the sequences identified by the down-regulated ATAC-seq peaks found that the CTCF/BORIS motif was almost 15-fold enriched in comparison to any other motif in all three cell lines tested (Fig 3.8C). Moreover, when considering the 888 ATAC-seq peaks that were found in all three cell lines (Fig 3.7D), the enrichment for CTCF motifs was 20-fold (not shown). In contrast, there was little enrichment among the up-regulated peaks.

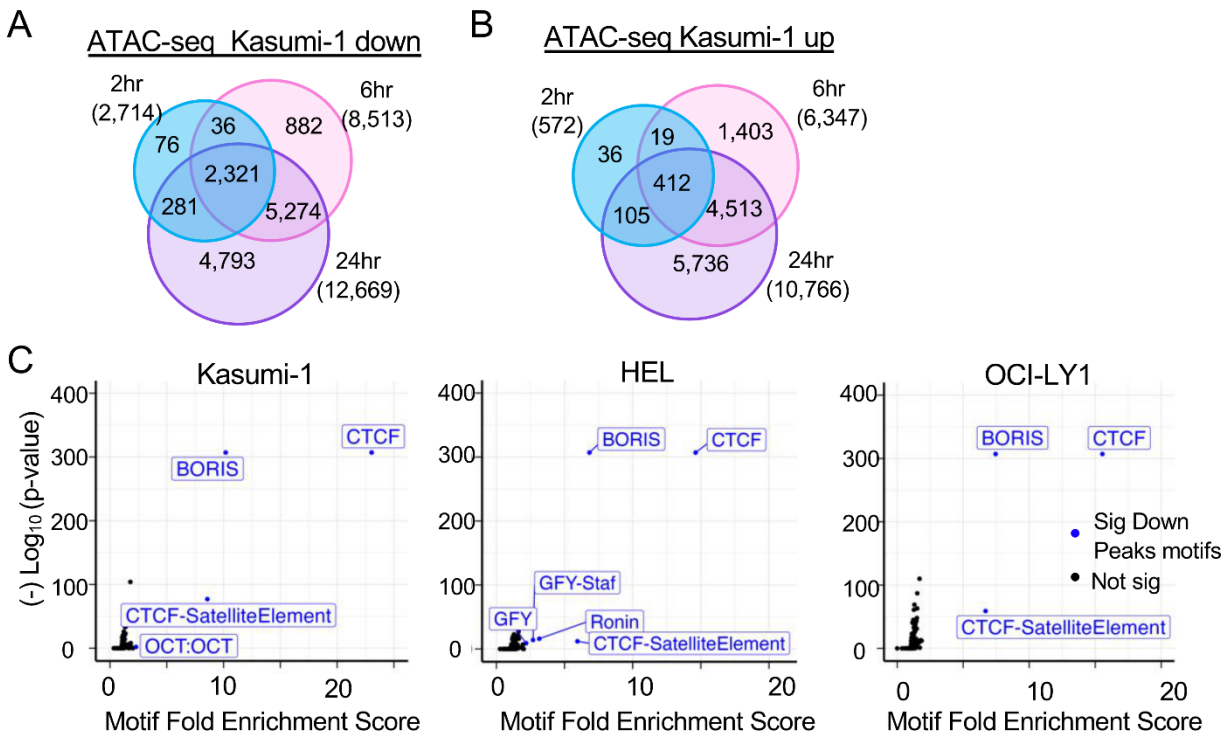

Figure 3.8: SMARCA5 degradation affects chromatin accessibility.

(A and B) Venn diagrams showing the loss of in ATAC-seq peaks (A) or gain of peaks (B) over the time course of degradation of SMARCA5-FKBP12^{F36V}-FLAG in Kasumi-1 cells. (C) Volcano plots of the negative Log_{10} of the p-value versus the motif enrichment score of ATAC-seq peaks 6hr after degradation of SMARCA5-FKBP12^{F36V}-FLAG in Kasumi-1, HEL, and OCI-LY1 cells. The blue dots represent motifs enriched by at least 2-fold.

3.2.5 SMARCA5 only modestly affects RNA polymerase dynamics

Next, we sought to determine whether the changes in chromatin accessibility observed upon SMARCA5 degradation were associated with changes in gene expression. Having a small molecule that rapidly degraded SMARCA5-FKBP12^{F36V} allowed us to assess gene expression changes in the first hours after SMARCA5 degradation to determine if it directly regulates transcription. Given that SMARCA5 has been linked to the control of RNA polymerase elongation (Jimeno-González et al., 2015), we performed precision nuclear run-on sequencing (PRO-seq) (Mahat et al., 2016) at 0, 2, 6, and 24hr after the addition of dTAG-47 to Kasumi-1 cells containing the SMARCA5-FKBP12^{F36V} fusion protein (Fig 3.9 and 3.10). This analysis identified very few genes that exhibited reduced or increased gene body (GB) transcription with 50 genes up and 63 genes down at least two-fold at 6hr after degrading SMARCA5 (Figure 3.9A). Moreover, there was very little overlap of the transcriptional changes over time with only 19 genes exhibiting increased gene body transcription and 23 genes exhibiting reduced gene body transcription at all 3 time points (Fig 3.9B). Roughly a third of these 6hr changed genes could be associated with a SMARCA5 binding site within +/- 25 KB of the TSS (Fig 3.9C). Interestingly, many more of these 6hr changed genes were associated with a nearby CTCF binding site (50 genes within 500 bp of the TSS; Fig 3.9D). While a similar number of changes were detected by RNA-seq analysis at 6hr after degrading SMARCA5, the changes in nascent transcription only affected a small number of mRNAs even at 24hr after degradation of SMARCA5 (40 down; 18 up; Fig 3.9E, 3.9F).

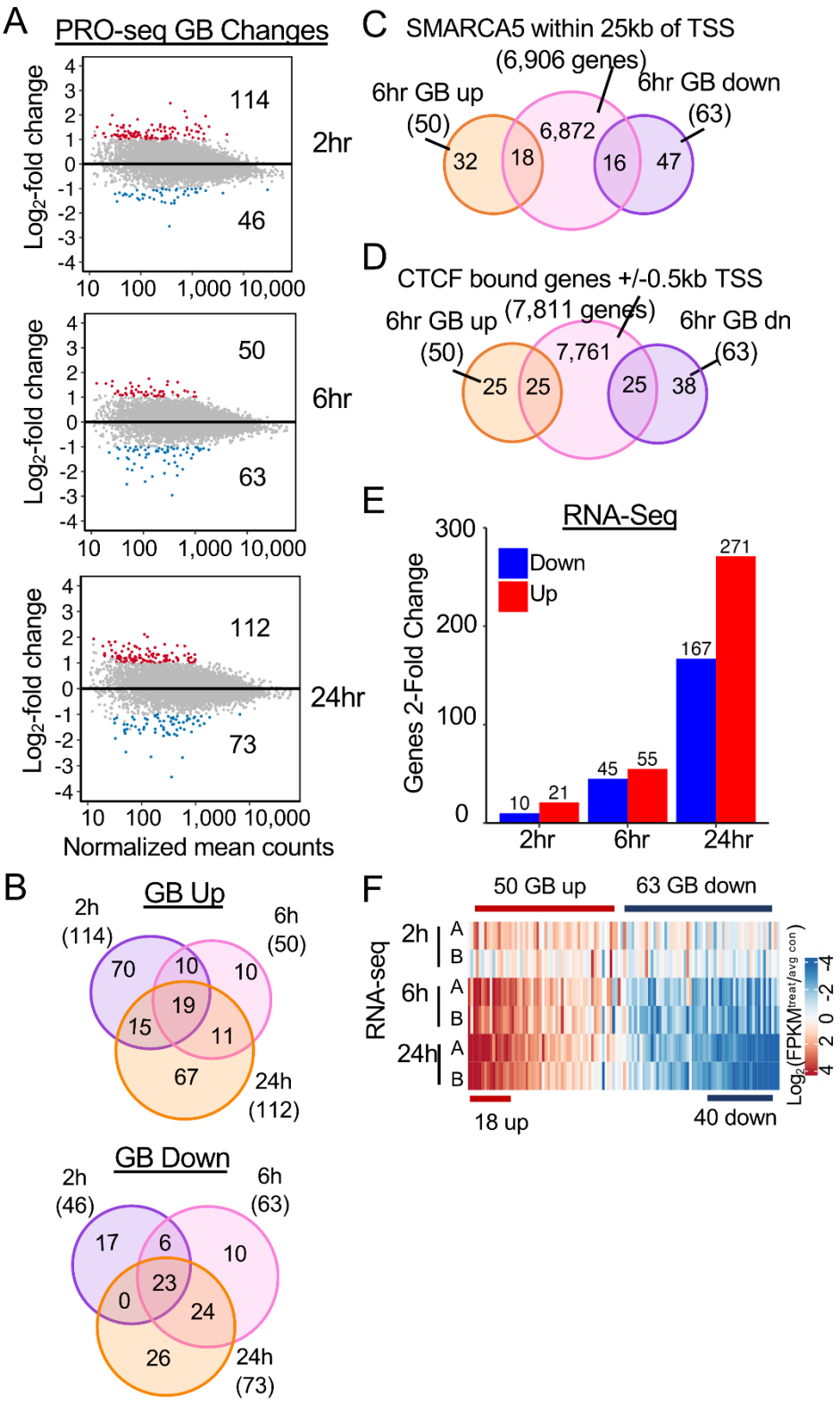
A
PRO-seq GB Changes
Log2-fold change
114
46
2hr
50
63
6hr
112
73
24hr
10 100 1,000 10,000
Normalized mean counts
B
GB Up
2h (114)
6h (50)
24h (112)
70
10
10
19
15
11
67
GB Down
2h (46)
6h (63)
24h (73)
17
6
10
23
0
24
26
C
SMARCA5 within 25kb of TSS
(6,906 genes)
6hr GB up (50)
6hr GB down (63)
32
18
6,872
16
47
D
CTCF bound genes +/-0.5kb TSS
(7,811 genes)
6hr GB up (50)
6hr GB dn (63)
25
25
7,761
25
38
E
RNA-Seq
Genes 2-Fold Change
Down
Up
10
21
45
55
167
271
2hr
6hr
24hr
F
50 GB up
63 GB down
RNA-seq
2h
6h
24h
A
B
Log2(FPKMtreat/avg con)
18 up
40 down

Figure 3.9: Degradation of SMARCA5 causes modest transcriptional changes.

(A) MA plots of the $\log_2$ fold change vs the normalized mean counts of the gene body. The colored dots are the significant (adjusted p-value < 0.05 and fold change > 2; red = Up, blue = Down) gene body changes. (B) Venn diagram showing the overlap between changes within the gene body of up- or down-regulated genes at 2, 6, 24hr. (C) Venn diagram showing the overlap between gene body changes of up- or down-regulated genes at 6hrs after SMARCA5 degradation intersected with SMARCA5 CUT&RUN peaks annotated to +/-25kb of the TSS of the nearest neighbor gene. (D) Venn diagram showing the overlap between genes up- or down-regulated 6hr after SMARCA5 degradation intersected with CTCF CUT&RUN peaks annotated to +/-500bp of the TSS of the nearest neighbor gene. (E) Bar graph of the changes in mRNA pools measured by RNA-seq at 2h, 6h and, 24h after dTAG-47 treatment to degrade SMARCA5 in Kasumi-1 cells. (F) Heatmap of $\log_2$ (FPKM/avgFPKM) of each RNA-seq replicate 2h, 6h, and 24h after SMARCA5 degradation for genes regulated in the PRO-seq analysis at 6h.

PRO-seq also provides information about RNA polymerase pausing around the TSSs of expressed genes. We quantified the promoter-proximal (pp) RNA polymerase density and found that rapid degradation of SMARCA5 resulted in only modest effects (Fig 3.10A-B) and only 26 of these genes showed a loss of expression in our RNA-seq datasets (Figure 3.10C). Interestingly, 53 of these pp down genes contained a CTCF CUT&RUN binding site, within 500bp of the transcription start site and 46 of these pp down genes were also bound SMARCA5 (Figure 3.10D-E), suggesting that these effects could be secondary to the loss of CTCF DNA binding.

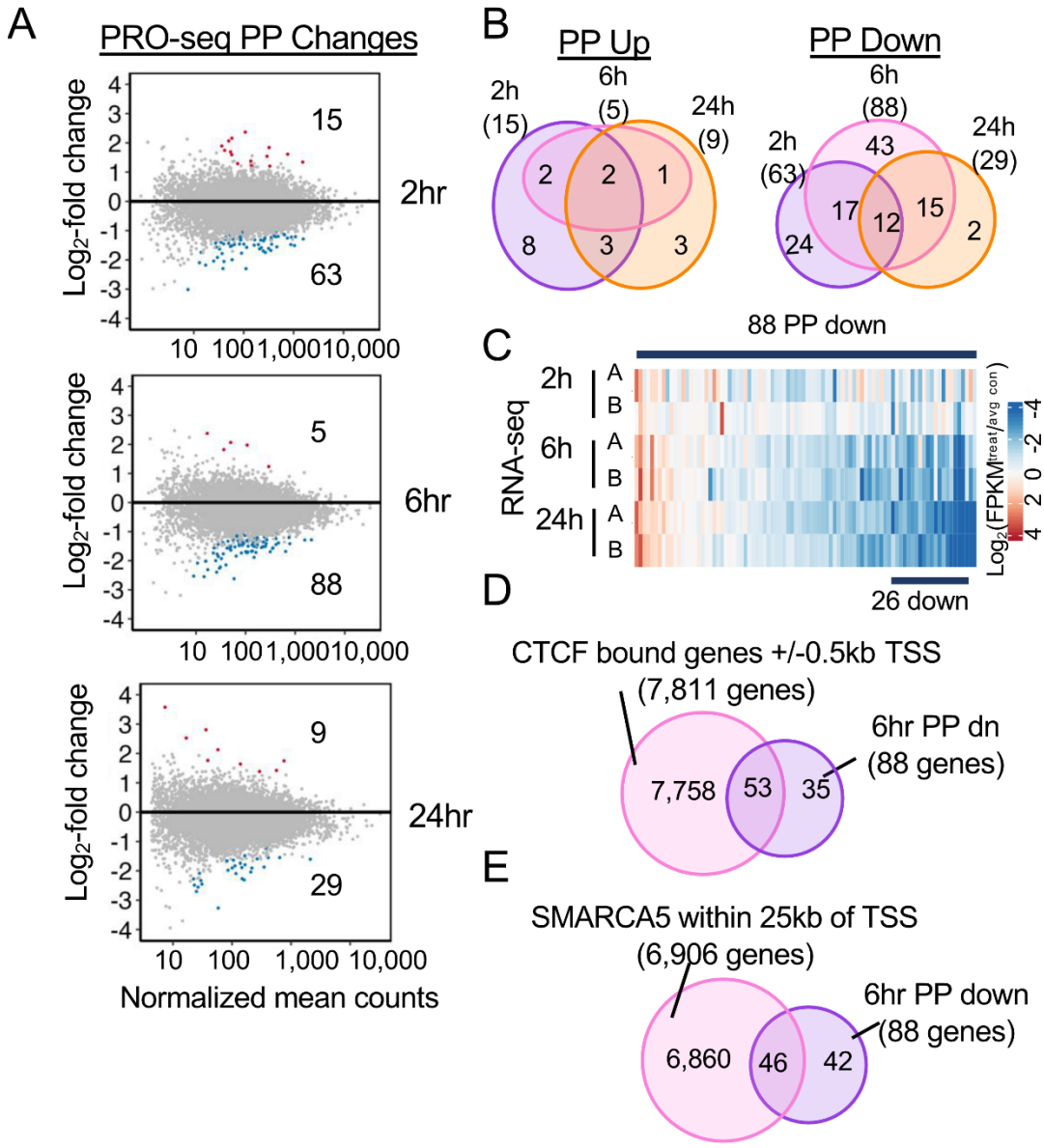
A
PRO-seq PP Changes
Log2-fold change
15
63
2hr
5
88
6hr
9
29
24hr
10 100 1,000 10,000
Normalized mean counts
B
PP Up
2h (15)
6h (5)
24h (9)
2
2
1
8
3
3
PP Down
2h (63)
6h (88)
24h (29)
43
17
12
15
24
2
C
88 PP down
RNA-seq
2h
6h
24h
A
B
Log2(FPKMtreat/avg con)
26 down
D
CTCF bound genes +/-0.5kb TSS
(7,811 genes)
6hr PP dn
(88 genes)
7,758
53
35
E
SMARCA5 within 25kb of TSS
(6,906 genes)
6hr PP down
(88 genes)
6,860
46
42

Figure 3.10: Degradation of SMARCA5 causes changes in RNA polymerase pausing.

(A) MA plots of the log2-fold change vs the normalized mean counts of the promoter proximal regions of all expressed genes. The colored dots are the significant promoter proximal gene changes (adjusted p-value < 0.05 and fold change > 2; red = Up, blue = Down). (B) Venn diagrams showing the overlap between promoter proximal changes of up- or down-regulated genes at 2, 6, 24hr. (C) Heatmap of Log_2 (FPKM/avgFPKM) of each RNA-seq replicate of the time course of SMARCA5 degradation. Shown are the relative changes of gene expression of the 88 genes that displayed a loss of paused polymerase at 6h after degradation of SMARCA5. (D) Venn diagram showing the overlap between promoter proximal genes up- or down-regulated 6hr after SMARCA5 degradation intersected with CTCF CUT&RUN peaks annotated to +/-500bp of the TSS of the nearest neighbor gene. (E) Venn diagram showing the overlap between promoter proximal changes of up- or down-regulated genes at 6hrs after SMARCA5 degradation intersected with SMARCA5 CUT&RUN peaks annotated to +/-25kb of the TSS of the nearest neighbor gene.

3.2.6 SMARCA5 loss affects chromatin architecture

Given that SMARCA5 can position nucleosomes, we further probed the ATAC-seq data to examine nucleosome repeat length as a measure of nucleosome compaction within the three different cell lines. The algorithm uses a local regression fit model with two smoothing parameters (only one shown, red; Fig 3.11A) to calculate the average nucleosome repeat length throughout the region of interest (Barisic et al., 2019). This analysis detected a lengthening in the nucleosome repeat length within 6hr of dTAG-47 treatment that was even further increased by 24hr in Kasumi-1 cells (Fig 3.11A and 3.11B). Analysis of the HEL and OCI-LY1 ATAC-seq data also showed a significant increase in global nucleosome repeat length 6hr after dTAG-47 treatment (Fig 3.11C-3.11F). While our ATAC-seq sequencing depth was not sufficient to measure nucleosome repeat length around individual features within the genome, it still appeared that SMARCA5 was regulating global nucleosome repeat length.

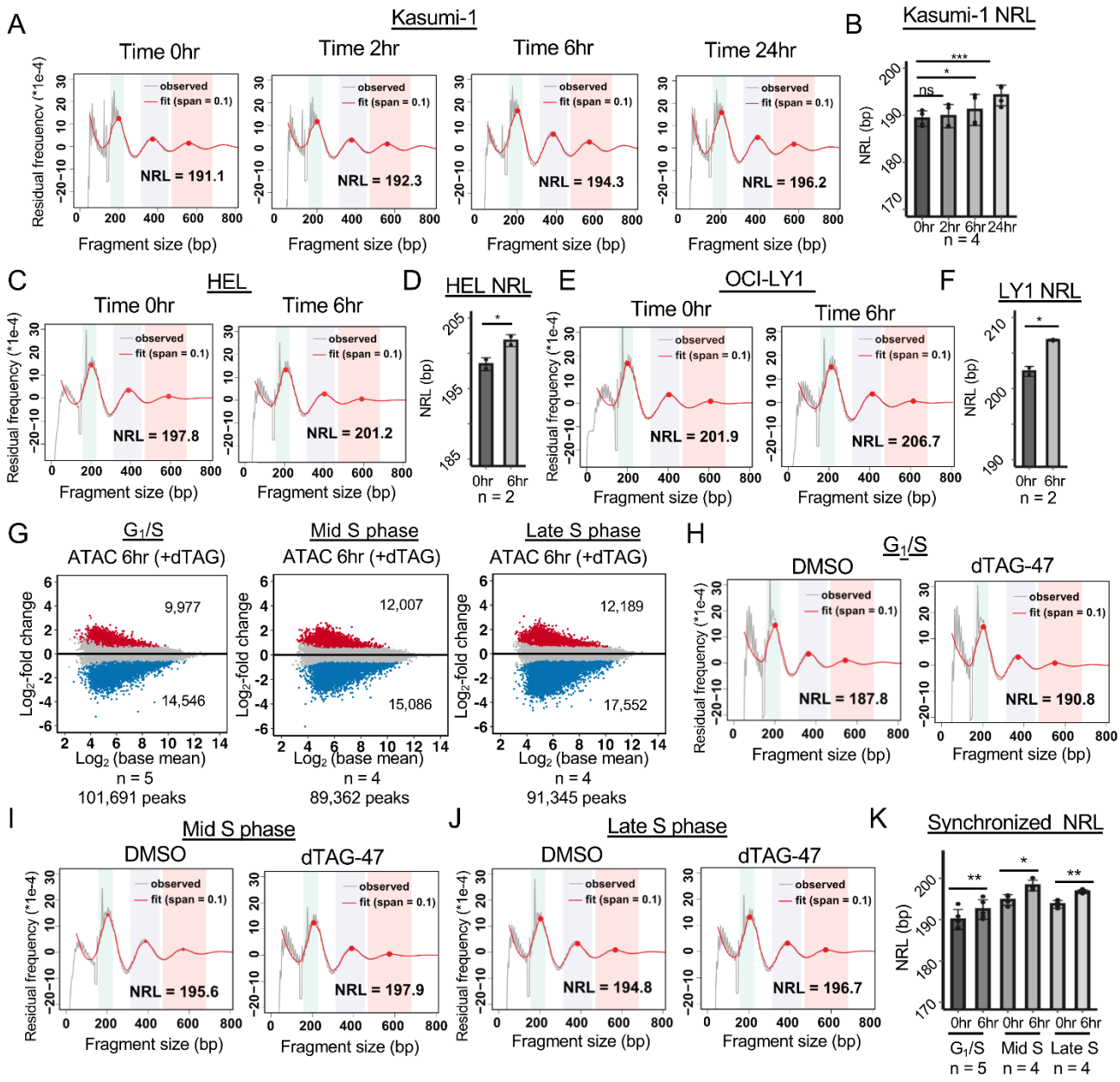
A
Kasumi-1
Time 0hr
Time 2hr
Time 6hr
Time 24hr
observed
fit (span = 0.1)
NRL = 191.1
NRL = 192.3
NRL = 194.3
NRL = 196.2
Residual frequency (*1e-4)
Fragment size (bp)
B
Kasumi-1 NRL
NRL (bp)
ns
*
***
0hr 2hr 6hr 24hr
n = 4
C
HEL
Time 0hr
Time 6hr
NRL = 197.8
NRL = 201.2
D
HEL NRL
0hr 6hr
n = 2
E
OCI-LY1
Time 0hr
Time 6hr
NRL = 201.9
NRL = 206.7
F
LY1 NRL
G
G1/S
Mid S phase
Late S phase
ATAC 6hr (+dTAG)
9,977
14,546
12,007
15,086
12,189
17,552
Log2-fold change
Log2 (base mean)
n = 5
101,691 peaks
n = 4
89,362 peaks
n = 4
91,345 peaks
H
G1/S
DMSO
dTAG-47
NRL = 187.8
NRL = 190.8
I
Mid S phase
NRL = 195.6
NRL = 197.9
J
Late S phase
NRL = 194.8
NRL = 196.7
K
Synchronized NRL
**
*
**
0hr 6hr 0hr 6hr 0hr 6hr
G1/S
Mid S
Late S
n = 5
n = 4
n = 4

Figure 3.11: SMARCA5 regulates nucleosome repeat length independent of the cell cycle.

(A) Nucleosome phasing graph showing changes in the nucleosome repeat length (NRL) over the time course of degradation of SMARCA5-FKBP12^{F36V}-FLAG. Red line shows the "fit" of the curve versus the observed ATAC-seq signal of one replicate (grey line) using the entire genome. (B) Bar graph of the NRL of all replicates (n=4) at each of the indicated time points. (C) Nucleosome phasing graph showing changes in the NRL at 6hr after degradation of SMARCA5-FKBP12^{F36V}-FLAG in HEL cells using the entire genome. (D) Bar graph of the NRL of all replicates (n=2) at 6hr after degradation of SMARCA5 in HEL cells. (E) Nucleosome phasing graph showing changes in the NRL at 6hr after degradation of SMARCA5-FKBP12^{F36V}-FLAG in OCI-LY1 cells using the entire genome. (F) Bar graph of the NRL of all replicates (n=2) at 6hr after degradation of SMARCA5 in OCI-LY1 cells. (G) MA plots showing changes in accessible peaks 6hr after degrading SMARCA5 at the G_1/S, mid S and late S phases of the cell cycle. The blue and red dots represent the changed ATAC peaks with an at least 1.5-fold change. (H-J) Nucleosome phasing graphs showing changes in the nucleosome repeat length in Kasumi-1 cells synchronized at G_1/S (H), Mid S phase (6hr after release from double thymidine block; I), or late S phase (12hr after release from block; J). (K) Bar graph of the NRL of individual replicates (dots) at each of the indicated phases of the cell cycle. For B, D, F and K, the significance was calculated with a two sampled paired t-test: ****=p<0.0001, ***=p<0.001, **=p<0.01, *=p<0.05; error bars represent standard deviation.

A major advantage of the dTAG system is the ability to probe the action of SMARCA5 during the cell cycle. We used a double thymidine block to synchronize Kasumi-1 cells at the G_1/S boundary prior to the degradation of SMARCA5-FKBP12^{F36V} and used BrdU incorporation to track the progression of cells through the S phase (Fig 3.12A-3.12C). At the initiation of the time course, the vast majority of the cells were at G_1/S and did not incorporate BrdU. However, within the first 2hr after removal of the thymidine block, the cells had re-entered the S phase, indicating that SMARCA5 was not required to initiate DNA replication. Within the first 6hr after release from the block, more than 60% of the cells had incorporated BrdU (Fig 3.12A-3.12C) and within 12hr the control cells showed a roughly 50% reduction in the percentage of cells in the S phase, as these cells traversed and then exited the S phase. In contrast, Kasumi-1 cells lacking SMARCA5 were retained in the late S phase 12hr after release into the S phase and displayed a significant lag in returning to the G_1 phase (Fig 3.12A-3.12C). These data indicate that SMARCA5 may contribute to the passage of the late S phase in Kasumi-1 cells.

Using this information, we then performed ATAC-seq on synchronized Kasumi-1 cells with or without a 6hr degradation of SMARCA5-FKBP12^{F36V} at G_1/S, mid S (6hr after release), or late S (12hr after release). First, we identified accessible peaks and performed differential analysis to define the effect of synchronization on general accessibility. This analysis showed more dramatic changes in chromatin accessibility with a greater loss of peaks than gain of peaks in each of the cell cycle phases in comparison to untreated cells (Fig 3.11G). The up- and down-regulated peaks also showed a high degree of overlap between these phases of the cell cycle (40-60%; Fig 3.12D-3.12E), and

a high degree of overlap with binding of CTCF at each phase of the cell cycle (Fig 3.12F-3.12H). This suggests that loss of CTCF was likely a key component of the loss of accessibility upon degradation of SMARCA5 at these different stages of the cell cycle.

Next, we analyzed the nucleosome repeat length of these populations of cells. We noted that nucleosome repeat length in the control cells changed as cells progressed from G_1/S into S phase with a longer distance between nucleosomes at mid and late S phase after release from the block (0hr time points, Fig 3.11H-3.11J). Even though the starting point for assessing the effects of SMARCA5 degradation was greater during the S phase, upon degradation of SMARCA5, we found a small but reproducible increase of the average length between nucleosomes at all three stages of the cell cycle assessed (Fig 3.11H-3.11K). Together, these results suggest that SMARCA5 regulated chromatin accessibility independent of the cell cycle, and that SMARCA5 was required for nucleosomal spacing and compaction in both late G_1 and the S phase of the cell cycle.

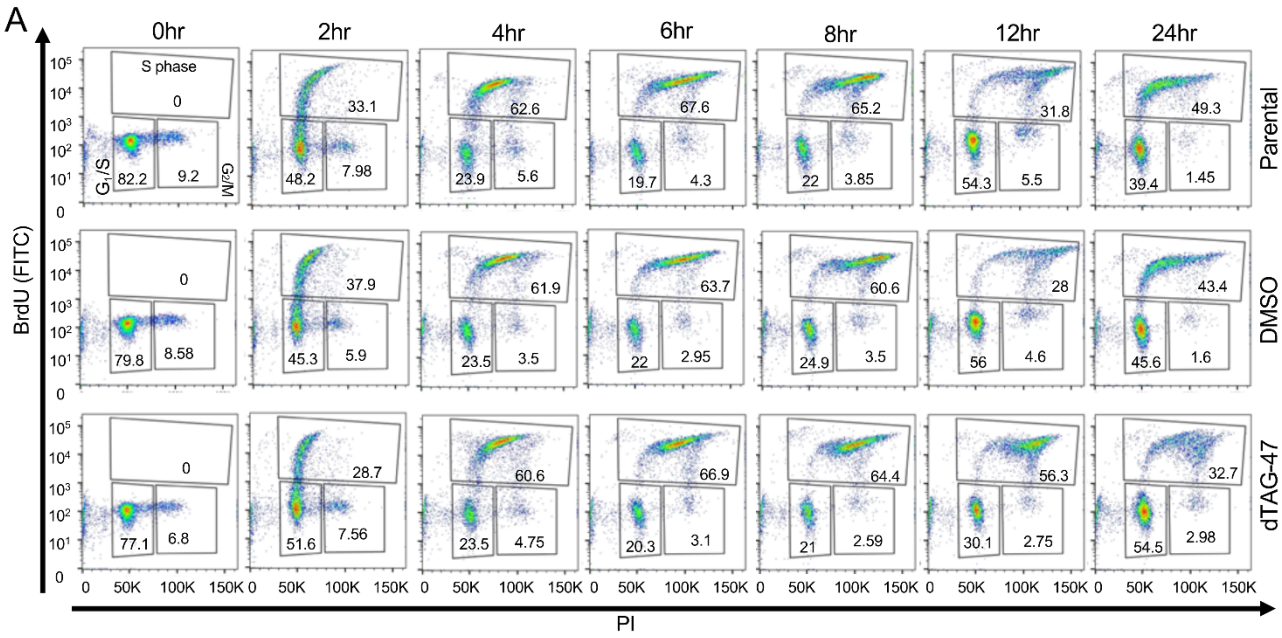
A
0hr
2hr
4hr
6hr
8hr
12hr
24hr
Parental
DMSO
dTAG-47
BrdU (FITC)
PI
S phase
G1/S
G2/M
B
S phase
Percent Average
T6hr
T8hr
T12hr
T24hr
C
G1/S phase
D
ATAC-seq down
G1/S (14,546)
T6hr Release (15,086)
T12hr Release (17,584)
E
ATAC-seq up
G1/S (9,977)
T6hr Release (12,007)
T12hr Release (12,262)
F
G1/S down (14,546)
CTCF down T6hr (11,683)
G
6hr Release down (15,086)
CTCF down 6h (11,683)
H
12hr Release Down (17,584)
CTCF down 6hr (11,683)

Figure 3.12: SMARCA5 loss affects chromatin accessibility during the cell cycle.

(A) Flow cytometry analysis of BrdU incorporation versus propidium iodide staining of Kasumi-1-SMARCA5- FKBP12^{F36V}-FLAG cells after a double thymidine block, followed by treatment with dTAG-47 for 2hr to ensure SMARCA5 degradation before release into the S phase for the times indicated above each plot. Parental cells are unmodified Kasumi-1 cells and DMSO are the *SMARCA5-FKBP12^{F36V}*-expressing cells that were mock treated with DMSO. (B) Bar graph of quantification of flow cytometry analysis of BrdU incorporation from cells released at 6h, 8h, 12h, and 24h (n=2). (C) Bar graph of quantification of flow cytometry analysis of G1 phase cells at 6h, 8h, 12h, and 24h after release (n=2). (D) Venn diagram of the down-regulated ATAC-seq peaks at G1S, 6h after release (mid S), and 12h after release (late S). (E) Venn diagram of the up-regulated ATAC-seq peaks at G1S, 6h after release (mid S), and 12h after release (late S). (F-H) Venn diagrams of the down-regulated CTCF peaks from Figure 3 intersected with the down-regulated ATAC-seq peaks at G1/S (F), 6h after release (G), and 12hr after release (H).

While ATAC-seq is a relatively rapid procedure, to deepen our analysis of chromatin structure and to examine different genomic features upon inactivation of SMARCA5, we performed micrococcal nuclease coupled with sequencing (MNase-seq) in Kasumi-1 cells (Barisic et al., 2019; Goldman et al., 2010; He et al., 2008; Tsukiyama et al., 1999; Whitehouse et al., 2003). Like the ATAC-seq analysis, MNase-Seq detected a subtle lengthening in the nucleosome repeat length within two hours after the addition of dTAG-47 that was more distinct by 6hr and continued to expand at the 24hr time point (Fig 3.13A and 3.13B). Using the MNase-seq datasets, we were also able to measure the nucleosome repeat length around other sites throughout the genome. We assessed the nucleosome repeat length around the TSS of all expressed genes and at CTCF sites. We found no significant change in nucleosome repeat length over time around the start sites of expressed genes (Fig 3.14A and 3.14B), which is consistent with the small effects that were found on transcription. Furthermore, there was no change in nucleosome positioning when we plotted the normalized nucleosome occupancy -0.5kb and + 1kb from the TSSs of expressed genes (Fig 3.14C).

Next, we used CUT&RUN to assess the location of H3K27me3 to mark closed chromatin and to assess the nucleosome repeat length around H3K27me3-marked sites. Although using subsets of the genome reduced the robustness of the phasing due to the inclusion of fewer reads in the analysis, it appeared that the nucleosome repeat length around H3K27me3 was longer than the global nucleosome repeat length at the beginning of the time course, which is consistent with a more closed chromatin conformation (Fig 3.13C and 3.13D). The nucleosome repeat length still increased in H3K27me3-marked regions upon degradation of SMARCA5 (Fig 3.13C and 3.13D).

H2A.Z marks open chromatin and has also been linked to SMARCA5 function, as H2A.Z was better at stimulating the ATPase activity of SMARCA5 than H2A (Goldman et al., 2010). Moreover, H2A.Z also co-localized to CTCF binding sites (Wen et al., 2020). Therefore, we used CUT&RUN to map H2A.Z genomic localization and assessed the co-localization of SMARCA5 and CTCF (Fig 3.14D). It appeared that all essentially all of the H2A.Z peaks were associated with at least a low amount of SMARCA5 that was lost upon degradation, and roughly 70% of the called SMARCA5 peaks co-localized with both H2A.Z and CTCF (Fig 3.14E). The majority of these SMARCA5 binding sites were found within the more intense peaks found in clusters 1 and 2 after k-means clustering of the H2A.Z signal (Fig 3.14D and data not shown). Likewise, essentially all of the H2A.Z peaks were associated with CTCF, and CTCF was rapidly lost from these H2A.Z peaks upon degradation of SMARCA5 with a larger decrease associated with clusters 1 and 2 (Fig 3.14D and 3.14F). The proportion of down-regulated CTCF binding sites that overlapped with an H2A.Z site remained consistent throughout the time course of SMARCA5 degradation (25-30%, Fig 3.14F).

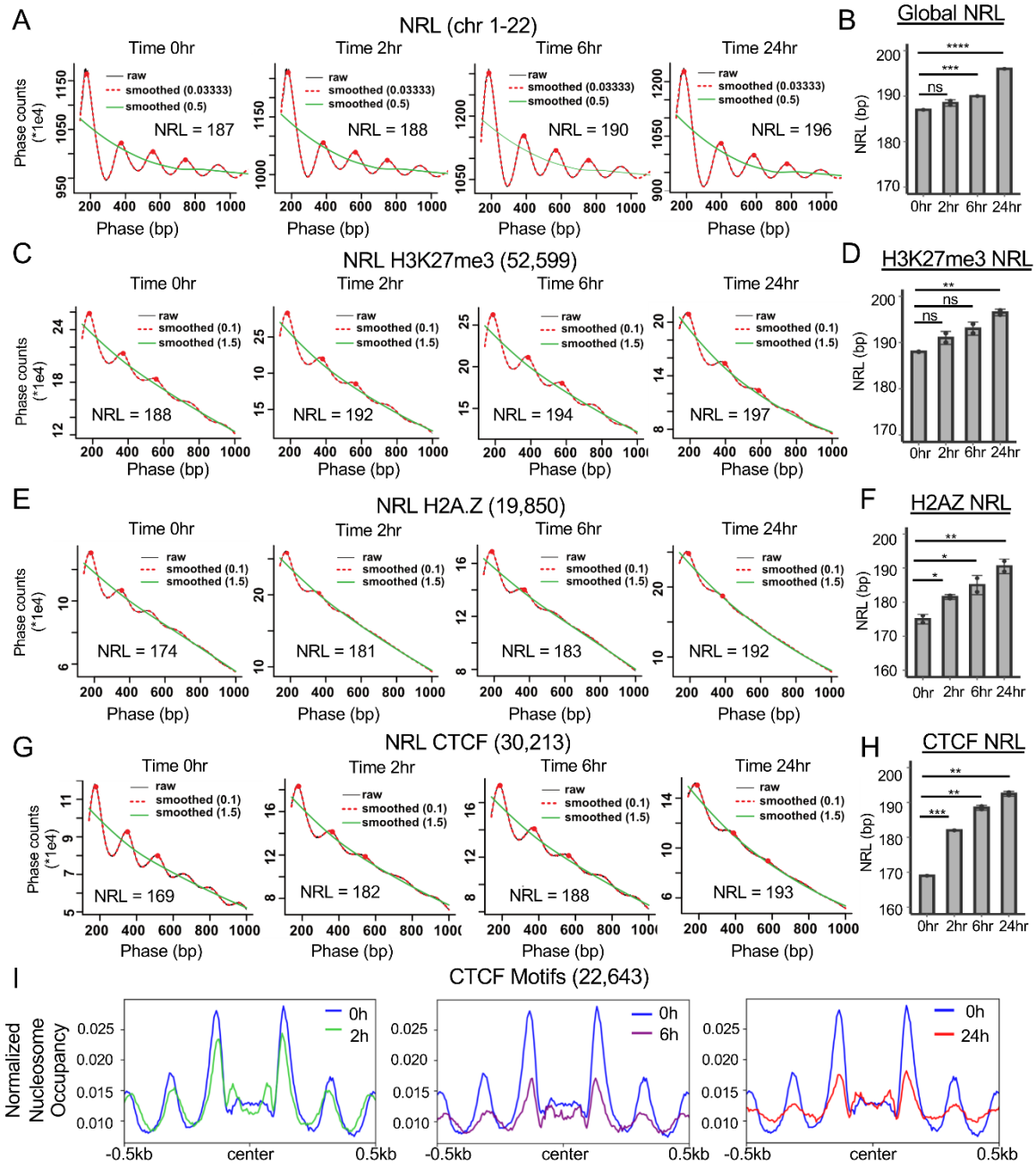
A
NRL (chr 1-22)
Time 0hr
Time 2hr
Time 6hr
Time 24hr
raw
smoothed (0.03333)
smoothed (0.5)
NRL = 187
NRL = 188
NRL = 190
NRL = 196
Phase counts (*1e4)
Phase (bp)
B
Global NRL
NRL (bp)
0hr 2hr 6hr 24hr
C
NRL H3K27me3 (52,599)
smoothed (0.1)
smoothed (1.5)
NRL = 188
NRL = 192
NRL = 194
NRL = 197
D
H3K27me3 NRL
E
NRL H2A.Z (19,850)
NRL = 174
NRL = 181
NRL = 183
NRL = 192
F
H2AZ NRL
G
NRL CTCF (30,213)
NRL = 169
NRL = 182
NRL = 188
NRL = 193
H
CTCF NRL
I
CTCF Motifs (22,643)
Normalized Nucleosome Occupancy
0h
2h
6h
24h
-0.5kb
center
0.5kb

Figure 3.13: SMARCA5 maintains nucleosome structure around CTCF motifs in Kasumi-1 cells.

(A) “Phasogram” of the nucleosome repeat length (NRL) during the time course of SMARCA5 degradation (red line, span = 0.1, and green line, span = 1.5) derived from the MNase-seq signal of a representative replicate (grey line) using the entire genome. (B) Bar graph quantification from biological replicates of the NRL calculated from the MNase-seq signal at each of the indicated time points. (C) “Phasogram” of the NRL of SMARCA5 degradation at indicated times around H3K27me3 peaks +/-1kb. (D) Bar graph quantification of NRL at H3K27me3 peaks from biological replicates calculated using the MNase-seq signal at each of the indicated time points. (E) “Phasogram” of the NRL during the time course of SMARCA5 degradation around H2A.Z CUT&RUN peaks +/-1kb. (F) Bar graph quantification of the NRL at H2A.Z peaks of biological replicates calculated from the MNase-seq signal at each of the indicated time points. (G) “Phasogram” of the NRL at the indicated times within 1kb of the peak center of CTCF CUT&RUN peak motifs. H. Bar graph quantification of the NRL of biological replicates within 1kb of CTCF motifs at the indicated time points. I. Histograms of nucleosome occupancy using MNase-seq data plotted around all CTCF motif centers (31,015) from the CTCF CUT&RUN data. The signal was plotted +/- 500bp from the motif center and the changes after SMARCA5 degradation at the indicated time points are shown. For B, D, F, and H, the significance was calculated with Welch’s two sample t-test: ****=$p<0.0001$, ***=$p<0.001$, **=$p<0.01$, *=$p<0.05$; error bars represent standard deviation.

The nucleosome repeat length was shorter at sites occupied by H2A.Z than the global nucleosome repeat length, consistent with a more open configuration of H2A.Z-containing nucleosomes (Fig 3.13E, 0hr). While the phasing was less distinct at these areas of open chromatin, there was a large increase in nucleosome repeat length over the 24hr time course of SMARCA5 degradation at these regions (Fig 3.13E and 3.13F). Similarly, the nucleosome repeat length around CTCF motifs showed a rather short nucleosome repeat length that increased dramatically upon degradation of SMARCA5 (Fig 3.13G and 3.13H). Because of the apparent large changes in chromatin structure around CTCF sites upon SMARCA5 inactivation, we plotted the normalized nucleosome occupancy centered on the CTCF DNA binding motif under the CTCF peaks (Fig 3.13I). We found that the nucleosome phasing around the CTCF motifs was beginning to erode even at the 2hr time point and the phasing was significantly affected within the first 6hr of SMARCA5 degradation (Fig 3.13I). These data suggest that SMARCA5 is required to maintain the regular nucleosome spacing in concert with CTCF and H2A.Z, and that SMARCA5 acts across the genome and at different stages of the cell cycle as a nucleosome ruler to prevent undue chromatin compaction.

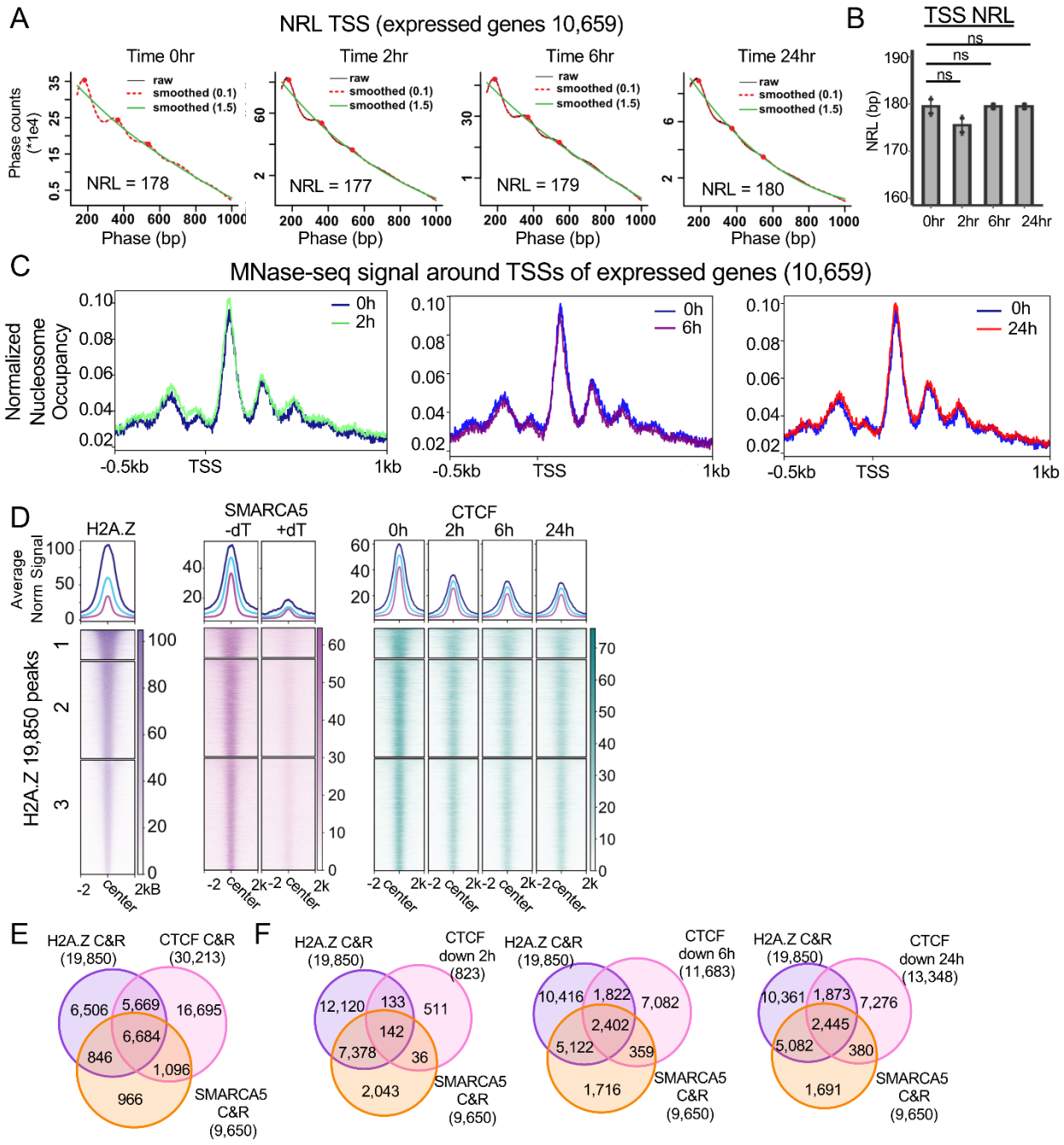
A
NRL TSS (expressed genes 10,659)
Time 0hr
Time 2hr
Time 6hr
Time 24hr
raw
smoothed (0.1)
smoothed (1.5)
Phase counts (*1e4)
NRL = 178
NRL = 177
NRL = 179
NRL = 180
Phase (bp)
B
TSS NRL
NRL (bp)
ns
0hr 2hr 6hr 24hr
C
MNase-seq signal around TSSs of expressed genes (10,659)
Normalized Nucleosome Occupancy
0h
2h
6h
24h
-0.5kb
TSS
1kb
D
H2A.Z
SMARCA5
-dT
+dT
CTCF
0h
2h
6h
24h
Average Norm Signal
H2A.Z 19,850 peaks
-2
center
2kB
2k
E
H2A.Z C&R (19,850)
CTCF C&R (30,213)
SMARCA5 C&R (9,650)
6,506
5,669
16,695
6,684
846
1,096
966
F
CTCF down 2h (823)
CTCF down 6h (11,683)
CTCF down 24h (13,348)
12,120
133
511
142
7,378
36
2,043
10,416
1,822
7,082
2,402
5,122
359
1,716
10,361
1,873
7,276
2,445
5,082
380
1,691

Figure 3.14: SMARCA5 and CTCF localize around H2A.Z sites.

(A) “Phasogram” of the nucleosome repeat length using the MNase-seq signal of a representative replicate after degradation of SMARCA5 at the indicated times +/-1kb from the TSSs of all expressed genes, as determined from the PRO-seq data in Figure 5. (B) Bar graph quantification of the nucleosome repeat lengths of biological replicates calculated using the around the TSS of the expressed genes at each of the indicated time points. (C) Histograms of nucleosome occupancy using MNase-seq data plotted from 500bp upstream to 1000bp downstream of all expressed TSSs from the PRO-seq data in Figure 5. (D) Heatmaps of H2A.Z, SMARCA5, and CTCF CUT&RUN at the indicated timepoints around H2A.Z – containing peaks +/- 2kb. H2A.Z sites were clustered by k-means using deepTools. (E) Venn diagram showing the overlap between H2A.Z and SMARCA5 with all CTCF binding sites. (F) Venn diagrams showing the overlap between H2A.Z and SMARCA5 with CTCF binding sites lost at 2, 6, and 24hrs in Kasumi-1 cells.

3.3 Discussion

SMARCA5 is an ATP-dependent helicase that is able to slide nucleosomes and order nucleosome arrays *in vitro* to act as a “ruler” (Aalfs et al., 2001). Genetic studies using gene deletion or siRNA inhibition have implicated SMARCA5 as being required for the control of gene expression, DNA replication and repair, chromatin accessibility, as well as nucleosome spacing and even higher order chromatin looping (Aydin et al., 2014; Barisic et al., 2019; Cavellán et al., 2006; Dluhosova et al., 2014; Hendy et al., 2022; Jimeno-González et al., 2015; Kokavec et al., 2017; Längst and Becker, 2001). However, these steady-state analyses performed a few days to weeks after inactivation of *SMARCA5* were unable to differentiate direct versus indirect effects. By engineering the endogenous alleles of SMARCA5 for rapid degradation in mammalian cells, we established that SMARCA5 is continuously required to maintain nucleosomal spacing throughout the genome. This implies that nucleosomes are constantly moving or being moved, and that SMARCA5 was required to push these nucleosomes together to maintain more relaxed or accessible chromatin. These effects were independent of the phase of the cell cycle or whether the chromatin was more open or closed and suggests that chromatin is highly dynamic. This is in agreement with recent findings that individual loci are far more dynamic than previously envisioned (Gabriele et al., 2022).

Genetic models have established CTCF as a central factor in SMARCA5-mediated chromatin regulation (Barisic et al., 2019; Dluhosova et al., 2014; Wiechens et al., 2016). CTCF is required to establish chromatin loops and transcription activation domains or TADs, which are a key feature of chromosome organization (Barisic et al., 2019; Gabriele et al., 2022; Khoury et al., 2020; Luan et al., 2021; Nora et al., 2017). However, because

these genetic models reached a steady state in which nearly all CTCF DNA binding was lost (Barisic et al., 2019), these genetic approaches did not have the ability to link SMARCA5 to the loss of CTCF. Our time course approach showed that rapid degradation of SMARCA5 led to a relatively quick decline in all CTCF DNA binding with about a third of the CTCF sites being significantly reduced within the first 6hr of dTAG-47 addition (Fig 3.5A-3.5C). In addition, our CUT&RUN analysis indicated that roughly 70% of the SMARCA5 peaks overlapped with a CTCF and a H2A.Z peak (Fig 3.14E) and 70-80% of the ATAC-seq peaks that decreased overlapped with a CTCF binding site (Fig 3.7F-3.7H). Interestingly, the ATAC-seq peaks that were gained were not linked to CTCF or another DNA binding factor, indicating that these sites were gained stochastically. Thus, we hypothesize that inactivation of SMARCA5 allowed nucleosomes, likely containing H2A.Z to be positioned over the CTCF binding sites causing a rapid loss of CTCF binding and eventually a loss of DNA looping as observed in mESCs (Barisic et al., 2019). This is consistent with a rapid loss of nucleosomal phasing around CTCF binding sites (Fig 3.13I).

The use of a degron tag allowed us to synchronize the cells prior to SMARCA5 degradation, which indicated that SMARCA5 contributed to maintaining chromatin accessibility at the G_1/S phase boundary, mid S phase and late S/G_2 phase (Fig 3.11). Thus, SMARCA5 may contribute to maintaining nucleosomal phasing around CTCF and H2A.Z sites during multiple phases of the cell cycle. Indeed, even during the S phase, the ATAC-seq peaks that were lost were strongly associated with CTCF binding motifs (data not shown). These data argue that the impairment that we observed in Kasumi-1 cells late in the S phase is unlikely to be due to an effect in the G_1 or G_2/M phases of the cell

cycle and point to a more direct role for SMARCA5 in the S phase (Barisic et al., 2019; Ding et al., 2021a, 2021c; Dluhosova et al., 2014; Levendosky and Bowman, 2019; Stopka et al., 2000; Zikmund et al., 2020). Interestingly, Kasumi-1 cells have lost an allele of *RAD21* (Ghandi et al., 2019), which raises the possibility that disruptions of nucleosome repeat length or CTCF binding could be synthetic lethal with cohesin mutations, which are prevalent in various types of leukemia and solid cancers. Given that the defect in S phase transit and the impaired growth were more evident in the t(8;21) cell line (Kasumi-1), than in the erythroleukemia cell line HEL or DLBCL cell line OCI-LY1 (Fig 3.1), it is possible that SMARCA5 could be a therapeutic target in t(8;21) AML or other types of cancer containing cohesin mutations (DepMap, 2017; Hanahan and Weinberg, 2011; Negrini et al., 2010; Tsherniak et al., 2017).

4. CHAPTER 4: Conclusions and Future Directions

4.1 Conclusions

The more we know about basic biology, biochemistry, and cellular functions of TFs and chromatin remodeling enzymes, the more equipped we are to understand the basic mechanisms underlying disease initiation and development. Chromatin structure and transcription are constantly dysregulated in cancer and other diseases. My dissertation work developed a more detailed understanding of the chromatin remodeling ATPase, SMARCA5, in human cells. To describe the work put into my dissertation, I would say, "perseverance is key," as nothing in science is as easy as it seems. I used a combination of new degron tags, CRISPR Cas9 editing, and next-generation sequencing (NGS) to answer the question of whether human SMARCA5 regulates transcription, chromatin structure, and transcription factor binding.

The new technology of PROTACs and degron tags provide innovative methods to determine the direct effects after loss of a specific protein (Erb et al., 2017; Weintraub et al., 2017). I successfully created a C-terminal degron-tagged SMARCA5 in three different cell lines: the acute myeloid leukemia cell line Kasumi-1, the erythroleukemia cell line HEL, and the diffuse large b-cell lymphoma cell line OCI-LY1. The endogenous human SMARCA5 protein was removed from the system after just two hours of dTAG-47 treatment, creating an efficient and rapid system to study SMARCA5 compared to genetic knockdown/knockout systems (Barisic et al., 2019; Bhaskara et al., 2013; Zikmund et al., 2019).

Incorporation of a FLAG tag into *SMARCA5FKBP12F36-FLAG* allowed me to determine SMARCA5 localization throughout the genome. The downside to determining

where SMARCA5 would be binding was that it does not contain a DNA-binding domain, unlike many TFs. Therefore, the CUT&RUN localization data had fewer peaks and higher background. However, I did find that the SMARCA5 peaks contained CTCF motifs, similar to previously published SMARCA5 ChIP-seq data and downregulated MNase-seq peaks after *Smarca5* knockdown or knockout (Barisic et al., 2019; Luan et al., 2021; Wiechens et al., 2016). My dissertation work validated that SMARCA5 binds and functions directly around CTCF motifs in human cell lines.

Chromatin remodeling enzymes can function at promoters, enhancers, exons, intergenic regions, or intronic regions of the genome to slide, evict, or insert nucleosomes along the DNA (Swygert and Peterson, 2014). This can lead to direct changes in transcription if the chromatin structure changes around genes and their transcription factor binding sites (Stengel et al., 2021). Our lab used PRO-seq and RNA-seq in combination with inhibitor treatments to study direct vs indirect transcriptional changes (Johnston et al., 2020; Sampathi et al., 2019; Wang et al., 2018; Zhao et al., 2016). The use of the degron tag with SMARCA5$^{FKBP12F36\text{-}FLAG}$ provided unique insights into the function of SMARCA5 during nascent transcription. Previous studies claimed that SMARCA5 regulates the transcription of thousands of genes, such as *MYC*, but these results were generated from knockout and overexpression assays (Barisic et al., 2019; Jimeno-González et al., 2015). By contrast, direct degradation of SMARCA5 over a short time course enabled us to measure small changes in nascent transcription and determine that few of these genes had overlapping SMARCA5-annotated peaks. This transcriptional outcome is similar to that of a recently published paper also utilizing PRO-seq, but with CTCF degradation with an AID system in the mouse erythroblast cell line G1E-ER4 (Luan

et al., 2021). These authors reported very few changes in transcription after CTCF degradation for 4 h (Luan et al., 2021). The knowledge that SMARCA5 binds and regulates CTCF motifs combined with the PRO-seq data sets suggests that SMARCA5 does not have a direct role in transcription and knockdown/knockout models measure secondary effects after days or weeks of SMARCA5 loss (Barisic et al., 2019; Wiechens et al., 2016).

As a chromatin remodeling ATPase, SMARCA5 slides nucleosomes along the DNA *in vitro*, and *Smarca5* knockout/knockdown reduces nucleosome phasing around CTCF motifs as measured by MNase-seq (Dluhosova et al., 2014; Oppikofer et al., 2017; Wiechens et al., 2016). We used a more straightforward method to measure NFRs via ATAC-seq after SMARCA5 degradation in all three cell lines, and observed significant overlaps between downregulated peaks, but not upregulated peaks, in all three cell lines. The upregulated peaks could be more cell-type specific and require more detailed investigation. The downregulated peaks were enriched for CTCF motifs by almost 15 to 20-fold, which is consistent with the *Smarca5* knockout data in mESCs.

SMARCA5 is associated with CTCF and RAD21, which are part of the cohesion complex (Dluhosova et al., 2014; Hakimi et al., 2002). *Smarca5* knockout in mESCs led to a global loss of topologically associated domains (TADs), which resembles the loss of CTCF (Barisic et al., 2019; Luan et al., 2021). SMARCA5 degradation resulted in a rapid loss of CTCF binding over time in all three cell lines. Interestingly, the heterozygous frameshift deletion in RAD21 did not affect the outcome of CTCF binding loss in the Kasumi-1 cell line compared to that in the OCI-LY1 or HEL cell lines. There was no loss in RUNX1 TF binding in Kasumi-1, and the loss of AML1-ETO binding was only at the

CTCF binding or CTCF motifs. These studies confirm the results of Barisic et al. (2019) in knockout mESCs and are consistent with the conclusion that SMARCA5 specifically regulates CTCF binding in mouse and human cells.

To regulate chromatin structure, SMARCA5 slides nucleosomes along the DNA *in vitro* and *in vivo* to create a set nucleosome repeat length (NRL) (Barisic et al., 2019; Clarkson et al., 2019; He et al., 2006; Levendosky and Bowman, 2019; Poot et al., 2000). I confirmed a global increase in NRL measured by ATAC-seq in all three cell lines after 6 h of SMARCA5 degradation. The observed increase in NRL was not specific to a specific cell cycle when NRL was measured in synchronous cells. We expected a specific increase in NRL at 6 h or 12 h after release due to the increased number of cells in S phase after BrdU incorporation and previously published data showing that SMARCA5 is important for DNA replication after knockdown (Bhaskara et al., 2013). My dissertation work demonstrated that there is a rapid increase in NRL over the SMARCA5 degradation time course of 2, 6, and 24 h, which is associated with SMARCA5 loss in human cells and is independent of the cell cycle.

CTCF binding sites with a higher binding affinity have a lower NRL, suggesting that CTCF requires remodeling around CTCF motifs for the best binding (Clarkson et al., 2019). SMARCA5 co-bound to CTCF binding sites also showed a reduction in the NRL compared to other chromatin remodeling ATPases bound to CTCF sites (Clarkson et al., 2019). We observed a significant increase in NRL around CTCF motifs after SMARCA5 degradation, which also resulted in decreased phasing around CTCF motifs.

SMARCA5 preferentially remodels nucleosomes containing the histone variant H2A.Z over those containing H2A (Goldman et al., 2010). CTCF binding sites are

enriched for H2A.Z rather than H2A histones, thereby stimulating CTCF binding (Wen et al., 2020). We observed that H2A.Z was localized around CTCF and SMARCA5 binding sites, and H2A.Z sites had an increase in NRL. These data suggest that SMARCA5 remodels nucleosomes containing the H2A.Z histone variant, thereby organizing nucleosomes around CTCF motifs for CTCF binding (Fig 4.1).

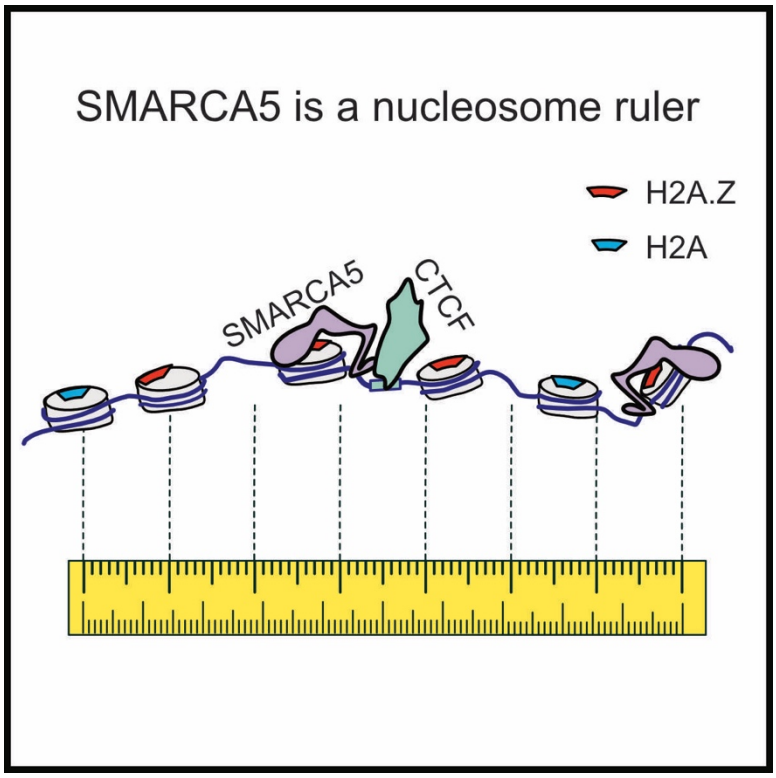

Figure 4.1: SMARCA5 functions as a nucleosome spacing ruler in the nucleus.

Graphical abstract of SMARCA5 sliding nucleosomes containing H2A.Z around CTCF motifs, thereby enabling the binding of CTCF.

4.2 Future directions

There are still many avenues to study to understand the mechanisms of SMARCA5 in chromatin structure and function more deeply. The PROTAC models developed in Chapter 3 provide a great tool to quickly degrade SMARCA5 in combination with sequencing and functional studies. The interesting questions to ask are: Does CTCF regulate SMARCA5 binding and function; Which comes first, H2A.Z or CTCF; Does SMARCA5 regulate chromatin condensation during mitosis; Do external treatments affect CTCF binding; What chromatin remodeling enzymes oppose SMARCA5.

SMARCA5 and CTCF function in maintaining chromatin structure are interrelated, however the dynamics of this relationship on a molecular and mechanistic level remain unclear. It is well established that SMARCA5 and CTCF associate, however we need to unravel how removing one protein from this interaction contributes to their function in cells (Dluhosova et al., 2014; Song et al., 2022). To start to determine whether CTCF regulates SMARCA5 function and binding, we would engineer an endogenous HALO-tagged CTCF in the SMARCA5-FKBP12-1xFLAG cell lines. The first experiment to test whether CTCF regulates SMARCA5 binding by performing CUT&RUN for the flag tag on the endogenous SMARCA5 after degradation of CTCF. Next, to compare the function of CTCF vs SMARCA5 in chromatin structure, we would perform ATAC-seq after a CTCF degradation time course. Comparing the differential changes in NFRs after CTCF degradation vs SMARCA5 degradation would provide information as to which protein regulates NFRs around CTCF sites. Furthermore, we can utilize this system to determine how H2A.Z localization changes around CTCF sites. CUT&RUN following degradation time courses of SMARCA5, CTCF or both can provide insight as to the importance of H2A.Z around

CTCF motifs. Together these experiments can answer the question does CTCF regulate SMARCA5 structure and function.

Wash-out experiments in the Kasumi-1, HEL, and OCI-LY1 cell lines could provide new information describing how CTCF or H2A.Z re-localizes to chromatin and how NFRs change. Performing washouts after CTCF degradation and/or SMARCA5 degradation will provide direct information as to how the chromatin structure and proteins around the CTCF motifs are re-established. This could determine whether there is a point of no return for chromatin structure after SMARCA5 degradation, or whether the effects of SMARCA5 loss can be reversed. We could also answer the question of which comes first H2A.Z or CTCF.

CTCF degradation before release from nocodazole treatment results in an inability to form A and B compartments of TADs in G_1, so it is possible SMARCA5 may play a role in chromatin condensation (Zhang et al., 2021). A nocodazole treatment combined with SMARCA5 degradation, and ATAC-seq would answer how SMARCA5 regulates NRL during and after chromatin condensation. The use of CUT&RUN of CTCF during and after release from the nocodazole treatment would determine if loss SMARCA5 creates the same phenotype as loss of CTCF in the A and B compartment formation. Overall it would provide a better understanding of what cell cycle SMARCA5 functions during.

HEL erythroblast cells can be treated with phorbol myristate acetate (PMA) to induce megakaryocyte differentiation. We saw Kasumi-1 myeloblast cells led to differentiation into $CD11b^+$ macrophages after a three-day SMARCA5 degradation period; however, we did not look at the differentiation of HEL cells into megakaryocytes as they did not slow growth till 5 or 6 days. The use of PMA treatment on HEL cells in combination with and

without SMARCA5 degradation would allow us to ask if SMARCA5 degradation can lead to increased differentiation after PMA treatment. This will start to determine if SMARCA5 degradation increases the rate of megakaryocyte differentiation ($CD41^+/CD42^+$), as SMARCA5 degradation in Kasumi-1 myeloblast cells leads to differentiation into $CD11b^+$ macrophages.

Not only can we look at differentiation effects after SMARCA5 degradation, but we can see how PMA-induced differentiation of HEL cells can affect NRL and CTCF localization. Performing CUT&RUN for CTCF with and without degradation of SMARCA5 and of PMA treatment could provide insight as to whether differentiation can change CTCF binding independent of SMARCA5. ATAC-seq would also provide NRL information with and without PMA-induced differentiation, asking the question does differentiation affect NRL independent of SMARCA5 degradation. These data could provide a foundation to determine how SMARCA5 functions in hematopoietic differentiation and if CTCF could also play a role.

There must be a chromatin remodeling enzyme opposing SMARCA5, which brings the nucleosomes closer together. The other ISWI family member, SMARCA1, is not expressed in the cell lines we used, so we can conclude something else is opposing SMARCA5 nucleosome sliding. The possible options are INO80, SMARCA4, or SMARCA2. We can easily test this with PROTACs against different chromatin remodeling factors, HALO tags, or knockouts of chromatin remolding factors before and after SMARCA5 degradation. The PROTAC ACBI-1 degrades SMARCA4 and SMARCA2, we can use this in combination with ATAC-seq to measure NRL with and without SMARCA5 degradation (Farnaby et al., 2019). To look directly at each chromatin remodeling

enzyme, we would first try to individually HALO tag each chromatin remodeling enzyme in the SMARCA5-FKBP12 cells line, creating 3 separate cell lines. We can secondly confirm HALO tag results by performing individual knockouts of INO80, SMARCA4, and SMARCA2. The loss of INO80, SMARCA4, or SMARCA2 in combination with ATAC-seq to measure NRL with and without SMARCA5 degradation will provide information on how they each affect NRL. Together, this data could inform us what chromatin remolding enzyme opposes SMARCA5 nucleosome sliding. Overall, this will start to untangle the intricate dynamics of chromatin remodeling enzymes in chromatin structure.

5. Appendix

5.1 Functions:

```
select_sig_genes <- function(gbpp)
  { gbpp %>%
  dplyr::filter(padj <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2FoldChange)))) %>%
  dplyr::filter(FoldChange >= 2) %>%
  dplyr::glimpse()}

select_sigup <- function(gbpp)
  { gbpp %>%
  dplyr::filter(padj <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2FoldChange)))) %>%
  dplyr::filter(FoldChange >= 2) %>%
  dplyr::filter(log2FoldChange > 0) %>%
  dplyr::glimpse()}

select_sigdn <- function(gbpp)
  { gbpp %>%
  dplyr::filter(padj <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2FoldChange)))) %>%
  dplyr::filter(FoldChange >= 2) %>%
  dplyr::filter(log2FoldChange < 0) %>%
```

```
  dplyr::glimpse()}

select_sigup_writefile <- function(gbpp, filename)
  { gbpp %>%
  dplyr::filter(padj <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2FoldChange)))) %>%
  dplyr::filter(FoldChange >= 2) %>%
  dplyr::filter(log2FoldChange > 0) %>%
  dplyr::glimpse() %>%
  write_tsv(filename) }

select_sigdn_writefile <- function(gbpp, filename)
  { gbpp %>%
  dplyr::filter(padj <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2FoldChange)))) %>%
  dplyr::filter(FoldChange >= 2) %>%
  dplyr::filter(log2FoldChange < 0) %>%
  dplyr::glimpse() %>%
  write_tsv(filename) }

select_sigup1.5_writefile <- function(gbpp, filename)
  { gbpp %>%
  dplyr::filter(padj <= 0.05) %>%
```

```
  dplyr::mutate(FoldChange = (2^(abs(log2FoldChange)))) %>%
  dplyr::filter(FoldChange >= 1.5) %>%
  dplyr::filter(log2FoldChange > 0) %>%
  dplyr::glimpse() %>%
  write_tsv(filename) }

select_sigdn1.5_writefile <- function(gbpp, filename)
  { gbpp %>%
  dplyr::filter(padj <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2FoldChange)))) %>%
  dplyr::filter(FoldChange >= 1.5) %>%
  dplyr::filter(log2FoldChange < 0) %>%
  dplyr::glimpse() %>%
  write_tsv(filename) }

select_sigupFDR <- function(gbpp, filename)
  { gbpp %>%
  dplyr::filter(FDR <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2fc)))) %>%
  dplyr::filter(FoldChange >= 2) %>%
  dplyr::filter(log2fc > 0) %>%
  dplyr::glimpse() %>%
  write_tsv(filename) }
```

```
select_sigdnFDR <- function(gbpp, filename)
  { gbpp %>%
  dplyr::filter(FDR <= 0.05) %>%
  dplyr::mutate(FoldChange = (2^(abs(log2fc)))) %>%
  dplyr::filter(FoldChange >= 2) %>%
  dplyr::filter(log2fc < 0) %>%
  dplyr::glimpse() %>%
  write_tsv(filename) }

select_clusters <- function(df,cluster, filename) {df %>%
    dplyr::filter(deepTools_group == cluster) %>%
    write_tsv(filename)}
```

6. References

Aalfs JD, Narlikar GJ, Kingston RE. 2001. Functional Differences between the Human ATP-dependent Nucleosome Remodeling Proteins BRG1 and SNF2H. *Journal of Biological Chemistry* **276**:34270–34278. doi:10.1074/jbc.M104163200

Abrams E, Neigeborn L, Carlson M. 1986. Molecular analysis of SNF2 and SNF5, genes required for expression of glucose-repressible genes in Saccharomyces cerevisiae. *Mol Cell Biol* **6**:3643–3651. doi:10.1128/mcb.6.11.3643

Aihara T, Miyoshi Y, Koyama K, Suzuki M, Takahashi E, Monden M, Nakamura Y. 1998. Cloning and mapping of SMARCA5 encoding hSNF2H, a novel human homologue of drosophila ISWI. *Cytogenet Cell Genet* **81**:191–193. doi:10.1159/000015027

Allen BL, Taatjes DJ. 2015. The Mediator complex: a central integrator of transcription. *Nature Reviews Molecular Cell Biology 2015 16:3* **16**:155–166. doi:10.1038/nrm3951

Allfrey VG, Faulkner R, Mirsky AE. 1964. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proceedings of the National Academy of Sciences* **51**:786 LP – 794. doi:10.1073/pnas.51.5.786

Avet-Loiseau H, Attal M, Moreau P, Charbonnel C, Garban F, Hulin C, Leyvraz S, Michallet M, Yakoub-Agha I, Garderet L, Marit G, Michaux L, Voillat L, Renaud M, Grosbois B, Guillerm G, Benboubker L, Monconduit M, Thieblemont C, Casassus P, Caillot D, Stoppa AM, Sotto JJ, Wetterwald M, Dumontet C, Fuzibet JG, Azais I, Dorvaux V, Zandecki M, Bataille R, Min-Vielle S, Harousseau JL, Facon T, Mathiot C, Durie BG, Kyle RA, Belch A, Bensinger W, Blade J, Boccadoro M, Child JA,

Comenzo R, Djulbegovic B, Fantl D, Gahrton G, Hungria V, Joshua D, Ludwig H, Mehta J, Morales AR, Morgan G, Nouel A, Oken M, Powles R, Rood-Man D, San Miguel J, Shimizu K, Singhal S, Sirohi B, Sonneveld P, Tricot G, Van Ness B. 2009. The IKZF3 (Aiolos) transcription factor is highly upregulated and inversely correlated with clinical progression in chronic lymphocytic leukaemia. *Br J Haematol* **144**:268–270. doi:10.1111/J.1365-2141.2008.07442.X

Aydin ÖZ, Marteijn JA, Ribeiro-Silva C, Rodríguez López A, Wijgers N, Smeenk G, van Attikum H, Poot RA, Vermeulen W, Lans H. 2014. Human ISWI complexes are targeted by SMARCA5 ATPase and SLIDE domains to help resolve lesion-stalled transcription. *Nucleic Acids Res* **42**:8473–8485. doi:10.1093/nar/gku565

Azmi IF, Watanabe S, Maloney MF, Kang S, Belsky JA, MacAlpine DM, Peterson CL, Bell SP. 2017. Nucleosomes influence multiple steps during replication initiation. *Elife* **6**:e22512. doi:10.7554/eLife.22512

Bakayev V v, Georgiev GP. 1976. Heterogeneity of Chromatin Subunits in Vitro and Location of Histone HI. *Nucleic Acids Res* **3**:477–492. doi:10.1093/nar/3.2.477

Baldi Sandro, Korber P, Becker PB. 2020. Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nat Struct Mol Biol* **27**:109–118. doi:10.1038/s41594-019-0368-x

Baldi S., Korber P, Becker PB. 2020. Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nat Struct Mol Biol* **27**:109–118. doi:https://doi.org/10.1038/s41594-019-0368-x

Barak O, Lazzaro MA, Lane WS, Speicher DW, Picketts DJ, Shiekhattar R. 2003. Isolation of human NURF: A regulator of Engrailed gene expression. *EMBO Journal* **22**:6089–6100. doi:10.1093/emboj/cdg582

Barisic D, Stadler MB, Iurlaro M, Schübeler D. 2019. Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors. *Nature* **569**:136–140. doi:10.1038/s41586-019-1115-5

Berbenetz NM, Nislow C, Brown GW. 2010. Diversity of Eukaryotic DNA Replication Origins Revealed by Genome-Wide Analysis of Chromatin Structure. *PLoS Genet* **6**:e1001092.

Bhaskara S, Jacques V, Rusche JR, Olson EN, Cairns BR, Chandrasekharan MB. 2013. Histone deacetylases 1 and 2 maintain S-phase chromatin and DNA replication fork progression. *Epigenetics Chromatin* **6**:1–21. doi:10.1186/1756-8935-6-27

Bomber ML, Wang J, Liu Q, Hodges E, Stengel KR, Correspondence SWH, Barnett KR, Layden HM, Hiebert SW. 2022. Human SMARCA5 is continuously required to maintain nucleosome spacing. *Mol Cell* **83**:2023. doi:10.1016/j.molcel.2022.12.018

Bozhenok L, Wade PA, Varga-Weisz P. 2002. WSTF-ISWI chromatin remodeling complex targets heterochromatic replication foci. *EMBO Journal* **21**:2231–2241. doi:10.1093/emboj/21.9.2231

Brenner S, Jacob F, Meselson M. 1961. An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. *Nature* **190**:576–581. doi:10.1038/190576a0

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-

binding proteins and nucleosome position. *Nat Methods* **10**:1213–1218. doi:10.1038/nmeth.2688

Cavellán E, Asp P, Percipalle P, Farrants AKÖ. 2006. The WSTF-SNF2h chromatin remodeling complex interacts with several nuclear proteins in transcription. *Journal of Biological Chemistry* **281**:16264–16271. doi:10.1074/jbc.M600233200

Chen X, Yin X, Li J, Wu Z, Qi Y, Wang X, Liu W, Xu Y. 2021. Structures of the human Mediator and Mediator-bound preinitiation complex. *Science (1979)* **372**. doi:10.1126/SCIENCE.ABG0635/SUPPL_FILE/ABG0635S9.MP4

Chen Y, Yamaguchi Y, Tsugeno Y, Yamamoto J, Yamada T, Nakamura M, Hisatake K, Handa H. 2009. DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. *Genes Dev* **23**:2765–2777. doi:10.1101/gad.1834709

Clackson T, Yang W, Rozamus LW, Hatada M, Amara JF, Rollins CT, Stevenson LF, Magari SR, Wood SA, Courage NL, Lu X, Cerasoli F, Gilman M, Holt DA. 1998. Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc Natl Acad Sci U S A* **95**:10437–10442. doi:10.1073/pnas.95.18.10437

Clapier CR, Cairns BR. 2012. Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes. *Nature* **492**:280–284. doi:10.1038/nature11625

Clarkson CT, Deeks EA, Samarista R, Mamayusupova H, Zhurkin VB, Teif VB. 2019. CTCF-dependent chromatin boundaries formed by asymmetric nucleosome arrays with decreased linker length. *Nucleic Acids Res* **47**:11181–11196. doi:10.1093/nar/gkz908

Collins N, Poot RA, Kukimoto I, García-Jiménez C, Dellaire G, Varga-Weisz PD. 2002. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nat Genet* **32**:627–632. doi:10.1038/ng1046

Crick F. 1970. Central Dogma of Molecular Biology. *Nature* **227**:561–563. doi:10.1038/227561a0

Crick F. 1958. On protein synthesis. *Symp Soc Exp Biol* **12**:138–163.

Dao HT, Dul BE, Dann GP, Liszczak GP, Muir TW. 2020. A basic motif anchoring ISWI to nucleosome acidic patch regulates nucleosome spacing. *Nat Chem Biol* **16**:134–142. doi:10.1038/s41589-019-0413-4

Deindl S, Hwang WL, Hota SK, Blosser TR, Prasad P, Bartholomew B, Zhuang X. 2013. ISWI Remodelers Slide Nucleosomes with Coordinated Multi-Base-Pair Entry Steps and Single-Base-Pair Exit Steps. *Cell* **152**:442–452. doi:10.1016/j.cell.2012.12.040

DepMap. 2017. The Cancer Dependency Map Consortium. *Nat Genet*. doi:10.1038/ng.3984

Ding Y, Li Y, Zhao Z, Cliff Zhang Q, Liu F. 2021a. The chromatin-remodeling enzyme smarca5 regulates erythrocyte aggregation via keap1-nrf2 signaling. *Elife* **10**. doi:10.7554/ELIFE.72557

Ding Y, Wang W, Ma D, Liang G, Kang Z, Xue Y, Zhang Yifan, Wang L, Heng J, Zhang Yong, Liu F. 2021b. Smarca5-mediated epigenetic programming facilitates fetal HSPC development in vertebrates. *Blood* **137**:190–202. doi:10.1182/blood.2020005219

Ding Y, Wang W, Ma D, Liang G, Kang Z, Xue Y, Zhang Yifan, Wang L, Heng J, Zhang Yong, Liu F. 2021c. Smarca5-mediated epigenetic programming facilitates fetal

HSPC development in vertebrates. *Blood* **137**:190. doi:10.1182/BLOOD.2020005219

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, Diao Y, Liang J, Zhao H, Lobanenkov V V., Ecker JR, Thomson JA, Ren B. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature 2015 518:7539* **518**:331–336. doi:10.1038/nature14222

Dluhosova M, Curik N, Vargova J, Jonasova A, Zikmund T, Stopka T. 2014. Epigenetic control of SPI1 gene by CTCF and ISWI ATPase SMARCA5. *PLoS One* **9**:e87448:1-11. doi:10.1371/journal.pone.0087448

Eaton JD, Davidson L, Bauer DLV, Natsume T, Kanemaki MT, West S. 2018. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes Dev* **32**:127–139. doi:10.1101/GAD.308528.117/-/DC1

Eberhardy SR, Farnham PJ. 2001. c-Myc Mediates Activation of the cad Promoter via a Post-RNA Polymerase II Recruitment Mechanism. *Journal of Biological Chemistry* **276**:48562–48571. doi:10.1074/jbc.M109014200

Elfring LK, Deuring R, McCallum CM, Peterson CL, Tamkun JW. 1994. Identification and characterization of Drosophila relatives of the yeast transcriptional activator SNF2/SWI2. *Mol Cell Biol* **14**:2225–2234. doi:10.1128/mcb.14.4.2225

Erb MA, Scott TG, Li BE, Xie H, Paulk J, Seo HS, Souza A, Roberts JM, Dastjerdi S, Buckley DL, Sanjana NE, Shalem O, Nabet B, Zeid R, Offei-Addo NK, Dhe-Paganon S, Zhang F, Orkin SH, Winter GE, Bradner JE. 2017. Transcription control by the ENL YEATS domain in acute leukaemia. *Nature* **543**:270–274. doi:10.1038/nature21688

Erdel F, Rippe K. 2011. Chromatin remodelling in mammalian cells by ISWI-type complexes - Where, when and why? *FEBS Journal*. doi:10.1111/j.1742-4658.2011.08282.x

Evrin C, Clarke P, Zech J, Lurz R, Sun J, Uhle S, Li H, Stillman B, Speck C. 2009. A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proceedings of the National Academy of Sciences* **106**:20240 LP – 20245. doi:10.1073/pnas.0911500106

Farnaby W, Koegl M, Roy MJ, Whitworth C, Diers E, Trainor N, Zollman D, Steurer S, Karolyi-Oezguer J, Riedmueller C, Gmaschitz T, Wachter J, Dank C, Galant M, Sharps B, Rumpel K, Traxler E, Gerstberger T, Schnitzer R, Petermann O, Greb P, Weinstabl H, Bader G, Zoephel A, Weiss-Puxbaum A, Ehrenhöfer-Wölfer K, Wöhrle S, Boehmelt G, Rinnenthal J, Arnhof H, Wiechens N, Wu MY, Owen-Hughes T, Ettmayer P, Pearson M, McConnell DB, Ciulli A. 2019. BAF complex vulnerabilities in cancer demonstrated via structure-based PROTAC design. *Nature Chemical Biology 2019 15:7* **15**:672–680. doi:10.1038/s41589-019-0294-6

Formosa T. 2012. The role of FACT in making and breaking nucleosomes. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1819**:247–255. doi:https://doi.org/10.1016/j.bbagrm.2011.07.009

Fu Y, Sinha M, Peterson CL, Weng Z. 2008. The Insulator Binding Protein CTCF Positions 20 Nucleosomes around Its Binding Sites across the Human Genome. *PLoS Genet* **4**:e1000138. doi:10.1371/JOURNAL.PGEN.1000138

Gabriele M, Brandão HB, Grosse-Holz S, Jha A, Dailey GM, Cattoglio C, Hsieh THS, Mirny L, Zechner C, Hansen AS. 2022. Dynamics of CTCF- and cohesin-mediated

chromatin looping revealed by live-cell imaging. *Science (1979)* **376**:476–501. doi:10.1126/science.abn6583

Gamarra N, Johnson SL, Trnka MJ, Burlingame AL, Narlikar GJ. 2018. The nucleosomal acidic patch relieves auto-inhibition by the ISWI remodeler SNF2h. *Elife* **7**. doi:10.7554/ELIFE.35322

Ghandi M, Huang FW, Jané-Valbuena J, Kryukov G v., Lo CC, McDonald ER, Barretina J, Gelfand ET, Bielski CM, Li H, Hu K, Andreev-Drakhlin AY, Kim J, Hess JM, Haas BJ, Aguet F, Weir BA, Rothberg M v., Paolella BR, Lawrence MS, Akbani R, Lu Y, Tiv HL, Gokhale PC, de Weck A, Mansour AA, Oh C, Shih J, Hadi K, Rosen Y, Bistline J, Venkatesan K, Reddy A, Sonkin D, Liu M, Lehar J, Korn JM, Porter DA, Jones MD, Golji J, Caponigro G, Taylor JE, Dunning CM, Creech AL, Warren AC, McFarland JM, Zamanighomi M, Kauffmann A, Stransky N, Imielinski M, Maruvka YE, Cherniack AD, Tsherniak A, Vazquez F, Jaffe JD, Lane AA, Weinstock DM, Johannessen CM, Morrissey MP, Stegmeier F, Schlegel R, Hahn WC, Getz G, Mills GB, Boehm JS, Golub TR, Garraway LA, Sellers WR. 2019. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**:503–508. doi:10.1038/s41586-019-1186-3

Giaimo BD, Ferrante F, Herchenröther A, Hake SB, Borggrefe T. 2019. The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin* **12**. doi:10.1186/S13072-019-0274-9

Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, Bentley DL. 2009. TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal

domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol Cell Biol* **29**:5455–5464. doi:10.1128/MCB.00637-09

Goldman JA, Garlick JD, Kingston RE. 2010. Chromatin remodeling by imitation switch (ISWI) class ATP-dependent remodelers is stimulated by histone variant H2A.Z. *Journal of Biological Chemistry* **285**:4645–4651. doi:10.1074/jbc.M109.072348

Greaves IK, Rangasamy D, Ridgway P, Tremethick DJ. 2007. H2A.Z contributes to the unique 3D structure of the centromere. *Proc Natl Acad Sci U S A* **104**:525–530. doi:10.1073/PNAS.0607870104/SUPPL_FILE/07870FIG7.JPG

Grüne T, Brzeski J, Eberharter A, Clapier CR, Corona DFV, Becker PB, Müller CW. 2003. Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Mol Cell* **12**:449–460. doi:10.1016/S1097-2765(03)00273-9

Guetg C, Lienemann P, Sirri V, Grummt I, Hernandez-Verdun D, Hottiger MO, Fussenegger M, Santoro R. 2010. The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. *EMBO Journal* **29**:2135–2146. doi:10.1038/emboj.2010.17

Hakimi MA, Bochar DA, Schmiesing JA, Dong Y, Barak OG, Spelcher DW, Yokomori K, Shiekhattar R. 2002. A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature* **418**:994–998. doi:10.1038/nature01024

Hamiche A, Sandaltzopoulos R, Gdula DA, Wu C. 1999. ATP-Dependent Histone Octamer Sliding Mediated by the Chromatin Remodeling Complex NURF. *Cell* **97**:833–842. doi:https://doi.org/10.1016/S0092-8674(00)80796-5

Hanahan D, Weinberg RA. 2011. Hallmarks of Cancer: The Next Generation. *Cell* **144**:646–674. doi:https://doi.org/10.1016/j.cell.2011.02.013

Hargreaves DC, Horng T, Medzhitov R. 2009. Control of Inducible Gene Expression by Signal-Dependent Transcriptional Elongation. *Cell* **138**:129–145. doi:10.1016/J.CELL.2009.05.047

He X, Fan HY, Garlick JD, Kingston RE. 2008. Diverse regulation of SNF2h chromatin remodeling by noncatalytic subunits. *Biochemistry* **47**:7025–7033. doi:10.1021/bi702304p

He X, Fan HY, Narlikar GJ, Kingston RE. 2006. Human ACF1 Alters the Remodeling Strategy of SNF2h. *Journal of Biological Chemistry* **281**:28636–28647. doi:10.1074/JBC.M603008200

Helfricht A, Wiegant WW, Thijssen PE, Vertegaal AC, Luijsterburg MS, van Attikum H. 2013. Remodeling and spacing factor 1 (RSF1) deposits centromere proteins at DNA double-strand breaks to promote non-homologous end-joining. *Cell Cycle* **12**:3070–3082. doi:10.4161/cc.26033

Hendy O, Serebreni L, Bergauer K, Muerdter F, Huber L, Nemčko F, Stark A. 2022. Developmental and housekeeping transcriptional programs in Drosophila require distinct chromatin remodelers. *Mol Cell* **82**:3598-3612.e7. doi:10.1016/J.MOLCEL.2022.08.019

Hirschhorn JN, Brown SA, Clark CD, Winston F. 1992. Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes Dev* **6**:2288–2298. doi:10.1101/gad.6.12a.2288

Hou H, Wang Y, Kallgren SP, Thompson J, Yates JR, Jia S. 2010. Histone variant H2A.Z regulates centromere silencing and chromosome segregation in fission yeast. *J Biol Chem* **285**:1909–1918. doi:10.1074/JBC.M109.058487

Hwang WL, Deindl S, Harada BT, Zhuang X. 2014. Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. *Nature* **512**:213–217. doi:10.1038/nature13380

Iestyn W, Chris S, Andrew F, D. SM, Tom O-H. 2003. Evidence for DNA Translocation by the ISWI Chromatin-Remodeling Enzyme. *Mol Cell Biol* **23**:1935–1945. doi:10.1128/MCB.23.6.1935-1945.2003

Ito T, Ando H, Suzuki T, Ogura T, Hotta K, Imamura Y, Yamaguchi Y, Handa H. 2010. Identification of a primary target of thalidomide teratogenicity. *Science (1979)* **327**:1345–1350. doi:10.1126/SCIENCE.1177319/SUPPL_FILE/ITO.SOM.REVISION1.PDF

Ito T, Bulger M, Pazin MJ, Kobayashi R, Kadonaga JT. 1997. ACF, an ISWI-Containing and ATP-Utilizing Chromatin Assembly and Remodeling Factor. *Cell* **90**:145–155. doi:10.1016/S0092-8674(00)80321-9

Jimeno-González S, Ceballos-Chávez M, Reyes JC. 2015. A positioned +1 nucleosome enhances promoter-proximal pausing. *Nucleic Acids Res* **43**:3068–3078. doi:10.1093/nar/gkv149

Johnston G, Ramsey HE, Liu Q, Wang J, Stengel KR, Sampathi S, Acharya P, Arrate M, Stubbs MC, Burn T, Savona MR, Hiebert SW. 2020. Nascent transcript and single-cell RNA-seq analysis defines the mechanism of action of the LSD1 inhibitor INCB059872 in myeloid leukemia. *Gene* **752**:144758. doi:10.1016/j.gene.2020.144758

Kaushal A, Mohana G, Dorier J, Özdemir I, Omer A, Cousin P, Semenova A, Taschner M, Dergai O, Marzetta F, Iseli C, Eliaz Y, Weisz D, Shamim MS, Guex N, Lieberman

Aiden E, Gambetta MC. 2021. CTCF loss has limited effects on global genome architecture in Drosophila despite critical regulatory functions. *Nature Communications 2021 12:1* **12**:1–16. doi:10.1038/s41467-021-21366-2

Khoury A, Achinger-Kawecka J, Bert SA, Smith GC, French HJ, Luu PL, Peters TJ, Du Q, Parry AJ, Valdes-Mora F, Taberlay PC, Stirzaker C, Statham AL, Clark SJ. 2020. Constitutively bound CTCF sites maintain 3D chromatin architecture and long-range epigenetically regulated domains. *Nat Commun* **11**:2041–1723. doi:10.1038/s41467-019-13753-7

Kokavec J, Zikmund T, Savvulidi F, Kulvait V, Edelmann W, Skoultchi AI, Stopka T. 2017. The ISWI ATPase Smarca5 (Snf2h) Is Required for Proliferation and Differentiation of Hematopoietic Stem and Progenitor Cells. *Stem Cells* **35**:1614–1623. doi:10.1002/stem.2604

Kornberg RD. 1977. Structure of chromatin. *Annu Rev Biochem*. doi:10.1146/annurev.bi.46.070177.004435

Krönke J, Udeshi ND, Narla A, Grauman P, Hurst SN, McConkey M, Svinkina T, Heckl D, Comer E, Li X, Ciarlo C, Hartman E, Munshi N, Schenone M, Schreiber SL, Carr SA, Ebert BL. 2014. Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science (1979)* **343**:301–305. doi:10.1126/SCIENCE.1244851/SUPPL_FILE/KRONKE-SM-REV.PDF

Kukimoto I, Elderkin S, Grimaldi M, Oelgeschläger T, Varga-Weisz PD. 2004. The Histone-Fold Protein Complex CHRAC-15/17 Enhances Nucleosome Sliding and Assembly Mediated by ACF. *Mol Cell* **13**:265–277. doi:10.1016/S1097-2765(03)00523-9

Labib K, Tercero JA, Diffley JFX. 2000. Uninterrupted MCM2-7 Function Required for DNA Replication Fork Progression. *Science (1979)* **288**:1643 LP – 1647. doi:10.1126/science.288.5471.1643

Längst G, Becker PB. 2001. Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors. *J Cell Sci* **114**:2561–2568. doi:10.1242/jcs.114.14.2561

Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**:1021–1036. doi:10.1016/j.cell.2013.01.052

Layden HM, Eleuteri NA, Hiebert SW, Stengel KR. 2021. A protocol for rapid degradation of endogenous transcription factors in mammalian cells and identification of direct regulatory targets. *STAR Protoc* **2**:100530. doi:10.1016/J.XPRO.2021.100530

Lee HS, Lee SA, Hur SK, Seo JW, Kwon J. 2014. Stabilization and targeting of INO80 to replication forks by BAP1 during normal DNA synthesis. *Nature Communications 2014 5:1* **5**:1–14. doi:10.1038/ncomms6128

LeRoy G, Loyola A, Lane WS, Reinberg D. 2000. Purification and Characterization of a Human Factor That Assembles and Remodels Chromatin*. *Journal of Biological Chemistry* **275**:14787–14790. doi:https://doi.org/10.1074/jbc.C000093200

LeRoy G, Orphanides G, Lane WS. 1998. Requirement of RSF and FACT for transcription of chromatin templates in vitro. *Science (1979)* **282**:1900–1904. doi:10.1126/science.282.5395.1900

Levendosky RF, Bowman GD. 2019. Asymmetry between the two acidic patches dictates the direction of nucleosome sliding by the ISWI chromatin remodeler. *Elife* **8**:e45472.1-e45472.18. doi:10.7554/eLife.45472

Li B, Carey M, Workman JL. 2007. The Role of Chromatin during Transcription. *Cell* **128**:707–719. doi:10.1016/J.CELL.2007.01.015

Li J, Liu Y, Rhee HS, Ghosh SKB, Bai L, Pugh BF, Gilmour DS. 2013. Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. *Mol Cell* **50**:711–722. doi:10.1016/j.molcel.2013.05.016

Lilley DMJ, Howarth OW, Clark VM. 1976. The existence of random coil N-terminal peptides — 'tails' — in native histone complexes. *FEBS Lett* **62**:7–10. doi:https://doi.org/10.1016/0014-5793(76)80004-X

Lipford JR, Bell SP. 2001. Nucleosomes Positioned by ORC Facilitate the Initiation of DNA Replication. *Mol Cell* **7**:21–30. doi:https://doi.org/10.1016/S1097-2765(01)00151-4

Long H, Zhang L, Lv M, Wen Z, Zhang W, Chen X, Zhang P, Li T, Chang L, Jin C, Wu G, Wang X, Yang F, Pei J, Chen P, Margueron R, Deng H, Zhu M, Li G. 2019. H2A.Z facilitates licensing and activation of early replication origins. *Nature 2019 577:7791* **577**:576–581. doi:10.1038/s41586-019-1877-9

Loyola A, Huang J-Y, LeRoy G, Hu S, Wang Y-H, Donnelly RJ, Lane WS, Lee S-C, Reinberg D. 2003. Functional Analysis of the Subunits of the Chromatin Assembly Factor RSF. *Mol Cell Biol* **23**:6759–6768. doi:10.1128/mcb.23.19.6759-6768.2003

Luan J, Xiang G, Gómez-García PA, Tome JM, Zhang Z, Vermunt MW, Zhang H, Huang A, Keller CA, Giardine BM, Zhang Y, Lan Y, Lis JT, Lakadamyali M, Hardison RC,

Blobel GA. 2021. Distinct properties and functions of CTCF revealed by a rapidly inducible degron system. *Cell Rep* **34**:108783: 1–21. doi:10.1016/j.celrep.2021.108783

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**:251–260. doi:10.1038/38444

Luk E, Ranjan A, FitzGerald PC, Mizuguchi G, Huang Y, Wei D, Wu C. 2010. Stepwise Histone Replacement by SWR1 Requires Dual Activation with Histone H2A.Z and Canonical Nucleosome. *Cell* **143**:725–736. doi:10.1016/J.CELL.2010.10.019

Luo Z, Lin C, Shilatifard A. 2012. The super elongation complex (SEC) family in transcriptional control. *Nature Reviews Molecular Cell Biology 2012 13:9* **13**:543–547. doi:10.1038/nrm3417

Mahat DB, Kwak H, Booth GT, Jonkers IH, Danko CG, Patel RK, Waters CT, Munson K, Core LJ, Lis JT. 2016. Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). *Nat Protoc* **11**:1455–1476. doi:10.1038/nprot.2016.086

Mardis ER. 2007. ChIP-seq: welcome to the new frontier. *Nat Methods* **4**:613–614. doi:10.1038/nmeth0807-613

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, Voigt P, Martin SR, Taylor WR, de Marco V, Pirrotta V, Reinberg D, Gamblin SJ. 2009. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature 2009 461:7265* **461**:762–767. doi:10.1038/nature08398

Mazza D, Abernathy A, Golob N, Morisaki T, McNally JG. 2012. A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Res* **40**:e119–e119. doi:10.1093/NAR/GKS701

Min S, Jo S, Lee HS, Chae S, Lee JS, Ji JH, Cho H. 2014. ATM-dependent chromatin remodeler Rsf-1 facilitates DNA damage checkpoints and homologous recombination repair. *Cell Cycle* **13**:666–677. doi:10.4161/cc.27548

Mizuguchi G, Shen X, Landry J, Wu W-H, Sen S, Wu C. 2004. ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. *Science (1979)* **303**:343–348. doi:https://doi.org/10.1126/science.1090701

Moon KJ, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. 2005. The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. *Mol Cell* **19**:523–534. doi:10.1016/J.MOLCEL.2005.06.027

Morisaki T, Müller WG, Golob N, Mazza D, McNally JG. 2014. Single-molecule analysis of transcription factor binding at transcription sites in live cells. *Nat Commun* **5**:4456. doi:10.1038/ncomms5456

MURRAY K. 1964. THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. *Biochemistry* **3**:10–15. doi:10.1021/bi00889a003

Navas TA, Zhou Z, Elledge SJ. 1995. DNA polymerase ϵ links the DNA replication machinery to the S phase checkpoint. *Cell* **80**:29–39. doi:https://doi.org/10.1016/0092-8674(95)90448-4

Negrini S, Gorgoulis VG, Halazonetis TD. 2010. Genomic instability an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* **11**:220–228. doi:10.1038/nrm2858

Nick McElhinny SA, Gordenin DA, Stith CM, Burgers PMJ, Kunkel TA. 2008. Division of Labor at the Eukaryotic Replication Fork. *Mol Cell* **30**:137–144. doi:https://doi.org/10.1016/j.molcel.2008.02.022

Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**:930–944. doi:10.1016/j.cell.2017.05.004

Okabe I, Bailey LC, Attree O, Srinivasan S, Perkel JM, Laurent BC, Carlson M, Nelson DL, Nussbaum RL. 1992. Cloning of human and bovine homologs of SNF2/SWI2: a global activator of transcription in yeast S. cerevisiae. *Nucleic Acids Res* **20**:4649–4655. doi:10.1093/nar/20.17.4649

Oppikofer M, Bai T, Gan Y, Haley B, Liu P, Sandoval W, Ciferri C, Cochran AG. 2017. Expansion of the ISWI chromatin remodeler family with new active complexes. *EMBO Rep* **18**:1697–1706. doi:10.15252/embr.201744011

Papamichos-Chronakis M, Peterson CL. 2008. The Ino80 chromatin-remodeling enzyme regulates replisome function and stability. *Nat Struct Mol Biol* **15**:338–345. doi:10.1038/nsmb.1413

Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. 2011. Global Regulation of H2A.Z Localization by the INO80 Chromatin-Remodeling Enzyme Is Essential for Genome Integrity. *Cell* **144**:200–213. doi:10.1016/J.CELL.2010.12.021

Perpelescu M, Nozaki N, Obuse C, Yang H, Yoda K. 2009. Active establishment of centromeric CENP-A chromatin by RSF complex. *J Cell Biol* **185**:397–407. doi:10.1083/jcb.200903088

Peterlin BM, Price DH. 2006. Controlling the Elongation Phase of Transcription with P-TEFb. *Mol Cell* **23**:297–305. doi:10.1016/j.molcel.2006.06.014

PHILLIPS DM, SIMSON P. 1962. Identification of some peptides from an arginine-rich histone and their bearing on the structure of deoxyribonucleohistone. *Biochem J* **82**:236–241. doi:10.1042/bj0820236

Poot RA, Bozhenok L, van den Berg DLC, Steffensen S, Ferreira F, Grimaldi M, Gilbert N, Ferreira J, Varga-Weisz PD. 2004. The Williams syndrome transcription factor interacts with PCNA to target chromatic remodelling by ISWI to replication foci. *Nat Cell Biol* **6**:1236–1244. doi:10.1038/ncb1196

Poot RA, Dellaire G, Hülsmann BB, Grimaldi MA, Corona DFV, Becker PB, Bickmore WA, Varga-Weisz PD. 2000. HuCHRAC, a human ISWI chromatin remodelling complex contains hACF1 and two novel histone-fold proteins. *EMBO Journal* **19**:3377–3387. doi:10.1093/emboj/19.13.3377

Pursell ZF, Isoz I, Lundström E-B, Johansson E, Kunkel TA. 2007. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* **317**:127–130. doi:10.1126/science.1144067

Roberts CWM, Orkin SH. 2004. The SWI/SNF complex - Chromatin and cancer. *Nat Rev Cancer* **4**:133–142. doi:10.1038/nrc1273

Sabantsev A, Levendosky RF, Zhuang X, Bowman GD, Deindl S. 2019. Direct observation of coordinated DNA movements on the nucleosome during chromatin remodelling. *Nature Communications 2019 10:1* **10**:1–12. doi:10.1038/s41467-019-09657-1

Sadeghifar F, Böhm S, Vintermist A, Östlund Farrants AK. 2015. The B-WICH chromatin-remodelling complex regulates RNA polymerase III transcription by promoting Max-dependent c-Myc binding. *Nucleic Acids Res* **43**:4477–4490. doi:10.1093/nar/gkv312

Sakamoto KM, Kim KB, Kumagai A, Mercurio F, Crews CM, Deshaies RJ. 2001. Protacs: Chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc Natl Acad Sci U S A* **98**:8554–8559. doi:10.1073/PNAS.141230798/ASSET/96E10B62-DC16-4F76-94AF-64F8B7059B8E/ASSETS/GRAPHIC/PQ1412307006.JPEG

Sampathi S, Acharya P, Zhao Y, Wang J, Stengel KR, Liu Q, Savona MR, Hiebert SW. 2019. The CDK7 inhibitor THZ1 alters RNA polymerase dynamics at the 5' and 3' ends of genes. *Nucleic Acids Res* **47**:3921–3936. doi:10.1093/nar/gkz127

Sánchez-Molina S, Mortusewicz O, Bieber B, Auer S, Eckey M, Leonhardt H, Friedl AA, Becker PB. 2011. Role for hACF1 in the G2/M damage checkpoint. *Nucleic Acids Res* **39**:8445–8456. doi:10.1093/nar/gkr435

Santoro R, Li J, Grummt I. 2002. The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nat Genet* **32**:393–396. doi:10.1038/ng1010

Schneekloth AR, Pucheault M, Tae HS, Crews CM. 2008. Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics. *Bioorg Med Chem Lett* **18**:5904–5908. doi:10.1016/J.BMCL.2008.07.114

Schwartz BE, Ahmad K. 2005. Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev* **19**:804–814. doi:10.1101/gad.1259805

Shi X, Zhai Z, Chen Y, Li J, Nordenskiöld L. 2022. Recent Advances in Investigating Functional Dynamics of Chromatin. *Front Genet* **13**:689. doi:10.3389/FGENE.2022.870640/XML/NLM

Shimada K, Oma Y, Schleker T, Kugou K, Ohta K, Harata M, Gasser SM. 2008. Ino80 Chromatin Remodeling Complex Promotes Recovery of Stalled Replication Forks. *Current Biology* **18**:566–575. doi:10.1016/J.CUB.2008.03.049

Shopland LS, Hirayoshi K, Fernandes M, Lis JT. 1995. HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites. *Genes Dev* **9**:2756–2769. doi:10.1101/gad.9.22.2756

Skene PJ, Henikoff JG, Henikoff S. 2018. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc* **13**:1006–1019. doi:10.1038/nprot.2018.015

Smith S, Stillman B. 1989. Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell* **58**:15–25. doi:https://doi.org/10.1016/0092-8674(89)90398-X

Song Y, Liang Z, Zhang J, Hu G, Wang J, Li Y, Guo R, Dong X, Babarinde IA, Ping W, Sheng Y-L, Li H, Chen Z, Gao M, Chen Y, Shan G, Zhang MQ, Hutchins AP, Fu X-D, Yao H. 2022. CTCF functions as an insulator for somatic genes and a chromatin remodeler for pluripotency genes during reprogramming. *Cell Rep* **39**:110626. doi:10.1016/J.CELREP.2022.110626

Stengel KR, Ellis JD, Spielman CL, Bomber ML, Hiebert SW. 2021. Definition of a small core transcriptional circuit regulated by AML1-ETO. *Mol Cell* **81**:530–545. doi:10.1016/j.molcel.2020.12.005

Stillman B. 2008. DNA polymerases at the replication fork in eukaryotes. *Mol Cell* **30**:259–260. doi:10.1016/j.molcel.2008.04.011

Stopka T, Skoultchi AI. 2003. The ISWI ATPase Snf2h is required for early mouse development. *Proc Natl Acad Sci U S A* **100**:14097–14102. doi:10.1073/pnas.2336105100

Stopka T, Zakova D, Fuchs O, Kubrova O, Blafkova J, Jelinek J, Necas E, Zivny J. 2000. Chromatin remodeling gene SMARCA5 is dysregulated in primitive hematopoietic cells of acute leukemia. *Leukemia* **14**:1247–1252. doi:10.1038/sj.leu.2401807

Strobino M, Wenda JM, Padayachy L, Steiner FA. 2020. Loss of histone H3.3 results in DNA replication defects and altered origin dynamics in C. elegans. *Genome Res* **30**:1740–1751. doi:10.1101/GR.260794.120

Strohner R, Nemeth A, Jansa P, Hofmann-Rohrer U, Santoro R, Längst G, Grummt I. 2001. NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. *EMBO J* **20**:4892–4900. doi:10.1093/emboj/20.17.4892

Suto RK, Clarkson MJ, Tremethick DJ, Luger K. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol* **7**:1121–1124. doi:10.1038/81971

Swift J, Coruzzi GM. 2017. A matter of time — How transient transcription factor interactions create dynamic gene regulatory networks. *Biochim Biophys Acta Gene Regul Mech*. doi:10.1016/j.bbagrm.2016.08.007

Swygert SG, Peterson CL. 2014. Chromatin dynamics: Interplay between remodeling enzymes and histone modifications. *Biochim Biophys Acta Gene Regul Mech*. doi:10.1016/j.bbagrm.2014.02.013

Thoma F, Koller T, Klug A. 1979. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *Journal of Cell Biology* **83**:403–427. doi:10.1083/jcb.83.2.403

Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S, Harrington WF, Pantel S, Krill-Burger JM, Meyers RM, Ali L, Goodale A, Lee Y, Jiang G, Hsiao J, Gerath WFJ, Howell S, Merkel E, Ghandi M, Garraway LA, Root DE, Golub TR, Boehm JS, Hahn WC. 2017. Defining a Cancer Dependency Map. *Cell* **170**:564-576.e16. doi:10.1016/J.CELL.2017.06.010

Tsukiyama T, Palmer J, Landel CC, Shiloach J, Wu C. 1999. Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes Dev* **13**:686–697. doi:10.1101/gad.13.6.686

Varga-Weisz PD, Wilm M, Bonte E, Dumas K, Mann M, Becker PB. 1997. Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388**:598–602. doi:10.1038/41587

Vermeulen M, Mulder KW, Denissov S, Pijnappel WWMP, van Schaik FMA, Varier RA, Baltissen MPA, Stunnenberg HG, Mann M, Timmers HTM. 2007. Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. *Cell* **131**:58–69. doi:10.1016/j.cell.2007.08.016

Vervoort SJ, Welsh SA, Devlin JR, Barbieri E, Knight DA, Offley S, Bjelosevic S, Costacurta M, Todorovski I, Kearney CJ, Sandow JJ, Fan Z, Blyth B, McLeod V,

Vissers JHA, Pavic K, Martin BP, Gregory G, Demosthenous E, Zethoven M, Kong IY, Hawkins ED, Hogg SJ, Kelly MJ, Newbold A, Simpson KJ, Kauko O, Harvey KF, Ohlmeyer M, Westermarck J, Gray N, Gardini A, Johnstone RW. 2021. The PP2A-Integrator-CDK9 axis fine-tunes transcription and can be targeted therapeutically in cancer. *Cell* **184**:3143-3162.e32. doi:10.1016/J.CELL.2021.04.022

Vos SM, Farnung L, Boehning M, Wigge C, Linden A, Urlaub H, Cramer P. 2018a. Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature 2018 560:7720* **560**:607–612. doi:10.1038/S41586-018-0440-4

Vos SM, Farnung L, Urlaub H, Cramer P. 2018b. Structure of paused transcription complex Pol II–DSIF–NELF. *Nature 2018 560:7720* **560**:601–606. doi:10.1038/s41586-018-0442-2

Wang J, Zhao Y, Zhou X, Hiebert SW, Liu Q, Shyr Y. 2018. Nascent RNA sequencing analysis provides insights into enhancer-mediated gene regulation. *BMC Genomics* **19**:633:1–18. doi:10.1186/s12864-018-5016-z

Weintraub AS, Li CH, Zamudio A v, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, Guo YE, Hnisz D, Jaenisch R, Bradner JE, Gray NS, Young RA. 2017. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **171**:1573–1588. doi:10.1016/j.cell.2017.11.008

Wen Z, Zhang L, Ruan H, Li G. 2020. Histone variant H2A.Z regulates nucleosome unwrapping and CTCF binding in mouse ES cells. *Nucleic Acids Res* **48**:5939–5952. doi:10.1093/NAR/GKAA360

Whitehouse I, Stockdale C, Flaus A, Szczelkun MD, Owen-Hughes T. 2003. Evidence for DNA Translocation by the ISWI Chromatin-Remodeling Enzyme. *Mol Cell Biol* **23**:1935–1945. doi:10.1128/mcb.23.6.1935-1945.2003

Wiechens N, Singh V, Gkikopoulos T, Schofield P, Rocha S, Owen-Hughes T. 2016. The Chromatin Remodelling Enzymes SNF2H and SNF2L Position Nucleosomes adjacent to CTCF and Other Transcription Factors. *PLoS Genet* **12**:e1005940:1-25. doi:10.1371/journal.pgen.1005940

Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H. 1999. NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to Repress RNA Polymerase II Elongation. *Cell* **97**:41–51. doi:https://doi.org/10.1016/S0092-8674(00)80713-8

Yang JG, Madrid TS, Sevastopoulos E, Narlikar GJ. 2006. The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. *Nat Struct Mol Biol* **13**:1078–1083. doi:10.1038/nsmb1170

Yeeles JTP, Deegan TD, Janska A, Early A, Diffley JFX. 2015. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**:431–435. doi:10.1038/nature14285

Zhang H, Lam J, Zhang D, Lan Y, Vermunt MW, Keller CA, Giardine B, Hardison RC, Blobel GA. 2021. CTCF and transcription influence chromatin structure re-configuration after mitosis. *Nature Communications 2021 12:1* **12**:1–16. doi:10.1038/s41467-021-25418-5

Zhao Y, Liu Q, Acharya P, Stengel KR, Sheng Q, Zhou X, Kwak H, Fischer MA, Bradner JE, Strickland SA, Mohan SR, Savona MR, Venters BJ, Zhou MM, Lis JT, Hiebert

SW. 2016. High-Resolution Mapping of RNA Polymerases Identifies Mechanisms of Sensitivity and Resistance to BET Inhibitors in t(8;21) AML. *Cell Rep* **16**:2003–2016. doi:10.1016/j.celrep.2016.07.032

Zhou Y, Santoro R, Grummt I. 2002. The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. *EMBO J* **21**:4632–4640. doi:10.1093/emboj/cdf460

Zikmund T, Kokavec J, Turkova T, Savvulidi F, Paszekova H, Vodenkova S, Sedlacek R, Skoultchi AI, Stopka T. 2019. ISWI ATPase Smarca5 Regulates Differentiation of Thymocytes Undergoing β-Selection. *The Journal of Immunology* **202**:3434–3446. doi:10.4049/jimmunol.1801684

Zikmund T, Paszekova H, Kokavec J, Kerbs P, Thakur S, Turkova T, Tauchmanova P, Greif PA, Stopka T. 2020. Loss of ISWI ATPase SMARCA5 (SNF2H) in acute myeloid leukemia cells inhibits proliferation and chromatid cohesion. *Int J Mol Sci* **21**:2073:1–13. doi:10.3390/ijms21062073

Žumer K, Maier KC, Farnung L, Jaeger MG, Rus P, Winter G, Cramer P. 2021. Two distinct mechanisms of RNA polymerase II elongation stimulation in vivo. *Mol Cell* **81**:3096-3109.e8. doi:10.1016/J.MOLCEL.2021.05.028