Characterizing the Genome Targeting Mechanisms of Caspase Activated DNase (CAD) during Cell Differentiation

Dalal Alsowaida

This thesis is submitted to the University of Ottawa in partial fulfillment of the requirements for the Ph.D. degree in Cellular and Molecular Medicine.

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

© Dalal Alsowaida, Ottawa, Canada, 2024

ABSTRACT

In skeletal muscle progenitor cells, caspase-3 activity engages the differentiation program, in part through the targeted activation of caspase activated DNase (CAD). Once activated, CAD nuclease induces transient DNA strand breaks, where CAD targeting of the promoter region of the cell cycle inhibitor p21, leads to p21 induction, which is itself a potent differentiation signal. However, the DNA strand breaks in differentiating muscle cells are numerous, suggesting that CAD may target multiple loci to influence cell fate. Here, we sought to identify the CAD targeted genome (and associated epigenetic alterations), in an effort to understand how this nuclease influences skeletal muscle cell differentiation. Using CUT & Tag (Cleavage Under Targets and Tagmentation), we mapped CAD genome wide association and identified two large cohorts of CAD binding preferences, promoter regions and to exon/intron boundaries. In particular, for the intragenic sites, we identified an enriched CAD binding to known translocation breakpoints in the *Pax7* and *Foxo1* genes. The resulting Pax7-Foxo1 fusion gene is the cause of alveolar rhabdosarcoma (aRMS) and our data show that *Pax-7, Foxo1* gene breaks are detected during muscle cell differentiation, along with evidence of the fusion transcript. *Pax7* expression must decline for the muscle cell differentiation program to proceed, and our data is consistent with a CAD mediated targeting of the *Pax7* gene to disrupt its expression and secure the differentiation program. We also noted robust association of CAD to a large number of genes (in differentiating muscle cells) that are implicated in chromosomal translocations for both leukemia and lymphoma. The CAD genome targeting of breakpoint translocations appears to be a conserved phenomenon, as CAD CUT & Tag in differentiating T cells revealed a large overlap with the CAD targets in muscle cells, suggesting that the origin of many cancer causing breakpoint translocations may develop from CAD activity in differentiating progenitor cell populations. Our observations also demonstrated

that CAD did not retain sequence specificity for genome targets, suggesting that other chromatin and/or epigenetic modifications may control CAD DNA targeting events. We reasoned that such epigenetic modifications may be repressive, as blocking CAD access to the genome would be a priority for the cell, as it would act to limit premature engagement of CAD induced differentiation. Enhancer of zeste homolog 2 (Ezh2) is a key component of the multiprotein polycomb repressive complex 2 (PRC2) that directs K27 lysine methylation to control gene expression, and CUT & Tag experiments revealed a significant overlap in the genome wide binding sites of Ezh2 and CAD in muscle cells. In a separate set of experiments, we also demonstrate that caspase-3 cleavage inactivates Ezh2, suggesting that caspase 3 protease may act to remove the epigenetic constraints that limit CAD genome targeting during muscle cell differentiation. Taken together, our observations suggest that CAD control of muscle cell differentiation is a complex genome reprogramming event, that may give rise to the origin of cancer causing break point translocations.

AUTHORIZATION

Permission to include published articles and figures as part of dissertation/thesis was obtained:

- "From [Benada et al. Self-inflicted DNA breaks in cell differentiation and cancer. Trends Cell Biol. 2023]. Reprinted with permission from Elsevier," to include full article, as a co-author.
- "From [Alsowaida and Megeney. Destruction as a creative process: CAD-induced DNA strand breaks promote macrophage differentiation. Cell Rep. 2024]. Reprinted with permission from Elsevier," to include full article, as a co-author.
- "From [Bell et al. Chromatin Reorganization during Myoblast Differentiation Involves the Caspase-Dependent Removal of SATB2. Cells. 2022]. Reprinted with permission from MDPI," to include full article, as a co-author.
- From [Larsen et al. Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress. Science. 2022]. Reprinted with permission from AAAS," to include figure 4.
- "From [Larsen et al. Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. Proc Natl Acad Sci U S A. 2010]. Reprinted with permission from the National Academy of Sciences of the United States of America," to include figures 1 and 2.
- "From [Ichim and Tait. A fate worse than death: apoptosis as an oncogenic process. Nat Rev Cancer. 2016]. Reprinted with permission from Springer Nature," to include figure 1.

LIST OF CONTENTS

ABSTRACT ii
AUTHORIZATION iv
LIST OF CONTENTS v
LIST OF FIGURES ix
LIST OF TABLES x
LIST OF ABBREVIATIONS xi
ACKNOWLEDGEMENT xiii
THESIS FORMAT xv
CHAPTER 1: INTRODUCTION 1
1. Caspases - Proteolytic Enzymes and Signal Transduction Regulators 2
1.1. Caspases – Key Enzymes for the Induction of Cell Death 2
1.2. Caspases as Regulators of Organism Homeostasis 7
1.3. Cell Autonomous DNA Fragmentation Mediated by Caspase-Activated DNase (CAD) 8
1.4. Caspase-3 as a Cell Fate Modulator 10
1.4.1. Caspase-3 Promotes a Non-Death Function: Attenuated Cell Death 10
1.4.2. Caspases Promote Cellular Differentiation without Enucleation or Organelle Removal 12
1.5. Skeletal Muscle Differentiation 14
1.5.1. Caspase-3 Activity During Skeletal Muscle Differentiation and Muscle Stem Cell Self-Renewal 15
1.5.2. CAD-Dependent Strand Breaks as an Inductive Cue for Differentiation of Skeletal Muscle Progenitor Cells 17
1.5.3. Repair of CAD-Mediated DNA Breaks in Myogenic Differentiation 20
1.6. CAD-Directed Strand Breaks Act as a Distinct Form of Cell Fate Regulation 20
1.7. Caspase/CAD-Mediated DNA Breaks in Cancer Progression 21
1.7.1. CAD-Mediated DNA Breaks in Cancer Therapy 25
1.8. Physiologic DNA Strand Breaks 29
1.8.1. Self-Inflicted DNA Breaks in Neuronal Development 29
1.8.2. Self-Inflicted DNA Breaks in Germline 31

1.8.3. Self-Inflicted DNA Breaks in Adaptive Immunity 31
1.9. Chromatin Landscape 32
1.9.1. Caspase as a Genome Reorganizing Factor 33
2. Rationale and Hypothesis 34
CHAPTER 2: MATERIALS and METHODS 36
1. Mice 37
2. Primary Myoblast Isolations and Culture 37
3. C2C12 Cells 38
4. ICAD Cloning and Site-Directed Mutagenesis 38
5. Plasmid Transfection and Generation of Stable Cell Lines 39
6. Caspase Inhibition Assays 39
7. CUT and Tag Protocol 40
7.1. T Cells Isolation from Thymus for CUT and Tag 44
8. RNA-Sequencing 45
9. RNA Isolation and Quantitative Reverse Transcription, Real Time PCR 45
10. Fluorescent in Situ Hybridization (FISH) 46
11. Immunoprecipitation-Mass Spectrometry 48
12. Immunoblotting 49
13. In Vitro Cleavage Assay 49
14. Immunohistochemistry 50
14.1. Immunohistochemistry of Adenoviruses Expressing CAD/ICAD 51
15. Bioinformatics Analysis for CUT and Tag and RNA-Seq 51
15.1. Data Processing for Next-Generation (high-throughput) Genome Sequencing ... 51
15.2. Data Processing for Differential Binding Analysis 52
15.3. Data Processing for Circa Plots 52
15.4. Data processing for the RNAseq 53
CHAPTER 3: RESULTS 54
Caspase Activated DNase (CAD) Targets Genomic Loci in Differentiating Muscle Cells that are Susceptible to Formation of Breakpoint Translocations 54
Abstract 56
Background 57
Results 60
3.1. Genome Wide Mapping of CAD Strand Breaks Targets in Differentiating Myoblasts 60
3.2. CAD has no Specific DNA Binding Motif in Differentiating Cells 70

3.3. CAD Binds Gene Promoters Correlated with Increased Gene Expression in Muscle Cell Differentiation 71
3.4. CAD Bound Genes that Comprise the Parent Genes for Breakpoint Translocations in Differentiating Muscle Cells 75
3.5. CAD Activity in Differentiating Muscle Cells Bind the Parent Genes that Give Rise to alveolar Rhabdomyosarcoma (aRMS) 83
3.6. Caspase-3/CAD Activity Regulates Pax7 Gene Expression during Skeletal Muscle Differentiation. 89
3.7. CAD Binding Cancer Breakpoint Genes in T Progenitors Cells 92
Summary 104
CHAPTER 4: RESULTS 105
Caspase-3 Targets the Epigenetic Regulatory Protein Ezh2 During Myogenesis 105
Abstract 106
Background 107
Results 111
4.1. Epigenetic Modifier Ezh2 Binds to CAD Targets in the Genome of Muscle Cells 111
4.2. Chromatin Modifier Ezh2 and H3k27me3 Histone Mark Overlap with CAD Targeting Sites 120
4.3. Ezh2 Genomic Targets are Sensitive to Caspase-3 Inhibitor during Differentiation of C2C12 Myoblasts 124
4.4. Detection of EZH2 Fragments during Differentiation of Muscle Cells 128
4.5. Ezh2 Protein is a Caspase-3 Substrate 132
4.6. Ezh1, EED, and SUZ12 are not Caspase-3 Targets 137
4.7. CAD form immunocomplex with Ezh2 137
4.8. Inhibition of caspas-3 intensified Ezh2 nuclear protein 139
Summary 142
CHAPTER 5: RESULTS 143
CAD Protein Interactions in Differentiating Muscle Cells 143
Abstract 144
Background 145
Results 148
5.1. ICAD/CAD Protein Localization in Differentiating Muscle Cells 148
5.2. Identification of Novel CAD Interacting Proteins 150
5.3. CAD Bind DNA Break/Repair Protein Complex 151
5.4. CAD Interacts with Enzymes that Regulate Replication, Transcription and Control of DNA Topology 157

5.5. Genomic Target of CAD Overlap with TopIIIβ 158
5.6. Genomic Target of CAD Overlap with RPA1/2 163
5.7. CAD binds to linker histone H1 variants 172
Summary 174
CHAPTER 6: DISCUSSION 175
6.1. Caspase-3/CAD Drives Divergent Cell Fate Outcomes 176
6.2. CAD Activity may Initiate Cancer Breakpoint Fusion Gene Formation 177
6.3. Cell Differentiation Initiates a CAD Association with Numerous Translocation Prone Cancer Fusion Genes 179
6.4. Radiation and Topoisomerase Poisons Contribute to Chromosomal Translocations .. 180
6.5. Chromosomal Aberrations are Detected in Non-Cancerous Contexts 182
6.5.1. Chromosomal Translocations 182
6.5.2. Chromothripsis 183
6.6. Is CAD Operative in Other Cell Fate Transition/Differentiation Other than Muscle Progenitor Cells? 184
6.6.1. Caspase/CAD Activity in the Initiation of Hematopoietic Differentiation 184
6.6.2. Does Caspase-3 Activate CAD to Drive Differentiation in the Neuronal Lineage? 185
6.7. Does the Epigenetic Landscape Influence CAD Genome Targeting during Cellular Differentiation? 186
6.8. Gene Repression by PRC Complexes Protect DNA Damage Sites 187
6.9. Caspases Proteolytic Activity may Target Chromatin Proteins in Differentiating Muscle Cells 187
6.10. Ezh2 Catalytic Activity Depends on the PRC2 Complex 188
6.11. Topoisomerase IIIβ Interaction with CAD; A Breakpoint Modifying Protein Complex? 189
6.12. DNA Repair Mechanisms Mediate CAD Induced DNA Damage 191
6.13. Concluding Remarks and Future Directions 192
6.14. Technical Consideration for CAD Genome 193
Appendices 196
Refrences 247

LIST OF FIGURES

Chapter 1

Fig. 1.1. Extrinsic and intrinsic apoptotic signaling pathways.

Fig. 1.2. Schematic of the stages of myogenic lineage progression.

Fig. 1.3. Transient DNA strand breaks are detected early during C2C12 myoblast differentiation.

Fig. 1.4. Summary of different modes of CAD activation and respective functional outputs.

Fig. 1.5. Self-inflicted breaks in differentiation and cancer development.

Fig. 1.6. Cancer cells hijack CAD nuclease to trigger a second wave of DNA breaks following DNA-damaging treatment.

Fig.1.7. Model of CAD-dependent G_2 phase checkpoint.

Chapter 2

Fig. 2.1. Workflow of CUT and Tag protocol.

Fig. 2.2. Chromosomal localization and rearrangement of *PAX7*, *FOXO1*, and *PAX7-FOXO1* FISH probes.

Chapter 3

Fig. 3.1. Genomic distribution of CAD binding sites.

Fig. 3.2. RNA-Seq. analysis of CAD promoter genes.

Fig. 3.3. CAD activity in differentiating muscle cells binds the parent genes that give rise to breakpoints in C2C12 cells.

Fig. 3.4. CAD activity in differentiating muscle cells bind to *Pax7* and *Foxo1* genes.

Fig. 3.5. Caspase-3/CAD activity regulates *Pax7* gene expression during skeletal muscle differentiation.

Fig. 3.6. Genomic distribution of CAD binding sites, from T progenitor cells.

Chapter 4

Fig. 4.1. Genomic distribution of CAD and Ezh2 binding sites.

Fig. 4.2. Ezh2 and H3k27me3 overlap with CAD targeting sites.

Fig. 4.3. Genomic distribution of Ezh2 binding sites.

Fig. 4.4. Ezh2 expression during differentiation of C2C12 myoblasts.

Fig. 4.5. Prospective caspase-3 cleavage sites of the Ezh2 protein complex.

Fig. 4.6. CAD forms an immunocomplex with Ezh2 protein.

Fig. 4.7. Inhibition of caspase-3 rescued Ezh2 expression.

Chapter 5

Fig. 5.1. Colocalization of ICAD/CAD in differentiating skeletal muscle C2C12 cells.

Fig. 5.2. CAD-interacting proteins, identified by MS from C2C12 cells.

Fig. 5.3. Genomic distribution of CAD and TopIIIβ binding sites.

Fig. 5.4. Genomic distribution of CAD and RPA1 binding sites.

Fig. 5.5. Genomic distribution of CAD and RPA2 binding sites.

Fig. 5.6. CAD binds to linker histone H1 variants.

LIST OF TABLES

Table 3.1. List of genes that are bound to CAD and implicated in chromosomal translocations.

LIST OF ABBREVIATIONS

AID	Activation-induced cytidine deaminase
AP	A basic sites
Apaf-1	Apoptotic peptide activating factor-1
APE	Apurinic Endonuclease
ATM	Ataxia-telagastia mutated
ATR	ATM and RAD3-related
BCL2	B-Cell CLL/Lymphoma 2
BER	Base Excision Repair
CAD	Caspase Activated Nuclease
cDNA	Complementary Deoxyribonucleic acid
ChIP	Chromatin Immunoprecipitation
BH3	BCL-2 homology domain 3
bHLH	Basic Helix Loop Helix
BSA	Bovine Serum Albumin
CAD	Caspase-activated DNase
CFC	Contextual fear conditioning
ChIP-Seq	ChIP couple to high throughput sequencing
CK II	Casein kinase II
CK2	Casein Kinase 2
CSR	Class Switch Recombination
DAPI	4,6-Diamidino-2-phenylindole
CTCF	CCCTC-binding factor
DFF	DNA Fragmentation Factor
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA-PK	DNA Protein Kinase
DSB	DNA double strand break
EDTA	Eythylenediaminetetraacetic acid
ESCs	Embryonic stem cells
FBS	Fetal Bovine Serum
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
HRR	Homologous Recombination Repair
IAP	Inhibitor of Apoptosis Protein
ICAD	Inhibitor of CAD
ISNT	In situ nick translation
KD	Knock down
KO	Knock out
LM-PCR	Ligation Mediated Polymerase Chain Reaction
MACs	Model Based Analysis of ChIP-seq
DSBs	Double strand breaks
FADD	Fas-associated death domain
HMGB1/2	High mobility group proteins 1/2
HSCs	Hematopoietic stem cells
HSP70	Heat shock protein 70

ICAD	Inhibitor of caspase-activated DNase
ID	Interdigital cell death
ISNT	In situ nick translation
LM-PCR	Ligation-mediated PCR
LMW	Low molecular weight
MHC	Myosin Heavy Chain
MHC	Myosin heavy chain
miR	microRNA
MLL	Mixed lineage leukemia
MRFs	Master Regulatory Factors
MST	Mammalian Sterile Twenty-like kinase
MOMP	Mitochondrial outer membrane permeabilization
NHEJ	Non-homologous End Joining
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
RAG(1-2)	Recombination-activating-gene protein 1-2
RNA	Ribonucleic Acid
RSS	Recombination Specific Sequence
RT-PCR	Reverse-transcriptase Polymerase Chain Reaction
SC	Satellite Cell
SHM	Somatic Hypermutation
ssDNA	DNA single strand break
t-AML	Secondary acute myeloid leukemia
TopoIIβ	Topoisomerase II Beta
NE	Nuclear envelope
NLS	Nuclear localization sequence
NPC	Nuclear core complex
PCD	Programmed cell death or apoptosis
PBTs	Breakpoint
RPA	Replication protein A
RT	Radiotherapy
Ruvbl1,2	RuvB like AAA ATPase 1,2
SSBs	Single strand breaks
TADs	Topologically associated domains
TCR	T-cell antigen receptor
UV	Ultraviolet
V(D)J	Variable(diversity)joining
z-DEVD-fmk	N-benzyloxycarbonyl-Asp-Glu-V al-Asp-fluomethylketone
γH2AX	H2AX histone variant

ACKNOWLEDGEMENT

First and foremost, I would like to thank Dr. Lynn Megeney for his guidance, mentorship, and encouragement throughout these years. Your passion and enthusiasm for this project truly enhanced my experience and intellect as a junior scientist. Every time I speak to you feeling low after a failed experiment, I leave your office with my spirit lifted, and encouraged to do more for science. This journey has fostered growth in me, on many levels, making me stronger and more resilient. Thank you, Lynn.

I would also like to extend my gratitude to my thesis advisory committee members: Dr. Marjorie Brand, Dr. Jeffery Dilworth, and Dr. Nadine Wiper-Bergeron, for their time, advice, and feedback, and most importantly, for being a great source of support. A special thanks to Drs. Brand and Dilworth for their collaboration on different aspects of my thesis.

I would like to acknowledge the former and present members of the Megeney lab: Dr. Mohammad (Mo) Abdul-Ghani, Dr. Ryan Bell, Steve Brunette, and Anupam Sharma, for their friendship and for making the lab environment enjoyable. I also extend my gratitude to Eric Areeza for his help with bioinformatic analysis—your insights were truly appreciated. Additionally, I would like to thank my friend Thomas Olender from the Brand lab for his technical advice on CUT and Tag, and most importantly, for your friendship.

I must express my deep gratitude to my husband, Dr. Bassam Albraidy, for his unconditional love, support, and encouragement. Thank you for listening to me on my rough days and helping me see the light at the end of the tunnel. To my beloved children, Fahad and Lubna, thank you for being

my light. You were the best distraction from my work, and your smiles, love, and laughter fueled me to thrive and do more.

A special thanks to my parents, Saleh and Salma, who set me off on this path. Without you, I wouldn't have been able to do it. Thank you for your continuous love, support, and prayers. To my siblings, thank you for your love and encouragement throughout this experience. To my best friend, Dr. Arwa Almutlaq, thank you for your endless support, I'm grateful for your friendship. Soon, we will celebrate our achievements and embark on a new journey of endless scientific collaboration.

Finally, I would like to thank King Saud University for their generous scholarship program, and Saudi Arabian Cultural Bureau (SACB) for administering this scholarship. This thesis is funded by CIHR grants.

THESIS FORMAT

This thesis is written in a monograph format. *Chapter 1* is a general introduction describing the theoretical background of the work presented herein. *Chapter 2* is a detailed methodological explanation of experimental approaches utilized in this thesis. *Chapter 3* is descriptive of the results, titled **"Caspase Activated DNase (CAD) Targets Genomic Loci in Differentiating Muscle Cells that are Susceptible to Formation of Breakpoint Translocations"**. *Chapter 4* is descriptive of the results, titled **"Caspase-3 Targets the Epigenetic Regulatory Protein Ezh2 During Myogenesis"**. *Chapter 5* is descriptive of the results, titled **"CAD Protein Interactions in Differentiating Muscle Cells"**. *Chapter 6* is a discussion of the overall findings of this Ph.D. thesis.

CHAPTER 1: INTRODUCTION

1. Caspases - Proteolytic Enzymes and Signal Transduction Regulators

Caspase proteases are an integral part of development that elicit divergent responses and functions. Caspase enzymes underlie various cell fate outcomes, while maintaining cellular homeostasis. Caspases are a family of cysteine-aspartic-proteases that function in cell death pathways, such as apoptosis or pyroptosis. Caspase proteases are highly conserved across metazoan evolution, and are fundamental endoproteases for maintaining homeostasis by regulating cell death and inflammation, which influence morphogenesis and organogenesis (Rothlin et al, 2021). In healthy cells, caspase proteins are catalytically inactive and are expressed as zymogens, referred to as procaspases. Following defined signaling events, procaspases undergo dimerization to induce catalytic activation of processed homodimers, which subsequently target cellular substrates through aspartic acid-directed cleavage events (Crawford & Wells, 2011). The only exception to this rule is caspase-9, which is a zymogen with basal activity that is further stimulated by dimerization via a scaffolding protein referred to as apoptotic peptide activating factor 1 (Apaf-1), in turn, oligomerizes and autocatalyzes caspase-9 itself (Hu et al, 2013). Furthermore, orthologs of Apaf-1, for example, CED-4 in *C. elegans* and Dapaf-1 (known as Dark) in *Drosophila* play a role in the activation of the caspase cascade. Additionally, the nematode proteins CED-9, which is anti-apoptotic, and a pro-apoptotic protein EGL-1, (functional homologs of the mammalian BCL-2 and BID/BIM proteins), shape the mitochondrial pore complex, which aid to release cytochrome c (Dick & Megeney, 2013; Bratton & Salvesen,2010).

1.1.Caspases – Key Enzymes for the Induction of Cell Death

Apoptosis or programmed cell death (PCD), is a highly regulated process that involves the systematic dismantling of dying cells, preventing unfavorable immune responses by cleaving

substrates that lead to the structural disassembly of the cell components. This is a key concept in cell development and tissue homeostasis. PCD is characterized by distinct morphological alterations of cells including plasma membrane blebbing, cytoplasmic and nuclear condensation, and degradation (Wyllie,1980). The phenomenon of cell death during normal embryonic development was first described by Glucksman in 1951 (Glucksman, 1951; Geske & Gerschenson, 2001). In 1972, Kerr et al. described the morphological definition of apoptosis. Using electron microscopy, Kerr observed cellular debris, referred to as apoptotic body formation, arising from cellular condensation and budding. Kerr defined apoptosis in two phases: i) nuclear and cytoplasmic condensation and cell dismantling into membrane bound fragments, ii) the apoptotic debris are degraded and phagocytosed (Kerr et al, 1972). In addition, they described the involvement of apoptosis in cell turnover in healthy adult tissues, targeted elimination of cells during normal embryonic development, and in physiological atrophy of various tissues and organs.

Following these early discoveries, scientists focused on understanding the mechanisms of cell death. These efforts expanded the base of knowledge; and ultimately in 2002, the Nobel committee awarded the Nobel Prize for Physiology or Medicine jointly to Sydney Brenner, H. Robert Horvitz and John E.Sulston, for their seminal discoveries on organ development and PCD in the nematode *Caenorhabditis elegans (C. elegans)*. Genetic studies of the PCD in *C. elegans* have identified two key genes, Ced-3 and Ced-4, which are responsible for PCD in *C. elegans*. Loss or mutation of Ced-3 results in elimination of PCD (Ellis & Horvitzet, 1986; Yuan & Horvitz,1990). In fact, there is only one caspase, Ced-3, which functions as initiator and effector caspase in this model (Green, et al 2022; Dick & Megeney, 2013). *C. elegans* was the first pivotal model to link a caspase protease to cell death. Following these studies, research explored more complex organisms for caspase activity apoptotic signaling. One model investigators explored was the fruit fly,

Drosophila melanogaster, in which it is controlled by caspases like proteases. In *Drosophila*, caspases include death-related ced-3/Nedd2-like caspase (Dredd) and death regulator Nedd2-like caspase (Dronc), which act as initiator caspases, along with death caspase-1 (Dcp1) and death-related ICE-like caspase (Drice), which resemble effector caspases (Abrams,1999; Steller 2008). To conclude, *Drosophila* model contributed to our understanding of mechanisms underlying PCD and revealed an evolutionary conserved cell death pathway with *C. elegans*.

Vertebrate caspases are categorized into three groups based on the sequence similarity and the mechanism of action. The first group includes caspases-1,-4, and -5 are known as human inflammatory caspases, which are implicated in inflammation-mediated cell death and necroptosis. The second group is termed the effector caspases, comprising caspase-3,-6, and -7 based on the commonalities of having a short pro-domain of 20-30 residues (Parrish et al, 2013; Shalini et al, 2015). These effector caspases, often referred to as executioner caspases, are capable of triggering apoptosis independently of other caspases once activated. The third group represent the initiator caspases, including caspase -8,-9, and -10. These caspase enzymes are characterized by having a long N-terminal prodomain, approximately 90 amino acids in length. Initiator caspases are autoactivated through dimerization and subsequently activate a cascade of downstream effector caspases, which then target a wide array of cellular substrates (Crawford and Wells, 2011). The initiator caspases can be activated either through the extrinsic or intrinsic cell death pathway, each pathway having distinct characteristics for engaging caspase activation in cells undergoing apoptosis (Fig.1.1). The extrinsic pathway involves the binding of a death receptor ligand, for example FAS, to a cognate death receptor such as TRAIL. This binding is mediated by Fas-associated death domain (FADD), and once bound, activates the initiator caspases-8 or -10. In

turn, active caspase-8 cleaves the executioner caspases-3 and -7, resulting in targeted cleavage of protein substrates and eventual cell death/apoptosis (Parrish et al, 2013; Shalini et al, 2015).

The intrinsic pathway, on the other hand, is initiated by various cell stimuli that leads to activation of BCL-2 homology domain 3 (BH3), which activates pro apoptotic proteins BAX and BAK (Fig.1.1). Following this, BAX and BAK trigger mitochondrial outer membrane permeabilization (MOMP) which facilitates the release of the respiratory chain component cytochrome c from the mitochondria to the cytosol. Next, cytochrome c then binds to Apaf-1 which aids in apoptosome assembly to recruit and stimulate caspase-9. Activated caspase-9 subsequently cleaves and activates downstream effector caspases, such as caspases-3, -6, and -7 (Parrish et al, 2013; Shalini et al, 2015). In a functional caspase, the active site for the enzyme contains a nucleophilic cysteine that cleaves a recognized amino acid motif. When caspases cleave intracellular substrates, they recognize a specific sequence and hydrolyze the peptide bond on the C terminal of a 4 amino acid stretch, P4-P3-P2-P1 mainly on the aspartic acid side in the substrate, and less frequently on the glutamate residue (Yao et al, 2021). Hundreds of substrates have been identified for caspase proteases, with most efforts focused on identifying and characterizing substrates for the executioner caspases, 3 and 7 (Green, 2022). Both caspase-3 and caspase-7 share preference for a substrate with a DXXD or DEVD sequence motif. Despite considerable overlap in their substrate targeting between caspase-3 and -7, each protease also has distinct substrate proteins (Walsh et al, 2008; Bell et al, 2022). To conclude, caspase proteolytic activity cleaves numerous cellular proteins underlying various cellular functions.

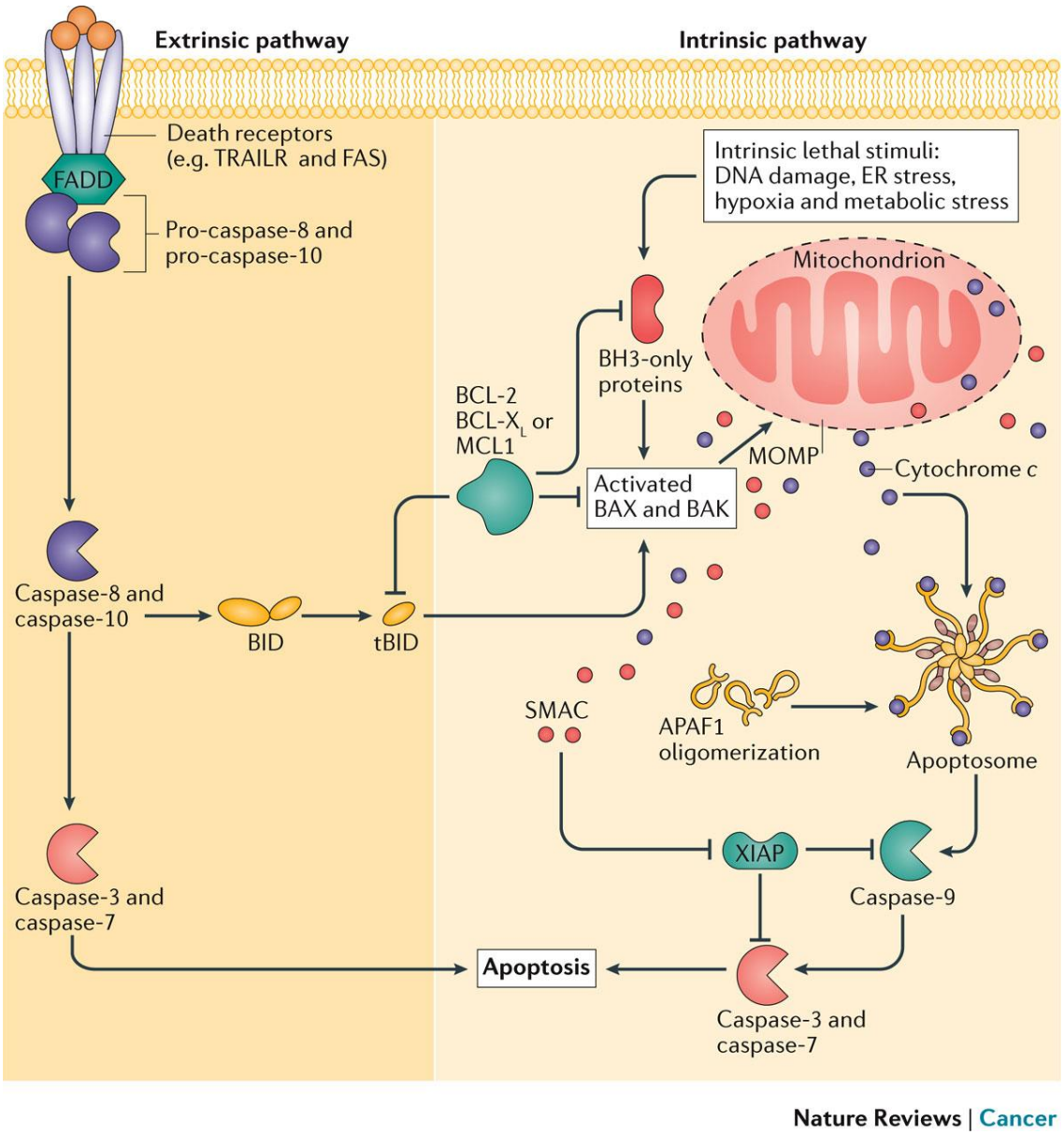


Fig. 1.1. Extrinsic and intrinsic apoptotic signalling pathways.
In the extrinsic apoptotic pathway, upon binding to their cognate ligand, death receptors such as tumour necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) receptor (TRAILR) and FAS can activate initiator caspases (caspase-8 and caspase-10) through dimerization mediated by adaptor proteins such as FAS-associated death domain protein (FADD). Active caspase-8 and caspase-10 then cleave and activate the effector caspase-3 and caspase-7, leading to apoptosis. The intrinsic (or mitochondrial) pathway of apoptosis requires mitochondrial outer membrane permeabilization (MOMP). Cell stresses engage BCL-2 homology domain 3 (BH3)-only protein activation, leading to BAX and BAK activity that triggers MOMP. Anti-apoptotic BCL-2 family proteins counteract this. Following MOMP, mitochondrial intermembrane space proteins such as second mitochondria-derived activator of caspases (SMAC) and cytochrome *c* are released into the cytosol. Cytochrome *c* interacts with apoptotic protease activating factor 1 (APAF1), triggering apoptosome assembly, which activates caspase-9. Active caspase-9, in turn, activates caspase-3 and caspase-7, leading to apoptosis. Mitochondrial release of SMAC facilitates apoptosis by blocking the caspase inhibitor X-linked inhibitor of apoptosis protein (XIAP). Caspase-8 cleavage of the BH3-only protein BH3-interacting death domain agonist (BID) enables crosstalk between the extrinsic and intrinsic apoptotic pathways. ER, endoplasmic reticulum; MCL1, myeloid cell leukaemia 1; tBID, truncated BID. "From [Ichim and Tait. A fate worse than death: apoptosis as an oncogenic process. Nat Rev Cancer. 2016]. Reprinted with permission from Springer Nature."

1.2.Caspases as Regulators of Organism Homeostasis

Caspase protease activation has long been considered a hallmark of apoptosis, yet these enzymes also function in other forms of cell death including pyroptosis and necroptosis. Pyroptosis is a an inflammatory form of cell death that typically occurs in macrophages in response to microbial infections and it is regulated by activation of caspase-1 and caspase-11 (Frank & Vince, 2019). Morphologically, it is characterized by cell swelling and plasma membrane pore formation. This will facilitate the release of proinflammatory cytokines, activating inflammasomes. When apoptosis is compromised, necroptosis is activated. This form of cell death is manifested by an increase in the cell volume, swelling of organelles, and rupture of membrane followed by a leakage of intracellular contents (Frank & Vince, 2019).

Developmental cell death plays a crucial homeostatic function, which aids in tissue organ development by the selective removal of dead or dying cells. A classic example of developmental morphogenesis is interdigital cell death (IDC), which is essential for the separation of digits during embryogenesis. Inter-digital mesenchyme undergoes apoptosis during the seventh and eighth week of gestation (Hernández-Martínez & Covarrubias, 2011; Jordan et al, 2012). Failure of apoptosis in the interdigital mesenchyme causes abnormalities as in syndactyly, which is the fusion of adjacent digits leading to hand deformities (Jordan et al, 2012). Another classic morphogenetic example is the regression or destruction of the male mammary tissue. Apoptosis is triggered in the mesenchyme surrounding the center of the mammary gland, which condenses, and then the epithelial cells are destroyed at 14-15 days of gestation in response to hormones (Dürnberger & Kratochwil, 1980; Kumar et al, 2000). Additionally, apoptosis serves to eliminate cells that are no longer needed to terminate a function. For example, the majority of thymocytes undergo apoptosis through intrinsic pathway during random selection of T-cell antigen receptor (TCR) gene

recombination in thymus (Rothlin et al, 2021). This elimination of autoreactive thymocytes is a critical developmental process for establishing immunological self-tolerance and functional immune system (Daley et al, 2017).

1.3. Cell Autonomous DNA Fragmentation Mediated by Caspase-Activated DNase (CAD)

DNA degradation is a characteristic of apoptosis and is regulated on many levels to maintain cellular homeostasis. In 1980, Andrew Wyllie describied the nuclear cleavage of DNA into a discrete 200 base pair fragments during the terminal stages of apoptosis, which was termed low molecular weight (LMW) fragmentation. Wyllie proposed that this degradation was mediated by an endogenous endonuclease that targets linker DNA between nucleosomes. Subsequent research linked the activation of caspases, particularly caspase-3, to the fragmentation of genomic DNA during apoptosis (Fernandes-Alnemri & Alnemri, 1994).

The central importance of DNA fragmentation in apoptotic cell death led to efforts to identify the key factors that triggered DNA cleavage, culminating in the discovery of CAD. The primary mechanism of CAD originated from an *in vitro* experiment reported by Liu in 1997. Here, incubation of nuclei extracted from HeLa cells with recombinant caspase-3 activated CAD and induced nucleosomal cleavage of DNA (Liu et al, 1997). CAD associates with its inhibitor (ICAD) forming a complex (Enari et al, 1998; Sakahira et al, 1998). In its inactive state, CAD nucleophilic attack is restricted by ICAD heterodimerization in the physiologic complex (ICAD/CAD) (McCarty et al, 1999). Upon apoptotic signaling, caspase-3 cleaves ICAD at two aspartic acid residues (D117 and D224), releasing CAD from its inhibition. This allows CAD monomers to dimerize into a scissor-like structure capable of DNA cleavage (Sakahira et al., 1998). The active

nuclease sites of the CAD dimer are oriented in parallel to accommodate double-stranded DNA, breaking the phosphodiester bonds and generating blunt ends with 3' hydroxyl groups and 5' phosphate residues, thereby inducing strand breaks (Samejima & Earnshaw, 2005).

The heterodimeric protein ICAD/CAD has two known functions in cells undergoing apoptosis: DNA fragmentation and chromatin condensation (Halenbeck et al, 1998; Enari et al, 1998). Further, CAD nuclease preferentially cleaves the linker DNA, generating mono- and oligonucleosomal fragments. ICAD exists in two isoforms generated by alternative splicing: ICAD-Long (ICAD-L) and ICAD-Short (ICAD-S). Of these, only ICAD-L is absolutely required to participate in proper folding of CAD. While the CAD/ICAD complex can bind to DNA, ICAD inhibits CAD's nuclease activity (Samejima & Earnshaw, 2000; Sakahira et al, 1998). Both CAD and ICAD-L possess nuclear localization signals (NLS) at their C termini, which facilitates nuclear accumulation of the heterodimer complex (Samejima & Earnshaw et al 2000). In *Drosophila*, a similar ICAD/CAD complex exists, but it requires cleavage by *Drosophila*-specific caspases to activate the DNase. Notably, *Drosophila* CAD lacks the C-terminal nuclear localization sequence present in its mammalian counterpart, resulting in the absence of DNA fragmentation activity (Yokoyama et al., 2000).

To best understand CAD function, it is essential to consider some of its key biochemical properties: CAD has a molecular weight of 40 KDa, it is a Mg^{2+} dependent endonuclease requiring pH 7.5 of optimal activity, and it is inhibited by Zn^{2+} (Widlak, 2000); The enzyme comprises 345 amino acids and is divided into three domains: i) CIDE-N domain contains 80 amino acids that are highly conserved with those of ICAD, ii) C2 consists of amino acids 86-131, and iii) the catalytic domain (13-345 amino acids), has the catalytic enzymatic activity. The C-terminal region holds the nuclear localization sequence (NLS) (Woo et al, 2004). CAD's nuclease activity is further regulated by

interactions with nuclear proteins such as histone H1, topoisomerase II, high mobility group proteins 1/2 (HMGB1/2), and heat shock protein 70 (HSP70), which bind to and stabilize its function (Kulbay et al., 2021).

Other DNases have been identified to fragment DNA in apoptosis, including DNase I, DNase II, and Endonuclease G (Endo G), yet CAD is one of the best characterized endonucleases (Samejima & Earnshaw, 2005). Endo G is a mitochondrial nuclease that is activated by apoptotic stimuli, translocate to the nucleus and cleave chromosomal DNA independent of caspase (LY et al, 2001). Both Endo G and CAD are ubiquitously expressed, while DNases I & II exhibit more restricted tissue distribution (Keyel, 2017). DNase I require Ca^{2+} and Mg^{2+} divalent cations, functions optimally at neutral pH, and create 5` phosphate following DNA cleavage. DNase II, conversely, operates best at acidic pH, does not require cofactors, and create 3` phosphate DNA end (Keyel, 2017). DNAse I is a secreted DNase and has been implicated in the pathophysiology of systemic lupus erythematosus (SLE) due to its role in chromatin cleavage and impaired chromatin degradation by DNase I is associated with SLE (Valle et al, 2008). DNase II is a lysosomal endonuclease and is also involved in DNA degradation (Bass et al, 2009). Overall, DNA degradation is essential for maintaining cellular health and ensuring proper apoptotic processes.

1.4. Caspase-3 as a Cell Fate Modulator

1.4.1. Caspase-3 Promotes a Non-Death Function: Attenuated Cell Death

A hallmark of a metazoan multicellular organism is cellular specialization and adaption for individualized function, which is achieved through a process known as cell differentiation. The fine-tuned differentiation process is established during development, which creates diversity of cell types to carry out distinct functions in multicellular organisms. Developing organisms have

different cell fates, including differentiation and apoptosis. Both processes, cells undergo cytoskeletal rearrangements, membrane fusion, and chromatin condensation.

Despite the difference in outcomes between apoptosis and differentiation, there is a high degree of morphological and biochemical similarities between these contrasting events (Fernando & Megeney, 2007). To illustrate, some developing, short-lived cell types degrade their nuclei, yet they terminally differentiate. This category of cells includes lens fiber epithelial cells, red blood cells, and skin keratinocytes (Burgon & Megeney, 2018). Interestingly, caspase-3 activity in lens fiber epithelial cells was demonstrated to be directly involved in the enucleation phenotype to induce terminal differentiation of these cells (Ishizaki et al,1998; Gheyas & Menko, 2023). Inhibition of caspase-3 blocked the enucleation process (Ishizaki et al, 1998). Shortly after, two other groups noted similar observations, where caspase-3 activity was noted to be involved in the loss of the nucleus in keratinocyte differentiation, (Weil et al, 1999), and erythroid differentiation (Zermati et al, 2001). This enucleation phenotype was regarded at the time as an attenuated or incomplete form of apoptosis, as the final step was nuclear removal, akin to apoptosis itself. Although morphology changes dramatically in differentiation-associated enucleation phenotypes, i.e. loss of organelles and chromatin condensation, these altered cells do not progress to death. Collectively, these publications led to the hypothesis that caspase-3 protease regulation of differentiation in keratinocytes, lens fiber epithelial cells, and red blood cells was a modified or attenuated form of cell death.

1.4.2. Caspases Promote Cellular Differentiation without Enucleation or Organelle Removal

While the concept of attenuated apoptosis was prevailing in the field over two decades ago, three pivotal studies emerged to refute this concept and replace it with the model for caspase control of cell differentiation. These definitive studies were conducted in cell lineages that did not display enucleation phenotypes upon differentiation. First, was a long-lived cell type, skeletal muscle, which required transient activation of caspase-3 to engage and complete the differentiation program (Fernando et al, 2002; See 1.5.1 for in-depth discussion). The second study, Sordet and colleagues demonstrated the activation of caspase-3 and caspase-9 in human peripheral blood monocytes without evidence of apoptosis. Both caspases induced differentiation of monocytes into mature macrophages in response to colony-stimulating factor (Sordet et al, 2002). The third study focused on the differentiation of *Drosophila melanogaster* spermatids, where caspase activity was found to be essential for spermatogenesis (Arama et al, 2003). Specifically, the effector caspase drICE (analogous to vertebrate caspases 3/7) was required for spermatid individualization by removing bulk cytoplasm, thus generating mature sperm. Inhibition of drICE blocked cytoplasm removal and prevented sperm maturation. Furthermore, the study identified the ubiquitin enzyme dBruce as crucial for protecting the sperm nucleus from hypercondensation and excessive caspase activity (Arama et al, 2003). A recent study from the same group noted that the drICE activation in spermatids is associated with a large protein complex localized at the outer mitochondrial membrane, where excessive protease activity is restrained to direct differentiation rather than cell death (Aram et al, 2016).

Following these seminal studies, a significant body of research emerged, demonstrating the involvement of effector caspase activity in the differentiation of various cell lineages across a wide

range of metazoan organisms. Indeed, caspases were shown to engage in osteogenic differentiation and differentiation programs for bone marrow stromal cells, (Mogi et al, 2003; Miura et al, 2004), and cardiomyocyte progenitors (Abdul-Ghani et al, 2011; Bulatovic et al, 2015). Furthermore, the differentiation of neuronal cell lineages, such as astrocytes, striatal and olfactory neurons, and oligodendrocytes, were all demonstrated to be caspase-dependent processes (Fernando et al, 2005; Kanuka, et al 2005; Ohsawa et al, 2010). These findings demonstrate the broad and essential role of caspases in regulating cell differentiation beyond the association with apoptosis.

Finally, it is interesting to note that caspase-3 may act at multiple junctures to drive a cell differentiation program. In a study reported by Fujita et al, the authors demonstrated increased caspase-3 activity following induction of differentiation of embryonic stem cells (ESCs). ESCs have the capacity to self-renew or commit to a differentiation program, and the authors observed that caspase-3 activity limited embryonic stem cell self-renewal by directly cleaving the pluripotency factor Nanog (Fujita et al, 2008). Conversely, inhibition of casapse-3 led to significant increase in the self-renewing ESC population (Fujita et al, 2008). The capacity of an effector caspase to limit stem cell self-renewal was also noted in hematopoietic stem cells (HSCs), where caspase-3 null mice retained very high numbers of long-term repopulating HSCs that had limited differentiation capacity or mature hematopoietic lineages (Janzen et al, 2008). Taken together, these studies suggest that caspase-3 can act on a continuum, where both commitment to- and the terminal completion of a differentiation program is controlled by activation of this protease (Abdul-Ghani & Megeney, 2008).

1.5. Skeletal Muscle Differentiation

Skeletal muscle represents the largest tissue in the body. Vertebrate skeletal muscle originates from paraxial mesoderm during embryogenesis which arises from progenitor cells in the somites (Buckingham et al, 2003). Skeletal muscle retains remarkable regenerative capacity, mainly because of its resident muscle stem cells known as satellite cells (SCs). SCs are a population of undifferentiated, quiescent myogenic cells characterized by their unique anatomical location between basal lamina and sarcolemma of the muscle fiber. Satellite cells are the primary muscle cell involved in regeneration (Murphy et al, 2011). In response to a stimuli or injury, SCs will exit their normal quiescent state and become activated. This activation subsequently leads to cell proliferation. Cells exit G1 phase of the cell cycle, migrate to a new location, and re-enter the cell cycle to proliferate. Some cells will have rounds of proliferation, express myogenic factors, then differentiate and fuse to form new myofibers to replace the damaged one. Other cells will remain as a reserve population underneath the basal lamina for future rounds of regeneration (Xin et al, 2015).

Muscle development is a multistep process involving the proliferation of myoblasts, their fusion into myotubes, and the subsequent differentiation of these myotubes into fully formed multinucleated myofibers. Skeletal muscle differentiation process is tightly regulated by a complex gene regulatory network, particularly by the expression of four helix loop-helix (bHLH) transcription factors named myogenic regulatory factors (MRFs), including MyoD, Myf5, Myogenin, and MRF4 (Berkes & Tapscott, 2005) (Fig.1.2). Following a muscle injury, resident muscle stem cells are stimulated to re-enter proliferation. At this stage, cells are named myoblasts and express Pax7, Myf5 and MyoD. After proliferation, myoblasts quiet the cell cycle and differentiate into a mature muscle cell. When SC exit the cell cycle, expression of Pax7 and Myf5

are downregulated, but MyoD expression is upregulated. Myogenin and MRF4 are induced in terminal differentiation, while the late stage of differentiation is associated with increased expression of myosin heavy chain (MHC) (Hernández et al, 2017).

1.5.1. Caspase-3 Activity During Skeletal Muscle Differentiation and Muscle Stem Cell Self-Renewal

Skeletal muscle serves as an ideal model system for studying cellular differentiation due to the well-defined hallmarks that can be quantified during the transition from a committed stem cell to a differentiating myoblast, to the final terminally differentiated muscle fiber (Chal & Pourquié, 2017). The Megeney group has demonstrated that the initiation and progression of skeletal muscle cell differentiation is heavily dependent on caspase-3 activity. They observed that caspase-3 activity peaked at early timepoints during low serum-induced differentiation in C2C12 myoblasts, then decreased in the later stages. Moreover, peptide inhibition of caspase-3 resulted in a dramatic reduction in differentiation with concurrent reductions in the expression of muscle fiber-specific proteins (Fernando et al, 2002). Similarly, they noted that caspase-3 null mice displayed impaired myoblast differentiation leading to a dramatic reduction in the formation of multinucleated myofibers (Fernando et al, 2002). Furthermore, Fernando and colleagues reported that caspase-3 cleaved mammalian sterile twenty-like kinase (MST1) to propel differentiation. Indeed, transfection of this cleaved and active MST1 protein partially restored differentiation in caspase-3 null myoblasts (Fernando et al, 2002). This study established the key role of an effector caspase in modulating proteins to induce myogenic differentiation. Subsequently, several groups validated these observations, and established that intrinsic pathway activation of caspase-9 leads to caspase-3 activation, which ultimately promotes muscle cell differentiation (Murray et al, 2008).

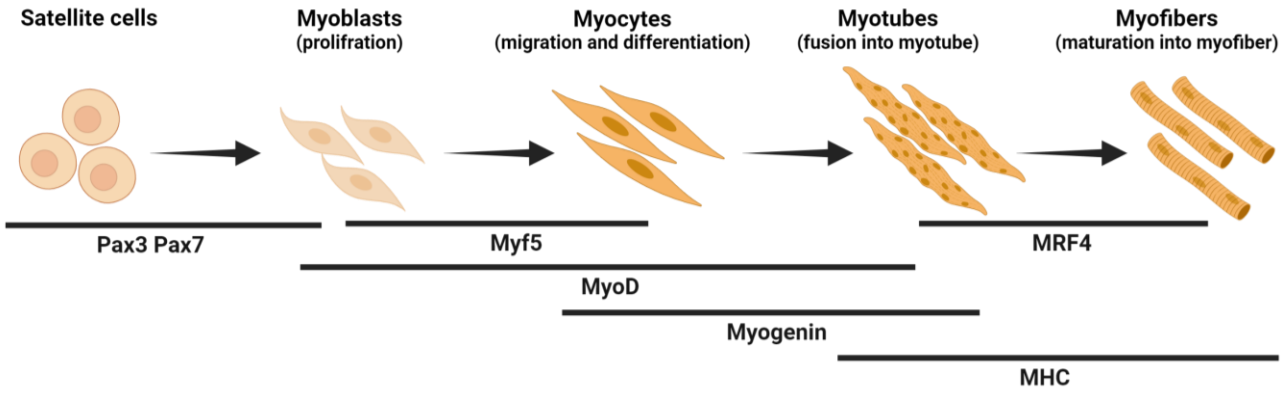


Fig. 1.2. Schematic of the stages of Myogenic lineage progression.
A quiescent satellite cells are activated and undergo several transitions, including activation, proliferation, differentiation, fusion, and maturation. Each stage is distinguished by the expression of various transcriptional factors, as indicated.

Finally, similar to the role uncovered for caspase-3 of modulating ESC self-renewal, the Megeney team uncovered a novel function for caspase-3 in muscle stem cell fate specification. Here, they observed that caspase-3 targets and degrades the Pax7 protein, limiting self-renewal of muscle stem cells (satellite cells) to facilitate the muscle cell differentiation program (Dick et al, 2015). Further, our lab delineated a mechanism for post translational modification of Pax7 (phosphorylation by casein kinase II (CK2) that renders it resistant to caspase cleavage and thereby protects the stem cell phenotype from inadvertent engagement of the differentiation program (Dick et al, 2015). This study provided further evidence that caspase-3 acts to control multiple decision points in a differentiation program by limiting stem cell self-renewal to the late induction of mature cell proteins.

1.5.2. CAD-Dependent Strand Breaks as an Inductive Cue for Differentiation of Skeletal Muscle Progenitor Cells

Given the degree of conservation in caspase signaling events between differentiation and PCD, the Megeney lab then hypothesized that the terminal step of PCD, which is CAD-directed DNA strand breaks/damage, may also be operative in the muscle cell differentiation program (Larsen et al, 2010). Subsequently, the authors observed a transient increase in DNA damage foci early in myoblast differentiation, as detected by *in situ* nick translation (ISNT) and single cell gel electrophoresis (comet assay) of DNA strand breaks. This transient DNA strand break phenomenon was followed by an active DNA repair process as detected by phosphorylation of the H2AX histone variant, γH2AX (Fig.1.3). Larsen et al noted that the damage was induced through the caspase-3 activation of CAD (Larsen et al, 2010), as 1) Use of a caspase-3 peptide inhibitor led to absence of DNA damage concurrent to a block in differentiation; 2) Inhibition of caspase-3 activity or reduction in CAD expression negatively affected the formation of DNA strand breaks

consequently restricting differentiation (Fig.1.3); and 3) Expression of an ICAD mutant (D117) resistant to caspase-3 cleavage, led to impaired muscle cell differentiation. Of interest, the authors noted that caspase-3 targeting of ICAD is distinct from what is observed during apoptosis, where ICAD is cleaved at aspartic acid residues D117 and D224 (Sakahira et al, 1998).

Finally, the authors provided experimental evidence for a mechanism by which CAD-inflicted DNA strand breaks may induce muscle progenitor cell differentiation. Using ligation-mediated PCR (LM-PCR) to map DNA strand breaks, they observed caspase-3/CAD-mediated breaks in the gene promoter of the cyclin-dependent kinase (CDK) inhibitor p21, which coincided with increased expression of p21 (Larsen et al, 2010). Moreover, CAD shRNA clones showed reduced p21(*Cdkn1a* gene) expression. The expression of p21 is an essential step for the differentiation of a wide variety of cell lineages (Kreis et al, 2019), and the results here suggest that breaks in the p21 promoter act as a priming event to induce p21 gene transcription (Larsen et al, 2010). Indeed, the mechanism of promoter-targeted DNA strand breaks have been observed in other cell types (Ju et al, 2006), suggesting that this regulated form of DNA damage may be a conserved mechanism to induce transcription independent of cell pathology or apoptosis (discussed in 1.6).

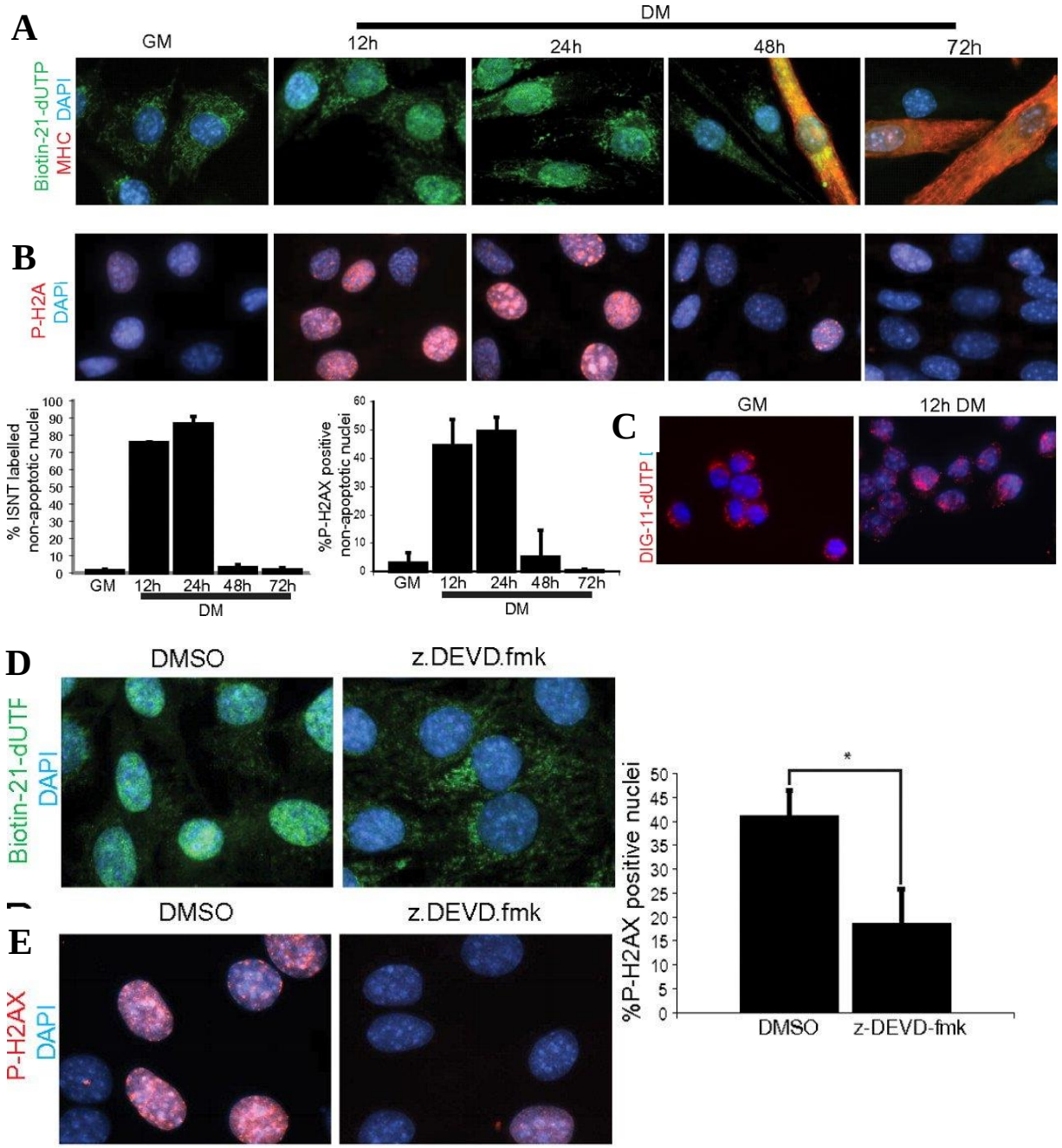


Fig. 1.3. Transient DNA strand breaks are detected early during C2C12 myoblast differentiation. (A) ISNT labels methanol-fixed C2C12 nuclei at 12 and 24 h after being switched from growth media (GM) to low-serum differentiation media (DM). Cells were stained for incorporated biotin-21-dUTP (green), myosin heavy chain (MHC; red), and DAPI (blue). (Error bars are SD.) (B) Phosphorylated H2AX foci are detected in differentiating C2C12 myoblasts at 12 and 24 h after low-serum–induced differentiation. Cells were fixed with 95% ethanol and 5% acetic acid and stained for P-H2AX (red) and DAPI (blue). (Error bars are SD.) (C) ISNT labels fixed primary myoblast nuclei at 12 h when switched to DM. Cells were stained for incorporated DIG-11-dUTP (red) and DAPI (blue). (D) DNA strand breaks are dependent on caspase 3 activity. The caspase 3 inhibitor, z.DEVD.fmk, blocks the formation of DNA strand breaks detectable by ISNT. C2C12 myoblasts were induced to differentiate under low-serum conditions treated with either 10 μM z.DEVD.fmk or DMSO. (E) Formation of P-H2AX foci at 24 h of low-serum differentiation are diminished when myoblasts are treated with z.DEVD.fmk. (Error bars are SD; $P < 0.001$.). "From [Larsen et al. Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. Proc Natl Acad Sci U S A. 2010]. Reprinted with permission from the National Academy of Sciences of the United States of America."

1.5.3. Repair of CAD-Mediated DNA Breaks in Myogenic Differentiation

Myoblast differentiation relies on the precise formation of DNA damage foci and strand breaks induced by CAD. This transient CAD activation suggests a concurrent DNA repair mechanism that enables the differentiation program to advance. The Megeney group demonstrated that during myoblast differentiation, DNA repair is mediated by X-ray repair cross complementing protein 1 (XRCC1), a key scaffold protein involved in base excision repair (BER) (Alkahalf et al, 2016). Deletion of XRCC1 resulted in a perinatal lethal phenotype, with severely impaired myoblast differentiation and unresolved DNA strand breaks. Thus, XRCC1 contributed to the progression of caspase-3 dependent myogenic differentiation. Moreover, XRCC1 was recruited to the CAD-induced strand break in the p21 promoter to increase p21 gene expression, whereas Cdkn1a gene expression was not induced in XRCC1-depleted muscle cells. These results suggest that the repair event was essential to ensure adequate transcription of the CAD-targeted promoter, and thereby complete the differentiation program (Alkahalf et al, 2016).

Furthermore, a study from the Fearnhead group identified that DNA-PK and RPA1/2, which have known importance for DNA repair, are required for the differentiation of myoblasts. Specifically, inhibition of DNA-dependent protein kinase (DNA-PK) blocked differentiation (Connolly & Fearnhead, 2016). CAD dependence on DNA-PK is evidence of DNA damage response (DDR) response to CAD-induced break. At present, this is the extent of understanding regarding repair mechanisms that modify CAD-directed strand breaks during myoblast differentiation.

1.6. CAD-Directed Strand Breaks Act as a Distinct Form of Cell Fate Regulation

As mentioned above, CAD targets the genome by inducing strand breaks to advance the differentiation program, demonstrating that controlled DNA breaks are not lethal and can play a

significant role in normal physiological processes. In particular, managed DNA strand breaks have been shown to induce alterations in gene expression. For example, the employment of strand breaks in the promoter region utilized by Topoisomerase IIβ (TopoIIβ) promoted induction of androgen receptor-responsive gene expression (Ju et al, 2006). Similar to androgen-mediated signaling in prostate cell lines, TopoIIβ-induced breaks in glucocorticoid receptor gene stimulated its expression (Haffner et al, 2010). Interestingly, caspasse-3/CAD appears to promote cancer activity (Liu et al, 2015; Cao et al, 2022; Haimovici et al, 2022) and increase its resistance to different therapeutic approaches (Larsen et al, 2022), as detailed in sections 1.7 and 1.7.1. Accordingly, all these studies suggest that self-inflicted DNA breaks are utilized to promote the disparate cellular events of differentiation and cancer (Fig.1.4) (Benada et al, 2023).

1.7. Caspase/CAD-Mediated DNA Breaks in Cancer Progression

Double strand breaks (DSBs) are a major hallmark of cancer, arising from various factors including ionizing radiation, oxidation, erroneous DNA repair, and other genotoxic stresses (Tubbs et al, 2017). If DSBs are left unrepaired, they promote the formation of chromosomal translocations and mutagenesis, and ultimately have serious consequences for normal cell function. A unique form of genome instability has been identified over the last decade, referred to as Chromothripsis. “Chromo” stands for chromosome, and “thripsis” stands for the shattering process (Stephens et al, 2011). This phenomenon is characterized by reshuffling of hundreds of DNA breaks localized to a single or a few chromosome(s) or subregion(s), through multiple chromosomal rearrangements (Rode et al, 2016). One single catastrophic event happens in 2-3% of all cancers, suggesting a mutual mechanism or source of origin (Stephens et al, 2011). Interestingly, this genome-wide reshuffling process was linked to apoptosis.

Although apoptosis is typically considered an anti-oncogenic process by eliminating damaged cells, evidence is emerging that apoptosis contributes to carcinogenesis under certain circumstances. One key aspect of apoptosis is the MOMP, which leads to the release of mitochondrial proteins, such as, cytochrome c (Shalini et al, 2015). Subsequently, cytochrome c leakage activates the apoptosome, which in turn activates downstream effector caspases. Once extensive MOMP has occurred, cell death soon follows, serving as a form of tumor suppression. However, Ichim et al. observed that a small amount of MOMP results in limited caspase activation that ultimately stimulates CAD-induced DNA damage to a degree insufficient to kill the cell, but instead promote tumorigenesis (Ichim et al, 2015). Thus, the extent of MOMP is a central determinant of whether mitochondrial apoptotic pathways exert tumor suppressor or oncogenic function (Ichim et al, 2015). Likewise, Cao et al have shown that the dysfunction of mitochondrial dynamics triggers a small amount of MOMP, which causes DNA damage in a caspase/CAD-dependent fashion (Fig.1.5) (Cao et al, 2022; Benada et al, 2023).

In this context, Liu et al showed that under cytotoxic stress, sublethal activation of caspase-3 promotes, rather than supresses, genetic instability and oncogenic transformation. Specifically, cells were exposed to ionizing radiation that sustained active caspase-3. In turn, caspase-3 activity coincided with activation of the Endo G DNase, which mediated persistent DNA breaks that caused skin carcinogenesis (Liu et al, 2015). Similarly, sublethal activation of caspase-3 and Endo G were implicated in Myc-provoked oncogenic activity (Cartwright et al, 2017).

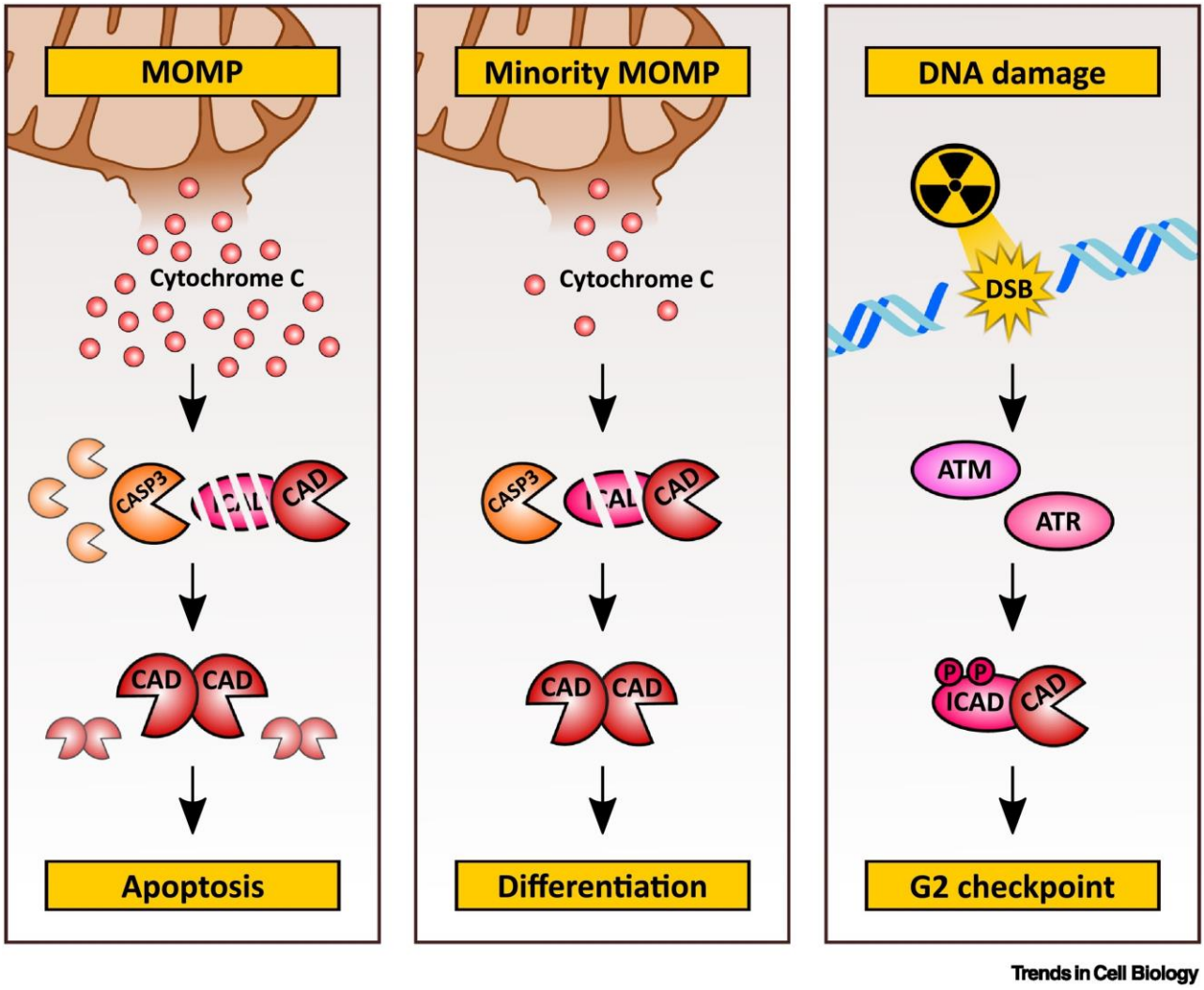


Fig. 1.4. Summary of different modes of CAD activation and respective functional outputs.
During the apoptosis, MOMP leads to the release of cytochrome C from mitochondria and subsequent activation of caspase-3. Caspase 3 then cleaves ICAD at D224 and D117, resulting in CAD dimerization. Next, to the full extent activated CAD homodimers fragment DNA, contributing to the cell death program. The situation of CAD activation in differentiation is reminiscent of the one in the apoptotic program, the main difference is in the extent of the activated processes. MOMP is only limited (termed minority MOMP) and low concentration of cytochrome C leads only to partial activation of Caspase 3. Partially active Caspase 3 cleaves ICAD less efficiently at a single D117 residue and partially active CAD generates a limited amount of DNA breaks which can trigger differentiation programs, but do not result in genome-wide DNA fragmentation. By contrast, activation of CAD-ICAD upon DNA damage is Caspase-independent and mediated by ATM/ATR phosphorylation of ICAD at S107 and S257. Activated CAD-ICAD heterodimer then facilitates G2 checkpoint maintenance via induction of CdSSB and consequent boost of checkpoint signaling. Abbreviations: ATM, ataxia-telangiectasia mutated; ATR, ataxia-telangiectasia and Rad3-related; CAD, caspase-activated DNase; DSB, double-strand break; ICAD, inhibitor of caspase-activated DNase; MOMP, mitochondrial outer membrane permeabilization. “From [Benada et al. Self-inflicted DNA breaks in cell differentiation and cancer. Trends Cell Biol. 2023]. Reprinted with permission from Elsevier.”

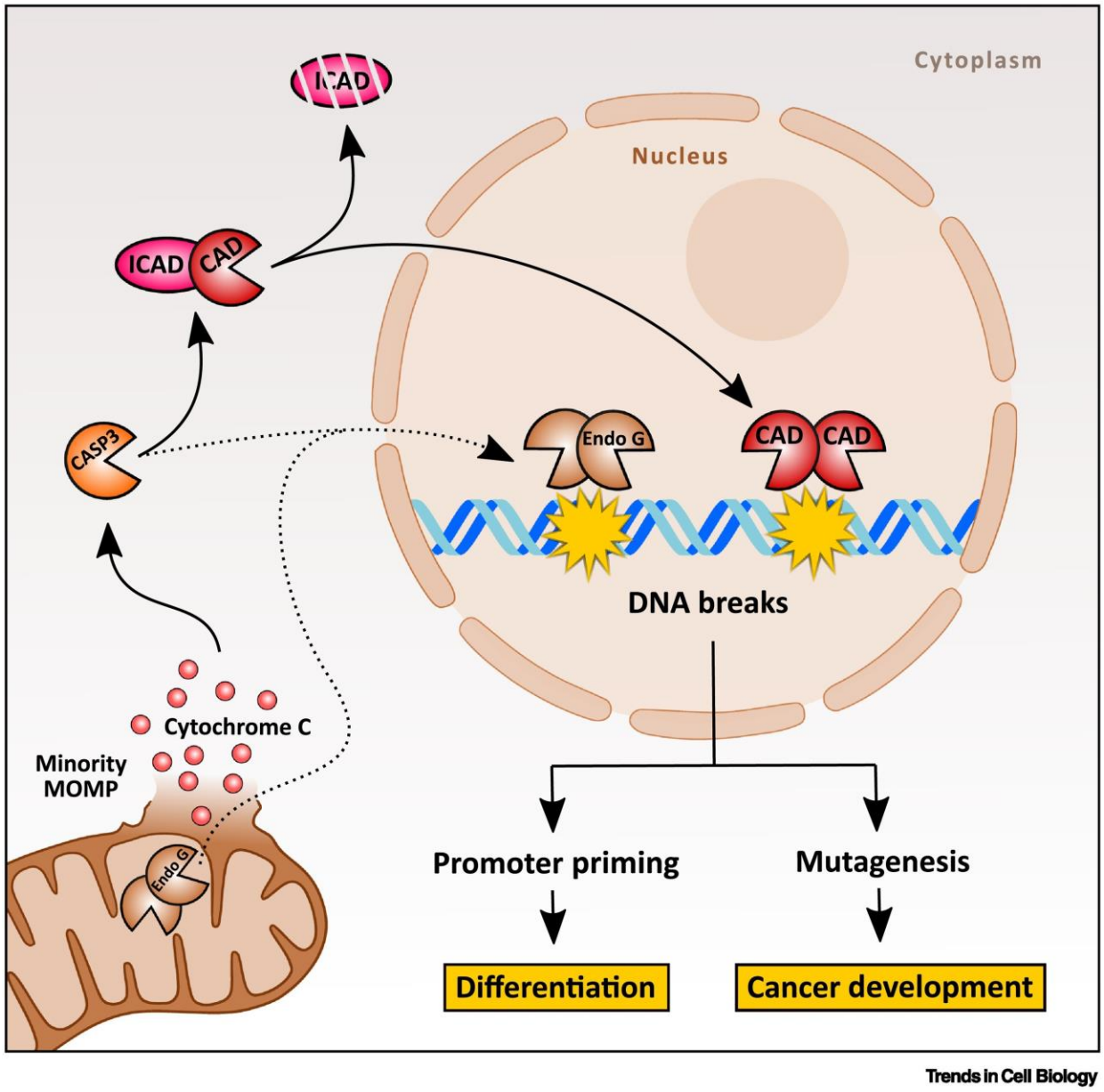


Fig. 1.5. Self-inflicted breaks in differentiation and cancer development.
Minority mitochondrial outer membrane permeabilization (minority MOMP) can lead to sublethal caspase activation. Activated CASP3 (caspase 3) mediates cleavage of inhibitor of caspase-activated DNase (ICAD) that leads to dimerization of caspase-activated DNase (CAD). CAD homodimer then causes DNA breaks. Of note, the caspase-dependent cleavage of ICAD is necessary for CAD dimerization, but both CAD homodimer and CAD-ICAD heterodimer can translocate to the nucleus. CAD-dependent selfinflicted breaks can have a physiological role, for example, transient breaks in promoter regions can prime them to engage gene expression and induce differentiation. By contrast, self-inflicted breaks have mutagenic potential and can drive cancer development. Similarly, upon minority MOMP, homodimeric Endo G nuclease translocates to the nucleus, where it mediates mutagenic breaks and translocations. The relationship of Endo G to caspase-mediated signaling remains unclear (broken lines). Abbreviation: ICAD, inhibitor of caspase-activated DNase. "From [Benada et al. Self-inflicted DNA breaks in cell differentiation and cancer. Trends Cell Biol. 2023]. Reprinted with permission from Elsevier."

Furthermore, partial or incomplete apoptosis promoted melanoma cancer cell invasiveness (Berthenet et al, 2020). Taken together, a growing body of research has emerged over the past decade suggesting the involvement of caspase-3 and their downstream targets CAD or Endo G in carcinogenesis (Fig.1. 5) (Benada et al, 2023).

1.7.1. CAD-Mediated DNA Breaks in Cancer Therapy

Cancer is one of the leading causes of death globally. Approximately 50% of all cancer patients receive radiotherapy (RT) as part of their treatment regimen. However, resistance to radiation is a major obstacle for achieving optimal treatment outcomes, often leading to treatment failure or tumor reoccurrence, and thus placing a large burden on the healthcare system (Baskar et al, 2012; Wu et al, 2023). A study indicated the appearance of temporal DNA damage secondary to the initial breaks inflicted by RT in cancerous cells (Cieślar-Pobuda et al, 2012). RT is a genotoxic cancer therapy using a high radiation dose which damages DNA to an extent that causes either cell death or dramatically slows cancer cell growth (Carlos-Reyes et al, 2021). RT-induced cancer cell death results from wide-ranging DNA damage in the form of single strand breaks (SSBs), DSBs and interstrand cross link formation (Huang & Zhou, 2020).

An important study by Larsen and colleagues investigated how cancer cells evade the RT-induced genotoxic stress. They exposed wild type (WT) human cells and colorectal cancer cells to radiation, then they measured DNA strand breaks by single-cell alkaline gel electrophoresis (Comet assay) (Larsen et al, 2022). The initial DNA damage was comparable between the WT and cancerous cells, indicating an activation of DDR to repair the damage. However, 24 hours following the RT, a secondary accumulation of DNA breaks emerged in cancer cells employing a checkpoint control, thus providing a mechanism for RT resistance. More importantly, they

identified CAD as the DNase responsible for inflicting genome-wide DNA breaks in these cells. They quantified DNA breaks using ISNT, which confirmed that the wave of DNA damage was CAD-dependent (Larsen, et al, 2022). The emergence of SSBs breaks following RT coincided with enrichment at CCCTC-binding factor (CTCF) regions, which are genomic elements that aid in higher order chromatin. Interestingly, this cancer-specific CAD activity was independent of caspase-3, instead regulated by the phosphorylation of ICAD through DNA damage response kinases. Here, ataxia-telangiectasia mutated (ATM)/ataxia-telangiectasia and Rad3-related (ATR)-dependent phosphorylation of ICAD contributed to CAD activation, leading to induction of DNA DSBs to control the cell cycle checkpoints and tumor cell survival (Fig.1.6; Fig.1.7) (Larsen et al, 2022). This study sheds light on the role of CAD/ICAD signaling in resisting genotoxic stress, suggesting that tumor cells induce regulated DNA strand breaks to escape therapy-related DNA damage, which lead to maintenance of tumor growth.

Another example of cancer therapy that inflicts DNA breaks are topoisomerase inhibitors. Topoisomerases are enzymes that have both endonuclease and ligase activity, which function in DNA topology (Pommier et al, 2016). Topoisomerase inhibitors are widely prescribed drugs for cancer chemotherapy, with their primary mode of action being inhibition of the re-ligation of broken DNA ends. Of note, therapeutic use of anticancer drugs that target topoisomerase can have negative consequences. For example, topoisomerase inhibitors induce DNA DSBs that are associated with therapy-related acute myeloid leukemia (t-AML). t-AML occurs secondary to treatment with TOP2 inhibitors across a wide range of primary neoplasias, in which they initiate chromosomal translocations and fusion of genes (Allan & Travis, 2005; Martín-Encinas et al, 2022).

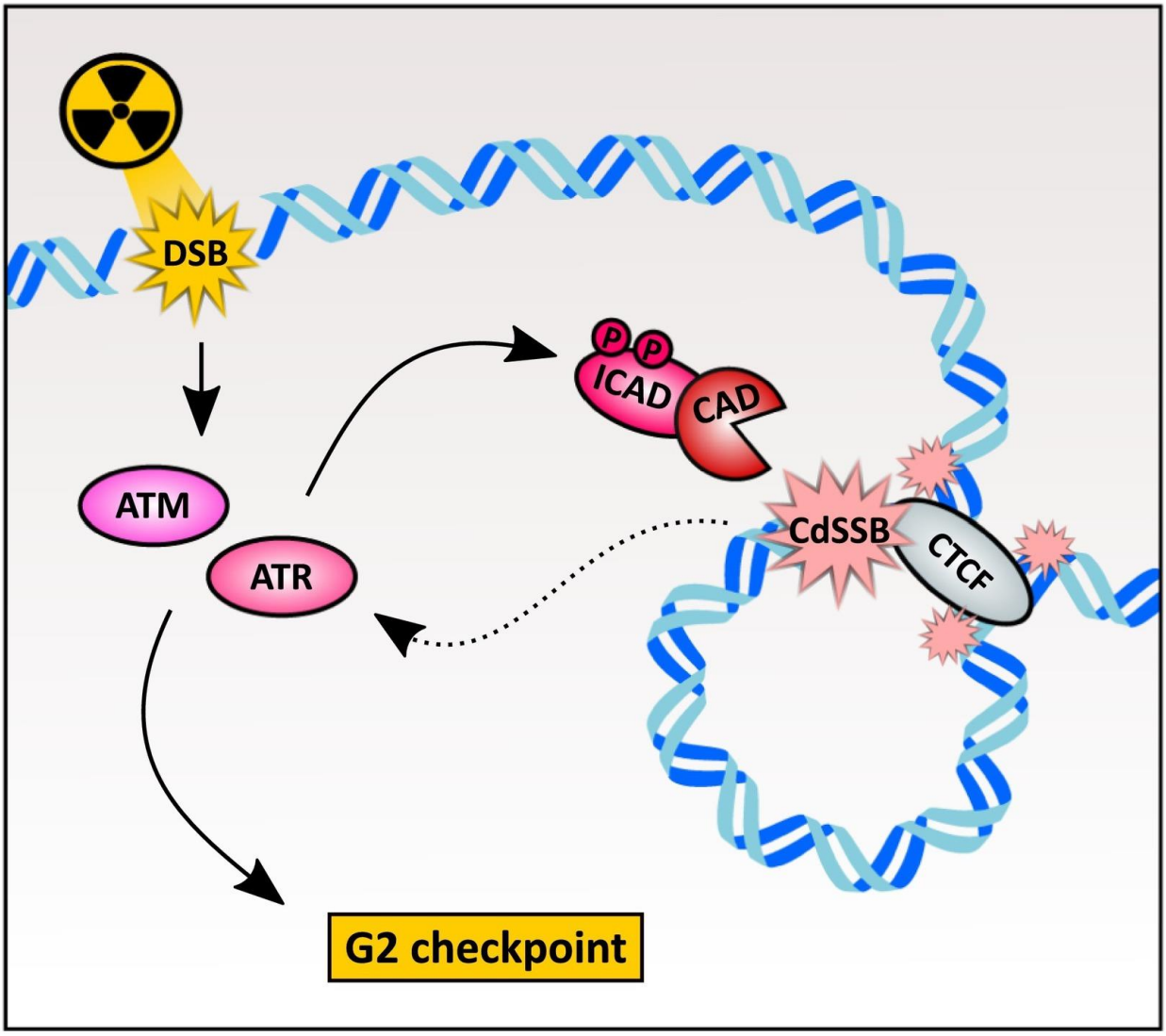


Fig. 1.6. Cancer cells hijack CAD nuclease to trigger a second wave of DNA breaks following DNA-damaging treatment. The initial DNA doublestrand breaks (DSBs) that are inflicted by radiotherapy activate ataxia-telangiectasia mutated and ataxia-telangiectasia and Rad3-related (ATM/ATR) kinases. ATM/ ATR directly phosphorylate inhibitor of caspase-activated DNase (ICAD) and thus prompt CAD-ICAD (caspase-activated DNase) dimer to chromatin. Subsequently, CAD-ICAD dimer inflicts predominantly DNA single-strand breaks [termed CAD-dependent SSBs (CdSSBs)], which are present at CCCTC-binding factor (CTCF) sites. CdSSBs signal in positive feedback loop back to ATM/ATR in order to boost the protective G2 checkpoint and therefore promote therapy resistance. Molecular factors regulating CdSSBs processing and signaling to ATM/ATR remain to be elucidated (broken line). Abbreviation: SSBs, single-strand breaks. "From [Benada et al. Self-inflicted DNA breaks in cell differentiation and cancer. Trends Cell Biol. 2023]. Reprinted with permission from Elsevier."

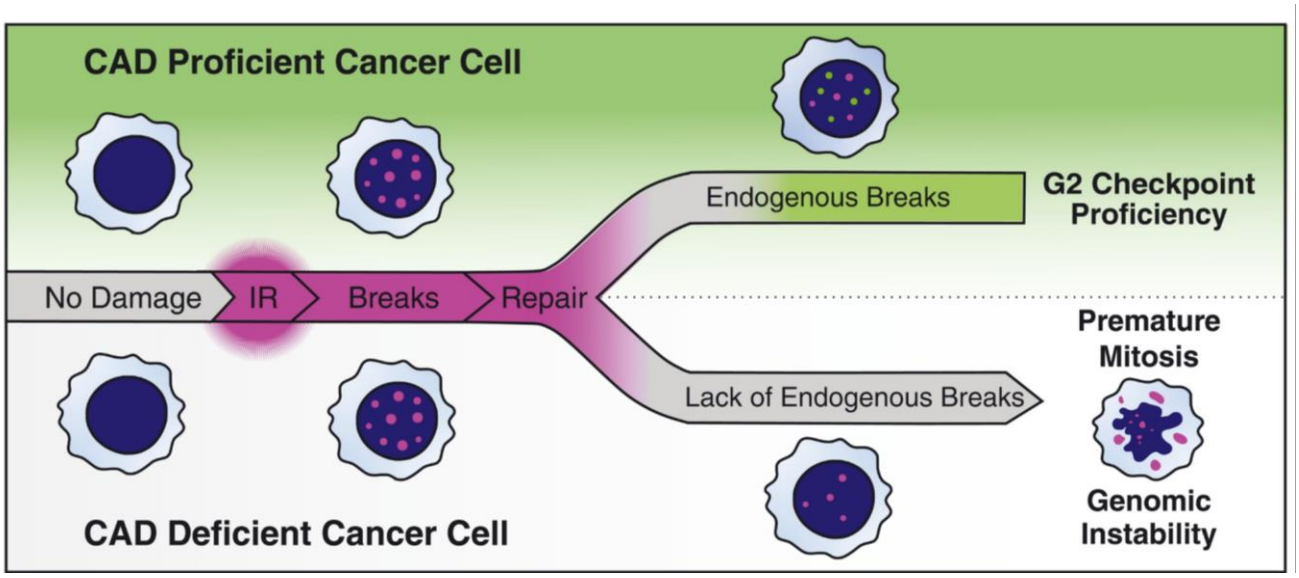


Fig.1.7. Model of CAD-dependent G_2 phase checkpoint.
Tumor cells can activate a nuclease that causes limited induction of DNA breaks at specific sites, which is coordinated with the process of DNA break repair. These self-inflicted DNA breaks trigger the G2 cell cycle checkpoint, preventing tumor cells from cycling and protecting them from death due to treatment-induced DNA damage. "From [Larsen et al. Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress. Science. 2022]. Reprinted with permission from AAAS."

In addition to well known DNA damaging agents, other small molecule cancer therapies appear to inflict DNA DSBs in surviving tumor cells. Namely, treatment with small molecule inhibitors of BRAF, KRAS, or EGFR led to induction of DNA breaks, concurrent with activation of executioner caspases-3 and -7. Additionally, this study confirmed activation of CAD and ATM kinase, which were correlated with enhanced tumor cell growth (Ali et al, 2023).

1.8. Physiologic DNA Strand Breaks

DNA breaks have been shown to control a wide range of normal processes. Physiologic DNA breaks arise from activation of nucleases, and this regulated DNA breakage can have beneficial outcomes for the cell. As noted earlier, this physiologic mechanism is exemplified by CAD induction of DNA damage that propels muscle progenitor cell differentiation. Likewise, DNA damage and subsequent repair has been implicated in normal brain activity, in which regulated strand breaks modify gene expression in response to neural and synaptic plasticity (Konopka & Atkin et al, 2022). In addition to the examples above, meiosis during gametogenesis, and antibody production within the adaptive immune system, both also represent well-studied models in which directed DNA strand breaks control physiologic outcomes rather than cell pathology (Keeney et al 2007; Suberbielle et al, 2013), as detailed in sections 1.8.1, 1.8.2 & 1.8.3.

1.8.1. Self-Inflicted DNA Breaks in Neuronal Development

Neurons are long-lived terminally differentiated cells. Given the longevity of these cells, their genome is subjected to a greater range of DNA damage events, some of which may be derived from characteristic oxidative stress in the brain secondary to its high metabolic demands. However, several laboratories have confirmed a role for programmed DNA breakage in neurons under physiological conditions. In fact, DNA damage and repair have been implicated in neuronal and

synaptic plasticity, which refers to the functional and structural modifications that modulate communication between neurons. DNA strand breaks have been documented to modify gene expression during brain development (Knopka et al, 2022). Likewise, transient DNA strand breaks arise spontaneously in the brain of young WT type mice concurrent with natural exploration of novel environments and encoding of new information. This was further demonstrated by the finding that increased neuronal activity via optogenetic stimulation of striatum results in elevated γH2AX foci across multiple brain regions especially within the dentate gyrus, which is a critical region for learning and memory (Elsa et al, 2013). Moreover, an increase in oxidative stress or activation of glutamate receptors generated DNA strand breaks in cortical neuronal culture (Crowe et al, 2006). DNA strand breaks were inferred by the formation of γH2AX foci that colocalized with the DNA repair protein Mre11 (Crowe et al, 2006). Similarly, another study demonstrated repair of oxidative DNA damage in cerebral cortical neurons following transient activation of glutamate receptor (Yang et al, 2010). Glutamate is an excitatory neurotransmitter that plays an important role in memory, cognition, and motor function (Yang et al, 2010). A recent study showed that mice exposed to contextual fear conditioning (CFC), a form of learning behaviour test, underwent normal physiological neuronal processing which resulted in widespread DNA strand breaks in prefrontal cortex and hippocampus (Stott et al, 2021).

Of interest, a specific role for CAD/ICAD complex was suggested for synaptic protein production in the development of the nervous system. Specifically, ICAD mutant mice manifested enhanced spatial learning and longer memory retention in novel object recognition task (Slane et al, 2002). Although not documented in this study, it is reasonable to conclude that loss of ICAD may have resulted in elevated CAD activity, leading to more widespread DNA strand breaks. Similar mechanisms have been reported in the *Drosophila* as well, where the CIDE-N domain protein

Drep-2 (cell death-inducing DNA fragmentation factor 45 (DFF45)-like effector), an analog of human CAD and ICAD, was identified as a novel synaptic regulatory protein that controls learning and memory. Together, these studies suggest the important role for known DNA strand break proteins and imply that self-inflicted DNA breakage may be a conserved feature for normal brain development. In turn, it has been hypothesized that dysfunction of these processes may lead to cognitive decline and neurodegenerative disease (Konopka & Atkin, 2022).

1.8.2. Self-Inflicted DNA Breaks in Germline

Programmed DNA strand break formation initiates the meiotic recombination exchange between homologous paternal and maternal chromosomes, creating genetic diversity. The parental chromosomes undergo crossover at corresponding segments resulting in a new combination of allele inheritance from haploid to diploid cells. The DNA strand break formation is catalyzed by Spo11 and its accessory proteins Rec114, Mer2, and Mei4. Spo11 is a topoisomerase-related protein that generates a transient DNA cut. This critical developmental process contributes to the passing of traits from parents to offspring and ensures chromosomal segregation during meiotic division (Keeney et al, 2007; Panizza et al, 2011; Yadav et al, 2021).

1.8.3. Self-Inflicted DNA Breaks in Adaptive Immunity

Another classic example of beneficial DNA breakage occurs in B and T lymphocytes in the vertebrate adaptive immune system. This branch of the immune system utilizes DNA breaks to establish genetic diversification in antigen receptors and antibodies to clear pathogenic infections. Here, the RAG1 and RAG2 (recombination activating genes 1 and 2) nucleases introduce DNA DSBs in variable (V), diversity (D), and joining (J) gene segments, as well as in recombination signal sequences (RSSs), in a process known as V(D)J recombination (Liu et al, 2022). The binding

of RAG proteins to RSS results in DNA cleavage, creating broken ends. RAG-mediated DNA breaks are then coupled with activation of ATM kinase-dependent DDR, triggering cell cycle checkpoint of the G1/S phase, followed by engagement of DNA repair machinery to resolve the RAG-induced breaks (Bredemeyer et al, 2006). In addition to the V(D)J system, immune cell-generated DNA breaks are employed for both class switch recombination (CSR) and somatic hypermutation (SHM). Both mechanisms employ an enzyme known as activation-induced cytidine deaminase (AID) to induce DNA breaks that generate isotope switching and increase the antibody repertoire (Feng et al, 2020). These are essential processes for generating long-lasting immune memory in response to past pathogens and are wholly dependent on the generation of regulated DNA strand breaks.

1.9. Chromatin Landscape

The human genome contains approximately 3.2 billion base pairs of DNA packaged into 23 chromosomes inside the nucleus. Genomic DNA is organized three-dimensionally within the nucleus where negatively charged DNA is wrapped around positively charged histone octamer proteins to form the nucleosome (García-Rodríguez et al, 2018). The histone octamer is comprised of two copies each of the following individual proteins: H2A, H2B, H3, and H4. Nucleosomes are the core unit of chromatin and are interconnected by linker DNA segments of 200 bp known as "beads on string". Chromatin is packaged in the nucleus, forming highly condensed chromatin fiber. Physical interactions between chromatin proteins and DNA change dynamically in response to developmental cues and environmental stimuli, leading to different degrees of chromatin compaction (Ho & Crabtree, 2010). The two distinct chromatin states are heterochromatin and euchromatin. Euchromatin is the less condensed form, known as a relaxed structure, which makes sites of DNA damage very accessible to repair factors. Heterochromatin on the other hand, has a

highly compacted structure that is very critical for genome stability (Janssen et al, 2018). Thus, interplay between the two distinct forms of chromatin is important, as cells must maintain a balance between providing repair factors with access to damaged loci, and counteracting additional DNA damage and chromosomal mutations.

1.9.1. Caspase as a Genome Reorganizing Factor

It is historically known that apoptosis involves cellular and molecular changes, one of which is chromatin condensation (Parrish et al, 2013). Indeed, caspase signaling mechanisms alter the chromatin environment prior to DNA cleavage and nuclear bifurcation. This is exemplified in imaging studies that demonstrate the compaction of chromatin in primary cortical neurons occurs prior to caspase-3 activation and induction of apoptosis (Renata et al, 2022). Likewise, caspase-6 has been shown to target the nuclear lamins A and C, which are necessary for chromatin condensation and nuclear disassembly during apoptotic execution (Ruchaud et al, 2002). The apoptotic machinery targets and breaks down the nuclear envelope (NE) through caspase-mediated cleavage of the nuclear core complex (NPC) (Lindenboim et al, 2020). There is compelling evidence that caspase targets a wide range of chromatin-associated proteins within the nucleus. For instance, caspase-6 mediates cleavage of special AT-rich sequence-binding proteins 1 (SATB1) during early T cells apoptosis. Efficient cleavage of SATB1 disrupts its dimerization which leads to its detachment from chromatin, ultimately facilitating nuclear degradation (Galande et al, 2001). More recently, a study from the Megeney group demonstrated that caspase-7 targets and cleaves the DNA binding protein SATB2 to expediate the differentiation program of skeletal muscle cells. Specifically, the removal of SATB2 by caspase-7 influenced chromatin remodeling-associated genes and impacted binding to chromatin loops and their topologically associated domains (TADs) (Bell et al, 2022). These studies imply that effector caspases can act as chromatin

organizing proteins. However, what is less clear is whether this chromatin-centric function impinges upon CAD-directed DNA strand breaks. Nevertheless, the observations do suggest that effector caspases and CAD act as potent genome modifiers outside the traditional cell death response.

2. Rationale and Hypothesis

We have previously demonstrated CAD induced transient DNA strand breaks peaking at 12 to 24 hours post serum induction of differentiation as a prerequisite for prompting skeletal muscle differentiation (Larsen et al 2010). Using LM-PCR to map DNA strand breaks, we observed caspase/CAD dependent strand breaks in the gene promoter of the cyclin-dependent kinase (CDK) inhibitor p21, which coincided with increased expression of p21. The extent of directed DNA breaks suggested that CAD may simultaneously target several genomic loci to manage the differentiation specific gene expression program, which led us to further elucidate CAD activity by genome wide mapping of these strand breaks during myogenesis. Additionally, CAD alters other aspects of cell biology in a manner that can facilitate genome instability and become a source of carcinogenesis. Studies from different groups have shown that the DNase activity of CAD mediated oncogenic strand breaks (Cao et al, 2022; Ichim et al, 2015; Haimovici et al, 2022). Furthermore, in collaboration between Dr. Claus Sørensen's lab and our lab, we recently demonstrated that radiotherapy triggered resistance in cancer cells through CAD inflicted genome wide DNA breaks (Larsen et al, 2022).

Given the critical and diverse roles of CAD, where i) CAD promotes differentiation of muscle cells; ii) CAD is implicated in cancer development; and iii) CAD inflicts widespread DNA strand breaks, **I explored the following hypothesis that CAD targeting of the genome is programmed**

and alters global gene expression to promote myogenesis, and that these programmed strand breaks may have significant overlap with the CAD targeted genome in cancer cells. In short, this thesis will greatly contribute to the understanding of CAD mediated breaks and their role in both cell differentiation and whether there is a shared history with cancer cell mutagenesis. **The specific aims of my thesis; i) Mapping CAD targeted genome, ii) Determining whether epigenetic alterations influence CAD accessibility to its targets, iii) Identifying DNA repair proteins that may enhance/limit CAD activity.**

CHAPTER 2: MATERIALS and METHODS

1. Mice

All mice were housed and treated at the University of Ottawa Animal Care and Veterinary Services in accordance with the Canadian Council on Animal Care (CCAC) guidelines and the University of Ottawa Animal Care Committee protocols.

2. Primary Myoblast Isolations and Culture

Primary myoblasts were isolated from 4–6 week-old mice. For the isolation, skeletal muscle from the hind limbs was extracted. The exposed hind limbs were rinsed in 1X phosphate-buffered saline (PBS) and then cut into small pieces until a liquefied muscle isolate was obtained. The liquefied muscle chunks were then cultured in a dispase/collagenase solution (dispase II, Roche, and 1% collagenase reconstituted in sterile PBS with 2.5 mM $CaCl_2$). To facilitate the enzymatic breakdown of the tissue, multiple rounds of trituration were performed by pipetting up and down. Following this, the mixture was centrifuged at 750 x g for 5 minutes, and the pellet was resuspended in Ham's F-10 Enriched Growth Media (Ham's F-10 medium from Gibco Life Technologies, 20% FBS, 2% Penicillin-Streptomycin (Pen-Strep), and 2.5 μg/mL βFGF). This process yielded a population of primary myoblasts, reducing contamination from other tissues. The cells were grown on collagen-coated plastic tissue culture plates and maintained by replacing the Ham's F-10 media every 48 hours to ensure optimal nutrient availability. Primary myoblasts were passaged at 70% confluency to prevent spontaneous differentiation. To induce differentiation, primary growth media was replaced with primary differentiation media consisting of Dulbecco's Modified Eagle Medium (DMEM) supplemented with 5% horse serum and 2% Pen-Strep. Cells were harvested by scraping off the plates and centrifuged to pellet the cells at various growth and differentiation time points.

3. C2C12 Cells

C2C12 cells were obtained from the American Type Culture Collection (ATCC) and were initially provided in DMEM supplemented with 10% Fetal Bovine Serum (FBS) and 10% Dimethylsulphoxide (DMSO). The cells were thawed in a 37°C water bath and washed with 1X PBS to remove excess DMSO. Following this, the cells were centrifuged at 750 x g, resuspended in growth media (DMEM supplemented with 10% FBS and 2% Pen-Strep), and plated on 10 cm tissue culture plates. The cells were incubated at 37°C with 5% CO_2. Passaging was performed when cells reached 70% confluency, and the media was replaced every 48 hours to maintain cell growth. For differentiation, cells were induced under low serum conditions. The growth media was aspirated, cells were washed with 1X PBS, and differentiation media (DMEM supplemented with 2% Horse Serum and 2% Pen-Strep) was added. Cells were differentiated according to experimental protocols, and the differentiation media was replaced every 48 hours.

4. ICAD Cloning and Site-Directed Mutagenesis

ICAD cloning constructs were previously described by Larsen et al. (2010). In brief, wild-type human ICAD was cloned into the reading frame A of the pcDNA3.1/HisC (Invitrogen) mammalian expression vector between the BamHI and NotI restriction sites. The PCR primers used included a BamHI site in the forward primer (5'-ATGGATCCGAGGTGACCGGGGACGCC-3') a NotI site in the reverse primer (5'-TAGCGGCCGCCTATGTGGGATCCTGTCTG-3'). A two-step megaprimer approach was employed to introduce the D117E mutation. The forward primer listed above and a reverse primer (5'-CTGCCCCGCTCTCTGTTTCATC-3') were used to generate the megaprimer, which was then purified and used with the original reverse primer to produce the full-length ICAD sequence containing the D117E mutation via PCR. The D224E mutation was

introduced in a similar manner, using a reverse primer (5'-CTGATACCCGTCTCTACTGCATC-3') to generate the megaprimer. The resulting PCR products were inserted between the BamHI and NotI restriction sites of pcDNA3.1/HisC as described above. All constructs were verified by sequencing.

5. Plasmid Transfection and Generation of Stable Cell Lines

Plasmid transfections were performed as described by Larsen et al. (2010). Briefly, primary myoblasts were cultured until they reached 50-60% confluency in the absence of antibiotics. Purified plasmid DNA was prepared by diluting 2 μg of DNA and transfected using Lipofectamine 2000 (Invitrogen). The transfection mixture was incubated for 3 hours at 37°C. Following incubation, 1 mL of antibiotic-free growth media was added, and cells were allowed to grow overnight at 37°C. The next day, the growth media was replaced with fresh antibiotic-free growth media, which was changed every 48 hours post-transfection. For antibiotic selection, G418 (Invitrogen) was added to the growth media at a concentration of 400 μg/mL. Transfected cultures were maintained and passaged to generate stable pooled cultures.

6. Caspase Inhibition Assays

For caspase-3 inhibition, cultured myoblasts were pre-treated with either 15 μM Z-DEVD-FMK (DEVD) from BioVision or 15 μM DMSO for 4 hours at 37°C. After pre-treatment, cells were induced to differentiate using low serum media or continued in growth media, both supplemented with 15 μM DEVD or DMSO. The media was replaced every 48 hours throughout the differentiation period.

Caspase inhibition was also achieved by infecting C2C12 cells with adenoviruses expressing p35 (Advs) with a GFP reporter. A GFP-expressing control Advs was used as a control for the p35-expressing Advs. Cells were collected at predetermined time points and processed as described.

7. CUT and Tag Protocol

The Cut and Tag experiment was performed using freshly harvested C2C12 cells (at room temperature) (Fig. 2.1), following the protocol outlined by the Henikoff Laboratory (Kaya-Okur, 2019). Cells were counted and centrifuged for 3 minutes at 600 × g. The pellet was resuspended in 1.5 mL of wash buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 0.5 mM spermidine) and transferred to a 2 mL Eppendorf tube, which was kept on ice until bead preparation.

To facilitate cell binding to concanavalin A-coated beads (conA; BP531, Bangs Laboratories), 10 µL of conA bead slurry was added to each 150,000 cells. The bead slurry was mixed with binding buffer (4 mM HEPES, pH 7.9, 2 mM KCl, 0.2 mM $CaCl_2$, and 0.2 mM $MnCl_2$) by pipetting up and down in a 2 mL tube. The tube was then placed on a magnet stand to clear the solution, and the liquid was completely removed. Finally, 10 µL of binding buffer was added to the beads per sample.

Cells were gently added to the bead slurry and incubated on a rotator for 5-10 minutes to allow binding. After incubation, the tubes were briefly centrifuged at low speed and then placed on a magnet stand to quickly remove the liquid. The cells were resuspended in ice-cold antibody buffer, which consisted of 8 µL of 0.5 M EDTA and 6.7 µL of 30% BSA mixed with 2 mL of 0.05% digitonin (Dig)-wash buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 0.5 mM spermidine). This suspension was divided into 50 µL aliquots in separate tubes.

Primary antibodies were diluted to 1:100 and added to each sample with gentle vortexing. The primary antibodies included CAD- Anti-Dff40 rabbit Ab (AB16926, Millipore), Ezh2 (D2C9) rabbit mAb (5246, Cell Signaling), Anti-Top3β pAb (PA5-101080, Invitrogen), RPA70/RPA1 rabbit Ab (2267, Cell Signaling), RPA32/RPA2 rabbit mAb (ab76420, Abcam), and Histone H3K27me3 rabbit pAb (39155, Active Motif). Samples were incubated overnight at 4°C on a nutator.

The next day, the tubes were placed on a magnet stand to remove the liquid. A secondary guinea pig anti-rabbit antibody (ab46540, Abcam) was then added at a 1:100 dilution in Dig-wash buffer and incubated on a nutator at room temperature for 1 hour. The tubes were briefly centrifuged, and the liquid was removed before washing three times with Dig-wash buffer.

The fusion protein A–Tn5 (pA-Tn5) adapter complex was prepared in Dig-300 buffer (0.01% digitonin, 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM spermidine, and 1 Roche Complete Protease Inhibitor EDTA-Free tablet) to achieve a final concentration of 1:250 per sample (100 μL). After a brief centrifugation, the tubes were placed on a magnet stand to remove the liquid. The pA-Tn5 samples were then gently vortexed to ensure that the solution dislodged most or all of the beads. The tubes were incubated on a nutator at room temperature for 1 hour. After another brief centrifugation, the tubes were placed on the magnet stand to remove the liquid. Subsequently, 1 mL of Dig-300 buffer was added, and the tubes were inverted 10 times or gently vortexed to dislodge the remaining beads. This washing step was performed three times in total. Finally, the tubes were briefly centrifuged, and the liquid was removed on the magnet stand.

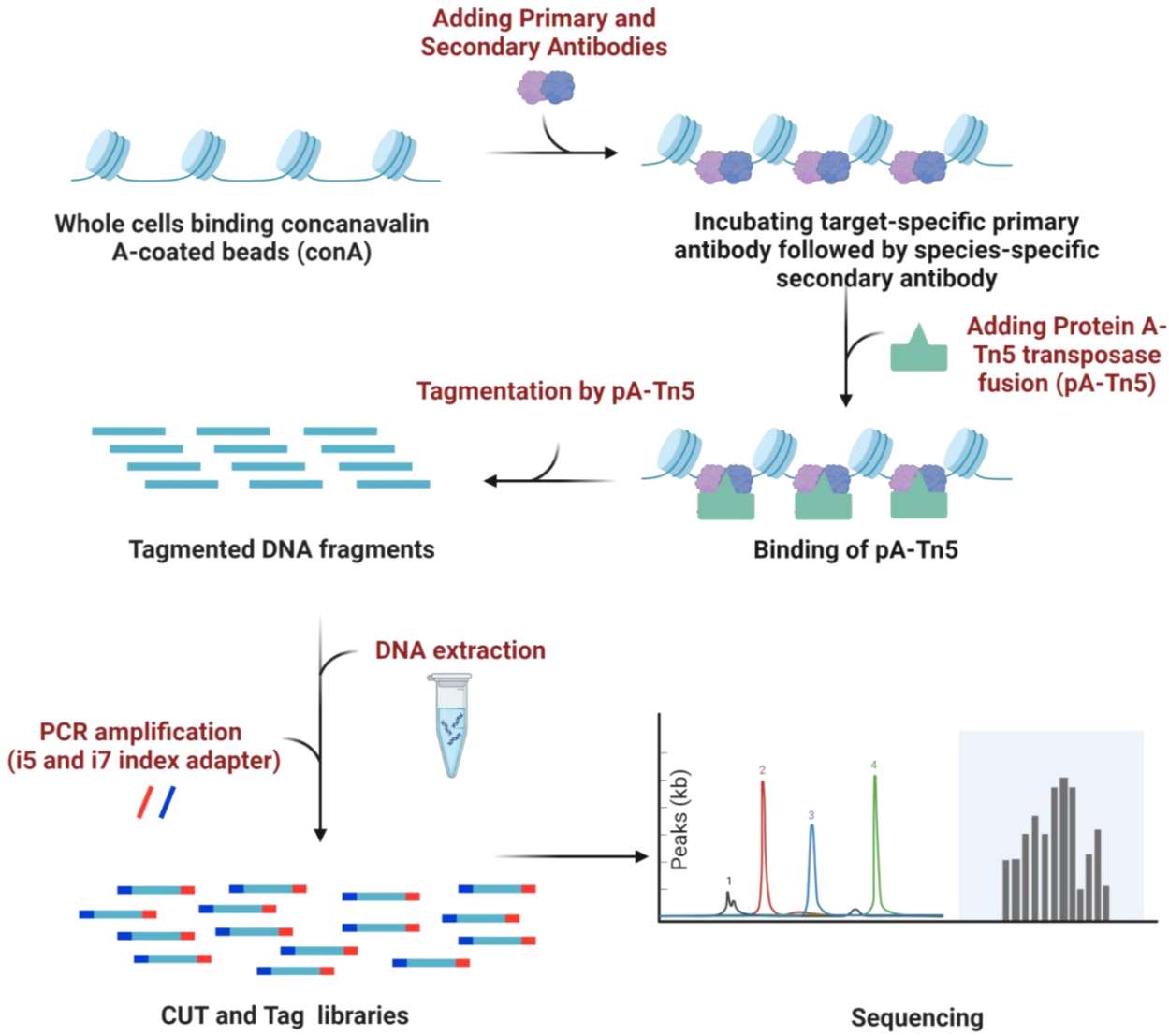


Fig. 2.1. Workflow of CUT and Tag protocol.

Next, 300 µL of tagmentation buffer was added while gently vortexing, and the samples were incubated at 37°C for 1 hour in a water bath or incubator. To halt tagmentation and solubilize the DNA fragments, 10 µL of 0.5 M EDTA, 3 µL of 10% SDS, and 2.5 µL of 20 mg/mL Proteinase K (5 µL of 10 mg/mL) were added to each sample. The samples were mixed by vortexing at full speed for ~2 seconds and incubated at 37°C overnight to reverse cross-links.

The following day, 300 µL of UltraPure™ Phenol: Chloroform: Isoamyl Alcohol (25:24:1, v/v) was added to each sample, mixed by vortexing at full speed, and then transferred to a phase-lock tube. The samples were centrifuged at room temperature at 16,000 × g. Next, 300 µL of chloroform was added to the samples, which were then inverted ~10 times to mix. The tubes were centrifuged again at room temperature at 16,000 × g for 3 minutes. The aqueous layer was carefully pipetted into a fresh 1.5 mL tube containing 750 µL of 100% ethanol and mixed by pipetting up and down. The samples were chilled on ice and centrifuged at 4°C at 16,000 × g for at least 10 minutes. In a laminar flow hood, the liquid was poured off, and the pellet was rinsed with 1 mL of 100% ethanol. The samples were then centrifuged at 4°C at 16,000 × g for 1 minute.

When the tubes were dry, 25-30 µL of elution buffer (1 mM Tris-HCl, pH 8, 0.1 mM EDTA) was added, and the tubes were vortexed at full speed to dissolve the pellet. Each sample was then tested by qPCR to determine the required number of PCR cycles. Following qPCR, 21 µL of DNA from each sample was combined with 2 µL of 10 µM barcoded i5 and i7 primers in a thin-wall 0.5 mL PCR tube. To this, 25 µL of NEBNext HiFi 2x PCR Master Mix was added, and the mixture was placed in a thermocycler with the cycles determined by the qPCR test.

After the PCR run, the tubes were allowed to cool and were then removed from the cycler. 65 µL of SPRI bead slurry was added to each tube and mixed by pipetting up and down. The tubes were

placed on a magnet stand to clear the solution, and the liquid was carefully withdrawn. Next, 200 µL of 80% ethanol was added to the tubes, and they were placed on the magnet stand to clear the solution before the liquid was carefully withdrawn. Subsequently, 25 µL of 10 mM Tris-HCl, pH 8, was added, and the samples were vortexed at full speed. The tubes were placed on the magnet stand to allow the solution to clear, and the liquid was transferred to a new tube. Finally, the samples were frozen at -20°C and sent to the Institut de recherches cliniques de Montréal (IRCM) sequencing facility (Montreal, Canada).

7.1.T Cells Isolation from Thymus for CUT and Tag

T cells were isolated from the thymus of wild-type mice. The thymus was gently crushed in media and filtered through a single-cell filter. To purify the T cells, the EasySep Mouse Kit for positive selection was used, which selectively binds CD4 and CD8 cells to the magnet. The isolated cells were labeled with PE (phycoerythrin)-conjugated antibodies from the mouse. The remaining cells, double negative for CD4 and CD8, were collected and subjected to flow cytometry using the BD LSRFortessa™ Cell Analyzer with the following antibodies: T cell marker Thy1.2 (CD90), CD4, CD8, and a live/dead marker. Flow cytometry (FACS) confirmed that 70% to 80% of T cells were at the double-negative (DN) stage.

The purified cells were then subjected to the CUT and Tag experiment using the CAD antibody mentioned earlier. The markers used for flow cytometry included CD4 (Anti-Mouse CD4, from Stemcell Technologies), CD8 (CD8 Monoclonal Antibody [53-6.7], from Invitrogen), and CD90.2 (a marker for thymus, BUV395 Rat Anti-Mouse from BD Biosciences).

8. RNA-Sequencing

Total RNA was isolated from primary myoblasts of of 8-12 weeks mice. For RNA isolation, the NucleoSpin RNA kit from Macherey-Nagel (cat. 740990.50) was used. Libraries were prepared and sent to the IRCM sequencing facility.

9. RNA Isolation and Quantitative Reverse Transcription, Real Time PCR

RNA was isolated from cultured primary myoblast cells using the RNeasy Mini Kit (Qiagen). A total of 1 μg RNA was then used to synthesize cDNA through reverse transcription with the iScript cDNA synthesis Kit (BioRad). Quantitative gene expression analysis of Pax7 and the Pax7-Foxo1 fusion was performed using quantitative RT-PCR (qPCR). The cDNA was subjected to qPCR using SYBR® Green Supermix (BioRad) with 5 μM forward and reverse primers in an Eco™ Real-Time PCR system (Illumina). qPCR primers were synthesized as follow: *Pax7*-forward 5'-GTT TAT CCA GCC GAC TCT GG-3'; *Pax7*-reverse 5'-ATC GGC ACA GAA TCT TGG AG-3'; *Foxo1*-forward 5'-TAT TGA GCG CTT GGA CTG TG-3'; *Foxo1*-reverse 5'- TGG ACT GCT CCT CAG TTC CT- 3'; *GAPDH*-forward 5'-TTG GGT TGT ACA TCC AAG CA-3'; and *GAPDH*-reverse 5'-CAA GAA ACA GGG GAG CTG AG-3'). For the amplification of the *Pax7-Foxo1* fusion transcript, *Pax7* forward primer and *Foxo1* reverse primer were used for the PCR. Gene expression analysis was performed using a two-step PCR cycle. The run began with a 5-minute hot start activation at 95°C, followed by 40 cycles of 10-second denaturation at 95°C and a 30-second combined annealing/extension step at 60°C. The critical threshold (CT) was set to encompass the exponential amplification curve of each sample. The delta CT (ΔCT) was calculated by subtracting the CT of the gene of interest from that of the GAPDH control, and values were

normalized as indicated in the corresponding legends. Statistical significance was determined using one-way ANOVA or two-way ANOVA, as indicated in the legend.

10. Fluorescent in Situ Hybridization (FISH)

Pax7 break-apart FISH probe kit (CT-PAC088, Cytotest), Vysis Foxo1 break-apart FISH probe kit (03N60-020, Abbott), and Pax7-Foxo1 fusion/translocation FISH probe kit (CT-PAC089, Cytotest) were used according to the manufacturer's protocols (Fig. 2.2). Briefly, primary myoblast cells were seeded on collagen-coated coverslips. Cells were fixed at both growth and 24 hours post-differentiation. The coverslips were then transferred to glass slides and immersed in denaturation solution (70% Formamide/2X SSC at 73 ± 1°C) for 5 minutes. The slides were dehydrated sequentially in 70% ethanol, 85% ethanol, and 100% ethanol for 1 minute each.

A probe mixture was prepared by combining 7 μL of hybridization buffer, 1 μL of probe, and 2 μL of purified water, which was vortexed, centrifuged, and incubated in a 73 ± 1°C water bath for 5 minutes. This mixture was applied to each coverslip to hybridize with the targeted DNA and incubated overnight (16 hours) at 37°C in a humidified chamber. After incubation, slides were washed with wash solution (2X SSC/0.1% NP-40) and air-dried in darkness before staining with DAPI (10 μL). For storage, slides were kept at 4°C and protected from light. For visualization, image z-stack acquisition was performed using a fluorescence microscope equipped with suitable filters. DAPI, SpectrumOrange/CytoOrange, and SpectrumGreen/CytoGreen fluorophores were used to visualize FISH signals. Images were then deconvolved using Zen Desk software.

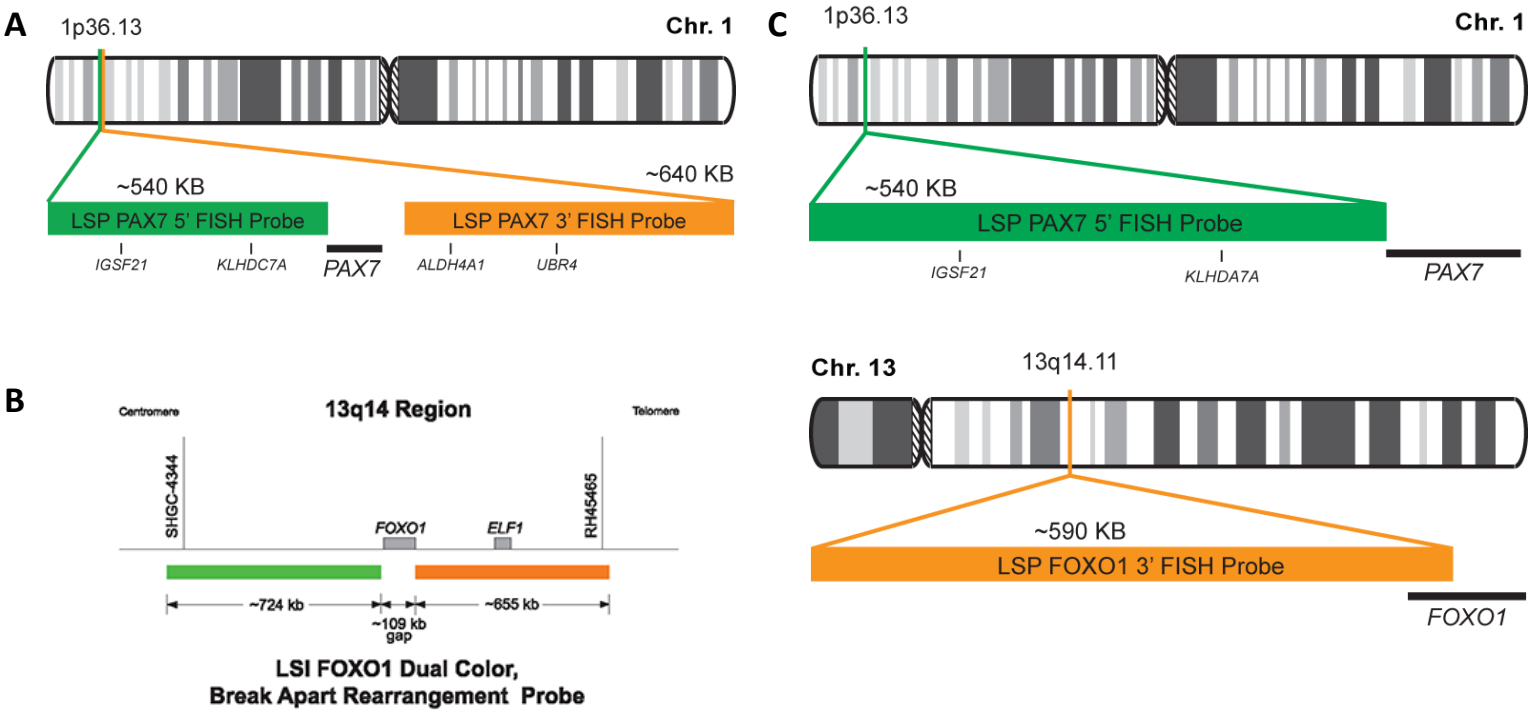


Fig. 2.2. Chromosomal localization and rearrangement of *PAX7, FOXO1,* and *PAX7-FOXO1* FISH probes. (A) PAX7 Break Apart FISH Probe Kit (CT-PAC088, Cytotest). The PAX7 break apart FISH probe is designed to detect rearrangements involving portions of the human *PAX7* gene located on chromosome band 1p36.13. (B) FOXO1 break apart FISH Probe Kit (Vysis 03N60-020, Abbott). The FOXO1 FISH probe is intended to detect chromosome rearrangements involving the *FOXO1* (FKH1, FKHR) gene located on chromosome 13q14. (C) PAX7-FOXO1 Fusion/Translocation FISH Probe Kit (CT-PAC089, Cytotest). The PAX7-FOXO1 fusion FISH probe set is designed to detect rearrangements between the human *PAX7* and *FOXO1* genes, located on chromosome band 1p36.13 and 13q14.11, respectively.

11. Immunoprecipitation-Mass Spectrometry

C2C12 cells in 100-mm dishes were washed twice with PBS and solubilized in cell lysis buffer (Cell Signaling) with protease inhibitors (~1 ml). Solubilized lysates were collected in microcentrifuge tubes and rotated end-over-end for 1 h at 4°C, and then centrifuged at 13,000 g for 15 min at 4°C to pellet insoluble material and cell debris. Protein concentration of lysates was measured using Bradford assay and BSA as a protein standard. 500 µg solubilized lysates were incubated overnight with (1:100) CAD- Anti-Dff40 rabbit Ab (AB16926, Millipore) or normal rabbit IgG at 4°C by end-to end rotation. At the end of incubation, 50 µL Dynabeads Protein A/G (Invitrogen) were added for additional two hours with rotation at 4C. Beads were pelleted, and supernatant were aspirated. Immunocomplex beads were washed with 1X PBS (total three washes), and samples were eluted with 2X laemmli loading buffer and stored at -20°C until used. Aliquots of 25 µg proteins used as input samples were collected.

Eluted immunoprecipitated (IP) samples and input samples were loaded onto 10% (v/v) SDS-PAGE gels followed by silver staining. The gel was fixed, and protein bands corresponding to the cleavage fragments were excised and processed for LC-MS/MS using trypsin, Arg C, and GluC digests. Mass spectral analysis was performed at the OHRI Proteomics Core Facility (Ottawa, Canada). Results were analyzed using Scaffold software. The dataset was queried for Ezh2 peptides produced from cleavage following aspartic acid (D) residues, and Scaffold was used to infer peptide and protein identities bound to CAD from the mass spectra.

12. Immunoblotting

Eluted samples were loaded onto SDS-PAGE gels and then transferred to PVDF membranes. The membranes were blocked with 5% (w/v) non-fat powdered milk in TBS-T (10 mM Tris, pH 7.4; 150 mM NaCl; 0.05% Tween-20) and incubated overnight at 4°C with primary antibodies: rabbit anti-Ezh2 (D2C9) mAb (5246, Cell Signaling); rabbit anti-cleaved-caspase3 (Asp175) mAb (9664, Cell Signaling); rabbit anti-SUZ12 (D39F6) mAb (3737, Cell Signaling); rabbit anti-Ezh1 (D7D5D) mAb (42088, Cell Signaling); rabbit anti-HIS tag (D3I10) mAb (12698, Cell Signaling); rabbit anti-α-tubulin (11H10) mAb (2125, Cell Signaling); mouse anti-β-tubulin (clone E7, Developmental Studies Hybridoma Bank); and mouse anti-MF20 (Myosin Heavy Chain, Developmental Studies Hybridoma Bank).

After incubation with primary antibodies, membranes were washed and incubated with HRP-conjugated secondary antibodies (goat anti-mouse or goat anti-rabbit, Bio-Rad). Protein expression was detected using an ECL detection kit (GE Healthcare). Relative protein levels were quantified using ImageJ software, with values normalized to α-tubulin as a loading control and expressed relative to growth conditions.

13. In Vitro Cleavage Assay

To assess the cleavage of Ezh2 protein by caspase-3, recombinant human active caspase-3 protein (ab52314, Abcam) was incubated with either recombinant human Ezh2/EED/Suz12 protein complex (31337, Active Motif) or recombinant human Ezh2 protein (132934, Abcam) in cleavage assay buffer (50 mM HEPES, pH 7.5; 0.5 M EDTA; 0.1 M NaCl; 10% glycerol; 5% CHAPS; 10 mM DTT). Reactions were incubated for 4 hours at 37°C, in the presence of either DMSO or the

chemical inhibitor of caspase-3, z.DEVD.fmk (15 μM). To stop the reactions, running buffer was added and the samples were subjected to SDS-PAGE. Western blot analysis was performed using anti-Ezh2 antibodies: rabbit anti-Ezh2 (D2C9) mAb (5246, Cell Signaling) and mouse anti-Ezh2 (AC22) mAb (3147, Cell Signaling).

14. Immunohistochemistry

C2C12 cells were cultured on collagen-coated 25 mm coverslips in 6-well Petri plates. At designated time points, the culture media was removed by aspiration and cells were washed three times with 1X PBS. Cells were fixed with 4% paraformaldehyde (PFA) in PBS for 10 minutes, then quickly washed twice with 1X PBS. Fixed cells were incubated with a blocking solution of 5% bovine serum albumin (BSA) in 1X PBS for 1 hour, followed by overnight incubation with the primary antibody at 4°C.

The next day, cells were washed and incubated with a fluorescently labeled species-specific secondary antibody for the corresponding primary antibody, as indicated in the figure legends. The secondary antibody was applied for 1 hour in the dark. Cells were then washed and counterstained with DAPI (1:10,000 in 1X PBS) for 10 minutes to stain the nuclei.

The labeled coverslips were mounted onto glass microscope slides using a drop of Dako Fluorescent Mounting Medium (Dako). Cells were visualized using a Zeiss Observer Z1 Inverted Microscope with appropriate filter settings.

14.1. Immunohistochemistry of Adenoviruses Expressing CAD/ICAD

Adenoviruses expressing CAD/ICAD were constructed for gene delivery, with a FLAG tag added to the N-terminal of WT-CAD and an HA tag added to the N-terminal of ICAD. After infecting C2C12 cells, cells were differentiated for 12- and 24-hours post-serum induction. Immunohistochemistry using anti-HA and anti-FLAG antibodies confirmed the successful overexpression of the tagged proteins. Co-expression of ICAD with CAD was found to temporally stabilize CAD and enhance its DNA binding capacity.

15. Bioinformatics Analysis for CUT and Tag and RNA-Seq

15.1. Data Processing for Next-Generation (high-throughput) Genome Sequencing

First, the quality of raw FASTQ files and subsequent processed files were assessed using FastQC with reports compiled using MultiQC. Illumina adapter sequences were trimmed from the raw reads using Cutadapt. The trimmed reads were then mapped to the reference genome with Bowtie2 to generate the alignment (.bam) files. The alignments were further processed with SAMtools and Picard, which sorted them by coordinate, marked duplicates, and filtered out low-quality alignments (below a default Q value of 10). The final filtered alignment file, containing only mapped reads excluding duplicates, was indexed using SAMtools. Coverage was computed with deepTools bamCoverage at thehighest resolution (bin size of 1), and normalized by Reads Per Genenome Count (RPGC) and by BPM (counts/bins per million mapped reads), resulting in bigwig and bedgraph files. Peak calling to identify binding sites was performed using MACS, SEACR, and GoPeaks. Alignments from two replicates for each condition were merged to have a representative condition sample prior to generating bigwigs and peak calling. Heatmap profiles

were created using deepTools computeMatrix with the "scale-regions" and "skipZeros" options, followed by visualization with plotHeatmap and plotProfile using the "perGroup" option.

15.2. Data Processing for Differential Binding Analysis

DiffBind was used to conduct pairwise differential and shared binding analysis from the alignment and peak files. The analysis began with an "occupancy analysis", where consensus peak sets were generated for each condition in the pairwise comparison. Peaks overlapping in at least 2 out of 3 peak callers were included in a caller consensus set for each replicate. Peaks that overlapped in at least 2 replicates from these consensus sets were retained as the final high-confidence consensus peak set for each condition. Differential comparisons of consensus peak sets between conditions were annotated with ChIPseeker using the UCSC database of known genes. Additionally, KEGG and GO analyses were performed using clusterProfiler.

15.3. Data Processing for Circa Plots

Data for circa plots were generated using the circlize R package, focusing on CAD genes within the COSMIC fusion gene set, referred to as "CAD-Fusion genes." COSMIC breakpoints were mapped to the mm10 genome assembly using the liftOver tool in R. All CAD peak locations within the CAD-Fusion gene regions were identified. Breakpoint locations from COSMIC within these regions were also obtained. Overlaps between CAD peaks and breakpoints in the CAD-Fusion gene regions were identified using the findOverlaps function from the ChIPseeker package. Fusion gene pairs where both genes were within the CAD-Fusion gene set were plotted in the circa plot. Upset plots were generated using the UpSetR package's upset function. To obtain values, CAD-

DEVD-24 genes' annotated peaks and COSMIC fusion genes were mapped to mm10 orthologs using bioMart, and intersections and unions of CAD and COSMIC genes were plotted.

15.4. Data processing for the RNAseq

RNA-seq data and statistics were obtained from DESeq2, performing pairwise comparisons of each differentiation timepoint against the growth timepoint as a reference. DESeq2 results provided log2 fold changes and p-values. Genes with an adjusted p-value < 0.05 across all timepoint comparisons were selected for expression trend analysis. An expression trend was determined based on whether the log2 fold change of subsequent timepoints increased or decreased continuously for each gene. Additional details regarding replicate numbers and conditions are described in the figure legend.

CHAPTER 3: RESULTS

Caspase Activated DNase (CAD) Targets Genomic Loci in Differentiating Muscle Cells that are Susceptible to Formation of Breakpoint Translocations

Caspase Activated DNase (CAD) Targets Genomic Loci in Differentiating Muscle Cells that are Susceptible to Formation of Breakpoint Translocations

Contribution of Collaborators

The experimental performance and preparation of this thesis chapter were conducted by the thesis author with the following exception. Dr. Brain Larsen (from the laboratory of Dr. Lynn Megeney) tested the expression of mutant ICAD in muscle cell differentiation and assayed Pax7-Foxo1 fusion transcript formation (as depicted in Fig. 3.4C, Fig. 3.5A-D). Steven Tur and Dr. Carmen Palii (from the laboratory of Dr. Marjorie Brand) isolated and purified T cells from thymus. Sarah Hachmer (from the laboratory of Dr. Jeffery Dilworth) processed primary muscle cells from a differentiation time course and extracted RNA for sequencing. Eric Areeza provided guidance with the bioinformatics analysis for the CUT and Tag sequenced genome libraries and RNA sequencing.

The preprint of this manuscript has been now published on bioRxiv. You can access it here for further details: PMID: 39386486, PMCID: PMC11463586, DOI: 10.1101/2024.09.24.614809

Abstract

CAD is an endogenous endonuclease that introduces DNA strand breaks during the differentiation of skeletal muscle cells, yet much of the CAD targeted genome and relevant biology remains undetermined. For example, how are CAD induced strand breaks utilized to regulate gene expression in differentiating muscle progenitor cells without causing deleterious effects? To elucidate CAD-mediated biology, I performed genome wide mapping of CAD bound targets during myogenesis/muscle cell differentiation using Cleavage Under Targets and Tagementation (CUT &Tag) methodology. CAD binding was enriched in promoters and within gene bodies, indicating a potential role in modifying gene expression. Further mapping of the CAD bound targets confirmed association to the *Foxo1* gene at intron 1 and the *Pax7* gene at intron 4, suggesting CAD activity may cause breaks in *Pax7* and *Foxo1* gene bodies, which would be disruptive to local transcription. Notably, reduction of *Pax7* expression is required for the induction of muscle cell differentiation, and we noted that impaired CAD activity, through overexpression of a caspase resistant ICAD (Inhibitor of CAD), allowed Pax7 expression to persist, blocking muscle cell differentiation. In addition to CAD control of a developmentally critical cell fate decision, comparative mapping of the CAD targeted *Pax7* and *Foxo1* loci revealed these regions to be the contributing loci that give rise to alveolar rhabdomyosarcoma (aRMS). Therefore, the regulation of the pax7 gene through CAD directed strand breaks raised an interesting possibility that CAD activity was the *de novo* source of the oncogenic Pax7-Foxo1 fusion gene. In support of this hypothesis, we observed that FISH based probes detected *Pax7* and *Foxo1* gene breaks and Pax-7-Foxo1 chimeric transcripts, exclusively in differentiating cells. We also identified CAD bound genes in differentiating muscle cells that align with cancer breakpoint translocations, known to cause leukemias and lymphomas. Notably, we identified a

cohort of these same genes as CAD targets in differentiating T-cell progenitors. Taken together, our data suggest that CAD targets translocation prone genes as a conserved biologic response to control the pace of cell differentiation. This unique regulatory mechanism may serve as the template for oncogene DNA strand break/gene fusions in other cell lineages and their nascent forms of cancer.

Background

Skeletal myogenesis serves as a paradigm for studying the cellular and molecular processes that control cell differentiation, wherein a complex array of signaling molecules and transcription factors work in synergy to instruct and develop mature muscle fibers (Braun & Gautel, 2011). The genetic control of muscle cell differentiation is governed by the MyoD, Myf5, myogenin, and MRF4 genes collectively which are referred to as myogenic regulatory factors (MRFs) (Rudnicki and Jaenisch 1995). However, the genetic initiation of myogenesis control is dominated by paired-homeobox transcription factors Pax3 and Pax7, which are key regulators in skeletal muscle development. *Pax3* is expressed at the onset of myogenesis and contributes to embryonic skeletal muscle development (Relaix et al, 2004). When *Pax3* expression is downregulated, *Pax7* compensates, which is essential for the specification of muscle satellite cells and governs muscle regeneration during postnatal muscle growth (Seale et al, 2000).

In addition to the transcriptional controls noted above, a seminal discovery from the Megeney laboratory demonstrating that skeletal muscle cell differentiation is dependent on caspase-3 signaling (Fernando et al, 2002). Subsequently, caspase-3 induction of differentiation was described across numerous cell lineages and species, suggesting that this might be an evolutionarily conserved mechanism to promote cell maturation in metazoan organisms

(Reviewed in Bell & Megeney, 2017, Burgon and Megeney, 2018). In 2010, follow-up studies from the Megeney group determined that muscle cell differentiation was due, in part, to the caspase activation of the endonuclease CAD, which induced DNA strand breaks in myonuclei to promote the differentiation process (Larsen et al, 2010). In muscle cells, CAD bound to the promoter of the cyclin dependent kinase inhibitor p21, inflicting strand breaks that preceded induction of p21 gene expression (Larsen et al, 2010). This observation suggested a viable target for CAD to induce non-death cell fate outcomes, as elevated expression of p21 is known to be essential for initiating cell differentiation in general.

CAD strand breaks in the p21 promoter are pro-myogenic, yet there is a large number of DNA damage foci in differentiating muscle cells, as detected by ISNT and comet assay (Larsen et al, 2010; Al-Khalaf et al, 2016). CAD induced breaks are likely to influence the course of muscle differentiation, and as such, identifying these loci is a priority. Interestingly, *Pax7* gene has been noted to be a target for strand break formation in other biologic settings, such as cancer initiation. For example, the Pax7 genetic loci form a unique fusion gene with a forkhead (FKHR) gene family member, leading to the production of a chimeric protein that is causative for a subtype of rhabdomyosarcoma (RMS), called alveolar rhabdomyosarcoma (aRMS) (Nguyen & Barr, 2018). RMS, which is the most common childhood soft tissue sarcoma, is an aggressive metastatic form of cancer, and is hypothesized to originate within muscle stem cells *per se* (Sesillo et al, 2019). At the molecular level, aRMS is characterized by chromosomal translocations of either Pax3-Foxo1 t(2;13) (q35; q14) or Pax7-Foxo1 t(1:13) (p36;q14), where these fusion genes create a novel transcription factor that imparts a dramatic increase in cell growth to enhance rhabdoid tumour formation and expansion (Manceau et al, 2022). The presence of the translocation is associated

with a worse prognosis independent of RMS histology (Kirk et al, 2010), yet little is known about the precise molecular event that propels the initial DNA strand break in aRMS prone loci.

Here, we sought to identify CAD bound loci using CUT and Tag, and how CAD targeted regions may influence muscle cell differentiation. Our observations indicate a genome wide distribution for CAD binding during early stages of skeletal muscle differentiation. Of note, we observed the *Pax7* gene body as a target of CAD, which was concurrent to decreased *Pax7* expression. Indeed, *Pax7* gene downregulation is required for the induction of muscle cell differentiation (Olguin & Olwin, 2004) and inhibition of CAD, through overexpression of a caspase resistant form of ICAD, led to sustained *Pax7* expression and inhibition of differentiation. Thus, the regulation of *Pax7* gene through CAD directed strand breaks implies a developmental mechanism to modulate cell fate by controlling the expression of nodal transcription factors. We also noted that both the *Pax7* and *Foxo1* genes were bound by CAD at gene regions syntenic to human aRMS gene loci, suggesting that the activity of this nuclease during cell differentiation might be the point of origin for this oncogenic fusion gene. Further analysis of CAD targeted genome indicated binding of this nuclease to dozens of loci, which align with fusion prone genes involved in the pathogenesis of leukemias and lymphomas. Binding of CAD to these regions suggest that CAD may act as a multi-lineage genome reprogramming nuclease. To understand whether these breakpoints are conserved, we performed CUT and Tag in early differentiating T cells and found that CAD bound to subgroup of the translocation prone genes identified in differentiating muscle cells. Taken together, our observations suggest that CAD activity in a differentiating cell environment may target gene loci to disrupt expression of factors that limit differentiation, in which many of these loci are predisposed to translocation, such as the reciprocal *Pax7 Foxo1* loci noted above.

Results

3.1.Genome Wide Mapping of CAD Strand Breaks Targets in Differentiating Myoblasts

To identify CAD genomic targets in muscle cells, I performed CUT and Tag, a technique that has an improved signal to noise ratio and requires fewer cell numbers to generate robust interaction libraries relative to ChIP-Seq (Heinkoff, 2019). CUT and Tag utilizes the use of protein A-fused Tn5 transposase (pA-Tn5), a primary antibody binds to the target chromatin protein in the genome, and a secondary antibody to enhance the tethering of pA-Tn5. I performed CUT and Tag on 150,000 fresh intact C2C12 cells using an antibody against endogenous CAD, at growth and 24 hours post induction of differentiation. The timepoint chosen is consistent with the peak formation of CAD induced DNA strand breaks, as published in (Larsen et al, 2010). Activation of pA-Tn5 tagments DNA producing fragments of 200-500 bp that are ready for PCR enrichment and DNA sequencing (Fig. 3.1A). Libraries were prepared and sent to the IRCM sequencing facility (Montreal PQ) using the Illumina NovaSeq 6000 platform.

To enhance the potential interactions for CAD with the DNA, cells were treated with z. DEVD-fmk, caspase-3 inhibitor. The rationale for this approach, was that prior studies revealed CAD can be bound to DNA while ICAD cleavage is blocked, suggesting a method to 'trap' CAD on its genomic targets. CAD nuclease has been shown to maintain strong binding with its substrate DNA, both as an active dimer and with greater binding kinetics when also inactive and in a complex with ICAD (Korn et al, 2005).

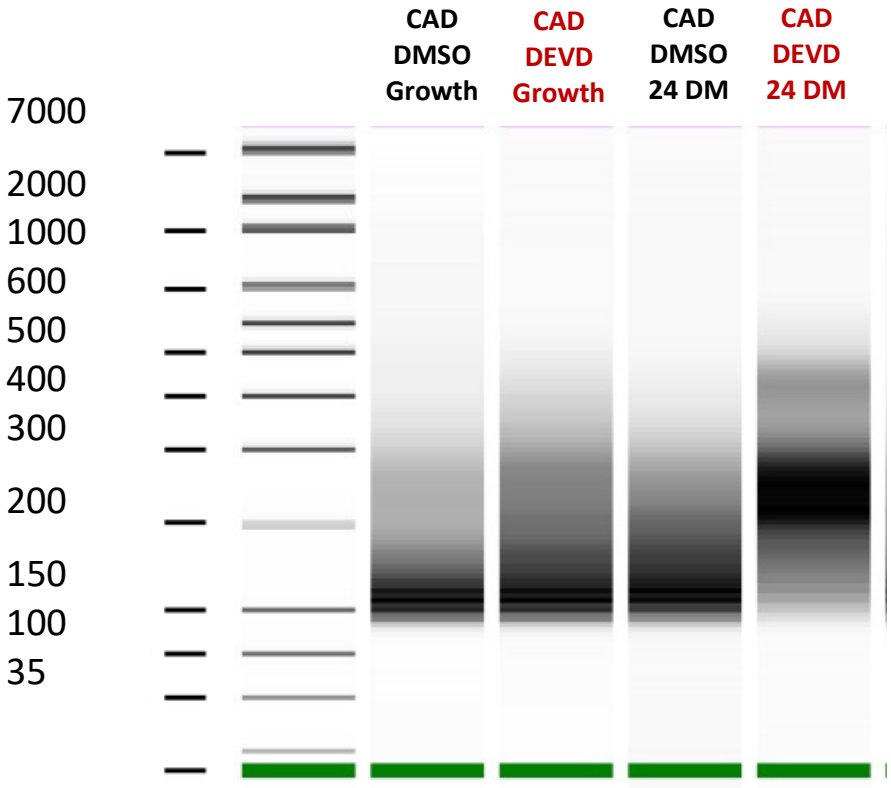
Here, we treated cells with either 15uM z. DEVD-fmk or 15u DMSO, and the results obtained from CUT and Tag demonstrated that the z. DEVD-fmk treatment improved the signal in differentiating cells, with a corresponding increase in binding targets near transcription start sites (TSS) in DEVD-

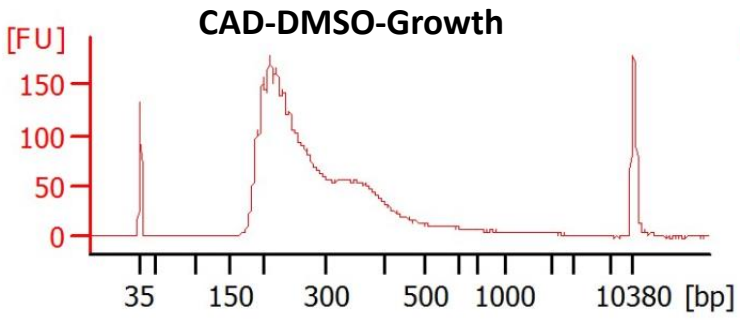
fmk treated cells compared to DMSO (as indicated in the plotHeatmap Fig. 3.1B). The plotProfile, which is represented as a single line graph, displays the rows condensed for the overall CAD binding site enrichment and strength of the signal for all genes (Fig. 3.1C). While no difference in growth conditions for CAD binding was noted in DEVD-fmk vs DMSO conditions, it is expected as caspase-3 targeting of the ICAD/CAD complex does not occur in growth condition leaving CAD inactive. A significant number of the CAD binding sites were identified within promoters at 52.26%, followed by exon-intron boundaries and distal intergenic regions at 44% (Fig. 3.1D), indicting a potential role for CAD in regulating gene expression.

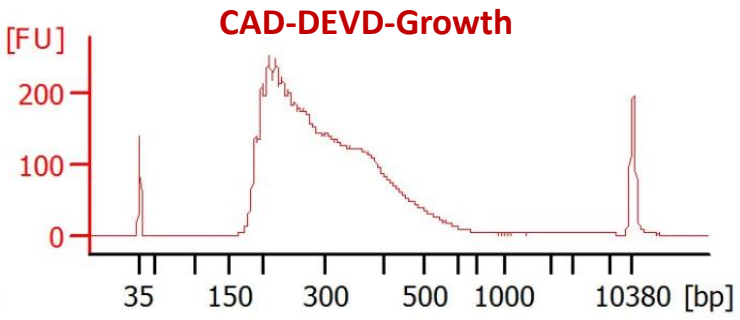
To identify which genes associated regions were bound by CAD and that were statistically enriched, KEGG and GO terms enrichment analysis were performed, listing the biological process and molecular functions by frequency of association (Fig. 3.1E-G). The top enriched terms included regulation of gene expression, regulation of differentiation of myeloid leukocyte and epithelial cell differentiation. In addition, DNA packaging, chromatin and nucleosome assembly were among the top terms identified for CAD bound genes/gene associated regions.

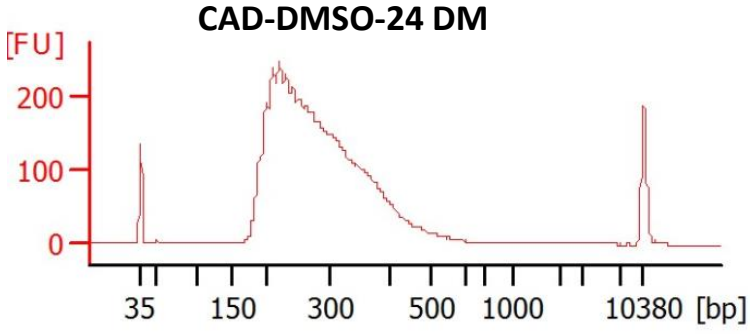
Library Size Distribution and Traces by Fragment Bioanalyzer

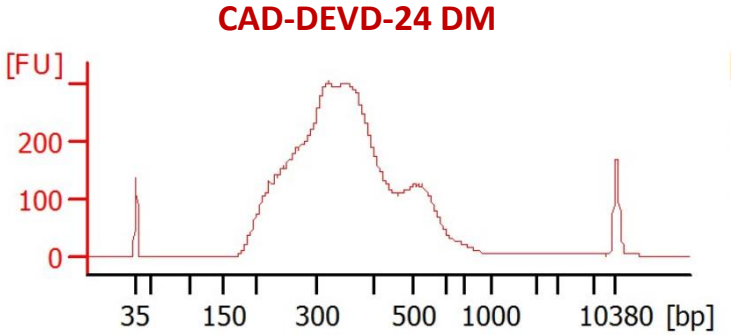
A





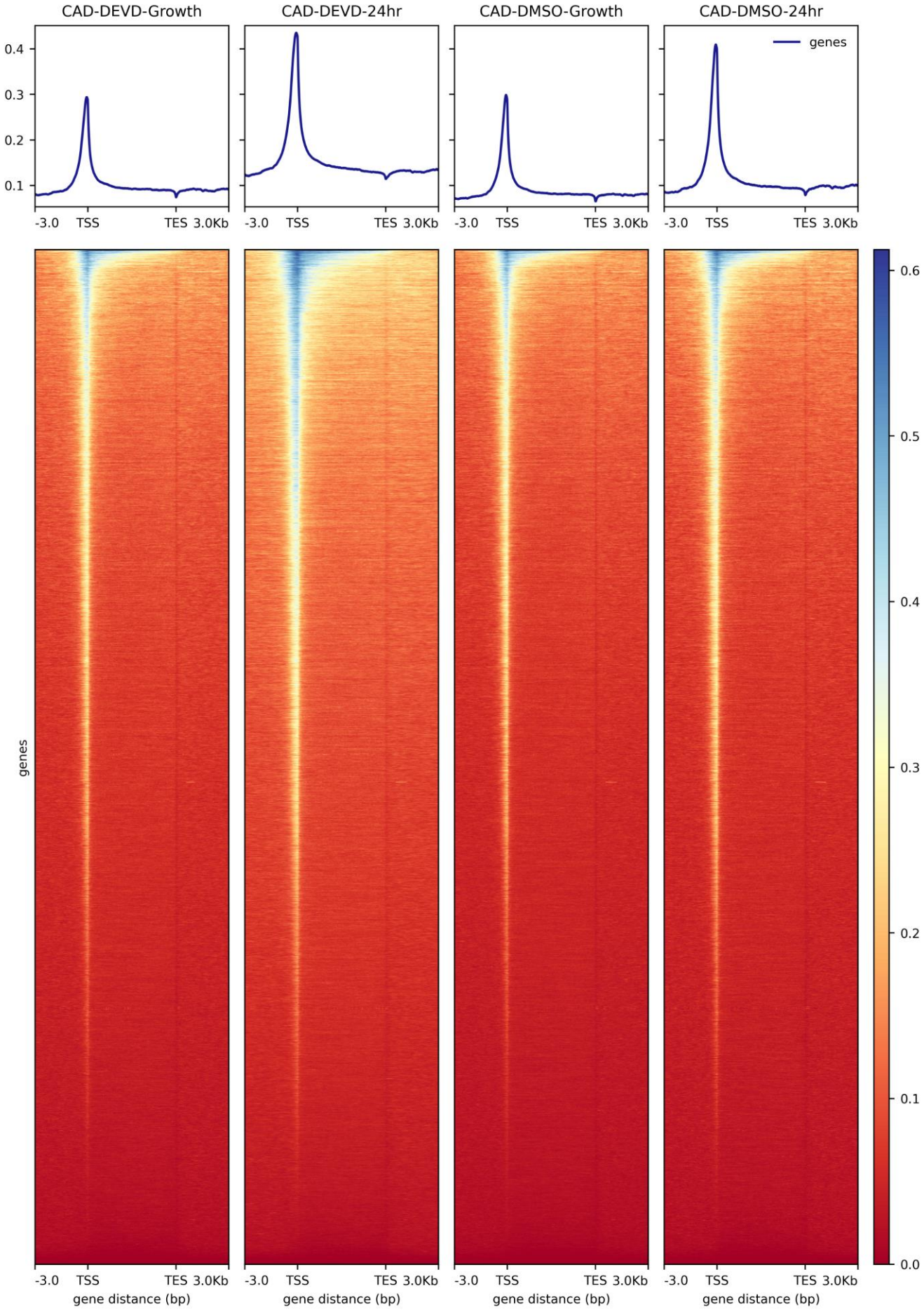






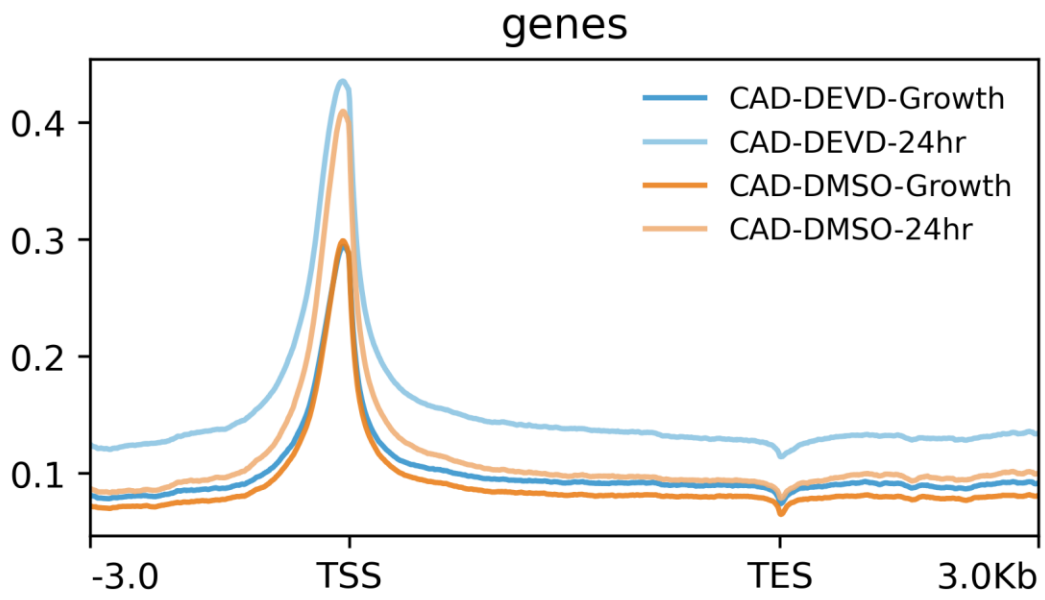
Genomic Distribution of CAD Binding Sites

B



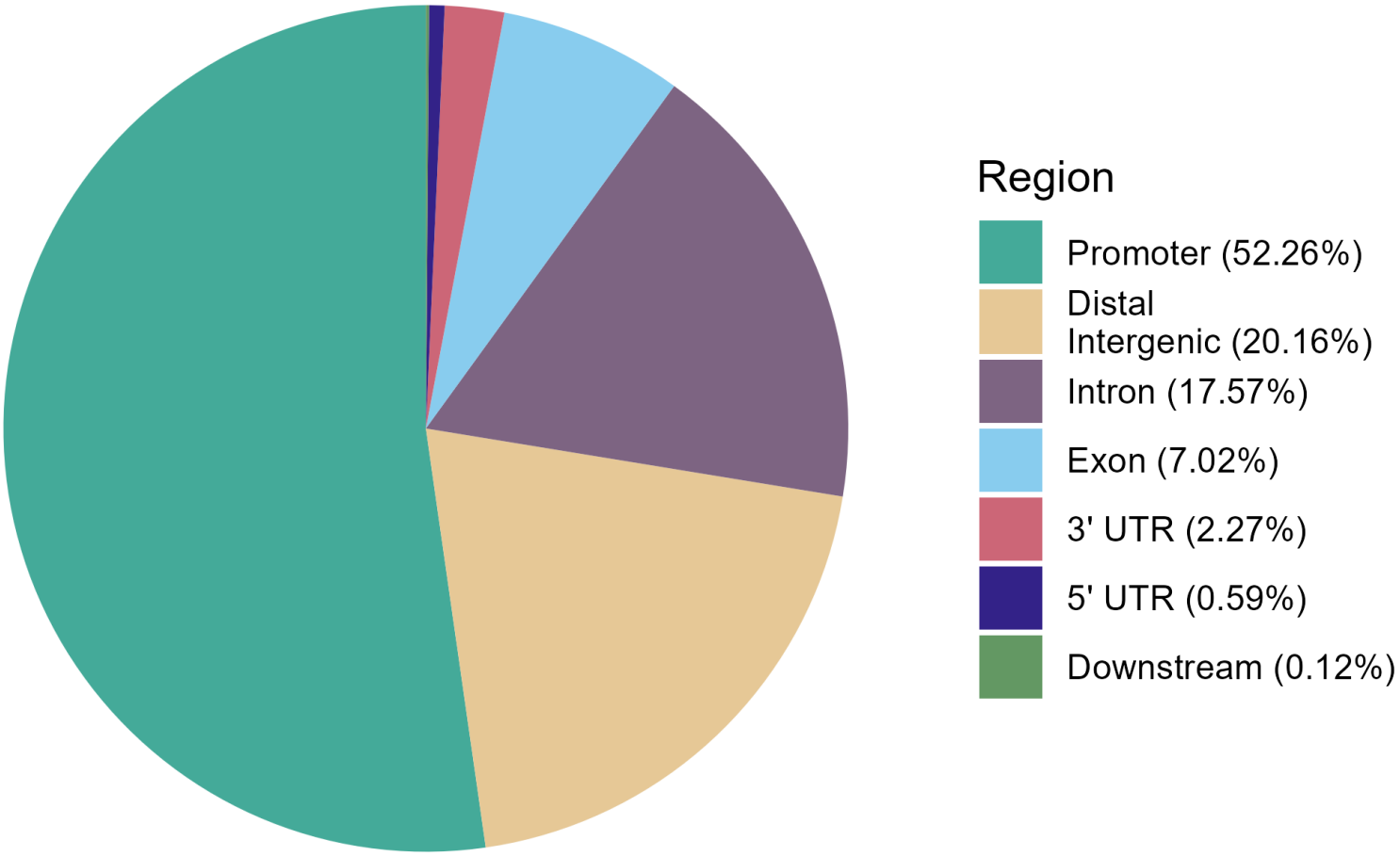
Genomic Distribution of CAD Binding Sites

C

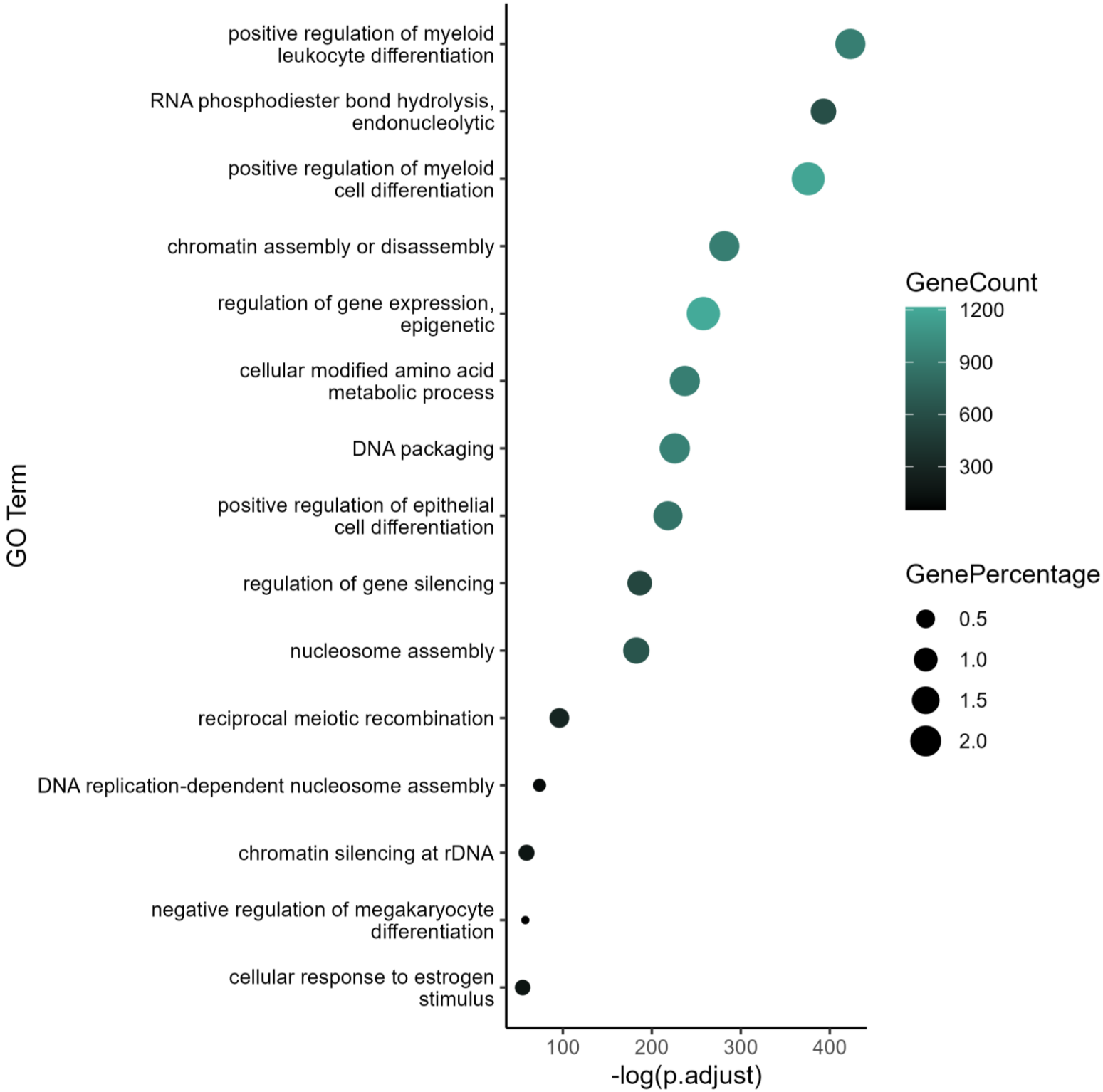


D

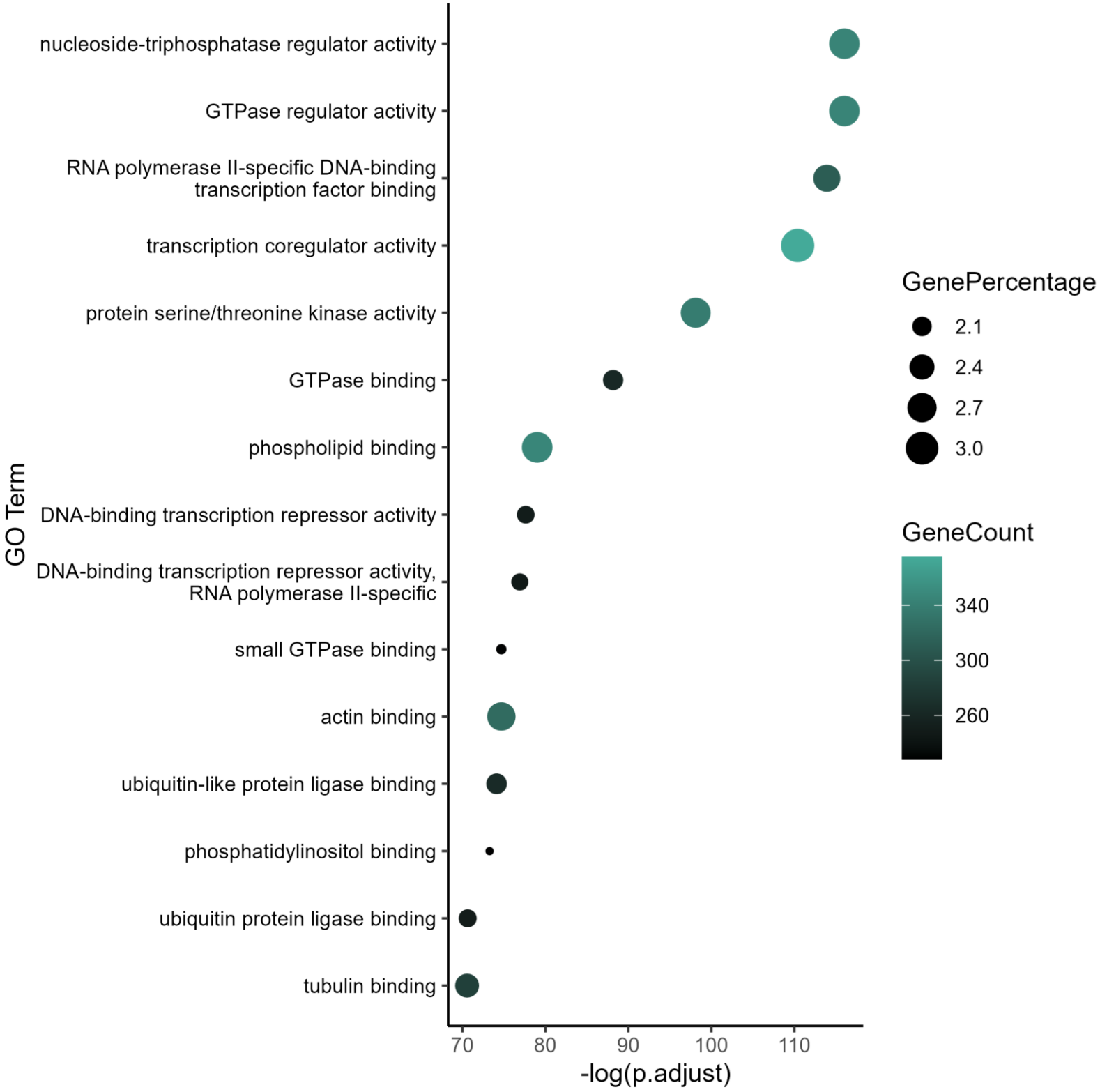
Genomic Distribution of CAD Binding Sites
CAD-DEVD-24 hr



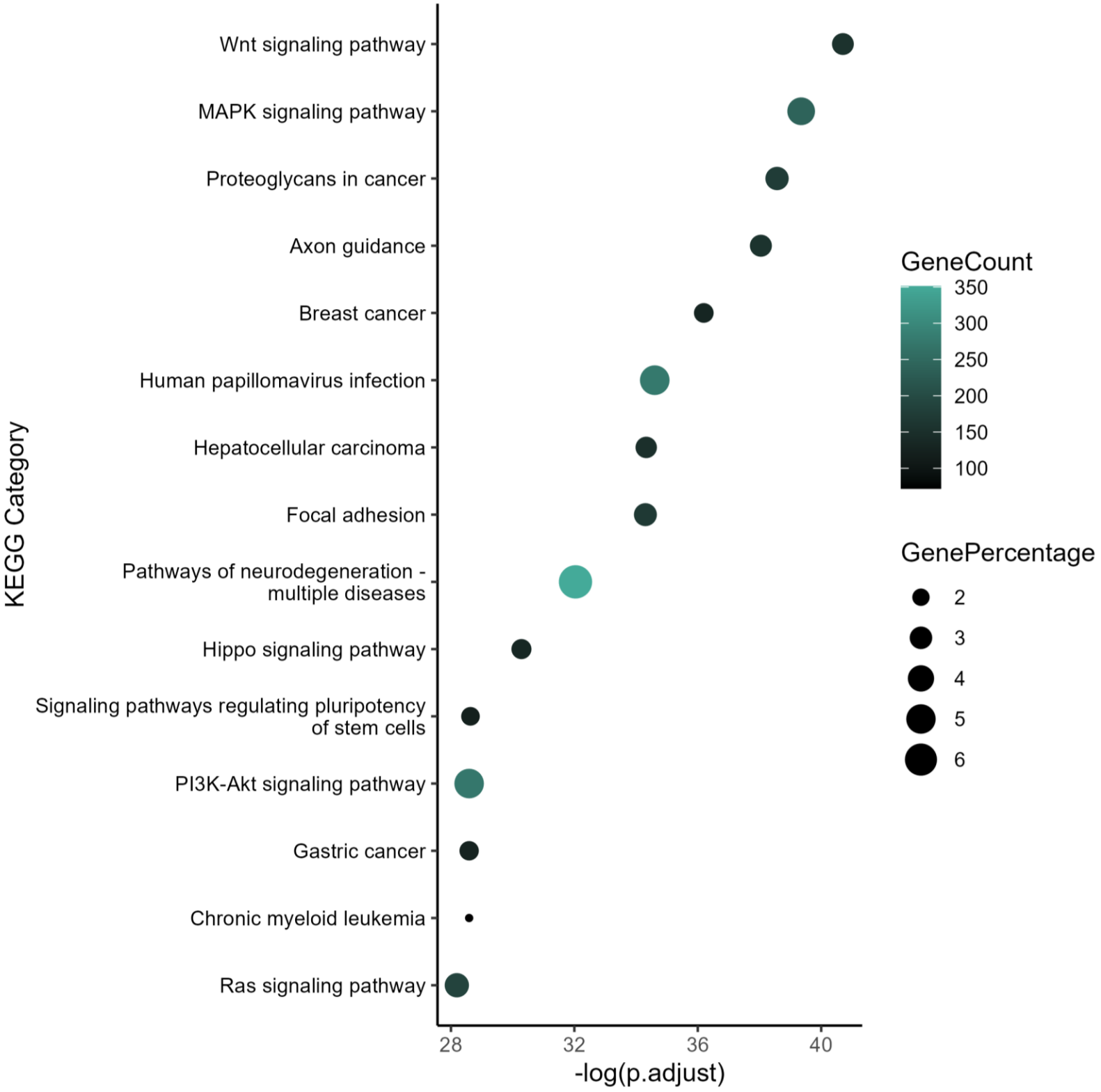
E **Gene Ontology (GO) Analysis Identified within CAD Functional Genes (Biological Process)**



F **Gene Ontology (GO) Analysis Identified within CAD Functional Genes (Molecular Function)**

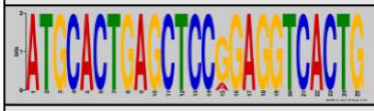
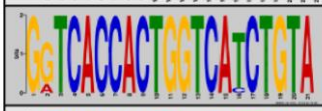
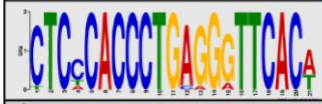


G **KEGG Pathway Analysis Identified within CAD Functional Genes**



H

Motif Enrichment Analysis of CAD Bound Genomic Regions

Discovered Motifs	E-Value	Sites	Width
	2.4e-035	7	25
	1.3e-011	11	25
	2.2e-059	10	25
	5.4e-066	13	25
	2.3e-029	9	21
	1.3e-011	11	25
	6.0e-047	11	21
	1.7e-075	16	21
	7.4e-035	8	25
	3.2e-028	9	25

I

Nucleotide Sequence of CAD Binding Sites

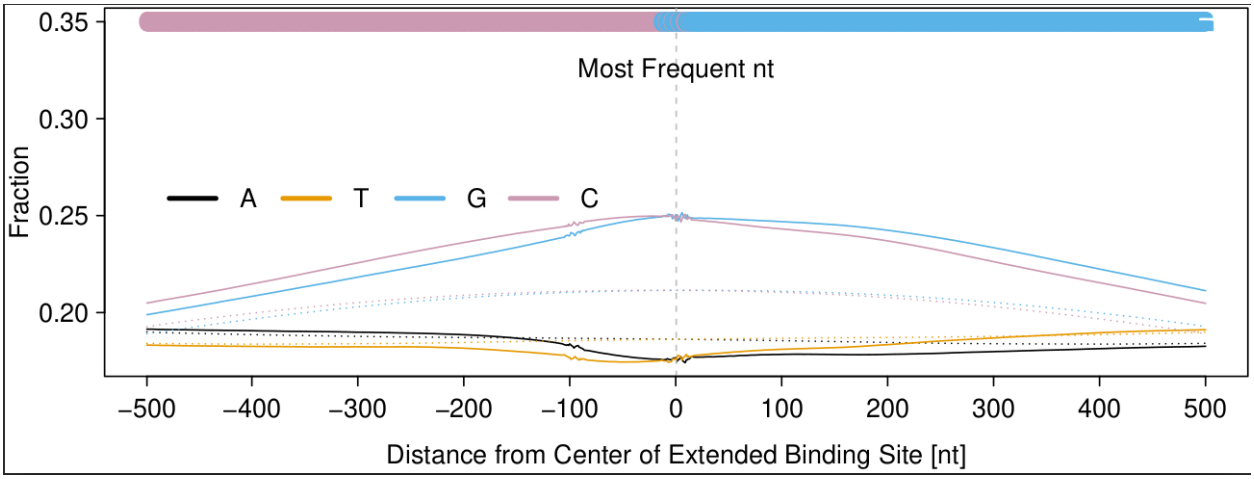


Fig. 3.1. Genomic distribution of CAD binding sites.
(A) Library size distribution and traces by fragment Bioanalyzer/Tape Station from CAD CUT and Tag concentration. Fragment size corresponding to each condition, and adapters, are indicated with an average expected fragment size of ~ 200-500 bp. Libraries were prepared from C2C12 cells using an antibody against endogenous CAD, at growth and 24 hours post induction of differentiation on cells treated with either z. DEVD-fmk or DMSO. (B) Genomic distribution of CAD signal for CAD binding sites. plotHeatmaps of CUT and Tag show count per million (CPM) normalized signals for reads across 3,000 bp genomic interval relative to the TSS. Genomic distribution of the binding sites of CAD at Growth, and 24 hours post differentiation from C2C12 muscle cells treated with DMSO or z.FMK-DEVD. Data sets were generated from merged two independent biological replicates for each condition. Heatmap profiles were generated using deepTools computeMatrix (Ramírez et al, 2016). Peak calling of binding sites was performed using Model-based Analysis of Chip-Seq. (MACS) (Zhanget al,2008), Sparse Enrichment Analysis for CUT & RUN (SEACR) (Meers et al, 2019), and GoPeaks (Yashar et al,2022). (C) PlotProfile is averaged profiles of the compressed matrix of occupancy/signal near TSS produced by computeMatrix. (D) Pie chart illustrating genomic distribution of the binding sites of CAD at 24 hours post differentiation from C2C12 muscle cells treated with z.FMK-DEVD. Different genomic regions displaying the genomic location of CAD binding sites across the genome including promoters, enhancers, distal intergenic regions, exons, introns, two untranslated regions (UTRs). This plot shows the percentage for each genomic location category. DiffBind tool was used to obtain the consensus peakset. AnnotatePeak function was used from ChIPseeker package on those peaks. Then a plotting package in R named ggplot2 that was used for the bar and pie charts, and the paletteer package for the color scheme. (E, F) Gene Ontology (GO) analysis identified within CAD functional genes Top 15 terms are enriched for biological processes (E) and molecular functions (F). (G) KEGG analysis identified within CAD functional genes. Top 15 terms are enriched. (H) Graphical representation of sequence motifs detected in CAD DNA sequences motif enrichment analysis identifies de novo motif showing the top enriched sequence motifs of CAD bound genomic regions. Motifs enrichment was calculated with HOMER tool. E-values of motifs are indicated. The statistical significance of the motif (E-value) is an estimate of the log ratio, width, site count, and the background letter frequencies. (I) All peaks for CAD are called by MACS regions extended +/- 500bp from their summits. Followed by mapping them to masked mm10 genome to get nucleotide sequences. The alignments have higher representation of GC-rich sequences in CAD binding genome.

3.2.CAD has no Specific DNA Binding Motif in Differentiating Cells

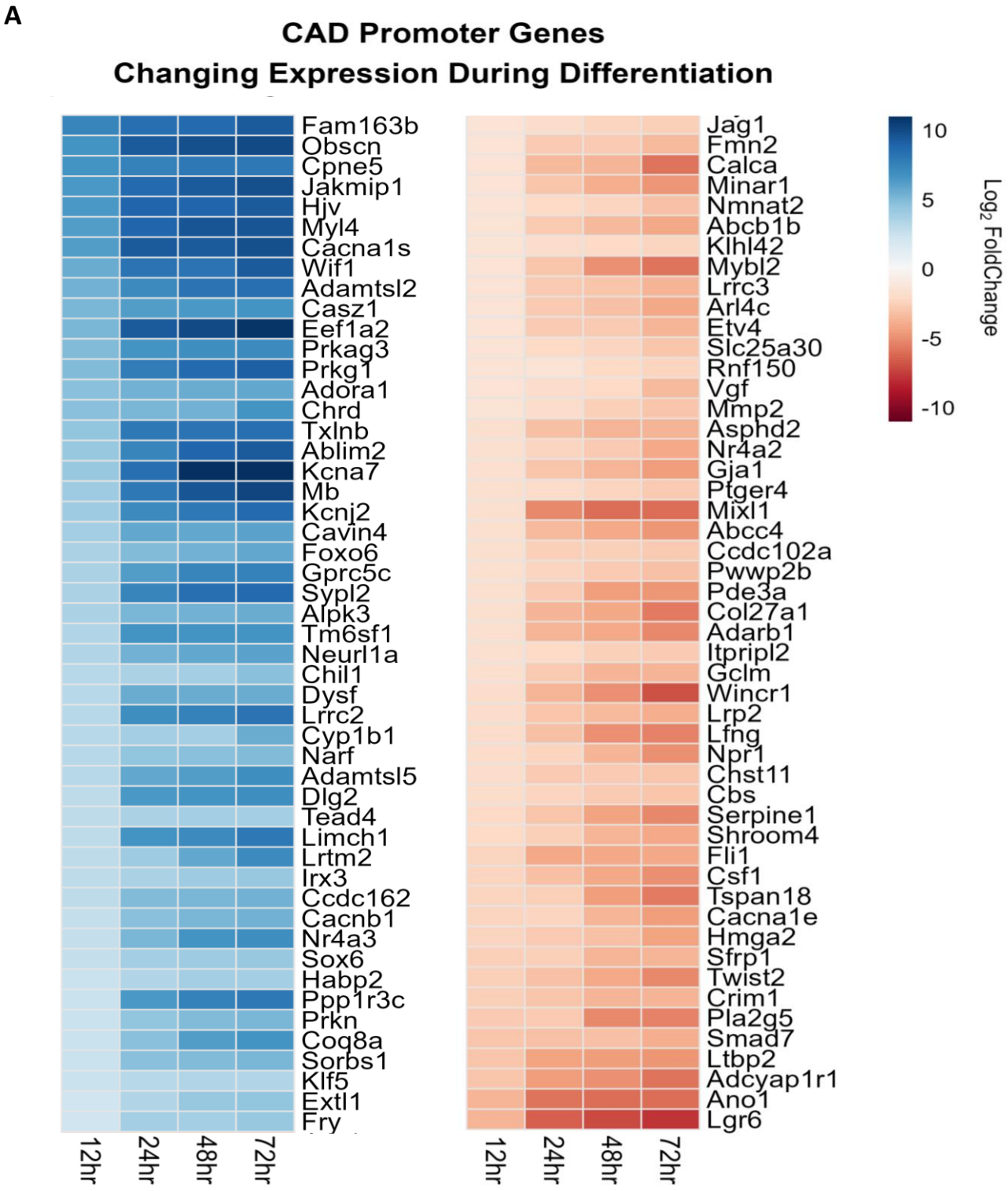
To further characterize the CAD bound regions and whether there is evidence for conserved binding motifs in differentiating muscle cells, I performed motif enrichment analysis using MEME-ChIP tool. The most frequent motifs detected in the DNA sequence within the full set of CAD peaks are graphically depicted in (Fig.3.1H). Most CAD bound loci lie in GC-rich regions (Fig.3.1I), where no obvious binding motif was identified. CAD interactions with a CTCF binding motif has been reported in irradiated cancer cells (Larsen et al, 2022), however no obvious recurrent CAD association with a CTCF motif is detected in differentiating myoblasts.

Investigation of DNA sequence motifs within the full set of CAD peaks revealed no conserved motifs on a genome-wide scale. This observation would suggest that other structural imperatives may manage or direct CAD binding and/or access to the genome. Indeed, our CUT and Tag results revealed a strong association between CAD binding sites and the genome binding regions affiliated with epigenetic enzymes. This is also represented in the GO terms of top enriched genes bound by CAD; chromatin regulation and gene silencing (Fig.3.1E). Thus, we speculated that the epigenetic modifications that occur at DNA damage sites have the potential to regulate CAD access to genomic loci in differentiating myoblasts (which will be further discussed in chapter 4). This is a reasonable assumption as CAD targeting of the genome is extensive and has the potential to redirect cell fate to either an apoptotic or cell differentiation outcome. Overlap binding of repressive epigenetic modifiers with CAD bound sites would ensure that a cell has additional levels of control over CAD's ability to propel cell fate choices, beyond the caspase 3 directed cleavage of ICAD alone.

3.3.CAD Binds Gene Promoters Correlated with Increased Gene Expression in Muscle Cell Differentiation

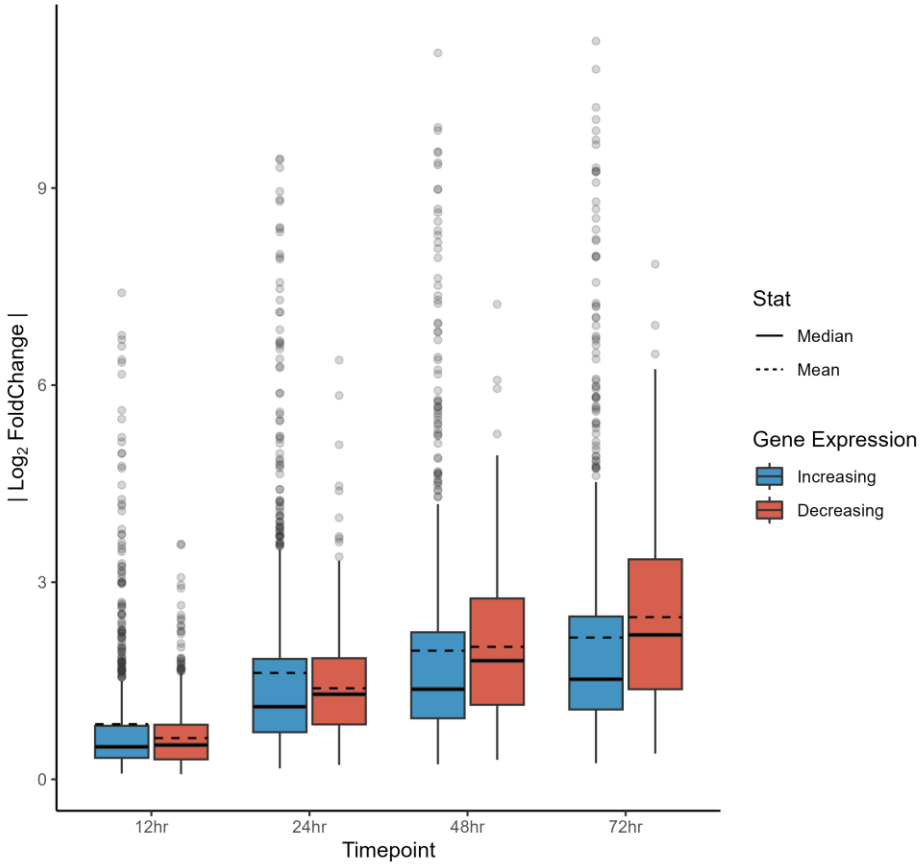
Prior observations from our group noted CAD bound to the p21 (CDKN1A) gene promoter and inflicted strand breaks that were synonymous with increased p21 gene expression, therefore, modifying cell differentiation (Larsen et al, 2010). In addition, here I found that 52% of the total CAD binding sites within core promoters (<1 KB from the TSS) or distal enhancers 1 to 3 KB 5' from the TSS (Fig. 3.1D). Peak calling across CAD bound genes generated a list of individual promoters, many of which are known to increase in expression and promote early stages of cell differentiation. RNA-Seq. analysis of differentiating murine primary myoblasts cells revealed that a significant percentage of these same CAD bound promoters are also associated with a respective increase in expression (Fig. 3.2A). This indicates CAD may act as a genome wide initiation cue for the adopting a differentiation specific gene expression profile on genes with single SSBs (Fig. 3.2A-E). Conversely, the cohort of decreased expressed genes in differentiating muscle cells is likely reflective of CAD induced DSBs, which are known to cause transcription repression (Fig. 3.2A, C-E) (He et al, 2024). Moreover, in 24 hours differentiation, CAD protomer genes seem to be more inductive relative to 12 hours as depicted in volcano plot (Fig. 3.2D, E).

A


CAD Promoter Genes
Changing Expression During Differentiation
Fam163b
Obscn
Cpne5
Jakmip1
Hjv
Myl4
Cacna1s
Wif1
Adamtsl2
Casz1
Eef1a2
Prkag3
Prkg1
Adora1
Chrd
Txlnb
Ablim2
Kcna7
Mb
Kcnj2
Cavin4
Foxo6
Gprc5c
Sypl2
Alpk3
Tm6sf1
Neurl1a
Chil1
Dysf
Lrrc2
Cyp1b1
Narf
Adamtsl5
Dlg2
Tead4
Limch1
Lrtm2
Irx3
Ccdc162
Cacnb1
Nr4a3
Sox6
Habp2
Ppp1r3c
Prkn
Coq8a
Sorbs1
Klf5
Extl1
Fry
12hr
24hr
48hr
72hr
Jag1
Fmn2
Calca
Minar1
Nmnat2
Abcb1b
Klhl42
Mybl2
Lrrc3
Arl4c
Etv4
Slc25a30
Rnf150
Vgf
Mmp2
Asphd2
Nr4a2
Gja1
Ptger4
Mixl1
Abcc4
Ccdc102a
Pwwp2b
Pde3a
Col27a1
Adarb1
Itpripl2
Gclm
Wincr1
Lrp2
Lfng
Npr1
Chst11
Cbs
Serpine1
Shroom4
Fli1
Csf1
Tspan18
Cacna1e
Hmga2
Sfrp1
Twist2
Crim1
Pla2g5
Smad7
Ltbp2
Adcyap1r1
Ano1
Lgr6
12hr
24hr
48hr
72hr
10
5
0
-5
-10
Log2 FoldChange

B

CAD Promoter Genes
Changing Expression During Differentiation



C

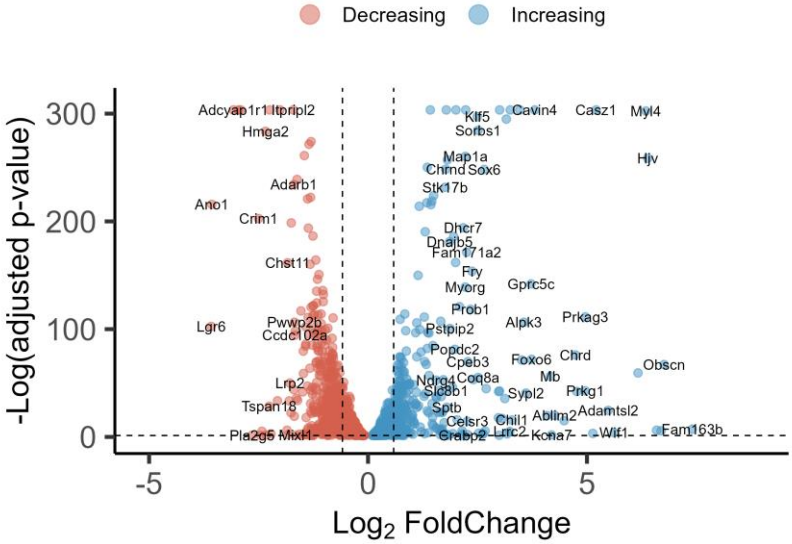
Upregulated CAD Promoter Genes at 12- and 24-hour Differentiating Cells

	12 DM		24 DM	
	Gene	Log Fold Change	Gene	Log Fold Change
1	Obscn	6.76	Obscn	9.45
2	Myl4	6.34	Myl4	8.83
3	Casz1	5.21	Casz1	6.26
4	Mypn	4.93	Mypn	6.97
5	Sgcd	3.79	Sgcd	8.69
6	Tead4	3.01	Tead4	3.71
7	Igf2	2.91	Igf2	5.25
8	Klf5	2.50	Klf5	3.26
9	Cdkn1a	2.01	Cdkn1a	2.80
10	Tbx15	2.00	Tbx15	3.71
11	Sox6	2.65	Sox6	3.79
12	Prkag3	4.96	Ucp2	2.60
13	Txlnb	4.48	Atp6v0a1	2.07
14	Cavin4	3.81	Acacb	2.78
15	Coq8a	2.51	Tgm2	3.70

D

CAD Promoter Genes
Changing Expression During Differentiation

t = 12hr



E

CAD Promoter Genes
Changing Expression During Differentiation

t = 24hr

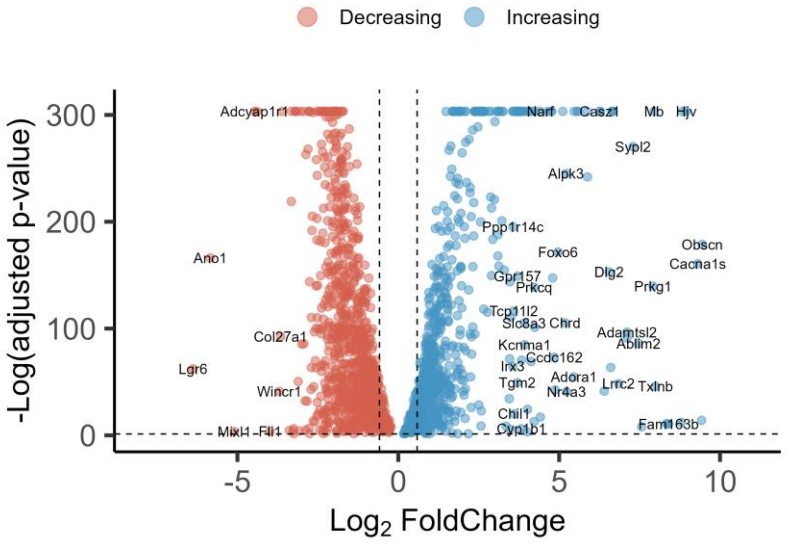


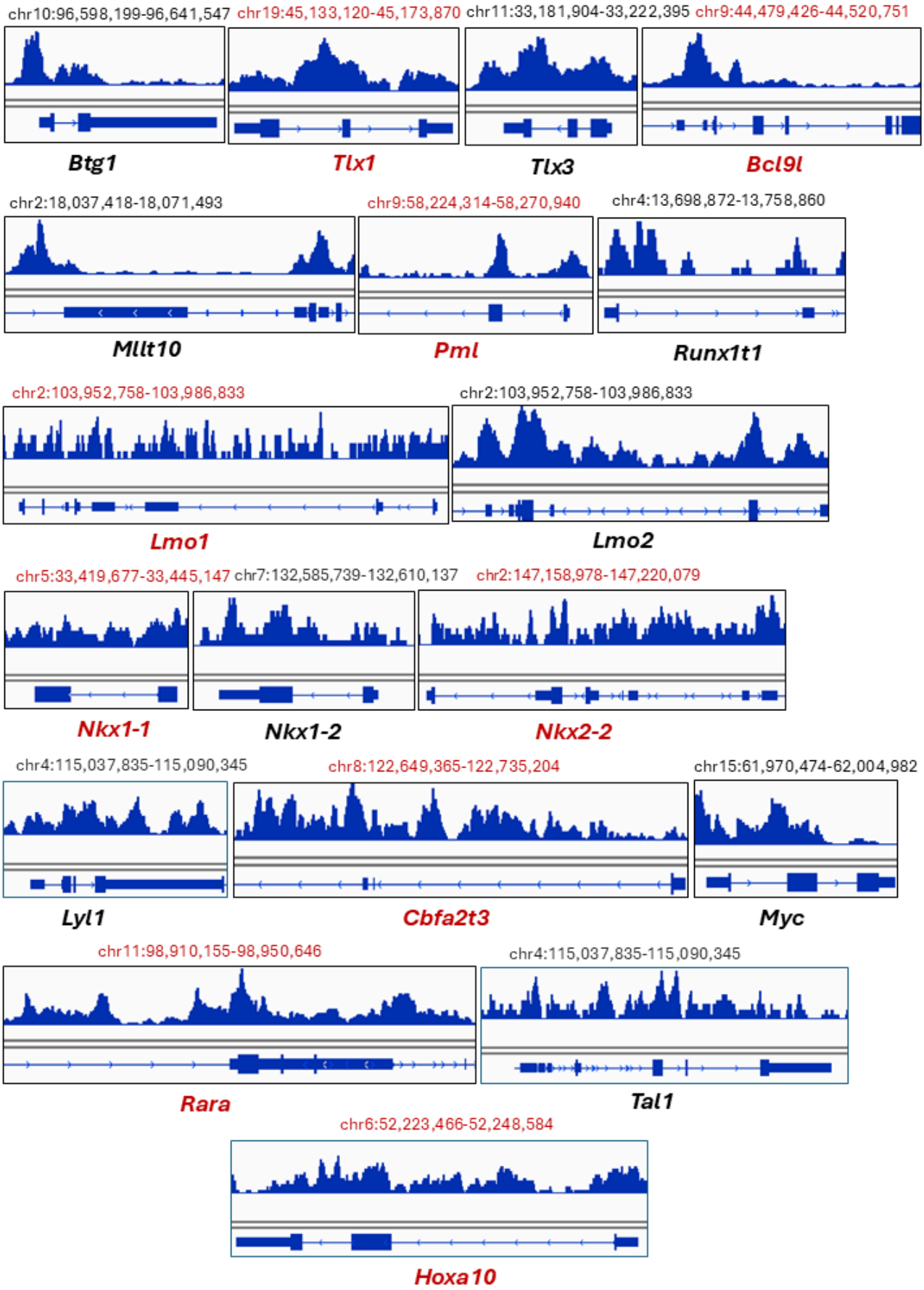
Fig. 3.2. RNA-Seq. analysis of CAD promoter genes.
(A) Heatmaps depicting changes in gene expression levels in CAD promoter genes during 12-, 24-, 48-, and 72-hours differentiation. Blue color indicates elevated mRNA levels, while red color indicates repressed levels. Color scale represents the log2fold change. The statistics for the RNA-Seq data come from DESeq2 results performing pairwise comparisons of each differentiation timepoint with the growth timepoint as reference. Three biological replicates for each timepoint were assessed, and DESeq2 results were obtained to get the log2FoldChange and p values. Genes with p. adjusted < 0.05 across all timepoint comparisons were kept for the expression trend analysis. (B) Box-plot diagram depicts the distributions of gene expression log2Fold change across different differentiation timepoints 12, 24, 48, and 72 hours. The central line in the box marking the median value. (C) A subset of CAD bound genes (promoters) listed with an increase in gene expression in C2C12 muscle cells at 12- and 24-hour differentiating cells. All genes were filtered first that the adjusted p-value for all timepoints is < 0.05. Then the median was taken for each timepoint 12 and 24 hours to be equal to or greater than 2 log2fold value. All genes listed are above the median. (D, E) The volcano plot shows the genes with trending expression changes at 12hr (D) and 24 hr (E) differentiation for which CAD is bound to those gene promoters.

3.4.CAD Bound Genes that Comprise the Parent Genes for Breakpoint Translocations in Differentiating Muscle Cells

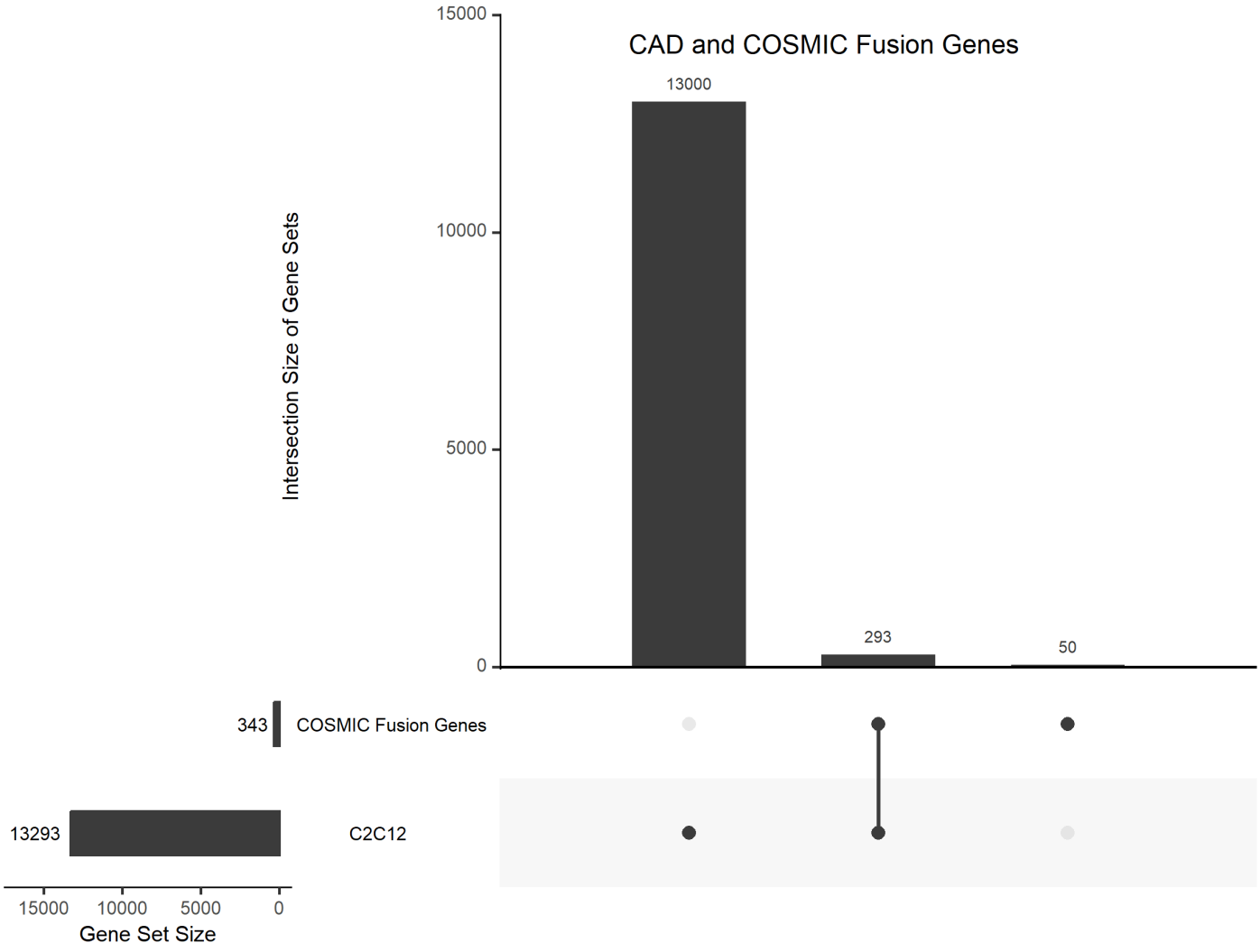
The CUT and Tag analysis also identified a common CAD binding event in differentiating C2C12 muscle cells (44% of the total CAD peaks called), specifically in exons-introns and distal intergenic (Fig. 3.1D). This concentration of CAD within gene bodies suggests a second function for this nuclease during cell differentiation, where DNA strand breaks within exons-introns are less likely to be gene inductive events and may be more inclined to create structural impediments for gene expression, as noted above (He S, et al, 2024). The Initial curation and inspection of the CAD intragenic targets also revealed enrichment of genes that give rise to oncogenic breakpoint translocations; where breaks in individual genes are mismatched with other broken genes in the subsequent DNA repair process, forming novel fusion genes (Rowley et al, 1977, Dalla-Favera et al, 1983, Cowell et al, 2023). These CAD targets include a broad range of known oncogenic breakpoints. Browser view of gene tracks of the respective genes is shown in (Fig. 3.3A), and listed in (Table 3.1). The list includes *Tal1, Lyl1, Tlx1, Tlx3, Pax7, Pax3, Foxo1, Myc, Runx1t1, Rara, Pml, Cbfa2t3, Mllt10, Lmo2, Btg1*, *Nkx1-1*, *Nkx1-2*, *Nkx2-2*, and *Bcl9l* genes, which are known translocation prone genes that trigger hematopoietic cancers. CAD targeting of breakpoints prone genes during cell differentiation may be limited in scope, impacting a small subset of genes, or CAD may range across the genome, targeting all breakpoints sensitive loci. To address the relative frequency of CAD targeting to breakpoint affected genes in differentiating murine muscle cells, we conducted a direct comparison of the CAD bound genome against the known compendium of described human cancer causing breakpoints, as curated by the Catalogue Of Somatic Mutations In Cancer (COSMIC; https://cancer.sanger.ac.uk/cosmic/about). The degree of CAD association with the COSMIC human breakpoint genes was very high in differentiating

mouse muscle cells, with CAD bound to 85% of the total breakpoint gene pool (293/343 genes) (UpSet plot Fig. Fig. 3.3B; Circablot Fig. 3.3C). This cross-species CAD gene binding in murine muscle cells with human breakpoint sensitive genes was also characterized by a significant degree of syntenic site overlap between CAD and the gene breakpoints (n=180 breakpoints) (Fig. 3.3C). These data suggest that the CAD nuclease is positioned within the genome to inflict breaks that may preconfigure breakpoint formation. Similar to the CAD bound promoters, CAD targeting of breakpoints sensitive genes may be part of inducing the differentiation program as noted above, i.e. breaking a gene through a DSB to reduce its expression. This is a reasonable assumption, as it is very unlikely that breakpoints would be a conserved alteration solely dedicated to the induction of a cancer mutation *per se*. A cross comparison of CAD bound fusion genes and CAD-bound breakpoint genes against the RNA-Seq revealed that a number of these CAD target genes have decreased levels of expression in differentiating muscle cells versus growing cells (Fig. 3.3D-G).

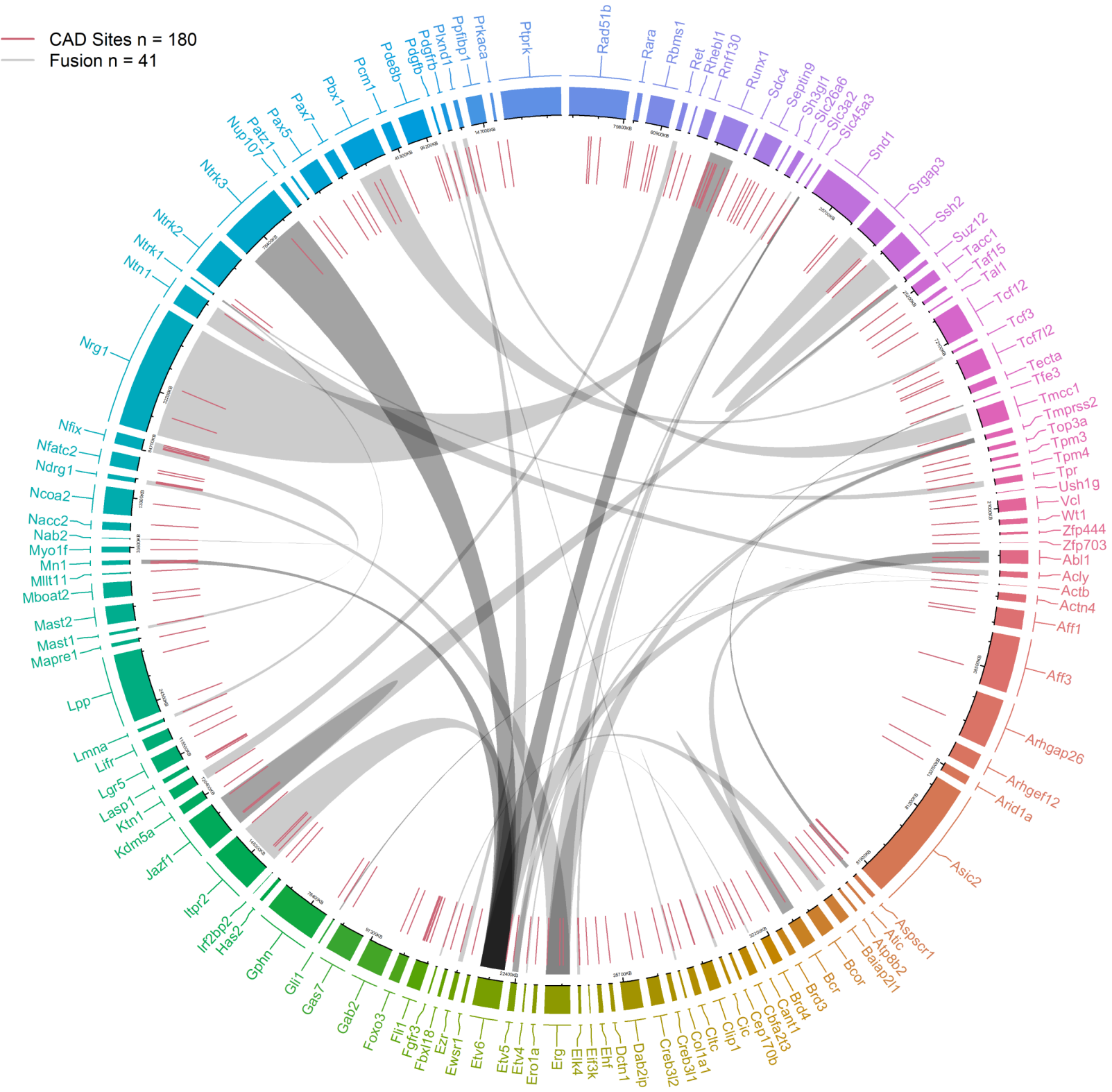
A **Detection of Gene Fusions Breakpoint in C2C12 Skeletal Muscle Cells**



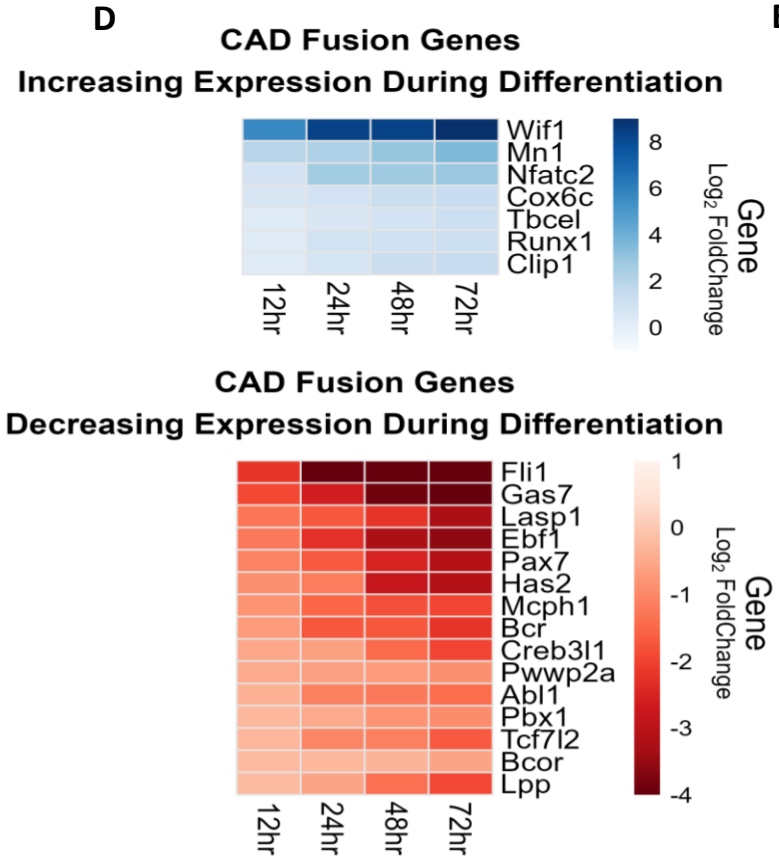
B **Intersection Size of Fusion Gene Shared between CAD from C2C12 Cells and COSMIC Database**



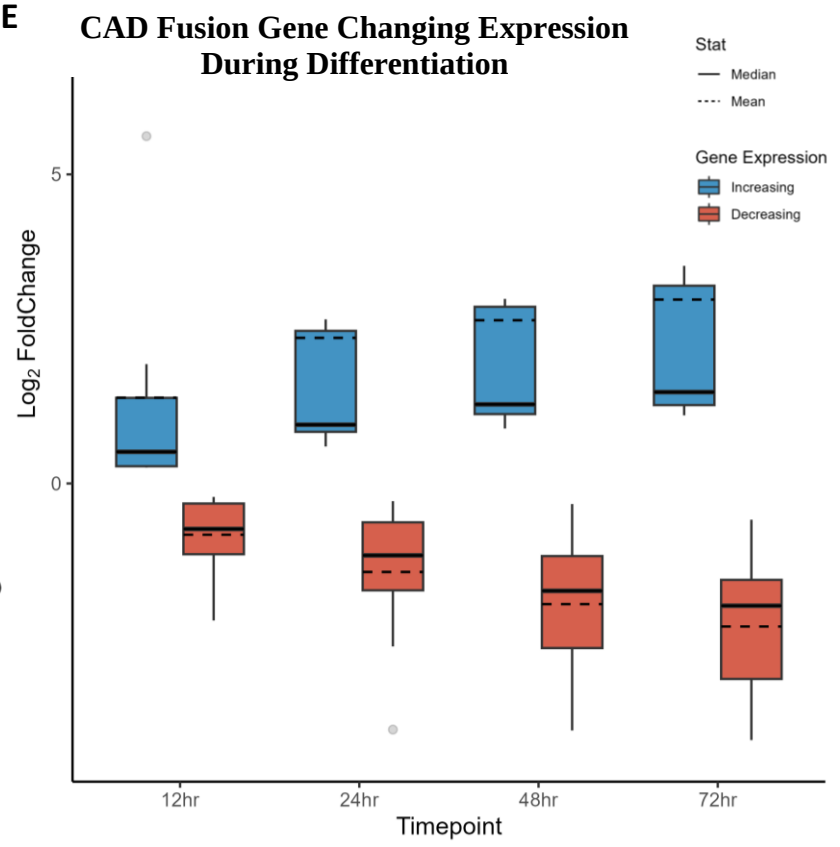
C **C2C12 CAD Sites Overlapping COSMIC Breakpoints in 125 COSMIC Fusion Genes**



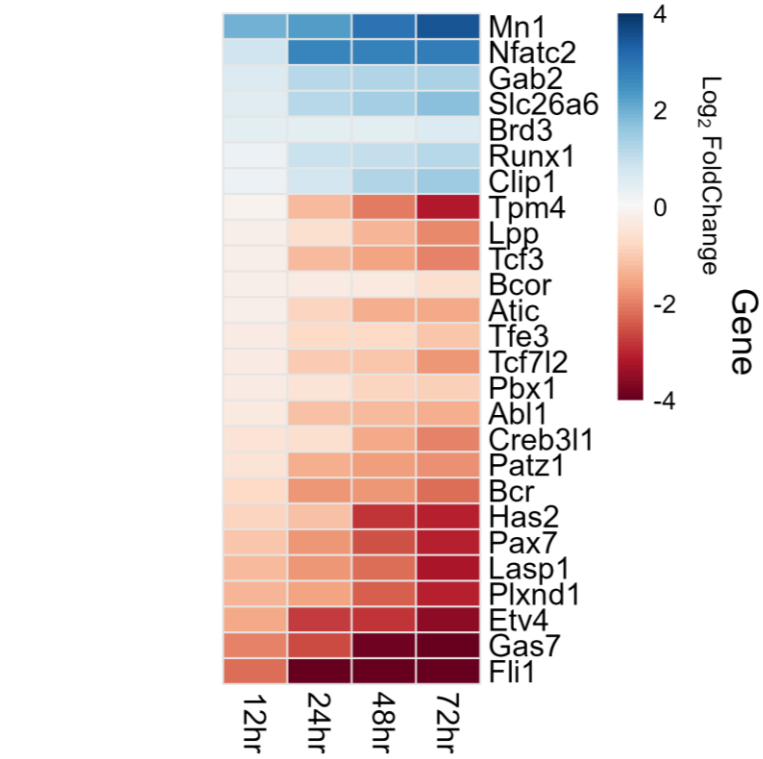
D


CAD Fusion Genes
Increasing Expression During Differentiation
Wif1
Mn1
Nfatc2
Cox6c
Tbcel
Runx1
Clip1
12hr
24hr
48hr
72hr
8
6
4
2
0
Log2 FoldChange
Gene
CAD Fusion Genes
Decreasing Expression During Differentiation
Fli1
Gas7
Lasp1
Ebf1
Pax7
Has2
Mcph1
Bcr
Creb3l1
Pwwp2a
Abl1
Pbx1
Tcf7l2
Bcor
Lpp
12hr
24hr
48hr
72hr
1
0
-1
-2
-3
-4
Log2 FoldChange
Gene

E


CAD Fusion Gene Changing Expression
During Differentiation
Stat
Median
Mean
Gene Expression
Increasing
Decreasing
Log2 FoldChange
5
0
12hr
24hr
48hr
72hr
Timepoint

F


CAD Breakpoint Genes
Changing Expression During Differentiation
Mn1
Nfatc2
Gab2
Slc26a6
Brd3
Runx1
Clip1
Tpm4
Lpp
Tcf3
Bcor
Atic
Tfe3
Tcf7l2
Pbx1
Abl1
Creb3l1
Patz1
Bcr
Has2
Pax7
Lasp1
Plxnd1
Etv4
Gas7
Fli1
12hr
24hr
48hr
72hr
4
2
0
-2
-4
Log2 FoldChange
Gene

G

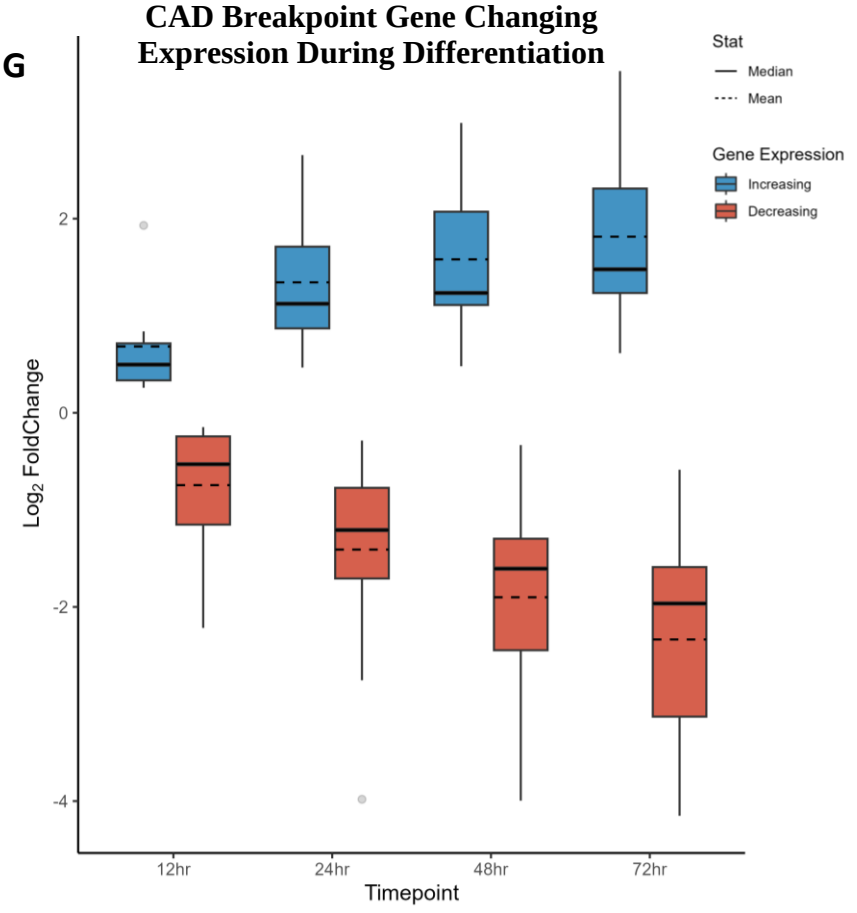

CAD Breakpoint Gene Changing
Expression During Differentiation
Stat
Median
Mean
Gene Expression
Increasing
Decreasing
Log2 FoldChange
2
0
-2
-4
12hr
24hr
48hr
72hr
Timepoint

Fig. 3.3. CAD activity in differentiating muscle cells binds the parent genes that give rise to breakpoints in C2C12 cells. (A) Integrative genomics viewer (IGV) snapshots depicting CAD CUT and Tag signal and associated peak call. (B) Upset plot visualizes the intersection size of gene sets between CAD (from C2C12) cells and fusion genes from COSMIC database. 293/343 sites of COSMIC genes are shared. Upset plots were made with the UpSetR package's upset function. COSMIC fusion genes were mapped to the mm10 orthologs from bioMart and obtained the intersection and union of the CAD and COSMIC genes to plot the Upset plot. (C) Circa plot visualizing cross-comparison between CAD-binding sites and COSMIC fusion genes set. Tracks from inside to the outside correspond to (1) gene synteny between chromosomes; (2) gene. For the circos plot, from outer to inner: Rainbow: represents COSMIC fusion genes for which CAD sites are bound (i.e. CAD-Fusion genes), Red/teal/purple lines represents CAD site locations where COSMIC Breakpoints overlap with the CAD sites (i.e. CAD-Breakpoints) in the fusion genes, lined-up with their respective genes in their gene ranges, and coloured for C2C12. The innermost links "Grey" indicate: Gene fusion pairs for CAD-Fusion genes in the circus. liftOver tool (in R) was used to map COSMIC breakpoints to map them to mm10. CAD peaks and breakpoint overlaps in the CAD-Fusion gene regions were obtained using ChIPseeker packages's find Overlaps OfPeaks function). Circa plots were made with the circlize package (R package). (D) Heatmaps depicting changes in gene expression levels in CAD fusion genes during during 12-, 24-, 48-, and 72-hours differentiation. Blue color indicates elevated mRNA levels, while red color indicates repressed levels. Color scale represents the log2fold change. Three biological replicates for each timepoint were assessed. (E) Boxplot representing CAD fusion genes changing expression during differentiation. (F) Heatmaps depicting changes in gene expression of CAD breakpoint genes during 12-, 24-, 48-, and 72-hours differentiation. Blue color indicates elevated mRNA levels, while red color indicates repressed levels. Color scale represents the log2fold change. (G) boxplot representing CAD breakpoint genes changing expression during differentiation.

Symbol	Gene name	pValue	qValue	SignalValue	Fusion genes associated with chromosomal rearrangement	References
Fus	fused in sarcoma	55.78	51.28	12.80	**FUS- CREB3L2** **FUS-DDIT3**	Panagopoulos et al, 2004, Rose et al, 2011, Zullow et al, 2022
KDM3B	KDM3B lysine (K)-specific demethylase 3B	41.71	37.48	9.66	**KDM3B-ETF1**	Hu et al, 2021
Ddit3	DNA-damage inducible transcript 3	38.85	34.69	6.66	**FUS-DDIT3**	Zullow et al, 2022
Set	SET nuclear oncogene	37.41	33.28	9.18	**SET-NUP214**	Chen et al, 2017
Stil	Scl/Tal1 interrupting locus	20.32	16.76	8.49	**STIL-TAL1**	Furness et al, 2018
Tal1	T cell acute lymphocytic leukemia 1	16.57	13.18	4.48	**STIL-TAL1** **MBNL1-TAL!**	Furness et al, 2018
Cbl	Casitas B-lineage lymphoma	15.65	12.31	6.39	**KMT2A-CBL**	Bataller et al, 2021
Baiap2l1	BAI1-associated protein 2-like 1	15.54	12.20	6.14	**CUX1-BAIAP2L1**	Pan et al, 2024
Pax7	Paired-box 7	**15.25**	**11.93**	**6.40**	**Pax7-Foxo1**	Dumont et al, 2012, Marshal et al, 2012
Nup214	nucleoporin 214	12.12	8.98	6.53	**NUP214-ABL1**	Chen et al, 2017
Erg	ETS transcription factor	12.02	8.89	5.61	**EWSR1-ERG**	Chen et al, 2015, Zou et al, 2023
Zbtb16	zinc finger and BTB domain containing 16	11.46	8.36	5.52	**ZBTB16-ABL1**	Chen et al, 2017
Mllt1	myeloid/lymphoid or mixed-lineage leukemia; translocated to, 1	11.27	8.19	5.82	**KMT2A-MLLT1**	De Matos et al, 2022
Foxo1	Forkhead box O1	**11.06**	**8.00**	**4.97**	**Pax7-Foxo1** **Pax3-Foxo1**	Dumont et al, 2012, Marshal et al, 2012
Creb3l2	cAMP responsive element binding protein 3-like 2	10.59	7.56	5.65	**FUS-CREB3L2**	Panagopoulos et al, 2004, Rose et al, 2011
Pax3	Paired-box 3	9.86	6.88	6.08	**Pax3-Foxo1**	Dumont et al, 2012, Marshal et al, 2012
Mbnl1	muscleblind like splicing regulator 1	9.45	6.51	4.54	**MBNL1-TAL1**	Chen et al, 2017
Mllt10	myeloid/lymphoid or mixed-lineage leukemia; translocated to, 10	9.23	6.30	5.69	**KMT2A-MLLT10**	Kim et al, 2023, Yamamoto, et al, 2023
Cux1	cut-like homeobox 1	8.55	5.67	4.12	**CUX1-BAIAP2L1** **CUX1-ALK**	Zhang et al, 2018
ETF1	eukaryotic translation termination factor 1	8.07	5.24	4.80	**KDM3B-ETF1**	Hu et al, 2021
Ewsr1	Ewing sarcoma breakpoint region 1	7.91	5.09	4.89	**EWSR1-PATZ1** **EWSR1-ERG**	Chen et al, 2015, Siegfried et al, 2019, Michal et al, 2021 Zou et al, 2023
Kmt2a	lysine (K)-specific methyltransferase 2A	7.37	4.61	4.93	**KMT2A-MLLT1** **KMT2A-MLLT10**	Kim et al, 2023
Notch1	notch 1	6.92	4.20	3.87	**NOTCH1-FBXW7**	Jenkinson et al, 2013
Patz1	POZ (BTB) and AT hook containing zinc finger 1	5.54	2.99	3.95	**EWSR1-PATZ1**	Siegfried et al, 2019, Michal et al, 2021
Fbxw7	F-box and WD-40 domain protein 7	5.36	2.83	3.98	**NOTCH1-FBXW7**	Jenkinson et al, 2013
ALK	anaplastic lymphoma kinase	4.84	9.76	6.79	**CUX1-ALK**	Zhang et al, 2018
Abl1	c-abl oncogene 1, non-receptor tyrosine kinase	4.65	2.22	3.78	**ZBTB16-ABL1**	Chen et al, 2017

Table 3.1. List of genes that are bound to CAD and implicated in chromosomal translocations.

3.5.CAD Activity in Differentiating Muscle Cells Bind the Parent Genes that Give Rise to alveolar Rhabdomyosarcoma (aRMS)

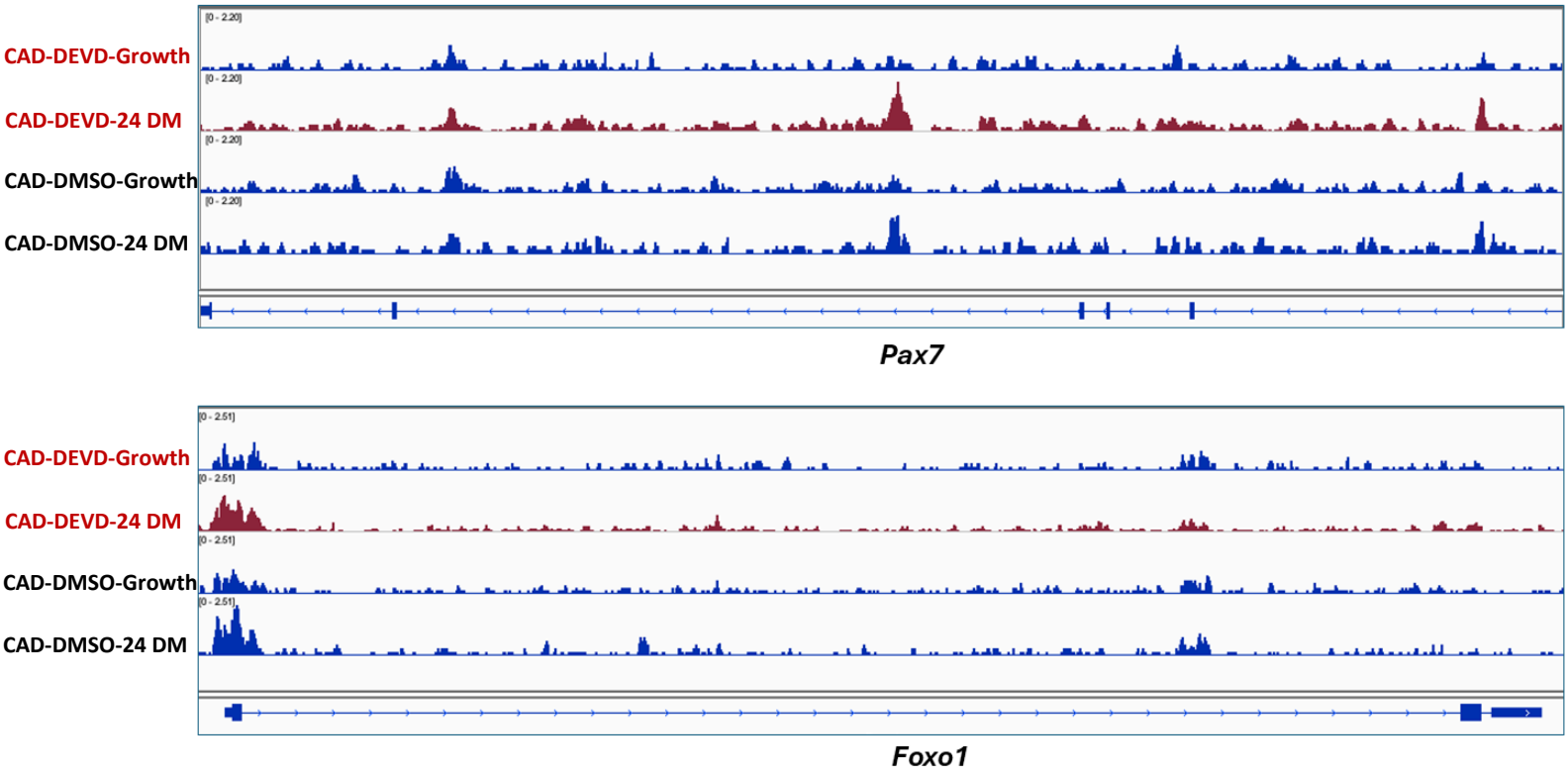
The analysis revealed CAD binding to pair of particular interest: *Pax7* and *Foxo1* gene, as shown in (Fig. 3.4A). We noted that CAD binding varied considerably at the *Pax7 Foxo1* loci between growth versus differentiation, the latter being very prominent (Fig. 3.4A). Of note, CAD was bound to the exon-intron boundary for both genes in C2C12 murine genome that is syntenic to the human *Pax7* gene intron 7 and intron 1 of *Foxo1a* (Fig. 3.4B). Thus, our data suggests that this fusion event possibly is evolutionarily conserved between murine and human genomes.

Furthermore, given that CAD associated to genes that give rise to a translocation syntenic to human genes, we speculated the detection of the translocation event and fusion transcript in murine muscle cells. Next, we extracted RNA from differentiating primary myoblasts and used primers designed for the 5' end of the *Pax7* transcript and the 3' end of the *Foxo1*. Interestingly, we were able to detect the full translocation event between Pax-7-Foxo1 fusion transcript by PCR during differentiation of primary derived skeletal muscle myoblasts at 24 hours, but not in growth (Fig. 3.4C). PCR product of the corresponding transcript band was excised and subjected to purification and sent for sequencing. As expected, we observed that the transcript had the 5' coding region of *Pax7* and 3' coding region of *Foxo1* (Fig. 3.4C). Similarly, we were able to detect the gene fusion by qPCR in differentiating primary myoblasts at 12 and increasing at 24 hours; significant increase of the transcript at 24 hours relevant to growth (Fig. 3.4D).

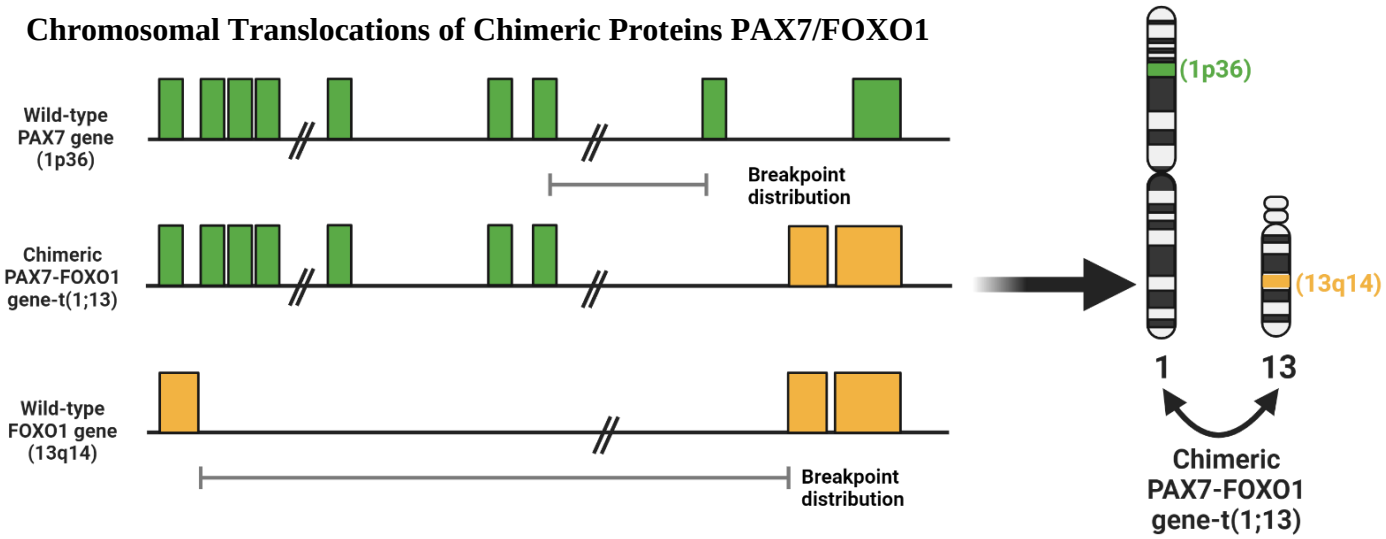
To document the fusion between *Pax7* and *Foxo1* genes, I performed fluorescence *in situ* hybridization (FISH), a standard molecular diagnostic technique used to detect chromosomal translocations by hybridizing to specific DNA sequences. The Pax7-Foxo1 fusion (translocation)

FISH probes identified two distinct (normal) *Pax7* and *Foxo1* alleles in primary myoblast nuclei during cell growth and differentiation (Fig. 3.4E). During differentiation, the signals were predominantly distinct (66%), but a small subset (34%) of cells exhibited a fused Pax7-Foxo1 chromosomal translocation event (Fig. 3.4E). To further explore these Pax7-Foxo1 fusion genes, break-apart FISH probes for *Foxo1* and *Pax7* were employed to determine if breakage events occurred, especially during differentiation. Overlapping, fused, intact (normal) FISH signals for *Foxo1* and *Pax7* were detected under both growth and differentiation conditions (Fig. 3.4F, G). However, both *Foxo1* and *Pax7* experienced breakage during primary muscle cell differentiation (Fig. 3.4F, G). Specifically, 54.5% of Foxo1 FISH signals were distinct (indicating breakage) during differentiation, compared to 45.5% of intact fused signals (Fig. 3.4F).

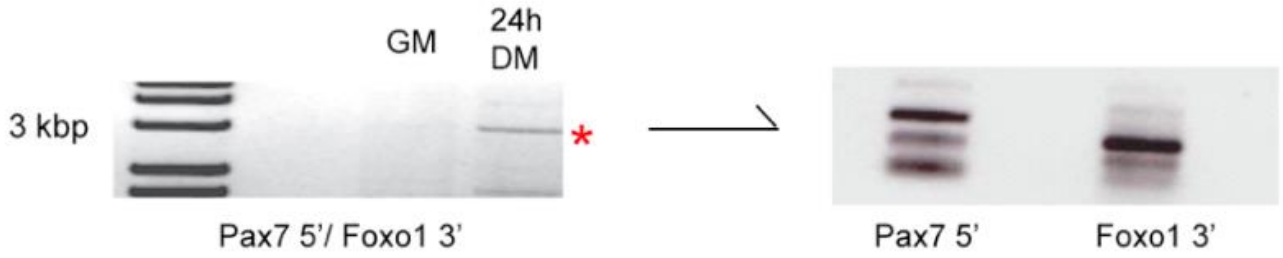
A **Enrichment of CAD at *Pax7* and *Foxo1* Genomic Loci**



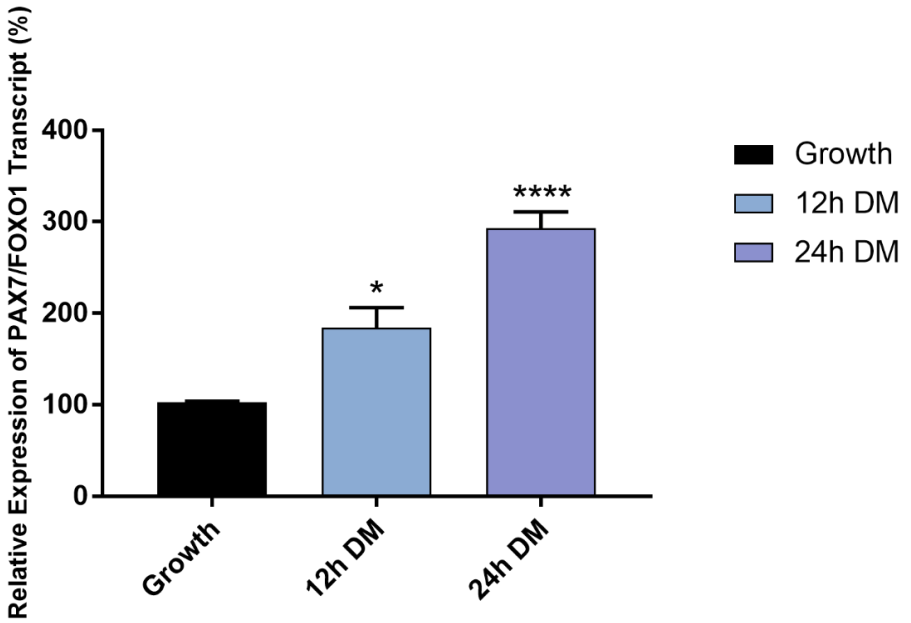
B **Chromosomal Translocations of Chimeric Proteins PAX7/FOXO1**



C **Pax7-Foxo1 Fusion Transcript in Differentiating Primary Myoblasts**

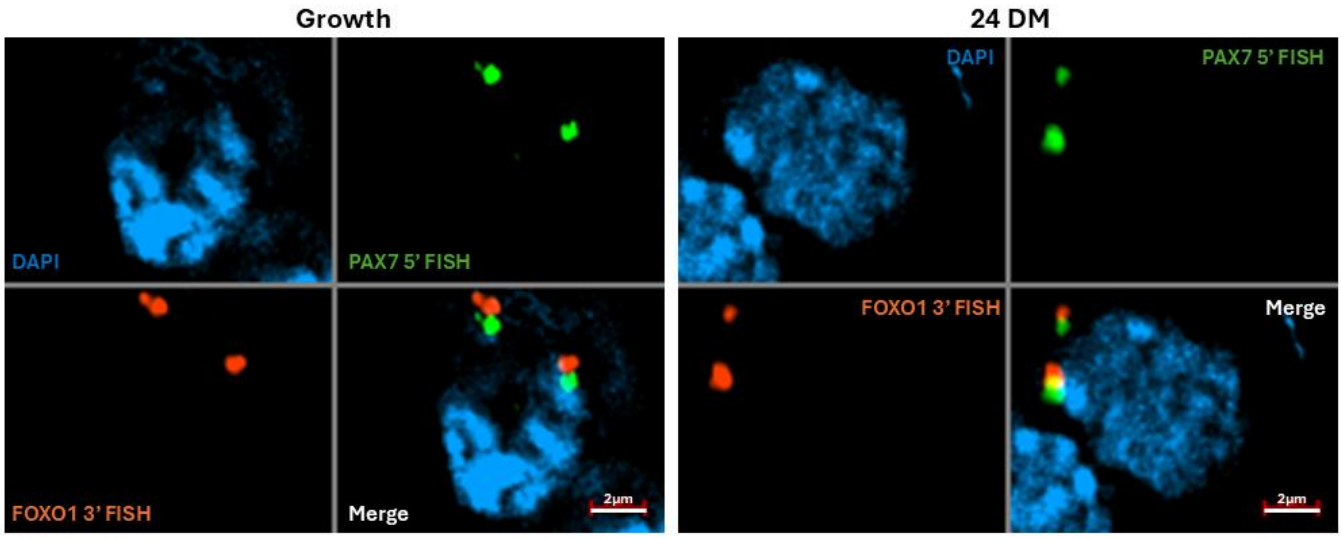


D **Quantitative PCR of Pax7-Foxo1 Gene Fusion in Differentiating Primary Myoblasts**

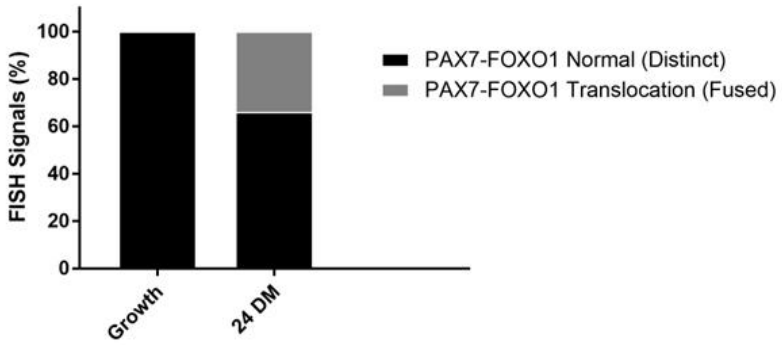


Pax7-Foxo1 FISH Probe in Differentiating Primary Myoblasts

E

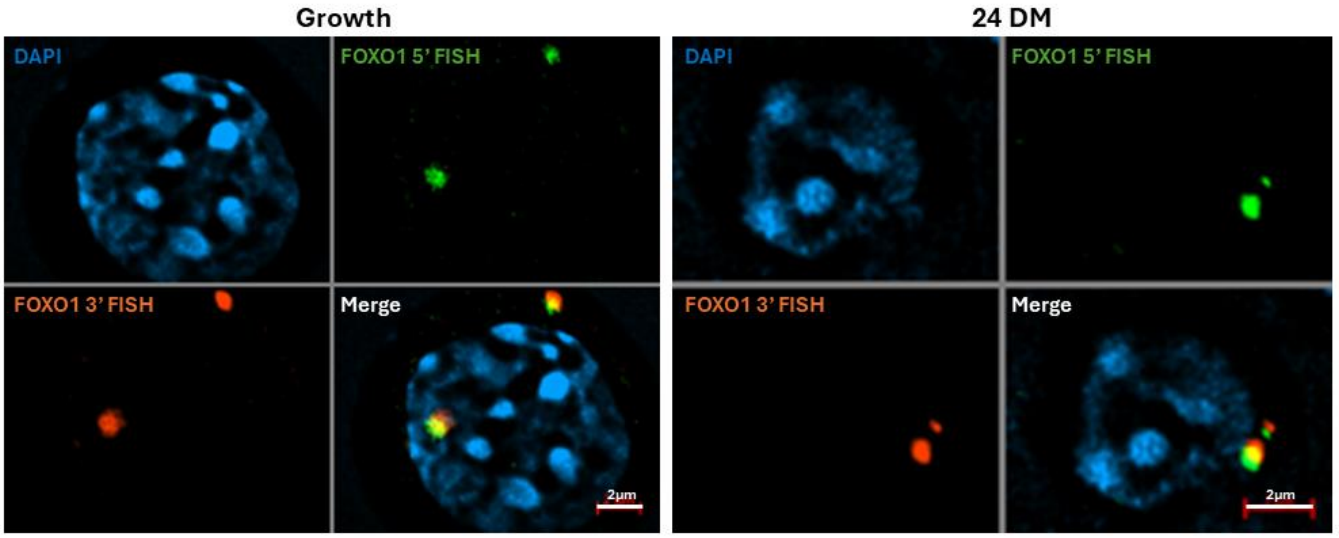


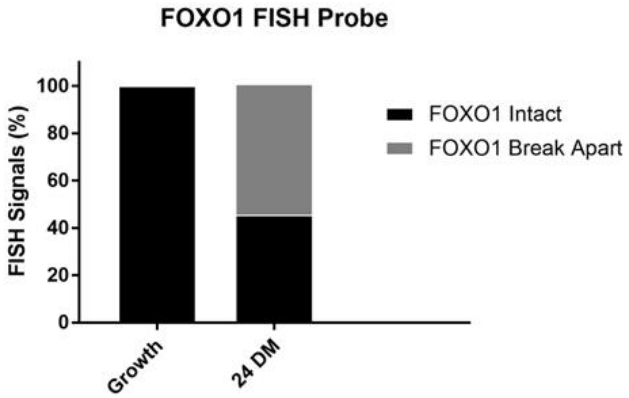
PAX7-FOXO1 FISH Probe



Foxo1 FISH Probe in Differentiating Primary Myoblasts

F





Pax7 FISH Probe in Differentiating Primary Myoblasts

G

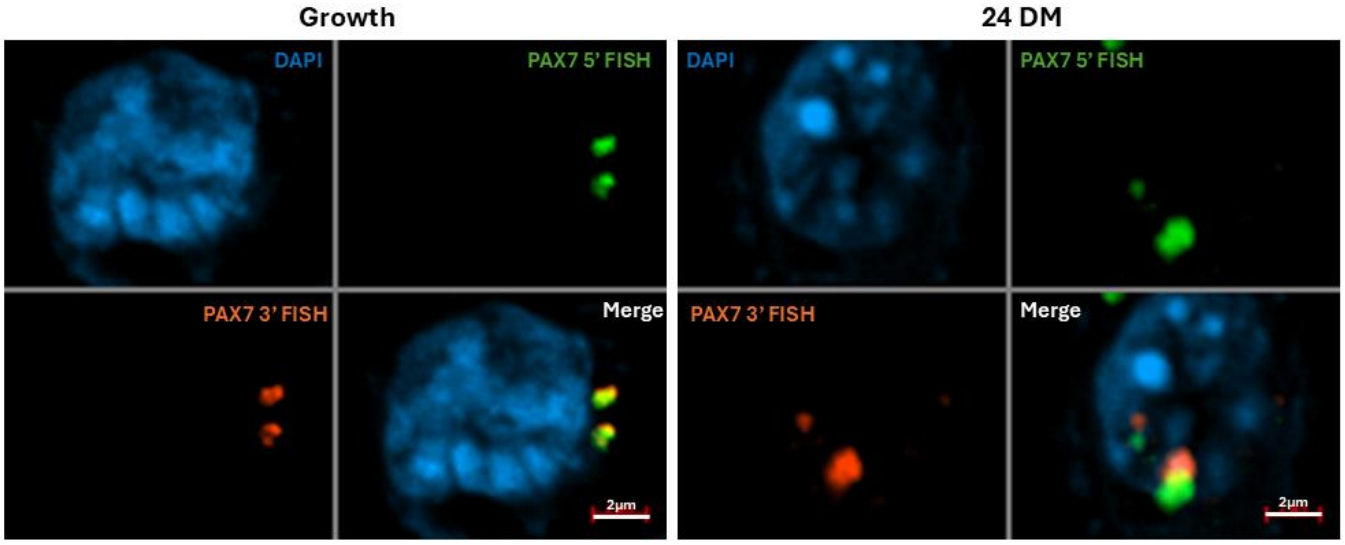


Fig. 3.4. CAD activity in differentiating muscle cells bind to *Pax7* and *Foxo1* genes.
(A) Integrative genome view representing enrichment of CAD at Pax7 and Foxo1 genomic loci. Genome tracks identified by MACS2 peak caller of CAD treated with DMSO or DEVD (growth and 24 hours post-differentiation). (B) Schematic outline depicting the chromosomal translocations producing the chimeric proteins PAX7/FOXO1 t(1;13). (C, D) Detection of Pax7-Foxo1 fusion transcript in differentiating primary myoblasts. Primers specific to the 5' region of the Pax7 transcript and the 3' region of the Foxo1 transcript were used to probe cDNA libraries of differentiating primary myoblasts for the appearance of a fusion transcript. (C) A PCR product was observed at ~3000 bp that corresponded to the expected size of a Pax7-Foxo1 transcript. Secondary PCR from the isolated 3000bp band was preformed for the 5' region of the Pax7 transcript and 3' region of the Foxo1 transcript. (N=2). (D) Quantitative PCR (N=4) of Pax7-Foxo1 gene fusion during 12- and 24 hours post-differentiation. *: $p<0.05$, ****: $p<0.0001$ when compared to growth using one-way ANOVA. (E, F, G) FISH analysis detect Pax7-Foxo1 gene fusions (E), Foxo1 (F), and PAX7 (G) in primary myoblast cells during cell growth and 24 hours post differentiation. For Pax7-Foxo1 FISH signals (E), 100 copies (n=50) of distinct FISH probes (100%) were quantified during growth condition. In differentiated condition (n=50), 66 copies of distinct FISH probes (66%) and 34 copies of fused (translocated) FISH signals (34%) were quantified. For Foxo1 FISH probe (F), 110 copies (n=55) of intact FISH probes (100%) were quantified during growth condition. In differentiated condition (n=55), 50 copies of intact FISH probes (45.5%) and 60 copies of break apart FISH signals (54.5%) were quantified.

3.6. Caspase-3/CAD Activity Regulates Pax7 Gene Expression during Skeletal Muscle Differentiation.

Pax7 maintains the identity of muscle precursor cells/satellite cells and promotes the symmetric expansion of this committed stem cell population. *Pax7* expression is rapidly decreased, which is essential for the differentiation of satellite cell derived myoblasts (Olguin & Olwin, 2004). The binding of CAD to the *Pax7* gene body suggested that the DNase may inflict strand breaks to modify the expression of Pax7. To explore this hypothesis, we compared peptide inhibition of caspase-3 to control conditions, and measured transcript levels of the *Pax7* gene by qPCR at 24 hours following low serum induction of differentiation. We observed significantly higher levels of *Pax7* transcript in the peptide inhibited primary myoblasts compared to the control DMSO treated myoblasts (Fig. 3.5A), implying a caspase-3/CAD regulation of the *Pax7* gene.

To further investigate whether the reduction of pax7 gene expression is directly related to CAD *per se*, we stably expressed a mutant version of ICAD constructs in primary myoblast cells. The reasoning for this approach was that previous work demonstrated expression of ICAD with a mutation in the caspase-3 cleavage site (mtI-CAD-D117E) leads to a caspase resistant ICAD protein, which has a stabilized interaction with CAD, leading to impaired nuclease activity (Larsen et al, 2010). Caspase-3 cleaves ICAD at both D117 and D224 aspartic acid residues, yet only the D117 cleavage mutant infers inhibition of CAD function (Larsen et al, 2010).

To validate the expression of ICAD constructs, immunoblotting analysis of primary myoblast cells transfected with wild-type ICAD, mt-ICAD-D117E, or mt-ICAD-D224E were performed, and result showed normal level of ICAD (Fig. 3.5B). Myoblasts expressing either wild type-ICAD, mt-ICAD-D224E or empty vector, showed normal differentiation, as demonstrated by the

expression of myosin heavy chain and multi nucleated myotubes (Fig. 3.5C). Alternatively, expression of mt-ICAD-D117E hindered differentiation as shown in (Fig. 3.5C). Next, we tested whether mt-ICAD would alter *Pax7* gene expression under the same conditions. During differentiation, *Pax7* expression was reduced in cells transfected with the empty vector, as well as in those expressing wild-type ICAD and mt-ICAD-D224E (Fig. 3.5D). However, expression of the mutant mt-ICAD-D117E had a more pronounced effect and maintained Pax7 expression (Fig. 3.5D), suggesting active CAD is required to bind and repress Pax7 gene expression, during differentiation.

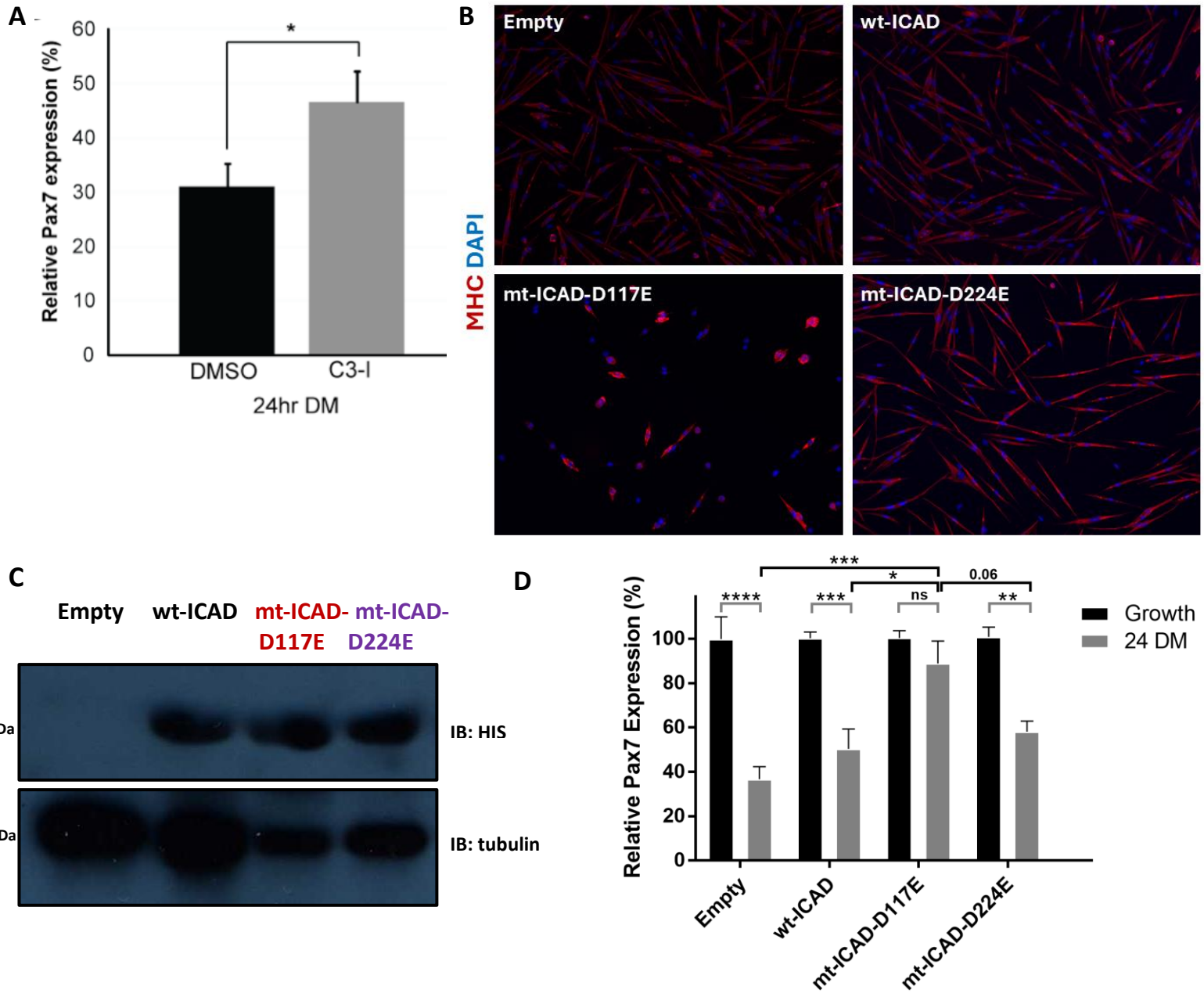


Fig. 3.5. Caspase-3/CAD activity regulates *Pax7* gene expression during skeletal muscle differentiation.
(A) Caspase-3 inhibition (C3-I) impairs Pax7 repression during differentiation. Primary myoblasts were induced to differentiate with either 15 μM of the caspase 3 inhibitor, z-DEVD- fmk, or the vehicle only control DMSO. RNA was collected 24 hours after the induction of differentiation and representative cDNA libraries were generated. Relative Pax7 expression at 24 hours was determined by qPCR and normalized to growth conditions. N=4 in triplicate. (Error bars are s.e.m.; P=0.0076) (B) Primary myoblasts expressing wt-hICAD, mt-hICAD-D117E, and mt-hICAD D224E were induced to differentiate under low serum conditions for 72 hours, fixed with 90% MeOH/10% PBS, and myosin heavy chain (MHC, red) was stained with mouse anti-MHC (MF20, 1:200, Developmental Hybridoma) followed by goat anti-mouse AlexaFluor 598 (1:2500, Molecular Probes), nuclei were counterstained with DAPI (blue). (C) His-tagged ICAD constructs are expressed in primary myoblasts. Protein lysates were collected from selected myoblast and subject to western blot analysis. His-tagged ICAD constructs were detected with a mouse anti-His (1:1000, Cell Signaling) primary antibody and subsequently a goat anti-mouse HRP secondary anti body. Tubulin was used to control for loading. (D) Stable expression of ICAD constructs in primary myoblast influences Pax7 depression following the induction of differentiation. Relative expression of Pax7 at 24 hours following the induction of differentiation was determined and normalized to the Pax7 expression in each individual cell line at growth. N=4 in triplicate. *: p<0.05, **: p<0.01, ***: p<0.001****: p<0.0001 using two-way ANOVA.

3.7.CAD Binding Cancer Breakpoint Genes in T Progenitors Cells

The majority of the genes identified above (section 3.4) are considered to be inaccessible in muscle cells, yet our observations indicate robust CAD binding on these gene loci. This observation suggests that CAD binding to these regions may occur in other cell lineages, i.e. CAD may retain a somewhat conserved function for initiating DNA strand breaks and gene fusions in other cell lineages. Thus, we speculate that CAD binding to sites that are prone to translocation events, might be a conserved biological mechanism. This is paralleled by the fact that caspases mediated cell differentiation is a conserved mechanism across different cell lineages (Burgon & Megeney, 2018; Bell & Megeney, 2017).

Most relevant to our observations, transient caspase activation is required to induce differentiation of several myeloid lineages, suggesting that caspase triggered CAD activity may be responsible for generating a portion (or all) of the leukemia and lymphoma oncogenic translocation genes. For example, caspase-3 activity drives the differentiation of erythroid cells (Carlile et al, 2004; Zermati et al, 2001), formation of proplatelets (Kaluzhny et al, 2001; De Botton et al, 2002), and differentiation of monocytes to the macrophage lineage (Sordet et al, 2002; Cathelin et al, 2006; reviewed in Solier et al, 2017). Caspase-3 is also essential during thymic development for the elimination of autoreactive lymphocytes, T cell proliferation (Kennedy et al, 1999), and in differentiation of mature T lymphocytes (Green et al, 2003; Krammer et al, 2007). Caspase-3 activation has also been shown to regulate early activation steps in nonapoptotic developing T and B cell lymphocytes (Chun et al. 2002; Misra et al. 2005; McComb et al. 2010).

Based on the collective strength of these observations, we reasoned that T cells may also retain a CAD binding activity during early stages of T cell differentiation and that this activity may target

translocation prone genes similar to what we have observed in differentiating muscle cells. To explore the hypothesis, I conducted CAD CUT and Tag experiments from primary derived early differentiating T cells that were extracted from murine thymus of 6-week-old wild type mice (Fig. 3.6A, B). To maximize the probability of generating an early progenitor pool, we purified primary derived thymic progenitors of T cells to double negative (DN) cells $CD4^{-}CD^{-}8$. Cells were subjected to FACS based purification using the marker selection as indicated in the legends to confirm that 70% to 80% of cell population were DN (early progenitors) and depleted of double positive (DP) cells (Fig. 3.6B). This DN selected/purified T cell pool was then used to perform CUT and Tag, using the same antibody directed against endogenous CAD (as detailed in section. 3.1). Sequenced libraries of CUT and Tag bound CAD in early differentiating T cell progenitors revealed that a majority of the total CAD binding sites reside in promoter regions, 42.89 %, followed by distal intergenic regions at 27.72 %, and 21.64 introns (Fig. 3.6C). This CAD bound genome distribution is similar to CAD binding in differentiating C2C12 muscle cells (Fig. 3.1D), suggesting a conserved CAD binding preference, at a gross genome level, between these distinct cell lineages. Cross comparison of the CAD bound genome against the human breakpoint atlas revealed that CAD bound to a large cohort of the parent translocation genes also bound in C2C12 muscle cells (85%) (Fig. 3.3B), with very few CAD binding events that were deemed to be lineage specific. However, the exact overlap in binding between the human breakpoints and CAD was higher in differentiating muscle cells versus T cells (180 sites versus 65 sites respectively), suggesting more active breakpoint targeting in skeletal muscle progenitor cells (Fig. 3.3C, Fig. 3.6D).

T cell acute lymphoblastic leukemia (T-ALL) is a heterogenous malignancy that originates from early T cell progenitors and is characterized by inflation of genomic lesions that influence the

development of T cells (De Bie et al, 2023). Group of genes display aberrant expression in pediatric and adult T-ALL, including T cell acute lymphocytic leukaemia 1 (*Tal1*), lymphoblastic leukaemia associated haematopoiesis regulator 1 (*Lyl1*), LIM domain only 1/2(*Lmo1, Lmo2*), NK2 homeo-box 1/2 (*Nkx2-1*) (*Nkx2-2*), and homeobox A (HOXA) genes (Dai et al, 2022; De Bie et al, 2023). We observed CAD binding to *Lyl1, Nkx2-2, Nkx2-1, Lmo1, Lmo2, Tal1, Hoxa10* genes in C2C12 cells (Fig. 3.3C). In T cells, CAD was also binding to *Tal1, Lyl1, Hoax10, Nkx1-1, Nkx1-2, Kmt2a* genes in T cells (Fig. 3.6D). Indeed, T-ALL is a dynamic area of research in the last decade. The WHO classified T-ALL subtypes in 2016, yet with the advances of sequencing and profiling of gene expression from T-ALL cases, this in turn enabled the expansion of T-ALL subgroups with further classification came in 2022 (Alaggio et al, 2022; De Bie et al, 2023).

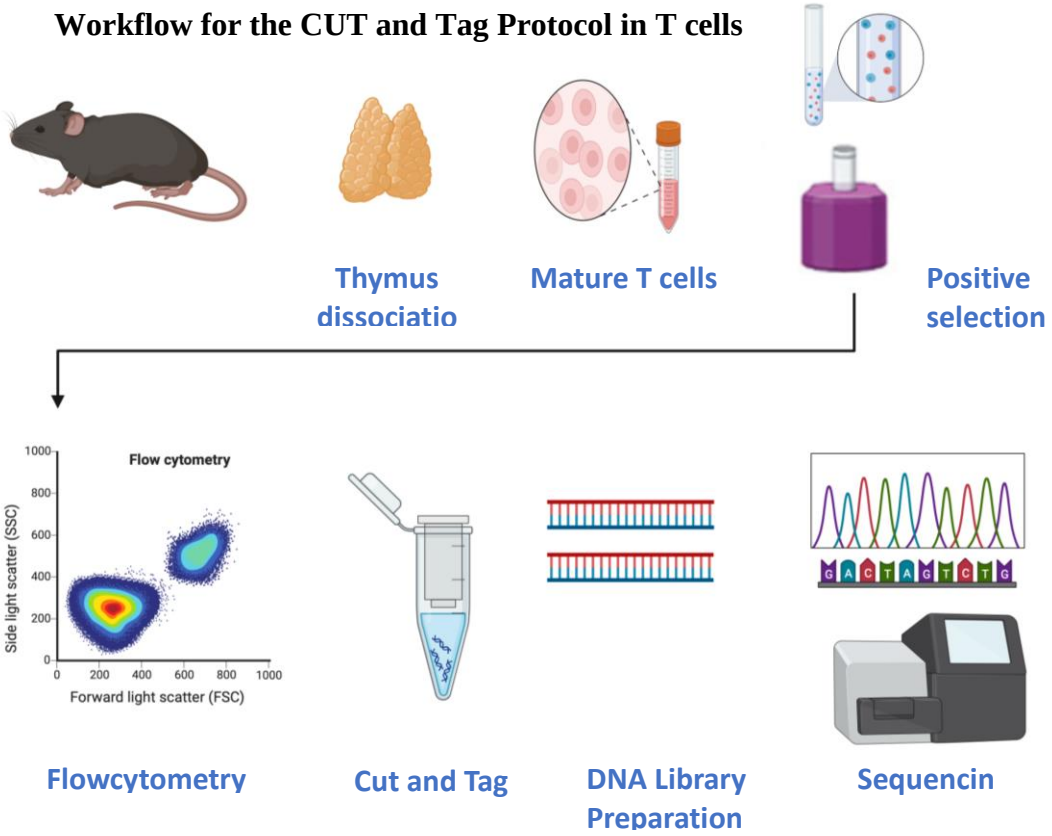
Next, I examined the degree of overlap in CAD bound translocation prone genes between early differentiating T cell progenitors and differentiating C2C12 muscle cells. This analysis revealed substantial overlap for CAD binding to *Tlx3*, *Tlx1*, *Mllt10*, *Lmo2, Myc, Rara, Runx3, Btg3, and Btg4* genes in both cell types (binding tracks shown in Fig. 3.6E, Circablot Fig. 3.6F). While this is not an identical overlap for CAD binding to translocation sensitive genes, the data suggest that several key translocation loci that produce fusion genes and similarly targeted by CAD binding.

To explore the similarities and differences for CAD binding preferences across the two cell lineages, I examined the distribution of binding widths (violin plot) for unique C2C12 muscle cell sites, unique T cells sites, and sites shared by both C2C12 and T cells (Fig. 3.6G). There are ~8 times more unique C2C12 targets than T cells targets. About 58% of T cells targets are differential sites from C2C12, whereas 92% of C2C12 targets are distinct sites from the CAD target regions in T cells (Fig. 3.6G).

To further delineate and characterize the gene distribution for CAD in differentiating T cells, we applied GO analysis (Fig. 3.6H). Application of these analytic tools revealed CAD enrichment to be associated with genes that regulate RNA polymerase II transcription binding, transcription regulation, lymphocyte differentiation, neuron differentiation, and cell fate commitment. This gene ontology analysis revealed functional overlap with the gene ontology analysis for CAD bound targets in differentiating C2C12 muscle cells, suggesting a shared or overlapping biology of CAD targets despite a much larger cohort of CAD bound targets that appear to be unique to the differentiating muscle cell lineage.

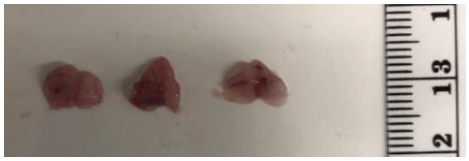
A

Workflow for the CUT and Tag Protocol in T cells



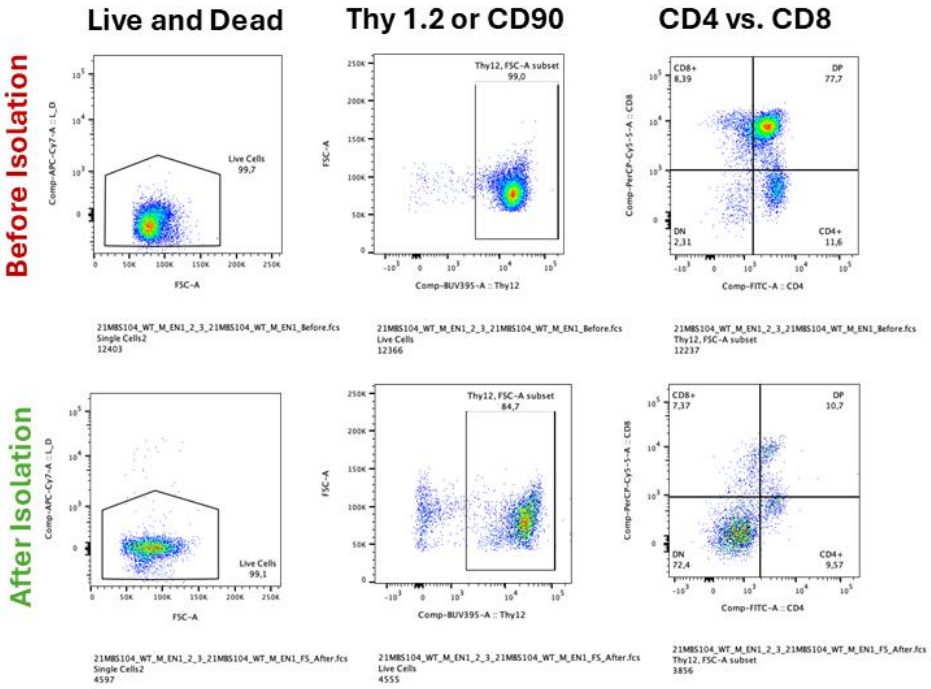
B

Extracted Thymus

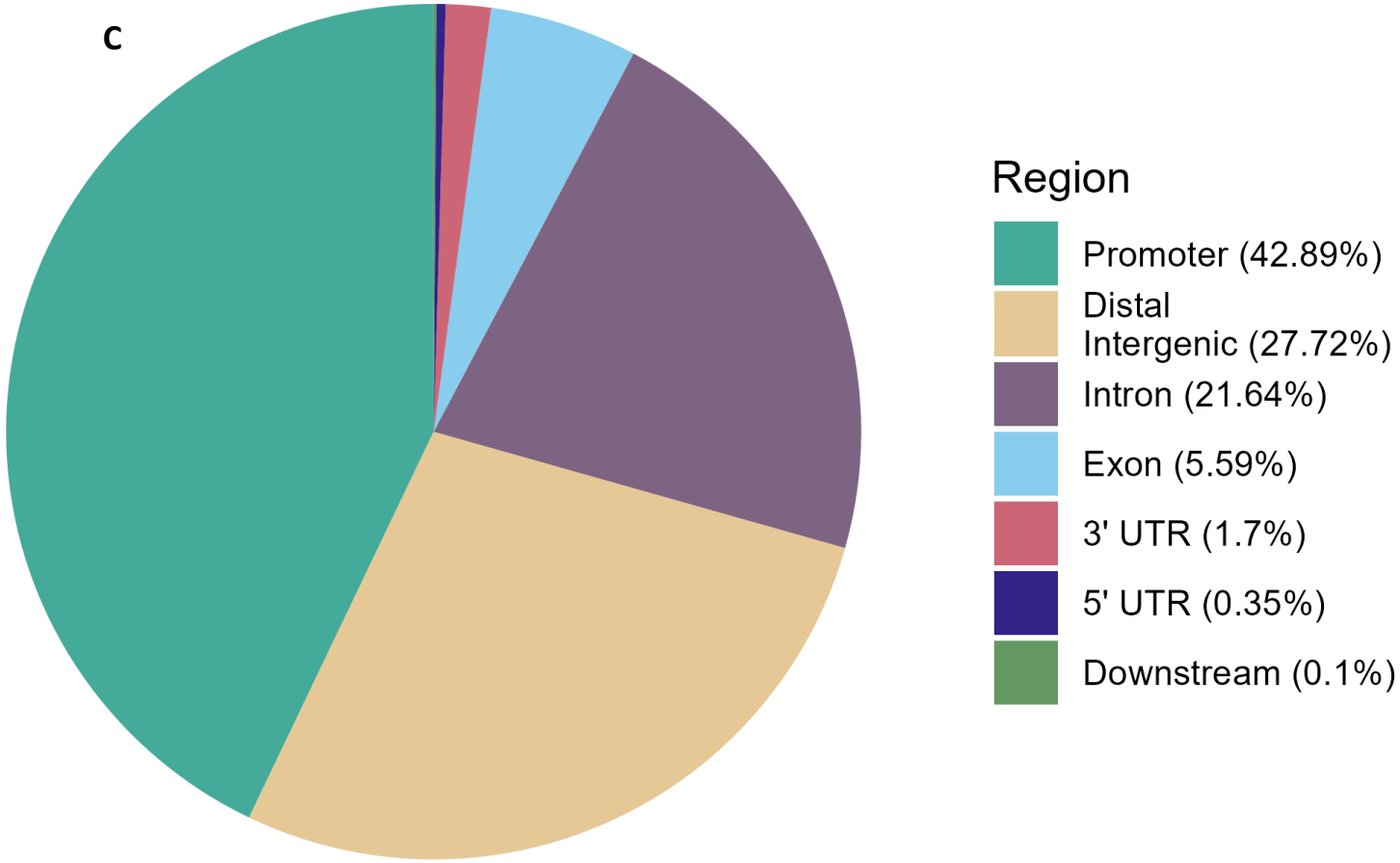


WT1 WT2 WT3

FACS Purification of T Cells (DN Stage)

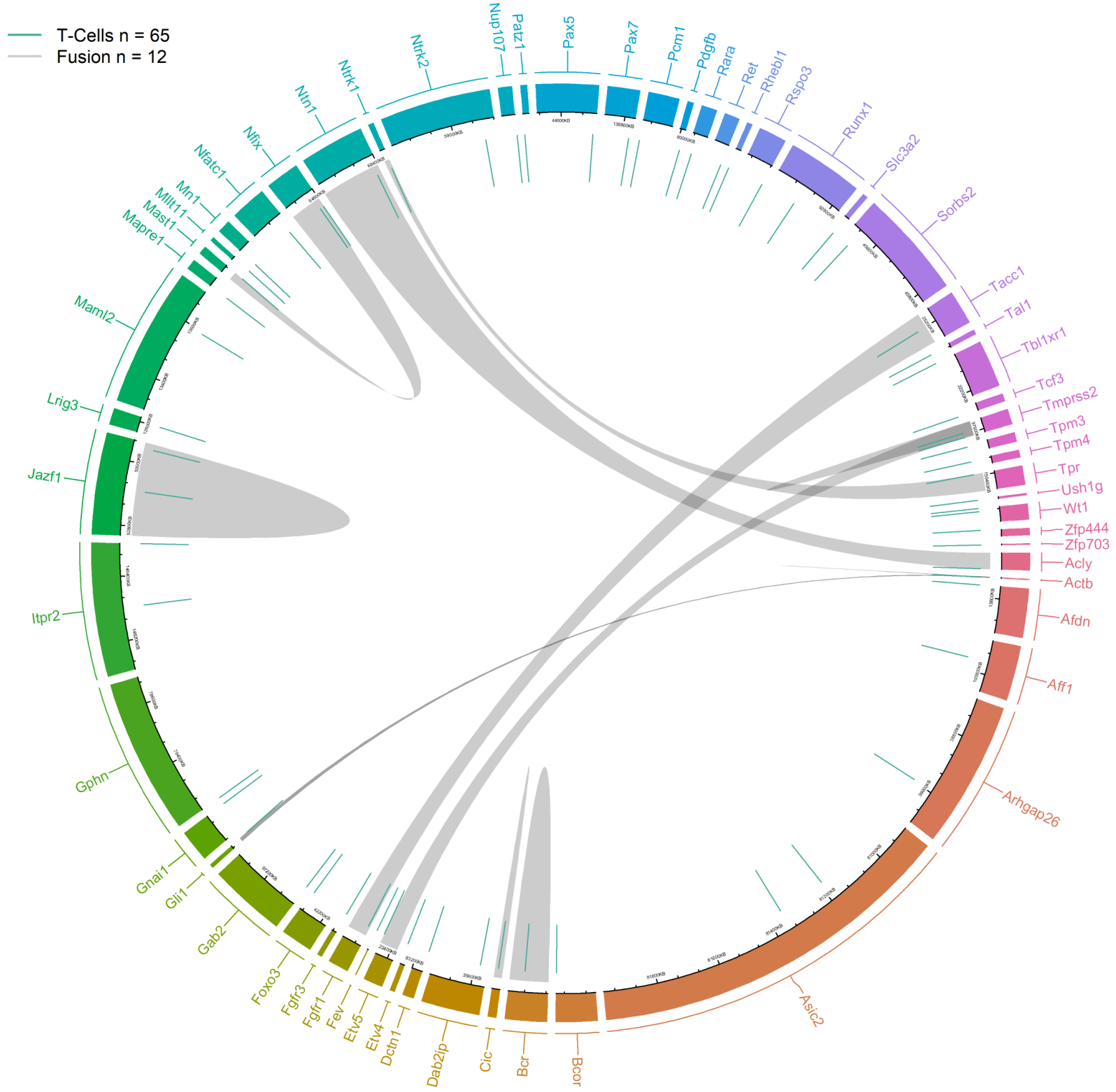


Genomic Distribution of CAD Binding Sites in T Cells



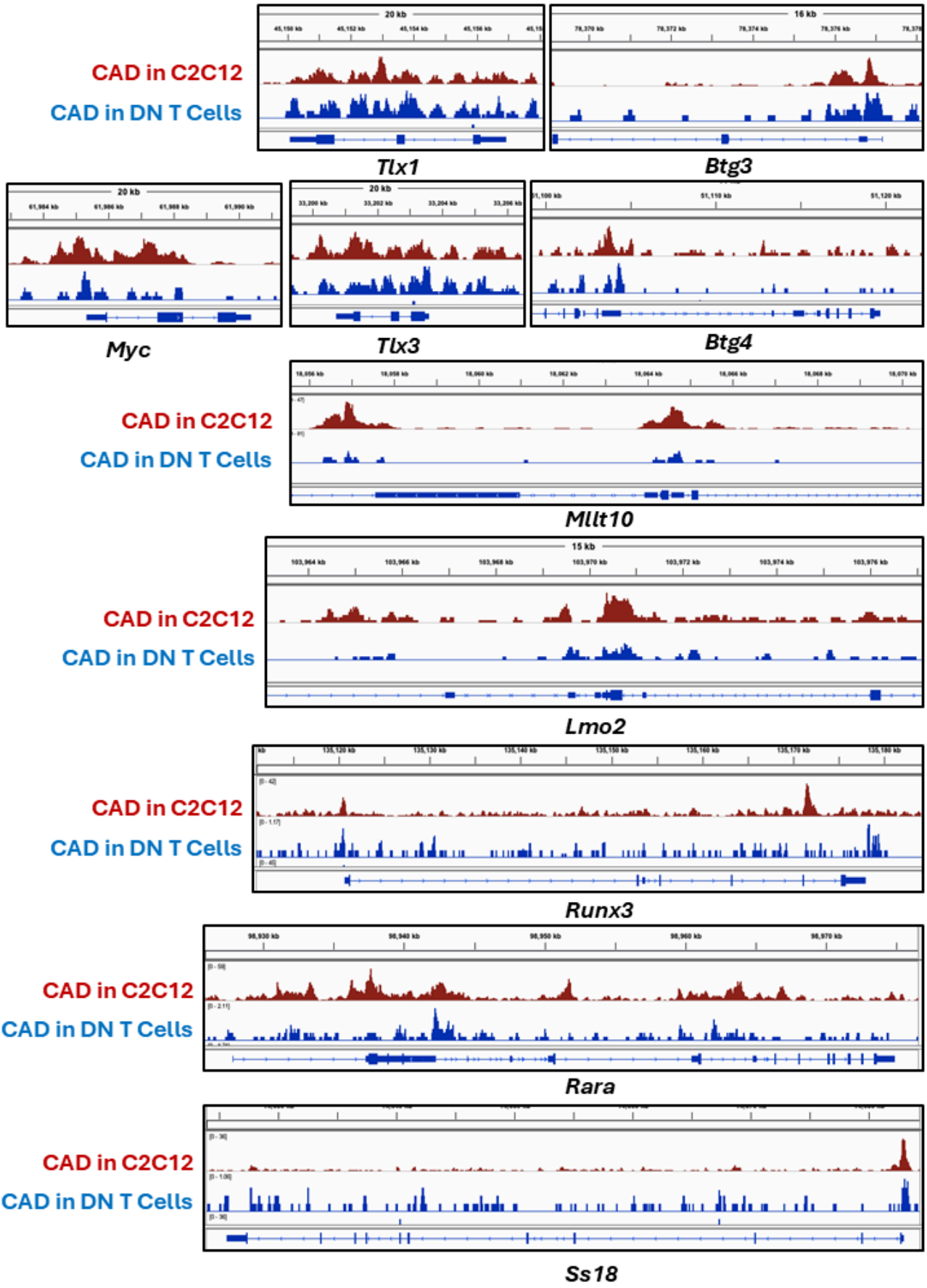
D

CAD Binding Sites Overlapping between T cells and COSMIC Breakpoints in 59 COSMIC Fusion Genes



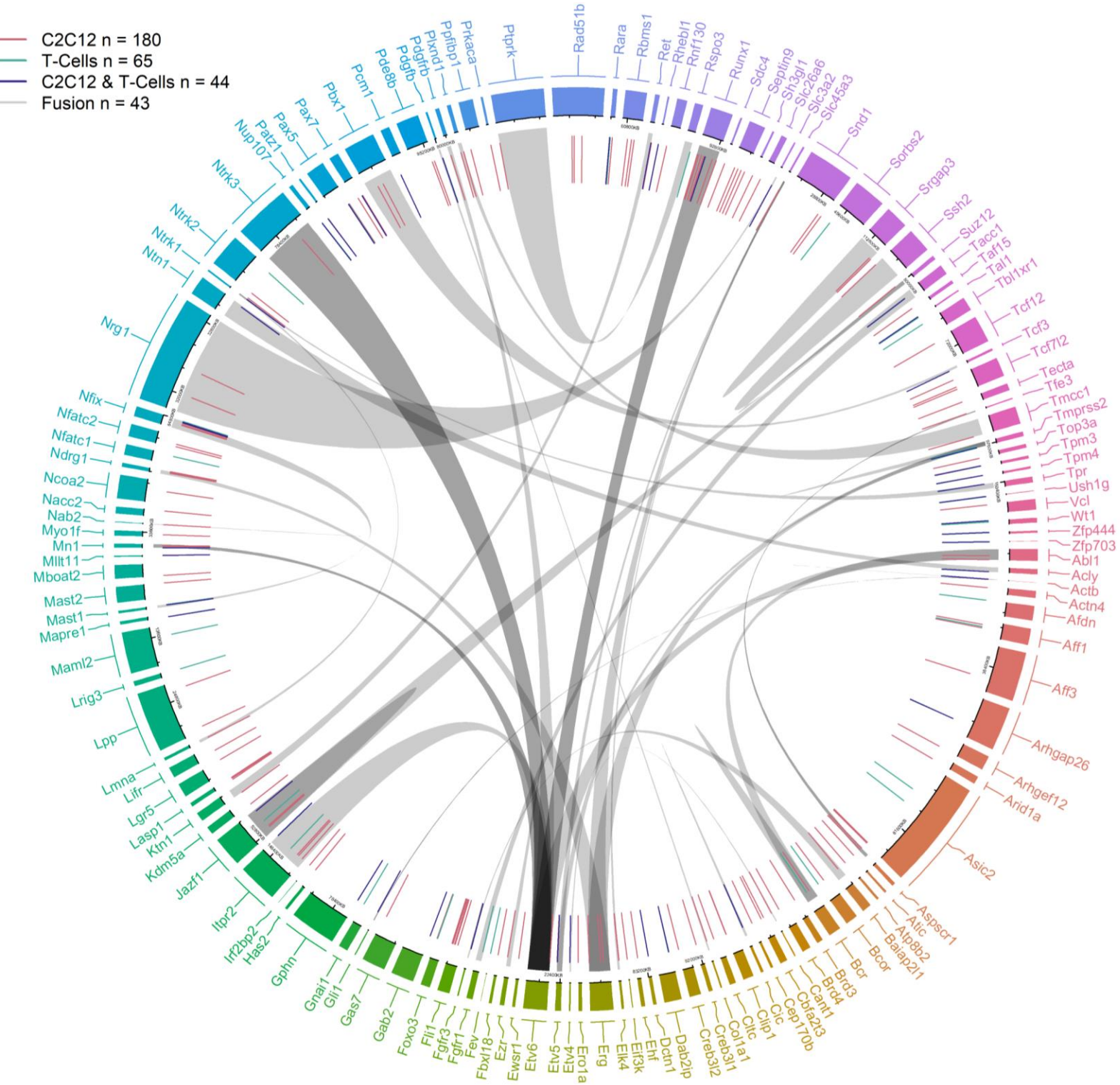
E

CAD Bind Fusion Genes in C2C12 and DN T Cells

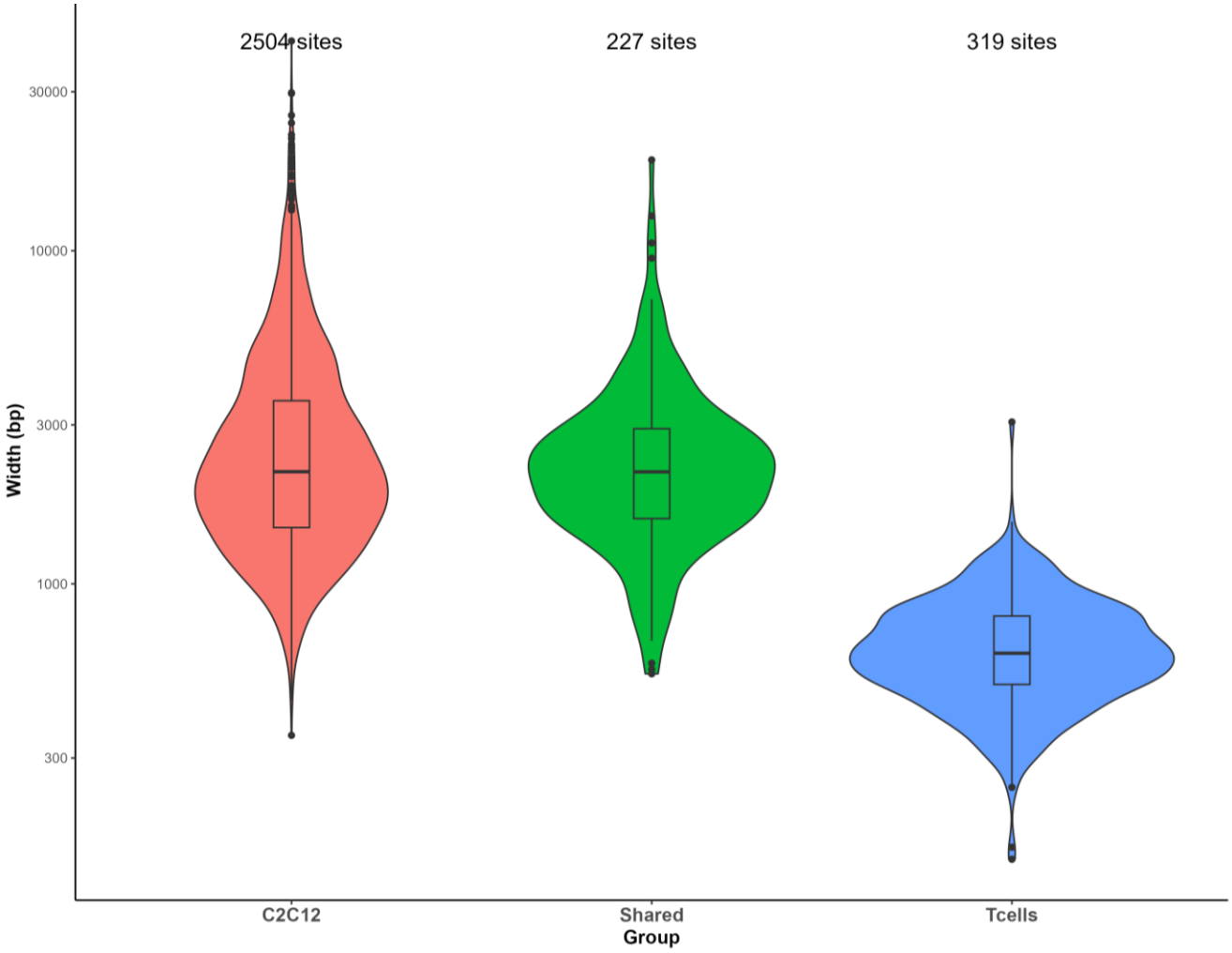


F

CAD Binding Sites Overlapping between C2C12, T Cells, and COSMIC Breakpoints in 135 COSMIC Fusion Genes



G **Distribution of CAD Binding Widths for C2C12 and T Cells**



H

Gene Ontology (GO) Analysis Identified within CAD Functional Genes in Differentiated T Cells

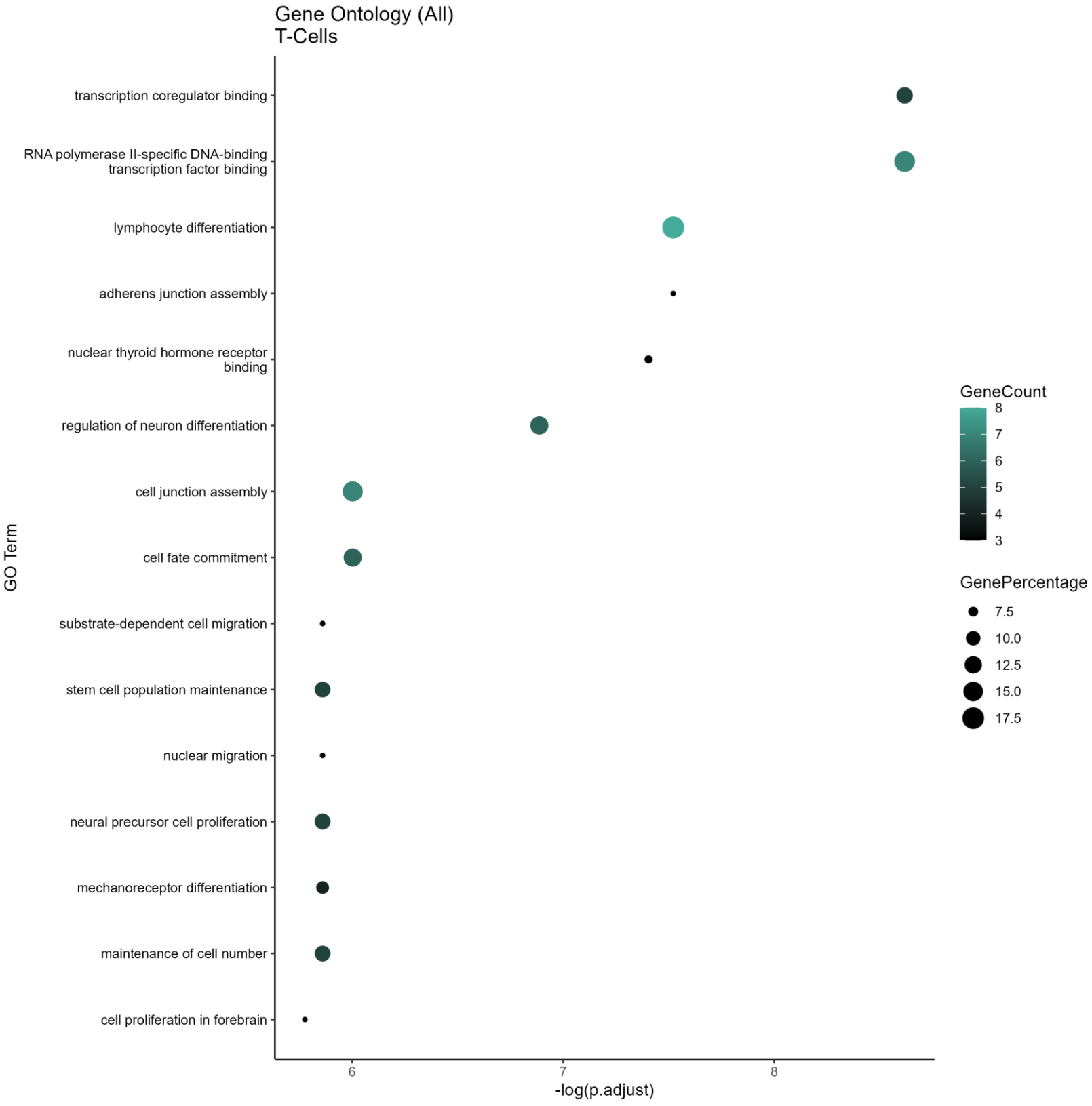


Fig. 3.6. Genomic distribution of CAD binding sites, from T progenitor cells.
(A) Schematic representation of the workflow for the CUT and Tag protocol applied to T cells. (B) Thymus from 6-week-old WT male mice was processed to extract T cells, followed by purification using FACS. (C) Pie chart illustrating the genomic distribution of CAD binding sites in T cells, showing the percentage for each genomic location category. N=3. (D) Circa plot visualizing the cross-comparison of CAD-binding sites between T cells and the COSMIC fusion gene set. (E) Integrative Genomics Viewer (IGV) snapshots depicting CAD binding to fusion genes in C2C12 and DN T cells. (F) Circa plot visualizing the cross-comparison of CAD-binding sites between C2C12, T cells, and the COSMIC fusion gene set. (G) Violin plot showing sites compiled through a 3-peak-caller-consensus and a 2-replicate-consensus peak set. Differential and shared sites were determined by counting overlapping sites between groups. (H) Gene Ontology (GO) analysis identifying the gene distribution for CAD in differentiating T cells, with the top 15 enriched terms.

Summary

In this chapter, we investigated the genome-wide mapping of CAD targets in skeletal muscle cells. We demonstrated that CAD binding to *Pax7* gene by inducing a physiologic DNA strand break that is associated with decreased *Pax7* expression. Our data suggest that this strand break i) is a developmental prerequisite to disrupt the transcription read through allowing the differentiation program to proceed, ii) the regulation of *Pax7* gene by CAD is a molecular mechanism that implicates CAD in the developmental origin of aRMS in differentiating cells. Expanding on these observations, we identified CAD binding several genes in differentiating myoblasts that are implicated in breakpoint translocations (e.g., leukemias and lymphomas), findings that also extended to T cell progenitors. Therefore, CAD biology in these breakpoints might be conserved phenomenon across cell lineages and may represent the point of the origin for breakpoints, thus providing new insights in the context of breakpoint translocations.

CHAPTER 4: RESULTS

Caspase-3 Targets the Epigenetic Regulatory Protein Ezh2 During Myogenesis

Caspase-3 Targets the Epigenetic Regulatory Protein Ezh2 During Myogenesis

Abstract

Epigenetic modifiers control cell fate and gene expression through post-translational modifications (PTMs) of histone proteins. A key epigenetic modifier is Enhancer of zeste homolog 2 (Ezh2), a component of the polycomb repressive complex 2 (PRC2), which methylates lysine 27 (H3k27me3) on histones to repress gene expression. Ezh2 is highly expressed in proliferative myoblasts and diminishes during differentiation, correlating with muscle gene activation and repression of proliferation genes. The reduction of Ezh2 activity is critical for myogenesis, yet the mechanisms controlling Ezh2 activity and its genomic targets during differentiation remain poorly understood. Our results in the previous chapter suggest that specific regions in the genome are targeted by CAD inducing DNA breaks during differentiation. Given that CAD has no defined DNA binding motif, it suggests that its access and/or activity on DNA is controlled by other features in the local chromatin environment. I hypothesized that the Ezh2/PRC2 complex might align with CAD genomic targets, acting as a methylated repressive signal to restrain CAD function, and thereby limit cell differentiation. CUT and Tag analysis revealed a significant overlap between Ezh2 and CAD targets in C2C12 muscle cells, with 90.5% of CAD regions shared with Ezh2 and 67.9% of Ezh2 binding sites aligning with CAD sites after 24 hours of differentiation. Additionally, Ezh2 contain conserved caspase 3 cleavage sites and herein I found that Ezh2 is cleaved by active caspase-3 during differentiation, leading to reduced Ezh2 protein levels and H3K27 methylation signals. Collectively, these observations suggest that Ezh2 target CAD genomic loci to limit its capacity to inadvertently engage gene expression and provide new insights into Ezh2 protein stability and the caspase signaling that coordinate the differentiation of skeletal muscle cells.

Background

Epigenetic signaling control cell fate and gene expression programs. It involves PMTs of histone proteins, which can profoundly shape local chromatin architecture to either activate or repress expression of target genes. While this regulatory mechanism primarily affects transcription factor accessibility, the local chromatin structure also influences DNase targeting, which affect gene expression without directly altering the core DNA sequence; a regulation known as epigenetic modification.

Ploycmob Group (PcG) proteins form chromatin remodeling complexes known as Polycomb Repressor Complexes 1 (PRC1) and PRC2. Structurally, PRC1 is comprised of BMI1, RING1A, and RING1B, which repress gene expression by modifying histone H2Aubiqitin (H2Aub) of Lys119 (Richly et al, 2011). PRC2 include core components of Ezh1/2, suppressor of zeste12 (SUZ12), and embryonic ectoderm development (EED) proteins (Margueron et al, 2008).

PRC1 and PRC2 are further categorized into canonical and non-canonical complexes, each with distinct pathways and cellular contexts that influence protein recruitment (Yu et al,2019). The canonical PRC2 complex is targeted to chromatin where Ezh1/2, via its conserved SET domain, catalyzes the mono-, di- and tri-methylation of Lys27 of H3 (H3K27me) (Yu et al, 2019). This methylation marks of H3k27me3 facilitates the binding of PRC1, which promotes chromatin compaction to repress gene expression (Cao et al, 2002; Ringrose et al, 2004). Additionally, JARIDII, a PRC2 protein component, binds to monoubiquitinated lysine 119 of histone H2A (H2AK119ub), facilitating cross-talk between PRC1 and PRC2 and maintain stable polycomb domains (Cooper et al,2016).

EZH1 and EZH2 are mammalian paralogs of the ancestral E(z) gene found in *Drosophila* (Whitcomb et al, 2007; Lee et al, 2022). Studies indicate that Ezh2, on its own, is catalytically inactive and requires interaction with the core PRC2 subunits, SUZ12 and EED, to elicit enzymatic targeting of K27 (Cooper et al, 2016; Holoch et al 2017). Structural studies have highlighted the importance of the interaction between Ezh2, EED, and the VESF domain of SUZ12 in maintaining the stability and proper positioning of the active SET domain (Justin et al, 2016). EED directly binds to trimethyl-lysine residues, forming the Ezh2-EED complex, which leads to the allosteric activation of the histone methyltransferase (HMT) activity of PRC2 (Margueron et al, 2009). Additionally, SUZ12 is essential for maintaining the HMT activity and the silencing function of the EED-EZH2 complex (Cao et al, 2004).

Interestingly, PRC2-Ezh1 and PRC2-Ezh2 complexes can independently form and target similar gene sets. In the skeletal muscle cell lineage, Ezh2 is more abundant in proliferative myoblasts than in differentiated myotubes, and its down-regulation correlates with activation of differentiation and muscle gene expression (Caretti et al. 2004; Dudakovic et al, 2015). Ezh2 has higher levels of histone lysine methyltransferase activity, and its inhibition affects global methylation levels, unlike Ezh1 repression, which does not affect global methylation in differentiating cell lineages (Margueron et al, 2008). PRC2-Ezh2 is responsible for establishes the repression state of chromatin, whereas PRC2-Ezh1 sustains this repressive state in the absence of Ezh2 to repress transcription and compact chromatin (Yu et al, 2019; Margueron et al, 2008). Therefore, this coordinated switch from PRC2-Ezh2 to PRC2-Ezh1 is essential for the successful transition from for myoblast to myotube transition (Stojic et al, 2011).

Furthermore, the activity of the Ezh2 protein itself is significantly influenced by PTMs. For instance, the degradation of Ezh2 is enhanced by phosphorylation at threonine 372, which is

targeted by p38α kinase and is required for muscle cell differentiation (Consalvi et al, 2017). This phosphorylation-driven degradation of Ezh2 during muscle cell differentiation is mediated by the Praja1 E3 ubiquitin ligase, which labels Ezh2 fragments for removal by the proteosome (Consalvi et al, 2017). Additionally, a shift from lysine methylation to phosphorylation also modulates Ezh2 protein degradation during hematopoiesis (Guo et al, 2024). Despite the critical role of Ezh2-PTMs in these processes, the specific protease mediating the initiation of Ezh2 degradation remains unknown.

Caspase proteolytic activity significantly impact the structural features within the cell and may function as a genome reorganizer. Research from Dr. Megeney's laboratory has demonstrated that the executioner caspase-7 is active during skeletal muscle differentiation and targets chromatin-associated proteins such as SATB2 (Bell et al, 2022). The removal of SATB2 leads to alterations in topologically associating domains (TADs) and chromatin loops, which in turn suppresses the expression of a set of chromatin genes crucial for driving the muscle cell differentiation program (Bell et al, 2022). Furthermore, Ring1B, a component of the E3 ubiquitin ligase complex responsible for histone H2A ubiquitination and transcriptional repression, has been identified as a direct substrate of caspase-3 (Wong et al., 2007). These findings suggest a broader role for caspases in modulating chromatin structure and gene expression during cellular differentiation processes.

Given these observations, it is reasonable to hypothesize that the PRC2 complex, though Ezh2 activity, may antagonize the pro-differentiation capacity of other genome modifying enzymes that promote skeletal muscle cell differentiation. Specifically, Ezh2 could inhibit the differentiation process by modulating CAD genome targeting and activity, via the suppressive trimethylation of H3K27me3. Considering that Ezh2 undergoes regulatory degradation during muscle cell

differentiation, I further explored whether caspase-3 directly cleaves Ezh2 protein, to delimit PRC2 capacity to target CAD binding and/or activity within the genome.

Additionally, my results from the previous chapter indicate that CAD targets specific genomic regions to induce DNA breaks during differentiation. Given that CAD has no defined DNA binding motif, it suggests that its access and/or activity on DNA is controlled by other features. Indeed, once CAD is active, it's DNase should be restrained sufficiently to access CAD sites, yet not to fracture the genome. Interestingly, research has shown that CAD-induced DNA damage sites may be restricted by protein complexes such as PRC1 and PRC2. For example, Ezh2, creates a repressive environment through the trimethylation of H3k27me3, thereby blocking differentiation (Caretti et al, 2004; Dudakovic et al, 2015).

Given this context, I sought to investigate whether CAD's genomic targets overlap with Ezh2. To address this, I performed "CUT and Tag" on 150,000 fresh intact C2C12 cells utilizing an antibody against endogenous Ezh2, harvested at growth and 24 hours post-serum induction of differentiation. The 24-hour timepoint was selected based on the hypothesis that Ezh2 is cleaved by caspase-3 at this stage. This timing also aligns with the observed peak formation of CAD-induced DNA strand breaks, as reported by Larsen et al. (2010).

Results

4.1.Epigenetic Modifier Ezh2 Binds to CAD Targets in the Genome of Muscle Cells

Bioinformatic analyses indicated extensive genetic overlap between CAD and Ezh2 genome-wide targets, in which 90.5% of CAD regions are shared with Ezh2, while 67.9% of Ezh2 regions are shared with CAD (Fig. 4.1A). Both CAD and Ezh2 share similar peak count frequency and distribution of transcription factor-binding loci relative to transcription start sites (TSS) (Fig. 4.1B). Additionally, functional annotation demonstrates that the shared targets between CAD and Ezh2 are similar, and are predominantly located in promoters <=1kb, followed by distal intergenic regions (Fig. 4.1C, D).

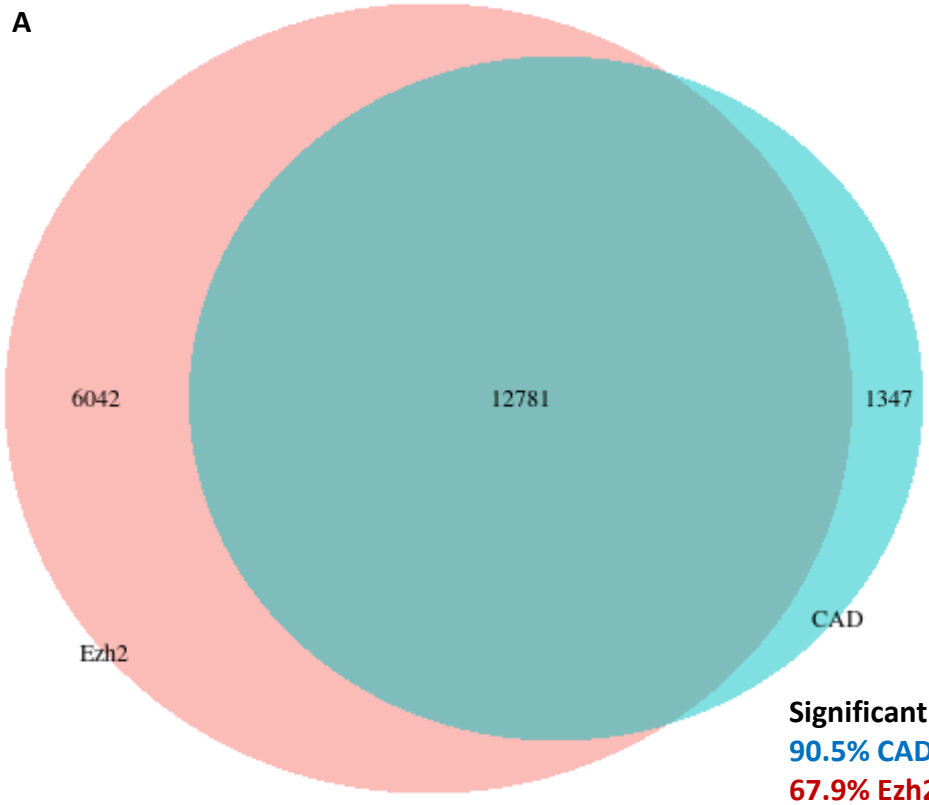
To explore the functional relevance of the shared genes between CAD and Ezh2, I conducted gene-set enrichment analysis. This analysis highlighted key biological processes with KEGG pathway analysis revealing significant enrichment in pathways associated with various cancers, including breast, pancreatic, colorectal, gastric, small cell lung, prostate, endometrial, and hepatocellular carcinoma (Fig. 4.1E). Additionally, Gene Ontology (GO) annotation identified associations with terms related to skeletal muscle cell differentiation, cardiac hypertrophy and differentiation, myeloid cell differentiation, and hematopoietic progenitor cell proliferation and differentiation (Fig. 4.1F). Pathways involved in cell cycle regulation and Wnt signaling were also enriched among the shared gene sets between CAD and Ezh2 (Fig. 4.1E, F).

The significant genetic overlap between CAD and Ezh2 prompted us to investigate whether these CAD and Ezh2 colocalize at breakpoint genes. Utilizing IGV, I analyzed a set of CAD genomic targets highlighted in the previous chapter. Notably, we observed a strong overlap between CAD and Ezh2 at specific genomic loci during early differentiation stages, particularly in genes

implicated in hematopoietic and T cell translocations. This cohort included fusion-prone genes such as *Pax7, Nkx2-1, Lmo2, Hoxa10, Tal1, Lyl1, Bcr, Pml, Bcl2, Foxo1, Rara, Runx1, Btg1, Pbx1, Cbfa2t3, Abl1, Bcl9l, Bcl11b, Hlf, Tlx1, and Tlx3* (Fig. 4.1G). These findings suggest that Ezh2 may play a role in repressing CAD's targeting or accessibility at specific genomic regions.

Venn Diagram of CAD and Ezh2 Binding Sites

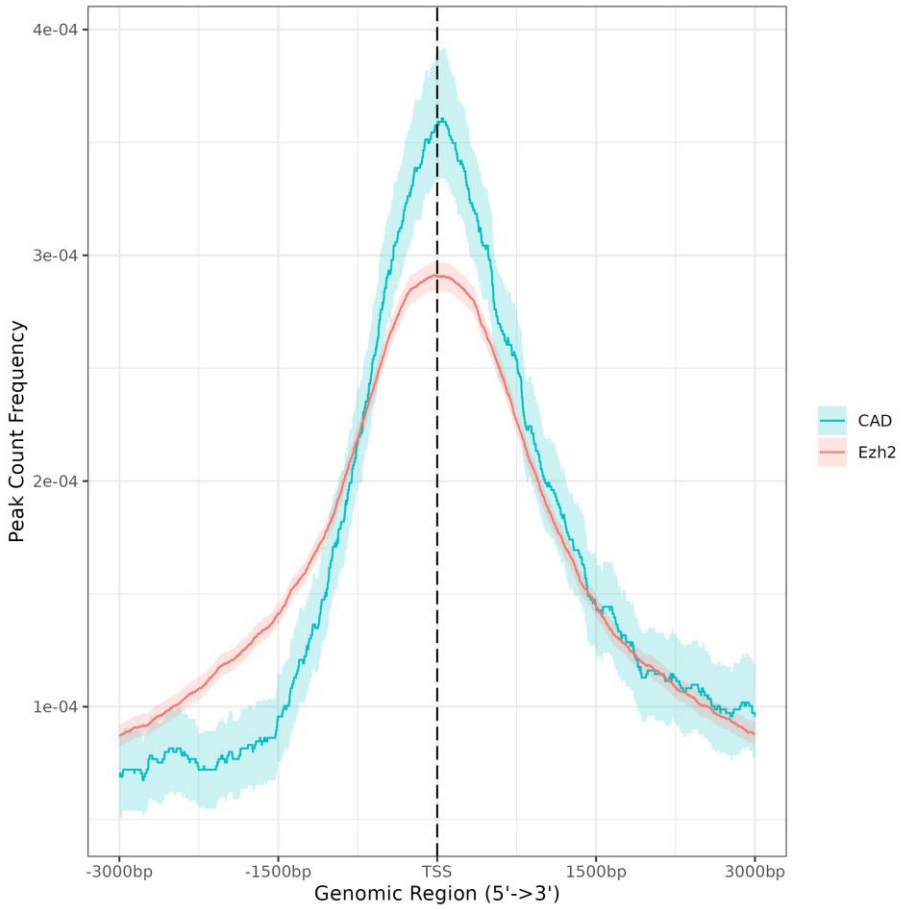
A



Significant Overlap between CAD and Ezh2
90.5% CAD sites are shared with Ezh2
67.9% Ezh2 sites are shared with CAD

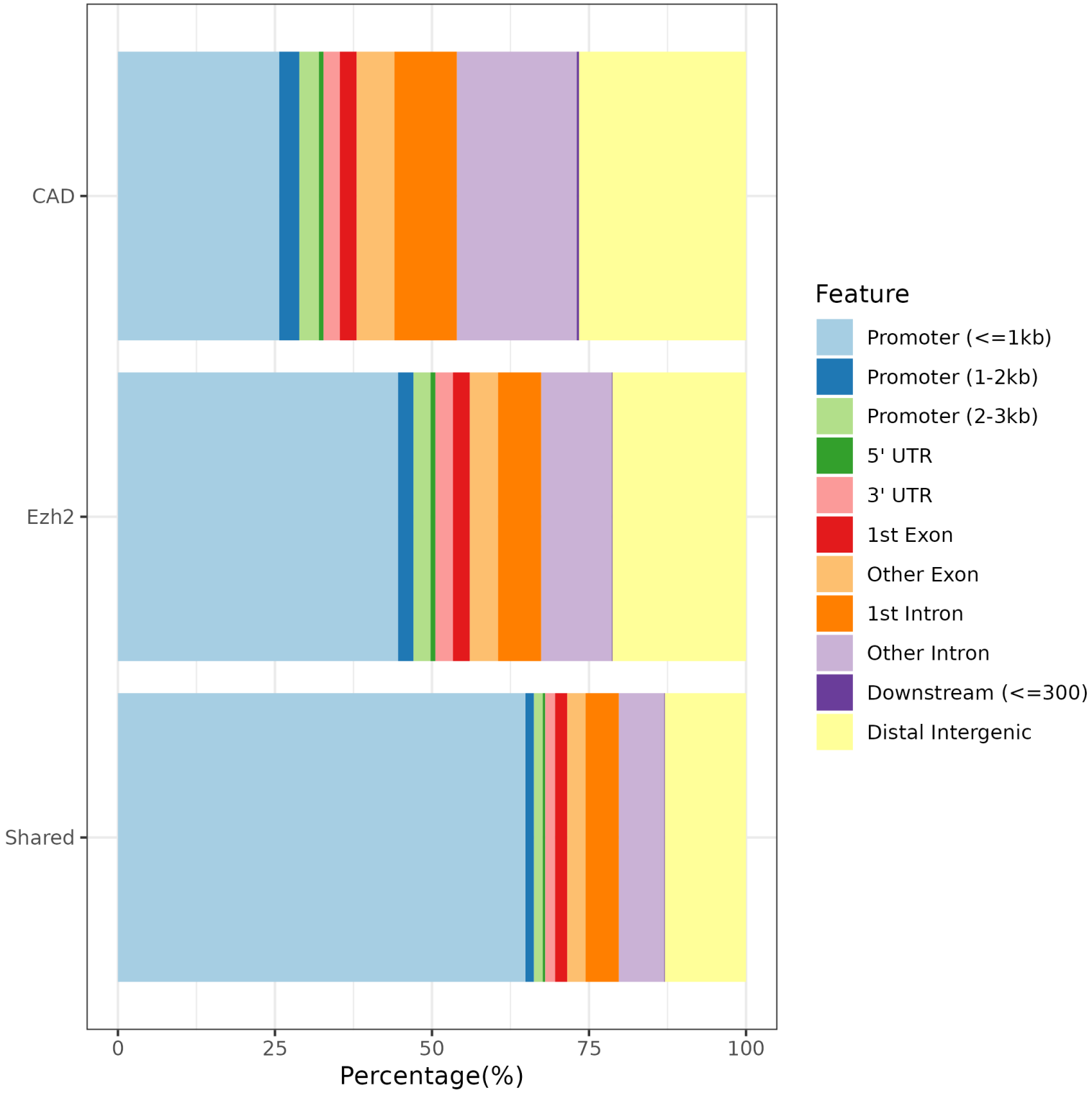
Genomic Regions of CAD and Ezh2 Binding Sites

B



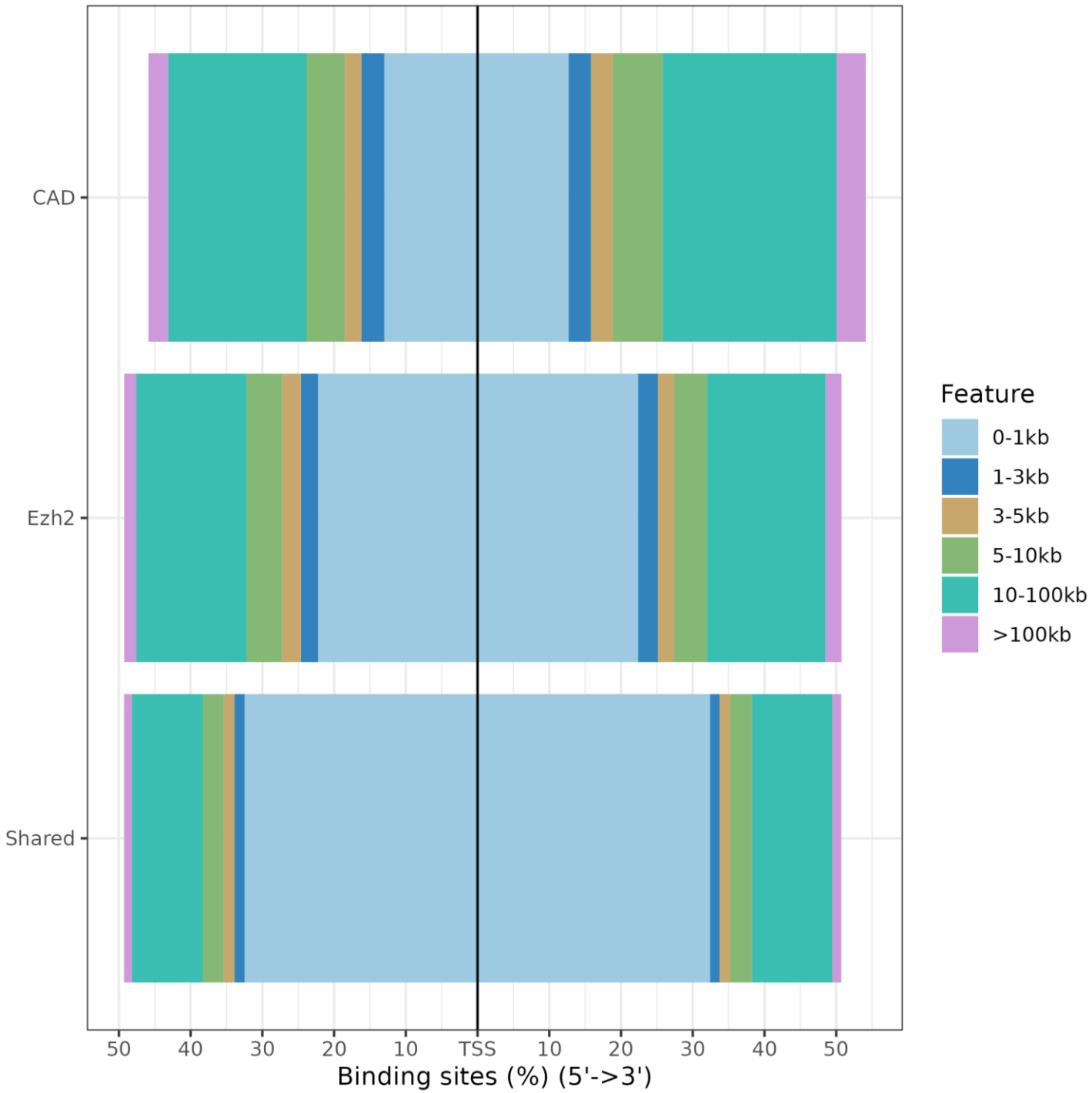
Genomic Region Distribution of CAD and Ezh2 Binding Sites

c



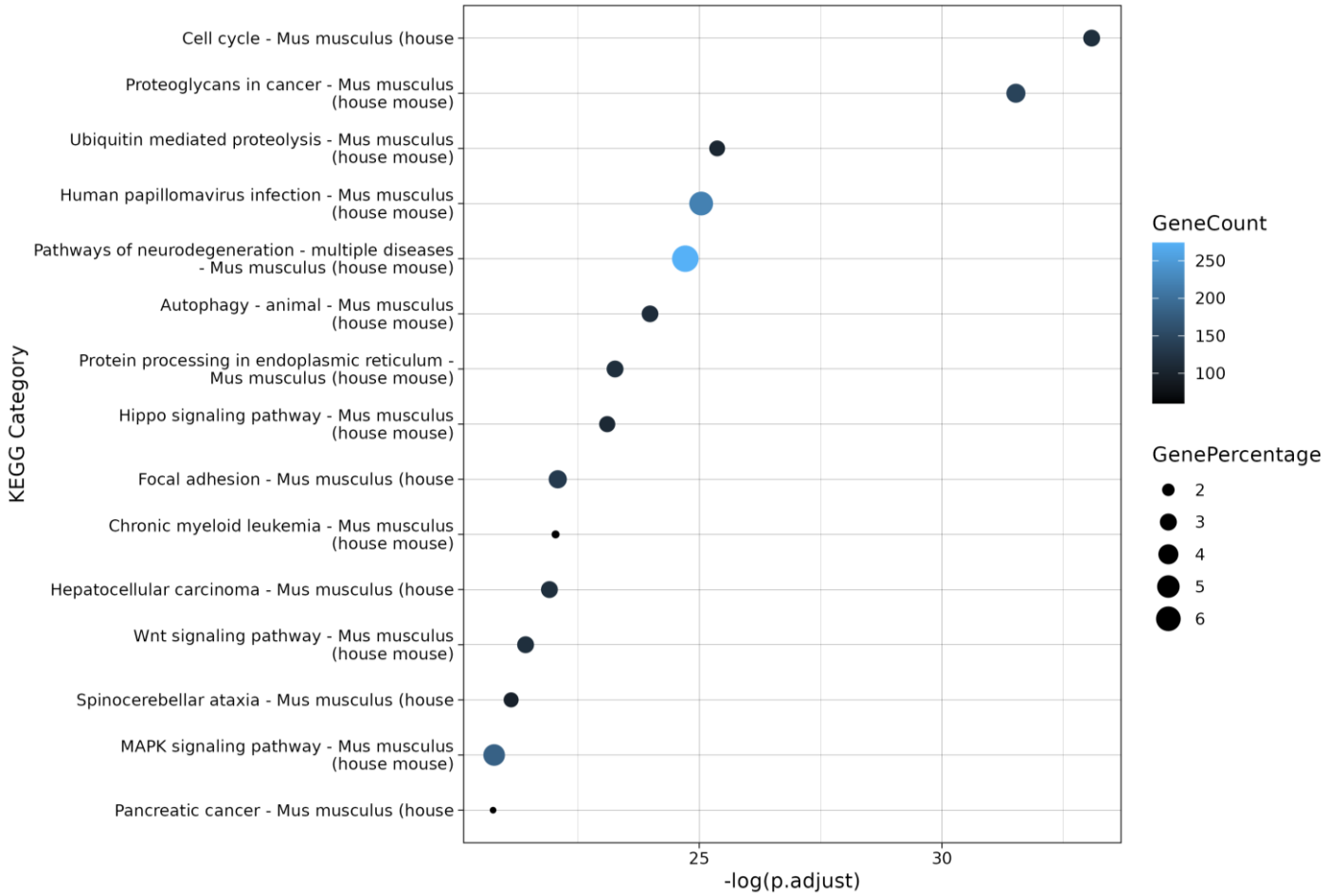
Distribution of CAD and Ezh2 Binding Sites relative to TSS

D



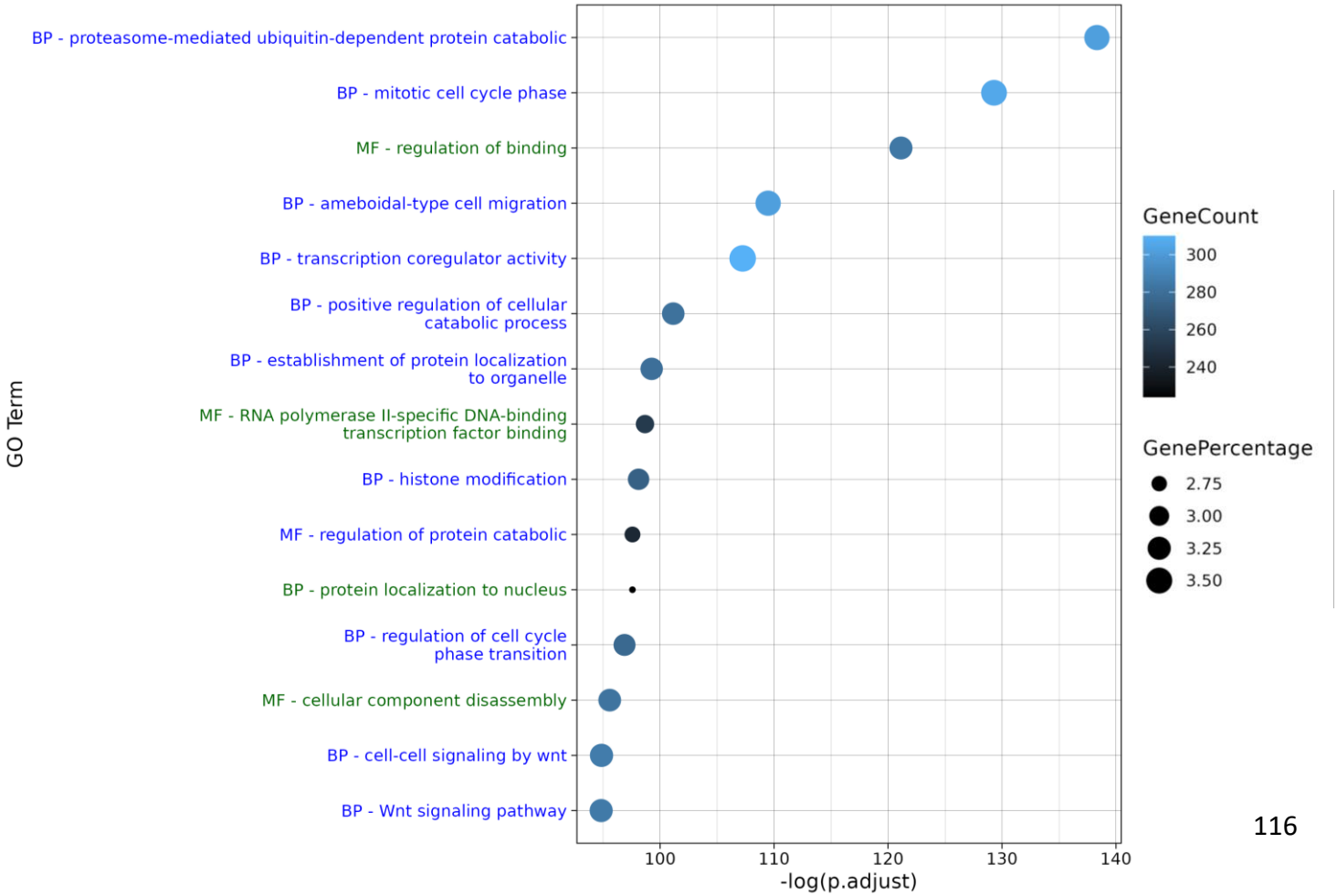
E

KEGG Pathway Analysis Identified within Shared Gene between CAD and Ezh2



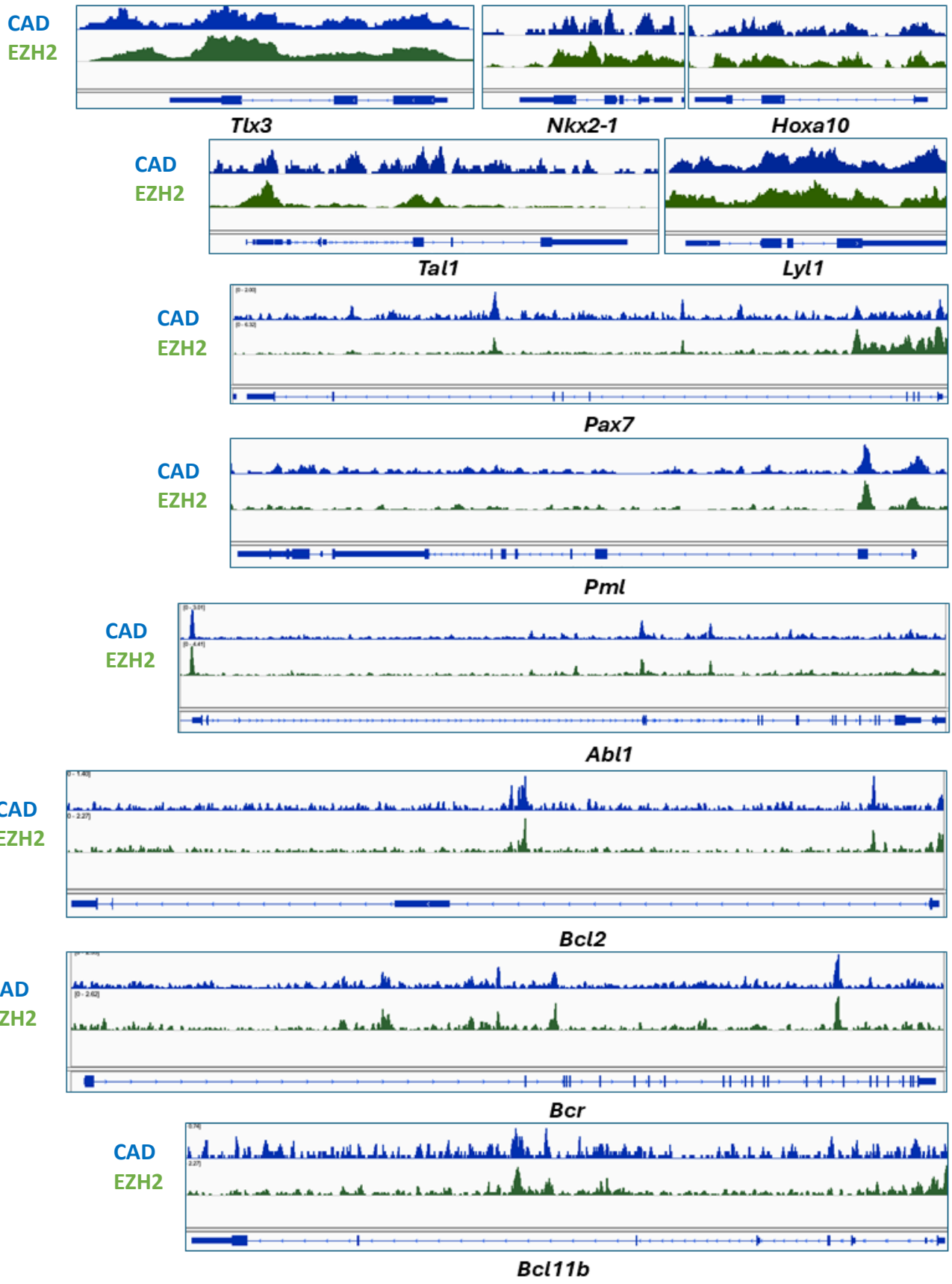
F

Gene Ontology (GO) Analysis Identified within Shared Gene between CAD and Ezh2



Detection of Breakpoint Genes Overlap between CAD and Ezh2

G



Detection of Breakpoint Genes Overlap between CAD and Ezh2

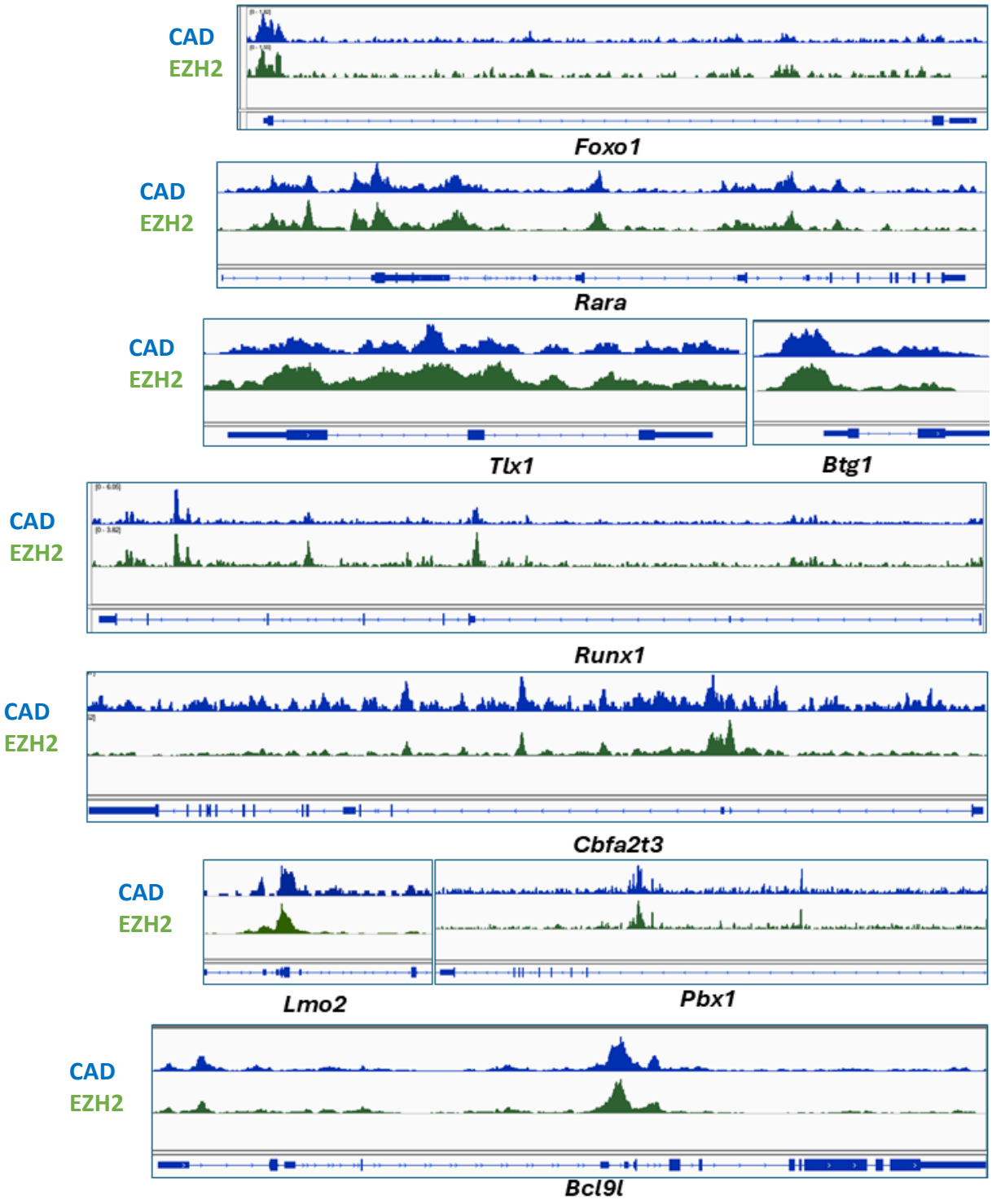


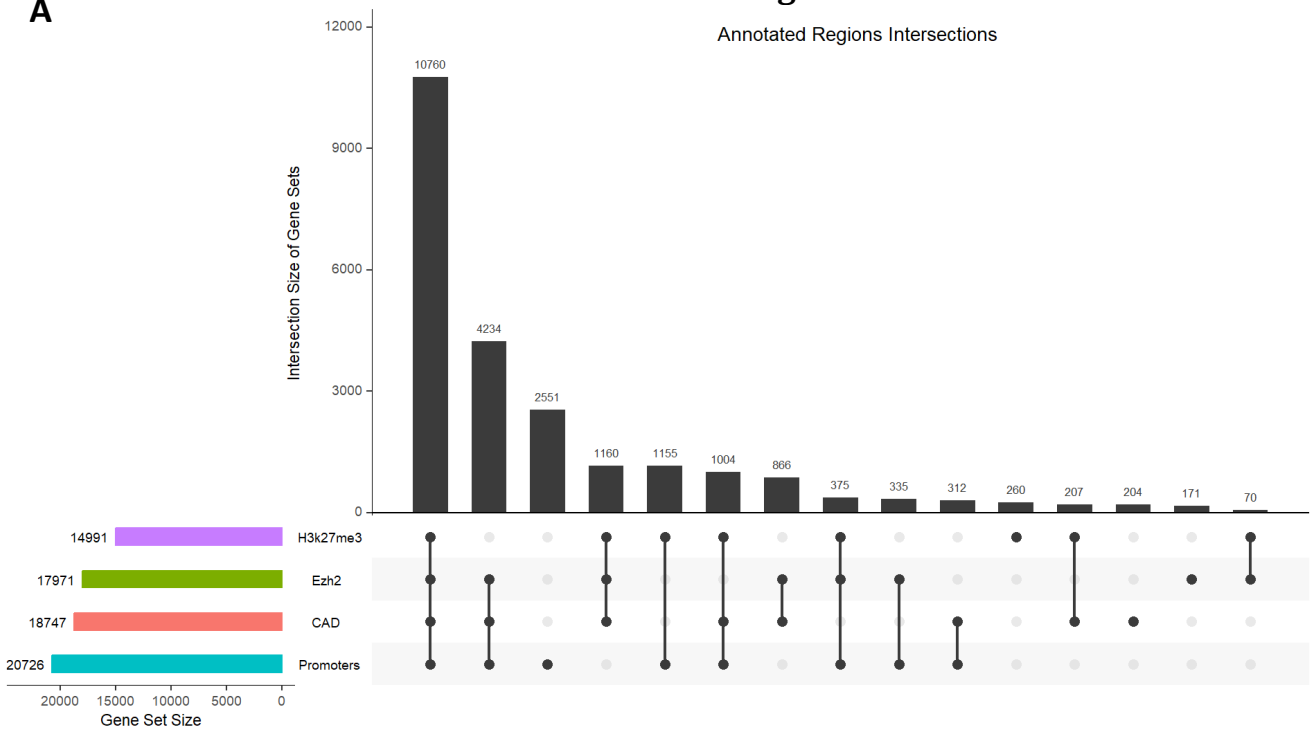
Fig. 4.1. Genomic distribution of CAD and Ezh2 binding sites.
(A) Venn diagram illustrating binding site overlap among functional gene lists from peak analysis between CAD and Ezh2. (B) Peak count frequency over genomic regions for CAD and Ezh2. (C) Genomic region distribution in the peaks falling in various categories displayed. (D) Distribution of transcription factor-binding loci relative to TSS in the peaks falling in various categories displayed. (E) KEGG analysis identified within the shared genes between CAD and Exh2. Top 15 terms are enriched. (F) Gene Ontology (GO) analysis identified within the shared genes between CAD and TopIIIβ. Top 15 terms are enriched. Biological Processes (BP), Molecular Functions (MF). (G) Integrative genome view representing enrichment of CAD and Ezh2 detected with different breakpoint genes.

4.2. Chromatin Modifier Ezh2 and H3k27me3 Histone Mark Overlap with CAD Targeting Sites

H3K27me3, an epigenetic mark deposited by Ezh2, is typically enriched within the promoter regions of repressed genes (Wiles & Selker 2018). As such, I analyzed the overlap between Ezh2 and H3k27me3 peaks, annotated to the nearest gene TSS. Specifically, we assessed the degree of similarity among CAD-Ezh2-H3k27me3 targets relative to promoters/enhancers. We considered that genomic subdomains may exist for CAD vs. Ezh2 against H3K27me3, that consist of overlap subdomains at enhancers only, promoters only, or both. The resulting UpSet plots illustrate the intersections of genes, revealing that the majority of shared genes between CAD, Ezh2, and H3K27me3 are concentrated at promoters (10,760 genes) (Fig. 4.2A), followed by enhancers (5,100 genes) (Fig. 4.2B), and both enhancers and promoters (11,125 genes) (Fig. 4.2C). These findings indicate a significant overlap in the gene promoters and enhancers targeted by CAD, Ezh2, and H3K27me3 in differentiating muscle cells (Fig. 4.2D).

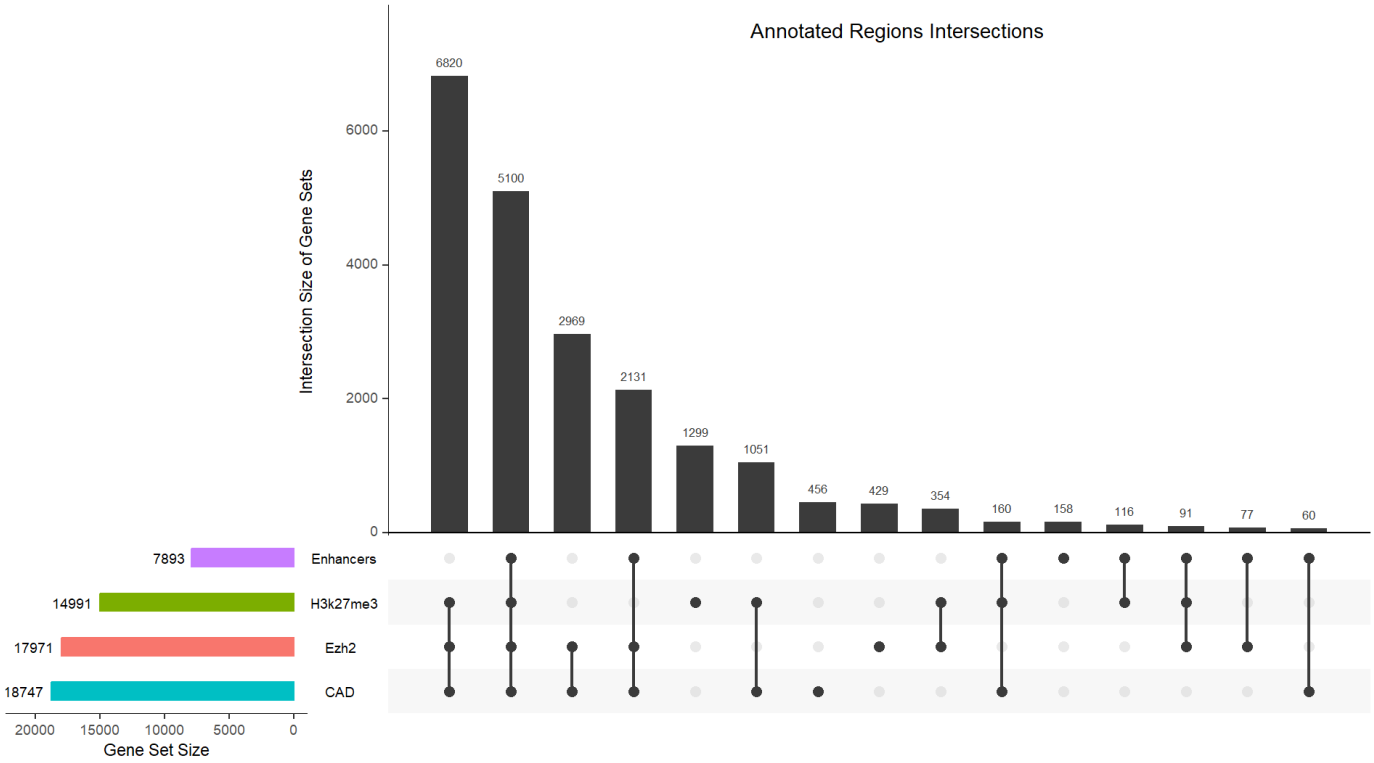
Intersection Size of Gene Shared between CAD, Ezh2, and H3K27me3, at Promoter Region

A

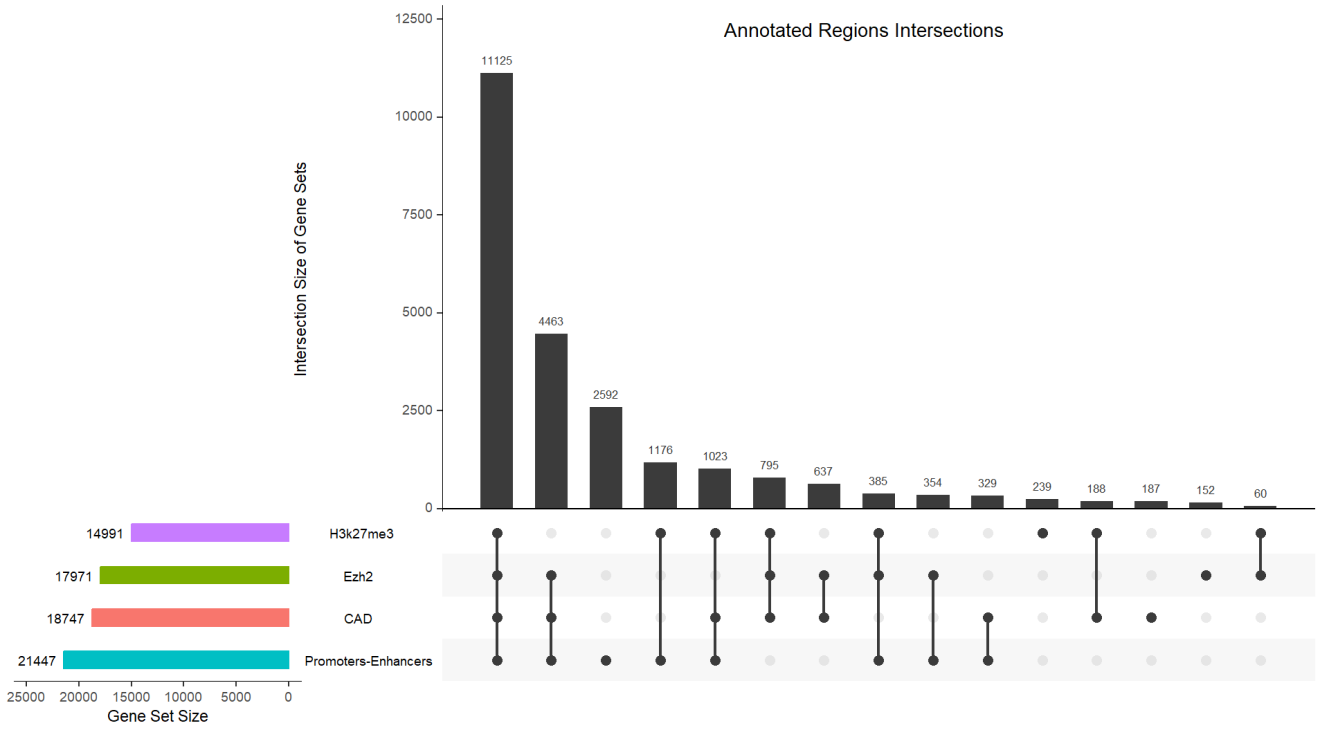


Intersection Size of Gene Shared between CAD, Ezh2, and H3K27me3, at Enhancer Region

B



C **Intersection Size of Gene Shared between CAD, Ezh2, and H3K27me3, at Promoter-Enhancer Regions**



D **Peak Over the Genome for CAD, Ezh2, and H3K27me3**



Fig. 4.2. Ezh2 and H3k27me3 overlap with CAD targeting sites.
(A-C): Upset plots visualizing the intersection size of gene sets between CAD, Ezh2, and H3K27me3 in C2C12 cells, categorized by enhancers only, promoters only, or both. Upset plots were generated using the upset function from the UpSetR package. (D) Peaks over genome represented between CAD, Ezh2, and H3K27me3. Created with ChIP seeker, covplot function.

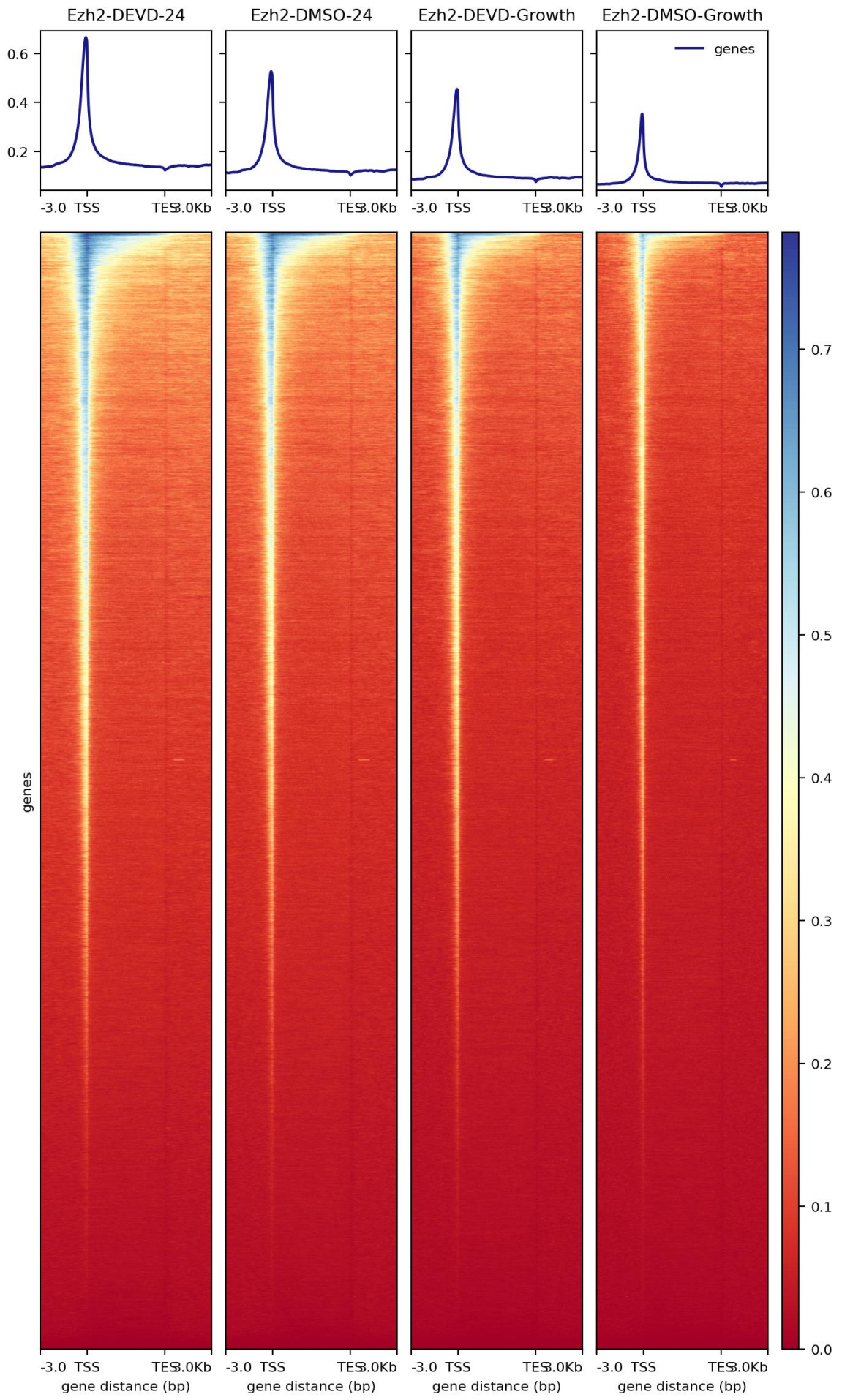
4.3.Ezh2 Genomic Targets are Sensitive to Caspase-3 Inhibitor during Differentiation of C2C12 Myoblasts

Several studies have highlighted the role of Ezh2 in myogenesis (Caretti et al, 2004; Stojic et al, 2011; Dudakovic et al, 2015), however, none have explored the role of Ezh2 as a chromatin modifier in relation to caspase-3 activity or how this protease might influence Ezh2's genomic loci. To investigate this, I conducted CUT & Tag experiments during caspase-3 inhibition, hypothesized that the loss of caspase-3 might lead to increased retention of EZH2. To directly address this, I treated C2C12 muscle cells with either 15uM z. DEVD-fmk, a chemical inhibitor of caspase-3, and -7, or 15uM DMSO, followed by CUT and Tag using an antibody against endogenous Ezh2.

Interestingly, caspase-3 inhibition enhanced the binding of Ezh2 genomic targets near TSS in z.DEVD-fmk treated conditions, both during growth and 24 hours post-differentiation, compared to DMSO treatment (Fig. 4.3A). These findings suggest that Ezh2 is a caspase-3 sensitive event. The plotProfile shows the overall binding site enrichment and strength of the signal for all genes (Fig. 4.3B). Ezh2 binding sites were enriched within promoters (35.83%), followed by distal intergenic regions (25.89%), and introns (17.1%) introns (Fig. 4.3C), with fewer binding events annotated in other genomic regions.

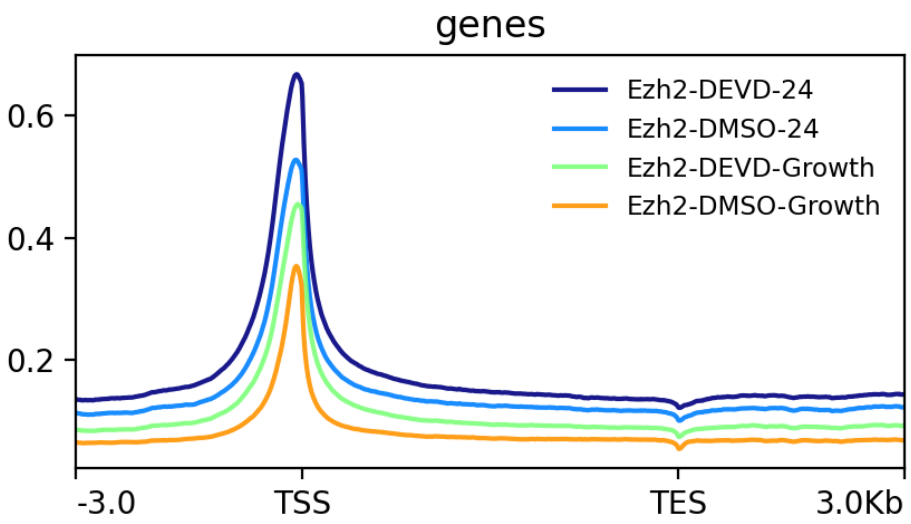
Further, I performed motif enrichment analysis using MEME-ChIP tool, revealing the most frequent motifs within the DNA sequence at Ezh2 peaks (Fig. 4.3D). The identified motifs align with commonly enriched motifs for Ezh2 previously reported, including SRY, SP1, SP2, KLF5, MEF2C, HLF, CEBPA, and STAT1 (Kim et al,2018).

A

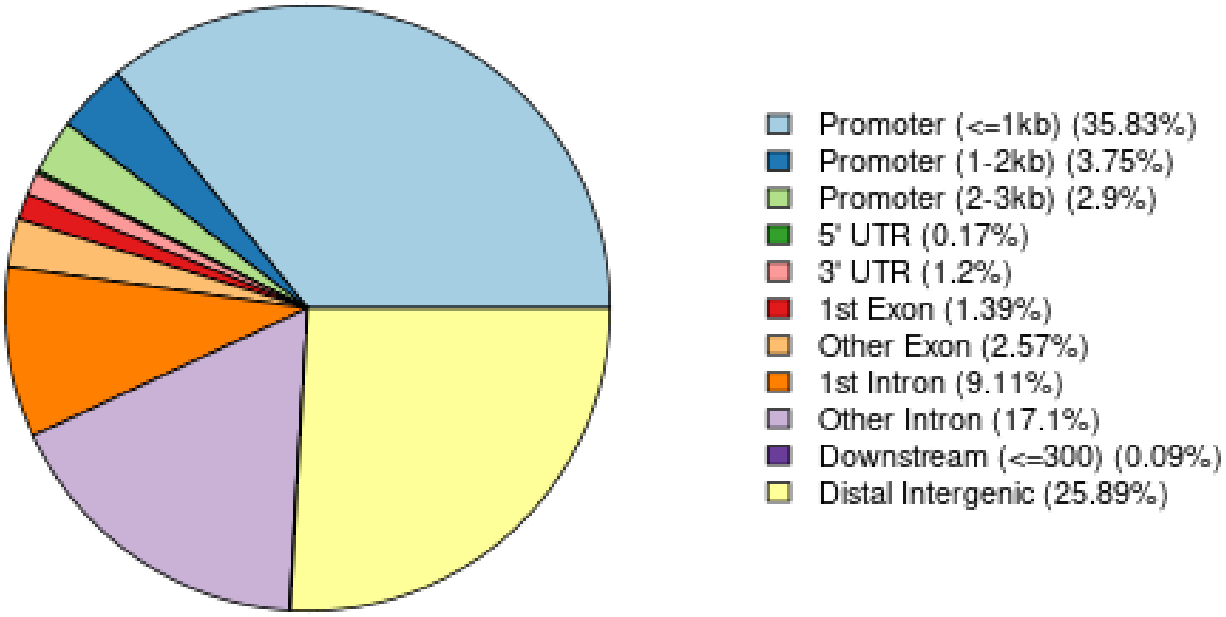

Genomic Distribution of Ezh2 Binding Sites
Ezh2-DEVD-24
Ezh2-DMSO-24
Ezh2-DEVD-Growth
Ezh2-DMSO-Growth
genes
0.6
0.4
0.2
-3.0 TSS TES 3.0Kb
genes
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
gene distance (bp)

B

Genomic Distribution of Ezh2 Binding Sites



C



D

Motif Enrichment Analysis of Ezh2 Bound Genomic Regions

Discovered Motifs

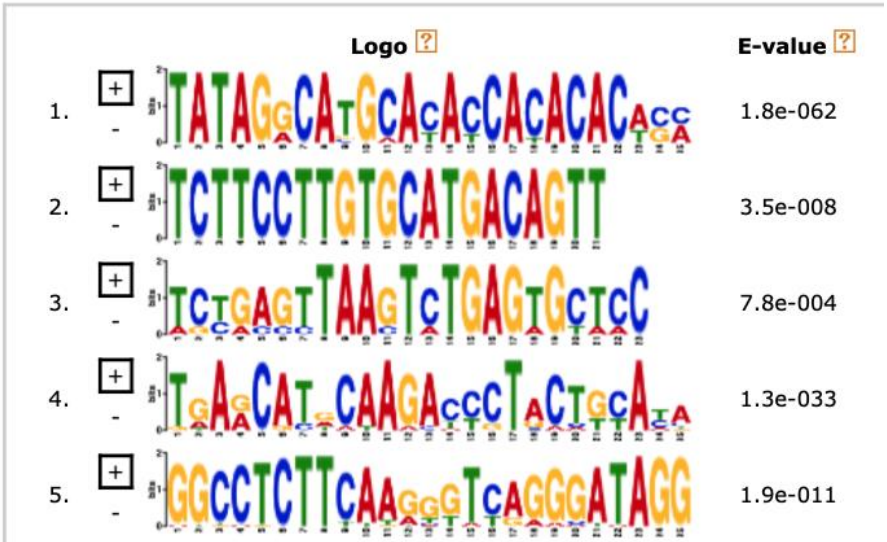


Fig. 4.3. Genomic distribution of Ezh2 binding sites.
(A) Genomic distribution of the binding sites of Ezh2 at Growth, and 24 hours post differentiation from C2C12 muscle cells treated with DMSO or z.FMK-DEVD. plotHeatmaps of Ezh2 CUT and Tag show count per million (CPM) normalized signals for reads across 3,000 bp genomic interval relative to the TSS. Data sets were generated from merged two independent biological replicates for each condition. Heatmap profiles were generated using deepTools computeMatrix (Ramírez et al, 2016). Peak calling of binding sites was performed using Model-based Analysis of Chip-Seq. (MACS) (Zhanget al,2008), Sparse Enrichment Analysis for CUT & RUN (SEACR) (Meers et al, 2019), and GoPeaks (Yashar et al,2022). (B) PlotProfile is averaged profiles of the compressed matrix of occupancy/signal near TSS produced by computeMatrix. (C) Pie chart illustrating genomic distribution of the binding sites of Ezh2 at 24 hours post differentiation from C2C12 muscle cells treated with z.FMK-DEVD. Different genomic regions displaying the genomic location of Ezh2 binding sites across the genome including promoters, enhancers, distal intergenic regions, exons, introns, two untranslated regions (UTRs). This plot shows the percentage for each genomic location category. DiffBind tool was used to obtain the consensus peakset. AnnotatePeak function was used from ChIPseeker package on those peaks. Then a plotting package in R named ggplot2 that was used for the bar and pie charts, and the paletteer package for the color scheme. (D) Graphical representation of sequence motifs detected in Ezh2 DNA sequences Motif enrichment analysis identifies de novo Motif showing the top enriched sequence motifs of Ezh2 bound genomic regions. Motifs enrichment was calculated with HOMER tool. E-values of motifs are indicated. The statistical significance of the motif (E-value) is an estimate of the log ratio, width, site count, and the background letter frequencies.

4.4.Detection of EZH2 Fragments during Differentiation of Muscle Cells

Given that Ezh2 genomic targets are sensitive to caspase-3, we aimed to determine whether Ezh2 targeting of the genome is directly or indirectly influenced by caspase-3 activity. An indirect mechanism of caspase sensitization of Ezh2 binding could be accounted for caspase 3 targeting other proteins, which in turn affect Ezh2 recruitment to the genome. Alternatively, but not exclusively, caspase-3 may directly target and cleave Ezh2 protein itself, which would have an obvious impact on Ezh2 recruitment kinetics during muscle cell differentiation. As previously discussed, (and throughout this thesis), caspase-3 is active during muscle cell differentiation and targets several nuclei specific proteins, making it a plausible candidate for proteolytic targeting of Ezh2. Interestingly, visual inspection of human and mouse Ezh2 amino acid sequences contain several potential caspase-3 cleavage sites (Fig. 4.4A).

Based on this, I monitored Ezh2 expression throughout a differentiation time course in C2C12 muscle cells. Full-length Ezh2 expression was highest in proliferative cells, but as differentiation progressed, full-length Ezh2 levels declined, reaching their lowest at 72 hours post-induction of differentiation, as previously reported (Fig. 4.4B) (Caretti et al, 2004).

Interestingly, during differentiation, I detected a probable fragment of Ezh2 (~53 kDa) (Fig. 4.4B), suggesting the possibility of an active proteolysis of the protein. The observation of this distinct fragment prompted us to examine whether caspase-3 directly targets the full-length Ezh2 during differentiation. I observed that cleaved (active) caspase-3 peaked at 24 and 48 hours (Fig. 4.4C). Staurosporine (STS), a standard apoptotic treatment, served as a positive control. These data suggest that the detection of active caspase-3 coincides with the decrease in full-length Ezh2 protein. It's noteworthy that Ezh2 has been reported to subject proteasome-mediated degradation

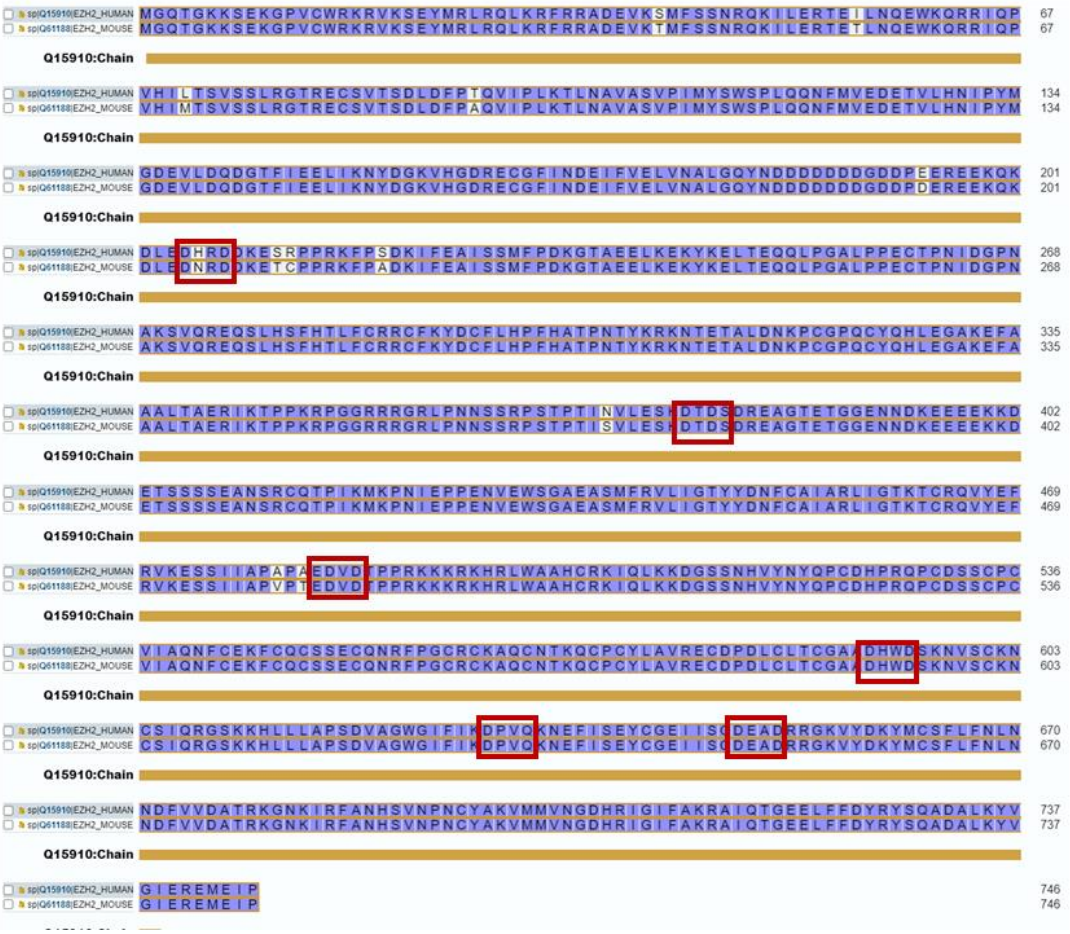
in differentiating muscle cells, particularly through phosphorylation at threonine 372 (Consalvi et al, 2017). Interestingly, their findings indicate that Ezh2 degradation is an early event in differentiation program, consistent with our observation of caspase-3-mediated cleavage of Ezh2 protein.

To assess the impact of caspase-3 inhibition on protein levels, I performed immunoblotting on cells infected with either GFP Advs or p35 Advs. The full-length Ezh2 decreased post-differentiation, accordingly more fragments were detected from the cleaved full length (Fig. 4.4D). Conversely, cells infected with p35 Advs, which inhibit caspase-3, showed elevated levels of full-length Ezh2, with the detection of fewer Ezh2 fragments. Taken together, these results suggests that Ezh2 is sensitive to caspase-3 during myogenesis.

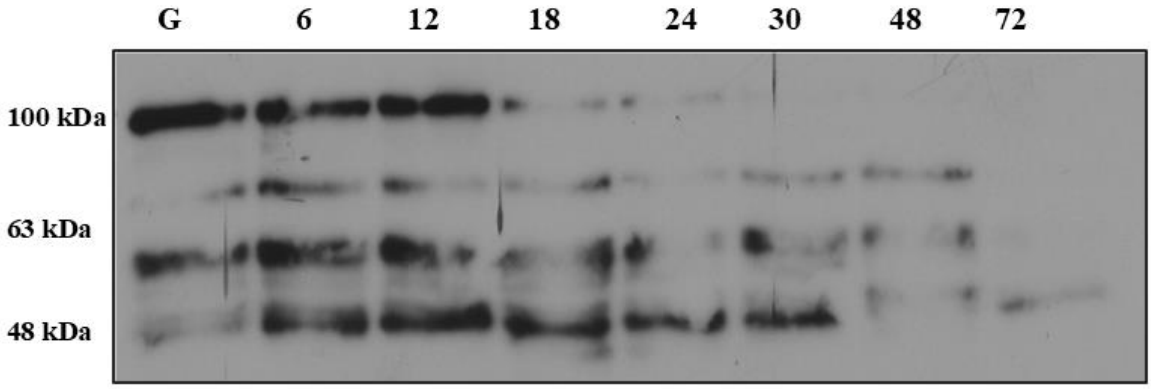
A

Caspase-3 Prospective Cleavage Sites in hEzh2 and mEzh2

DHRD, DHWD, EDVD, DPVQ, DEAD, KDTDS



B



IB: Ezh2 (full length)

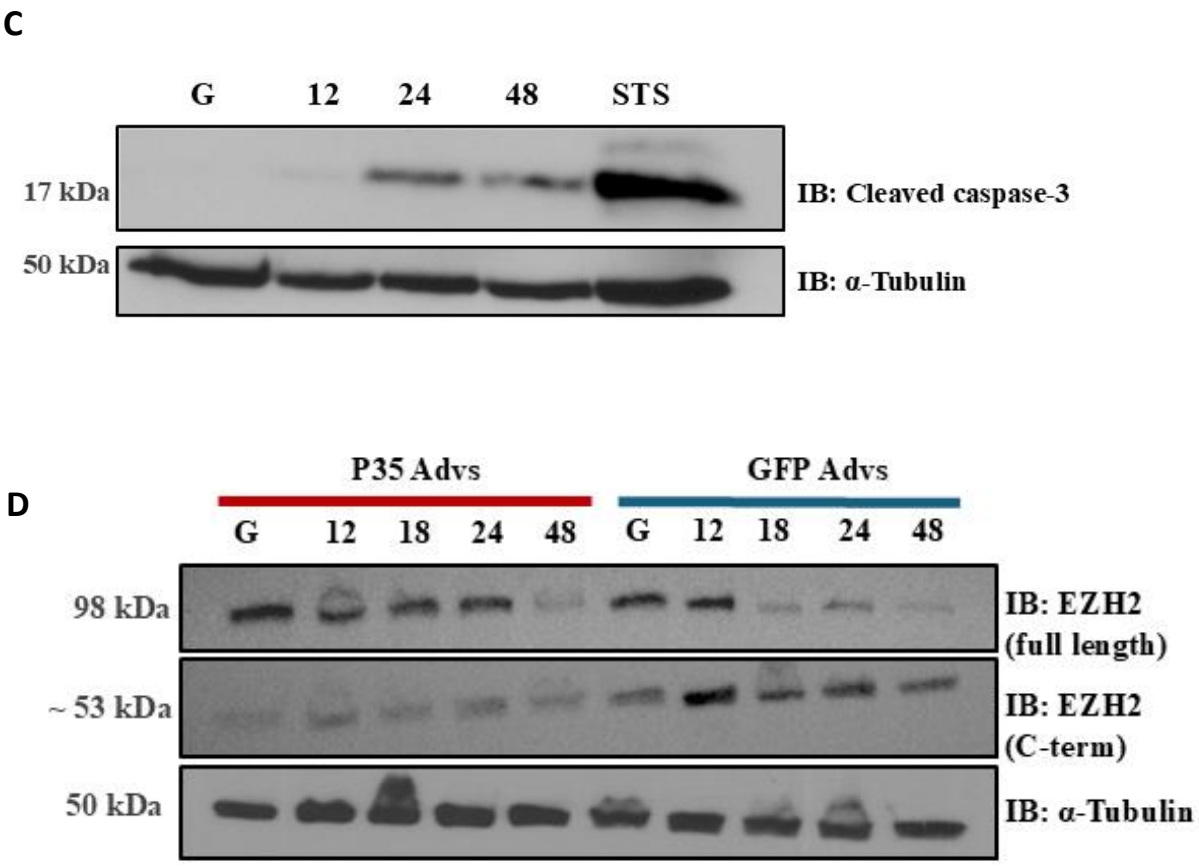


Fig. 4.4. Ezh2 expression during differentiation of C2C12 myoblasts.
(A): Amino acid sequence of the Ezh2 protein (human and mouse) showing several potential caspase-3 cleavage sites. (B): Western blot analysis of Ezh2 expression in C2C12 cells at growth (G) and during differentiation at 6, 12, 18, 24, 36, 48, and 72 hours. Blots were probed with an Ezh2 antibody. A total of 50 µg of lysate was loaded per lane. N=3 biological replicates. (C): Western blot showing active caspase-3 levels in C2C12 myoblasts induced to differentiate, with samples taken during proliferation (G) and post-serum starvation at 12, 24, and 48 hours. N= 3 biological replicates. (D): Representative image of C2C12 myoblasts expressing either p35 Advs or GFP Advs, serum-starved to induce differentiation, and then immunoblotted with antibodies targeting the N-terminus or middle of the Ezh2 sequence. A total of 50 µg of lysate was loaded per lane. N= 3 biological replicates.

4.5.Ezh2 Protein is a Caspase-3 Substrate

Ezh2 protein is biologically present with the other proteins of the PRC2 complex. To determine whether caspase-3 directly targets Ezh2, I performed an in *vitro* cleavage assay using an Ezh2-SUZ12-EED recombinant protein complex and active recombinant caspase-3. The proteins were incubated in cleavage assay buffer at 37°C for 4 hours with either 15 μM DMSO or z.DEVD. The samples were then immunoblotted and probed with an anti-Ezh2 antibody. The addition of active caspase-3 to the Ezh2-SUZ12-EED complex resulted in cleavage products of approximately 65 kDa, 30 kDa, and 23 kDa (Fig. 4.5A), indicating potential cleavage sites within Ezh2-SUZ12-EED. The z.DEVD inhibitor prevented the generation of these cleavage fragments, suggesting that these fragments are indeed products of caspase-3 activity (Fig. 4.5A). To further characterize and map the caspase-3 cleavage of Ezh2 (within the complex), I subjected the *in vitro* cleavage assay samples to SDS-PAGE and silver staining (Fig. 4.5B). The protein bands corresponding to the probable cleavage fragments were excised and analyzed by LC-MS/MS following Chymotrypsin digest. This analysis identified potential cleavage sites at aspartic acid residues $\mathbf{D_{652}}E_{653}A_{654}D_{655}$, $E_{485}\mathbf{D_{486}}V_{487}D_{488}$, and $\mathbf{D_{208}}H_{209}R_{210}D_{211}$ of Ezh2 protein (Fig. 4.5B). To validate these cleavage sites and exclude non-specific endopeptide generation, cleavage assays were conducted, and probable cleavage fragments were digested with other endoproteases: Trypsin. Trypsin digestion revealed an additional cleavage site at $\mathbf{D_{630}}P_{631}V_{632}Q_{633}$ that was not present in the caspase-inhibited samples (Fig. 4.5C). Collectively, these data suggest that caspase-3 can cleave Ezh2 within its native protein complex.

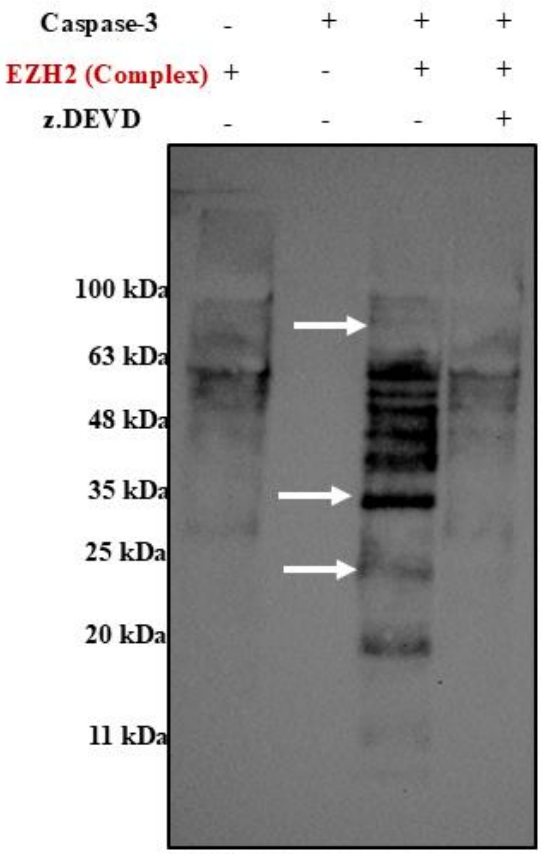
To verify the specificity of this interaction, I conducted additional *in vitro* caspase-3 cleavage assays using recombinant Ezh2 protein alone (without the complex). Immunoblotting revealed that Ezh2 was cleaved into multiple fragments (Fig. 4.5D), indicating that the fragments observed in

the protein complex experiment are likely due to specific cleavage events on Ezh2. Further MS analysis using Arg C digests identified the cleavage site at **D_{630}**$P_{631}V_{632}Q_{633}$, which was also detected in the cleavage assay for the Ezh2-SUZ12-EED complex (Fig. 4.5E). In summary, proteomics-based analysis, including digests with Chymotrypsin, Trypsin, and Arg C, identified potential cleavage sites at D208 and D486 (Fig. 4.5F). Notably, the identification of D630 and D652—residing in the catalytic SET domain of Ezh2— highlights these sites as significant targets for caspase-3 cleavage (Fig. 4.5F).

The cleavage assay represented in (Fig. 4.5B) was performed on two purified proteins, caspase-3 and it's substrate. As such, the generation of cleavage sites with or without the addition of z.DEVD-FMK suggests that these sites are protease cleavage events specific to Ezh2 substrate. It is possible that the cleavage sites indicated in (Fig. 4.4A) are cryptic sites, or hierarchical sites, yet that maybe context evident and requires further experimentation. Construction of a non-cleavable version of Ezh2 protein will elucidate this question.

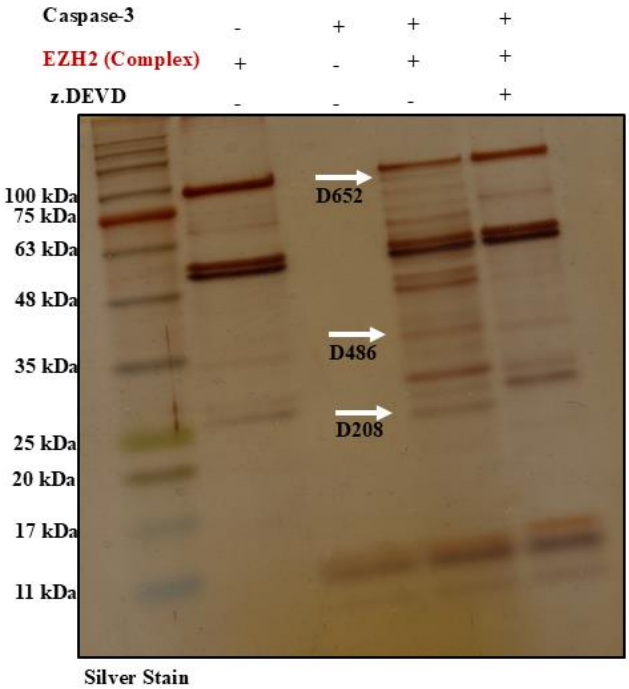
In Vitro Cleavage Assay of Ezh2 in PRC2 Complex

A



IB: Ezh2
Ezh2 86 kDa full length

B



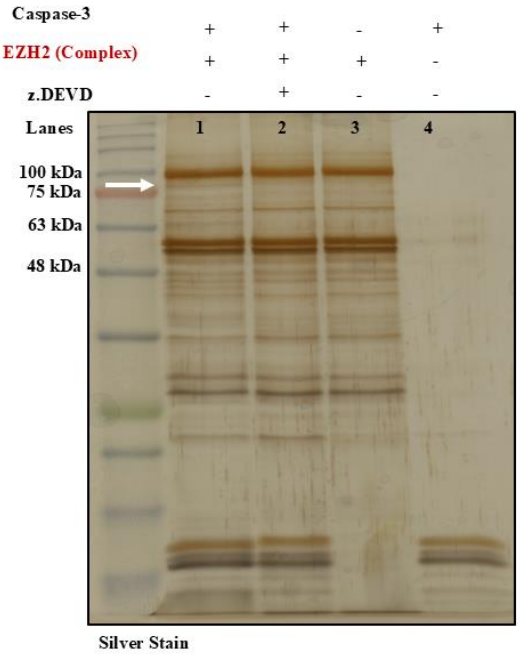
Silver Stain

EZH2 (complex) digested with Chymotrypsin.

```
  1 MGQTGKKSEK GPVCWRKRVK SEYMRLRQLK RFRRADEVKS MFSSNRQKIL ERTEILNQEW
 61 KQRRIQPVHI LTSVSSLRGT RECSVTSDLD FPTQVIPLKT LNAVASVPIM YSWSPLQQNF
121 MVEDETVLHN IPYMGDEVLD QDGTFIEELI KNYDGKVHGD RECGFINDEI FVELVNALGQ
181 YNDDDDDDDG DDPEEREEKQ LEDHRDDK ESRPPRKFPS DKIFEAISSM FPDKGTAEEL
241 KEKYKELTEQ QLPGALPPEC TPNIDGPNAK SVQREQSLHS FHTLFCRRCF KYDCFLHPFH
301 ATPNTYKRKN TETALDNKPC GPQCYQHLEG AKEFAAALTA ERIKTPPKRP GGRRRGRLPN
361 NSSRPSTPTI NVLESKDTDS DREAGTETGG ENNDKEEEEK KDETSSSSEA NSRCQTPIKM
421 KPNIEPPENV EWSGAEASMF RVLIGTYYDN FCAIARLIGT KTCRQVYEFR VKESSIIAPA
    PAEDVDTPPR KKKRKHRLWA AHCRKIQLKK DGSSNHVYNY QPCDHPRQPC DSSCPCVIAQ
541 NFCEKFCQCS SECQNRFPGC RCKAQCNTKQ CPCYLAVREC DPDLCLTCGA ADHWDSKNVS
601 CKNCSIQRGS KKHLLLAPSD VAGWGIFIKD PVQKNEFISE YC   QDE ADRRGKVYDK
661 YMCSFLFNLN NDFVVDATRK GNKIRFANHS VNPNCYAKVM MVNGDHRIGI FAKRAIQTGE
721 ELFFDYRYSQ ADALKYVGIE REMEIP
```

Digest: Chymotrypsin
D652 in DEAD motif
D486 in EDVD motif
D208 in DHRD motif

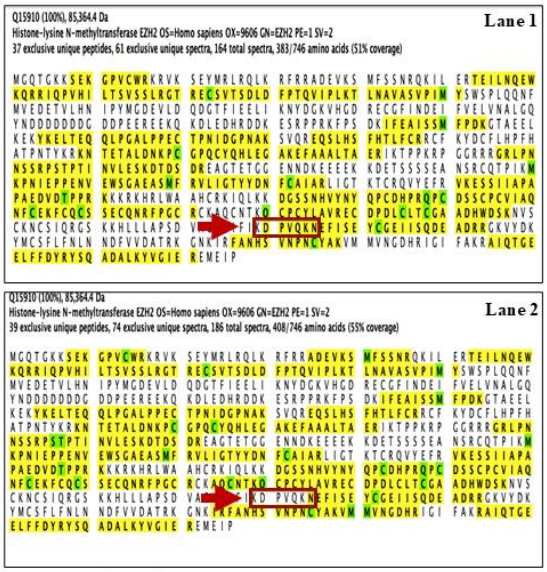
C



Silver Stain

EZH2 (complex) digested with trypsin.

EZH2 amino acid sequence with the identified peptides indicated in yellow.

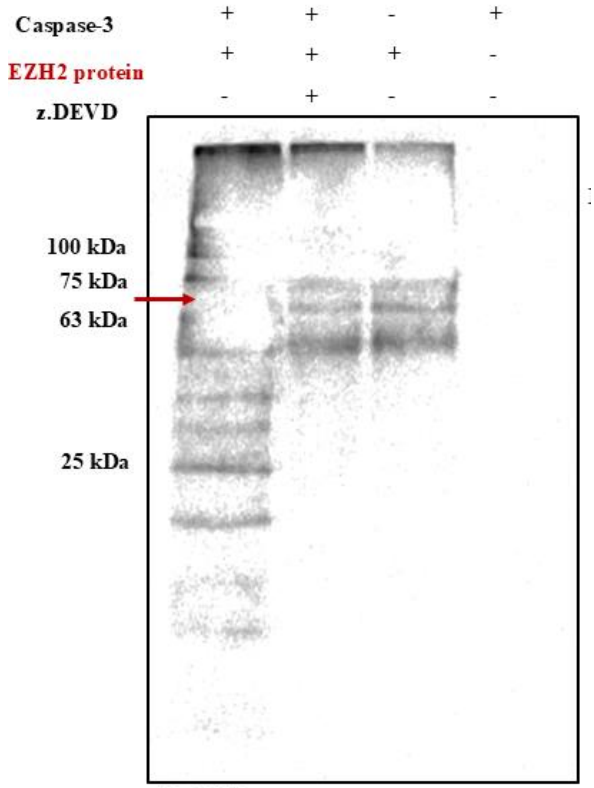


Digest: Trypsin

D630 in DPVQ motif

In Vitro Cleavage Assay of Ezh2

D



IB: Ezh2

Ezh2 86 kDa full length

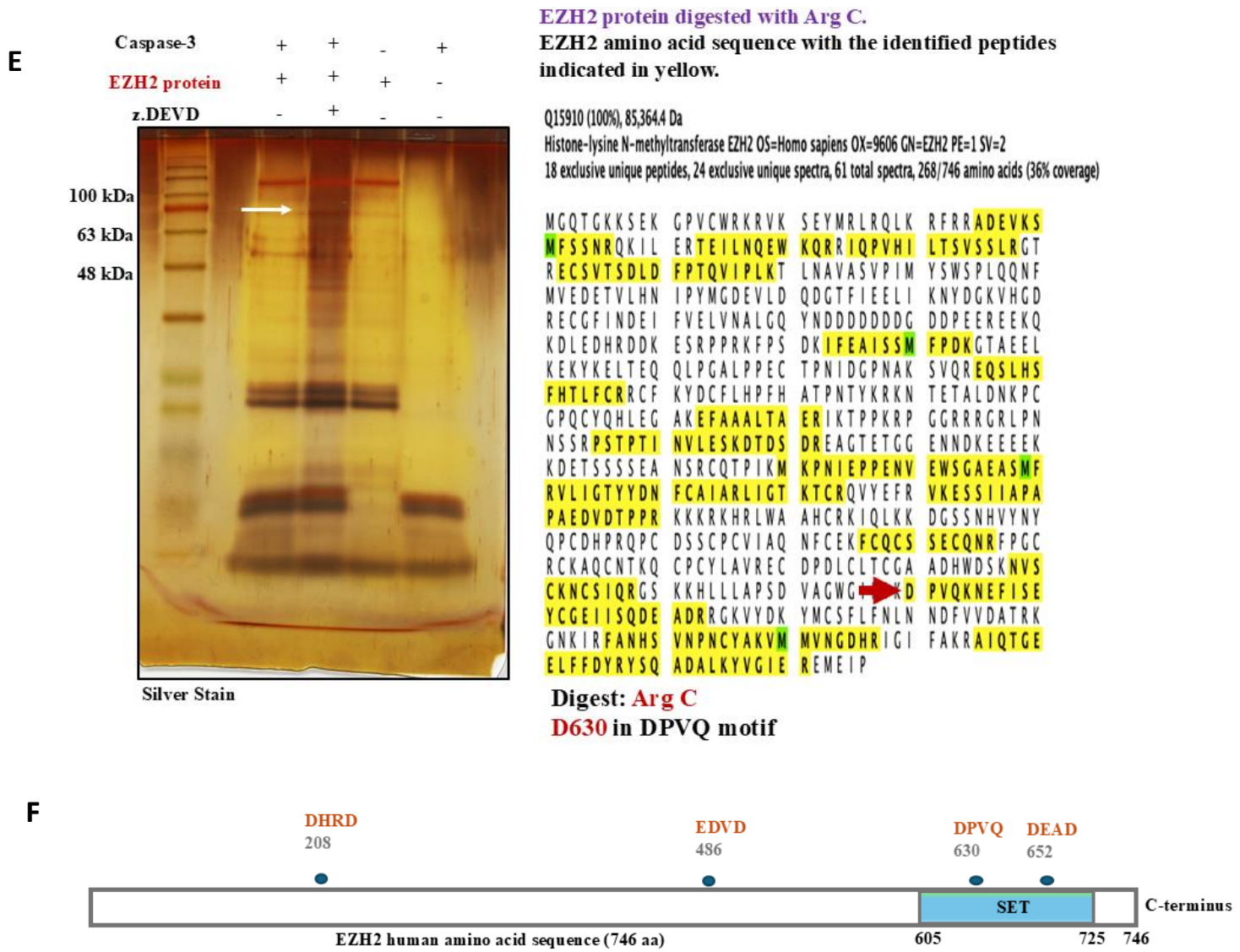


Fig. 4.5. Prospective caspase-3 cleavage sites of the Ezh2 protein complex.
(A-C): Cleavage assay using recombinant Ezh2-SUZ12-EED protein complex and active Caspase-3. Proteins were incubated in cleavage buffer at 37°C for 4 hours. The reaction was terminated by adding Laemmli buffer, and samples were subjected to SDS-PAGE followed by immunoblotting with an Ezh2 antibody (A) or silver staining (B, C). The corresponding cleavage fragments were excised, digested, and identified by mass spectrometry. The Ezh2 amino acid sequence with the identified peptides is indicated in yellow. The fragment marked with a red arrow indicates a probable cleavage site. (D, E): Cleavage assay using recombinant Ezh2 protein and active Caspase-3. Proteins were treated under the same conditions and analyzed as described above, by immunoblotting with an Ezh2 antibody (D) or silver staining (E). (F) schematic of caspase-3 cleavage sites identified in the Ezh2.

4.6.Ezh1, EED, and SUZ12 are not Caspase-3 Targets

Next, I aimed to assess whether caspase-3 affects other PRC2 core components beside Ezh2. To examine the normal biological expression of PRC2 core subunits, I measured their protein levels in proliferating C2C12 muscle cells throughout differentiation. Expression of Ezh1 and EED remained constant during differentiation, implying it is not targeted by caspase-3 (Fig. 4.6A). While the protein levels of SUZ12 did not change from proliferation to early differentiation, the full-length showed a slight decrease at 48 hours, yet the presence of non-specific band of the antibody made it challenging to discern specific fragmentation (Fig. 4.6B). Taken together, among PRC2 complex, Ezh2 is the target to caspase-3 enzyme.

4.7.CAD form immunocomplex with Ezh2

As the results from overlapping gene between Ezh2 and CAD, I ought to assess the interaction at protein level. I performed immunoprecipitation by looking at endogenous levels of expression of CAD from C2C12 cells, followed by immunoblotting with Ezh2. The results show that CAD forms native immunocomplex with Ezh2 at Growth, 12, and 24 hours (Fig. 4.6C). Thus, concurrent caspase targeting events to both Ezh2 and CAD promotes the differentiation program.

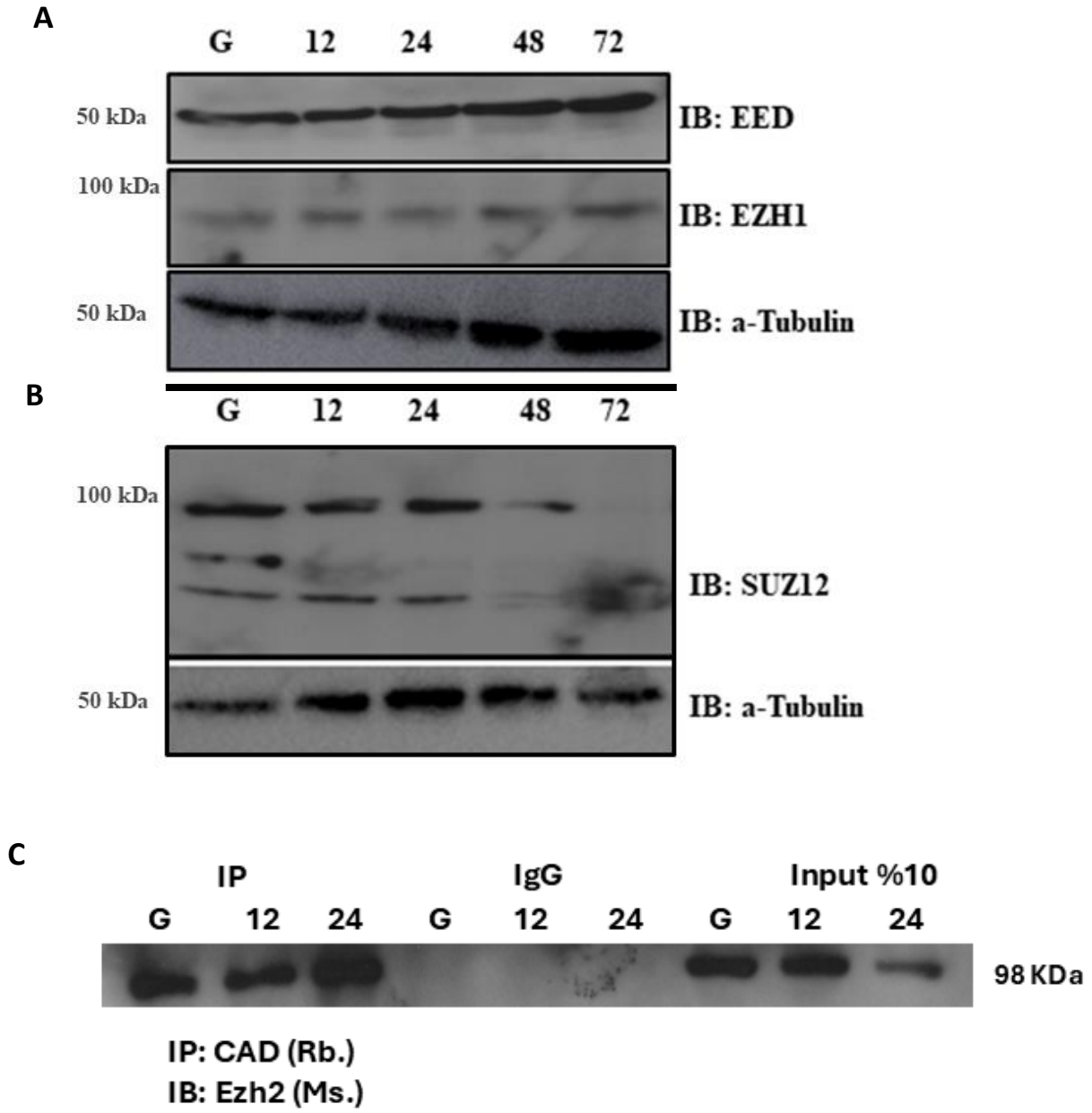
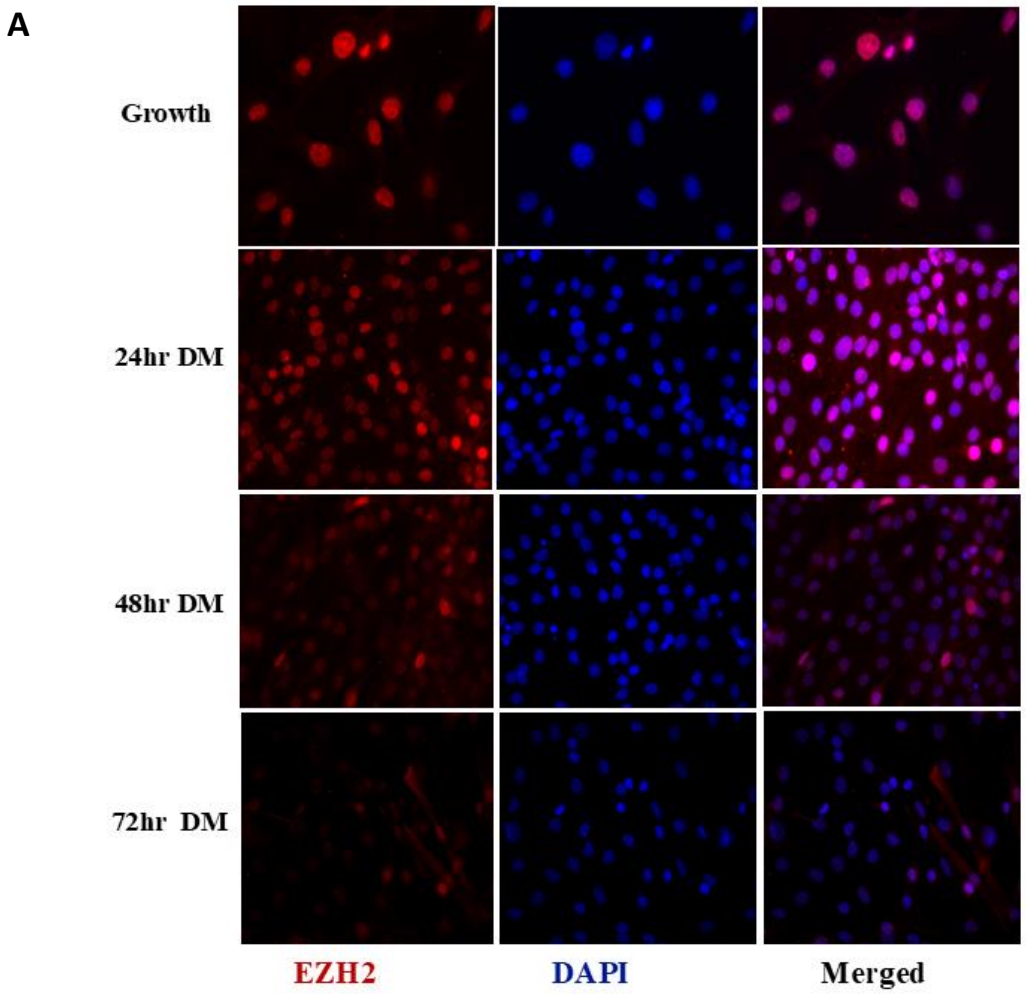
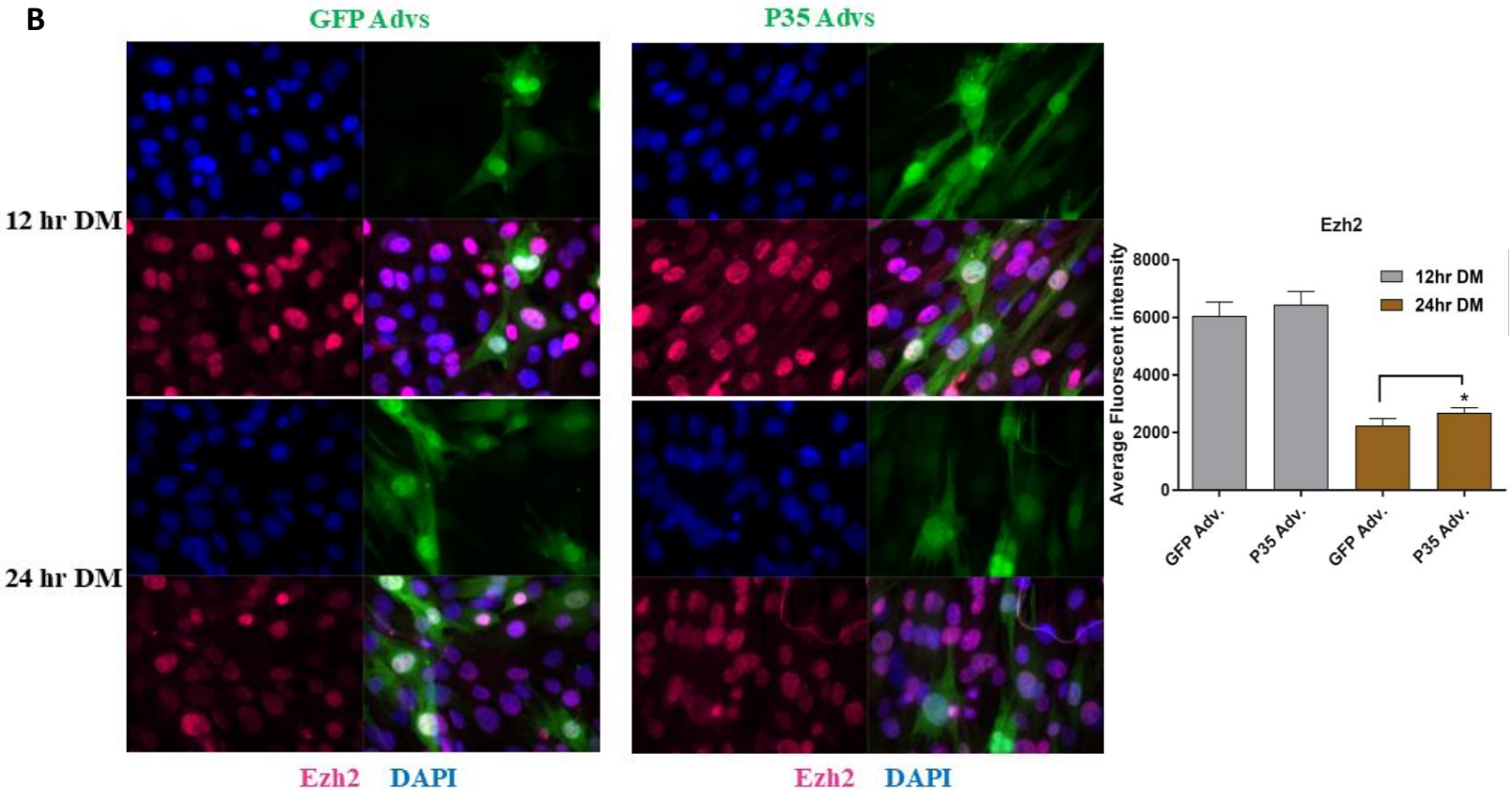


Fig. 4.6. CAD forms an immunocomplex with Ezh2 protein.
(A, B): Representative immunoblot of C2C12 cells in growth conditions (G) and during differentiation at 12, 24, 48, and 72 hours. A total of 50 µg of lysate was loaded per lane, and the blots were probed with antibodies against Ezh2, EED, and Suz12. N=2 biological replicates. (C): Immunoprecipitation of endogenous CAD from C2C12 cells at growth (G), and 12- and 24-hours post-differentiation. Immunoprecipitated samples were subjected to SDS-PAGE and immunoblotted with an anti-Ezh2 antibody. N=3 biological replicates.

4.8.Inhibition of caspas-3 intensified Ezh2 nuclear protein

To visualize Ezh2 in the nucleus of proliferating and differentiating C2C12 cells, I performed immunofluorescence staining at growth (proliferating myoblast), 24 hr (myocyte formation), 48 hr (fusion and myotube formation), and 72hr (terminal differentiation) post-differentiation. Ezh2 exhibited prominent nuclear localization at the growth stage, which decreased as differentiation progressed (Fig. 4.7A). Subsequently, I examined Ezh2 expression in C2C12 cells transduced with baculoviral p35 adenoviruses (p35 Advs.), which inhibit all caspases except caspase-9 (Vier et al, 2000). Nuclear staining of Ezh2 was observed to gradually decreased from 12 to 24 hrs (Fig. 4.7B). Inhibition of caspase-3 in cells infected with p35 intensified Ezh2 expression in 24-hour differentiated myoblasts compared to GFP (Fig. 4.7B). Given that H3K27 is a target for Ezh2 methylation, I also evaluated H3K27me3 protein levels in the nucleus. Cells expressing p35 showed intensified H3K27me3 expression in 24-hour differentiated myoblasts relative to control (Fig. 4.7C). Taken together, inhibition of caspase-3 enhances nuclear protein expression of Ezh2 and H3k27me3.


A
Growth
24hr DM
48hr DM
72hr DM
EZH2
DAPI
Merged


B
GFP Advs
P35 Advs
12 hr DM
24 hr DM
Ezh2 DAPI
Ezh2 DAPI
Ezh2
Average Fluorscent intensity
8000
6000
4000
2000
0
12hr DM
24hr DM
*
GFP Adv.
P35 Adv.
GFP Adv.
P35 Adv.

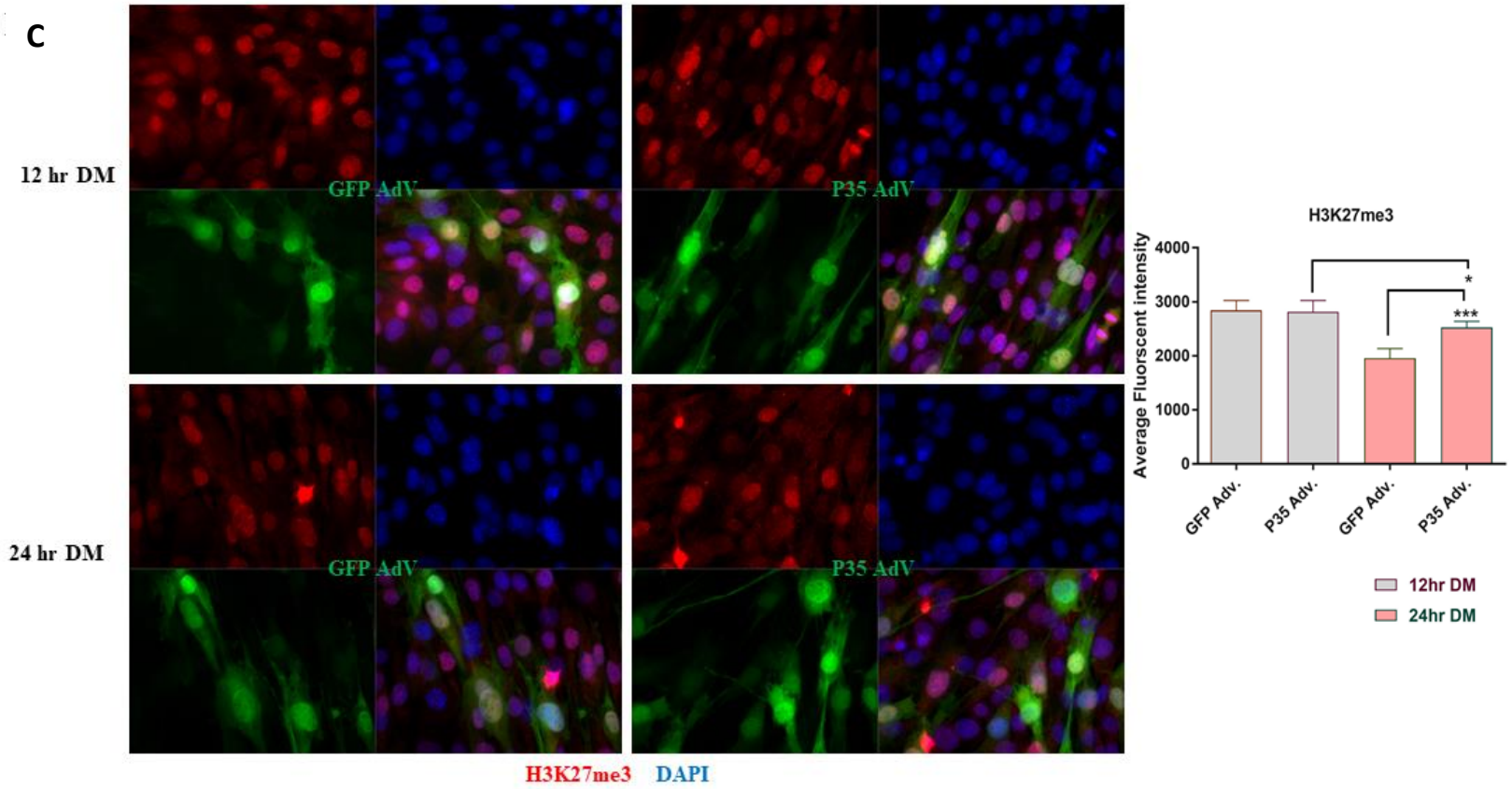


Fig. 4.7. Inhibition of caspase-3 rescued Ezh2 expression.
(A) Representative images of immunostained C2C12 cells in growth conditions and at 24-, 48-, and 72-hours post-differentiation. Cells were stained with anti-Ezh2 (Alexa Fluor 594, red) and counterstained with DAPI (blue) to visualize nuclei. (B) Representative images of immunostained C2C12 cells infected with p35 Advs or GFP Advs, followed by serum starvation for 12 or 24 hours. Cells were stained with anti-Ezh2 (Cy5) and DAPI. Fluorescence intensity was quantified based on the average Ezh2 signal in transfected cells (n=25 cells per condition) within a 63x field of view. Statistical significance was determined using one-way ANOVA; *$p < 0.05$. (C) Representative images of immunostained C2C12 cells infected with p35 Advs or GFP Advs, followed by serum starvation for 12 or 24 hours. Cells were stained for anti-H3K27me3 (Alexa Flour 594, red) and DAPI. Fluorescence intensity was quantified based on the average H3K27me3 signal in transfected cells (n=25 cells per condition) within a 63x field of view. Statistical significance was determined using one-way ANOVA; *$p < 0.05$, ***$p < 0.001$.

Summary

In this chapter, we investigated the role of the epigenetic regulator Ezh2 in muscle gene expression, focusing on its influence on chromatin dynamics around CAD sites. Our findings revealed i) a significant overlap between Ezh2 and CAD-bound genomic loci, suggesting that local chromatin architecture affects CAD accessibility, and ii) demonstrated that inhibiting caspase-3 enhances the enrichment of Ezh2 genomic targets. We also identified that caspase-3 targets and cleaves Ezh2 protein, with LC-MS/MS analysis uncovering multiple caspase-3 cleavage sites.

Caspase-3 targeting of ICAD releases CAD activity, and it is likely that caspase-3 targets Ezh2 in close temporal proximity. This is supported by our data indicating that Ezh2 and CAD may bind to the same loci in the genome. It's challenging to determine which event happens first, but it seems likely that the two processes occur in tandem. If Ezh2 methylates the sites where CAD binds, it could inhibit CAD activity. Therefore, restraining Ezh2, potentially by cleavage via caspase-3, may prevent ongoing methylation or re-methylation at these loci, and ensuring that CAD can function effectively. Collectively, these results provide new insights into Ezh2 protein stability and the caspase signaling events that coordinate the differentiation of skeletal muscle cells.

CHAPTER 5: RESULTS

CAD Protein Interactions in Differentiating Muscle Cells

CAD Protein Interactions in Differentiating Muscle Cells

Abstract

Caspase-3/CAD dependent activity has been implicated in the regulation of muscle cell differentiation. CAD activity is known to be managed by the interaction with ICAD (inhibitor of caspase activated DNase), yet it remains unknown whether CAD is targeted by other protein interactions that may influence CAD nuclease function or access to the chromatin environment in cell differentiation. To identify CAD protein-protein interactions, IP-MS was performed from early differentiating C2C12 cells using an antibody against endogenous CAD. This experimental approach revealed novel CAD protein interactors. As CAD is constitutively nuclear/peri-nuclear, we focused our analysis on the nuclear-specific protein interactions, which would inform the modulation of CAD function more accurately than cytosolic protein-protein interactions. Within this group of prospective nuclear CAD interactions, we noted a cohort of proteins that are known to manage the DNA repair process, including replication protein A (RPA) RPA1, RPA2, and XRCC1, all of which are implicated in SSBs and DSBs repair and base excision repair (BER). Prior observations in the Megeney lab had identified XRCC1 as a critical component of the CAD strand break induced DNA repair process. This suggests that CAD may be associated with DNA repair machinery to ensure the eventual resolution of strand breaks and mitigate the consequences of unrepaired DNA damage. In addition to DNA repair machinery, CAD bound other chromatin modifying enzymes, including helicases, RNA splicing ligase, DNA binding RNA polymerase, and topoisomerase IIIB (TopIIIβ). Topoisomerases are enzymes that relax supercoiling and torsional stress by altering the topological state. They have been implicated in formation of strand breaks that give rise to gene breakpoints prone to form human oncogenic translocations. Therefore, we reasoned that Topo IIIβ may act in concert with CAD to alter the local DNA topological state

and allow CAD to access genomic targets. As a first step to address this hypothesis, we performed CUT&Tag with an antibody directed against endogenous Topo IIIβ to map and compare the Topo IIIβ targeted genome versus the CAD targeted genome. The comparative analysis revealed a high overlap between Topo IIIβ, and CAD bound sites, where 74% of Topo IIIβ sites are shared with CAD bound sites, and 54% of CAD regions are shared with Topo IIIβ in differentiating C2C12 muscle cells. Taken together, these preliminary observations imply that CAD may exist within a larger protein complex, that may act to modify DNA ultrastructure for nuclease activity/access and ensure effective DNA repair immediately following CAD induced DNA breaks.

Background

CAD activity is restricted by an interaction with its inhibitory protein ICAD, which is inactivated by caspase-3 cleavage. In apoptosis, elevated caspase-3 activity leads to widespread DNA damage, where release of caspase directed cleavage of ICAD frees CAD to initiate DNA fragmentation (Sakahira et al, 1998). However, during myogenic differentiation, transient caspase-3 activity cleaves ICAD in a temporally restricted manner, leading to limited DNA strand break formation, which promotes differentiation (Larsen et al, 2010, Larsen & Megeney, 2010). In apoptotic cells, studies have identified different protein-protein interactions that enhance CAD, including histone H1 and all its variants, HMG-box-containing chromatin proteins (HMGB) 1 and 2, Histone H2AX, heat-shock protein (HSP) 70, and topoisomerase IIα (Samejima et al 2001, Gu et al, 1999, McCarty et al, 1999, and reviewed in Kulbay, et al 2021). CAD and DNA topoisomerases are both enzymes that can affect the topology of the DNA. For instance, DNA Topoisomerase II alpha (TopIIα) was shown to bind to CAD, and both cooperate to condense chromatin in the execution phase of apoptosis, which is initiated by cleavage of caspase-3 (Gómez-Herreros et al, 2019). Inhibition of CAD and TopIIα, at the same time terminate the apoptotic chromatin condensation. Therefore, the

presence of either CAD or TopIIα was adequate to condense chromatin (Durrieu et al, 2000), but not to induce strand breaks *per se*.

The DNA damage induced by CAD suggests that DNA repair proteins may reside in close proximity to repair CAD sites and limit the genome toxicity of sustained DNA damage. Indeed, as cells are continually exposed to exogenous or endogenous genotoxic stress, they employ a wide variety of surveillance mechanisms known collectively as a DNA damage response (DDR). Initially, DDR protein complexes sense DNA damage, activating ATM and ATR kinases (Maréchal et al. 2013). Subsequently, an effector protein, such as p53, is activated to mediate the cell response, which involves arresting cell cycle progression to allow repair of the DNA damage (Serrano et al, 2013).

Concurrent to p53 mediated cell cycle withdrawal, the recruitment of ATM & ATR leads to directed phosphorylation of the gH2AX variant, which spreads along the DNA strand to stabilize recruitment of additional repair proteins (d'Adda di Fagagna, 2008). One of these targets, replication protein A (RPA), binds and stabilizes ssDNA to prevent it from degradation or formation of a secondary structure, and aids further recruitment and assembly of proteins involved in DNA replication, repair and recombination (Dueva et al,2020). Interestingly, observations from the Fearnhead group revealed DNA-PK, not ATM nor ATR, targeted gH2AX phosphorylation during caspase 3 induced muscle cell differentiation, suggesting that this kinase may manage the DNA repair process following CAD induced damage (Connolly& Fearnhead, 2016). Earlier work from our laboratory observed extensive DNA damage /transient DNA strand breaks during early stages of muscle cell differentiation (Larsen et al,2010). A subsequent study from our laboratory identified XRCC1 as a repair factor for Caspase-3/CAD dependent myogenic differentiation (AL-

Khalaf et al, 2016). However, it is reasonable to hypothesize that other proteins are involved in the repair process of CAD breaks.

To explore this, and to enhance our understanding of CAD regulation, we characterized CAD protein interactions in differentiating muscle cells using mass spectrometry-based proteomics. The results revealed that CAD binds to a large number of proteins, suggesting, though not proving, that CAD may function within a larger protein complex. Specifically, we investigated whether the known DNA strand break repair factors, RPA1/2 participate in a CAD-centric damage/repair complex by examining the overlap (if any) in their respective genomic targets.

In addition, another prospective CAD protein interaction of high interest was TopIIIβ. A previous study reported that the abortive activity of topoisomerase II (TOP2) induced chromosomal translocations (Canela et al., 2019), an activity that we documented for CAD itself (Chapter 3). This led us to hypothesize that CAD might work in cooperation with TopIIIβ to facilitate CAD access to its genomic targets.

Results

5.1.ICAD/CAD Protein Localization in Differentiating Muscle Cells

To study CAD protein interactions, I first sought to test CAD within the subcellular localization. Given that CAD is constitutively bound in the nucleus, yet ICAD is mostly bound to CAD in a dimer complex acting as a chaperone to fold CAD properly to be catalytically competent nuclease (Sakahira et al, 1999). Therefore, it is essential to test the colocalization of both proteins. Here, we used the overexpression system of adenoviruses (Advs.) for gene delivery expressing ICAD/CAD. We introduced a FLAG-tag on the regulatory domain N-terminal of CAD Advs. and a HA tag on the N-terminal of ICAD Advs. Following that, C2C12 myoblasts were transfected with these Advs. and fixed at 24 hours of differentiation. immunohistochemistry of the HA and Flag tags confirmed the overexpression (Fig. 5.1). Since ICAD is an essential chaperone of CAD, I optimized the ratios of co-expression of ICAD:CAD ratio and with different multiplicity of infection (MOI). I found that 2:1 ratio of ICAD:CAD with 450 MOI is optimal to infect efficiently more cells, with minimal cytotoxicity (Fig. 5.1). We observed that the localization of CAD is predominantly in the nucleus while ICAD is localized to the cytoplasmic compartment. In addition, ICAD/CAD complex appears to be localized in the perinuclear at 12 and 24 hours of differentiation, suggesting ready and immediate access of CAD to nuclear targets.

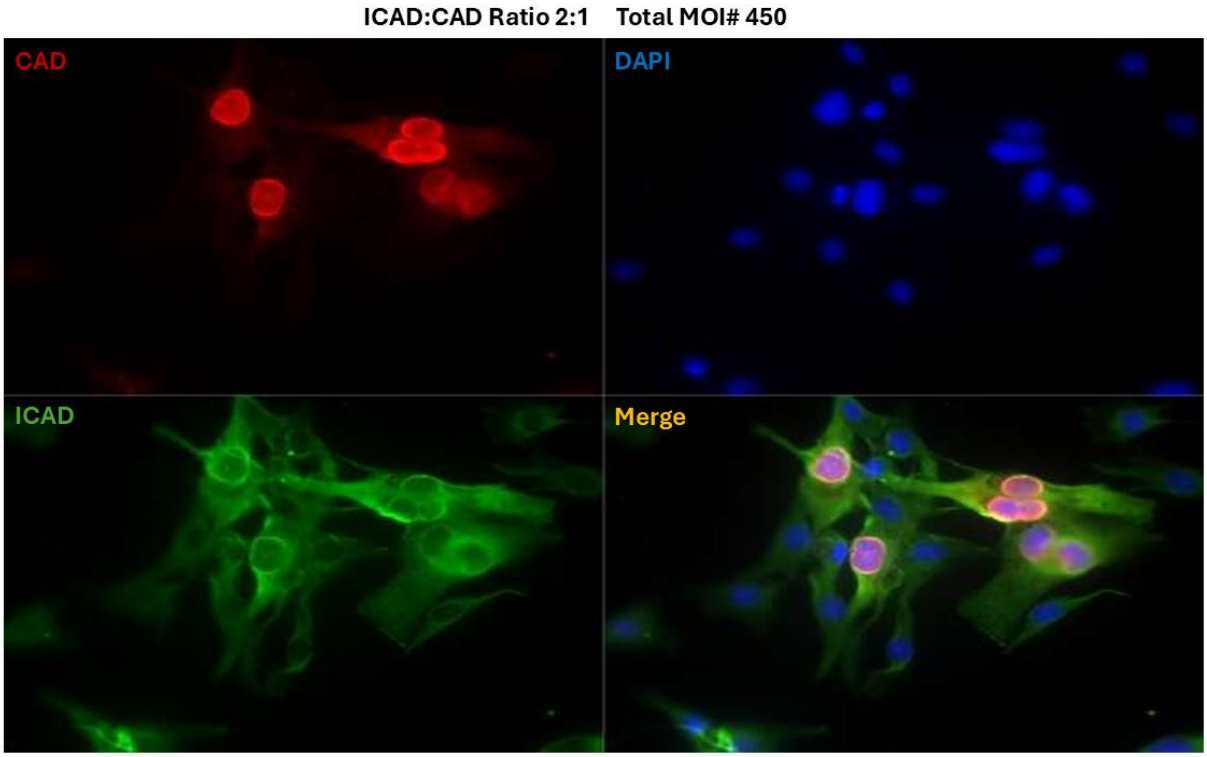


Fig. 5.1. Colocalization of ICAD/CAD in differentiating skeletal muscle C2C12 cells.
Representative image for C2C12 cells infected with co-expression of CAD-FLAG and ICAD-HA adenoviruses and induced for differentiation (24 hr). Cells then were immunostained with anti-HA (green), anti-FLAG (red), and counterstained with DAPI. MOI: Multiplicity of infection (MOI).

5.2.Identification of Novel CAD Interacting Proteins

To explore CAD protein interactors, I performed immunoprecipitation of endogenous CAD derived from lysates of 18 hours differentiated C2C12 cells. This time was selected as its likely to capture the peak of CAD activity and its interactions with protein partners, giving that CAD is active between 12-24 hours. Samples were subjected to SDS-PAGE and silver stain followed by LC-MS/MS (Fig.5.2A) (Orbitrap Fusion Lumos, OHRI Proteomics Core Facility). Total of 794 proteins were identified using scaffold software, with a 99% protein threshold and 95% peptide threshold. This set of proteins may associate with cytosolic function, that has been reported for CIDE related proteins, or lipid metabolism, which appears entirely distinct from CAD's role as a nuclease (Park, 2015). No interaction was detected between CAD and ICAD, which is expected given that ICAD is cleaved by caspase 3 during differentiation, destabilizing any CAD/ICAD interaction (Larsen et al. 2010).

Given that CAD is predominantly a nuclear protein in differentiating muscle cells, we curated a list of nuclear resident proteins bound to CAD. We used Cystoscape and STRING Protein-Protein Interactions (PPI) to visualize this interaction data as represented in (Fig.5.2B, C). CAD interactome proteins were subjected to GO terms and KEGG category analysis as shown in (Fig.5.2E, F). The corresponding top enriched terms are identified for the regulation of DNA recombination, DNA replication, base excision repair, mismatch repair, DNA packaging, chromosome condensation, chromatin remodeling, structural constituents of chromatin, and histone modification. Importantly, our results show that CAD binds to other caspase-3 targets previously identified, including casein kinase II (Fernando et al, 2002), gelsolin (Putinski et al, 2018), XRCC1 (AL-Khalaf et al, 2016), TDP43 and Prohibitin (Brunette et al. 2023) (Fig.5.2.C). This may suggest that caspase-3 targets could be in close proximity within the nucleus, and may

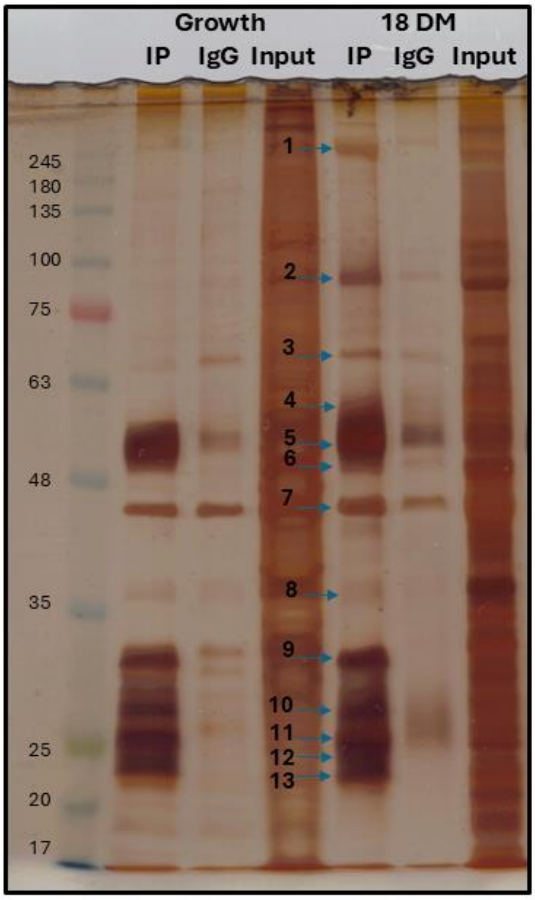
exist as part of larger protein complex providing new mechanistic and insights into the regulation of this crucial nuclease.

5.3.CAD Bind DNA Break/Repair Protein Complex

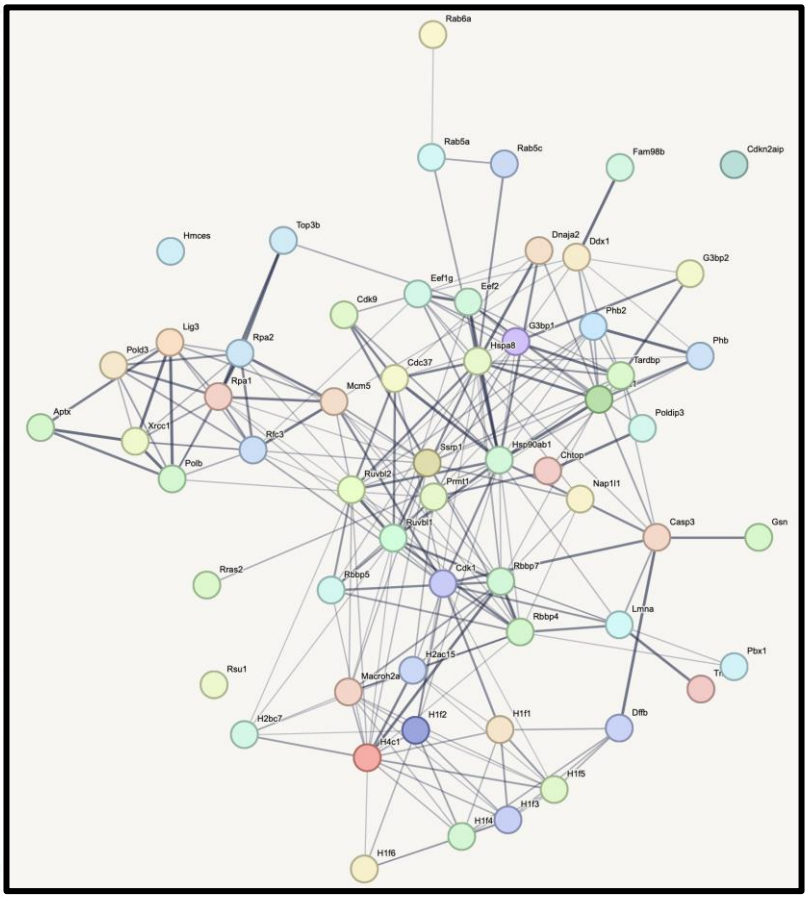
Among the potential CAD interactome, I identified a number of DNA repair proteins bound to CAD including RPA1, RPA2, XRCC1, RuvB like AAA ATPase 1,2(Ruvbl1, Ruvbl2), aprataxin (Aptx), and 5-Hydroxymethylcytosine Binding, ES Cell-Specific (Hmces) (Fig.5.2.C,D). Identification of XRCC1 is consistent with the previously identified role of XRCC1 as a DNA repair member to resolve CAD mediated DNA breaks in differentiating myoblasts (Al-Khalaf et al. 2016). Additionally, a CAD interaction was noted with RPA1/2, which are DNA-dependent protein kinases (DNA-Pks). Interestingly, these kinases have been shown to phosphorylate DNA-PK, which is required for the differentiation of myoblasts (Connolly and Fearnhead 2016).

Furthermore, among the list of DNA repair proteins, I noted a CAD interaction with Aptx, Ruvbl1,2, and Hmces proteins. Aptx primarily repairs DNA SSBs (Fig.5.2.C,D) (Mosesso et al, 2005); conversely, Ruvbl1,2 repairs DNA DSBs which is crucial for functioning of the homologous recombination (HR) and non-homologous end-joining (NHEJ) pathways (Tian et al, 2024). In addition, Hmces has a DNA repair role in BER as it binds and protect abasic sites to prevent SSBs from forming DSBs (Thompson et al, 2020). To the best of our knowledge, this is the first report in the literature to suggest a direct interaction of DNA repair complex proteins with the CAD nuclease. A reasonable interpretation of these data suggest that CAD may co-exist within a larger protein complex comprised of DNA repair factors to ensure that CAD mediated DNA strand breaks are rapidly resolved. This concept is consistent with the function of this nuclease to induce transient genome wide breaks, to promote a cell fate change without a full dissolution of the genome, that is a characteristic of apoptosis *per se*.

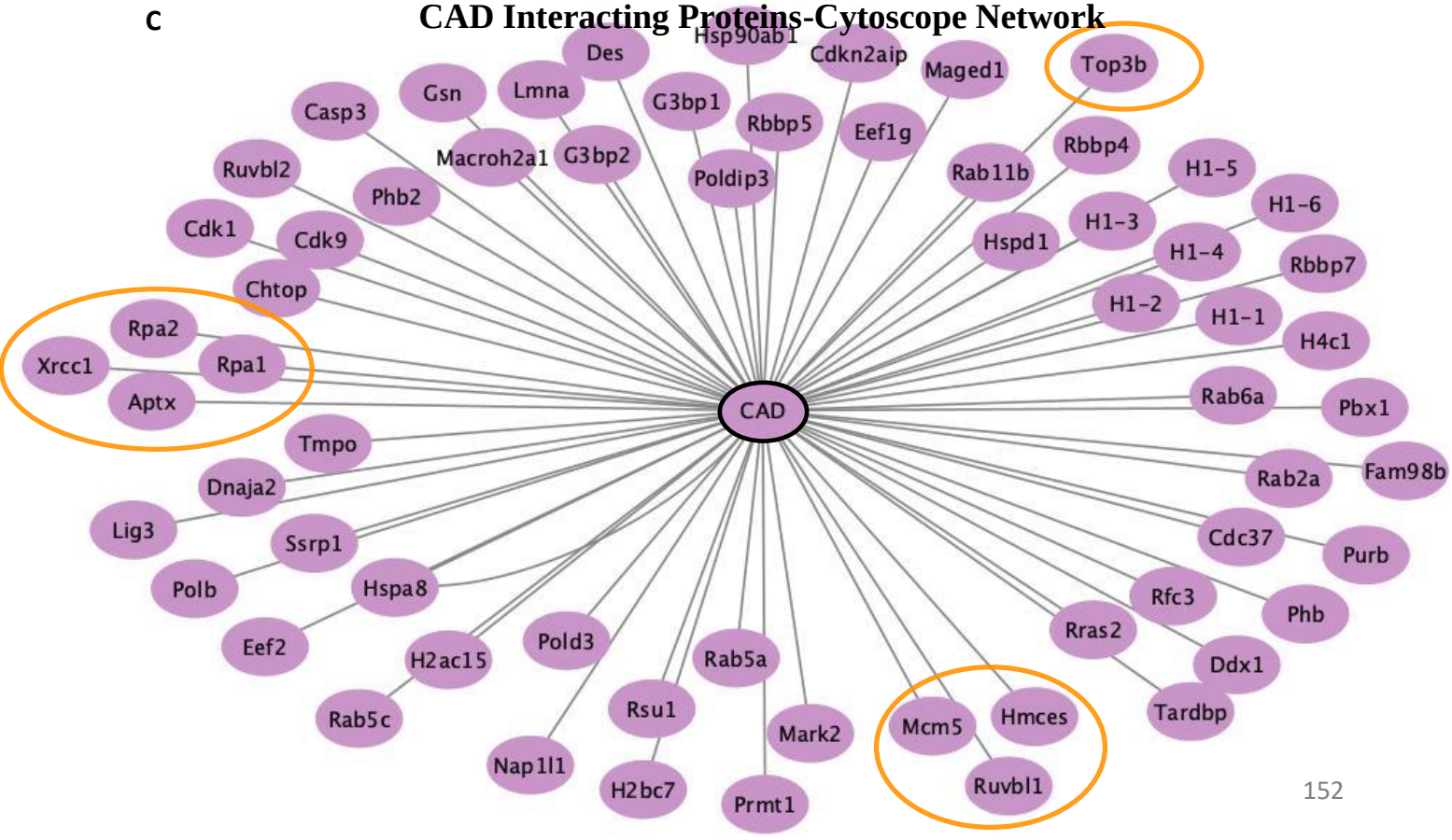
A



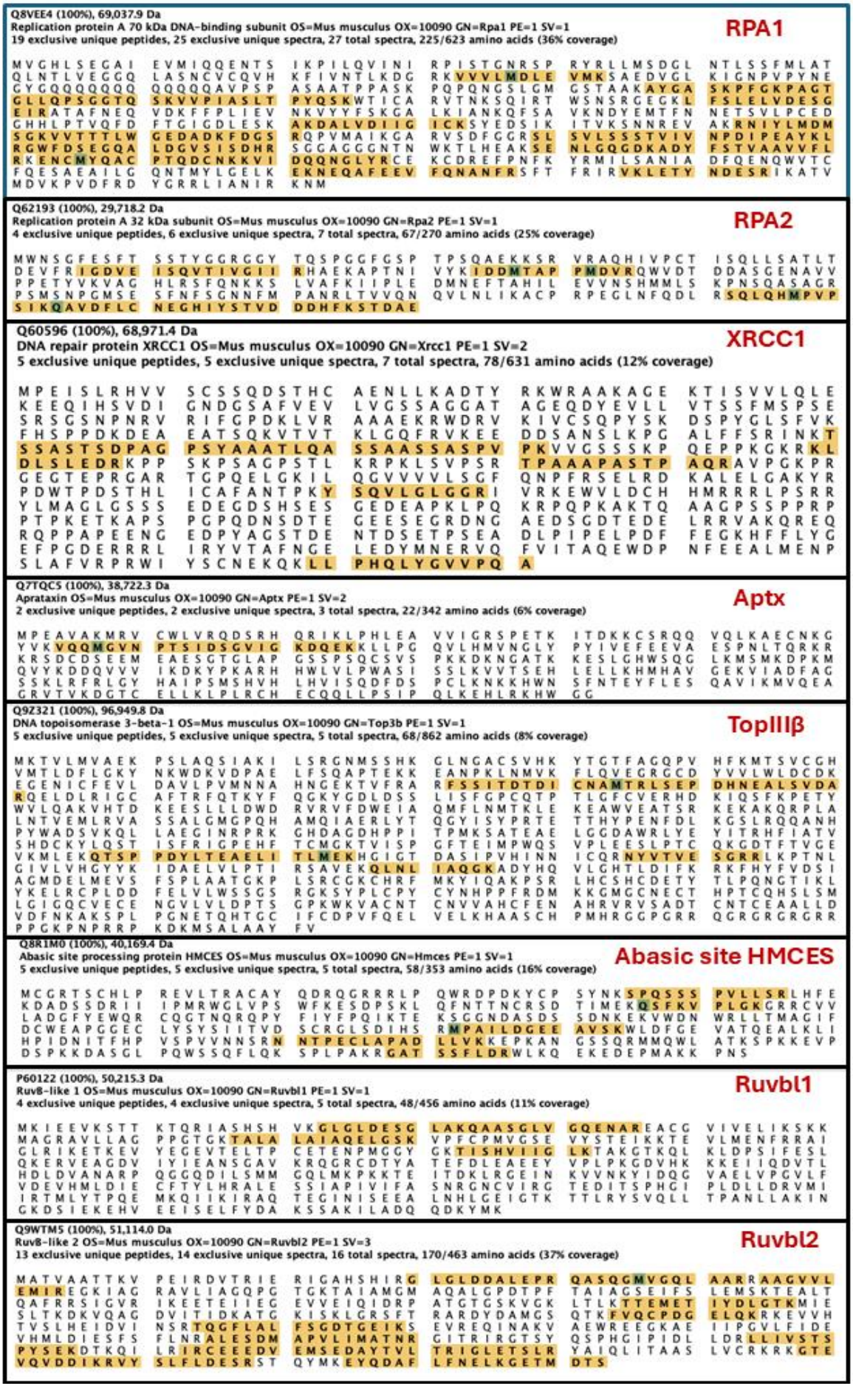
B **CAD Interacting Proteins-STRING Network**



C **CAD Interacting Proteins-Cytoscope Network**

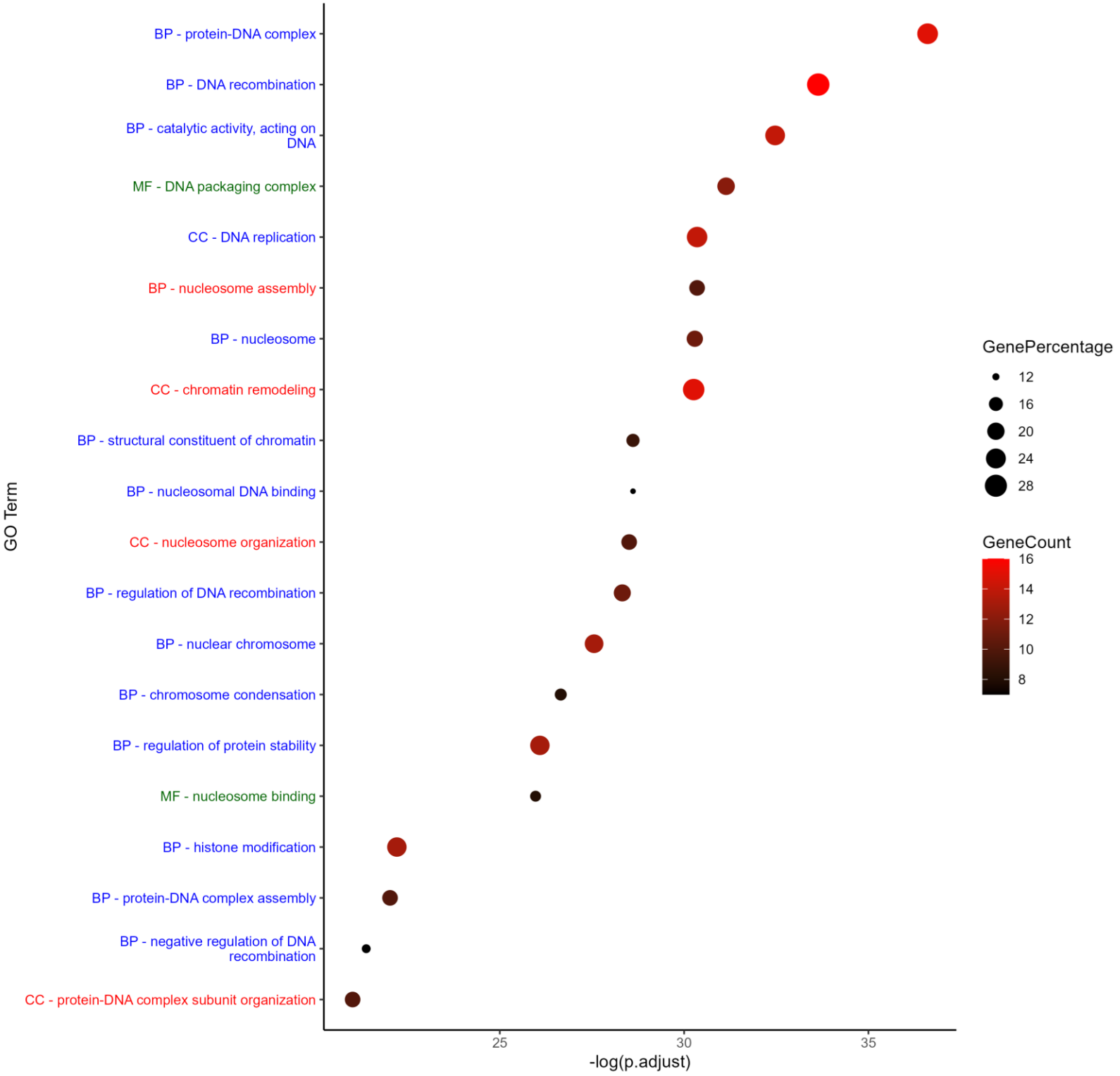


Proteins/Peptides of potential CAD Protein Interactors



E

Gene Ontology (GO) Analysis Identified within CAD-Interacting Proteins



F **KEGG Pathway Analysis Identified within CAD-Interacting Proteins**

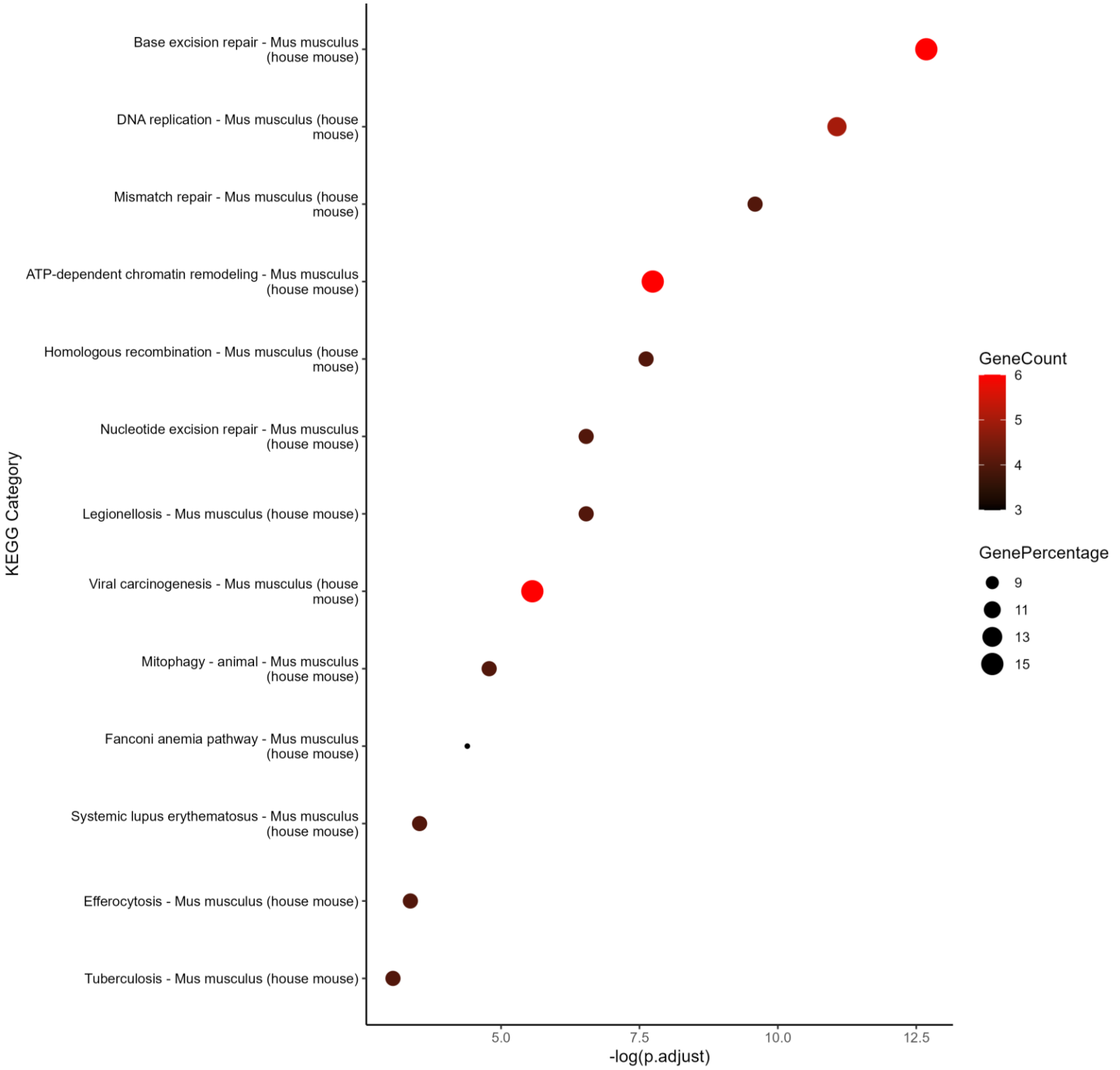


Fig. 5.2. CAD-interacting proteins, identified by MS from C2C12 cells.
(A) Differentiating C2C12 cells (18 hr) were homogenized, and immunoprecipitated against endogenous CAD, and samples were subjected to SDS-PAGE and silver stain. Bands with indicated arrows from IP lane were excised and subjected for LC-MS/MS. (B) Representation of CAD interactors using STRING Protein-Protein Interaction Network. Network nodes (colored circles) represent proteins. Lines between the nodes (edges) indicate the interaction between proteins. (C) Nuclear proteins interacting with CAD are displayed by Cytoscope. Network nodes interact directly with CAD. (D) Example of peptides coverage for some of CAD binding proteins. Analysis of identified proteins were generated using Scaffold software. The peptides highlighted in yellow are detected via LC-MS/MS for the corresponding protein. (E) Gene Ontology (GO) analysis identified within CAD interacting proteins. Top 15 terms are enriched. Biological Processes (BP), Molecular Functions (MF), Cellular Components (CC). (F) KEGG analysis identified within CAD interacting proteins. 15 terms are enriched.

5.4.CAD Interacts with Enzymes that Regulate Replication, Transcription and Control of DNA Topology

Within the broader category of typical nuclear resident interacting proteins, CAD interactions were observed with helicases, ligases, isomerases, DNA-binding RNA polymerase, and RNA splicing ligase (Fig. 5.2C). For instance, helicases are involved in DNA replication and transcription machinery (Brosh et al, 2013), while ligases catalyze the formation of a phosphodiester bond between DNA fragments crucial for sealing the DNA breaks during repair, for both SSBs and DSBs (Timson et al, 2000). RNA splicing ligase catalyzes the splicing and the subsequent ligation of RNA molecules, which is important for various processes, including gene expression regulation (Shkreta & Chabot, 2015). In addition, CAD bound to minichromosome maintenance complex component 5 (MCM5), a part of MCM hexameric complex (Slaymaker et al, 2013). This complex functions as a DNA helicase within the replication machinery, regulating DNA replication timing and initiation during the cell cycle, and interacts with DNA repair proteins to coordinate the response to replication stress (Yadav et al, 2024, Wei et al, 2016).

One of the most compelling enzymes identified in the CAD interactome was TopIIIβ (Fig. 5.2C, D). Topoisomerases resolve the topological strain that arises during transcription and replication by unwinding tangled DNA (Pommier et al, 2016). However, earlier studies have reported that topoisomerase II inflicted DNA strand breaks to regulate transcription. For example, TopIIβ induced promoter targeted strand breaks to initiate transcription in androgen-mediated signaling in prostate cancer cells (Ju et al. 2006; Haffner et al, 2010). Similarly, TopIIβ has been observed to be recruited to the promoter of glucocorticoid responsive genes (Ju et al, 2006). Given these observations, it is reasonable to suggest that TopIIIβ and CAD may have a coordinated function,

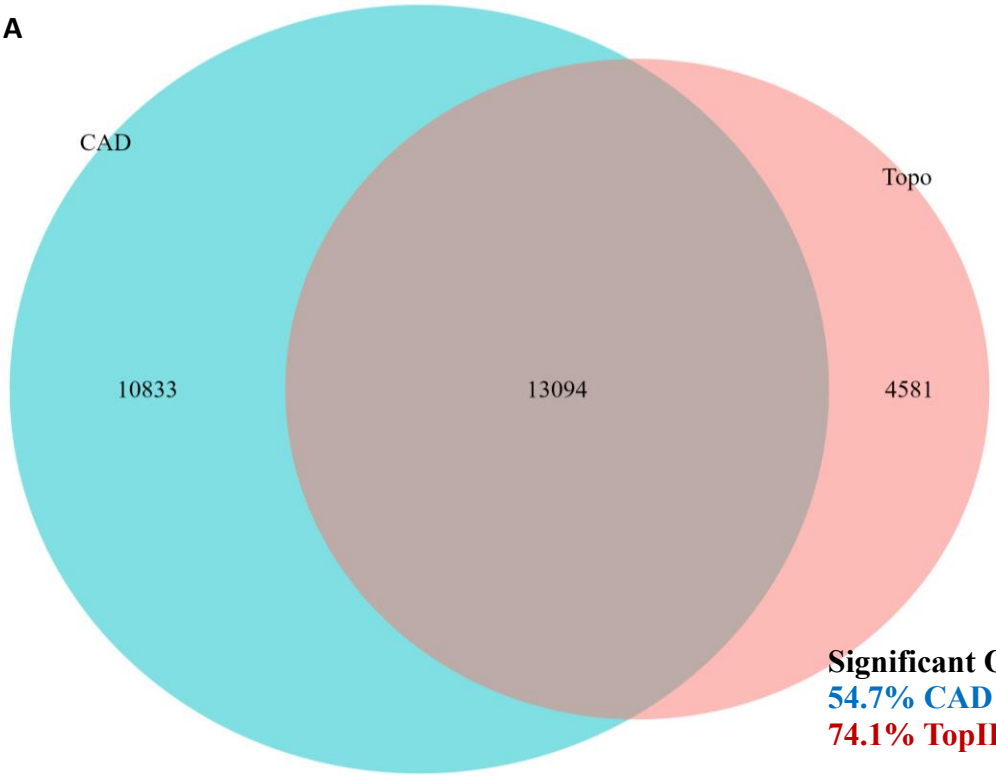
where the isomerase alters the local DNA structure to allow for CAD activity/access to genomic targets.

5.5.Genomic Target of CAD Overlap with TopIIIβ

Following the mass spectrometry data that identified TopIIIβ as a CAD-interacting protein (Fig. 5.2D), I aimed to further investigate this interaction and determine whether there is coordinated targeting of the same genomic regions. To do this, I performed CUT and Tag using an antibody against endogenous TopIIIβ in C2C12 treated with z.DEVD-fmk. The analysis revealed significant genomic overlap between TopIIIβ and CAD, with 74% of TopIIIβ regions shared with CAD, and 54% of CAD sites overlap with TopIIIβ (Fig. 5.3A). Functional annotation indicated that the shared targets between CAD and TopIIIβ are predominantly located in promoters <=1kb, followed by distal intergenic regions (Fig.5.3B). Both plotHeatmap and plotProfile analysis demonstrated enrichment signals for CAD and TopIIIβ binding sites near transcription start sites (TSS), and CAD shared peak over the genome with TopIIIβ (Fig. 5.3C-E). The extensive overlap suggests (but does not prove) that TopIIIβ may be essential to unwind the DNA strand for CAD to inflict strand breaks. Further examination of shared genes between TopIIIβ and CAD revealed that both proteins bind to other enzymes of topoisomerases family, including Top1, TopIIα, TopII β, TopIIIa (Fig.5.3F). Shared targets between CAD and TopIIIβ were subjected to KEGG and GO analysis, and the identified terms are indicted in (Fig.5.3G, H). These targets were enriched in the common signaling pathways such as Wnt, Hippo, AMPK signaling, and cell cycle regulation and transition. In addition, both CAD and TopIIIβ were enriched in cancer such as hepatocellular, pancreatic carcinoma, and in chronic myeloid leukemia, which is reminiscent of the CAD targeting in the breakpoint biology, described in Chapter 3.

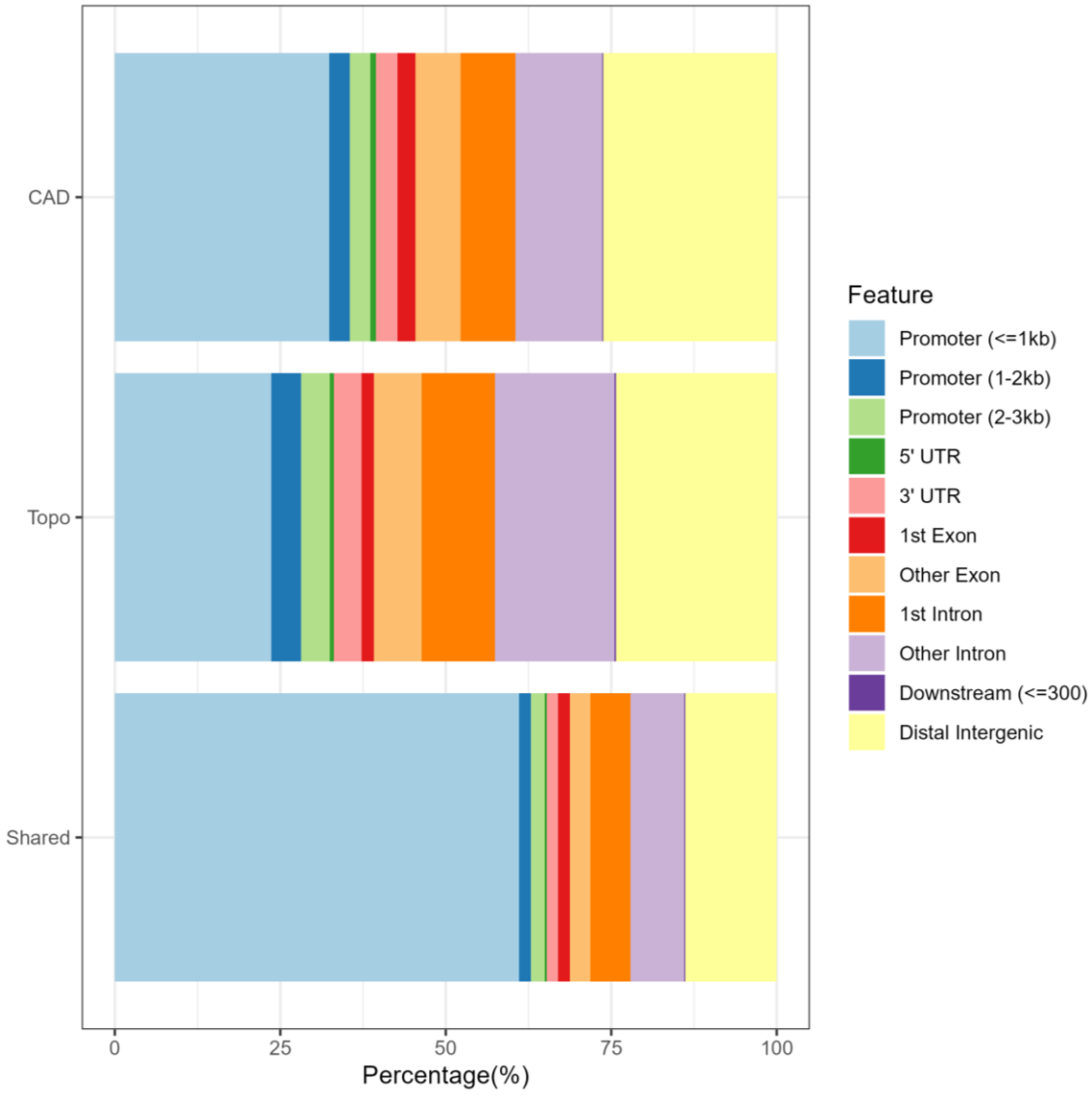
Genomic Distribution of CAD and TopIIIβ Binding Sites

A

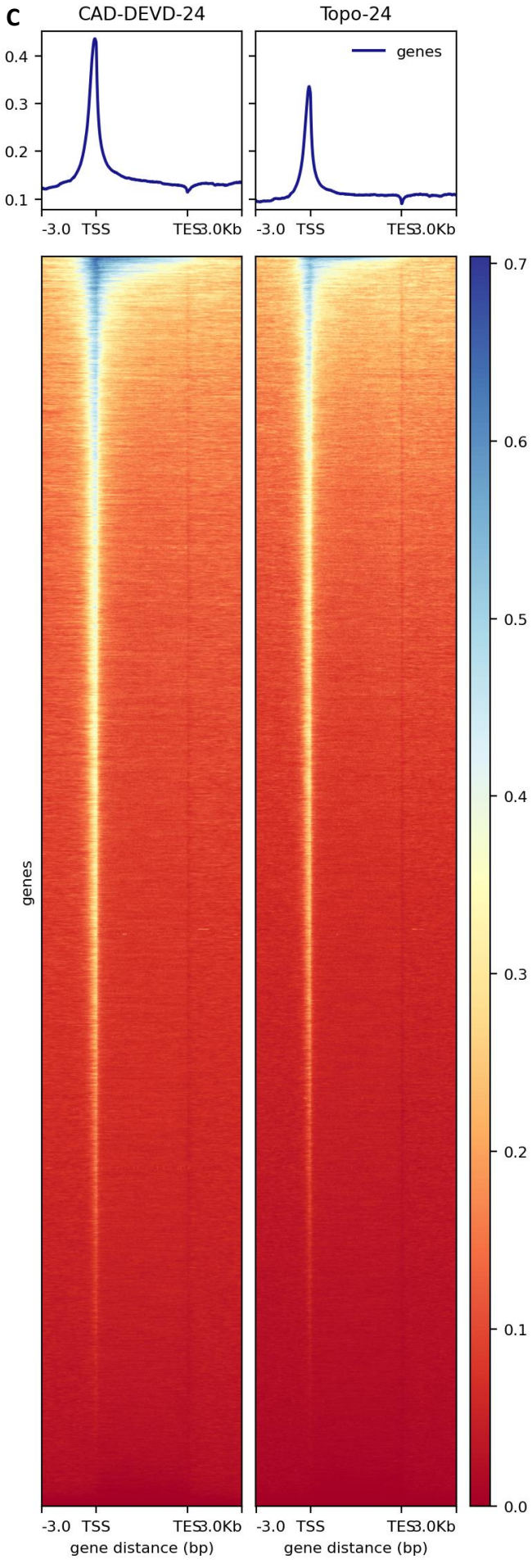


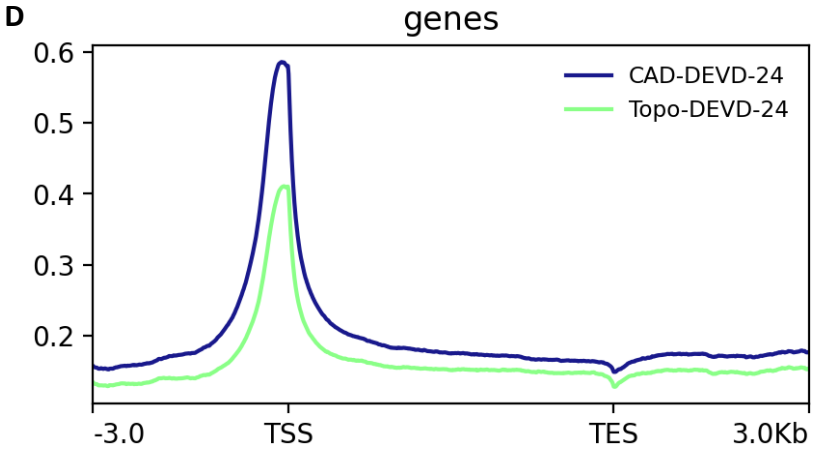
Significant Overlap between CAD and TopIIIβ
54.7% CAD sites are shared with TopIIIβ
74.1% TopIIIβ sites are shared with CAD

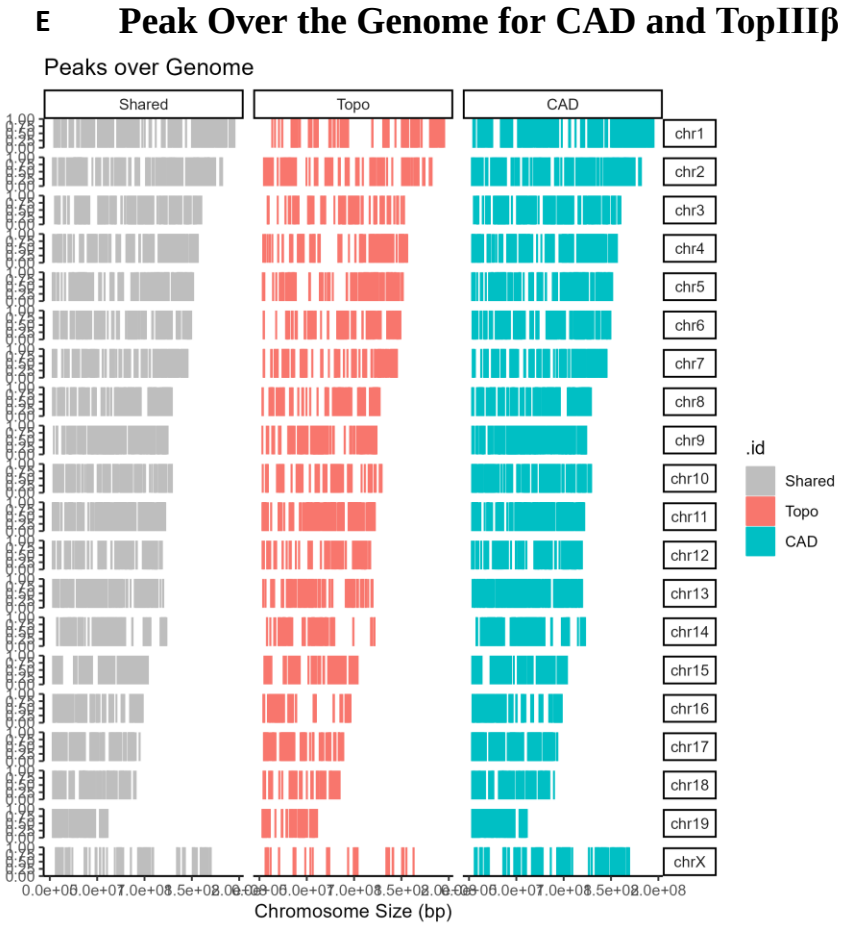
B



Genomic Distribution of CAD and TopIIIβ Binding Sites

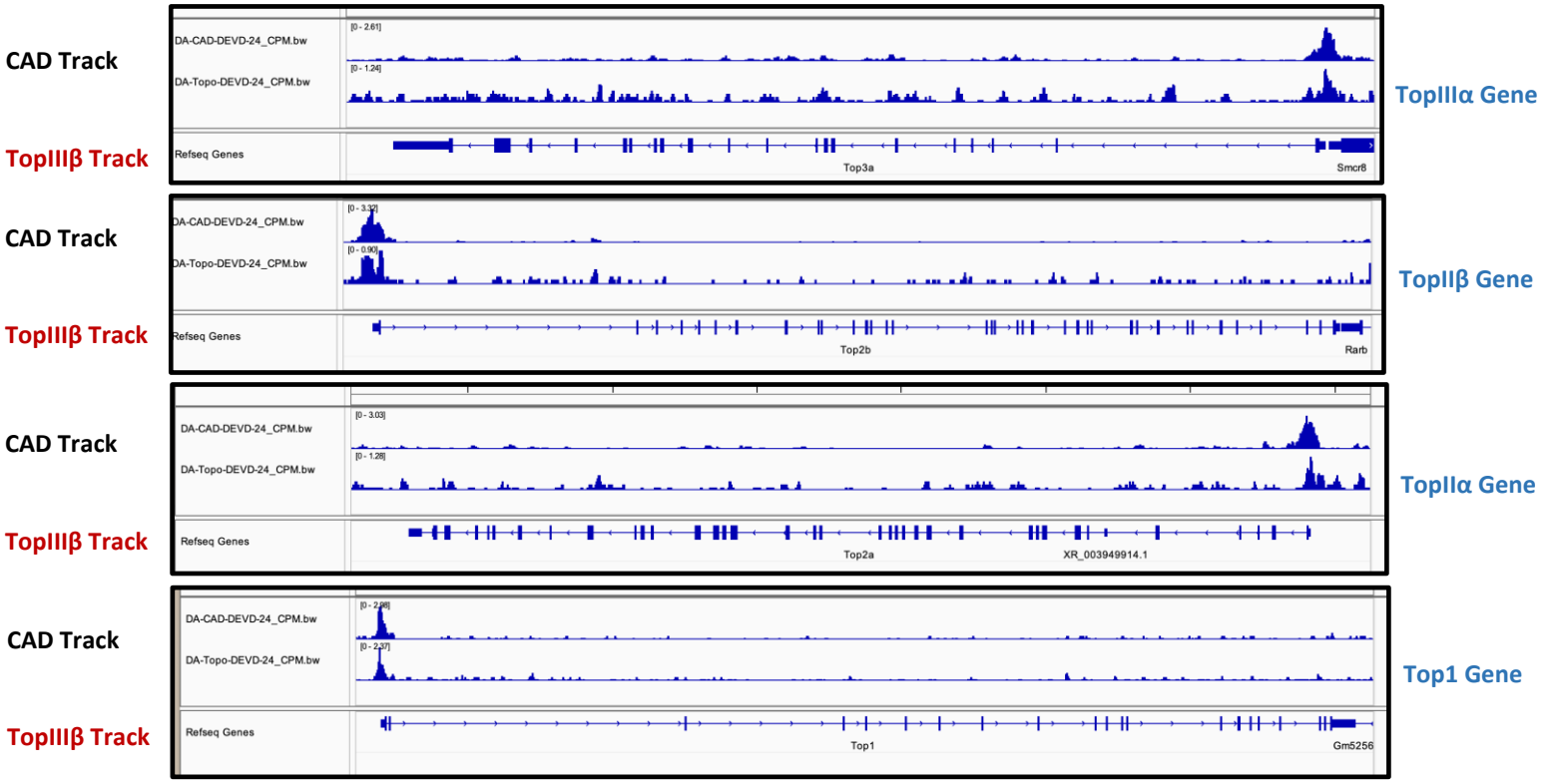






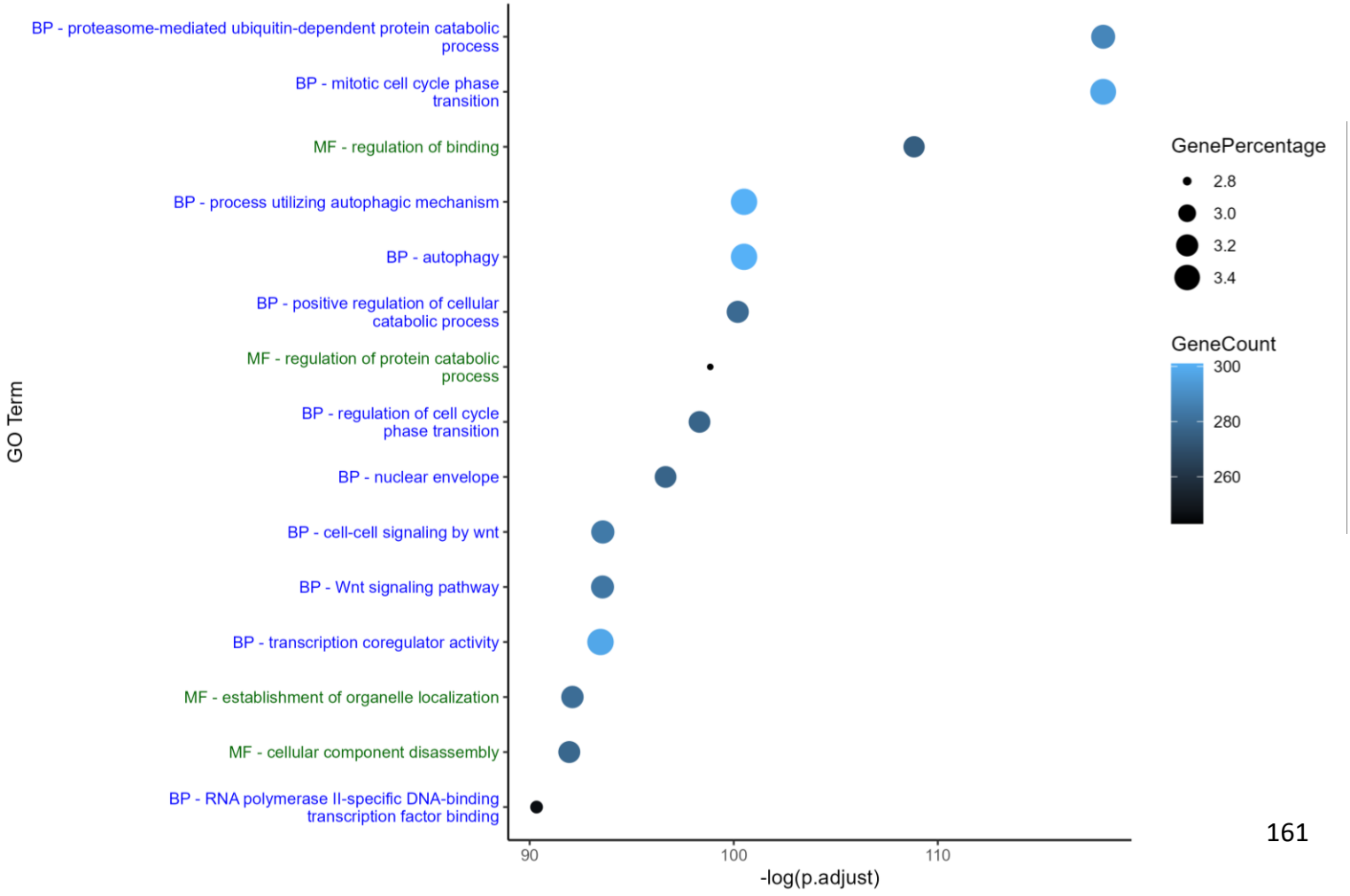
F

CAD and TopIIIβ Overlap on other Topoisomerases Gene Loci's



G

Gene Ontology (GO) Analysis Identified Shared Targets between CAD and TopIIIβ



H **KEGG Pathway Analysis Identified Shared Targets between CAD and TopIIIβ**

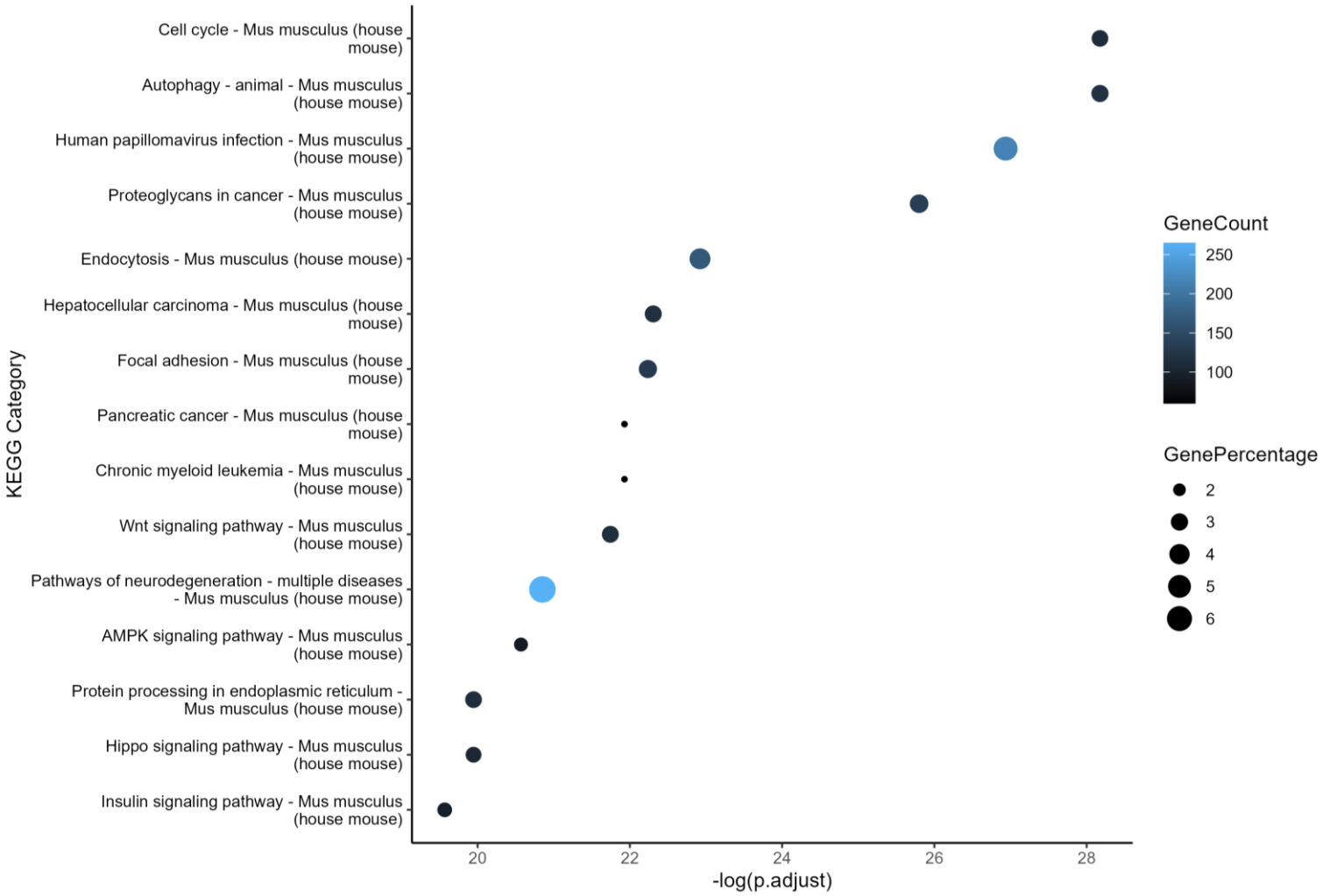


Fig. 5.3. Genomic distribution of CAD and TopIIIβ binding sites.
(A) Venn diagram illustrating binding site overlap among functional gene lists from peak analysis between CAD and TopoIIIβ. (B) Genomic region distribution in the peaks falling in various categories displayed. (C) plotHeatmaps of CUT and Tag show count per million (CPM) normalized signals for reads across 3,000 bp genomic interval relative to the TSS. Genomic distribution of the binding sites of CAD and TopoIIIβ at 24 hr post differentiation from C2C12 muscle cells. Data sets were generated from merged two independent biological replicates for each condition. Heatmaps were plotted using computeMatrix and plotHeatmap in deepTools. Peaks were called and ordered for each dataset using the consensus of MACS, SEACR, and GoPeaks. (D) PlotProfile is the averaged profiles of the compressed matrix of occupancy/signal near TSS produced by computeMatrix. Peak count frequency over genomic regions for CAD and TopoIIIβ. (E) Peaks over genome represented between CAD and TopoIIIβ. Created with ChIP seeker, covplot function. (F) Integrative genome browser view (IGV) of genes binding loci to CAD. Peaks were called and ordered for each dataset using the consensus of MACS, SEACR, and GoPeaks. Data sets were generated from merged two independent biological replicates. Tracks shown are for CAD and TopIIIB from 24 hr differentiation C2C12 cells. (G) Gene Ontology (GO) analysis identified within the shared genes between CAD and TopoIIIβ. Top 15 terms are enriched. Biological Processes (BP), Molecular Functions (MF), Cellular Components (CC). (H) KEGG analysis identified within the shared genes between CAD and TopoIIIβ. Top 15 terms are enriched.

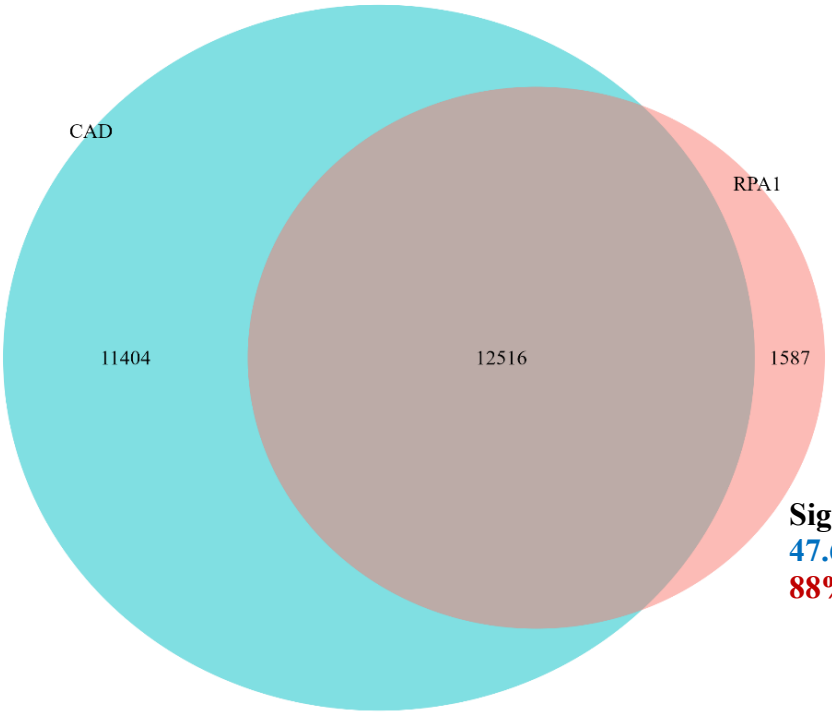
5.6.Genomic Target of CAD Overlap with RPA1/2

Following mass spectrometry data that identified RPA1 and RPA2 proteins as CAD partners (Fig. 5.2C, D), I conducted CUT and Tag analysis using antibodies against endogenous RPA1 and RPA2 in C2C12 treated with z.DEVD-fmk. during murine myoblast differentiation to explore this interaction at the genomic level. Interestingly, 88% of RPA1 binding sites are shared with CAD, while 47.6% of CAD binding sites shared with RPA1, indicating a significant overlap (Fig. 5.4A). On the other hand, 58% of RPA2 sites are shared with CAD, while 35% of CAD sites are overlap with RPA2 (Fig. 5.5A) These findings suggest that CAD genome targets have a greater overlap with RPA1, potentially indicating a preferred role for RPA1 in differentiation-associated DNA repair responses. Functional annotation revealed that the shared targets between CAD and RPA1, and CAD and RPA2, are predominantly located in promoters <=1kb (Fig.5.4B; Fig. 5.5B). Both plotHeatmap and plotProfile analysis demonstrated enrichment signals for CAD and RPA1, and CAD and RPA2 binding sites near transcription start sites (TSS) (Fig.5.4C, D; Fig. 5.5C, D). Additionally, CAD shared peak over the genome with RPA1 and RPA2 (Fig. 5.4E; Fig. 5.5E).

To gain an overview of the binding sites, GO and KEGG analysis were performed for the shared targets between CAD and RPA1 and CAD and RPA2. Similar and distinct pathways were observed for CAD and RPA1 and CAD and RPA2. Both targets were enriched in pathways such as ubiquitin proteasome-mediated processes, cell cycle regulation, and Wnt and Hippo signaling (Fig. 5.4F, G; (Fig. 5.5F, G). However, distinct pathway enrichments were noted: CAD and RPA1 targets were associated with terms related to hepatocellular carcinoma, pancreatic cancer, and chronic myeloid leukemia, while CAD and RPA2 targets were linked to small cell lung cancer (Fig. 5.4F, G; Fig. 5.5F, G).

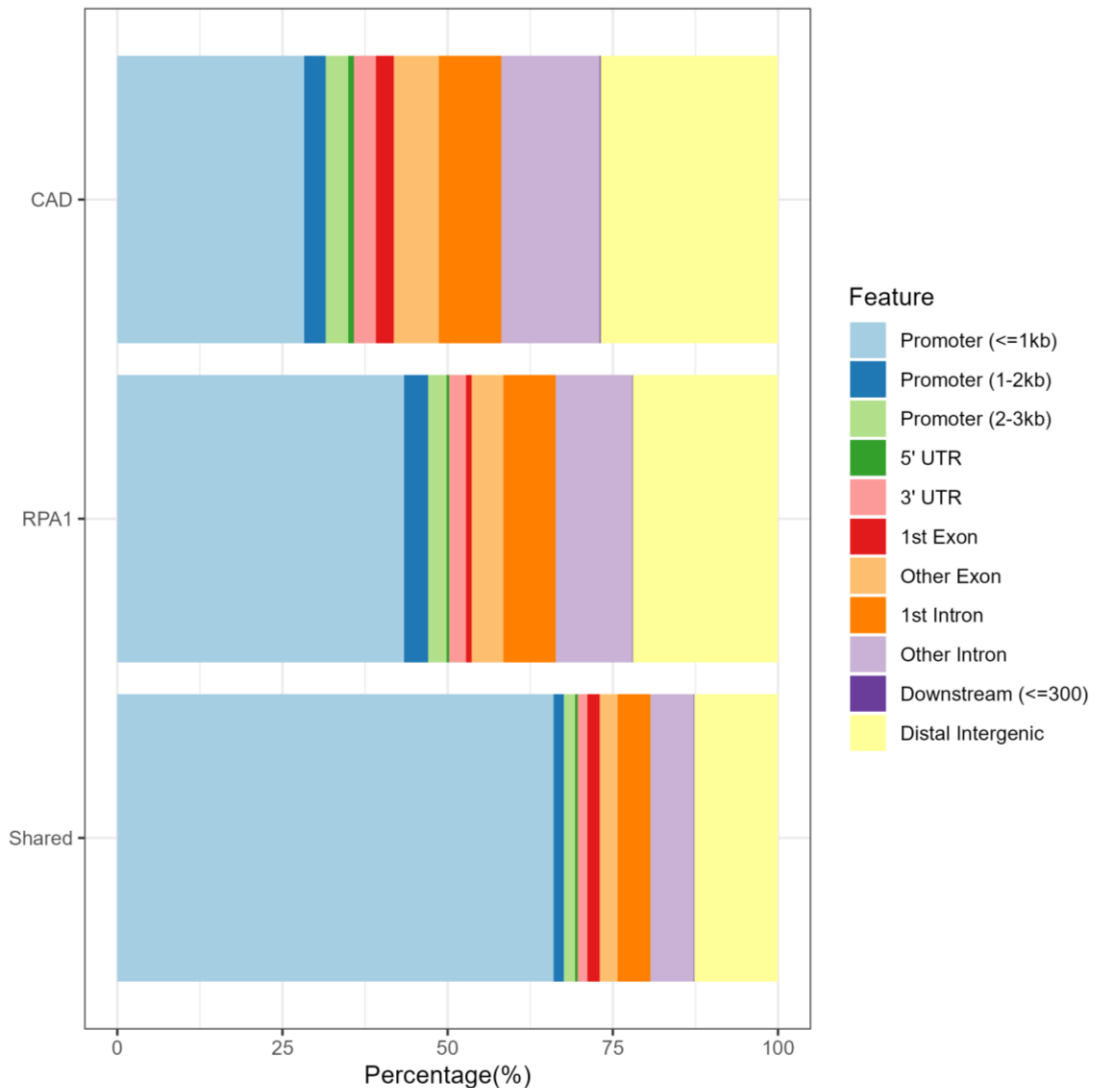
Genomic Distribution of CAD and RPA1 Binding Sites

A

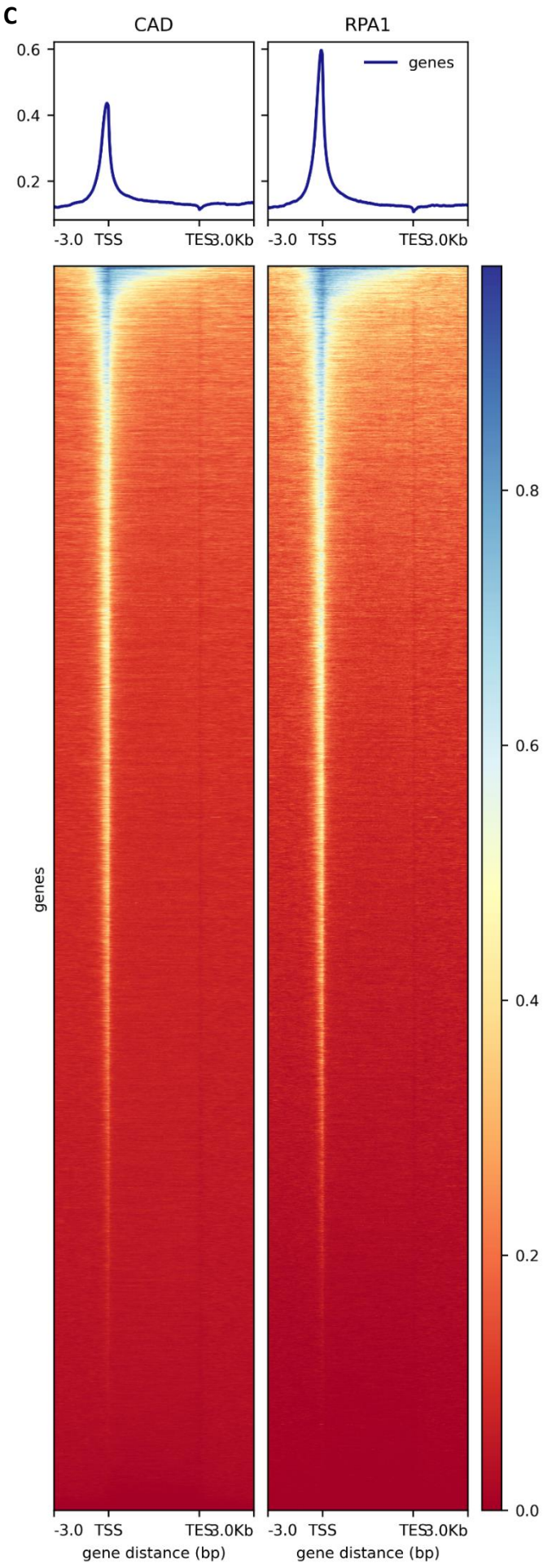


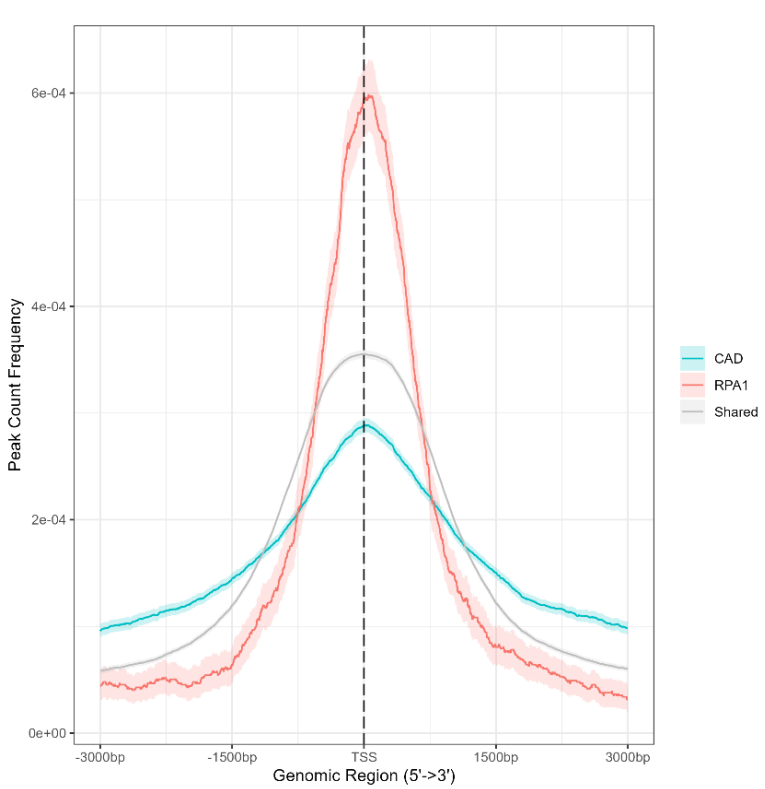
Significant Overlap between CAD and RPA1
47.6% CAD sites are shared with RPA1
88% RPA1 sites are shared with CAD

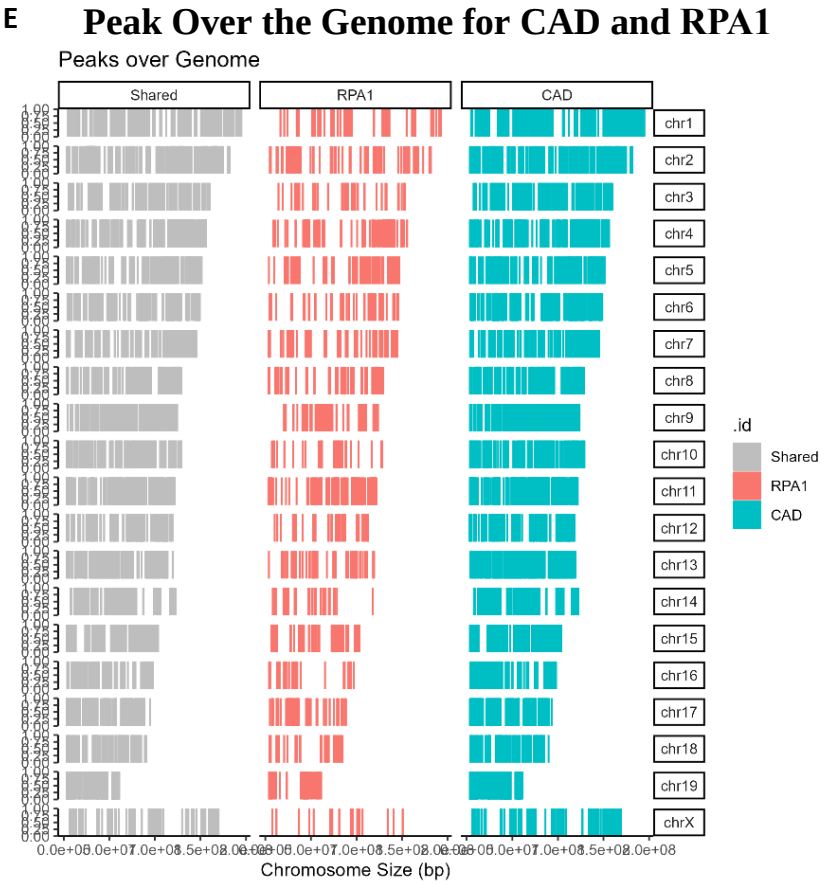
B



Genomic Distribution of CAD and RPA1 Binding Sites

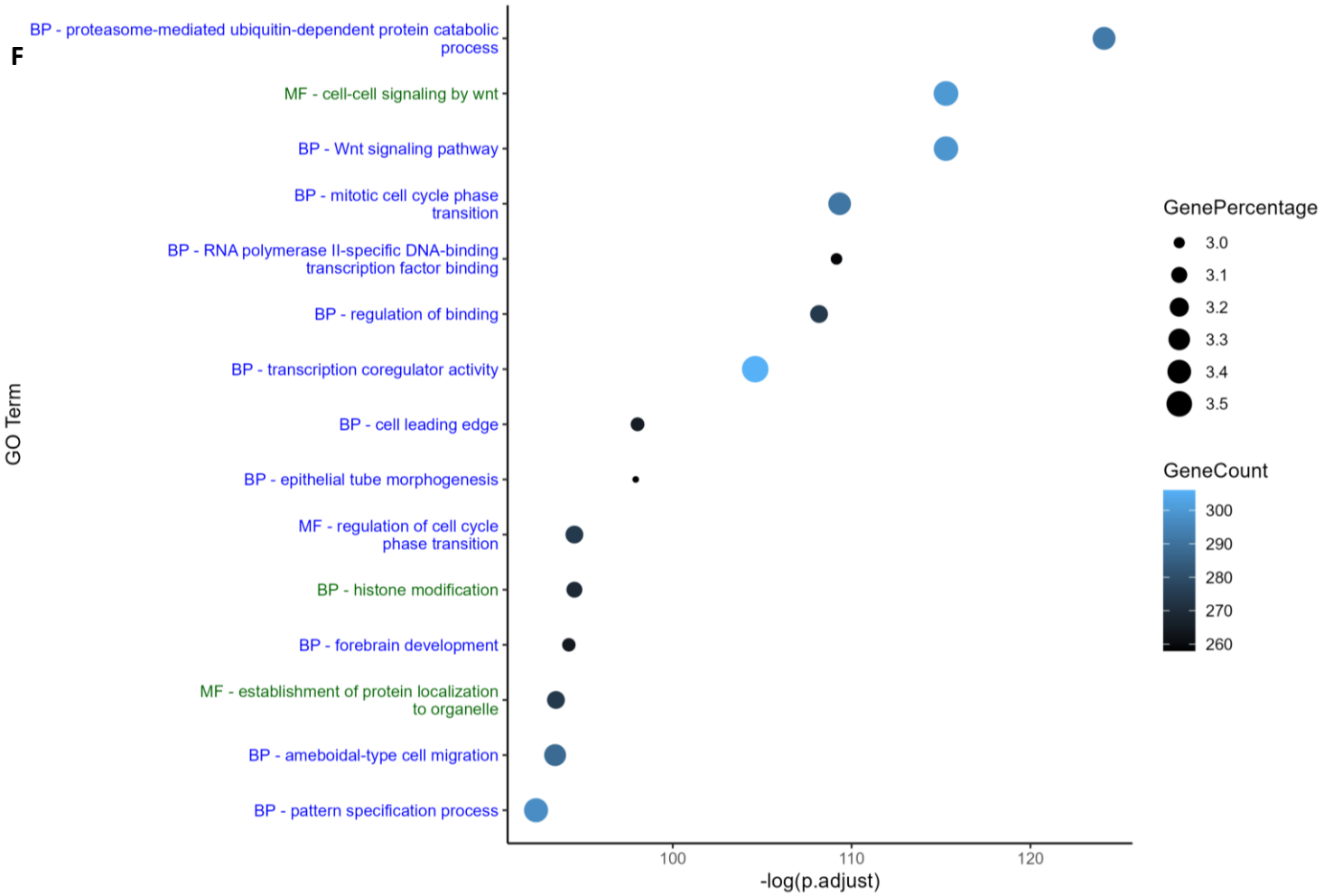






Gene Ontology (GO) Analysis Identified Shared Targets between CAD and RPA1

F



KEGG Pathway Analysis Identified Shared Targets between CAD and RPA1

G

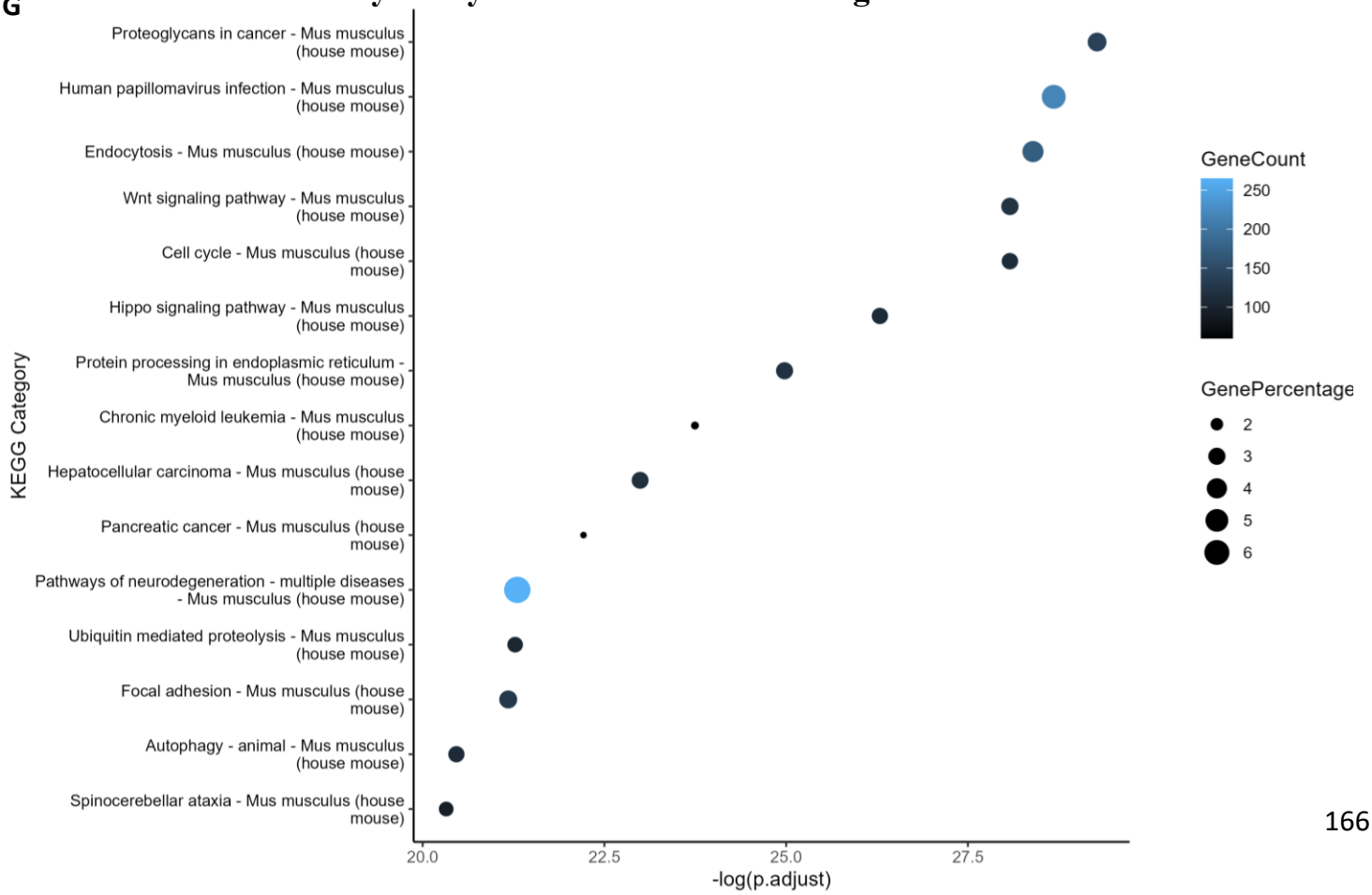
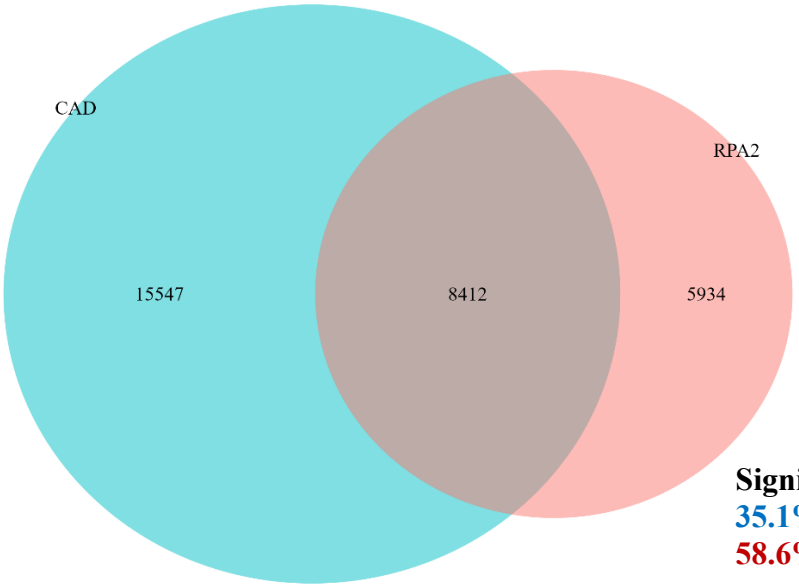


Fig. 5.4. Genomic distribution of CAD and RPA1 binding sites.
(A) Venn diagram illustrating binding site overlap among functional gene lists from peak analysis between CAD and PRA1. (B) Genomic region distribution in the peaks falling in various categories displayed. (C) plotHeatmaps of CUT and Tag show count per million (CPM) normalized signals for reads across 3,000 bp genomic interval relative to the TSS. Genomic distribution of the binding sites of CAD and RPA1 at 24 hr post differentiation from C2C12 muscle cells. Data sets were generated from merged two independent biological replicates for each condition. Heatmaps were plotted using computeMatrix and plotHeatmap tools in deepTools. Peaks were called and ordered for each dataset using the consensus of MACS, SEACR, and GoPeaks. (D) Peak count frequency over genomic regions for CAD and RPA1. (E) Peaks over genome represented between CAD and RPA1. Created with ChIP seeker, covplot function. (F) Gene Ontology (GO) analysis identified within the shared genes between CAD and RPA1. Top 15 terms are enriched. Biological Processes (BP), Molecular Functions (MF), Cellular Components (CC). (G) KEGG analysis identified within the shared genes between CAD and RPA1. Top 15 terms are enriched.

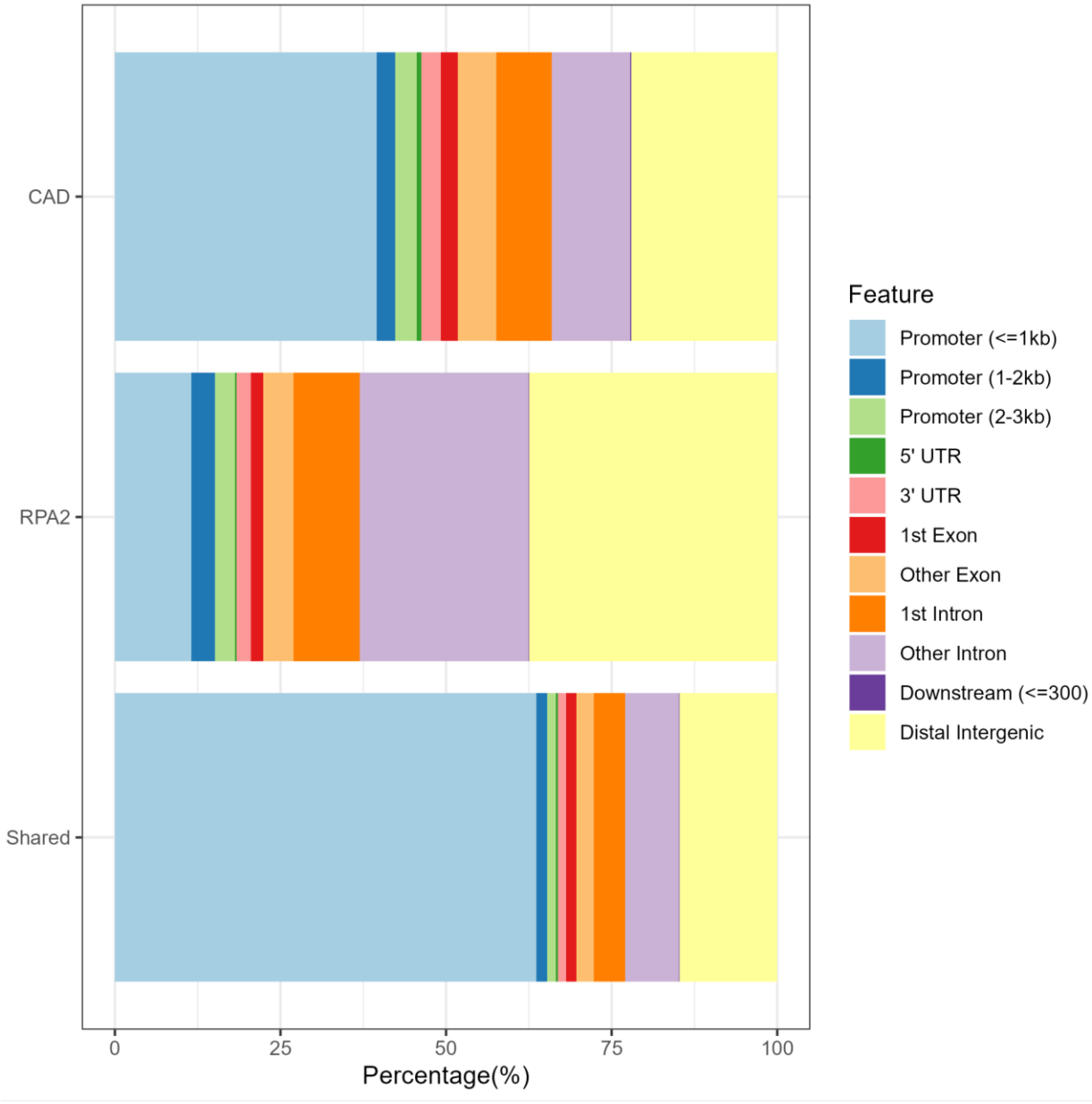
Genomic Distribuation of CAD and RPA2 Binding Sites

A

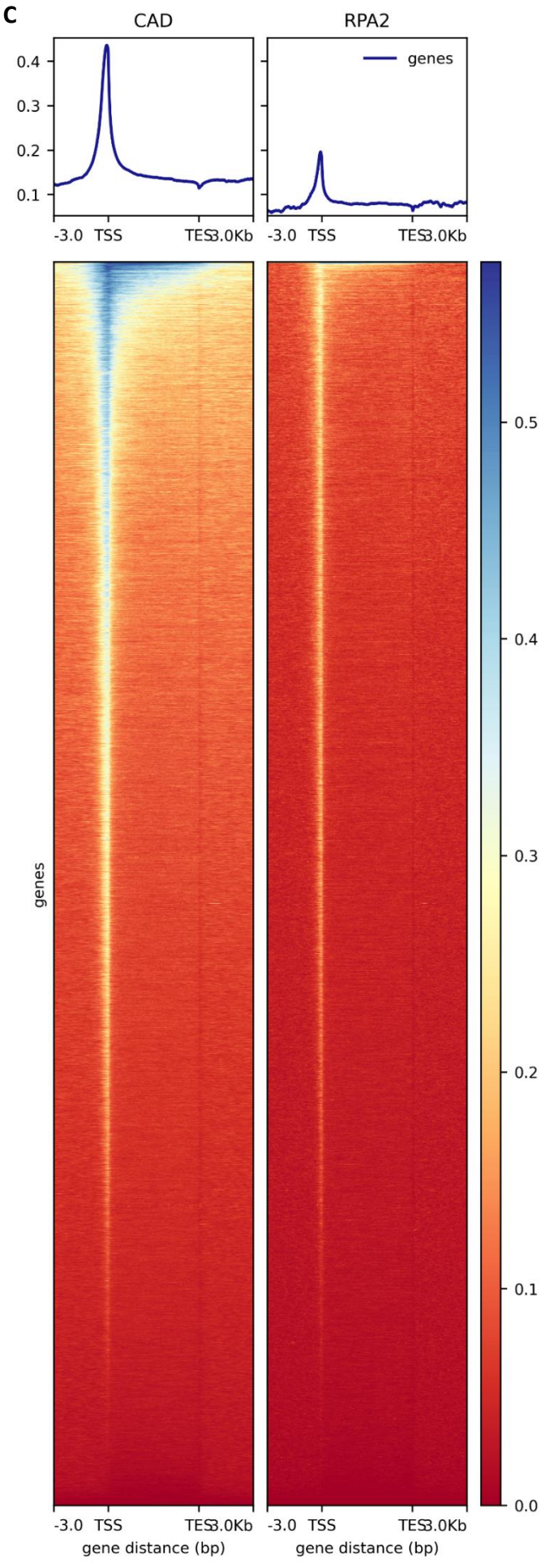


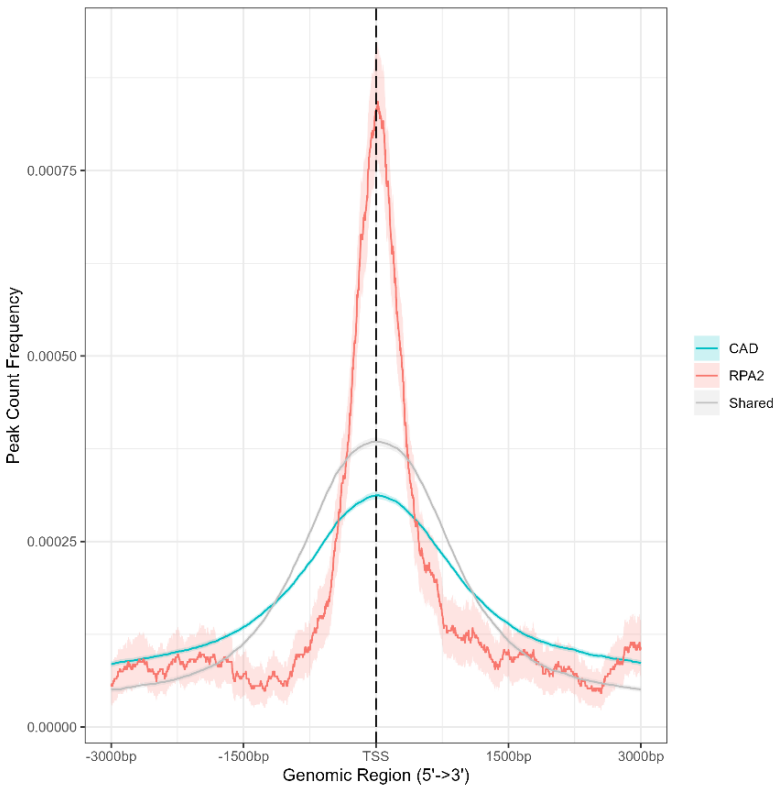
Significant Overlap between CAD and RPA2
35.1% CAD sites are shared with RPA2
58.6% RPA2 sites are shared with CAD

B



Genomic Distribution of CAD and RPA2 Binding Sites

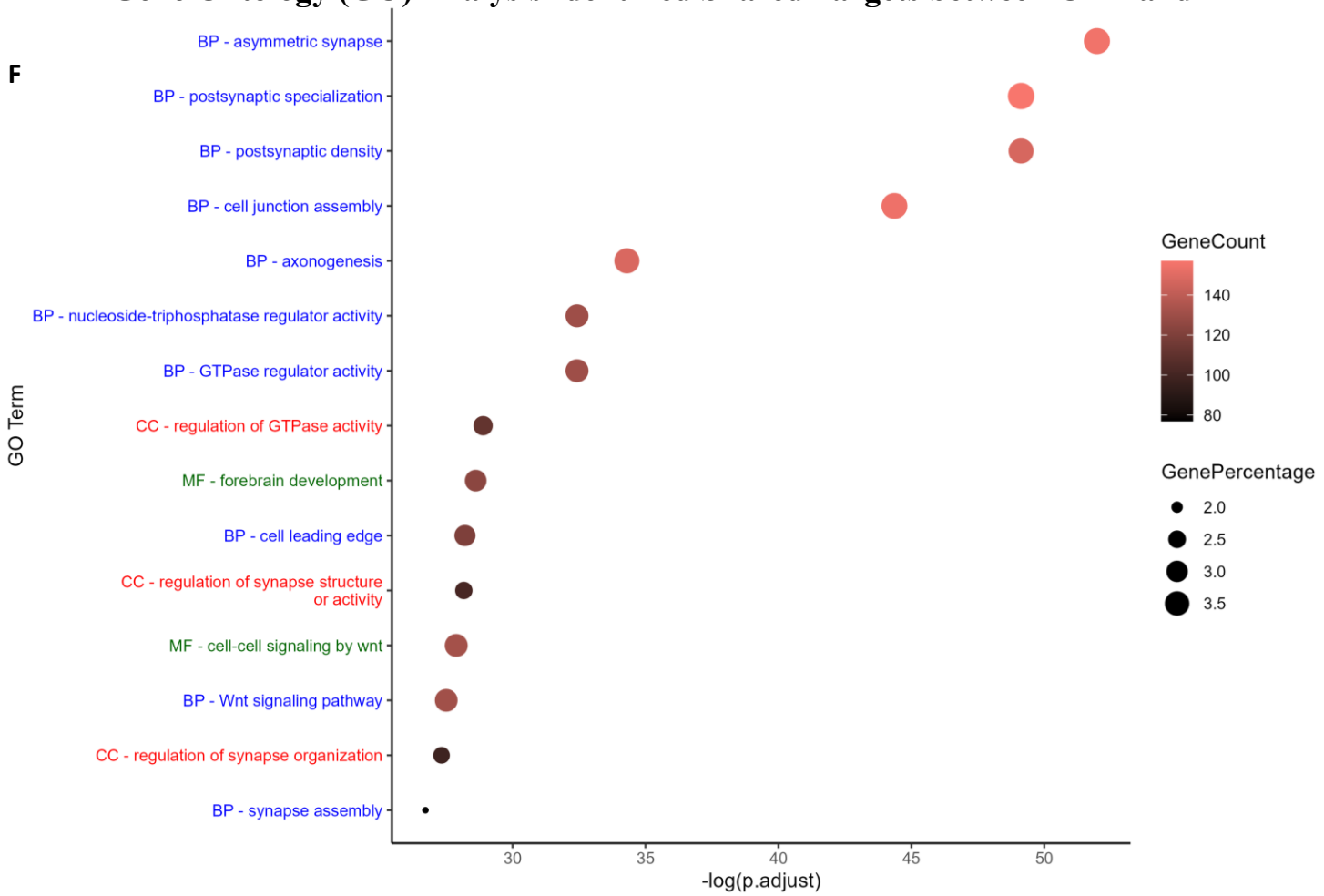






Gene Ontology (GO) Analysis Identified Shared Targets between CAD and RPA2

F



KEGG Pathway Analysis Identified Shared Targets between CAD and RPA2

G

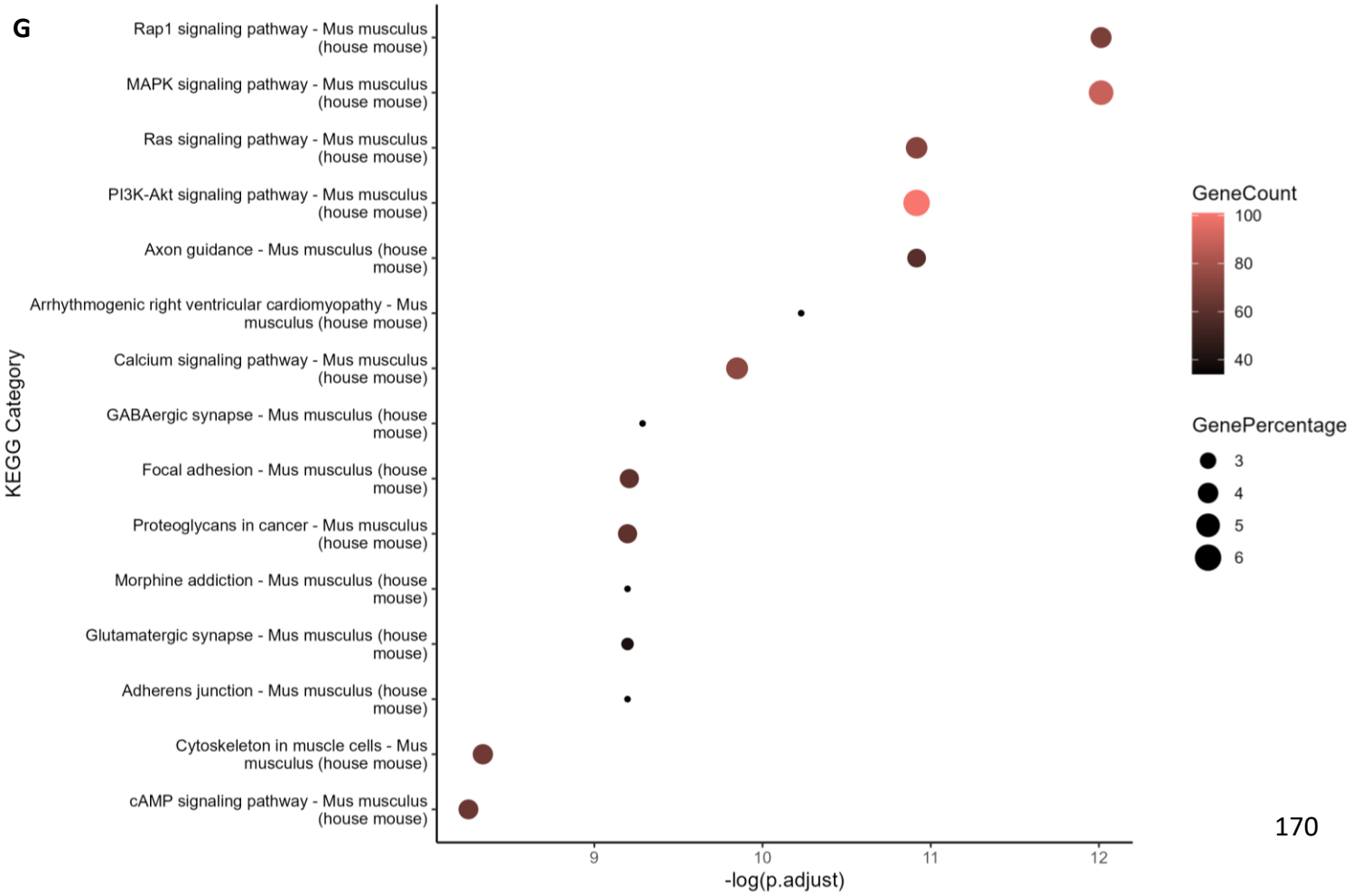


Fig. 5.5. Genomic distribution of CAD and RPA2 binding sites.
(A) Venn diagram illustrating binding site overlap among functional gene lists from peak analysis between CAD and PRA2. **(B)** Genomic region distribution in the peaks falling in various categories displayed. (C) plotHeatmaps of CUT and Tag show count per million (CPM) normalized signals for reads across 3,000 bp genomic interval relative to the TSS. Genomic distribution of the binding sites of CAD and RPA2 at 24 hr post differentiation from C2C12 muscle cells. Data sets were generated from merged two independent biological replicates for each condition. Heatmaps were plotted using computeMatrix and plotHeatmap tools in deepTools. Peaks were called and ordered for each dataset using the consensus of MACS, SEACR, and GoPeaks. (D) Peak count frequency over genomic regions for CAD and RPA2. (E) Peaks over genome represented between CAD and RPA2. Created with ChIP seeker, covplot function. (F) Gene Ontology (GO) analysis identified within the shared genes between CAD and RPA2. Top 15 terms are enriched. Biological Processes (BP), Molecular Functions (MF), Cellular Components (CC). (G) KEGG analysis identified within the shared genes between CAD and RPA2. Top 15 terms are enriched.

5.7.CAD binds to linker histone H1 variants

Histone H1 is an architectural protein located in internucleosomal linker regions of chromatin. It has been previously shown that during apoptosis, all H1 variants directly interact with CAD, stimulating its nuclease activity, which is necessary for CAD function (Widlak, et al 2005, Liu et al, 1999). H1 is a critical linker histone that interacts with linker DNA between nucleosomes and functions in chromatin compaction. Additionally, recent findings suggest that linker histone H1 binds to DNA to protect it from SSBs induced by CAD in cancer cells (Larsen et al, 2022). Interestingly, histone H1 and all its variants H1.1, H1.3, H1.4, and H1.5 were identified among CAD interactome via CAD IP-MS during differentiation of skeletal muscle cell differentiation (Fig. 5.6A). CAD CUT and Tag analysis also revealed CAD binding to all histone H1 variants; H1.5 Linker Histone (Hist1h1b), H1.2 Linker Histone (Hist1h1c), H1.3 linker Histone (Hist1h1d) and H1.4 Linker Histone (Hist1h1e) (Fig. 5.6B). The enrichment was stronger in 24 hours condition relative to growth. These findings suggests that H1 protein should remain intact to stabilize higher-order chromatin structure and CAD activity.

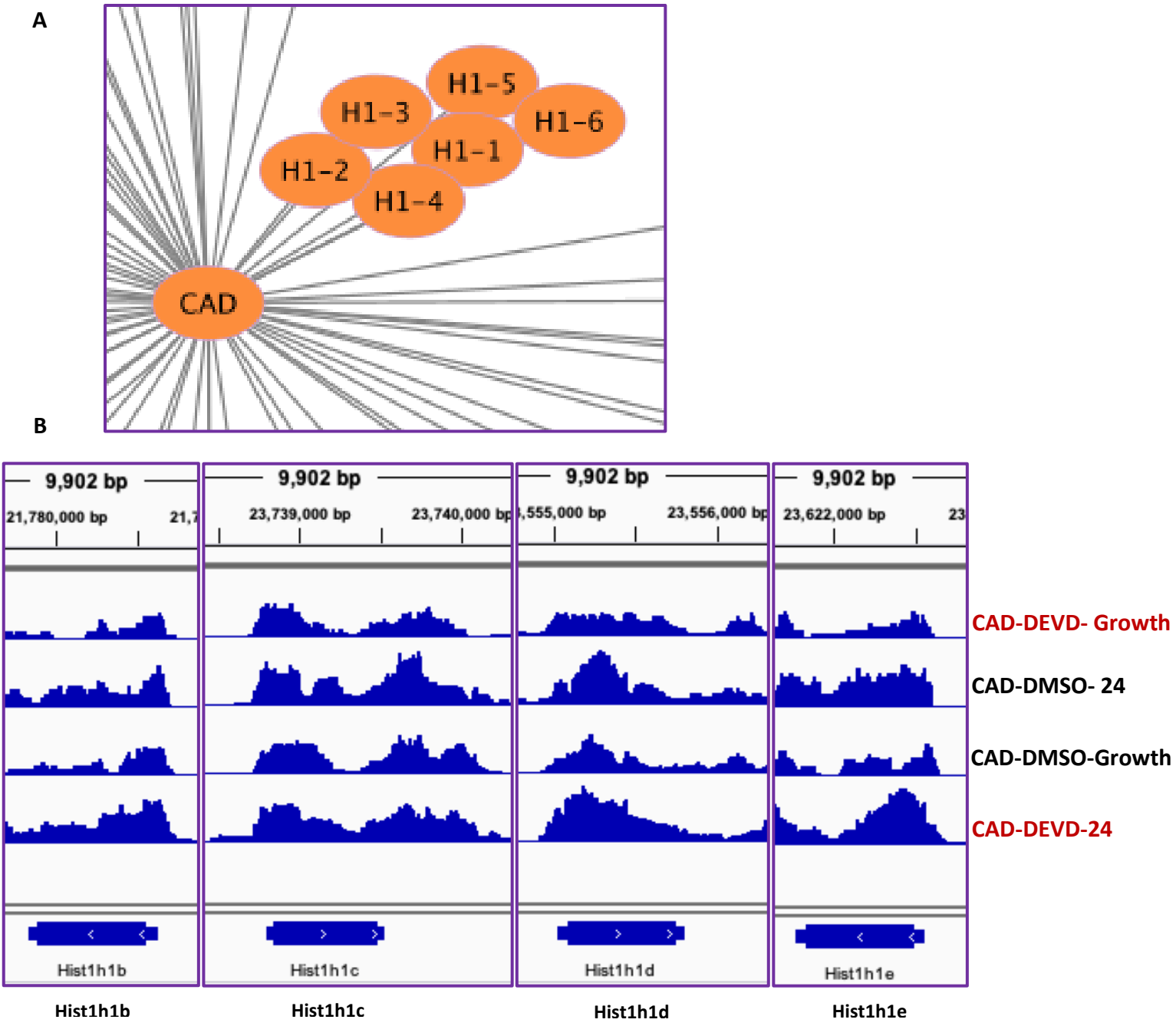


Fig. 5.6. CAD binds to linker histone H1 variants.
(A) CAD interacts with Histone H1 identified by IP-MS, from C2C12 cells. Created with Cyotscape.
(B) Representation of genomic view of CAD binding to Histone H1 variants identified by CAD CUT and Tag.

Summary

The CAD interactome identified several protein interactors involved in DNA repair and replication machineries, emphasizing the importance of future studies to elucidate the mechanisms by which these novel proteins mitigate CAD-mediated DNA breaks. Despite CAD and TopIIIβ having distinct functions, the significant overlap in their preferred genomic targets suggests a coordinated effort between an isomerase and a DNase to manage genomic adaptation during cell differentiation. Additionally, the greater overlap between CAD with RPA1, but not RPA2, targets may indicate a preferred role for RPA1 in differentiation-associated DNA repair responses.

CHAPTER 6: DISCUSSION

My thesis work focused on genome-wide mapping of CAD targets during the early differentiation of skeletal muscle cells, revealing CAD as a critical genome modifier. We observed that CAD binds to a broad cohort of gene promoters, extending beyond the previously identified p21 promoter, which is essential for skeletal muscle cell differentiation (Larsen et al, 2010). Surprisingly, our findings showed that CAD target numerous genes involved in cancer breakpoint gene fusions during the early differentiation of murine skeletal muscle cells. Most compelling, we identified that CAD binds to a similar subset of these breakpoint genes in T cell progenitors. Collectively, these observations suggest that CAD's role in the differentiation of skeletal muscle cells and T cells represents a conserved mechanism. We hypothesize that this conserved CAD function may extend to other cell lineages, potentially controlling the pace of cellular differentiation through directed DNA strand breaks. Additionally, it is reasonable to conclude that CAD binding to the parent genes that form breakpoint translocations may seed oncogenesis in other cell lineages, which is likely to occur during the cell differentiation process.

6.1. Caspase-3/CAD Drives Divergent Cell Fate Outcomes

Caspase-3/CAD are implicated in divergent cell fate outcomes (Bell & Megeney, 2017; Burgon & Megeney, 2018). During development, Caspase-3/CAD are crucial for driving the differentiation of murine skeletal muscle cells (Fernando et al, 2002; Larsen et al, 2010). Additionally, the differentiation of *Drosophila* macrophage cell types is dependent on CAD activation (Maurya et al, 2024). On the other hand, CAD has attracted significant attention from scientists in the cancer research field. Several studies have emerged over the preceding decade implicating caspase-3 in cancer development and tumorigenesis (Cartwright et al, 2017; Liu et al, 2017; reviewed in Benada et al, 2023). More recently, CAD nuclease has been shown to induce DNA breaks that contribute to therapy-induced cancer cell resistance (Larsen et al, 2022).

Additionally, other studies have demonstrated that CAD can regulate the response to inflammation-induced DNA damage following viral infection (Moeed et al, 2024; Haimovici et al, 2024). However, the study did not investigate whether CAD caused strand breaks *per se* or whether CAD influenced major inflammatory kinase signaling pathways such as NF-κB, which itself can target DNA (Moeed A, et al 2024). Another important study from the same group observed that the initiation of senescence was CAD dependent and could be engaged in response to various stressful stimuli such as aging (Haimovici, et al 2024).

Despite these discoveries, a genome-wide understanding of CAD's DNA targeting behavior has remained elusive. Does CAD manifest cell-specific targeting preferences? Does CAD bind to a conserved DNA element? Is CAD activity in cell fate transitions, such as cell differentiation unique or does it share features with CAD function in cancer cell emergence and/or responsiveness? These and other questions regarding CAD biology have remained unanswered. The confluence of these unknown aspects of CAD function served as the impetus for my thesis work.

6.2. CAD Activity may Initiate Cancer Breakpoint Fusion Gene Formation

In chapter 3, I observed that CAD binds to differentiation-responsive promoters and breakpoint fusion genes during normal skeletal muscle cell differentiation. These observations suggest that CAD may play a crucial physiological role in positively regulating the differentiation process, extending the original observations of Larsen et al (2010), where CAD-mediated breaks in the p21 promoter were identified as transcriptional induction cues for p21 itself. Additionally, a significant cohort of CAD targets during muscle cell differentiation comprised genes involved in breakpoint translocations. Here, CAD appeared to preferentially induce DNA strand breaks, leading to the

reduced expression of these genes, which was prerequisite for effective completion of the differentiation program.

One of the prominent CAD bound translocation prone genes was *Pax7,* as discussed in chapter 3. The regulation of *Pax7* through CAD-directed strand breaks raises the intriguing possibility that the initial step in the formation of the oncogenic *Pax7-Foxo1* fusion gene—the DNA strand break—could be a normal biological mechanism that evolved to repress the expression of genes that would otherwise inhibit the differentiation process. The oncogenic activity of the *Pax7-Foxo1* fusion protein may originate from a self-imposed regulatory event, where the fusion protein becomes resistant to natural degradation mechanisms that typically target Pax proteins in muscle stem and progenitor cells (Crose et al, 2014; Thalhammer et al, 2015). However, our results suggest that caspase-3/CAD targets the *Pax7* and *Foxo1* loci during the normal differentiation process, possibly to downregulate *Pax7* expression. This downregulation facilitates the differentiation program to proceed by introducing a physiological break at the exon-intron boundary, thereby disrupting transcription.

Given the observations of CAD binding to the translocation/fusion-prone *Pax7-Foxo1* loci, we investigated whether CAD targets/binds other genes in differentiating muscle cells known to contain translocation-sensitive regions. Unexpectedly, we found that CAD bound to a large cohort of genes associated with translocations in various cancer types, including sarcomas, leukemias, and lymphomas. This suggests that the genesis of oncogenic breakpoints may be a direct outcome of CAD activity, which is conserved to induce cell fate transitions such as differentiation. Thus, it is reasonable to propose that CAD-directed differentiation, through controlled DNA strand breaks, may be a common feature across all stem cell populations during cell fate transitions.

Other studies have indicated that CAD-like nuclease activity may promote cancer cell persistence. For example, limited MOMP has been shown to provoke caspase-dependent DNA damage, which can trigger oncogenic transformation and cancer (Ichim et al, 2015; Liu et al, 2017). Similarly, caspase-3 activation has been linked to persistent DNA breaks that contribute to tumorigenesis, although this has been attributed to the activity of the nuclease Endo G (Liu et al, 2015; Liu et al, 2017). While these studies demonstrated a cause-and-effect relationship between caspase-3 in cancer cell activation, the exact mechanism behind the strand breaks remained unresolved. Although the Endo G link is intriguing, this nuclease is primarily mitochondrial, and there is limited information about its direct nuclear DNA targets. Our observations suggest that CAD may be the operative DNase in MOMP-related cancer development, in addition to its prominent role in cell fate transitions.

6.3. Cell Differentiation Initiates a CAD Association with Numerous Translocation Prone Cancer Fusion Genes

In this thesis, we found that CAD bound to a cohort of translocation breakpoint parent genes in differentiating muscle cells and T cells, including *Tlx1, Tlx3, Btg2, Btg3, Cbfa2t3,* and *Runx1* (as discussed in chapter 3). Additionally, cross-comparison of CAD-bound breakpoint genes against RNA-seq data from muscle cells revealed that many of these CAD target genes have decreased expression levels in differentiating muscle cells. Interestingly, some of these genes also play roles in regulating differentiation cell fate. For example, *Tlx1* and *Tlx3* are homeobox transcriptional repressors that are frequently misexpressed and translocated in T cell acute lymphoblastic leukemia (T-ALL). Disruption of *Tlx1* and *Tlx3* genes has been reported to interfere with T cell differentiation, contributing to the development of a subset of T-ALL (Dadi et al, 2012; King et al, 2012). Moreover, B-cell translocation genes (BTGs), known for their anti-proliferative

functions, regulate the cell cycle and differentiation. Expression of both *Btg2* and *Btg3* was induced during myogenic differentiation of murine C2C12 cells but declined after 48 hours, indicating their role in myogenic differentiation (Feng et al, 2007). Furthermore, *CBFA2T3*, a transcriptional corepressor in hematopoiesis, is involved in myeloid differentiation and is associated with acute myeloid leukemia (AML) through the production of the *CBFA2T3-RUNX1* fusion in therapy-related AML (t-AML) linked to topoisomerase II inhibitors (Ottone et al, 2009; Steinauer et al, 2020). Our data also demonstrated CAD binding to the Runt-related transcription factor 1 (*Runx1*), which has been reported to translocate to *Runx1t1* and to modulate myogenic differentiation (Deng et al, 2023), as well as neuronal differentiation (Linqing et al, 2015). Additionally, c-Myc is known to regulate the balance between stem cell self-renewal and differentiation in hematopoietic system (Wilson et al, 2004), and its downregulation is crucial for the initiation of terminal differentiation in both epidermal and neural cell lineages (Gandarillas & Watt, 1997; Chin et al, 1995; Hurlin et al, 1995).

Taken together, these findings suggest that some differentiation genes may be targeted by an induction of strand breaks to disrupt transcription read-through and allow muscle cell differentiation to proceed. Perhaps CAD-mediated generation of strand breaks in differentiating progenitor cell is a normal biology that can initiate/precipitate translocation breakpoints. Additionally, it is important to note that our data do not support the hypothesis that breakpoint translocation originate from a cancer stem cell.

6.4. Radiation and Topoisomerase Poisons Contribute to Chromosomal Translocations

Cancer, in its essence, is a genome/genetic disease. In 1914, Theodor Boveri postulated that chromosome abnormalities could cause cancer, leading to the discovery of the first chromosomal translocation, known as the "Philadelphia chromosome." This chromosomal aberration was the

first genetic translocation linked to a specific cancer—chronic myeloid leukemia (CML). Boveri's work laid the groundwork for cancer cytogenetics, which was further developed in the 1960s, associating cancer to chromosomal aberrations (Nowell et al, 2007). Despite this, few studies have directly investigated the agents or causes of the initial breakpoints that lead to chromosomal translocations. Two factors have emerged as molecular explanations for the breakpoints, radiation exposure or topoisomerase II poisons, yet neither of these DNA disruptors can explain the genome wide impact or conservation of breakpoint formation.

Early studies demonstrated that treatment of breast cancer with topoisomerase II poisons generated breakpoints within the *PML* and *RARA* genes, leading to translocations associated with acute promyelocytic leukemia (Mistry et al, 2005). Although this study established a direct link between topoisomerase poisons and breakpoint formation, the underlying molecular mechanism remained unclear. Specifically, it was not well understood how the inhibition of a DNA isomerase led to predictable, site-specific DNA breaks without the involvement of a participating DNase.

Another major factor linked to chromosomal rearrangement is radiation. For example, the prevalence of papillary thyroid carcinomas (PTC) was notably high in patients exposed to radiation from the Chernobyl reactor disaster. After exposure at longer intervals with a period of 4 years latency, radiation induced chromosomal rearrangements specifically between NRTK1 and RET genes was associated with elevated rates of thyroid cancer (Rabes et al, 2000). Similarly, children from Belarus exposed to the Chernobyl nuclear accident exhibited an increased frequency of *RET* gene rearrangements in thyroid carcinoma (Fugazzola et al, 1995). Another rearrangement reported from the same incident, was involving *NTRK3-ETV6* in a Ukrainian-American population which was linked to radiation-induced thyroid cancer PTC (Leeman-Neill et al, 2014). Although these studies highlighted a significant global health issue and established a correlation between

radiation exposure and chromosomal rearrangements, they did not investigate the underlying basic mechanisms.

6.5. Chromosomal Aberrations are Detected in Non-Cancerous Contexts

Our observations in chapter 3 led us to question how CAD can bind to cancer-inducing breakpoint genes in a developmental model? Is there a precedent for DNA-damaging proteins to associate with DNA without causing strand breaks or precipitating cancer formation? Remarkably, there is a body of literature indicating that chromosomal aberrations can be detected in healthy individuals and tissues.

6.5.1. Chromosomal Translocations

The association between chromosomal translocations or breakpoints and cancer has been known for several decades (Bunting et al, 2013). Translocations are structural genome variants (SVs) that occur when two chromosomes in close proximity break and the incorrect ends rejoin, generating a fusion gene with oncogenic potential that deregulates cell growth (Ramsden et al, 2021). Although non-random chromosomal translocations are frequently linked to cancer (Mitelman et al, 2007), they are not exclusive to cancer; as they are also detected in newborn samples and healthy individuals. For example, screening of newborn cord blood samples revealed the presence of *TEL-AML1* and *AML1-ETO* fusion gene mRNA during normal fetal development. The authors suggested that these fusion genes likely occur in utero and are present in circulating blood at birth, but their high frequency alone is insufficient to trigger leukemia (Mori, 2002). Moreover, chromosomal translocation *t(14;18)(q32;q21)*, implicated in B-cell leukemia/lymphoma-2 (BCL2), has been detected in lymphoid tissues of healthy individuals (Summers et al, 2001). These

t(14;18)-harboring clones are long-lived and their prevalence increases with age, yet without overt evidence of cancer (Liu et al, 1994; Schüler et al, 2009).

The fusion gene product *bcr-abl*, characteristic of chronic myeloid leukemia (CML), results from the reciprocal translocation between chromosomes 9 and 22, known as the Philadelphia (Ph) chromosome. *Bcr-abl* RNA expression was also detected at low levels in samples from healthy subjects (Biernaux et al, 1995). These findings, indicating the presence of translocations in healthy individuals, led to the hypothesis that additional mutations or genetic hits are required to trigger malignancy (Aplan et al, 2006). Interestingly, the biological drivers of these gene breaks remain unknown, both in cancer tissues and in otherwise healthy individuals.

6.5.2. Chromothripsis

Chromothripsis refers to the shattering of one or multiple chromosomes and their subsequent reassembly in a single catastrophic event. Although this extreme pattern of chromosomal rearrangement has been linked to cancer, it has also been reported in samples from healthy populations (Gundy & Varga et al, 1983). Additionally, chromothripsis has been described in benign tumors, such as uterine fibroids and uterine leiomyoma cells—a benign tumor of the uterine myometrium (Mehine, 2013; Mehine, 2014)—as well as in meningiomas (Baltus, 2019).

The presence of chromothripsis in healthy individuals can appear phenotypically normal, suggesting that the genome can tolerate extreme reshuffling of chromosomes without triggering cell death or causing dramatic phenotypic changes. However, this phenomenon can still impact offspring, as it may be transmitted from mother to child, increasing the prevalence of miscarriages and severe congenital diseases (Pagter, 2015).

The precise mechanisms driving chromothripsis remain unclear, but the involvement of a caspase-activated CAD process could potentially explain the simultaneous genome-wide chromosome shuffling and segmentation. It is reasonable to speculate that CAD drives the process of chromothripsis which is triggered by DNA strand breaks, a known function of CAD. In addition, CAD activity has been observed in cancer cells, suggesting a potential involvement of CAD in this phenomenon. Additionally, in this context, there is one evidence linking CAD to micronuclei production in cancer cells (Haimovici et al., 2022). Micronuclei are DNA fragments of nuclei and chromosomes appear like chromosomal fragmentation seen in chromothripsis, implying a possible involvement to CAD activity. Despite this connection, this hypothesis still requires experimental validation.

6.6. Is CAD Operative in Other Cell Fate Transition/Differentiation Other than Muscle Progenitor Cells?

6.6.1. Caspase/CAD Activity in the Initiation of Hematopoietic Differentiation

In chapter 3, we identified numerous genes implicated in the hematopoietic system. Caspase-3, for instance, has been shown to drive erythrocyte differentiation (Boehm et al, 2013; Zermati et al, 2001). Further studies have identified few proteins cleaved by caspases to propel the differentiation of erythroblasts, including PARP, lamin B, Acinus (reviewed in Solier et al, 2017), and GATA1 (De Maria et al., 1999; Ribeil et al., 2006). Although several groups identified protein substrates to be cleaved by caspases during differentiation of hematopoietic cells, yet the underlying mechanism has remained undetermined. It is tempting to expect that CAD upon activation of caspases mediated the differentiation of mammalian hematopoietic cells akin to the skeletal muscle system. In addition, differentiation of Drosophila macrophage progenitor cells is dependent on

caspase/CAD mediated DNA strand breaks. However, CAD genome targets in the Drosophila macrophage lineage remains undetermined (Maurya, et al 2024).

Examining other aspects of erythropoiesis, in the terminal stages, mammalian erythrocyte nuclei polarize within the cell, progressively condense, and are then extruded from the cytoplasm in a process known as enucleation (Ji et al, 2011; Keerthivasan et al, 2011; Mei et al, 2021). This process is physiologically significant as it facilitates an increase in hemoglobin levels and allows red cells to acquire a biconcave shape, which likely provides flexibility for blood circulation through capillaries. However, unresolved questions remain, particularly regarding the precise mechanism of nucleus degradation during this attenuated form of cell death. For example, caspase-3 has been documented to participate in the enucleation process, with studies indicating that chromatin condensation and histone modifications, influenced by caspase-3, assist in red cell maturation (Zhao et al, 2016; Baron & Barminko, 2016). This mechanism mirrors that of apoptosis, driven by caspase activation. Additionally, phosphorylation of H2A.X during terminal erythroid maturation (the late stage of erythropoiesis), which coincides with the timing of nuclear extrusion (Jeffery, 2021; Zhao et al, 2016; Baron & Barminko, 2016), suggests that phosphorylated H2A.X may mark sites of DNA repair following CAD activation, akin to the differentiation process observed in skeletal muscle cells.

6.6.2. Does Caspase-3 Activate CAD to Drive Differentiation in the Neuronal Lineage?

Caspase signaling is conserved and has been reported to drive neural progenitor cell differentiation (Fernando et al, 2005; Kanuka et al, 2005; D'Amelio et al, 2010; Colon-Plaza et al, 2022). Studies have also shown an increase in DSBs formation during neurodevelopment, with neuron promoters being targeted by DSBs, leading to changes in gene expression which is crucial for synaptic

plasticity in the brain (Wu, 2021; Su et al, 2015; Michel et al, 2022). Additionally, we observed CAD binding to several genes implicated in brain-related fusion events, including Braf (BRAF), nuclear receptor coactivator 1 (Ncoa1), neuropilin 1 (Nrp1), cAMP responsive element modulator (CREM), and cullin 2 (Cul2). Although CAD genome targeting in the brain has never been explored, it is reasonable to hypothesize that this DNase may play a role in modulating neural cell differentiation. This hypothesis is supported by: 1) caspase-3 is active and drives the differentiation process of neurons; 2) that CAD binds to genes implicated in brain development; and 3) the detection of widespread DNA strand breaks in both developing and adapting neuronal populations. Further experimentation is required to integrate these events into a single narrative, yet these observations are collectively supportive of a role for CAD mediated genome wide breaks as a mechanism to drive neural differentiation.

6.7. Does the Epigenetic Landscape Influence CAD Genome Targeting during Cellular Differentiation?

Our findings in chapter 3 suggests that CAD target specific regions without a conserved binding motif in it's binding sites. CAD clearly is not acting as a transcription factor *per se* by targeting specific DNA sequences, neither targeting to CTCF sites as CAD behavior in cancer cells (Larsen et al, 2022). It seems possible that there are other structural features that shape the environment to facilitate CAD accessibility to it's genomic targets. Given that CAD must be strictly regulated to maintain a balance between transiently cleaving its targets, yet not destroying/fracturing the genome during differentiation, we speculate that CAD sites are likely restricted by a repression mechanism. This mechanism would control the positioning of CAD-induced strand breaks and limit CAD's ability to inadvertently engage gene expression.

6.8. Gene Repression by PRC Complexes Protect DNA Damage Sites

Diverse cell types within an organism share the same DNA content and sequence, yet it is the variation in gene expression that defines cell identity and function. Protein complexes bind to DNA to activate or repress gene expression through proximal or distal promoters and enhancers. Among these, PRC1 and PRC2 complexes are crucial epigenetic regulators involved in gene silencing. Strong evidence indicates that PRC complexes, including Ezh2, the E3 ubiquitin ligase-containing BMI1, and RING1B, are recruited to DSB sites, where they induce H3K27me3 to promote transcriptional repression (Miller et al, 2012). A proteomic study combining isotopic labeling with chromatin fractionation found that, in response to UV laser-induced DNA damage, PcG proteins including Ezh2 are recruited to DNA lesions (Chou et al, 2010). This recruitment was not associated with the common DNA damage response (DDR) pathways, such as ATM, ATR, or H2AX (Chou et al, 2010). The mechanisms by which PRC2 targets genomic loci for methylation are not fully understood and vary based on cell type and health status compared to disease conditions (Wiles & Sekler, 2017). Additionally, there is no specific DNA sequence to inform PRC2 to target the regions to be methylated. However, it's thought to be the chromatin environment including histone modifications, DNA methylation, transcription, and the chromatin structure are all interconnected to prime PRC2 to distribute H3k27me pattern (Wiles & Sekler, et al, 2017).

6.9. Caspases Proteolytic Activity may Target Chromatin Proteins in Differentiating Muscle Cells

In Chapter 4, we present evidence of caspase-3 regulating the chromatin protein Ezh2 by cleaving it in early differentiating muscle cells. This finding further confirms the role of caspases as genome

reorganizers through their proteolytic activity. Similarly, observations from our laboratory have reported that caspase-7 regulates chromatin by cleaving the SATB2 protein, thereby promoting the differentiation program in skeletal muscle cells (Bell et al, 2022). In this context, it is likely that a similar proteolytic targeting event modifies the Ezh2 protein. Indeed, Ezh2 has been reported to undergo degradation via posttranslational modification (PMT) of the ubiquitin system during muscle cell differentiation (Consalvi et al, 2017). However, our observations do not align with these findings, and it seems unlikely that such a critical protein for the differentiation program would be present exclusively to be subject for degradation.

In addition, other studies have reported that Ezh2 levels are targeted by microRNAs (miRNAs), leading to its degradation. Specifically, miR-26a, miR-214, and miR-221/222 have been shown to regulate gene expression during the differentiation of skeletal muscle cells by repressing Ezh2 levels (Wong et al, 2008; Juan et al, 2009; Ge et al, 2011).

6.10. Ezh2 Catalytic Activity Depends on the PRC2 Complex

It is important to note that Ezh2 alone lacks enzymatic activity; its incorporation with other PRC2 complex subunits is crucial for the deposition of the methylation mark H3K27me3 (Tan et al, 2014; Yu et al, 2019). Furthermore, recombinant Ezh2 protein has 1000-fold lower activity compared to the HMTase within the tetrameric PRC2 complex, highlighting the importance of non-catalytic subunits in the overall methylation activity of the complex (Nekrasov et al, 2005). This is manifested in the *Drosophila* model where the enzymatic activity of histone methyltransferase and nucleosome binding requires modulation by different PRC2 components (Nekrasov et al, 2005). For instance, Nurf55 and Esc are homologs of RbAp46 and RbAp48 in the mammalian PRC2. In

Drosophila, Suz12, Nurf55, and Esc work together to facilitate E(z) HMTase activity (Nekrasov et al, 2005).

Furthermore, Suz12, a subunit of PRC2, is essential for the HMT activity of PRC complexes, specifically for the stability of Ezh2-containing complexes, as reported in mouse embryogenesis and human differentiated cells (Pasini et al, 2004). Additionally, di- and tri-methylation of H3K27 showed a significant reduction in Suz12 knockout embryos, indicating a critical role for Suz12 in the enzymatic activity of Ezh2 on H3K27 (Pasini et al, 2004). The intact trimeric PRC2 complex, including EED, Suz12, and Ezh2, binds to H3K27me3 to maintain transcriptional repression independently of the Ezh2 catalytic SET domain (Hansen et al, 2008). In turn, local chromatin compaction paves the way to the distribution of histone H3K27 methylation, thereby maintaining the repressed state (Yuan et al, 2012).

6.11. Topoisomerase IIIβ Interaction with CAD; A Breakpoint Modifying Protein Complex?

Another group of DNA-modifying enzymes linked to the generation of DNA strand breaks is the topoisomerase family. Several studies have reported that breakpoints originate from topoisomerase enzymes, which alleviate torsional strain in DNA by converting a supercoiled state to a relaxed state (Pommier et al, 2022). For example, recurrent translocations associated with topoisomerase II cleavage sites include MLL-AF9, MLL-AF4, MLL-ETO, and PML-RARA (Rowley et al, 2002; Cowell et al, 2012; Pommier et al, 2022). Earlier studies identified topoisomerase I and II binding to oncogenic fusion genes that trigger Burkitt's lymphoma, mixed-lineage leukemia (MLL), and acute myeloid leukemia (AML) (Kato et al, 1991; Domer et al, 1993). More recently, topoisomerase IIβ has been linked to topologically prone breakpoints that align with both

chromosome loop structures and regions of paused RNA polymerase activity (Canela et al, 2017; Gothe et al, 2019).

Topoisomerases are molecular targets of several chemotherapies and anticancer agents (Buzun et al, 2020; Martín-Encinas et al, 2022). Treatment with these chemotherapies triggered chromosomal translocation that are causative of fusion genes. Despite these studies, the mechanisms proposed in most of them are predicting chromosomal breakpoints originating from topoisomerase based on drug-induced action generated DNA breaks (Pendleton M et al. 2014). Not only that, but also some of these studies compare in vitro topoisomerase II DNA-induced cuts with chromosomal breakpoints that are derived from leukemia patients, which is a comparison that is skeptical and requires further investigation (Mistry et al, 2005; Felix et al, 2006).

Our data in Chapter 5 indicates that CAD physically interacts with topoisomerase IIIβ protein, and that both targets share a significant number of the same genomic sites. This strongly suggests that CAD works in conjunction with topoisomerase, where topoisomerase facilitates the twisting of DNA while CAD inflicts DNA breaks. Unlike the enzymatic activity of CAD, which robustly catalyzes the cleavage of phosphodiester bonds in DNA under optimal pH and cofactor conditions, isomerases primarily function to catalyze the rearrangement of atoms within molecules to enable their interconversion. Equally intriguing, we speculate that topoisomerase IIIβ may also function as an inhibitory factor for CAD, potentially restraining CAD from inducing DNA strand breaks. In this hypothetical scenario, topoisomerase IIIβ might regulate CAD's access to DNA, thereby controlling the induction of breaks. Taken together, it is tempting to propose that topoisomerase IIIβ may open the local chromatin environment, facilitating CAD's ability to inflict DNA breaks.

6.12. DNA Repair Mechanisms Mediate CAD Induced DNA Damage

"*We totally missed the possible role of ... [DNA] repair although ... I later came to realise that DNA is so precious that probably many distinct repair mechanisms would exist.*" Francis Crick (Nature, 1974). Thus, it is essential to address a DNA repair process when discussing DNA damage factor, such as CAD. DSBs are a hallmark of cancer and can arise from external factors such as ionizing radiation and oxidation, as well as from genotoxic stress including replication, transcription, and erroneous DNA repair (Valente et al, 2022). When DSBs remain unresolved, they can lead to chromosomal translocations and mutagenesis, and cells activate various DNA repair pathways depending on the type of damage/lesion. For instance, base excision repair (BER) corrects chemical modifications of DNA bases, nucleotide excision repair (NER) removes bulky DNA lesions such as lesions caused by UV light, and mismatch repair (MMR) replaces mispaired bases. DSBs are repaired through homologous recombination (HR) or nonhomologous end joining (NHEJ) (Chatterjee et al, 2017; Ciccia et al, 2010; Xu & Xu, 2020). Effective repair allow cells to resume normal function without compromising genomic integrity. This process involves a range of enzymatic activities, including polymerases, helicases, recombinases, ligases, demethylases, phosphatases, topoisomerases, and glycosylases (Ciccia et al, 2010).

The prevalence of CAD-mediated DNA strand breaks in differentiating cells raises important questions about how cells survive this challenge and suggests that CAD activation may broadly influence the DNA repair process. Notably, three key studies have highlighted that DNA damage and repair are critical for activating muscle-specific genes (Larsen et al, 2010; Al-Khalaf et al, 2016; Connolly & Fearnhead, 2016). Additionally, observations presented in Chapter 5 reveal that CAD binds to several proteins involved in DNA repair and replication machineries. CAD's interaction with proteins implicated in SSBs and DSBs is particularly intriguing. It remains to be

determined whether CAD operates independently of DNA repair mechanisms or is integrated into the repair machinery during muscle cell differentiation. Further studies are needed to explore how CAD's protein interactions direct or control its activity, how CAD resolves differentiation-associated DNA damage, and regulates gene expression.

6.13. Concluding Remarks and Future Directions

In this thesis, we present data supporting the hypothesis that CAD exhibits widespread distribution across the genome, targeting both promoters and exon-intron boundaries. CAD binds to promoters, where it induces DNA strand breaks to positively regulate gene expression and advance cellular differentiation. Conversely, CAD also binds to exon-intron boundaries within gene bodies, potentially to target and suppress gene expression maybe as part of inducing the differentiation program. These observations are particularly relevant to cancer biology and the understanding of translocation breakpoints.

The observations presented in this thesis suggest that CAD binding to genes with decreased expression in differentiating muscle cells is presumably indication of induced DSBs to disrupt transcription. In contrast, SSBs mediated by CAD are probably associated with transcriptional induction. Therefore, genome-wide mapping to precisely identify the sites of SSBs versus DSBs introduced by CAD would be a valuable approach for distinguishing the types of breaks induced by CAD. Investigating the nature of DSBs versus SSBs will help validate the origins of translocations and promoters bound genes induced by CAD or other comparable enzymes.

As outlined in chapter 3, we conducted a cross-species comparison of CAD gene binding between murine muscle cells and human breakpoint regions from the COSMIC database. This analysis revealed significant syntenic overlap between CAD binding sites and gene breakpoints, with CAD

binding to nearly 80% of genes in differentiating mouse cells to all curated data from various human cell lineages in the COSMIC database. This suggests that CAD plays a conserved biological role. The next logical step would be to apply CAD CUT and Tag to human muscle cells to directly assess CAD binding to breakpoints in humans. This approach will provide a better understanding of CAD's role in breakpoint translocations across different cell lineages and their nascent forms of cancer.

Finally, our data suggest that CAD is intricately involved in regulating the local chromatin environment, with a specific focus on the Ezh2 protein. Our findings suggest concurrent caspase targeting events to control CAD directed strand breaks and the surrounding chromatin environment to regulate CAD sites accessibility. Ezh2 plays a critical role in the differentiation of muscle progenitor cells; however, the precise mechanisms by which Ezh2 is controlled require further investigation. A key approach to elucidate this is by constructing a non-cleavable version of Ezh2. This can be achieved by introducing single- or double-point mutations at the cleavage sites identified in Chapter 4, using site-directed mutagenesis. We hypothesize that a mutant version of Ezh2, which can no longer be cleaved by caspase-3, by substituting the expected aspartic acid residue with alanine will disrupt myogenesis. Specifically, we anticipate that the PRC2 complex will remain functional, maintaining gene expression which reduce the ability of muscle cells to differentiate.

6.14. Technical Consideration for CAD Genome

Despite the critical role of CAD, a single study by Larsen et al (2022) noted CAD activation that occurs following radiotherapy. Here, using a methodology named GLOE-seq. (which is genome-wide ligation of 3′-OH ends followed by sequencing) to map DNA single strand breaks, they

observed that CAD genome binding in human colon cancer cell line (HCT116) was enriched at CTCF sites (Larsen et al, 2022). They explained that CAD targeting of CTCF as a signal to induce cell cycle termination to protect cancer cells from the DNA damaging consequences of radiotherapy. Despite the excitement of these observations, still the CAD DNA targeting activity they reported in cancer cells is unlikely to be similar to the CAD binding in differentiating cells. In differentiating progenitor cells, CAD activity originated from caspase-3 cleavage of ICAD, which in turn release ICAD from the DNA bound CAD/ICAD complex (Larsen et al. 2010). However, in cancer cells, CAD activity is derived from direct phosphorylation of ICAD, and independent of caspase activity.

Despite the advances in sequencing, elucidating the genome of CAD was quiet challenging. As mentioned above, the DNA retention and libraries were very low with GLOE-seq. in cancer cells. To tackle CAD genome before implementing the Cut and Tag technique, I initially utilized an alternative method referred to as Breaks Labeling *In Situ* and Sequencing "BLISS" to label strand breaks (Yan et al, 2017). However, this method yielded limited success due to several factors, most notably the difficulty in detecting strand breaks under physiological conditions, likely due to CAD transient activity. Indeed, most of the studies that investigated DSBs utilizing this technique were employed under conditional induction of strand breaks using restriction enzymes or CRISPR, and not endogenous DSBs as the case of CAD.

We asked whether CAD binding it's genomic targets is constitutive on DNA or transient? To test this hypothesis, we constructed adeno viruses expressing CAD/ ICAD for gene delivery of mouse wild-type (WT) ICAD, WT CAD nuclease, and dead version of CAD. To generate nuclease dead version of CAD, we introduced a mutation of histidine residues to alanine in H263A, which revealed very little DNase activity. Next, C2C12 cells were infected with these adenoviruses and

subjected for CUT and Tag experiment. Unfortunately, all the libraries generated failed, most likely the expression of the adenoviruses interfered with the tagmentation process. Despite the challenges discussed above, the use of chemical inhibition of caspase-3 revealed robust CAD genome binding in C2C12 muscle cells. Additionally, it would be informative to filter the CAD/EZH2/RPA CUT&Tag data against an IgG control. While modest changes in the DNA-binding profile are expected after IgG filtering (due to the removal of weaker binding sites), the overall peak intensity in CUT&Tag data is unlikely to change significantly, particularly when compared to less robust genome-binding methods like ChIP-seq.

Appendices

Cold Spring Harbor Laboratory

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

HOME | SUBMIT | FAQ | BLOG
| ALERTS / RSS | RESOURCES
| ABOUT | CHANNELS

Search

Advanced Search

New Results

Follow this preprint

Caspase-Activated DNase localizes to cancer causing translocation breakpoints during cell differentiation

Dalal Alsowaida, Brian D. Larsen, Sarah Hachmer, Mehri Azimi, Eric Arezza, Steve Brunette, Steven Tur, Carmen G. Palii, Bassam Albraidy, Claus S. Sorensen, Marjorie Brand, F. Jeffrey Dilworth, Lynn A. Megeney

doi: https://doi.org/10.1101/2024.09.24.614809

This article is a preprint and has not been certified by peer review [what does this mean?].

0 0 0 1 0 0 10

Abstract | Full Text | Info/History | Metrics

Preview PDF

Abstract

Caspase activated DNase (CAD) induced DNA breaks promote cell differentiation and therapy-induced cancer cell resistance. CAD targeting activity is assumed to be unique to each condition, as differentiation and cancer genesis are divergent cell fates. Here, we made the surprising discovery that a subset of CAD-bound targets in differentiating muscle cells are the same genes involved in the genesis of cancer-causing translocations. In muscle cells, a prominent CAD-bound gene pair is Pax7 and Foxo1a, the mismatched reciprocal loci that give rise to alveolar rhabdomyosarcoma. We show that CAD-targeted breaks in the Pax7 gene are physiologic to reduce Pax7 expression, a prerequisite for muscle cell differentiation. A cohort of these CAD gene targets are also conserved in early differentiating T cells and include genes that spur leukemia/lymphoma translocations. Our results suggest the CAD targeting of translocation prone oncogenic genes is non-pathologic biology and aligns with initiation of cell fate transitions.

Competing Interest Statement

The authors have declared no competing interest.

Copyright The copyright holder for this preprint is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

bioRxiv and medRxiv thank the following for their generous financial support:

The Chan Zuckerberg Initiative, Cold Spring Harbor Laboratory, the Sergey Brin Family Foundation, California Institute of Technology, Centre National de la Recherche Scientifique, Fred Hutchinson Cancer Center, Imperial College London, Massachusetts Institute of Technology, Stanford University, University of Washington, and Vrije Universiteit Amsterdam.

Back to top

Previous Next

Posted September 26, 2024.

Download PDF
Print/Save Options
Email
Share
Citation Tools
Get QR code

Post Like 0

COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv

Subject Area

Genomics

Subject Areas

All Articles

Animal Behavior and Cognition
Biochemistry
Bioengineering
Bioinformatics
Biophysics
Cancer Biology
Cell Biology
Clinical Trials*
Developmental Biology



Self-inflicted DNA breaks in cell differentiation and cancer

Author: Jan Benada,Dalal Alsowaida,Lynn A. Megeney,Claus S. Sørensen

Publication: Trends in Cell Biology

Publisher: Elsevier

Date: October 2023

© 2023 Elsevier Ltd. All rights reserved.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK CLOSE WINDOW

© 2024 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Data Security and Privacy | For California Residents | Terms and ConditionsComments? We would like to hear from you. E-mail us at customercare@copyright.com

Privacy - Terms

CellPress

Trends in Cell Biology

Review

Self-inflicted DNA breaks in cell differentiation and cancer

Jan Benada,[1] Dalal Alsowaida,[2,3] Lynn A. Megeney,[2,*] and Claus S. Sørensen [1,*]

Self-inflicted DNA strand breaks are canonically linked with cell death pathways and the establishment of genetic diversity in immune and germline cells. Moreover, this form of DNA damage is an established source of genome instability in cancer development. However, recent studies indicate that nonlethal self-inflicted DNA strand breaks play an indispensable but underappreciated role in a variety of cell processes, including differentiation and cancer therapy responses. Mechanistically, these physiological DNA breaks originate from the activation of nucleases, which are best characterized for inducing DNA fragmentation in apoptotic cell death. In this review, we outline the emerging biology of one critical nuclease, caspase-activated DNase (CAD), and how directed activation or deployment of this enzyme can lead to divergent cell fate outcomes.

Highlights

Nucleases are employed physiologically to inflict DNA lesions that promote processes such as apoptosis, the evolution of the germline, and adaptation of the immune system.

Increasing evidence suggests that nucleases also inflict regulated DNA damage in other processes such as during cell differentiation.

During cancer development, nuclease-mediated genomic lesions promote tumorigenesis through enhancing genome instability.

Cancer cells can hijack nucleases such as caspase-activated DNase (CAD) to activate DNA damage response checkpoints to induce cell-cycle arrest and thus promote cancer cell survival.

Self-inflicted DNA breaks

DNA macromolecules are stable carriers of the genetic information for our development and function, even though it is a pivotal molecule, DNA breaks occur in a large variety of settings. Sources can be exogenous impacts such as radiation or reactive chemicals; moreover, endogenous cellular processes are frequently involved in the creation of nuclear strand breaks. Endogenously derived breaks can be induced by enzyme catalysis or via chemical impacts, for example, through reactive oxidative species generated during metabolic processes. Enzyme-mediated DNA breakage is catalyzed by nucleases that participate in multiple cellular and physiological processes; thus, **self-inflicted DNA breaks** (see Glossary) occur in multiple settings [1–6]. In addition, nuclease activities are integrated into DNA repair pathways to remove damaged bases or to trim DNA ends, a mechanism that is covered extensively in other reviews [7–9]. When DNA strand breaks arise as a consequence of exposure to genotoxic agents or endogenous reactive chemical species, outcomes depend on lesion load and the DNA repair capacity of the cell. Such lesions generally occur in a sequence-independent manner as they are unregulated and stochastic by nature. In normal cells, unrepaired DNA breaks can lead to acquired mutations that promote proliferation or may separately induce cell death or senescence [10,11].

In this review, we focus on the emerging biology of specific endonuclease-mediated endogenous break formation in differentiation and cancer. The following discussion is predominantly centered on discoveries from several laboratories that link CAD-mediated DNA breaks with cancer development and responses to therapy [12–14]. While CAD is best understood from its canonical function in cell death, where it catalyzes the cleavage of nuclear DNA into internucleosomal fragments [15,16], this nuclease has emerged as a nonlethal regulator of physiological cell function [13,17]. Several additional nucleases with essential roles also operate through self-inflicted breaks. This is prominent in the generation of genetic diversity in the germ line and immune system by forming DNA breaks that entrain self-different cells in the organism. For example, during meiosis DNA **double-strand breaks (DSBs)** are induced by the SPO11 nuclease, driving chromosomal

[1]Biotech Research and Innovation Centre, University of Copenhagen, Ole Maaløes Vej 5, Copenhagen 2200 N, Denmark
[2]Sprott Centre for Stem Cell Research, Ottawa Hospital Research Institute and the Departments of Medicine and Cellular and Molecular Medicine, University of Ottawa, Ottawa, K1H 8L6, Canada
[3]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia

*Correspondence:
lmegeney@ohri.ca (L.A. Megeney) and claus.storgaard@bric.ku.dk (C.S. Sørensen).

© 2023 Elsevier Ltd. All rights reserved.

exchanges during gametogenesis [1]. In addition to meiosis, the human adaptive immune system represents possibly the most studied case of the crucial role of nucleases in development. The diversification in antigen receptor and Ig genes is governed by precisely targeted DSBs generated during V(D)J recombination, somatic hypermutation (SHM), and class switch recombination (CSR). V(D)J recombination directs the assembly of Ig variable region in B cells and of T cell receptor variable region exons [18]. The process is initiated by recombination-activating gene (RAG) endonuclease, which comprises RAG1 and RAG2 subunits. RAG creates targeted DSBs between recombination signal sequences (RSSs) and the coding sequence of the V, D, and J segments [19,20]. Subsequently, RAG2 directs the repair of generated breaks into the alternative nonhomologous end-joining (NHEJ) pathway. Further Ig gene diversification is driven by CSR and SHM, which are initiated by activation-induced cytidine deaminase (AID) that promotes DNA break formation [18]. Here, we will compare and contrast these genome remodeling events with the newly described roles of CAD nuclease.

CAD-dependent DNA breaks in cellular differentiation

Following closely on the discovery of the protein machinery that mediated apoptosis/programmed cell death, a number of groups independently confirmed a role for these same factors in the induction of cellular differentiation. The initial examination of caspase proteases in this context was predicated on the observation that differentiation and apoptosis shared a cohort of morphologic features, including cytoskeletal reorganization, membrane fusion events, chromatin alteration, and so forth [21]. Caspase 3 was shown to be required for the differentiation of erythrocytes, keratinocytes, skeletal myoblasts, and macrophages in mammalian organisms as well as neural differentiation and spermatid maturation in Drosophila (reviewed in [22,23]).

CAD in differentiation

As noted earlier, CAD is the DNA shearing termini for cell death signaling pathways and is engaged by a caspase 3 directed cleavage of a nascent inhibitor, inhibitor of caspase-activated DNase (ICAD) [24]. Despite this death-centric function, CAD has been shown to be a primary caspase 3 substrate that induces the differentiation process in skeletal muscle progenitor cells [17]. Here, caspase 3 activation engages CAD activity through direct cleavage of ICAD, which mirrors the canonical signaling process in cell death. Once activated, CAD inflicts transient strand breaks in the p21 promoter, which act as a priming event to engage p21 gene expression [17] (Figure 1). The elevation of p21, promotes differentiation across a broad spectrum of cell lineages by limiting cell cycle progression [25]. This gene induction capacity of CAD through a promoter-directed strand break is reminiscent of the gene induction strategy utilized by topoisomerase IIβ (Top2β) in androgen-mediated prostate cancer cell gene expression. Here, Top2β induces DSBs in promoter regions, which act as assembly points for DNA repair machinery and transcription factors to coordinate and induce target gene expression [26]. In addition to acting as a transcription initiation foci for basal transcription machinery, CAD-directed promoter strand breaks may also act to unmask binding sites for individual transcription factors, thereby priming gene expression of selected targets. CAD-induced breaks have been reported to robustly activate NFKB [27], which has known binding sites on the p21 promoter [28].

Unresolved questions to CAD function in differentiation

What remains unknown is whether CAD-induced breaks are autonomous from DNA repair or whether CAD integrates more directly with the repair process during muscle cell differentiation, akin to the role of Top2β? Certainly, the kinetics of CAD-induced DNA damage and repair are temporally linked, with a rapid transition from measurable damage to the assembly of repair protein foci constituted by the base excision repair protein XRCC1 and the NHEJ regulatory kinase DNA-PK, both of which are essential for resolving the CAD/differentiation induced strand breaks

Glossary

Ataxia-telangiectasia mutated/ataxia-telangiectasia and Rad3 related (ATM/ATR) kinases: the two DNA damage response kinases ATM and ATR, which act as apical regulators of the DNA damage response.

Caspase-activated DNase–inhibitor of CAD (CAD-ICAD): CAD is a nuclease that cuts the DNA during apoptosis or promotes cell differentiation. (ICAD) binds CAD to inhibit its dimerization, caspases cleave ICAD to promote formation of highly active DNA DSB-inducing CAD homodimers.

CAD-dependent SSBs (CdSSBs): CAD-ICAD-dependent DNA single-strand breaks that are induced several hours after radiation.

Doxorubicin (Adriamycin): a topoisomerase II (Top2) inhibitor used in clinics as an anticancer drug. Mechanistically, doxorubicin inhibits Top2 in religation step and thus causes double-strand breaks. In addition, doxorubicin acts also as DNA intercalator and can generate oxidative stress.

Double-strand breaks (DSBs) and single-strand breaks (SSBs): two of the most common genomic lesions are the DNA DSBs and SSBs, where both or only one of the DNA strands are broken.

Etoposide: a topoisomerase II (Top2) inhibitor used in clinics as an anticancer drug. Etoposide inhibits Top2 in religation step and induces double-strand breaks.

Mitochondrial outer membrane permeabilization (MOMP): part of the apoptosis process where pores are forming in the outer mitochondrial membrane. This releases proteins that mediate the activation of proteases termed caspases, whose activities advance the cell death process. In case only a minority of mitochondria undergo MOMP, we speak of 'minority MOMP'. Minority MOMP leads only to limited activation of caspases, insufficient to facilitate the apoptotic program.

Self-inflicted DNA breaks: self-inflicted DNA breaks are DSBs or SSBs generated by nucleases without the purpose of canonical DNA processing during DNA repair.

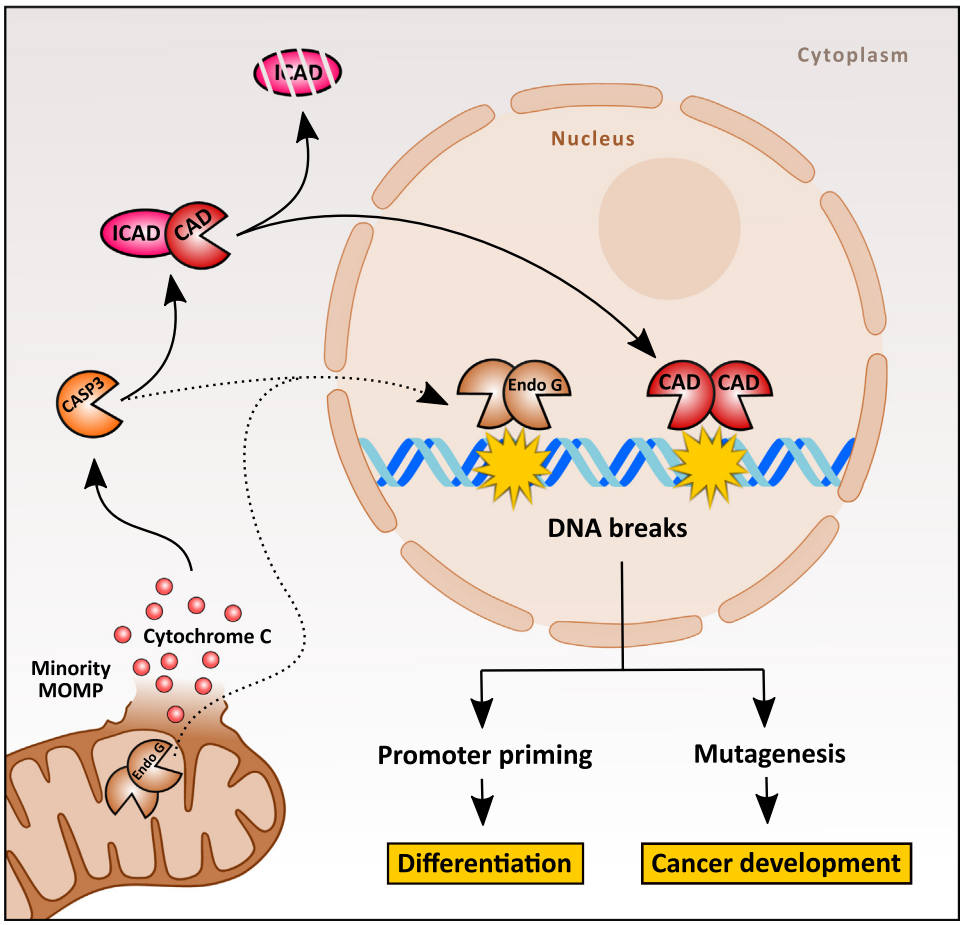


Figure 1. Self-inflicted breaks in differentiation and cancer development. Minority mitochondrial outer membrane permeabilization (minority MOMP) can lead to sublethal caspase activation. Activated CASP3 (caspase 3) mediates cleavage of inhibitor of caspase-activated DNase (ICAD) that leads to dimerization of caspase-activated DNase (CAD). CAD homodimer then causes DNA breaks. Of note, the caspase-dependent cleavage of ICAD is necessary for CAD dimerization, but both CAD homodimer and CAD-ICAD heterodimer can translocate to the nucleus. CAD-dependent self-inflicted breaks can have a physiological role, for example, transient breaks in promoter regions can prime them to engage gene expression and induce differentiation. By contrast, self-inflicted breaks have mutagenic potential and can drive cancer development. Similarly, upon minority MOMP, homodimeric Endo G nuclease translocates to the nucleus, where it mediates mutagenic breaks and translocations. The relationship of Endo G to caspase-mediated signaling remains unclear (broken lines). Abbreviation: ICAD, inhibitor of caspase-activated DNase.

[29,30]. A direct physical association between CAD and DNA repair proteins in this cell lineage is speculative, yet ICAD does accommodate interaction with critical DNA repair-associated kinases in cancer cells (see the section on CAD in cancer therapy responses). However, unlike **CAD-ICAD** function in cancer cell biology, CAD-associated muscle cell differentiation is dependent on caspase 3-mediated cleavage of ICAD, which implies that any direct communication between a DDR process and CAD may require direct physical association of repair proteins and CAD, rather than ICAD per se. Further characterization of CAD-ICAD protein complexes will be critical to address the mechanism by which this DNase bridges promoter-directed strand breaks to the initiation of gene expression during muscle cell differentiation. In addition, it will be of interest to define the relationship that governs CAD accessibility to DNA, and whether the nuclease is influenced by separate modifications in chromatin architecture that are coincident to caspase-

mediated activation. Further to this point, caspase 7 targets and removes the matrix attachment region (MAR) protein SATB2 to alter chromatin loop structure, which then leads to derepression of differentiation-inducing factors [31]. As SATB2 cleavage is concurrent to CAD activation, a potential model may be envisioned where convergent caspase activities ensure proficient CAD genome targeting by removing potential structural impediments akin to SATB2.

The extensive number of CAD-induced strand breaks and repair foci in differentiating muscle cells support the contention that CAD broadly targets the genome to elicit cell fate outcomes [17,29]. This observation, coupled with the chromatin dissolution capacity of CAD during apoptosis, suggests CAD is likely to engage other strand break events that elicit differentiation, beyond a cohort of transcription-sensitive promoters such as p21. These CAD-targeted sites are likely to involve strand breaks in intragenic regions or chromatin boundaries, similar to the CAD-directed genome fracturing that occurs during apoptosis, which disrupts gene expression within broader chromatin domains [24,32]. Alternatively, CAD may target specific gene loci, where intragenic strand breaks could facilitate impaired transcription read-through, thereby reducing gene expression that may limit the differentiation program. Mapping the CAD-targeted genome will be an essential next step in defining the full repertoire of CAD reprogramming activity that manages the cell differentiation program, and whether this is integrated with or distinct from the target domain of chromatin modifiers such as SATB2.

Finally, it is germane to consider that CAD may be a crucial genome modifier for cell differentiation in general. As noted earlier, caspase 3-induced cell differentiation is a broadly conserved phenomenon, required for multi-lineage cell differentiation across the metazoan phyla [23]. While caspase-induced CAD activity has only been directly implicated in skeletal muscle progenitor cell differentiation, recent observations in other cell lineages are strongly suggestive of a caspase/CAD-mediated genome reprogramming event. Late-stage erythropoiesis is dependent on caspase-initiated chromatin condensation, which is coincident to nuclear envelope alterations and engagement of DNA repair responses such as histone H2A.X phosphorylation [33,34]. Presumably, CAD is the operative DNase engaged in this process. Neural development offers another intriguing model to explore regarding CAD-dependent genome reprogramming. Caspase-mediated signaling events have been well described as a conserved inductive cue for neuronal differentiation [35,36]. Furthermore, neuron-specific promoters have been shown to be frequent targets of DNA strand break/repair in differentiation and maturation [37–39]. While CAD activity has not been directly investigated in this context, the fact that neural differentiation is caspase 3-mediated and concurrent to extensive DNA strand breaks is a remarkably similar phenomenon to the caspase/CAD activity that manages genome reprogramming in muscle cell differentiation. Finally, one must consider whether CAD functions outside cell cycle exit/termination. For example, replication stress is known to be a source of extensive DNA damage in actively dividing cells [40]. However, CAD is operative in differentiating muscle cells at 12–24 h post low serum induction of differentiation, suggesting that the temporally linked DNA strand breaks are not a replication-related event.

Self-inflicted breaks in cancer development

Chromosomal rearrangements

The extensive genome remodeling in physiological processes is crucial; however, it also comes with a potential cost in damaging genomic lesions. Illegitimate structural variants (SVs) are potentially pathologic genomic changes of chromosomes including inversions, deletions, and translocations which are features of cancer. Unbalanced translocation is among the most common SVs in human cancers. Both unbalanced and reciprocal translocations are cut-and-paste class of SVs which are reconstructed from inaccurate religation of DNA ends [41]. SVs influence the chromosome topology and the epigenetic landscape in a tumor type-specific way [42], but not

necessarily by altering gene expression [42,43]. The spatial organization of chromosomes inside the nucleus, the intrachromosomal organization of genes, and the GC genome sequence of chromosome territories are all factors that can trigger chromosomal rearrangements leading to cancer [44]. A subgroup of chromosomal translocations occurs at a specific defined 20–600 base pair (bp) zones named fragile zones that are more prone to DSBs; several factors explaining the breakage at these sites are reviewed in [45]. Translocations involved in fragile zones implicate the separation of DNA double strands under torsional or transcription stress. This is followed by exposure to damaging agents particularly AID or increased reactive oxygen species (ROS) causing persistent DNA lesions in mammalian fragile zones [46].

V(D)J recombination is a well-studied example of important DNA break-induced chromosomal translocations that provide the template for antibody diversification. Translocations themselves are the operative mechanism where an exchange of chromosomal segments generates SVs to produce novel antibody-producing loci. Despite the importance of this physiological process, an error in V(D)J recombination can lead to deleterious RAG breaks that may induce chromosomal translocations and genome instabilities. Aberrant RAG activity has been shown to be the trigger of mutagenic SVs in lymphoid blast crisis (LBC), which is the advanced stage of chronic myeloid leukemia (CML) [47]. Moreover, p53-deficient mice lacking RAG2 displayed unique genomic structures and gene amplifications generated by a mechanism known as breakage-fusion-bridge (BFB). This mechanism gives rise to unstable and complex rearrangements provoking malignancy in T cell lymphoma [48]. It is also established that RAG recombinase activity is the main driver of ETV6-Runx1 fusion in acute lymphoblastic leukemia (ALL) [49].

DNA fragmentation by endonucleases

More recently, caspase activity has been identified in the context of DNase-mediated oncogenic strand break formation and SVs. This phenomenon was reported by two groups independently, in reference to spontaneous DSBs (spDSBs) in cancer cells. Here, each group noted that the sublethal activation of caspase 3, and the generation of spDSBs, was coincident with the phenomenon of cancer cell survival and expansion [50,51]. Ichim *et al.* and others [14,50,52,53] have noted that the cancer cell activation of caspase 3 and the resulting DNA strand breaks were driven by a transient **mitochondrial outer membrane permeabilization** (referred to as minority **MOMP**), leading to engagement of the intrinsic/mitochondrial cell death pathway (Figure 1). This cancer cell rheostatic response is reminiscent of the mitochondrial-driven caspase 3 activity that promotes cell differentiation [54–56], suggesting that these widely divergent cell behaviors derive from identical molecular signals.

The work from Liu *et al.* observed that engagement of caspase 3 was concurrent to activation of the mitochondrial Endo G nuclease, which translocated to the nucleus to inflict the growth advantageous spDSBs [51]. Endo G resides within the mitochondria and impacts the replication of this organelle's DNA content [57] (Figure 1). This group also noted that loss of Endo G provided protection from both radiation and Myc-induced oncogenic transformation [51,58]. The role of Endo G nuclease as the operative agent in cancer DNA strand break formation is further strengthened by the observation that Endo G cleaves the MLL gene at a breakpoint cluster region (*MLLbcr*) to resolve stalled forks [58]. The MLLbcr locus is a well-defined leukaemogenic region, characterized by cancer-causing strand break and translocation events. In addition, cleavage of MLL gene has been reported in treatment-related breakpoints, reviewed in [24]. For instance, Top2β is the main inducer of **etoposide** breaks in MLL gene [59].

Although nuclear Endo G provides a strand break mechanism, its relationship to caspase-mediated signaling events in cancer cells is not entirely clear. For example, Endo G is not

caspase-sensitive, and it remains unclear how this nuclease gains access to a nuclear domain in a challenged cancer cell. Accordingly, CAD has been identified as a major nuclease in the development of cancer cell DNA SVs in response to increased caspase 3 function [14,53,60] (Figure 1). Here, CAD mediates extensive DNA strand break and micronuclei formation, along with chromosomal misalignments in both zebrafish and murine models [14]. Collectively, these observations support CAD as the de facto nuclease in caspase-driven cancer cell genome organization; yet, how CAD targets the genome to augment cell growth remains undefined. Future genome-wide studies of CAD and Endo G-mediated strand breaks in cancer genomes will reveal the molecular etiology of the breaks. What remain interesting open questions are the DNA sequences, chromatin structures, and epigenetic factors that influence CAD and Endo G to inflict strand breaks in cancer.

Self-inflicted breaks in cancer treatment and therapy evasion

Many cytotoxic therapies rely on the generation of DNA damage via a direct source (such as ionizing radiation) or indirectly by modulating the action of endonucleases. A prime example of the latter is the topoisomerase-targeting therapies. Topoisomerases are nucleases that possess both endonuclease and ligase activity. They play a crucial role in the regulation of DNA topology, and this is particularly vital for dividing cells. To this date, a large number of topoisomerase-targeting drugs have been discovered and several are used in cancer treatment. Their mechanism of action is commonly via blocking the relegation of nuclease-generated breaks. Since this topic has been extensively covered by several recent reviews [61–63], we will focus our attention on the emerging role of the CAD nuclease in therapy evasion.

Self-inflicted breaks in evasion from DNA-damaging therapy

Cancer cell response to irradiation has been linked with a precisely timed, yet unexplained wave of DNA damage occurring hours after the initial breaks that are inflicted directly by radiation energy [64]. Unexpectedly, this temporally delayed DNA damage has been shown to be a nuclease-driven break event that promotes resistance of cancer cells to radiotherapy and **doxorubicin**-based chemotherapy [13]. Cancer cells can hijack CAD nuclease to trigger a second wave of DNA breaks following DNA-damaging treatment. Secondary CAD-induced breaks signal to strengthen the G2 cell cycle checkpoint, preventing premature mitotic entry and thus protecting cancer from therapy-evoked DNA damage (Figure 2). This cancer cell-specific adaptation of CAD function appears caspase independent. Additionally, in contrast to CAD canonical function in apoptotic DNA fragmentation or cell differentiation, CAD can operate in form of the CAD-ICAD dimer which inflicts predominantly **single-strand breaks (SSBs)**. As opposed to DSBs created by initial DNA-damaging therapy, the **CAD-dependent SSBs (CdSSBs)** can be repaired with fast kinetics. CdSSBs can serve as temporary cell cycle checkpoint boosters, providing the cancer cells with time to repair harmful DSBs before they enter mitosis [13].

In the radiotherapy response, CdSSBs do not occur at random, but are enriched at CCCTC-binding factor (CTCF) sites of chromatin association [13]. The CTCF-binding sites are important genomic elements that help to maintain higher-order chromatin architecture and epigenetically regulated domains [65]. CdSSBs at these sites may contribute to the global chromatin response upon DNA damage. This break location at CTCF sites also appears to be linked to other signaling mechanisms to modify the chromatin microenvironment and impact DSB formation, that is, CAD-dependent breaks signal to **ataxia-telangiectasia mutated/ataxia-telangiectasia and Rad3 related (ATM/ATR)** to phosphorylate chromatin-associated KRAB-associated protein-1 (KAP1) on Ser 824 [13]. Prior studies have shown that KAP1 phosphorylation controls DSB-induced chromatin decondensation and its loss sensitizes cells to DNA-damaging agents [66]. The ATM/ATR-dependent phosphorylation of chromatin targets, such as KAP1, governs the duration of

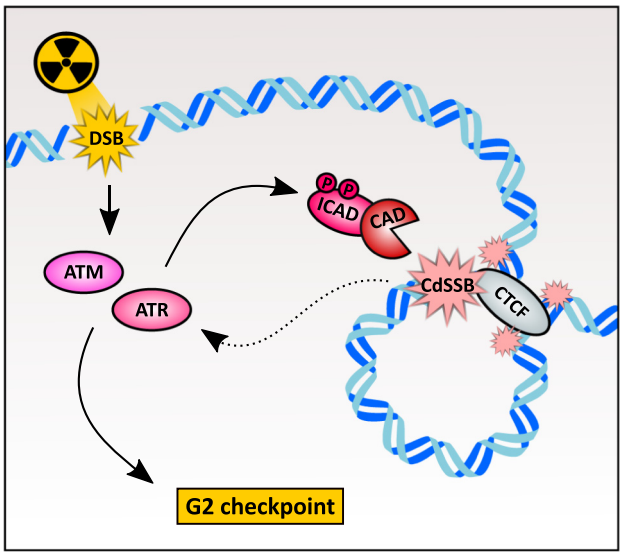


Figure 2. Cancer cells hijack CAD nuclease to trigger a second wave of DNA breaks following DNA-damaging treatment. The initial DNA double-strand breaks (DSBs) that are inflicted by radiotherapy activate ataxia-telangiectasia mutated and ataxia-telangiectasia and Rad3-related (ATM/ATR) kinases. ATM/ATR directly phosphorylate inhibitor of caspase-activated DNase (ICAD) and thus prompt CAD-ICAD (caspase-activated DNase) dimer to chromatin. Subsequently, CAD-ICAD dimer inflicts predominantly DNA single-strand breaks [termed CAD-dependent SSBs (CdSSBs)], which are present at CCCTC-binding factor (CTCF) sites. CdSSBs signal in positive feedback loop back to ATM/ATR in order to boost the protective G2 checkpoint and therefore promote therapy resistance. Molecular factors regulating CdSSBs processing and signaling to ATM/ATR remain to be elucidated (broken line). Abbreviation: SSBs, single-strand breaks.

the G2 checkpoint [67], which CAD itself is intricately linked to, that is, localization of a CAD-ICAD complex to chromatin is actively directed by ATM/ATR-dependent phosphorylation of ICAD [13].

Residual tumor cells represent a serious challenge in cancer treatment, potentially enabling the emergence of treatment-resistant tumor cells. Given the aforementioned information, the integrated signaling function of the CAD-ICAD complex may represent a cancer-specific vulnerability that could be exploited to curtail therapy resistance and enhance radio/chemosensitization. Indeed, healthy cells are predicted to be much less affected by targeting CAD as, in contrast to the majority of cancer cells, they harbor functional p53 and retinoblastoma (Rb) proteins providing them with a functional G1 checkpoint [68]. This is supported by a murine xenograft model of radiotherapy, where the deletion of CAD in tumor cells led to effective inhibition of tumor growth [13].

Self-inflicted breaks in the context of non-DNA-damaging therapies

Interestingly, induction of CAD-dependent breaks was also reported upon non-DNA-damaging small-molecule targeted therapies [12]. Treatment with small-molecule inhibitors of EGFR, ALK, KRAS, or BRAF resulted in sublethal CAD-dependent induction of DNA damage and activation of ATM, which in turn promoted tumor cell survival upon these therapies. The precise nature of the CAD-dependent breaks induced in this context has not been investigated in detail; however, the induction is caspase dependent and breaks are presumably predominantly DSBs [12]. Thus, it seems the regulation of and targeting of CAD-dependent breaks may differ upon acute DNA-damaging therapy versus in non-DNA-damaging therapies. Nevertheless, in this context, CAD-dependent breaks conferred an ability to -target resistant cells with ATM inhibitor. Taken together, this provides a rationale for the design of combinatorial treatments with oncogene-targeted therapies along with the use of inhibitors of the DNA damage response [12].

The regulation of CAD-dependent break formation is not well understood

The ability of CAD to mediate DNA damage response signaling via break formation evokes a number of so far unresolved mechanistic questions. In particular, it is not clear how cells regulate the timing of CdSSBs. The wave of CdSSBs was shown to peak at approximately 24 h after irradiation [13]. By contrast, activation of ATM/ATR occurs rapidly within seconds to minutes after

irradiation [69]. For that reason, besides the ATM/ATR-dependent phosphorylation of ICAD, there is likely another layer of regulation that restricts the timing of CdSSBs. Moreover, to activate ATR-dependent checkpoint signaling, at least a subset of CdSSBs needs to undergo further processing leading to the generation of more extended ssDNA. Finally, it is an intriguing possibility that a small subset of CdSSBs can be converted to DSBs, which would evoke ATM-mediated response. Therefore, it is of great interest to identify molecular factors and pathways which contribute to the regulation of signaling from CdSSBs.

Concluding remarks

The identification of CAD in numerous DNA break-mediated cell behaviors was unanticipated and not without controversy, as releasing activation of a known apoptotic nuclease would be assumed to inflict catastrophic and irreversible genomic damage on a cell. This begs a series of questions (see Outstanding questions); chief among these is how CAD activity is fine-tuned to achieve such divergent outcomes (Figure 3)? Clearly, there is a rheostat-like mechanism allowing

Outstanding questions

How common are CAD-induced breaks in differentiation processes, and are such breaks integrated with DNA repair pathways?

What are the molecular features and pathways that determine the timing and genomic locations of CAD and Endo G-mediated DNA breaks? Moreover, how are CdSSBs converted to DNA damage checkpoint signaling lesions?

Can CAD and Endo G be targeted therapeutically? Will such pharmacological targeting provide new options in cancer therapy, potentially by acting as tumor-specific radiosensitizing agents?

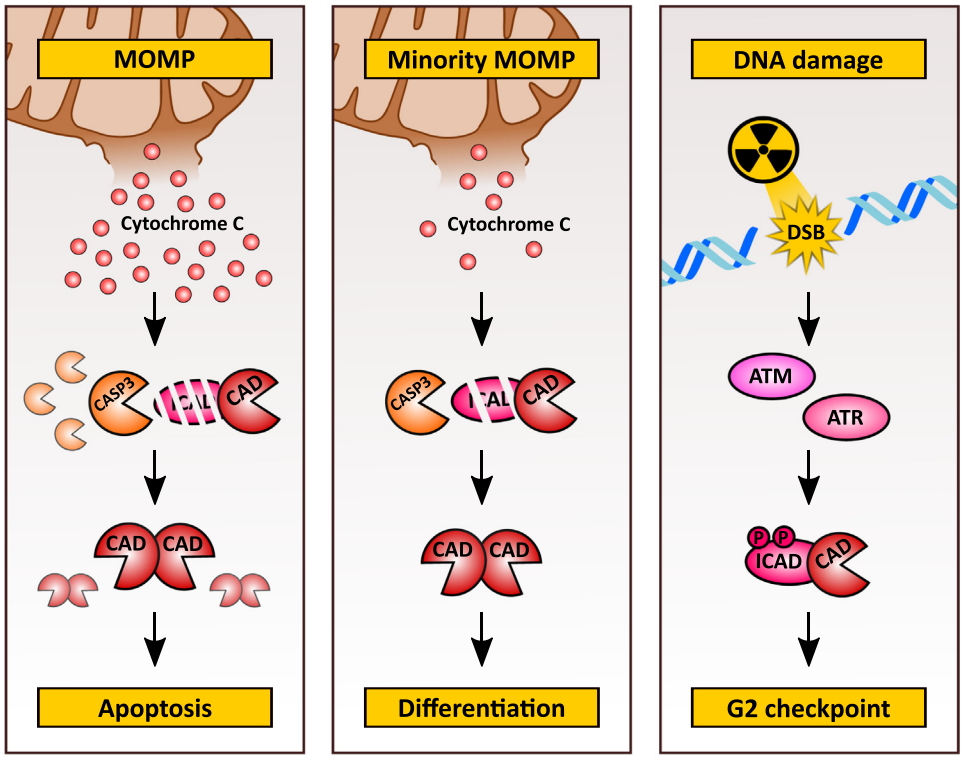


Figure 3. Summary of different modes of CAD activation and respective functional outputs. During the apoptosis, MOMP leads to the release of cytochrome C from mitochondria and subsequent activation of caspase-3. Caspase 3 then cleaves ICAD at D224 and D117, resulting in CAD dimerization. Next, to the full extent activated CAD homodimers fragment DNA, contributing to the cell death program. The situation of CAD activation in differentiation is reminiscent of the one in the apoptotic program, the main difference is in the extent of the activated processes. MOMP is only limited (termed minority MOMP) and low concentration of cytochrome C leads only to partial activation of Caspase 3. Partially active Caspase 3 cleaves ICAD less efficiently at a single D117 residue and partially active CAD generates a limited amount of DNA breaks which can trigger differentiation programs, but do not result in genome-wide DNA fragmentation. By contrast, activation of CAD-ICAD upon DNA damage is Caspase-independent and mediated by ATM/ATR phosphorylation of ICAD at S107 and S257. Activated CAD-ICAD heterodimer then facilitates G2 checkpoint maintenance via induction of CdSSB and consequent boost of checkpoint signaling. Abbreviations: ATM, ataxia-telangiectasia mutated; ATR, ataxia-telangiectasia and Rad3-related; CAD, caspase-activated DNase; DSB, double-strand break; ICAD, inhibitor of caspase-activated DNase; MOMP, mitochondrial outer membrane permeabilization.

for either a transient caspase-dependent or independent activation of CAD. In the cancer resistance scenario (which follows irradiation), this appears to be managed through the integration of ATM/ATR-dependent signals which phosphorylate ICAD to control CAD per se. In the cancer-origin scenario, this is managed through a failed completion of the canonical mitochondrial cell death pathway, where death-escaping cells derive the benefits of CAD genome reorganization. The control of CAD in differentiation is much less clear. CAD may utilize the same controls that govern mitochondrial-mediated activation of caspase 3, yet the extensive break formation in this cell fate must integrate other chromatin modifications without promoting complete genome dissolution.

Finally, we must consider whether a second CAD-ICAD protein complex exists in mammalian cells, as occurring in Drosophila, where two functional CAD-ICAD pairs have been identified [70]. CAD-like homology, referred to as CIDE domains [cell death-inducing DNA fragmentation factor 45 (DFF45)-like effector] have been identified in a larger family of mammalian proteins, where these factors have been linked to other cellular functions such as lipid metabolism [71]. To date, no DNA targeting activity has been explored for the larger mammalian CIDE protein family; however, such a function is reasonable based on the observations in Drosophila. The existence of interchangeable CIDE proteins may explain the lack of overt phenotype for the CAD gene-targeted mouse model, which does not have pronounced deficits in cell differentiation [72]. Presumably, gene-targeted deletion of CAD may lead to a compensatory increase in the expression of related CIDE proteins, masking functional deficits that may occur with loss of CAD per se. Answering these outstanding aspects of CAD biology will be an essential step to further our understanding of cell fate control and may also serve as a conduit for the rational design of strategies to limit therapy-induced cancer cell resistance.

Acknowledgments

We thank our funding sources for generous support of our laboratories (L.A.M. and C.S.S.). We thank Brian Larsen for helpful comments on the manuscript. Figures were created with the freeware Inkscape (Inkscape.org).

Declaration of interests

The authors declare no competing interests.

References

1. Betermier, M. *et al.* (2020) Coupling DNA damage and repair: an essential safeguard during programmed DNA double-strand breaks? *Trends Cell Biol.* 30, 87–96
2. Luo, S. *et al.* (2022) DNA damage response and repair in adaptive immunity. *Front. Cell Dev. Biol.* 10, 884873
3. Thada, V. and Greenberg, R.A. (2022) Unpaved roads: how the DNA damage response navigates endogenous genotoxins. *DNA Repair (Amst)* 118, 103383
4. Schubert, I. (2021) Boon and bane of DNA double-strand breaks. *Int. J. Mol. Sci.* 22, 5171
5. Lezaja, A. and Altmeyer, M. (2021) Dealing with DNA lesions: when one cell cycle is not enough. *Curr. Opin. Cell Biol.* 70, 27–36
6. Ma, J.Y. *et al.* (2022) Germline cell de novo mutations and potential effects of inflammation on germline cell genome stability. *Semin. Cell Dev. Biol.* Published online November 12, 2022. https://doi.org/10.1016/j.semcdb.2022.11.003
7. Paull, T.T. (2021) Reconsidering pathway choice: a sequential model of mammalian DNA double-strand break pathway decisions. *Curr. Opin. Genet. Dev.* 71, 55–62
8. Gnugge, R. and Symington, L.S. (2021) DNA end resection during homologous recombination. *Curr. Opin. Genet. Dev.* 71, 99–105
9. Manils, J. *et al.* (2022) Exonucleases: degrading DNA to deal with genome damage, cell death, inflammation and cancer. *Cells* 11, 2157
10. Clarke, T.L. and Mostoslavsky, R. (2022) DNA repair as a shared hallmark in cancer and ageing. *Mol. Oncol.* 16, 3352–3379
11. Niedernhofer, L.J. *et al.* (2018) Nuclear genomic instability and aging. *Annu. Rev. Biochem.* 87, 295–322
12. Ali, M. *et al.* (2022) Small-molecule targeted therapies induce dependence on DNA double-strand break repair in residual tumor cells. *Sci. Transl. Med.* 14, eabc7480
13. Larsen, B.D. *et al.* (2022) Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress. *Science* 376, 476–483
14. Haimovici, A. *et al.* (2022) Spontaneous activity of the mitochondrial apoptosis pathway drives chromosomal defects, the appearance of micronuclei and cancer metastasis through the Caspase-Activated DNAse. *Cell Death Dis.* 13, 315
15. Kulbay, M. *et al.* (2021) The role of the DFF40/CAD endonuclease in genomic stability. *Apoptosis* 26, 9–23
16. Sakahira, H. *et al.* (1998) Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. *Nature* 391, 96–99
17. Larsen, B.D. *et al.* (2010) Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4230–4235
18. Feng, Y. *et al.* (2020) AID in antibody diversification: there and back again. *Trends Immunol.* 41, 586–600
19. Liu, C. *et al.* (2022) Structural insights into the evolution of the RAG recombinase. *Nat. Rev. Immunol.* 22, 353–370

20. Zhang, Y. *et al.* (2022) The role of chromatin loop extrusion in antibody diversification. *Nat. Rev. Immunol.* 22, 550–566
21. Fernando, P. *et al.* (2002) Caspase 3 activity is required for skeletal muscle differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11025–11030
22. Baena-Lopez, L.A. *et al.* (2018) Non-apoptotic Caspase regulation of stem cell properties. *Semin. Cell Dev. Biol.* 82, 118–126
23. Bell, R.A.V. and Megeney, L.A. (2017) Evolution of caspase-mediated cell death and differentiation: twins separated at birth. *Cell Death Differ.* 24, 1359–1368
24. Larsen, B.D. and Sorensen, C.S. (2017) The caspase-activated DNase: apoptosis and beyond. *FEBS J.* 284, 1160–1170
25. Parker, S.B. *et al.* (1995) p53-independent expression of p21Cip1 in muscle and other terminally differentiating cells. *Science* 267, 1024–1027
26. Ju, B.G. *et al.* (2006) A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science* 312, 1798–1802
27. Liu, X. *et al.* (2017) Self-inflicted DNA double-strand breaks sustain tumorigenicity and stemness of cancer cells. *Cell Res.* 27, 764–783
28. Nicolae, C.M. *et al.* (2018) NFkappaB regulates p21 expression and controls DNA damage-induced leukemic differentiation. *Oncogene* 37, 3647–3656
29. Al-Khalaf, M.H. *et al.* (2016) Temporal activation of XRCC1-mediated DNA repair is essential for muscle differentiation. *Cell Discov.* 2, 15041
30. Connolly, P.F. and Fearnhead, H.O. (2016) DNA-PK activity is associated with caspase-dependent myogenic differentiation. *FEBS J.* 283, 3626–3636
31. Bell, R.A.V. *et al.* (2022) Chromatin reorganization during myoblast differentiation involves the caspase-dependent removal of SATB2. *Cells* 11, 966
32. Larsen, B.D. and Megeney, L.A. (2010) Parole terms for a killer: directing caspase3/CAD induced DNA strand breaks to coordinate changes in gene expression. *Cell Cycle* 9, 2940–2945
33. Jeffery, N.N. *et al.* (2021) Histone H2A.X phosphorylation and caspase-initiated chromatin condensation in late-stage erythropoiesis. *Epigenetics Chromatin* 14, 37
34. Zhao, B. *et al.* (2016) Erythropoietin-regulated oxidative stress negatively affects enucleation during terminal erythropoiesis. *Exp. Hematol.* 44, 975–981
35. Fernando, P. *et al.* (2005) Neural stem cell differentiation is dependent upon endogenous caspase 3 activity. *FASEB J.* 19, 1671–1673
36. Mukherjee, A. and Williams, D.W. (2017) More alive than dead: non-apoptotic roles for caspases in neuronal development, plasticity and disease. *Cell Death Differ.* 24, 1411–1421
37. Ballarino, R. *et al.* (2022) An atlas of endogenous DNA double-strand breaks arising during human neural cell fate determination. *Sci. Data* 9, 400
38. Michel, N. *et al.* (2022) Transcription-associated DNA DSBs activate p53 during hiPSC-based neurogenesis. *Sci. Rep.* 12, 12156
39. Wu, W. *et al.* (2021) Neuronal enhancers are hotspots for DNA single-strand break repair. *Nature* 593, 440–444
40. Ahuja, A.K. *et al.* (2016) A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. *Nat. Commun.* 7, 10660
41. Li, Y. *et al.* (2020) Patterns of somatic structural variation in human cancer genomes. *Nature* 578, 112–121
42. Akdemir, K.C. *et al.* (2020) Somatic mutation distributions in cancer genomes vary with three-dimensional chromatin structure. *Nat. Genet.* 52, 1178–1188
43. Ghavi-Helm, Y. *et al.* (2019) Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* 51, 1272–1282
44. Federico, C. *et al.* (2021) Chromosomal rearrangements and altered nuclear organization: Recent Mechanistic Models in Cancer. *Cancers (Basel)* 13, 5860
45. Pannunzio, N.R. and Lieber, M.R. (2018) Concept of DNA lesion longevity and chromosomal translocations. *Trends Biochem. Sci.* 43, 490–498
46. Pannunzio, N.R. and Lieber, M.R. (2017) AID and reactive oxygen species can induce DNA breaks within human chromosomal translocation fragile zones. *Mol. Cell* 68, 901–912.e903
47. Thomson, D.W. *et al.* (2020) Aberrant RAG-mediated recombination contributes to multiple structural rearrangements in lymphoid blast crisis of chronic myeloid leukemia. *Leukemia* 34, 2051–2063
48. Bianchi, J.J. *et al.* (2019) Breakage-fusion-bridge events trigger complex genome rearrangements and amplifications in developmentally arrested T cell lymphomas. *Cell Rep.* 27, 2847–2858.e2844
49. Papaemmanuil, E. *et al.* (2014) RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat. Genet.* 46, 116–125
50. Ichim, G. *et al.* (2015) Limited mitochondrial permeabilization causes DNA damage and genomic instability in the absence of cell death. *Mol. Cell* 57, 860–872
51. Liu, X. *et al.* (2015) Caspase-3 promotes genetic instability and carcinogenesis. *Mol. Cell* 58, 284–296
52. Berthenet, K. *et al.* (2020) Failed apoptosis enhances melanoma cancer cell aggressiveness. *Cell Rep.* 31, 107731
53. Cao, K. *et al.* (2022) Mitochondrial dynamics regulate genome stability via control of caspase-dependent DNA damage. *Dev. Cell* 57, 1211–1225.e1216
54. Arama, E. *et al.* (2003) Caspase activity and a specific cytochrome C are required for sperm differentiation in Drosophila. *Dev. Cell* 4, 687–697
55. Dehkordi, M.H. *et al.* (2020) Apoptosome-dependent myotube formation involves activation of caspase-3 in differentiating myoblasts. *Cell Death Dis.* 11, 308
56. Murray, T.V.A. *et al.* (2008) A non-apoptotic role for caspase-9 in muscle differentiation. *J. Cell Sci.* 121, 3786–3793
57. Bock, F.J. and Tait, S.W.G. (2020) Mitochondria as multifaceted regulators of cell death. *Nat. Rev. Mol. Cell Biol.* 21, 85–100
58. Gole, B. *et al.* (2015) Endonuclease G initiates DNA rearrangements at the MLL breakpoint cluster upon replication stress. *Oncogene* 34, 3391–3401
59. Cowell, I.G. *et al.* (2012) Model for MLL translocations in therapy-related leukemia involving topoisomerase IIbeta-mediated DNA strand breaks and gene proximity. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8989–8994
60. Cartwright, I.M. *et al.* (2017) Essential roles of Caspase-3 in facilitating Myc-induced genetic instability and carcinogenesis. *eLife* 6, e26371
61. Buzun, K. *et al.* (2020) DNA topoisomerases as molecular targets for anticancer drugs. *J. Enzyme Inhib. Med. Chem.* 35, 1781–1799
62. Martin-Encinas, E. *et al.* (2022) The design and discovery of topoisomerase I inhibitors as anticancer therapies. *Expert Opin. Drug Discovery* 17, 581–601
63. Pommier, Y. *et al.* (2022) Human topoisomerases and their roles in genome stability and organization. *Nat. Rev. Mol. Cell Biol.* 23, 407–427
64. Cieslar-Pobuda, A. *et al.* (2012) PARP-1 inhibition induces a late increase in the level of reactive oxygen species in cells after ionizing radiation. *Mutat. Res.* 732, 9–15
65. Dehingia, B. *et al.* (2022) CTCF shapes chromatin structure and gene expression in health and disease. *EMBO Rep.* 23, e55146
66. Ziv, Y. *et al.* (2006) Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nat. Cell Biol.* 8, 870–876
67. Jaiswal, H. *et al.* (2017) ATM/Wip1 activities at chromatin control Plk1 re-activation to determine G2 checkpoint duration. *EMBO J.* 36, 2161–2176
68. Engeland, K. (2022) Cell cycle regulation: p53-p21-RB signaling. *Cell Death Differ.* 29, 946–960
69. Hartlerode, A.J. *et al.* (2015) Recruitment and activation of the ATM kinase in the absence of DNA-damage sensors. *Nat. Struct. Mol. Biol.* 22, 736–743
70. Ha, H.J. and Park, H.H. (2022) Molecular basis of apoptotic DNA fragmentation by DFF40. *Cell Death Dis.* 13, 198
71. Morishita, Y. *et al.* (2021) Cell death-associated lipid droplet protein CIDE-A is a noncanonical marker of endoplasmic reticulum stress. *JCI Insight* 6, e143980
72. Yan, B. *et al.* (2009) Increased skin carcinogenesis in caspase-activated DNase knockout mice. *Carcinogenesis* 30, 1776–1780

RightsLink

DA



Destruction as a creative process: CAD-induced DNA strand breaks promote macrophage differentiation

Author: Dalal Alsowaida,Lynn A. Megeney
Publication: Cell Reports
Publisher: Elsevier
Date: 22 October 2024

© 2024 The Author(s). Published by Elsevier Inc.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK CLOSE WINDOW

© 2024 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Data Security and Privacy | For California Residents | Terms and ConditionsComments? We would like to hear from you. E-mail us at customercare@copyright.com

Privacy - Terms

Cell Reports

CellPress
OPEN ACCESS

Preview

Destruction as a creative process: CAD-induced DNA strand breaks promote macrophage differentiation

Dalal Alsowaida[1,2] and Lynn A. Megeney[1,*]
[1]Sprott Centre for Stem Cell Research, Ottawa Hospital Research Institute and the Departments of Medicine and Cellular and Molecular Medicine, University of Ottawa, Ottawa K1H 8L6, Canada
[2]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia
*Correspondence: lmegeney@ohri.ca
https://doi.org/10.1016/j.celrep.2024.114816

In a recent issue of *Cell Reports*, Maurya et al.[1] demonstrated that Drep4, the *Drosophila* homolog of caspase-activated DNase (CAD), drives DNA strand breaks during myeloid-like/macrophage cell differentiation and that this Drep4/CAD activity is essential for the differentiation process.

Over the last two decades, an accumulated wealth of data has demonstrated that caspase-dependent signaling pathways, initially discovered as the *de facto* mediators of apoptosis, are also conserved inductive cues that promote cell differentiation.[2–4] Caspase proteases have been the primary focus of study for this dual-use biology, but the enzyme that degrades DNA in latter stages of apoptosis, caspase-activated DNase (CAD), has received much less attention. CAD and Drep4 were initially discovered as the operative DNases that disassemble chromatin/DNA during apoptosis.[5,6] CAD is activated by caspase-3-targeted cleavage of its nascent inhibitor, ICAD (inhibitor of CAD), which releases the DNase to form active dimers and target DNA.

Despite the prominent link to apoptosis, caspase-3-stimulated CAD was reported to promote *in vitro* muscle cell differentiation through regulated DNA-strand-break activity.[7] Whether CAD influenced the differentiation program *in vivo* or in any other cell lineage or organism has remained uncorroborated. In a recent *Cell Reports* paper, Maurya et al.[1] provided evidence for the physiological relevance of CAD in cell-fate determination. The paper is a tour de force that showcases the strength of the *Drosophila* model in interrogating molecular controls that guide cell fate. In their study, Maurya et al. made use of multiple markers of DNA damage/repair as well as robust genetic reporter systems to document DNA-strand-break activity in differentiating macrophage progenitor cells. The team then used a series of rigorous models (genetic mutants and RNAi expression) to confirm the epistatic relationship of Dronc/caspase 9 activation to Drice/caspase 3, which in turn activates Drep4/CAD by cleaving its nascent inhibitor ICAD, and these all coordinate to drive macrophage differentiation. Less prominently featured in the paper, but no less important, is the observation that the DNases EndoG and DNAse II do not inflict differentiation-specific DNA breaks.

The Jnk pathway activation of caspase/CAD signaling is an equally compelling discovery documented by Maurya et al.[1] The data clearly show that an Akt/Ask1/Jnk signaling conduit is concurrent to and stimulatory for caspase/CAD activation, although the point at which this kinase signaling intersects with caspase activation was not explored. In an unrelated study, Dronc/caspase 9 activation was reported to directly impair Jnk signaling activity in *Drosophila* macrophage-like cells.[8] The work of Maurya et al. and Xu et al.[8] might appear contradictory in regard to Dronc/caspase 9 modulation of Jnk, but it is important to note that Maurya et al. examined cell differentiation, whereas Xu et al. investigated Dronc/caspase 9 control of Jnk in a tumor-like model. Elucidating the precise Akt/Ask1/Jnk point of control over Dronc/caspase 9 induction could explain the divergence in these respective cell niches.

As with all compelling science, the work of Maurya et al.[1] answered a number of outstanding questions pertaining to CAD biology while opening the door to new "unknowns." For example, how does CAD-mediated DNA strand breakage engage the genome to promote macrophage cell differentiation? Does CAD simply fracture the genome through DNA double-strand breaks (DSBs) akin to its role in apoptosis? Or does CAD engage a less disruptive form of genome stress, by promoting DNA single-strand breaks (SSBs) in target regions, which then alters gene expression to sway cell fate? Perhaps CAD manages a combination of DNA SSBs and DSBs, each to distinct loci, which then act in concert to propel cell differentiation. Exclusive high-frequency CAD-inflicted DSBs in a differentiating cell would appear less likely. Inflicting very large numbers of DSBs would be an effective mechanism to disrupt gene expression, yet such genome disruption would likely usher in apoptosis.

Maurya et al.[1] have provided compelling evidence that the CAD-mediated strand breaks in the differentiating macrophage progenitors are detected by a modified *in situ* nick-translation assay, which preferentially labels DNA SSBs. This observation is similar to the work of Larsen et al.,[7] where CAD-inflicted DNA breaks in muscle cell differentiation were also labeled by *in situ* nick translation. More recently, unique CAD-induced strand break activity was documented in cancer cells, where the formation of CAD breaks are permissive for the development of radiotherapy

This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

resistance.[9] Here, the genome targets of CAD were documented and were consistent with regions that bind CTCF, a protein modifier that influences gene expression through higher order chromatin structure. Whether similar mechanisms are operational in the CAD-directed cell differentiation as described by Maurya et al.[1] is unknown.[10]

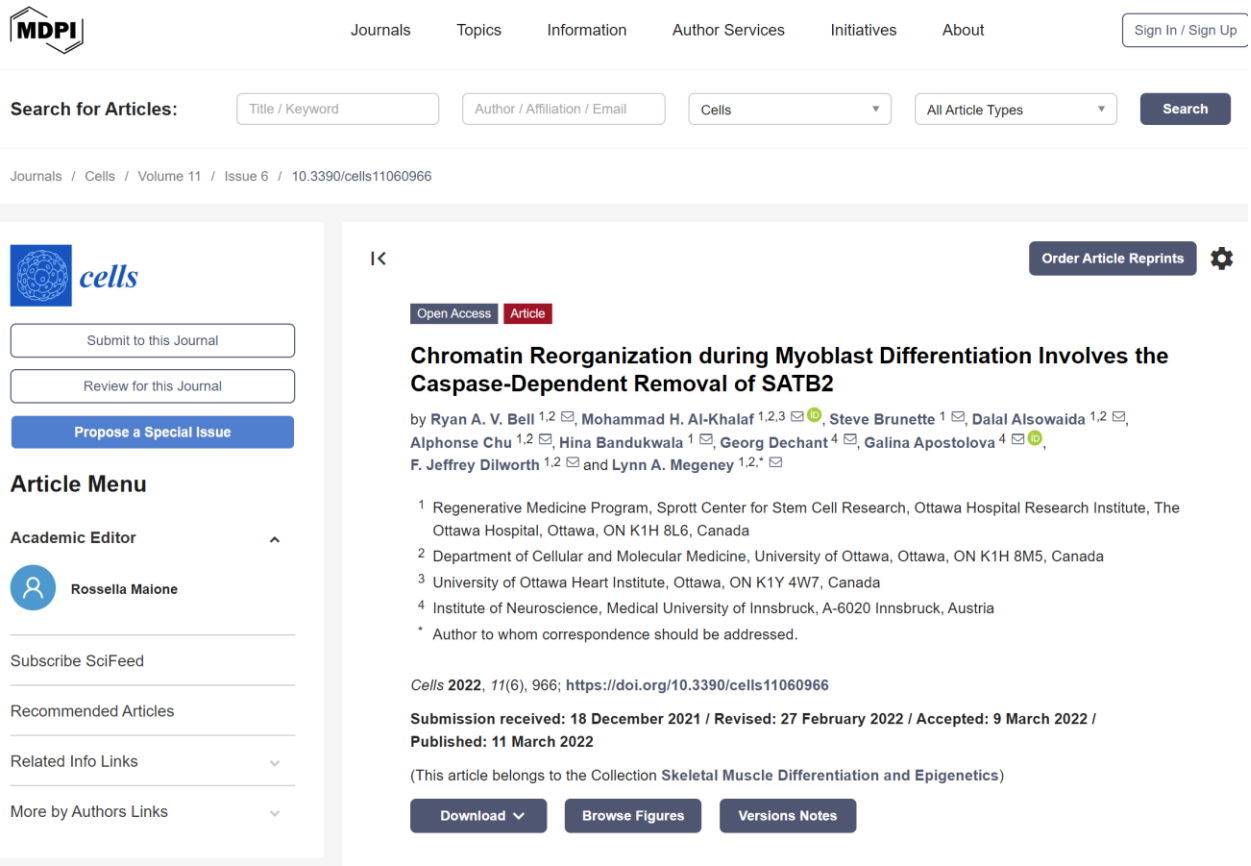
A priority for those in the field should be to construct genome-wide maps of the CAD binding sites, because this information would provide a more fulsome understanding of how this DNase drives cell fate/cell differentiation. In addition, cell and developmental biologists should pay close attention to this effort, as the observations of Maurya et al.[1] suggest caspase/CAD control of cell differentiation is conserved across metazoan organisms, in very divergent cell lineages. Confirming a universal CAD-directed DNA-strand-break mechanism as an inductive for cell differentiation would be a fundamental reordering of our understanding of how progenitor cells manage cell-fate transitions.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Maurya, D., Rai, G., Mandal, D., and Mondal, B.C. (2024). Transient caspase-mediated activation of caspase-activated DNase causes DNA damage required for phagocytic macrophage differentiation. Cell Rep. *43*, 114251.
2. Arama, E., Agapite, J., and Steller, H. (2003). Caspase activity and a specific cytochrome C are required for sperm differentiation in Drosophila. Dev. Cell *4*, 687–697.
3. Baena-Lopez, L.A., Arthurton, L., Xu, D.C., and Galasso, A. (2018). Non-apoptotic caspase regulation of stem cell properties. Semin. Cell Dev. Biol. *82*, 118–126.
4. Fernando, P., Kelly, J.F., Balazsi, K., Slack, R.S., and Megeney, L.A. (2002). Caspase 3 activity is required for skeletal muscle differentiation. Proc.Natl.Acad.Sci.USA. *99*, 11025–11030.
5. Enari, M., Sakahira, H., Yokoyama, H., Okawa, K., Iwamatsu, A., and Nagata, S. (1998). A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. Nature *391*, 43–50.
6. Inohara, N., Koseki, T., Chen, S., Wu, X., and Núñez, G. (1998). CIDE, a novel family of cell death activators with homology to the 45kDa subunit of the DNA fragmentation factor. EMBO J. *17*, 2526–2533.
7. Larsen, B.D., Rampalli, S., Burns, L.E., Brunette, S., Dilworth, F.J., and Megeney, L.A. (2010). Caspase 3/Caspase activated DNase promote cell differentiation by inducing DNA strand breaks. Proc.Nat.Acad.Sci.USA. *107*, 4230–4235.
8. Xu, D.C., Wang, L., Yamada, K.M., and Baena-Lopez, L.A. (2022). Non-apoptotic activation of Drosophila caspase-2/9 modulates JNK signaling, the tumor microenvironment, and growth of wound-like tumors. Cell Rep. *39*, 110718.
9. Larsen, B.D., Benada, J., Yung, P.Y.K., Bell, R.A.V., Pappas, G., Urban, V., Ahlskog, J.K., Kuo, T.T., Janscak, P., Megeney, L.A., et al. (2022). Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress. Science *376*, 476–483.
10. Benada, J., Alsowaida, D., Megeney, L.A., and Sørensen, C.S. (2023). Self-inflicted DNA breaks in cell differentiation and cancer. Trends Cell Biol. *33*, 850–859.



"From [Bell et al. Chromatin Reorganization during Myoblast Differentiation Involves the Caspase-Dependent Removal of SATB2. Cells. 2022]. Reprinted with permission from MDPI," to include full article, as a co-author.

MDPI Open Access Information and Policy

All articles published by MDPI are made immediately available worldwide under an open access license. This means:

- everyone has free and unlimited access to the full-text of *all* articles published in MDPI journals;
- everyone is free to re-use the published material if proper accreditation/citation of the original publication is given;
- open access publication is supported by the authors' institutes or research funding agencies by payment of a comparatively low **Article Processing Charge (APC)** for accepted articles.

Permissions

No special permission is required to reuse all or part of article published by MDPI, including figures and tables. For articles published under an open access Creative Common CC BY license, any part of the article may be reused without permission provided that the original article is clearly cited. Reuse of an article does not imply endorsement by the authors or MDPI. Furthermore, no special permission is required for authors to submit their work to external repositories. This policy extends to all versions of a paper: submitted, accepted, and published.

Copyrights

Copyright and Licensing

For all articles published in MDPI journals, copyright is retained by the authors. Articles are licensed under an open access Creative Commons CC BY 4.0 license, meaning that anyone may download and read the paper for free. In addition, the article may be reused and quoted provided that the original published version is cited. These conditions allow for maximum use and exposure of the work, while ensuring that the authors receive proper credit.

In exceptional circumstances articles may be licensed differently. If you have specific condition (such as one linked to funding) that does not allow this license, please mention this to the editorial office of the journal at submission. Exceptions will be granted at the discretion of the publisher.

cells

MDPI

Article

Chromatin Reorganization during Myoblast Differentiation Involves the Caspase-Dependent Removal of SATB2

Ryan A. V. Bell [1,2], Mohammad H. Al-Khalaf [1,2,3], Steve Brunette [1], Dalal Alsowaida [1,2], Alphonse Chu [1,2], Hina Bandukwala [1], Georg Dechant [4], Galina Apostolova [4], F. Jeffrey Dilworth [1,2] and Lynn A. Megeney [1,2,*]

1 Regenerative Medicine Program, Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute, The Ottawa Hospital, Ottawa, ON K1H 8L6, Canada; rybell@ohri.ca (R.A.V.B.); mal-khalaf@ottawaheart.ca (M.H.A.-K.); sbrunette@ohri.ca (S.B.); dalsowaida@ohri.ca (D.A.); alphonse.chu@yahoo.ca (A.C.); hbandukwala@ohri.ca (H.B.); jdilworth@ohri.ca (F.J.D.)
2 Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON K1H 8M5, Canada
3 University of Ottawa Heart Institute, Ottawa, ON K1Y 4W7, Canada
4 Institute of Neuroscience, Medical University of Innsbruck, A-6020 Innsbruck, Austria; georg.dechant@i-med.ac.at (G.D.); galina.apostolova@i-med.ac.at (G.A.)
* Correspondence: lmegeney@ohri.ca

Abstract: The induction of lineage-specific gene programs are strongly influenced by alterations in local chromatin architecture. However, key players that impact this genome reorganization remain largely unknown. Here, we report that the removal of the special AT-rich binding protein 2 (SATB2), a nuclear protein known to bind matrix attachment regions, is a key event in initiating myogenic differentiation. The deletion of myoblast SATB2 in vitro initiates chromatin remodeling and accelerates differentiation, which is dependent on the caspase 7-mediated cleavage of SATB2. A genome-wide analysis indicates that SATB2 binding within chromatin loops and near anchor points influences both loop and sub-TAD domain formation. Consequently, the chromatin changes that occur with the removal of SATB2 lead to the derepression of differentiation-inducing factors while also limiting the expression of genes that inhibit this cell fate change. Taken together, this study demonstrates that the temporal control of the SATB2 protein is critical in shaping the chromatin environment and coordinating the myogenic differentiation program.

Keywords: SATB2; chromatin remodeling; caspase 7; myogenic differentiation

check for updates

Citation: Bell, R.A.V.; Al-Khalaf, M.H.; Brunette, S.; Alsowaida, D.; Chu, A.; Bandukwala, H.; Dechant, G.; Apostolova, G.; Dilworth, F.J.; Megeney, L.A. Chromatin Reorganization during Myoblast Differentiation Involves the Caspase-Dependent Removal of SATB2. *Cells* **2022**, *11*, 966. https://doi.org/10.3390/cells11060966

Academic Editor: Rossella Maione

Received: 18 December 2021
Accepted: 9 March 2022
Published: 11 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

One of the key alterations that characterize a cell's progression through differentiation is the restructuring of the nuclear landscape to allow for the expression of lineage-specific genes. Indeed, chromatin conformation undergoes significant restructuring when cells progress from replicating to differentiating phenotypes, with changes occurring to repress or open access at specific gene loci [1,2]. These genome alterations are facilitated by two general mechanisms—loci-specific modifications of DNA and histones and the targeting of structural proteins that control the higher-order structure of chromatin [3,4]. Progress through the myogenic differentiation program is no exception to this paradigm, with extensive changes to genome structure during myogenesis that allow for muscle-specific genes to be expressed [5]. A number of proteins that govern DNA and histone methylation changes have been implicated in the control of myogenesis [6], yet limited information exists on the role of proteins that manage higher-order reorganization such as the formation of chromatin loops and the delineation of topologically associating domains (TADs).

Matrix attachment region (MAR) binding proteins represent a group of factors proposed to act as modifiers of the chromatin loop structure to influence gene expression during cell differentiation. MAR binding proteins are unique in that they integrate both repressive and inductive signals for gene expression, a feature that would be valuable for

the management of cell fate transitions. This duality of MAR protein function depends on the proximity of gene promoters and insulator regions relative to the position where the MAR protein anchors the DNA to the chromatin scaffold [7,8]. Scaffold attachment factor b1 (Safb1) has been shown to activate pro-differentiation gene expression in skeletal muscle cells, yet it does not appear to directly repress the expression of genes in muscles [9].

The special AT-rich binding proteins (SATB1 and SATB2) are a family of MAR binding proteins that may play an active role in global chromatin organization during cell differentiation. Although SATB proteins were originally defined as classic chromatin structure factors, a number of studies have suggested their direct influence on gene expression, akin to the biology of a transcription factor. For example, SATB2 has been found to control the expression of several genes involved in stem cell fate determination and cancer cell proliferation, acting in concert through physical association with other transcription factors [10–16]. However, one recent observation has shown that SATB2 can modify gene expression in cortical neurons through its association with the inner nuclear membrane protein, LEMD2 [17]. This latter study is more consistent with the structural role of SATB2, and the plethora of developmental abnormalities that accompany SATB2 mutation in humans suggest a genome-spanning function for this MAR protein [18]. Given these observations, the regulation of the SATB2 protein content may be a critical step in managing the duality of gene expression during cell maturation, i.e., the simultaneous repression and activation of targeted gene loci.

The mechanism by which SATB2 manages such a diverse biologic response is unknown. Likewise, there are no studies that have characterized the SATB2-dependent genome-wide architecture or the mode by which the SATB2 protein is targeted or displaced from its associated genomic targets. However, the caspase-mediated cleavage of SATB1 has been shown to be an important step in the control of gene expression that precedes T-cell apoptosis, where the cleavage of SATB1 results in its release from MARs [19]. Interestingly, caspase proteases have prominent roles in non-death processes such as differentiation, inflammation, remodeling, and cell survival [20]. The role of caspases in differentiation is particularly well-established, with both initiator and executioner caspases being involved in the development of a variety of tissues [21]. With respect to skeletal muscle differentiation, catalytically active caspase 3 and 9 appear to be required for progenitor progression through the myogenic program [22–24]. As such, the integral nature of caspase activity during myoblast differentiation and the possibility that caspases could act on chromatin organizing proteins have led us to investigate the behavior of SATB2 within muscle progenitor cells and the potential of caspase enzymes to mediate its function.

2. Materials and Methods

2.1. Mice and In Vivo Procedures

All mice were housed and treated at the University of Ottawa Animal Care and Veterinary Services. Mice used in our studies were housed and cared for according to Canadian Council on Animal Care (CCAC) guidelines and University of Ottawa Animal Care Committee protocols.

In order to determine the effects of SATB2 removal from muscle satellite cells, *Pax7/CreER Satb2*$^{fl/fl}$ and *Satb2*$^{fl/fl}$ (control) mice were fed a diet supplemented with tamoxifen (40 mg/kg body weight; Envigo) at 3 weeks old. These mice were monitored weekly for evidence of the potential effects of SATB2 ablation from muscle stem cells. After 8–10 months, three *Pax7/CreER Satb2*$^{fl/fl}$ and *Satb2*$^{fl/fl}$ mice (female) were euthanized by CO_2 asphyxiation and cervical dislocation. These mice were then assessed for any structural abnormalities, and their tibialis anterior (TA) muscle was excised and fixed in 10% neutral buffered formalin. After 24 h, the samples were then placed in 70% ethanol and then paraffin embedded and cut into 4 µm sections.

2.2. Cell Culture

Cells of the immortalized mouse cell line, C2C12, were grown on non-collagen coated cell culture plates in Dulbecco's Modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (growth medium). Once confluent, the cells were differentiated in DMEM supplemented with 2% horse serum and 1% penicillin/streptomycin (differentiation medium).

Primary myoblasts were isolated, as described by Fernando et al. [22]. Following isolation, myoblasts were cultured in 1:1 (v/v) of Ham's F10:DMEM, supplemented with 20% FBS, 1% penicillin/streptomycin, and 10 ng/mL basic fibroblast growth factor. Myoblasts were differentiated in DMEM supplemented with 5% horse serum and 1% penicillin/streptomycin.

2.3. Protein Extraction and Western Blotting

C2C12 cells were lysed using a modified RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 1% NP-40, 1% glycerol, and a cocktail of protease inhibitors) for whole-cell extracts, or with a modified NE-PER Nuclear and Cytoplasmic Extraction Kit, as per the manufacturer's instructions (Thermo Scientific, Waltham, MA, USA). Extracted proteins were then separated via SDS-PAGE and then transferred to a 0.45 µM polyvinylidene fluoride (PVDF) membrane (Millipore, Burlington, MA, USA) on a TRANS-BLOT SD apparatus (Bio-Rad, Hercules, CA, USA). Membranes were blocked with Tris-buffered saline plus 0.1% Tween-20 (TBST) containing 5% skim milk for 1 h at room temperature. Membranes were then incubated overnight at 4 °C in primary antibody made with the blocking solution. The primary antibodies used in this study included mouse SATB2 (ab51502; Abcam, Cambridge, UK; 1:1000), rabbit cleaved caspase 7 (9491S; Cell Signaling, Danvers, MA, USA; 1:1000), mouse myosin heavy chain (Development Studies Hybridoma Bank, DSHB Iowa City, IA, USA; MHC; 1:250), rabbit lamin A/C (2032; Cell Signaling; 1:1000), mouse tubulin (DSHB; 1:200), and mouse glyceraldehyde 3-phosphate dehydrogenase (#2118; GAPDH; 1:4000). Western blots were performed on three independent cell preparations in all cases (n = 3).

2.4. Immunofluorescence

C2C12 cells or primary myoblasts were cultured and fixed in paraformaldehyde on 25 mm coverslips at the desired time points. The cells were then incubated for 10 min at room temperature with a permeabilization solution containing 0.5% Triton-X 100 in PBS. Subsequently, after washing in PBS, cells were incubated for 1 h in a blocking solution consisting of 5% horse serum in PBS. After blocking, the cells were incubated for 2 h at room temperature or overnight at 4 °C in primary antibody that was diluted in blocking solution. The primary antibodies used were mouse SATB2 (1:50; ab51502, Abcam), rabbit heterochromatin protein 1 alpha (HP1α; 1:200; #2616, Cell Signaling), rabbit GAPDH (1:400; #2118, Cell Signaling), rabbit desmin (1:400; ab15200, Abcam), and mouse MHC (1:50; DSHB). After primary antibody incubation, the cells were washed three times in PBS and incubated in secondary antibody (2 mg/mL Alexa Fluor 488, Invitrogen, Waltham, MA, USA; 1:1500, 2 mg/mL Alexa Fluor 594, Invitrogen, 1:1500; 2 mg/mL Alexa Fluor 568, Invitrogen, 1:1000) diluted in PBS for 1.5 h at room temperature. After incubation, the cells were washed 2× in PBS and then counterstained with 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI; 1:10,000; Sigma) for 10 min at room temperature. After incubation, the cells were washed 2× in PBS, and the coverslips were mounted on microscope slides using Dako Fluorescent Mounting Medium. The cells were then visualized using a Zeiss Observer Z1 inverted fluorescence microscope, Zeiss Canada, Toronto, ON, Canada. All images were captured using the AxioVision 4.8 software. Quantification of immunofluorescence was performed by densitometry analysis using the ImageJ/Photoshop C3 software (Public Domain). Immunofluorescence experiments were performed on n = 3–5 independent cell preparations depending on the experiment (detailed in the figure captions).

2.5. Immunohistochemistry

Mouse TA muscle was fixed in 10% formalin for 24–48 h before being embedded in paraffin and sectioned into 4 µm thick sections. The sections were then stained for Pax7 (DSHB hybridoma antibody; 1:3.5), SATB2 (Abcam; 1:300), or with hematoxylin and eosin (H&E). The control for IHC staining was staining that used just secondary antibody that originated from the same species as the primary antibody used. The sectioning and staining were conducted by the Pathology and Laboratory Medicine (PALM) Histology Core Facility at the University of Ottawa, Canada. To determine muscle fiber areas, H&E stained sections were assessed under an Observer A1 microscope (Zeiss), and areas were calculated using the ImageJ software.

2.6. Chromatin Immunoprecipitation (ChIP) Assay

C2C12 cells were grown on 15 cm plates and allowed to reach 100% confluence before three aliquots of twenty million proliferating cells were collected (technical replicate $n = 3$; biological replicate $n = 1$). Cells were fixed for 10 min using 1% formaldehyde in DMEM. Cell fixation was quenched by removing the fixation solution, rinsing the plates with PBS, and then pouring a solution of 0.125 M glycine in PBS onto the cells and incubating them for 5 min. Subsequently, the plates were washed 2× with PBS, and then cells were scraped from the plates, pelleted, and stored at −80 °C until used.

Given that SATB2 is a nuclear matrix attachment protein, it required an alternative protocol for chromatin shearing as compared to standard preparations. First, a commercial hypotonic solution (Active Motif, Carlsbad, CA, USA) was used to lyse cellular membranes but retain intact the nuclear membranes. We subsequently performed additional lysis using a handheld glass Dounce homogenizer. After centrifugation, the supernatant was discarded, and the pelleted nuclei were treated with Active Motif's Pro-enzymatic Digestion Nuclear Extraction Solution. After incubation at 37 °C for 5 min, we supplemented the reaction with 1 U micrococcal nuclease (New England Biolabs Canada, Toronto, ON, Canada) and allowed it to incubate for 10 min. We then supplemented the reaction with SDS to a final volume of 2% SDS-chromatin solution. We proceeded to the second stage of shearing using sonication via a Covaris M220 Focused-Ultrasonicator instrument, setting the parameters to produce sheared DNA of 200 bp. After sonication, we performed a final centrifugation at 16,000× *g* at 4 °C for 20 min and collected the supernatant containing the sheared chromatin.

We performed immunoprecipitations with 5 µg 1°Ab (Target: SATB2, Abcam; positive control: RNA pol II, Active Motif; negative control: mouse IgG, Santa Cruz, Dallas, TX, USA) using magnetic beads (Active Motif) diluted in a commercial ChIP-buffer (Active Motif) overnight at 4 °C. Captured chromatin was isolated using magnetic stands to pull down the beads. Bound DNA was subsequently eluted via commercial elution, reverse cross-linking, protein digestion, and RNA digestion solutions (Active Motif). The resulting samples of DNA were further purified using phenol:chloroform extraction procedures. Final DNA was quantified using a NanoDrop spectrophotometer, with a 25 µL final volume at concentrations of 75–100 ng/µL.

2.7. ChIP-Sequencing and Bioinformatics

For the genome-wide analysis, immunoprecipitated DNA was amplified and a 75 bp single-read sequencing was performed on an Illumina HiSeq 2500 at the Next-Generation Sequencing Facility at The Centre for Applied Genomics in The Hospital for Sick Children (Toronto, ON, Canada). The data were summarized, and basic comparisons were performed using the Excel spreadsheet program (Microsoft, Redmond, WASH, USA). Reads were aligned to the GRCm38 (UCSC mm10, Dec/2011) mouse genome from the UCSC genome browser with bowtie2 (v2.2.6) using the default parameters, and peaks were called using MACS version 2.1.0, again with default parameters and using the IgG ChIP-seq data as a control (Supplementary Table S4).

The peaks identified by MACS were analyzed in R using the ChIPseeker package [25] to identify sites of SATB2 enrichment relative to genes (promoter, exon, intron, UTR, etc.).

Genes with one or more SATB2 binding sites in a region from −5 kb to + 1 kb around the transcription start site (TSS) were identified as being associated with Satb2 binding. This gene list was analyzed with ClusterProfiler [26] to identify Gene Ontology (GO) terms for which genes bound by SATB2 are statistically enriched above the expected level using the hypergeometric test. HOMER v4.11.1 [27] was used to search for enriched DNA sequence motifs within the full set of Satb2 peaks, using the default parameters for the findMotifsGenome.pl tool.

2.8. RNA-Sequencing and Bioinformatics

Total RNA was isolated with an RNeasy Kit (QIAGEN, Hilden, Germany) using an on-column DNase digestion (RNase-Free DNase Set, QIAGEN) to avoid genomic DNA contamination. Library preparation and 126-bp paired-end RNA-seq was performed by the Next-Generation Sequencing Facility at The Centre for Applied Genomics in The Hospital for Sick Children. RNA integrity was assessed using the Bioanalyzer platform (Agilent Technologies, Inc. Santa Clara, CA, USA). Sequencing was performed using standard procedures for the Illumina HiSeq 2500 platform.

RNA-seq reads were mapped to the GRCm38 (mm10) mouse genome assembly using tophat2 v2.1.1 [28], and the mapped reads were assigned to transcripts from the GENCODE vM23 annotation using FeatureCounts v1.5.2 [29]. Read count data for 24 h differentiated siControl and siSATB2 C2C12 cells, with three replicates each, were loaded into R (v4.0.2), and the differential expression was assessed using DESeq2 v1.30.1 [30]; expression differences between the siSatb2 and siControl replicates were calculated using the DESeq2 lfcShrink function and applying the apeglm method (v 1.12.0. [31]). Multiple testing correction was performed using the Benjamini Hochberg method, and lists of significantly differentially expressed (DE) genes were identified using a q-value (i.e., a corrected p value) cut-off of 0.05.

The matrix of gene expression values was transformed with the DESeq2 rlog function, and the DESeq2 plotPCA function was used to run a principal component analysis of the 500 most variable genes and to plot the results of the first two principal components. Sets of genes that were significantly higher or lower in abundance in siSatb2 cells (2013 and 1914 genes, respectively) were analyzed using the clusterProfiler package to identify the GO terms for which genes in either list were significantly over-represented using a hypergeometric test. Gene lists were also uploaded to the gProfiler tool [32] to identify enriched pathways in GO datasets.

GSEA was performed using the R package fgsea v1.14.0. In brief, a ranking statistic was first derived for all differentially expressed genes identified by DESeq2. This was performed by multiplying the sign of the Log2FoldChange with the respective log10 transformed p value for each gene. After arranging the genes using the signed p value statistic, GSEA was performed with the ranked gene list and the REACTOME Canonical Pathway v7.5 from the Molecular Signature Database (MSigDb). Parameters of note that were used with the fgsea function include eps = 0.0, minSize = 15, and maxSize = 500 [33].

2.9. Chromosome Conformation Capture, Hi-C

siControl and siSatb2-treated C2C12 cells were prepared for Hi-C experiments, as described below in the section illustrating the siRNA-mediated knockdown of SATB2. Only proliferating C2C12 cells were assessed via Hi-C. Once the cell treatments were complete, C2C12s were counted using trypan blue and a hemocytometer so that each sample would have 1×10^7 cells/sample. Two independent samples were collected for each treatment. Each aliquot of cells was pelleted at 500× g for 5 min and then resuspended in 10 mL PBS. Cells were then fixed in 8% formaldehyde for 10 min, after which the cross-linking reaction was quenched in 0.65 M glycine for 5 min. This mixture was allowed to sit in wet ice for 15 min before the cells were pelleted at 500× g for 5 min. C2C12s were subsequently resuspended in 5 mL PBS and then pelleted once again at 500× g for 5 min. The supernatant was then aspirated, and the pellet was frozen on dry ice for

10 min before being stored at –80 °C. The samples were then shipped on dry ice to Arima Genomics (http://arimagenomics.com (accessed on 30 April 2021) San Diego, CA, USA), which performed the Hi-C analysis according to the Arima-HiC + protocols described in the Arima-HiC kit (P/N: A510008). After the Arima-HiC protocol, Illumina-compatible sequencing libraries were prepared by first shearing the purified Arima-HiC proximally ligated DNA and then size-selecting the DNA fragments of ~200–600 bp using SPRI beads. The size-selected fragments were then enriched for biotin and converted into Illumina-compatible sequencing libraries using the KAPA Hyper Prep kit (P/N: KK8504). After adapter ligation, the DNA was PCR amplified and purified using SPRI beads. The purified DNA underwent standard quality control (qPCR and Bioanalyzer) and was sequenced on the HiSeq X, following the manufacturer's protocols.

2.10. Hi-C Data Processing

Hi-C data were analyzed with Juicer v1.6 [34] using a custom restriction site file for the GRCm38 mouse genome assembly provided by Arima Genomics (GATC and GANTC sites). Data from the replicates (n = 2) were combined for the siControl and siSatb2-treated samples and subsequently used for any downstream analysis. Chromatin loops were identified using HiCCUPS from Juicer v1.6 with the default settings. Loops were delineated at resolutions of 25,000, 10,000, and 5000 bp and thereafter merged to generate a final loop set. HiCexplorer v3.6 was used to identify topologically associated domains (TADs; 100 kb, 500 kb, and 1 Mb resolution) and sub-TAD contact domains (10 kb resolution). With respect to TADs, only data from the 100 kb resolution was depicted. The function *hicFindTADs* was used with the Knight–Ruiz (KR) normalized Hi-C matrices with a p value cut-off of 0.05 and FDR 0.01. ChIPseeker v1.24.0 was used for the functional annotation of the domains.

SATB2 binding sites and chromatin loops/TADs were compared to repressive (H3K27me3) and permissive (H3K4me3 and H3K36me3) histone marks that had been generated by Asp et al. [35] from proliferating C2C12 myoblasts. The accession numbers for the associated ChIP-seq data are: H3K4me3 (GEO accession: GSM721292), H3K27me3 (GEO accession: GSM721294), and H3K36me3 (GEO accession: GSM721296). Single-end reads were downloaded from the GEO database and mapped to mm10 using Bowtie2 v2.4.2. After removing duplicates with Picard Tools (openjdk v1.8.0_292) and MAPQ-based filtering (-q 10) with Samtools v1.12, broadPeaks were called using Macs3 v3.0.0a6 (–broad-cutoff 0.05 -*p* 0.05 –keep-dup 1).

Similarly, SATB2 ChIP-seq peaks and chromatin structural features determined here by Hi-C were compared to known CTCF binding sites. Processed data for CTCF ChIP-seq data in mouse C2C12 cells were obtained directly from ENCODE (ENCSR000AIJ). The processed data included: (1) reads aligned to mm10 in BAM format (ENCSR000AIJ), and (2) IDR thresholded peaks in bed narrowPeak format (ENCFF784ASD).

2.11. Caspase Cleavage Assays

Recombinant SATB2 protein (250–500 ng; Abnova) and recombinant active caspase 3 (0.5 µg; Chemicon, Thermo Fisher Scientific, Waltham, MA, USA) or recombinant active caspase 7 (0.5 µg; BioVision, Milpitas, CA, USA) were incubated for 3 h in a cleavage assay buffer (50 mM Hepes, pH 7.5, 0.1 M NaCl, 10% (v/v) glycerol, 0.1% Chaps, 10 mM dithiothreitol) containing either dimethyl sulphoxide (DMSO) or z.DEVD.fmk (20 µM; BioVision; [36,37]), as indicated. Reactions were stopped by the addition of the Laemmli sample buffer and subjected to SDS-PAGE. Mass spectrometry was performed at the Ottawa Hospital Research Institute Proteomics Core Facility (Ottawa, Canada). MASCOT 2.3.01 software (Matrix Science, Boston, MA, USA) was used to infer peptide and protein identities from the mass spectra.

2.12. Caspase Inhibition Assays

For caspase 3/7 inhibition, cultured C2C12s were pre-treated with either 15 µM z-DEVD-fmk (DEVD) from BioVision or 15 µM DMSO from Sigma for 2 h at 37 °C. After

pre-treatment, the cells were induced to differentiate using low serum media or continued in growth media both containing 15 μM DEVD or DMSO as a vehicle-only control. The inhibition or control media were changed every 48 h until the end of the time course. Cells were collected at the predetermined time points and analyzed as described. Caspase inhibition assays were performed on $n = 3$ independent cell preparations.

2.13. siRNA-Mediated Depletion of SATB2 and Caspase 7 Gene Expression

siRNA duplexes were used to suppress *Satb2* and caspase 7 gene expression in C2C12 cells. C2C12 cells were transfected at 25% confluence with 10 nM siRNA (siSatb2, siCasp7, or siControl) and the Lipofectamine RNAiMAX reagent, as directed by the manufacturer's protocol (Invitrogen). After an overnight incubation, fresh media were added onto the cells, and the cells were re-transfected. This continued until the cells reached 100% confluence, after which differentiation media were added to the cells. Media change and re-transfection occurred every 24 h throughout the indicated time course.

2.14. Statistical Analysis

Statistical analysis of three or more data sets was performed using one-way analysis of variance (ANOVA), with the post hoc test being indicated in the figure caption. For comparison between two sample sets, an unpaired, two-tailed Student's *t*-test was performed, unless otherwise stated, and a $p < 0.05$ was considered statistically significant. Data represented by bar graphs are all mean ± standard error of the mean (SEM).

3. Results

3.1. SATB2 Restrains the Induction of Muscle Cell Differentiation

Western blot and immunofluorescent analyses identified SATB2 as a nuclear protein within proliferating C2C12 muscle cells, which was markedly reduced in expression during the early stages of differentiation (Figures 1A,B and S1A). This signal was not confounded by potential cross reactivity with the closely related protein, SATB1, as this protein was not detected in proliferating or differentiated myoblasts (Supplementary Figure S1F). When expressed, SATB2 was mainly relegated to the euchromatic nuclear space; however, there was a significant subfraction of SATB2 co-staining with HP1α occurring in the more condensed regions of the nucleus, which declined as the differentiation program proceeded (Figure 1C). Having established that the SATB2 protein is reduced during myogenesis, we next sought to determine the role of SATB2 in muscle cell proliferation and differentiation. To this end, we initially designed CRISPR-Cas9 guide RNAs to inhibit SATB2 gene expression in replicating myoblasts. However, infection of C2C12 muscle cells with an adenovirus expressing enhanced Cas9 [38] led to a notable increase in SATB2 protein content in post-differentiated myoblasts, at a time when SATB2 would otherwise be in significant decline (see Supplementary Figure S1D). Given this unexpected impact of Cas9 on our protein of interest, we chose to pursue the siRNA-mediated targeting of SATB2 as a means of addressing its biologic role in cell culture models. The siRNA targeting of SATB2 (siSATB2) did not produce any noticeable alteration in cell viability or growth (Supplementary Figure S1B,C), although it provided an effective decrease in SATB2 expression (Figure 1D). However, compared to wildtype cells, siSATB2-treated cells displayed altered differentiation kinetics, with the accelerated formation of multinucleated myotubes (Figure 1E) and the increased expression of the contractile protein, myosin heavy chain (MHC) (Supplementary Figure S1E).

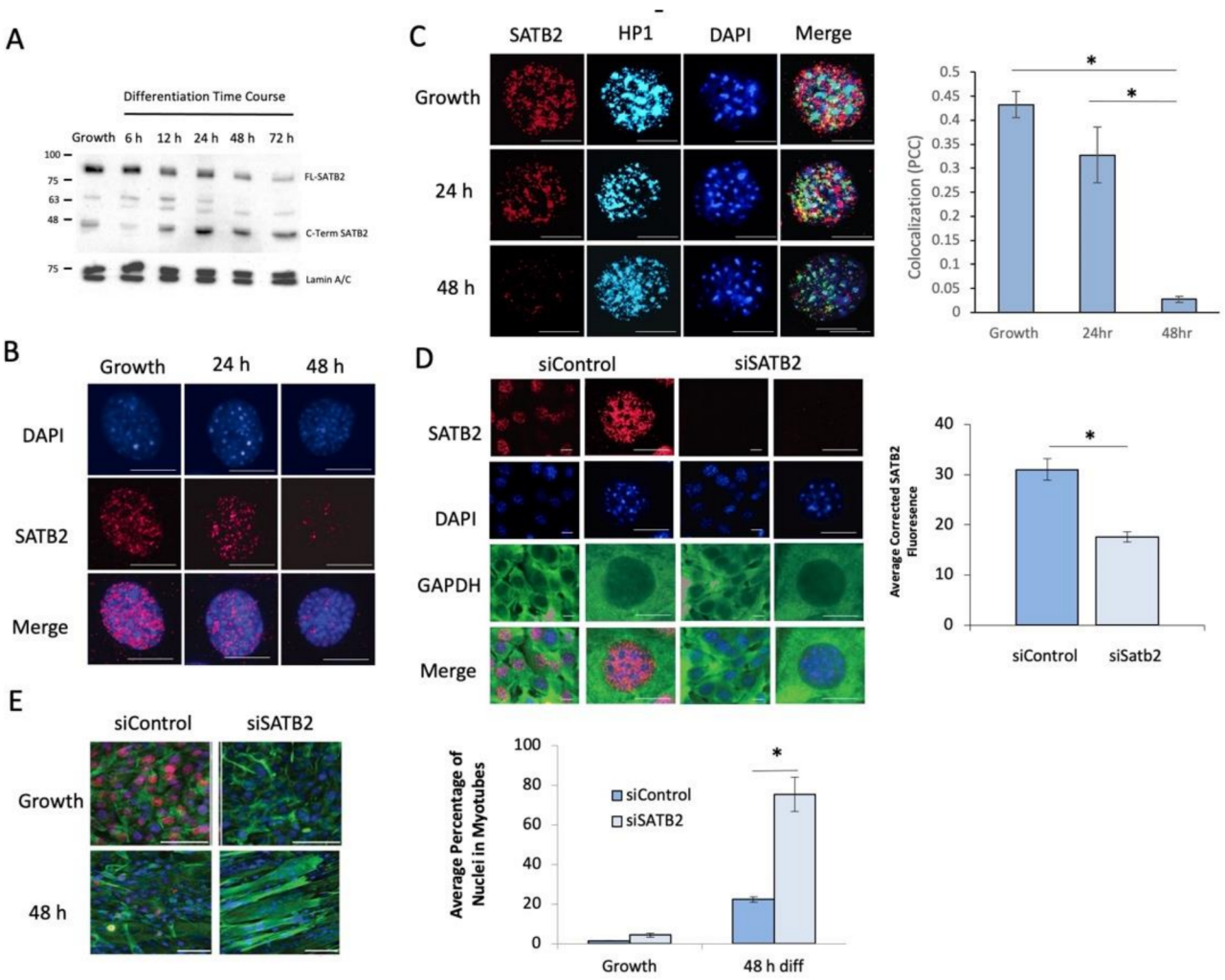


Figure 1. Loss of SATB2 expression accelerates myoblast differentiation. (**A**) Representative western blot showing the decreased expression of full-length SATB2 (FL-SATB2) during a differentiation time course for C2C12 cells (n = 3). This was accompanied by a concomitant increase in a C-terminal fragment of SATB2 (C-Term SATB2; identified in Supplementary Table S1). (**B**) Representative immunofluorescent staining of SATB2 showing a decrease in SATB2 expression during early C2C12 cell differentiation. Images are representative of n = 5 independent samples at each time point. DAPI (blue), SATB2 (red); scale = 10 μm. (**C**) Representative immunofluorescent images depicting SATB2 localization in relation to heterochromatin protein 1α (HP1α). Images are representative of n = 3 independent determinations for each time point. Scale bar = 10 μm. DAPI (blue), SATB2 (red), HP1α (teal). Bar graph depicts colocalization of SATB2 and HP1 within the nucleus of proliferating and differentiating myoblasts. Data are the quantification of immunofluorescent images to the left of the graph. * indicates a significant change in colocalization, determined using a one-way ANOVA with a post hoc Tukey test ($p < 0.01$). Data are means ± SEM for n = 3 independent determinations. (**D**) Immunofluorescent images depicting the efficacy of SATB2 depletion using siRNA in proliferating C2C12 cells. Muscle cell cytoplasm was counterstained with an anti-GAPDH antibody. Images are representative of n = 3 independent determinations. DAPI (blue), GAPDH (green), and SATB2 (red). Scale bar for the low magnification column = 10 μm. Scale bar for the higher magnification column = 5 μm. Bar graph depicts the quantification of SATB2 expression from the immunofluorescent images. * indicates a significant difference in SATB2 expression between siControl and siSATB2 conditions using the Student's t-test ($p < 0.05$). Data are means ± SEM for n = 3 independent determinations. (**E**) Immunofluorescent images indicating that the depletion of SATB2 led to early myotube formation after C2C12 cells were induced to differentiate. Images are representative of n = 3 independent determinations. DAPI (blue), SATB2 (red), and desmin (green). Scale bar = 50 μm. Bar graph depicts the

fusion index for proliferating (Growth) and 48 h differentiated myoblasts. Data were a consequence of assessing immunofluorescent images depicted in Figure 1E. * indicates a significant change in the fusion index following siSATB2 treatment as compared to siControl-treated cells, determined by the Student's *t*-test ($p < 0.05$). Data are means ± SEM for $n = 3$ independent determinations.

3.2. Caspase 7 Cleaves SATB2 during Early Myogenesis

Evidence from apoptotic nuclei suggests that nuclear structural proteins are targeted by caspase proteases [19,39,40]. Moreover, our identification of a major C-terminal fragment of SATB2 that becomes more prominent during muscle cell differentiation suggests that a proteolytic event may target the SATB2 protein (Figures 1A and S2D). Prior studies from our laboratory have shown that caspase 3 plays a prominent role in myogenesis, targeting and cleaving a number of substrate proteins to engage the differentiation program [22,24,36]. To examine whether SATB2 proteolysis occurred, we suppressed endogenous effector caspase activity (caspase 3 and 7) with the peptide inhibitor z.DEVD.fmk and reassessed SATB2 protein expression during muscle cell differentiation. Caspase inhibition during differentiation led to the sustained expression of the SATB2 protein in the nucleus compared to control cells (Figure 2A,B). In an attempt to attribute SATB2 cleavage to either caspase 3 or 7 (or both), an in vitro cleavage assay was performed. This assay demonstrated that caspase 7 activity was very robust at targeting SATB2 protein, whereas caspase 3 did not induce measurable cleavage (Figure 2C). A mass spectrometry analysis of the caspase 7-mediated SATB2 fragments mapped a putative caspase cleavage site at D477 (Supplementary Figure S2C and Table S1).

To corroborate caspase 7 as a targeting protease of SATB2 and by inference as a regulatory enzyme that may promote muscle cell differentiation, we utilized siRNA to suppress caspase 7 expression. siRNA repression of caspase 7 (siCasp7) led to a concomitant accumulation of the SATB2 protein during C2C12 muscle cell differentiation when compared to siControl-treated cells (Figures 2D,E and S2A). Investigating caspase 7 more closely, we determined that the expression of nuclear, active caspase 7 was inversely correlated with SATB2 expression, with nuclear caspase 7 levels increasing during early myoblast differentiation (Figures 2F and S2E). To assess the broad effect of this protease on muscle differentiation, we compared siCasp7 cells (where siRNA treatment reduced caspase 7 expression by ~60%; Figure 2D) vs. siControl-treated cells and noted that siCasp7 cultures displayed a significant impairment in the low serum induction of differentiation, with a dramatic reduction in the expression of MHC and a near complete inhibition of multinucleate myotube formation (Figure 2G). Interestingly, the simultaneous suppression of caspase 7 and SATB2 via RNA interference caused a modest increase in MHC expression from levels observed when only caspase 7 was repressed (Supplementary Figure S2B).

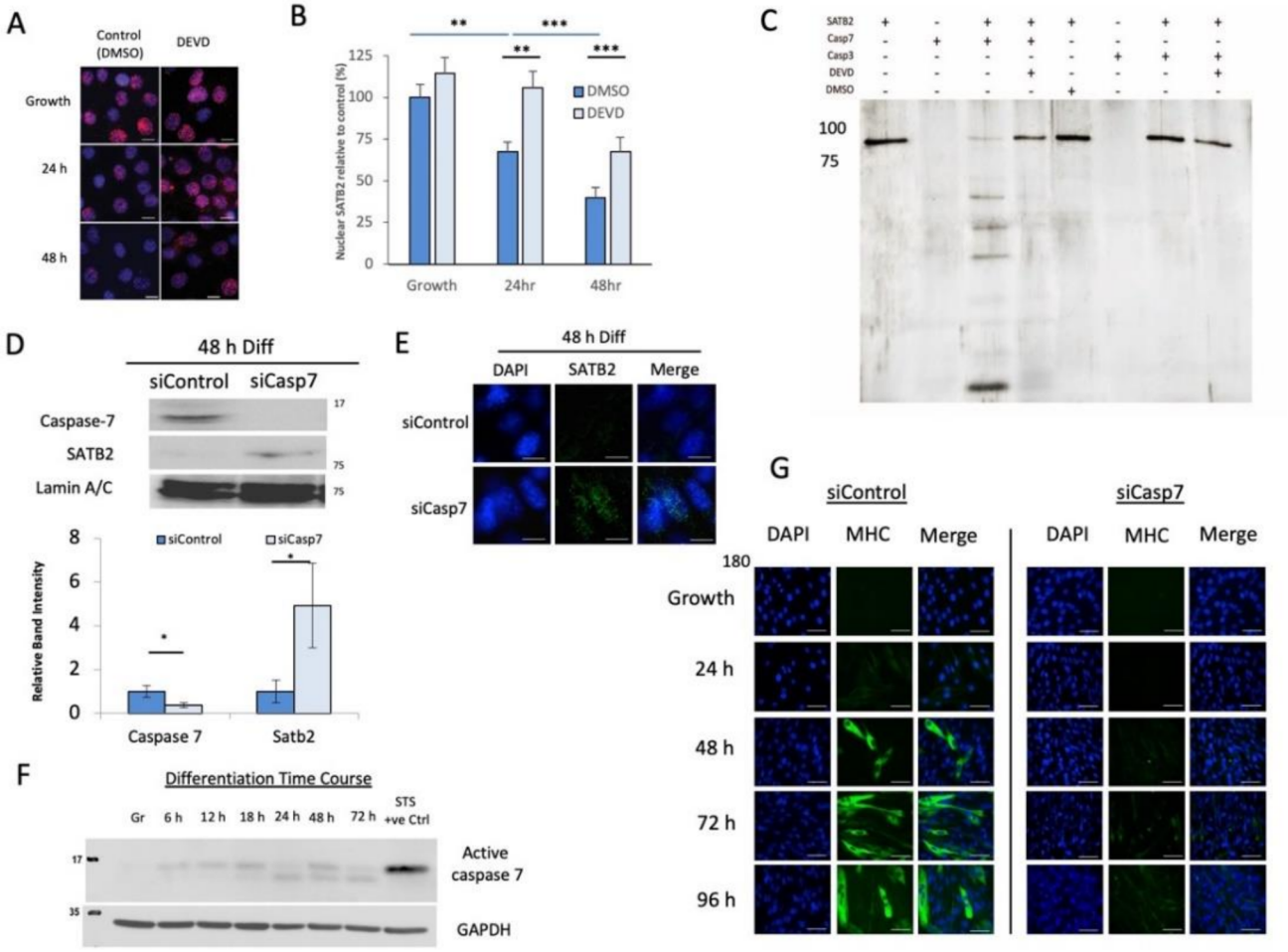


Figure 2. Caspase 7 cleaves SATB2 during early skeletal muscle differentiation and is necessary for myogenesis. (**A**) Immunofluorescent staining showing the maintenance of SATB2 expression following caspase inhibition by z-DEVD.fmk (DEVD) in differentiating C2C12 cells. Images are representative of $n = 3$ determinations on independent samples. DAPI (blue); SATB2 (red); scale: 10 μm. (**B**) Quantification of SATB2 nuclear expression following DEVD-mediated suppression of caspase activity. Data are the means ± SEM from $n = 3$ independent samples; 75–100 cells per condition/time point were analyzed. ** indicates $p < 0.01$ and *** indicates $p < 0.001$, as determined by the Student's *t*-test. (**C**) Silver stained gel of an in vitro caspase cleavage assay showing SATB2 cleavage by caspase 7. Reactions include combinations of recombinant SATB2 protein, active caspase 3/7 recombinant proteins, caspase chemical inhibitor (DEVD), or the control chemical DMSO ($n = 3$). (**D**) Caspase 7 and SATB2 expression following siCasp7 and siControl treatments of C2C12 cells. Both the histogram and the representative western blots ($n = 3$) were the expression levels of those two proteins 48 h post-induction of differentiation. * indicates a significant difference in caspase 7 or SATB2 expression, as determined by the Student's *t*-test ($p < 0.05$). (**E**) Immunofluorescent images showing the sustained expression of SATB2 after the depletion of caspase 7 in C2C12 cells differentiated for 48 h. Images are representative of $n = 3$ independent determinations. Scale bars = 10 μm. (**F**) Representative western blot depicting the cleaved active fragment of caspase 7 found in the nuclear fraction of C2C12 cells during their proliferative phase (Gr) and at various stages of differentiation (6–72 h post-induction; $n = 3$). Staurosporine (STS) was administered to proliferating C2C12 cells to induce apoptosis as a positive control. (**G**) Immunofluorescent images showing the reduction in myosin heavy chain (MHC) expression in differentiating C2C12 myoblasts. Panel is representative of $n = 3$ determinations on independent samples. DAPI (blue), MHC (green). Scale bars = 50 μm.

3.3. ChIP- and RNA-seq Data Support the Role of SATB2 in Regulating Muscle Satellite Cell Differentiation and Chromatin Reorganization

To clarify the role of SATB2 in the regulation of genes important to myogenesis and the associated chromatin remodeling, we performed both ChIP- and RNA-seq analyses. Analysis of the ChIP-seq data indicated that SATB2 was bound throughout the proliferating muscle stem cell genome (Figure 3A) and was mainly positioned in the intergenic spaces (Figure 3C). That said, a significant number of SATB2 peaks were found within or adjacent to genes, including binding sites within promoters, introns, exons, and 3′ untranslated regions (Figure 3B for the binding site distance to the transcription start site (TSS) of the nearest gene, and Figure 3C for the binding site genomic context). An investigation into SATB2 binding motifs within these satellite cells revealed no conserved motifs on a genome-wide scale (Supplementary Table S2).

Given the canonical role of SATB proteins in regulating the chromatin structure and the proximity of SATB2 to the coding regions of many genes, a more thorough investigation into its impact on chromatin remodeling-associated gene expression changes was warranted. To this end, RNA-seq was performed on 24 h differentiated C_2C_{12} cells that had been treated with either scrambled or SATB2 siRNA. Prior to our in-depth analyses of gene expression changes following SATB2 loss, we confirmed via RNA-seq that our siSatb2 treatment successfully downregulated *Satb2* mRNA (Supplementary Figure S3A). Using DESeq2 with an adjusted p value of 0.05 as a cut-off, our RNA-seq analyses revealed that the siRNA-mediated knockdown of SATB2 led to the reduced expression of 1510 genes and the increased expression of 1595 genes in 24 h differentiated myoblasts. Among the genes that were differentially expressed when SATB2 expression was altered were several prominent chromatin remodeling factors. These genes included *Setbp1*, *Hdac4/5/9/11*, and *Gata3* (Figure 3D).

GO categorization of the genes that increased in expression following *Satb2* knockdown included various signaling genes and those involved in cell differentiation (Figure 3E); this included several chromatin modifying genes known to accelerate differentiation such as *Smarca4* [41–43], *Med21* [44], *Kat2b* [45], and *Bdnf* [46]. Alternatively, the GO categorization of those genes significantly downregulated upon SATB2 removal included genes involved in muscle development (Figure 3F). Several of these genes are known repressors of differentiation, such as *Hdac4* [47], *Hdac5* [47], *Ncoa1* [48], and *Mylk* [49], and may act to limit satellite cell activation and muscle growth. However, several of the suppressed genes are known mediators of differentiation, such as *Nfatc4*. Given that the removal of SATB2 accelerates myoblast differentiation, these changes may be a compensatory response to a myogenic program that is well underway.

To obtain a more granular view of SATB2′s role in regulating myogenic progression, we assessed the expression changes of genes that possessed a SATB2 binding site located −5 kb to + 1 kb from the gene TSS. GO categorization revealed that SATB2 was bound near the TSS of many genes involved in chromatin organization and modification (Supplementary Figure S3E). However, several genes associated with cell differentiation or proliferation were found to have SATB2 near its TSS (Supplementary Figure S3D), suggesting a limited role for SATB2 in managing gene expression within promoter boundaries. Finally, we undertook a complimentary unbiased bioinformatic assessment of gene expression in wildtype versus siSATB2 cells using GSEA (gene set enrichment analysis) coupled to the Reactome cluster profiling program. Similar to the GO categorization, GSEA identified an up regulation in a number of gene groups/pathways associated with chromatin/chromosome modification, including cell cycle checkpoint and chromatid cohesion categories, along with the down regulation of a subcluster of skeletal muscle contractile gene sets. Interestingly, unlike the GO categorization, GSEA revealed an upregulation in three gene categories known to separately enhance myoblast cell fusion/differentiation, including IL4 signaling, hyaluronan metabolism, and Rho GTPases/formin activation [50–52].

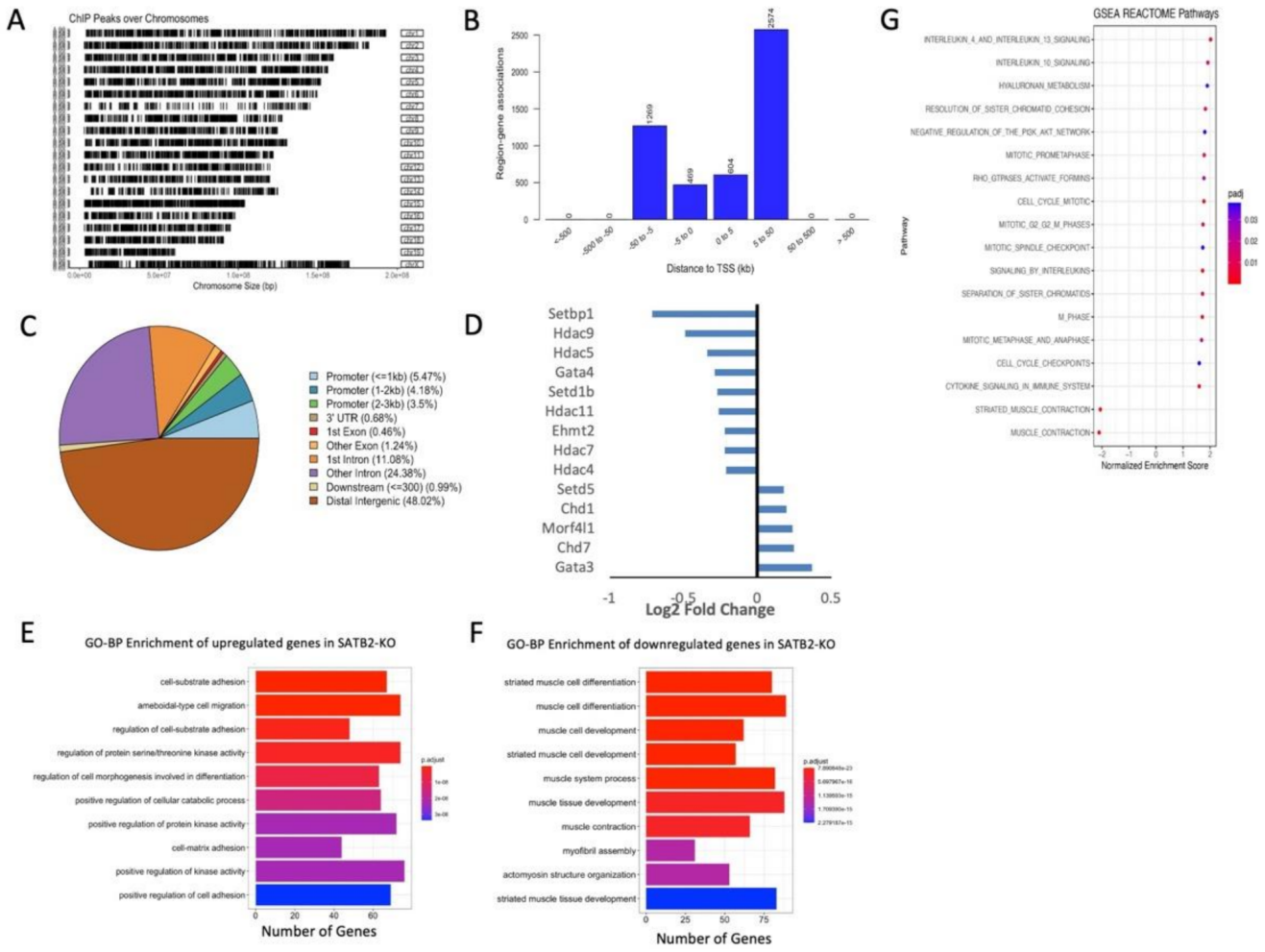


Figure 3. SATB2 regulates the expression of genes associated with chromatin remodeling and cell differentiation. (**A**) Chromosome distribution of SATB2 binding sites as determined from our ChIP-seq analyses (using ChIPseeker) on proliferating C2C12 cells. (**B**) Distribution of SATB2 binding sites within the muscle satellite cell genome with respect to the closest associated gene transcription start site (TSS). (**C**) Genomic context of SATB2 binding sites as determined by ChIPseeker. (**D**) Differentially expressed chromatin remodeling genes after inhibition of *Satb2* expression in 24 h differentiated myoblasts, as determined using DESeq2 following our RNA-seq experiments. (**E**) Top GO biological process (BP) categorization of the genes upregulated following *Satb2* repression in 24 h differentiated myoblasts. (**F**) Top GO-BP categorization of the downregulated genes following *Satb2* siRNA-mediated repression in 24 h differentiated myoblasts. Analyses in (**E**,**F**) were provided by ClusterProfiler. (**G**) Gene Set Enrichment Analysis (GSEA) for differentially expressed genes identified by RNA-seq. The Reactome bioinformatic tool was used for pathway analysis and identified a number of gene clusters known to promote myoblast fusion/differentiation.

3.4. SATB2 Regulates Chromatin Organization around Genes Related to Myogenesis

SATB2 was initially characterized as a MAR protein; however, there are no genome-wide data examining the role of SATB2 as a chromatin organizing factor and how this protein may control genome structure. To address this issue directly, we performed a chromatin conformation capture using the Hi-C method and examined this genome structure data in relation to our SATB2 ChIP-seq data. This combined genome-wide analysis indicated that SATB2 binding sites were enriched near the ends of genetic loops and slightly enriched near the ends of topologically associated domains (TADs) (Figure 4A,B show the results for siControl samples). Interestingly, SATB2 binding sites are near loop ends but

are not coincident with loop anchor points, as evidenced by the lack of direct overlap with CTCF binding motifs (data taken from ENCODE (ENCSR000AIJ) and the representative images shown in Figure 4C,D; Supplementary Figure S4B shows an absence of overlap for the SATB2 and CTCF binding sites), which are known markers of chromatin loop boundaries [53]. Given SATB2's localization at the base of these three-dimensional genetic structures, one might predict that the loss of SATB2 in muscle cells would cause chromatin remodeling at various loci. Indeed, the loss of SATB2 in proliferating myoblasts caused a ~12% reduction in the number of loop structures and only a ~1% reduction in TADs (Figure 4A,B).

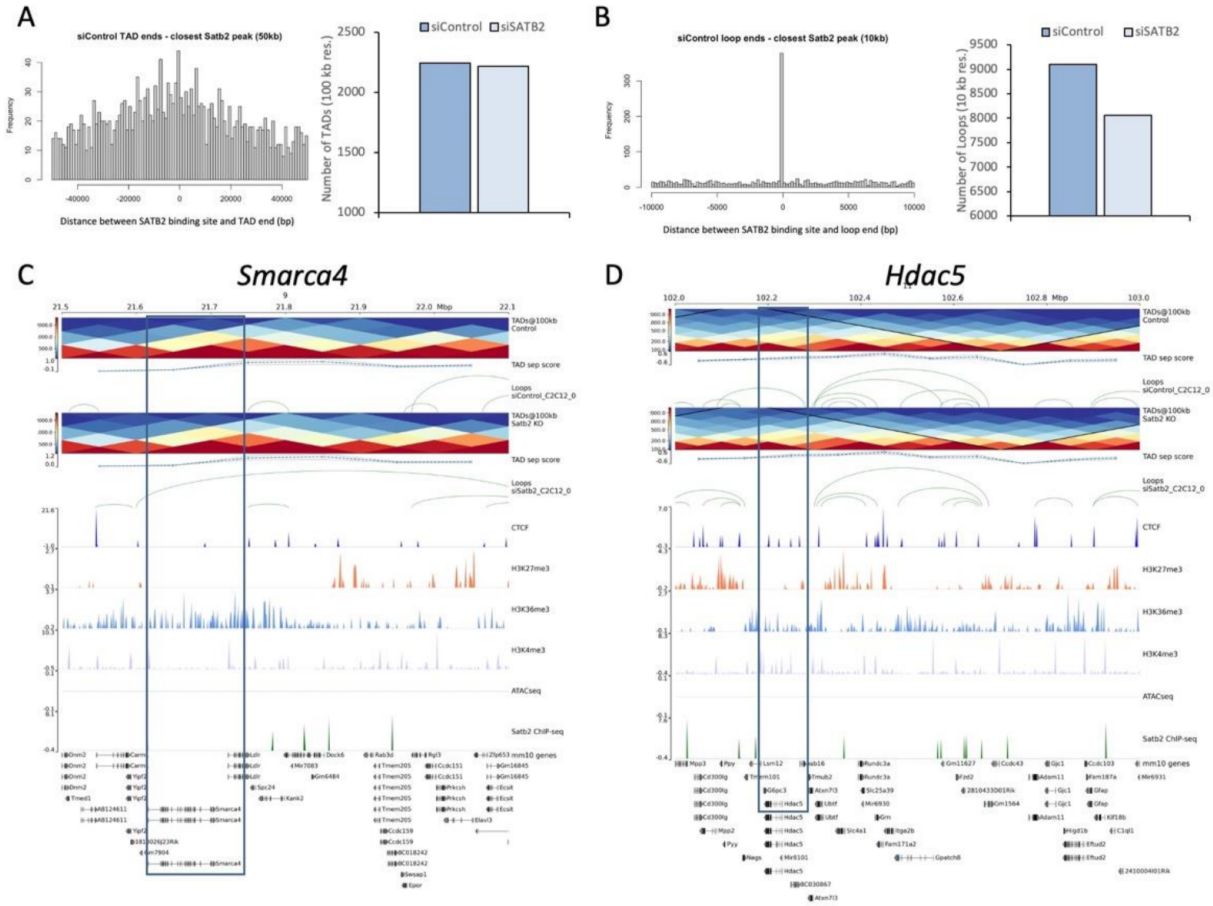


Figure 4. SATB2 is associated with chromatin loops and regulates loop dynamics in a subset of myogenesis-related genes. (**A**) Distribution of SATB2 binding sites with respect to TAD ends (50 kb resolution), as determined by Juicer v.1.6. The bar graph depicts the number of TADs detected at 100 kb resolution between siControl- and siSATB2-treated proliferating C2C12 samples. (**B**) Distribution of SATB2 binding sites with respect to loop ends (10 kb resolution), as determined by Juicer v.1.6. The bar graph depicts the number of loops detected at this resolution between siControl- and siSATB2-treated samples. (**C**) Chromatin structural landscape around the *Smarca4* locus. TADs and loops were assessed at 100 and 10 kb resolutions, respectively, using the tools described in (**A**,**B**). CTCF binding motifs were obtained from CTCF ChIP-seq data in mouse C2C12 cells from ENCODE (ENCSR000AIJ). The location of repressive (H3K27me3) and permissive (H3K4me3 and H3K36me3) histone marks were obtained by Asp et al. [35] from proliferating C2C12 myoblasts. The accession numbers for the associated ChIP-seq data are: H3K4me3 (GEO accession: GSM721292), H3K27me3 (GEO accession: GSM721294), and H3K36me3 (GEO accession: GSM721296). (**D**) Chromatin structural landscape around the *Hdac5* locus. The structural aspects depicted in the panel were obtained as described above.

The observation that SATB2 removal from muscle myoblasts affects chromatin loops and TADs suggests that SATB2 may manage gene expression through the direct modification of large-scale chromatin structures. In turn, the changes in chromatin organization may prepare muscle stem cells for differentiation by establishing the architecture needed for the efficient expression or repression of genes that manage the myogenic program. Indeed, the genetic locus surrounding the pro-differentiation factor, *Smarca4*, undergoes a transformation when SATB2 is lost, whereby a chromatin loop forms at the *Smarca4* site where one did not previously exist (Figure 4C). This loop structure is likely to facilitate *Smarca4* expression when cells are induced to differentiate. Conversely, for *Hdac5*, a chromatin loop containing the gene disappears following SATB2 ablation, which is concurrent with the repression of *Hdac5* expression during accelerated myogenesis (Figure 4D). This, however, may not be the only mechanism by which SATB2 contributes to gene expression changes, as several gene loci undergo subtle or limited local loop remodeling (Supplementary Figure S4). In these cases, SATB2 removal at more distal sites may have a broader effect on chromatin conformation that affects gene expression. An analysis of sub-TAD sized contact domains revealed alterations near several of the genes that lacked SATB2-derived chromatin loop remodeling. For instance, we observed changes in the number or width of contact domains present near *Ncoa1*, *Mylk*, and *Hdac4* (Supplementary Table S3). Interestingly, these types of changes were not observed in the genes that possessed local chromatin loop remodeling (i.e., *Smarca4*, *Hdac5*, and *Med21*).

The myogenic gene expression program has also been shown to be heavily influenced by epigenetic modifications, the most prominent of which is histone methylation, which acts to induce or repress expression within specific genes [54,55]. However, an investigation into the relationship between SATB2 binding sites and known muscle regulatory methylation sites such as the H3K27me3, H3K36me3, and H3K4me3 marks in proliferating myoblasts (taken from [35]) indicates that there is no significant correlation between histone methylation and SATB2 binding.

3.5. In Vivo Depletion of SATB2 in Muscle Satellite Cells Decreases Muscle Fiber Area and the Number of Pax7-Expressing Satellite Cells

To examine whether the loss of SATB2 would alter the muscle cell differentiation program in vivo, we conducted a preliminary analysis using a muscle progenitor cell-specific gene-targeted deletion of SATB2. Here, we generated the satellite cell-specific deletion of SATB2, under the control of tamoxifen induction, by crossing the *Satb2*$^{fl/fl}$ mouse strain [56] with the *Pax7CreER* mouse strain [53] to generate *Pax7CreER*/*Satb2*$^{fl/fl}$ mice. *Pax7CreER*/*Satb2*$^{fl/fl}$ mice (and requisite controls, *Satb2*$^{fl/fl}$ strain) were given tamoxifen at three weeks of age and monitored continuously (as per University of Ottawa Animal Care and Veterinary Service (ACVS) guidelines). The gross assessment of motor function and the physical state indicated that there were no substantial motor effects stemming from the depletion of SATB2. The disruption of SATB2 expression was confirmed via immunohistochemistry as well as western blots probing for SATB2 in myoblasts isolated from tamoxifen-treated mice (Supplementary Figure S5A,B). A quantitative assessment of muscle (extensor digitorum longus and TA) weights showed no significant differences between control and *Pax7CreER*/*Satb2*$^{fl/fl}$ mice (Supplementary Figure S5C). However, a histologic/morphologic analysis revealed notable alterations in skeletal muscle structure between control and *Pax7CreER*/*Satb2*$^{fl/fl}$ mice. For example, measurement of the fiber area within the TA from control and *Pax7CreER*/*Satb2*$^{fl/fl}$ mice indicated that the removal of SATB2 from muscle satellite cells led to a significant decrease in fiber area concurrent with an increase in fiber number per μm^2 as compared to the control strains (Figure 5A–C). Moreover, following tamoxifen treatment, the number of Pax7-expressing satellite cells within *Pax7CreER*/*Satb2*$^{fl/fl}$ TA muscle decreased substantially when compared to the *Satb2*$^{fl/fl}$ control muscle (Figure 5D). We also showed that primary myoblasts isolated from *Satb2*$^{fl/fl}$ mice and treated with Cre adenovirus displayed enhanced differentiation kinetics (Figure 5E) that were similar to that observed in siSatb2-targeted C2C12 cells (Figure 1E).

Indeed, these observations are consistent with the muscle hypotonia that often accompanies human SATB2 mutations [57].

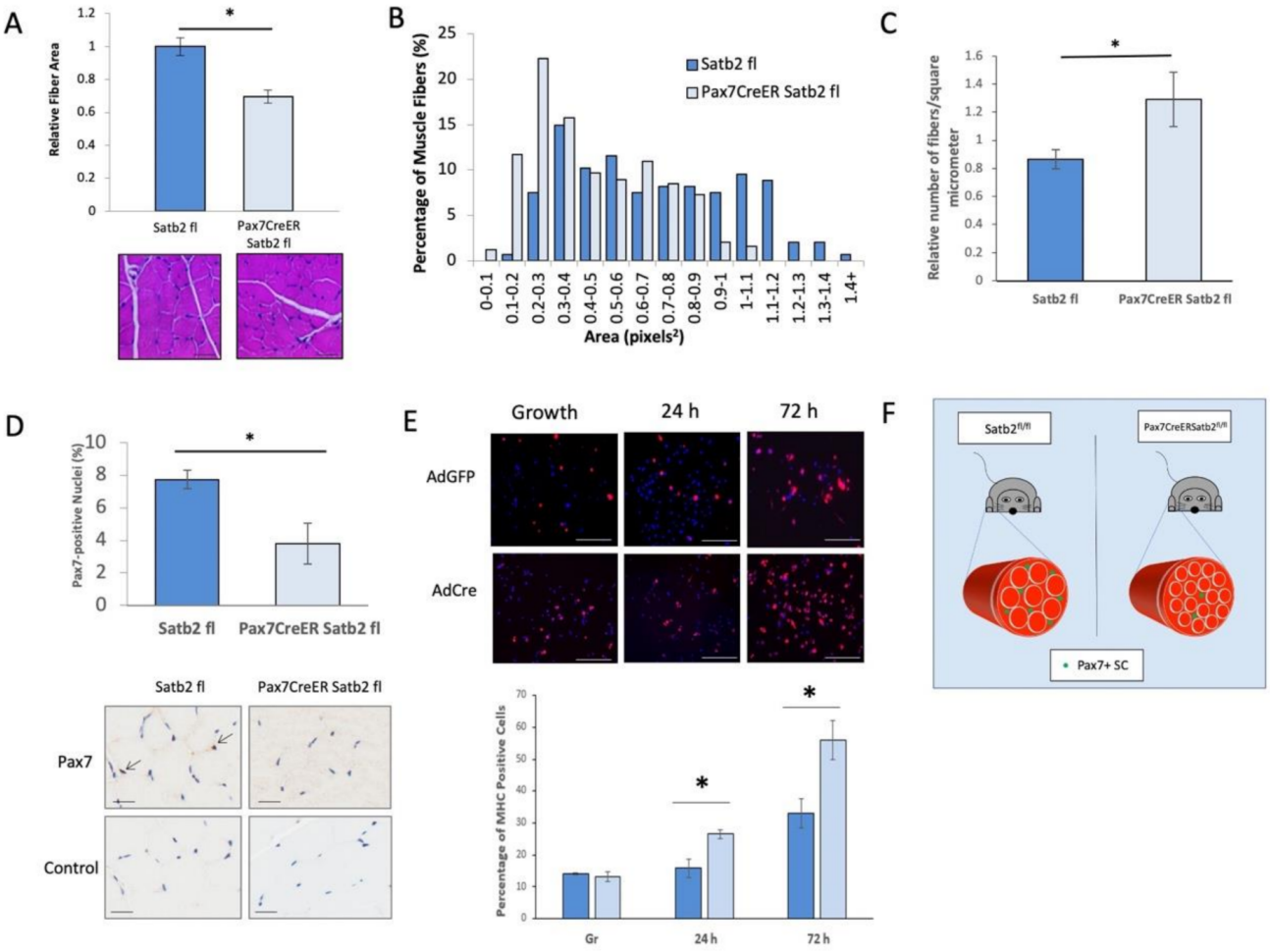


Figure 5. The in vivo reduction of SATB2 in muscle satellite cells decreases muscle fiber area and the number of Pax7-expressing satellite cells. (**A**) Tibialis anterior fiber areas decreased in *Pax7CreER/Satb2*$^{fl/fl}$ mice as compared to control (*Satb2*$^{fl/fl}$) mice. ImageJ was used to measure the fiber areas of hematoxylin and eosin (H&E) stained tissue sections. Data are the means ± SEM, $n = 3$ independent determinations on separate tissue samples. H&E stained sections are representative of $n = 3$ independent determinations. Scale bar: 100 µm. (**B**) The distribution of muscle fiber sizes across 15 bins. Over 200 muscle fibers were measured for both control and *Pax7CreER/Satb2*$^{fl/fl}$ mice. (**C**) Bar graph indicating the significant increase in the relative number of muscle fibers per µm^2. * indicates a significant difference between *Pax7CreER/Satb2*$^{fl/fl}$ and *Satb2*$^{fl/fl}$ mice, as determined by the Student's *t*-test ($p < 0.05$). Data are the means ± SEM, $n = 3$ independent determinations on separate tissue samples. (**D**) The percentage of nuclei that were Pax7-positive in control (Satb2 fl) and Satb2-ablated (Pax7CreER Satb2 fl) mouse muscle. The panels below the bar graph are representative tissue sections stained with a Pax7 antibody (Pax7) or no antibody (Control); scale: 50 µm. Arrows indicate positively stained nuclei. Data are representative of $n = 3$ independent determinations on separate tissue samples. * indicates a significant difference between *Pax7CreER/Satb2*$^{fl/fl}$ and *Satb2*$^{fl/fl}$ mice, as determined by the Student's *t*-test ($p < 0.05$). Data are means ± SEM for $n = 3$ independent determinations. (**E**) Immunofluorescent images (with accompanying bar graph) depicting an increase in myosin heavy chain (MHC) expression in Cre recombinase adenovirus (AdCre)-treated *SATB2*^{fl} myoblasts as compared AdGFP-treated controls. Data are representative of three independent determinations on separate myoblast isolations. MHC (red); DAPI (blue). The histogram below the images represents the quantification of the number of cells that expressed MHC relative to the total

number of cells counted. * indicates a statistically significant difference between AdGFP- and AdCre-treated samples using the Student's *t*-test ($p < 0.05$). Data are means ± SEM for $n = 3$ independent determinations. Scale bars = 100 μm. (**F**) Graphical representation of the observed biology following in vivo SATB2 ablation (see Supplementary Figure S5A,B for confirmation of SATB2 knockdown) from muscle satellite cells in vivo. More specifically, with no change in muscle size (see Supplementary Figure S5C), SATB2 loss caused a significant reduction in muscle fiber size as well as the number of Pax7-expressing satellite cells.

4. Discussion

Chromatin remodeling plays a central role in stem cell differentiation as it facilitates the dramatic shift in gene expression profiles that accompanies the exit from the cell cycle and the commitment to a particular lineage [1,4,58–62]. These changes are dependent upon both transcription mediated change concurrent with a shift in the epigenetic landscape, which may allow or limit gene expression in the relevant regions of the genome. One mechanism that will influence lineage-dependent transcription and epigenetic change is the presence (or absence) of higher-order MAR proteins that mediate chromatin structure and thereby physical access to key genetic loci [10,63,64]. While changes to the chromatin architecture within skeletal muscle progenitor cells have been observed [5], the proteins that govern these changes are not well-known. Here, we identified SATB2 as a chromatin organizer that plays a key role in mediating the progression of myoblasts toward the myogenic differentiation program.

Interestingly, the reduction of SATB2 expression did not affect myoblast proliferation (Supplementary Figure S1B,C), indicating that SATB2 acts to block cells from prematurely entering into a terminally differentiated state rather than maintain a proliferation competent milieu. The reduced expression of SATB2 during early myogenesis appears to accelerate the differentiated phenotype, as evidenced by the siRNA-mediated knockdown of SATB2, which in turn hastens the expression of MHC (Supplementary Figure S1E), accelerates the formation of myotubes during early differentiation (Figure 1E), which is concurrent with the upregulation of factors that promote myoblast fusion (Figure 3G) [50–52]. This coincides with our supposition that SATB2 may sequester differentiation inductive genes and, once removed, primes myoblasts for to engage the myogenic program. While our observations indicate the clear anti-myogenic effect of SATB2 expression, others have shown that SATB2 may enhance muscle differentiation in transformed cell lines such as Sol8 cells [65]. However, this same study also reported a strong synergy between SATB2 and BMP-induced transcription activity, which is known to be robustly anti-myogenic, suggesting in fact that SATB2 does antagonize differentiation akin to our observations reported here.

Collectively, our observations suggest that the loss of SATB2 is coincident to the progress of the differentiation program. Therefore, the mechanism by which the cell manages the repression of SATB2, and whether this occurs through transcriptional or post-translational mechanisms, is of paramount interest. Our identification of a major C-terminal fragment of SATB2, which becomes more prominent during muscle cell differentiation, suggests that regulated proteolysis may target the SATB2 protein and that this may be a critical step in managing cell differentiation (Figures 1A and S2D). Indeed, prior studies from our laboratory have shown that caspase 3 plays a prominent role in myogenesis, targeting and cleaving a number of proteins to engage the differentiation program [22,24,36,37]. In addition, evidence from apoptotic nuclei suggests that nuclear structural proteins are targeted by caspase proteases as part of a chromatin dissolution mechanism [19,39,40].

Despite the prominent role of caspase 3 in the differentiation process, we have noted that SATB2 is cleaved exclusively by the effector caspase, caspase 7 (Figure 2C). Suppression of this protease in vitro led to the sustained expression of SATB2 (Figure 2D,E), with a concomitant reduction in myogenesis (Figure 2G). Cytosolic caspase 7 activity has been suggested to promote odontogenesis [66,67], yet nuclear caspase 7 disposition is considered to be an exclusive hallmark of apoptosis [67,68]. For example, remodeling the chromatin

micro-environment through targeted protein cleavage events is considered to be a conserved feature of apoptosis, as exemplified by the effector caspase cleavage of scaffold attachment factor b1 [69], lamina-associated polypeptide 2α [70], and SMARCA2 and SMARCA4 [40]. Nevertheless, our observations support a novel model whereby an effector caspase, caspase 7, targets a protein substrate (SATB2) in order to prime the nuclear matrix to engage differentiation, independent of cell death. Presumably, during muscle cell differentiation, the activation of caspase 7 is mediated by the same pathway that leads to the induction of caspase 3 via the engagement of the mitochondrial intrinsic cell death pathway [23]. The pattern of caspase 7 activation is remarkably similar to caspase 3 in differentiating myoblasts (Figures 2F and S2E), which does support a common signaling origin. What is more speculative is whether a level of integration may exist between caspase 3 and caspase 7 activity, where the targeting of their respective substrates is coordinated to drive the same biologic alteration.

Central to SATB2′s role in controlling gene accessibility and expression is its impact on the chromatin structure during myoblast differentiation. Our ChIP-seq analyses indicate that SATB2 is bound throughout the mouse myoblast genome (Figure 3A), mainly in the intergenic spaces, but also near or within the coding regions of many genes (Figure 3C and integrated genomics viewer/IGV display for SATB2 targets, Supplementary Figure S5). The widespread distribution of SATB2 is consistent with the localization generally observed for matrix attachment region proteins [71]. However, the binding sites of SATB2 are positioned near but independent from chromatin loop anchor points (Figures 4B and S4B), indicating a hitherto unappreciated role in maintaining the three-dimensional chromatin architecture. This observation is somewhat in contrast to recent work, where neuronal SATB2 was found to be associated with the nuclear lamina via LEMD2, and the ablation of either protein resulted in substantial changes to nuclear organization and gene regulation [17]. Accordingly, our study adds critical insight into the entities that govern chromatin organization and cell function, suggesting that in a myogenic cell, SATB2 actively modifies chromatin architecture by establishing gene expression boundaries through the formation of loops and TAD-like domains. Whether SATB2-dependent loops and TADs follow a specific subnuclear distribution, i.e., preferential location at the nuclear edge, and whether this occurs in other cell lineages will require further investigation.

The loss of myoblast SATB2 affects the genetic landscape in several ways. Most directly, SATB2 ablation led to changes in three-dimensional chromatin folding, as evidenced by the reduction in chromatin loops following its suppression (Figure 4B). SATB2′s primary association with loop ends, as opposed to TAD ends, adheres to the original structural role described for the SATB family of proteins [72–75]. For example, SATB1 has been shown to mediate the formation of chromatin in helper T cells and is necessary for the expression of key cytokines following T cell activation [74]. In addition to affecting gene accessibility through chromatin loop remodeling, SATB2 loss appears to affect the expression of several important chromatin modifiers such as *Hdac4/5/7/9* (Figure 3D), which themselves can further alter the genetic architecture during myogenesis and modify the activity of key transcription factors such as the Mef2 family. These effects likely form the foundation for many of the gene expression changes observed in this study and may serve to open/close genetic locales that would support the progression of the myogenic program. A prime example of this would be for *Smarca4*, which typically supports cell differentiation [41–43]. When SATB2 is suppressed, *Smarca4* becomes part of a new loop structure, which may facilitate its expression during myogenesis (Figure 4C). Conversely, *Hdac5* exits a SATB2-mediated loop structure when the latter is ablated, which may serve to help repress its expression and aid in the progression of the myogenic program (Figure 4D).

While changes in local chromatin folding may be the primary driver of SATB2-associated gene expression changes, our data suggest that some changes may occur by means of an independent mechanism, potentially through changes in sub-TAD contact domains or changes in the DNA looping that influences distal target genes [76]. For instance, the local chromatin loops around *Ncoa1* and *Hdac4* appear virtually unaltered between

siControl and siSATB2 treatments (Supplementary Figure S4C). However, the sub-TAD domains at these loci changed following SATB2 suppression, with either domains being lost or altered in size (Supplementary Table S3). These sub-TAD domains can be critical regulators of gene expression and may lie within a larger TAD superstructure that remains unaltered [77,78]. Moreover, evidence exists that changes in sub-megabase contact domains may be critical for the fidelity of gene expression during embryonic stem cell and neural progenitor cell fate determination [79]. Interestingly, those genes (*Smarca4*, *Hdac5*, and *Med21*) that possessed changes in local loop structures did not possess contact domains at the resolutions tested, suggesting that these modes of regulation may be mutually exclusive. Taken together, SATB2′s impact on genome loop structures and sub-TAD level contact domains appear to hamper myoblast differentiation, necessitating SATB2 removal for the induction of chromatin reorganization that facilitates the myogenic differentiation program.

In addition to alterations in local chromatin environments, others have noted that skeletal muscle differentiation coincides with macro nuclear alterations, where transcriptions factors and chromatin remodeling proteins, working in concert, reduce inter-chromosomal contacts at critical differentiation-specific loci [33]. We have noted that the loss of SATB2 expression is concurrent with a global reduction in the number of inter-chromosomal links (5.25 ± 0.22 vs. $5.90 \pm 0.01\%$, $p = 0.04$, see Supplementary Table S5), suggesting that SATB2 may similarly modify this macro nuclear structure, perhaps reducing inter-chromosomal links that repress the differentiation program. However, an accurate identification of such SATB2 sensitive loci will require alternative 3C experiments that provide greater sequence depth compared to our current Hi-C approach [80,81].

In closing, our data do not favor a model whereby SATB2 binds to any traditional chromatin anchor. Rather, our data suggest that SATB2 binds within the looped chromatin, proximal to CTCF and its cohesin-linked protein complexes, which together anchor the loop structure (as defined in [82]). Furthermore, the very modest association of SATB2 at anchor points implies a loop-oriented binding event rather than a direct physical association with CTCF per se (we also see SATB2 bind within a cohort of subTAD domains, which again reinforces the loop-binding propensity of this protein). Intra-chromatin loop protein binding (and the regulation of gene expression) has not been reported in the literature previously; however, our data are consistent with this novel and exciting hypothesis. Indeed, a very recent publication has shown that the Jpx RNA can bind directly within a chromatin loop, modify its structure, separately impair CTCF binding kinetics, which all together leads to altered gene expression [83]. This aforementioned study together with our observations suggest a new and unappreciated mechanism for controlling the local chromatin structure, affecting gene expression, and influencing cell fate decisions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cells11060966/s1, Figure S1: SATB2 expression represses myoblast differentiation, Figure S2: Caspase 7 cleaves SATB2 during early skeletal muscle differentiation and is necessary for myogenesis, Figure S3: Quality control and validation data related to ChIP- and RNA-seq experiments, Figure S4: Validation data for SATB2 genomic analysis, Figure S5: Integrative Genomics Viewer (IGV) tracks depicting select genes (blue) related to proliferation/differentiation and the SATB2 binding sites within that area of the myoblast genome (grey), Figure S6: Validation of SATB2 loss following tamoxifen treatment of *Satb2*$^{fl/fl}$ and *Pax7CreER/Satb2*$^{fl/fl}$ mice, Table S1: Mass spectrometry analysis of the major cleavage fragment of SATB2, Table S2: Top 10 enriched DNA sequence motifs within the full set of SATB2 ChIP peaks, Table S3: Sub-TAD contact domains identified near *Ncoa1*, *Hdac4*, and *Mylk* genes for siControl- and siSATB2-treated proliferating C2C12 cells, Table S4: ChIP-seq statistics for proliferating C2C12 cells, Table S5: HiC replicate statistics for siControl- and siSATB2-treated proliferating C2C12 cells

Author Contributions: R.A.V.B.: conception and design of the work, acquisition, analysis, and interpretation of data; M.H.A.-K.: conception and design of the work, acquisition, analysis, and interpretation of data; S.B.: acquisition and analysis of data; D.A.: acquisition and analysis of data; A.C.: acquisition, analysis, and interpretation of data; H.B.: acquisition, analysis, and interpretation of data; G.D.: conception and design of the work; G.A.: conception and design of the work; F.J.D.:

conception and design of the work; L.A.M.: conception and design of the work, analysis and interpretation of data. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Canadian Institutes of Health Research grants to Lynn Megeney and Jeffrey Dilworth (156120 for L.M. and FDN-143330 for J.D.).

Institutional Review Board Statement: The study was conducted in accordance with the Canadian Council on Animal Care (CCAC) guidelines and the University of Ottawa Animal Care Committee protocols.

Informed Consent Statement: Not applicable.

Data Availability Statement: Mass spectrometry data are available in the MassIVE repository under the identifier MSV000085254. Genomics (ChIP-seq, RNA-seq, and Hi-C) data are available using GEO accession GSE185437.

Acknowledgments: The authors would like to thank members of the Megeney laboratory and Michael Rudnicki for the helpful discussions. We would also like to thank Lawrence Puente for the mass spectrometry analysis (Proteomics Core Facility, Ottawa Hospital Research Institute) and Chris Porter for assistance with the bioinformatic analysis of the ChIP-seq, RNA-seq, and Hi-C genome libraries.

Conflicts of Interest: The authors declare no conflict of interest.

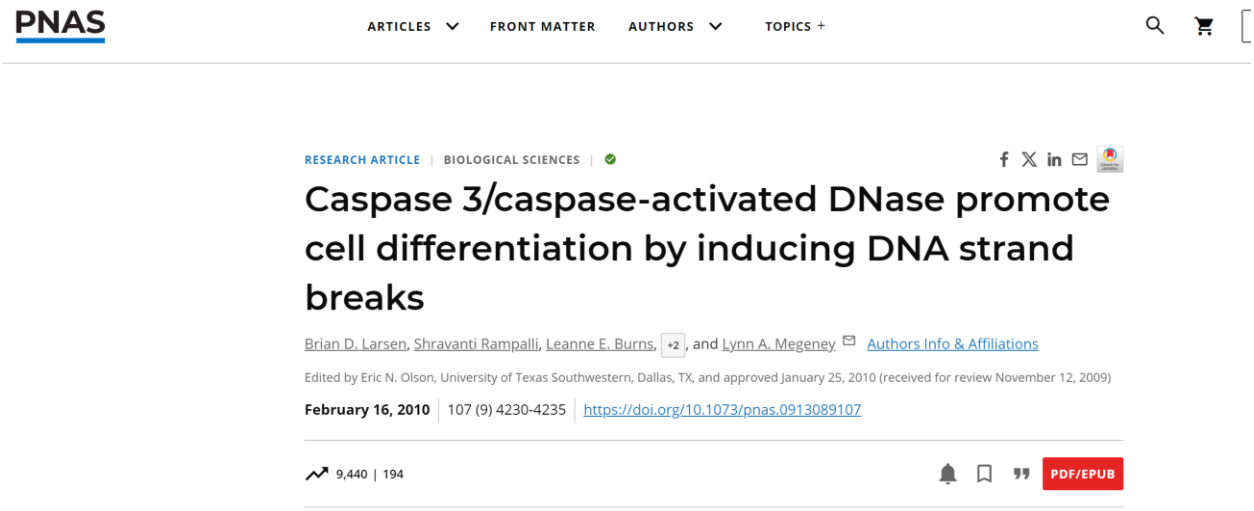
References

1. Müller, C.; Leutz, A. Chromatin remodeling in development and differentiation. *Curr. Opin. Genet. Dev.* **2001**, *11*, 167–174. [CrossRef]
2. Forcales, S.V.; Albini, S.; Giordani, L.; Malecova, B.; Cignolo, L.; Chemov, A.; Coultinho, P.; Saccone, V.; Consalvi, S.; Williams, R.; et al. Signal-dependent incorporation of MyoD-BAF60c into Brg1-based SWI/SNF chromatin-remodeling complex. *EMBO J.* **2012**, *31*, 301–316. [CrossRef] [PubMed]
3. Gómez-Díaz, E.; Corces, V.G. Architectural proteins: Regulators of 3D genome organization in cell fate. *Trends Cell Biol.* **2014**, *24*, 703–711. [CrossRef] [PubMed]
4. Hota, S.K.; Bruneau, B.G. ATP-dependent chromatin remodeling during mammalian development. *Development* **2016**, *143*, 2882–2897. [CrossRef]
5. Doynova, M.D.; Markworth, J.F.; Cameron-Smith, D.; Vickers, M.H.; O'Sullivan, J.M. Linkages between changes in the 3D organization of the genome and transcription during myotube differentiation in vitro. *Skelet. Muscle* **2017**, *7*, 5. [CrossRef]
6. Robinson, D.C.L.; Dilworth, J.F. Epigenetic regulation of adult myogenesis. *Curr. Top. Dev. Biol.* **2018**, *126*, 235–284.
7. Bushey, A.M.; Dorman, E.R.; Corces, V.G. Chromatin insulators: Regulatory mechanisms and epigenetic inheritance. *Mol. Cell* **2008**, *32*, 1–9. [CrossRef]
8. Arope, S.; Harraghy, N.; Pjanic, M.; Mermod, N. Molecular characterization of a human matrix attachment region epigenetic regulator. *PLoS ONE* **2013**, *8*, e79262. [CrossRef]
9. Hernández-Hernández, J.M.; Mallappa, C.; Nasipak, B.T.; Oesterreich, S.; Imbalzano, A.N. The scaffold attachment factor b1 (Safb1) regulates myogenic differentiation by facilitating the transition of myogenic gene chromatin from a repressed to an activated state. *Nucleic Acid Res.* **2013**, *41*, 5704–5716. [CrossRef]
10. Savarese, F.; Dávila, A.; Nechanitzky, R.; De La Rosa-Velazquez, I.; Pereira, C.F.; Engelke, R.; Takahashi, K.; Jenuwein, T.; Kohwi-Shigematsu, T.; Fisher, A.G.; et al. Satb1 and Satb2 regulate embryonic stem cell differentiation and *Nanog* expression. *Genes Dev.* **2009**, *23*, 2625–2638. [CrossRef]
11. Dong, W.; Zhang, P.; Fu, Y.; Ge, J.; Cheng, J.; Yuan, H.; Jiang, H. Roles of SATB2 in site-specific stemness, autophagy and senescence of bone marrow mesenchymal stem cells. *J. Cell. Physiol.* **2015**, *230*, 680–690. [CrossRef]
12. Britanova, O.; Akopov, S.; Lukyanov, S.; Gruss, P.; Tarabykin, V. Novel transcription factor SATB2 interacts with matrix attachment region DNA elements in a tissue-specific manner and demonstrates cell-type-dependent expression in the developing mouse CNS. *Eur. J. Neurosci.* **2005**, *21*, 658–668. [CrossRef]
13. Dobreva, G.; Chahrour, M.; Dautzenberg, M.; Chirivella, L.; Kanzler, B.; Fariñas, I.; Karsenty, G.; Grosschedl, R. SATB2 is a multifunctional determinant of craniofacial patterning and osteoblast differentiation. *Cell* **2006**, *125*, 971–986. [CrossRef]
14. Han, H.J.; Russo, J.; Kohwi, Y.; Kohwi-Shigematsu, T. SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis. *Nature* **2008**, *452*, 187–193. [CrossRef]
15. Leone, D.P.; Heavner, W.E.; Ferenczi, E.A.; Dobreva, G.; Huguendard, J.R.; Grosschedl, R.; McConnell, S.K. Satb2 regulates the differentiation of both callosal and subcerebral projection neurons in the developing cerebral cortex. *Cereb. Cortex* **2015**, *25*, 3406–3419. [CrossRef]

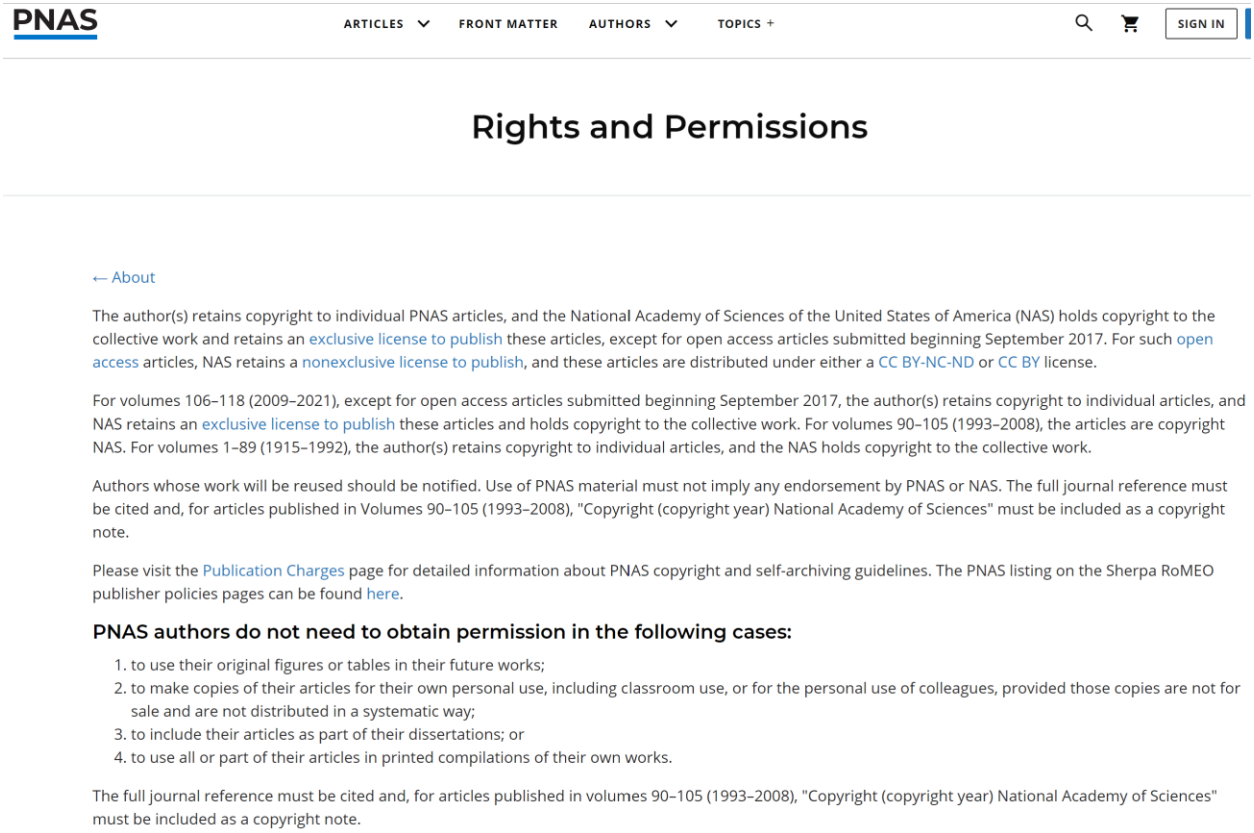
16. Urrutia, G.A.; Ramachandran, H.; Cauchy, P.; Boo, K.; Ramamoorthy, S.; Boller, S.; Dogan, E.; Clapes, T.; Trompouki, E.; Torres-Padilla, M.-E.; et al. ZFP451-mediated SUMOylation of SATB2 drives embryonic stem cell differentiation. *Genes Dev.* **2021**, *35*, 1142–1160. [CrossRef]
17. Feurle, P.; Abentung, A.; Cera, I.; Wahl, N.; Ablinger, C.; Bucher, M.; Stefan, E.; Sprenger, S.; Teis, D.; Fischer, A.; et al. SATB2-LEMD2 interaction links nuclear shape plasticity to regulation of cognition-related genes. *EMBO J.* **2021**, *40*, e103701. [CrossRef]
18. Bengani, H.; Handley, M.; Alvi, M.; Ibitoye, R.; Lees, M.; Lynch, S.A.; Lam, W.; Fannemel, M.; Nordgren, A.; Malmgren, H.; et al. Clinical and molecular consequences of disease-associated de novo mutations in $SATB_2$. *Nature* **2017**, *19*, 900–908. [CrossRef]
19. Galande, S.; Dickinson, L.A.; Milan, I.S.; Sikorska, M.; Kohwi-Shigematsu, T. SATB1 cleavage by caspase 6 disrupts PDZ domain-mediated dimerization, causing detachment from chromatin early in T-cell apoptosis. *Mol. Cell Biol.* **2001**, *21*, 5591–5604. [CrossRef]
20. Unsain, N.; Barker, P.A. New Views on the Misconstrued: Executioner caspases and their diverse non-apoptotic roles. *Neuron* **2015**, *88*, 461–474. [CrossRef]
21. Bell, R.A.V.; Megeney, L.A. Evolution of caspase-mediated cell death and differentiation: Twins separated at birth. *Cell Death Diff.* **2017**, *24*, 1359–1368. [CrossRef]
22. Fernando, P.; Kelly, J.F.; Balazsi, K.; Slack, R.S.; Megeney, L.A. Caspase 3 activity is required for skeletal muscle differentiation. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11025–11030. [CrossRef]
23. Murray, T.V.A.; McMahon, J.M.; Howley, B.A.; Stanley, A.; Ritter, T.; Mohr, A.; Zwacka, R.; Fearnhead, H.O. A non-apoptotic role for caspase-9 in muscle differentiation. *J. Cell Sci.* **2008**, *121*, 3786–3793. [CrossRef]
24. Larsen, B.D.; Rampalli, S.; Burns, L.E.; Brunette, S.; Dilworth, F.J.; Megeney, L.A. Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4230–4235. [CrossRef]
25. Yu, G.; Wang, L.-G.; He, Q.-Y. ChIPseeker: An R/Bioconductor Package for Chip Peak Annotation, Comparison and Visualization. *Bioinformatics* **2015**, *31*, 2382–2383. [CrossRef]
26. Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **2012**, *16*, 284–287. [CrossRef]
27. Heinz, S.; Benner, C.; Spann, N.; Bertolino, E.; Lin, Y.C.; Laslo, P.; Cheng, J.X.; Murre, C.; Singh, H.; Glass, C.K. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **2010**, *38*, 576–589. [CrossRef] [PubMed]
28. Kim, D.; Pertea, G.; Trapnell, C.; Pimental, H.; Kelley, R.; Salzberg, S.L. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **2013**, *14*, R36. [CrossRef] [PubMed]
29. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923–930. [CrossRef] [PubMed]
30. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]
31. Zhu, A.; Ibrahim, J.G.; Love, M.I. Heavy-tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics* **2019**, *35*, 2084–2092. [CrossRef]
32. Raudvere, U.; Kolberg, L.; Kuzmin, I.; Arak, T.; Adler, P.; Peterson, H.; Vilo, J. g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucl. Acids Res.* **2019**, *47*, W191–W198. [CrossRef]
33. Korotkevich, G.; Sukhov, V.; Budin, N.; Shipak, B.; Artyomov, M.N.; Sergushichev, A. Fast gene set enrichment analysis. *bioRxiv* **2016**. [CrossRef]
34. Durand, N.C.; Shamim, M.S.; Machol, I.; Rao, S.S.P.; Huntley, M.H.; Lander, E.S.; Aiden, E.L. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **2016**, *3*, 95–98. [CrossRef]
35. Asp, P.; Blum, R.; Vethantham, V.; Parisi, F.; Micsinai, M.; Cheng, J.; Bowman, C.; Kluger, Y.; Dynlacht, D. Genome-wide remodeling of the epigenetic landscape during myogenic differentiation. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E149–E158. [CrossRef]
36. Dick, S.A.; Chang, N.C.; Dumont, N.A.; Bell, R.A.V.; Putinski, C.; Kawabe, Y.; Litchfield, D.W.; Rudnicki, M.A.; Megeney, L.A. Caspase 3 cleavage of Pax7 inhibits self-renewal of satellite cells. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E5246–E5252. [CrossRef]
37. Al-Khalaf, M.H.; Blake, L.E.; Larsen, B.D.; Bell, R.A.V.; Brunette, S.; Parks, R.J.; Rudnicki, M.A.; McKinnon, P.J.; Dilworth, F.J.; Megeney, L.A. Temporal activation of XRCC1-mediated DNA repair is essential for muscle differentiation. *Cell Discov.* **2016**, *12*, 15041. [CrossRef]
38. Slaymaker, I.M.; Gao, L.; Zetsche, B.; Scott, D.A.; Yan, W.X.; Zhang, F. Rationally engineered Cas9 nucleases with improved specificity. *Science* **2016**, *351*, 84–88. [CrossRef]
39. Sun, Y.; Wang, T.; Su, Y.; Yin, Y.; Xu, S.; Ma, C.; Han, X. The behavior of SATB1, a MAR-binding protein, in response to apoptosis stimulation. *Cell Biol. Int.* **2006**, *30*, 244–247. [CrossRef]
40. Dudek, A.H.; Pfaff, F.; Bolte, H.; Waguia Kontchou, C.; Schwemmie, M. Partial inactivation of chromatin remodelers SMARCA2 and SMARCA4 in virus-infected cells by caspase-mediated cleavage. *J. Virol.* **2018**, *92*, e00343-18. [CrossRef]
41. Yu, Y.; Chen, Y.; Kim, B.; Wang, H.; Zhao, C.; He, X.; Liu, L.; Liu, W.; Wu, L.M.N.; Mao, M.; et al. Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell* **2013**, *152*, 248–261. [CrossRef]

42. Griffin, C.T.; Brennan, J.; Magnuson, T. The chromatin-remodeling enzyme BRG1 plays an essential role in primitive erythropoiesis and vascular development. *Development* **2008**, *135*, 493–500. [CrossRef]
43. Albini, S.; Toto, P.C.; Dall'Agnese, A.; Malecova, B.; Cenciarelli, C.; Felsani, A.; Caruso, M.; Bultman, S.J.; Puri, P.L. Brahma is required for cell cycle arrestand late muscle gene expression during skeletal myogenesis. *EMBO Rep.* **2015**, *16*, 1037–1050. [CrossRef]
44. Oda, Y.; Chalkley, R.J.; Burlingame, A.L.; Bikle, D.D. The transcriptional coactivator DRIP/mediator complex is involved in vitamin D receptor function and regulates keritinocyte differentiation. *J. Investig. Dermatol.* **2010**, *130*, 2377–2388. [CrossRef]
45. Sen, R.; Pezoa, S.A.; Shull, L.C.; Hernandez-Lagunas, L.; Niswander, L.A.; Artinger, K.B. Kat2a and Kat2b acetyltransferase activity regulates craniofacial cartilage and bone differentiation in zebrafish and mice. *J. Dev. Biol.* **2018**, *6*, 27. [CrossRef]
46. Hromadkova, L.; Bezdekova, D.; Pala, J.; Schedin-Weiss, S.; Tjernberg, L.O.; Hoschi, C.; Ovespian, S.V. Brain-derived neurotrophic factor (BDNF) promotes molecular polarization and differentiation of immature neuroblastoma cells into definitive neurons. *Biochim. Biophys. Acta* **2020**, *1867*, 118737. [CrossRef]
47. McKinsey, T.A.; Zhang, C.-L.; Lu, J.; Olson, E.N. Signal-dependent nuclear export of a histone deactylase regulates muscle differentiation. *Nature* **2000**, *408*, 106–111. [CrossRef]
48. Wang, L.; Li, W.; Li, K.; Guo, Y.; Liu, D.; Yao, Z.; Lin, X.; Li, S.; Jiang, Z.; Liu, Q.; et al. The oncogenic roles of nuclear receptor coactivator 1 in human esophageal carcinoma. *Cancer Med.* **2018**, *7*, 5205–5216. [CrossRef]
49. Zhou, X.; Liu, Y.; You, J.; Zhang, H.; Zhang, X.; Ye, L. Myosin light-chain kinase contributes to the proliferation and migration of breast cancer cells through cross-talk with activated ERK1/2. *Cancer Lett.* **2008**, *270*, 312–327. [CrossRef]
50. Horsely, V.; Jansen, K.M.; Mills, S.T.; Paviath, G.K. IL-4 acts as a myoblast recruitment factor dueing mammalian muscle growth. *Cell* **2003**, *113*, 483–494. [CrossRef]
51. Poveda-Reyes, S.; Moulisova, V.; Sanmartin-Masia, E.; Quintanilla-Sierra, L.; Salmeron-Sanchez, M.; Ferrer, G.G. Gelatin-hyaluronic acid hydrogels with tuned stiffness counterbalance cellular forces and promote cell differentiation. *Macromol. Biosci.* **2016**, *16*, 1311–1324. [CrossRef] [PubMed]
52. Gopinath, S.D.; Narumiya, S.; Dhawan, J. The RhoA effector mDiaphanous regulates MyoD expression and cell cycle progression via SRF-dependent and SRF-independent pathways. *J. Cell Sci.* **2007**, *120*, 3086–3098. [CrossRef] [PubMed]
53. Nishijo, K.; Hosoyama, T.; Bjornson, C.R.R.; Schaffer, B.S.; Prajapati, S.I.; Bahadur, A.N.; Hansen, M.S.; Blandford, M.C.; McGleish, A.T.; Rubin, B.P.; et al. Biomarker system for studying muscle, stem cells, and cancer in vivo. *FASEB J.* **2009**, *23*, 2681–2690. [CrossRef] [PubMed]
54. Seenundun, S.; Rampalli, S.; Liu, Q.C.; Aziz, A.; Palii, C.; Hong, S.; Blais, A.; Brand, M.; Ge, K.; Dilworth, F.J. UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO J.* **2010**, *29*, 1401–1411. [CrossRef]
55. Faralli, H.; Wang, C.; Nakka, K.; Benyoucef, A.; Sebastian, S.; Zhuang, L.; Chu, A.; Palii, C.G.; Liu, C.; Camellato, B.; et al. UTX demethylase activity is required for satellite cell-mediated muscle regeneration. *J. Clin. Investig.* **2016**, *126*, 1555–1565. [CrossRef]
56. Jaitner, C.; Reddy, C.; Abentung, A.; Whittle, N.; Rieder, D.; Delekate, A.; Korte, M.; Jain, G.; Fischer, A.; Sananbenesi, F.; et al. Satb2 determines miRNA expression and long-term memory in the adult central nervous system. *eLife* **2016**, *5*, e17361. [CrossRef]
57. Zhao, X.; Qu, Z.; Tickner, J.; Xu, J.; Dai, K.; Zhang, X. The role of SATB2 in skeletogenesis and human disease. *Cytokine Gr. Factor Rev.* **2014**, *25*, 35–44. [CrossRef]
58. de la Serna, I.L.; Ohkawa, Y.; Imbalzano, A.N. Chromatin remodeling in mammalian differentiation: Lessons from ATP-dependent remodelers. *Nat. Rev. Genet.* **2006**, *7*, 461–473. [CrossRef]
59. Keenen, B.; de la Serna, I.L. Chromatin remodeling in embryonic stem cells: Regulating the balance between pluripotency and differentiation. *J. Cell. Physiol.* **2008**, *219*, 1–7. [CrossRef]
60. Fisher, C.L.; Fisher, A.G. Chromatin states in pluripotent, differentiated, and reprogrammed cells. *Curr. Opin. Genet. Dev.* **2011**, *21*, 140–146. [CrossRef]
61. Chen, T.; Dent, S.Y.R. Chromatin modifiers: Regulators of cellular differentiation. *Nat. Rev. Genet.* **2014**, *15*, 93–106. [CrossRef]
62. Ye, Y.; Li, M.; Gu, L.; Chen, X.; Shi, J.; Zhang, X.; Jiang, C. Chromatin remodeling during in vivo neural stem cells differentiating to neurons in early *Drosophila* embryos. *Cell Death Diff.* **2017**, *24*, 409–420. [CrossRef]
63. Hawkins, S.M.; Kohwi-Shigematsu, T.; Skalnik, D.G. The matrix attachment region-binding protein SATB1 interacts with multiple elements within the gp91phox promoter and is down-regulated during myeloid differentiation. *J. Biol. Chem.* **2001**, *276*, 44472–44480. [CrossRef]
64. Asanoma, K.; Kubota, K.; Chakraborty, D.; Renaud, S.J.; Wake, N.; Fukushima, K.; Soares, M.J.; Rumi, M.A. SATB homeobox proteins regulate trophoblast stem cell renewal and differentiation. *J. Biol. Chem.* **2012**, *287*, 2257–2268. [CrossRef]
65. Addison, W.N.; Hall, K.C.; Kokabu, S.; Matsubara, T.; Fu, M.M.; Gori, F.; Baron, R. Zfp423 Regulates Skeletal Muscle Regeneration and Proliferation. *Mol. Cell. Biol.* **2019**, *39*, e00447-18. [CrossRef]
66. Matalova, E.; Lesot, H.; Svandova, E.; Vanden Berghe, T.; Sharpe, P.T.; Healy, C.; Vandenabeele, P.; Tucker, A.S. Caspase-7 participates in differentiation of cells forming dental hard tissues. *Dev. Growth Differ.* **2013**, *55*, 615–621. [CrossRef]
67. Svandova, E.; Vesela, B.; Tucker, A.S.; Matalova, E. Activation of pro-apoptotic caspases in non-apoptotic cells during odontogenesis and related osteogenesis. *Front. Physiol.* **2018**, *9*, 174. [CrossRef]
68. Zhivotovsky, B.; Samali, A.; Gahm, A.; Orrenius, S. Caspases: Their intracellular localization and translocation during apoptosis. *Cell Death Diff.* **1999**, *6*, 644–651. [CrossRef]

69. Lee, Y.B.; Colley, S.; Norman, M.; Biamonti, G.; Uney, J.B. SAFB re-distribution marks steps of the apoptotic process. *Exp. Cell Res.* **2007**, *313*, 3914–3923. [CrossRef]
70. Gotzmann, J.; Vicek, S.; Foisner, R. Caspase-mediated cleavage of the chromosome-binding domain of lamina-associated polypeptide 2α. *J. Cell Sci.* **2000**, *113*, 3769–3780. [CrossRef]
71. Narwade, N.; Patel, S.; Alam, A.; Chattopadhyay, S.; Mittal, S.; Kulkarni, A. Mapping of scaffold/matrix attachment regions in human genome: A data mining exercise. *Nucleic Acids Res.* **2019**, *47*, 7247–7261. [CrossRef]
72. Yasui, D.; Miyano, M.; Cai, S.; Varga-Weisz, P.; Kohwi-Shigematsu, T. SATB1 targets chromatin remodeling to regulate genes over long distances. *Nature* **2002**, *419*, 641–645. [CrossRef]
73. Galande, S.; Purbey, P.K.; Notani, D.; Kumar, P. The third dimension of gene regulation: Organization of dynamic chromatin loopscape by SATB. *Curr. Opin. Genet. Dev.* **2007**, *17*, 408–414. [CrossRef]
74. Cai, S.; Lee, C.C.; Kohwi-Shigematsu, T. SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat. Genet.* **2006**, *38*, 1278–1288. [CrossRef]
75. Kumar, P.; Bischof, O.; Purbey, P.K.; Notani, D.; Urlaub, H.; Dejean, A.; Galande, S. Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. *Nat. Cell Biol.* **2007**, *9*, 45–56.
76. Bouwman, B.A.M.; de Laat, W. Getting the genome in shape: The formation of loops, domains and compartments. *Genome Biol.* **2015**, *16*, 154. [CrossRef]
77. Filippova, D.; Patro, R.; Duggal, G.; Kingsford, C. Identification of alternative topological domains in chromatin. *Algorithms Mol. Biol.* **2014**, *9*, 14. [CrossRef]
78. Yu, W.; He, B.; Tan, K. Identifying topologically associating domains and subdomains by Gaussian mixture model and proportion test. *Nat. Commun.* **2017**, *8*, 535. [CrossRef] [PubMed]
79. Phillips-Cremins, J.E.; Sauria, M.E.G.; Sanyal, A.; Gerasimova, T.I.; Lajoie, B.R.; Bell, J.S.K.; Ong, C.-T.; Hookway, T.A.; Guo, C.; Sun, Y.; et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **2013**, *153*, 1281–1295. [CrossRef] [PubMed]
80. Maass, P.G.; Barutcu, A.R.; Weiner, C.L.; Rinn, J.L. Inter-chromosomal contact properties in live-cell imagin and in Hi-C. *Mol. Cell* **2018**, *69*, P1039–P1045. [CrossRef] [PubMed]
81. Xiong, K.; Ma, J. Revealing Hi-C subcompartments by imputing inter-chromosomal chromatin interactions. *Nat. Commun.* **2019**, *10*, 5069. [CrossRef]
82. Abbas, A.; He, X.; Niu, J.; Zhou, B.; Zhu, G.; Ma, T.; Song, J.; Gao, J.; Zhang, M.Q.; Zeng, J. Integrating Hi-C and FISH data for modeling of the 3D organization of chromosomes. *Nat. Commun.* **2019**, *10*, 2049. [CrossRef]
83. Oh, H.J.; Aguilar, R.; Kesner, B.; Lee, H.-G.; Kriz, A.J.; Chu, H.-P.; Lee, J.T. Jpx RNA regulates CTCF anchor site selection and formation of chromosome loops. *Cell* **2021**, *184*, P6157–P6173. [CrossRef]

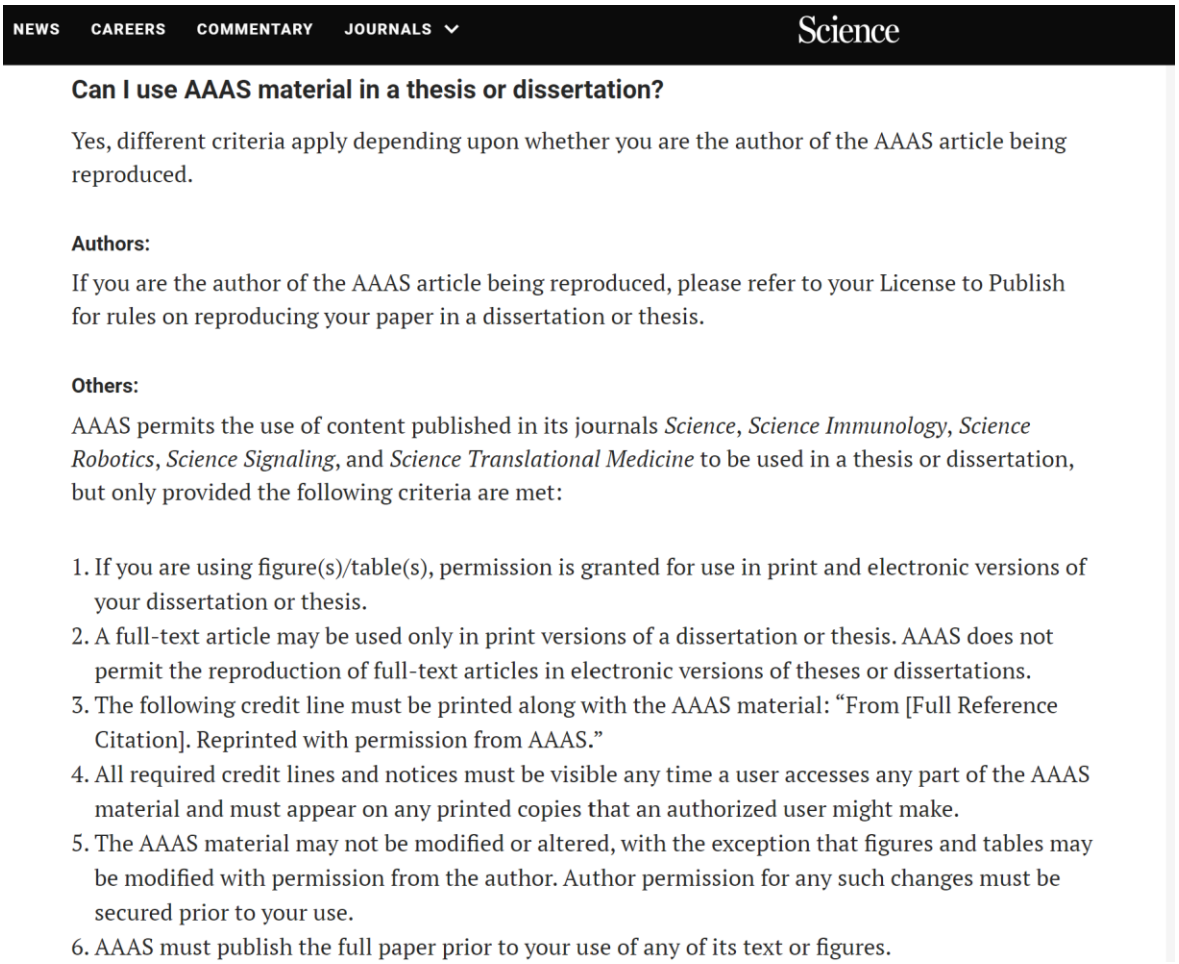


Reprinted with permission from the National Academy of Sciences of the United States of America," to include figures 1 and 2, in the introduction.





From [Larsen et al. Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress. Science. 2022]. Reprinted with permission from AAAS," to include figure 4, in the introduction.



SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Oct 21, 2024

This Agreement between Dalal Alsowaida ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5893590481018
License date	Oct 21, 2024
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Cancer
Licensed Content Title	A fate worse than death: apoptosis as an oncogenic process
Licensed Content Author	Gabriel Ichim et al
Licensed Content Date	Jul 1, 2016
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations

Number of figures/tables/illustrations	1
Would you like a high resolution image with your order?	no
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title of new work	Characterizing the Genome Targeting Mechanisms of Caspase Activated DNase (CAD) during Cell Differentiation
Institution name	University of Ottawa
Expected presentation date	Nov 2024
Portions	Figure 1
The Requesting Person / Organization to Appear on the License	Dalal Alsowaida
Requestor Location	Mrs. Dalal Alsowaida [redacted] [redacted] [redacted] Attn: Dalal Alsowaida

Billing Type	Invoice
Billing Address	Mrs. Dalal Alsowaida [redacted] [redacted] Attn: Dalal Alsowaida
Total	0.00 USD

Terms and Conditions

Springer Nature Customer Service Centre GmbH Terms and Conditions

The following terms and conditions ("Terms and Conditions") together with the terms specified in your [RightsLink] constitute the License ("License") between you as Licensee and Springer Nature Customer Service Centre GmbH as Licensor. By clicking 'accept' and completing the transaction for your use of the material ("Licensed Material"), you confirm your acceptance of and obligation to be bound by these Terms and Conditions.

1. Grant and Scope of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, non-sublicensable, revocable, world-wide License to reproduce, distribute, communicate to the public, make available, broadcast, electronically transmit or create derivative works using the Licensed Material for the purpose(s) specified in your RightsLink Licence Details only. Licenses are granted for the specific use requested in the order and for no other use, subject to these Terms and Conditions. You acknowledge and agree that the rights granted to you under this License do not include the right to modify, edit, translate, include in collective works, or create derivative works of the Licensed Material in whole or in part unless expressly stated in your RightsLink Licence Details. You may use the Licensed Material only as permitted under this Agreement and will not reproduce, distribute, display, perform, or otherwise use or exploit any Licensed Material in any way, in whole or in part, except as expressly permitted by this License.

1. 2. You may only use the Licensed Content in the manner and to the extent permitted by these Terms and Conditions, by your RightsLink Licence Details and by any applicable laws.

1. 3. A separate license may be required for any additional use of the Licensed Material, e.g. where a license has been purchased for print use only, separate permission must be obtained for electronic re-use. Similarly, a License is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the License.

1. 4. Any content within the Licensed Material that is owned by third parties is expressly excluded from the License.

1. 5. Rights for additional reuses such as custom editions, computer/mobile applications, film or TV reuses and/or any other derivative rights requests require additional permission and may be subject to an additional fee. Please apply to journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

2. Reservation of Rights

Licensor reserves all rights not expressly granted to you under this License. You acknowledge and agree that nothing in this License limits or restricts Licensor's rights in or use of the Licensed Material in any way. Neither this License, nor any act, omission, or statement by Licensor or you, conveys any ownership right to you in any Licensed Material, or to any element or portion thereof. As between Licensor and you, Licensor owns and retains all right, title, and interest in and to the Licensed Material subject to the license granted in Section 1.1. Your permission to use the Licensed Material is expressly conditioned on you not impairing Licensor's or the applicable copyright owner's rights in the Licensed Material in any way.

3. Restrictions on use

3. 1. Minor editing privileges are allowed for adaptations for stylistic purposes or formatting purposes provided such alterations do not alter the original meaning or intention of the Licensed Material and the new figure(s) are still accurate and representative of the Licensed Material. Any other changes including but not limited to, cropping, adapting, and/or omitting material that affect the meaning, intention or moral rights of the author(s) are strictly prohibited.

3. 2. You must not use any Licensed Material as part of any design or trademark.

3. 3. Licensed Material may be used in Open Access Publications (OAP), but any such reuse must include a clear acknowledgment of this permission visible at the same time as the figures/tables/illustration or abstract and which must indicate that the Licensed Material is not part of the governing OA license but has been reproduced with permission. This may be indicated according to any standard referencing system but must include at a minimum 'Book/Journal title, Author, Journal Name (if applicable), Volume (if applicable), Publisher, Year, reproduced with permission from SNCSC'.

4. STM Permission Guidelines

4. 1. An alternative scope of license may apply to signatories of the STM Permissions Guidelines ("STM PG") as amended from time to time and made available at https://www.stm-assoc.org/intellectual-property/permissions/permissions-guidelines/.

4. 2. For content reuse requests that qualify for permission under the STM PG, and which may be updated from time to time, the STM PG supersede the terms and conditions contained in this License.

4. 3. If a License has been granted under the STM PG, but the STM PG no longer apply at the time of publication, further permission must be sought from the Rightsholder. Contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

5. Duration of License

5. 1. Unless otherwise indicated on your License, a License is valid from the date of purchase ("License Date") until the end of the relevant period in the below table:

Reuse in a medical communications project	Reuse up to distribution or time period indicated in License
Reuse in a dissertation/thesis	Lifetime of thesis
Reuse in a journal/magazine	Lifetime of journal/magazine
Reuse in a book/textbook	Lifetime of edition
Reuse on a website	1 year unless otherwise specified in the License
Reuse in a presentation/slide kit/poster	Lifetime of presentation/slide kit/poster. Note: publication whether electronic or in print of presentation/slide kit/poster may require further permission.
Reuse in conference proceedings	Lifetime of conference proceedings
Reuse in an annual report	Lifetime of annual report
Reuse in training/CME materials	Reuse up to distribution or time period indicated in License
Reuse in newsmedia	Lifetime of newsmedia
Reuse in	

coursepack/classroom materials	Reuse up to distribution and/or time period indicated in license

6. Acknowledgement

6. 1. The Licensor's permission must be acknowledged next to the Licensed Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract and must be hyperlinked to the journal/book's homepage.

6. 2. Acknowledgement may be provided according to any standard referencing system and at a minimum should include "Author, Article/Book Title, Journal name/Book imprint, volume, page number, year, Springer Nature".

7. Reuse in a dissertation or thesis

7. 1. Where 'reuse in a dissertation/thesis' has been selected, the following terms apply: Print rights of the Version of Record are provided for; electronic rights for use only on institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/) and only up to what is required by the awarding institution.

7. 2. For theses published under an ISBN or ISSN, separate permission is required. Please contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

7. 3. Authors must properly cite the published manuscript in their thesis according to current citation standards and include the following acknowledgement: '*Reproduced with permission from Springer Nature*'.

8. License Fee

You must pay the fee set forth in the License Agreement (the "License Fees"). All amounts payable by you under this License are exclusive of any sales, use, withholding, value added or similar taxes, government fees or levies or other assessments. Collection and/or remittance of such taxes to the relevant tax authority shall be the responsibility of the party who has the legal obligation to do so.

9. Warranty

9. 1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. **You are solely responsible for ensuring that the material you wish to license is original to the Licensor and does not carry the copyright of another entity or third party (as credited in the published version).** If the credit line on any part of the Licensed Material indicates that it was reprinted or adapted with permission from another source, then you

should seek additional permission from that source to reuse the material.

9. 2. EXCEPT FOR THE EXPRESS WARRANTY STATED HEREIN AND TO THE EXTENT PERMITTED BY APPLICABLE LAW, LICENSOR PROVIDES THE LICENSED MATERIAL "AS IS" AND MAKES NO OTHER REPRESENTATION OR WARRANTY. LICENSOR EXPRESSLY DISCLAIMS ANY LIABILITY FOR ANY CLAIM ARISING FROM OR OUT OF THE CONTENT, INCLUDING BUT NOT LIMITED TO ANY ERRORS, INACCURACIES, OMISSIONS, OR DEFECTS CONTAINED THEREIN, AND ANY IMPLIED OR EXPRESS WARRANTY AS TO MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, PUNITIVE, OR EXEMPLARY DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE LICENSED MATERIAL REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION APPLIES NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

10. Termination and Cancellation

10. 1. The License and all rights granted hereunder will continue until the end of the applicable period shown in Clause 5.1 above. Thereafter, this license will be terminated and all rights granted hereunder will cease.

10. 2. Licensor reserves the right to terminate the License in the event that payment is not received in full or if you breach the terms of this License.

11. General

11. 1. The License and the rights and obligations of the parties hereto shall be construed, interpreted and determined in accordance with the laws of the Federal Republic of Germany without reference to the stipulations of the CISG (United Nations Convention on Contracts for the International Sale of Goods) or to Germany´s choice-of-law principle.

11. 2. The parties acknowledge and agree that any controversies and disputes arising out of this License shall be decided exclusively by the courts of or having jurisdiction for Heidelberg, Germany, as far as legally permissible.

11. 3. This License is solely for Licensor's and Licensee's benefit. It is not for the

benefit of any other person or entity.

Questions? For questions on Copyright Clearance Center accounts or website issues please contact springernaturesupport@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777. For questions on Springer Nature licensing please visit https://www.springernature.com/gp/partners/rights-permissions-third-party-distribution

Other Conditions:

Version 1.4 - Dec 2022

Questions? customercare@copyright.com.

References

Abdul-Ghani M, Megeney LA. Rehabilitation of a contract killer: caspase-3 directs stem cell differentiation. Cell Stem Cell. 2008 Jun 5;2(6):515-6. doi: 10.1016/j.stem.2008.05.013. PMID: 18522841.

Abrams JM. An emerging blueprint for apoptosis in Drosophila. Trends Cell Biol. 1999 Nov;9(11):435-40. doi: 10.1016/s0962-8924(99)01646-3. PMID: 10511707.

Alaggio R, Amador C, Anagnostopoulos I, Attygalle AD, Araujo IBO, Berti E, Bhagat G, Borges AM, Boyer D, Calaminici M, Chadburn A, Chan JKC, Cheuk W, Chng WJ, Choi JK, Chuang SS, Coupland SE, Czader M, Dave SS, de Jong D, Du MQ, Elenitoba-Johnson KS, Ferry J, Geyer J, Gratzinger D, Guitart J, Gujral S, Harris M, Harrison CJ, Hartmann S, Hochhaus A, Jansen PM, Karube K, Kempf W, Khoury J, Kimura H, Klapper W, Kovach AE, Kumar S, Lazar AJ, Lazzi S, Leoncini L, Leung N, Leventaki V, Li XQ, Lim MS, Liu WP, Louissaint A Jr, Marcogliese A, Medeiros LJ, Michal M, Miranda RN, Mitteldorf C, Montes-Moreno S, Morice W, Nardi V, Naresh KN, Natkunam Y, Ng SB, Oschlies I, Ott G, Parrens M, Pulitzer M, Rajkumar SV, Rawstron AC, Rech K, Rosenwald A, Said J, Sarkozy C, Sayed S, Saygin C, Schuh A, Sewell W, Siebert R, Sohani AR, Tooze R, Traverse-Glehen A, Vega F, Vergier B, Wechalekar AD, Wood B, Xerri L, Xiao W. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. Leukemia. 2022 Jul;36(7):1720-1748. doi: 10.1038/s41375-022-01620-2. Epub 2022 Jun 22. Erratum in: Leukemia. 2023 Sep;37(9):1944-1951. doi: 10.1038/s41375-023-01962-5. PMID: 35732829; PMCID: PMC9214472.

Al-Khalaf MH, Blake LE, Larsen BD, Bell RA, Brunette S, Parks RJ, Rudnicki MA, McKinnon PJ, Jeffrey Dilworth F, Megeney LA. Temporal activation of XRCC1-mediated DNA repair is essential for muscle differentiation. Cell Discov. 2016 Jan 12;2:15041. doi: 10.1038/celldisc.2015.41. PMID: 27462438; PMCID: PMC4860966.

Allan JM, Travis LB. Mechanisms of therapy-related carcinogenesis. Nat Rev Cancer. 2005 Dec;5(12):943-55. doi: 10.1038/nrc1749. PMID: 16294218.

Aplan PD. Causes of oncogenic chromosomal translocation. Trends Genet. 2006 Jan;22(1):46-55. doi: 10.1016/j.tig.2005.10.002. Epub 2005 Oct 28. PMID: 16257470; PMCID: PMC1762911.

Aram L, Yacobi-Sharon K, Arama E. CDPs: caspase-dependent non-lethal cellular processes. Cell Death Differ. 2017 Aug;24(8):1307-1310. doi: 10.1038/cdd.2017.111. PMID: 28695898; PMCID: PMC5520448.

Arama E, Agapite J, Steller H. Caspase activity and a specific cytochrome C are required for sperm differentiation in Drosophila. Dev Cell. 2003 May;4(5):687-97. doi: 10.1016/s1534-5807(03)00120-5. PMID: 12737804.
Baron MH, Barminko J. Chromatin Condensation and Enucleation in Red Blood Cells: An Open Question. Dev Cell. 2016 Mar 7;36(5):481-2. doi: 10.1016/j.devcel.2016.02.014. PMID: 26954541.

Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: current advances and future directions. Int J Med Sci. 2012;9(3):193-9. doi: 10.7150/ijms.3635. Epub 2012 Feb 27. PMID: 22408567; PMCID: PMC3298009.
Bass BP, Tanner EA, Mateos San Martín D, Blute T, Kinser RD, Dolph PJ, McCall K. Cell-autonomous requirement for DNaseII in nonapoptotic cell death. Cell Death Differ. 2009 Oct;16(10):1362-71. doi: 10.1038/cdd.2009.79. Epub 2009 Jun 26. PMID: 19557011; PMCID: PMC2770252.

Bataller A, Guijarro F, Caye-Eude A, Strullu M, Sterin A, Molina O, Chevallier P, Zaliova M, Zuna J, Mozas P, Magnano L, Grardel N, Cornillet-Lefebvre P, Fu JF, Shih LY, Boneva T, Nacheva E, Beà S, López-Guerra M, Bueno C, Menéndez P, Esteve J, Larghero P, Meyer C, Marschalek R. KMT2A-CBL rearrangements in acute leukemias: clinical characteristics and genetic breakpoints. Blood Adv. 2021 Dec 28;5(24):5617-5620. doi: 10.1182/bloodadvances.2021005736. PMID: 34638131; PMCID: PMC8714718.

Bell RAV, Al-Khalaf MH, Brunette S, Alsowaida D, Chu A, Bandukwala H, Dechant G, Apostolova G, Dilworth FJ, Megeney LA. Chromatin Reorganization during Myoblast Differentiation Involves the Caspase-Dependent Removal of SATB2. Cells. 2022 Mar 11;11(6):966. doi: 10.3390/cells11060966. PMID: 35326417; PMCID: PMC8946544.

Benada J, Alsowaida D, Megeney LA, Sørensen CS. Self-inflicted DNA breaks in cell differentiation and cancer. Trends Cell Biol. 2023 Oct;33(10):850-859. doi: 10.1016/j.tcb.2023.03.002. Epub 2023 Mar 28. PMID: 36997393.

Berkes CA, Tapscott SJ. MyoD and the transcriptional control of myogenesis. Semin Cell Dev Biol. 2005 Aug-Oct;16(4-5):585-95. doi: 10.1016/j.semcdb.2005.07.006. PMID: 16099183.

Berthenet K, Castillo Ferrer C, Fanfone D, Popgeorgiev N, Neves D, Bertolino P, Gibert B, Hernandez-Vargas H, Ichim G. Failed Apoptosis Enhances Melanoma Cancer Cell Aggressiveness. Cell Rep. 2020 Jun 9;31(10):107731. doi: 10.1016/j.celrep.2020.107731. PMID: 32521256.

Boehm D, Mazurier C, Giarratana MC, Darghouth D, Faussat AM, Harmand L, Douay L. Caspase-3 is involved in the signalling in erythroid differentiation by targeting late progenitors. PLoS One. 2013 May 2;8(5):e62303. doi: 10.1371/journal.pone.0062303. PMID: 23658722; PMCID: PMC3642196.

Boscolo Sesillo F, Fox D, Sacco A. Muscle Stem Cells Give Rise to Rhabdomyosarcomas in a Severe Mouse Model of Duchenne Muscular Dystrophy. Cell Rep. 2019 Jan 15;26(3):689-701.e6. doi: 10.1016/j.celrep.2018.12.089. PMID: 30650360; PMCID: PMC6853774.

Braun T, Gautel M. Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. Nat Rev Mol Cell Biol. 2011 Jun;12(6):349-61. doi: 10.1038/nrm3118. PMID: 21602905.

Brosh RM Jr. DNA helicases involved in DNA repair and their roles in cancer. Nat Rev Cancer. 2013 Aug;13(8):542-58. doi: 10.1038/nrc3560. Epub 2013 Jul 11. PMID: 23842644; PMCID: PMC4538698.

Brunette S, Sharma A, Bell R, Puente L, Megeney LA. Caspase 3 exhibits a yeast metacaspase proteostasis function that protects mitochondria from toxic TDP43 aggregates. Microb Cell. 2023 Jul 10;10(8):157-169. doi: 10.15698/mic2023.08.801. PMID: 37545643; PMCID: PMC10399456.

Bunting SF, Nussenzweig A. End-joining, translocations and cancer. Nat Rev Cancer. 2013 Jul;13(7):443-54. doi: 10.1038/nrc3537. Epub 2013 Jun 13. PMID: 23760025; PMCID: PMC5724777.

Burgon PG, Megeney LA. Caspase signaling, a conserved inductive cue for metazoan cell differentiation. Semin Cell Dev Biol. 2018 Oct;82:96-104. doi: 10.1016/j.semcdb.2017.11.009. Epub 2017 Nov 11. PMID: 29129746.
Canela A, Maman Y, Huang SN, Wutz G, Tang W, Zagnoli-Vieira G, Callen E, Wong N, Day A, Peters JM, Caldecott KW, Pommier Y, Nussenzweig A. Topoisomerase II-Induced Chromosome Breakage and Translocation Is Determined by Chromosome Architecture and Transcriptional Activity. Mol Cell. 2019 Jul 25;75(2):252-266.e8. doi: 10.1016/j.molcel.2019.04.030. Epub 2019 Jun 12. PMID: 31202577; PMCID: PMC8170508.

Cao K, Riley JS, Heilig R, Montes-Gómez AE, Vringer E, Berthenet K, Cloix C, Elmasry Y, Spiller DG, Ichim G, Campbell KJ, Gilmore AP, Tait SWG. Mitochondrial dynamics regulate genome stability via control of caspase-dependent DNA damage. Dev Cell. 2022 May 23;57(10):1211-1225.e6. doi: 10.1016/j.devcel.2022.03.019. Epub 2022 Apr 20. PMID: 35447090; PMCID: PMC9616799.

Cao R, Zhang Y. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol Cell. 2004 Jul 2;15(1):57-67. doi: 10.1016/j.molcel.2004.06.020. PMID: 15225548.
Caretti G, Di Padova M, Micales B, Lyons GE, Sartorelli V. The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. Genes Dev. 2004 Nov 1;18(21):2627-38. doi: 10.1101/gad.1241904. Erratum in: Genes Dev. 2005 Mar 15;19(6):768. PMID: 15520282; PMCID: PMC525543.

Carlile GW, Smith DH, Wiedmann M. Caspase-3 has a nonapoptotic function in erythroid maturation. Blood. 2004 Jun 1;103(11):4310-6. doi: 10.1182/blood-2003-09-3362. Epub 2004 Feb 19. PMID: 14976035.

Carlile GW, Smith DH, Wiedmann M. Caspase-3 has a nonapoptotic function in erythroid maturation. Blood. 2004 Jun 1;103(11):4310-6. doi: 10.1182/blood-2003-09-3362. Epub 2004 Feb 19. PMID: 14976035.

Cartwright IM, Liu X, Zhou M, Li F, Li CY. Essential roles of Caspase-3 in facilitating Myc-induced genetic instability and carcinogenesis. Elife. 2017 Jul 10;6:e26371. doi: 10.7554/eLife.26371. PMID: 28691902; PMCID: PMC5550274.

Cathelin S, Rébé C, Haddaoui L, Simioni N, Verdier F, Fontenay M, Launay S, Mayeux P, Solary E. Identification of proteins cleaved downstream of caspase activation in monocytes undergoing macrophage differentiation. J Biol Chem. 2006 Jun 30;281(26):17779-88. doi: 10.1074/jbc.M600537200. Epub 2006 Apr 24. PMID: 16636047.

Chal J, Pourquié O. Making muscle: skeletal myogenesis in vivo and in vitro. Development. 2017 Jun 15;144(12):2104-2122. doi: 10.1242/dev.151035. PMID: 28634270.

Chen B, Jiang L, Zhong ML, Li JF, Li BS, Peng LJ, Dai YT, Cui BW, Yan TQ, Zhang WN, Weng XQ, Xie YY, Lu J, Ren RB, Chen SN, Hu JD, Wu DP, Chen Z, Tang JY, Huang JY, Mi JQ, Chen SJ. Identification of fusion genes and characterization of transcriptome features in T-cell acute lymphoblastic leukemia. Proc Natl Acad Sci U S A. 2018 Jan 9;115(2):373-378. doi: 10.1073/pnas.1717125115. Epub 2017 Dec 26. Erratum in: Proc Natl Acad Sci U S A. 2020 Sep 15;117(37):23192. doi: 10.1073/pnas.2017617117. PMID: 29279377; PMCID: PMC5777070.

Chin L, Schreiber-Agus N, Pellicer I, Chen K, Lee HW, Dudast M, Cordon-Cardo C, DePinho RA. Contrasting roles for Myc and Mad proteins in cellular growth and differentiation. Proc Natl Acad Sci U S A. 1995 Aug 29;92(18):8488-92. doi: 10.1073/pnas.92.18.8488. PMID: 7667316; PMCID: PMC41182.

Chou DM, Adamson B, Dephoure NE, Tan X, Nottke AC, Hurov KE, Gygi SP, Colaiácovo MP, Elledge SJ. A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage. Proc Natl Acad Sci U S A. 2010 Oct 26;107(43):18475-80. doi: 10.1073/pnas.1012946107. Epub 2010 Oct 11. PMID: 20937877; PMCID: PMC2972950.

Cieślar-Pobuda A, Saenko Y, Rzeszowska-Wolny J. PARP-1 inhibition induces a late increase in the level of reactive oxygen species in cells after ionizing radiation. Mutat Res. 2012 Apr 1;732(1-2):9-15. doi: 10.1016/j.mrfmmm.2012.01.005. Epub 2012 Feb 2. PMID: 22321899.

Colon-Plaza S, Su TT. Non-Apoptotic Role of Apoptotic Caspases in the Drosophila Nervous System. Front Cell Dev Biol. 2022 Feb 9;10:839358. doi: 10.3389/fcell.2022.839358. PMID: 35223857; PMCID: PMC8863954.

Connolly PF, Fearnhead HO. DNA-PK activity is associated with caspase-dependent myogenic differentiation. FEBS J. 2016 Oct;283(19):3626-3636. doi: 10.1111/febs.13832. PMID: 27513301.

Consalvi S, Brancaccio A, Dall'Agnese A, Puri PL, Palacios D. Praja1 E3 ubiquitin ligase promotes skeletal myogenesis through degradation of EZH2 upon p38α activation. Nat Commun. 2017 Jan 9;8:13956. doi: 10.1038/ncomms13956. PMID: 28067271; PMCID: PMC5423270.

Cooper S, Grijzenhout A, Underwood E, Ancelin K, Zhang T, Nesterova TB, Anil-Kirmizitas B, Bassett A, Kooistra SM, Agger K, Helin K, Heard E, Brockdorff N. Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. Nat Commun. 2016 Nov 28;7:13661. doi: 10.1038/ncomms13661. PMID: 27892467; PMCID: PMC5133711.

Cowell IG, Austin CA. DNA fragility at the KMT2A/MLLlocus: insights from old and new technologies. Open Biol. 2023 Jan;13(1):220232. doi: 10.1098/rsob.220232. Epub 2023 Jan 11. PMID: 36629017; PMCID: PMC9832561.

Crawford ED, Wells JA. Caspase substrates and cellular remodeling. Annu Rev Biochem. 2011;80:1055-87. doi: 10.1146/annurev-biochem-061809-121639. PMID: 21456965.

Crick F. The double helix: a personal view. Nature. 1974 Apr 26;248(5451):766-9. doi: 10.1038/248766a0. PMID: 4599081.

Crick, F. The double helix: a personal view. Nature 248, 766–769 (1974).
Crose, L.E. et al. Alveolar rhabdomyosarcoma-associated PAX3-FOXO1 promotes tumorigenesis via Hippo pathway suppression. J Clin Invest., 124, 285-296 (2014).

Crowe SL, Movsesyan VA, Jorgensen TJ, Kondratyev A. Rapid phosphorylation of histone H2A.X following ionotropic glutamate receptor activation. Eur J Neurosci. 2006 May;23(9):2351-61. doi: 10.1111/j.1460-9568.2006.04768.x. PMID: 16706843; PMCID: PMC1534119.

d'Adda di Fagagna F. Living on a break: cellular senescence as a DNA-damage response. Nat Rev Cancer. 2008 Jul;8(7):512-22. doi: 10.1038/nrc2440. PMID: 18574463.

Dadi S, Le Noir S, Payet-Bornet D, Lhermitte L, Zacarias-Cabeza J, Bergeron J, Villarèse P, Vachez E, Dik WA, Millien C, Radford I, Verhoeyen E, Cosset FL, Petit A, Ifrah N, Dombret H, Hermine O, Spicuglia S, Langerak AW, Macintyre EA, Nadel B, Ferrier P, Asnafi V. TLX homeodomain oncogenes mediate T cell maturation arrest in T-ALL via interaction with ETS1 and suppression of TCRα gene expression. Cancer Cell. 2012 Apr 17;21(4):563-76. doi: 10.1016/j.ccr.2012.02.013. PMID: 22516263.

Daley SR, Teh C, Hu DY, Strasser A, Gray DHD. Cell death and thymic tolerance. Immunol Rev. 2017 May;277(1):9-20. doi: 10.1111/imr.12532. PMID: 28462532.

Dalla-Favera R, Martinotti S, Gallo RC, Erikson J, Croce CM. Translocation and rearrangements of the c-myc oncogene locus in human undifferentiated B-cell lymphomas. Science. 1983 Feb 25;219(4587):963-7. doi: 10.1126/science.6401867. PMID: 6401867.

D'Amelio M, Cavallucci V, Cecconi F. Neuronal caspase-3 signaling: not only cell death. Cell Death Differ. 2010 Jul;17(7):1104-14. doi: 10.1038/cdd.2009.180. Epub 2009 Dec 4. PMID: 19960023.

De Botton S, Sabri S, Daugas E, Zermati Y, Guidotti JE, Hermine O, Kroemer G, Vainchenker W, Debili N. Platelet formation is the consequence of caspase activation within megakaryocytes. Blood. 2002 Aug 15;100(4):1310-7. doi: 10.1182/blood-2002-03-0686. PMID: 12149212.

De Maria R, Zeuner A, Eramo A, Domenichelli C, Bonci D, Grignani F, Srinivasula SM, Alnemri ES, Testa U, Peschle C. Negative regulation of erythropoiesis by caspase-mediated cleavage of GATA-1. Nature. 1999 Sep 30;401(6752):489-93. doi: 10.1038/46809. PMID: 10519553.

de Matos RRC, Ferreira GM, Meyer C, Marschalek R, Larghero P, Ribeiro RC, Liehr T, Othman M, Bizarro MTSM, Sobral da Costa E, Land MGP, Abdelhay E, Binato R, Silva MLM. KMT2A-MLLT1 and the Novel SEC16A-KMT2A in a Cryptic 3-Way Translocation t(9;11;19) Present in an Infant With Acute Lymphoblastic Leukemia. J Pediatr Hematol Oncol. 2022 Apr 1;44(3):e719-e722. doi: 10.1097/MPH.0000000000002386. PMID: 34966090.

Deng K, Liu Z, Su Y, Zhang Z, Fan Y, Zhang Y, Wang F. RUNX1T1 modulates myogenic differentiation by regulating the calcium signaling pathway and the alternative splicing of ROCK2. FASEB J. 2023 Jul;37(7):e23044. doi: 10.1096/fj.202300677R. PMID: 37342905.

Dick SA, Chang NC, Dumont NA, Bell RA, Putinski C, Kawabe Y, Litchfield DW, Rudnicki MA, Megeney LA. Caspase 3 cleavage of Pax7 inhibits self-renewal of satellite cells. Proc Natl Acad Sci U S A. 2015 Sep 22;112(38):E5246-52. doi: 10.1073/pnas.1512869112. Epub 2015 Sep 8. PMID: 26372956; PMCID: PMC4586827.

Dick SA, Megeney LA. Cell death proteins: an evolutionary role in cellular adaptation before the advent of apoptosis. Bioessays. 2013 Nov;35(11):974-83. doi: 10.1002/bies.201300052. Epub 2013 Aug 14. PMID: 23943356.

Dudakovic A, Camilleri ET, Xu F, Riester SM, McGee-Lawrence ME, Bradley EW, Paradise CR, Lewallen EA, Thaler R, Deyle DR, Larson AN, Lewallen DG, Dietz AB, Stein GS, Montecino MA, Westendorf JJ, van Wijnen AJ. Epigenetic Control of Skeletal Development by the Histone Methyltransferase Ezh2. J Biol Chem. 2015 Nov 13;290(46):27604-17. doi: 10.1074/jbc.M115.672345. Epub 2015 Sep 30. PMID: 26424790; PMCID: PMC4646012.

Dueva R, Iliakis G. Replication protein A: a multifunctional protein with roles in DNA replication, repair and beyond. NAR Cancer. 2020 Sep 25;2(3):zcaa022. doi: 10.1093/narcan/zcaa022. PMID: 34316690; PMCID: PMC8210275.

Dürnberger H, Kratochwil K. Specificity of tissue interaction and origin of mesenchymal cells in the androgen response of the embryonic mammary gland. Cell. 1980 Feb;19(2):465-71. doi: 10.1016/0092-8674(80)90521-8. PMID: 7357614.

Ellis HM, Horvitz HR. Genetic control of programmed cell death in the nematode C. elegans. Cell. 1986 Mar 28;44(6):817-29. doi: 10.1016/0092-8674(86)90004-8. PMID: 3955651.

Enari M, Sakahira H, Yokoyama H, Okawa K, Iwamatsu A, Nagata S. A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. Nature. 1998 Jan 1;391(6662):43-50. doi: 10.1038/34112. Erratum in: Nature 1998 May 28;393(6683):396. PMID: 9422506.

Feng Z, Tang ZL, Li K, Liu B, Yu M, Zhao SH. Molecular characterization of the BTG2 and BTG3 genes in fetal muscle development of pigs. Gene. 2007 Nov 15;403(1-2):170-7. doi: 10.1016/j.gene.2007.08.009. Epub 2007 Aug 25. PMID: 17890019.

Fernandes-Alnemri T, Litwack G, Alnemri ES. CPP32, a novel human apoptotic protein with homology to Caenorhabditis elegans cell death protein Ced-3 and mammalian interleukin-1 beta-converting enzyme. J Biol Chem. 1994 Dec 9;269(49):30761-4. PMID: 7983002.

Fernando P, Brunette S, Megeney LA. Neural stem cell differentiation is dependent upon endogenous caspase 3 activity. FASEB J. 2005 Oct;19(12):1671-3. doi: 10.1096/fj.04-2981fje. Epub 2005 Aug 15. PMID: 16103108.

Fernando P, Kelly JF, Balazsi K, Slack RS, Megeney LA. Caspase 3 activity is required for skeletal muscle differentiation. Proc Natl Acad Sci U S A. 2002 Aug 20;99(17):11025-30. doi: 10.1073/pnas.162172899. Epub 2002 Aug 12. PMID: 12177420; PMCID: PMC123204.

Fernando P, Megeney LA. Is caspase-dependent apoptosis only cell differentiation taken to the extreme? FASEB J. 2007 Jan;21(1):8-17. doi: 10.1096/fj.06-5912hyp. Epub 2006 Nov 8. PMID: 17093139.

Frank D, Vince JE. Pyroptosis versus necroptosis: similarities, differences, and crosstalk. Cell Death Differ. 2019 Jan;26(1):99-114. doi: 10.1038/s41418-018-0212-6. Epub 2018 Oct 19. PMID: 30341423; PMCID: PMC6294779.
Fujita J, Crane AM, Souza MK, Dejosez M, Kyba M, Flavell RA, Thomson JA, Zwaka TP. Caspase activity mediates the differentiation of embryonic stem cells. Cell Stem Cell. 2008 Jun 5;2(6):595-601. doi: 10.1016/j.stem.2008.04.001. PMID: 18522852; PMCID: PMC2494585.

Gandarillas A, Watt FM. c-Myc promotes differentiation of human epidermal stem cells. Genes Dev. 1997 Nov 1;11(21):2869-82. doi: 10.1101/gad.11.21.2869. PMID: 9353256; PMCID: PMC316650.

Ge Y, Chen J. MicroRNAs in skeletal myogenesis. Cell Cycle. 2011 Feb 1;10(3):441-8. doi: 10.4161/cc.10.3.14710. Epub 2011 Feb 1. PMID: 21270519; PMCID: PMC3115018.
Gerschenson LE, Geske FJ. Virchow and apoptosis. Am J Pathol. 2001 Apr;158(4):1543. doi: 10.1016/S0002-9440(10)64105-3. PMID: 11290572; PMCID: PMC1891904.

Gheyas R, Menko AS. The involvement of caspases in the process of nuclear removal during lens fiber cell differentiation. Cell Death Discov. 2023 Oct 21;9(1):386. doi: 10.1038/s41420-023-01680-y. PMID: 37865680; PMCID: PMC10590423.

GLUCKSMANN A. Cell deaths in normal vertebrate ontogeny. Biol Rev Camb Philos Soc. 1951 Feb;26(1):59-86. doi: 10.1111/j.1469-185x.1951.tb00774.x. PMID: 24540363

Green DR, Droin N, Pinkoski M. Activation-induced cell death in T cells. Immunol Rev. 2003 Jun;193:70-81. doi: 10.1034/j.1600-065x.2003.00051.x. PMID: 12752672.

Green DR. Caspases and Their Substrates. Cold Spring Harb Perspect Biol. 2022 Mar 1;14(3):a041012. doi: 10.1101/cshperspect.a041012. PMID: 35232877; PMCID: PMC8886984. (B)

Green DR. Inflammasomes and Other Caspase-Activation Platforms. Cold Spring Harb Perspect Biol. 2022 Jul 1;14(7):a041061. doi: 10.1101/cshperspect.a041061. PMID: 35777830; PMCID: PMC9248825. (A)

Guo P, Lim RC, Rajawasam K, Trinh T, Sun H, Zhang H. A methylation-phosphorylation switch controls EZH2 stability and hematopoiesis. Elife. 2024 Feb 12;13:e86168. doi: 10.7554/eLife.86168. PMID: 38346162; PMCID: PMC10901513.

Haffner MC, Aryee MJ, Toubaji A, Esopi DM, Albadine R, Gurel B, Isaacs WB, Bova GS, Liu W, Xu J, Meeker AK, Netto G, De Marzo AM, Nelson WG, Yegnasubramanian S. Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. Nat Genet. 2010 Aug;42(8):668-75. doi: 10.1038/ng.613. Epub 2010 Jul 4. PMID: 20601956; PMCID: PMC3157086.

Haimovici A, Höfer C, Badr MT, Bavafaye Haghighi E, Amer T, Boerries M, Bronsert P, Glavynskyi I, Fanfone D, Ichim G, Thilmany N, Weber A, Brummer T, Spohr C, Öllinger R, Janssen KP, Rad R, Häcker G. Spontaneous activity of the mitochondrial apoptosis pathway drives chromosomal defects, the appearance of micronuclei and cancer metastasis through the Caspase-Activated DNAse. Cell Death Dis. 2022 Apr 7;13(4):315. doi: 10.1038/s41419-022-04768-y. PMID: 35393399; PMCID: PMC8990075.

Haimovici A, Rupp V, Amer T, Moeed A, Weber A, Häcker G. The caspase-activated DNase promotes cellular senescence. EMBO J. 2024 Jul 8. doi: 10.1038/s44318-024-00163-9. Epub ahead of print. PMID: 38977850
Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K. A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol. 2008 Nov;10(11):1291-300. doi: 10.1038/ncb1787. Epub 2008 Oct 19. Erratum in: Nat Cell Biol. 2008 Dec;10(12):1484. PMID: 18931660.

Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K. A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol. 2008 Nov;10(11):1291-300. doi: 10.1038/ncb1787. Epub 2008 Oct 19. Erratum in: Nat Cell Biol. 2008 Dec;10(12):1484. PMID: 18931660.

Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K. A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol. 2008 Nov;10(11):1291-300. doi: 10.1038/ncb1787. Epub 2008 Oct 19. Erratum in: Nat Cell Biol. 2008 Dec;10(12):1484. PMID: 18931660.

He S, Huang Z, Liu Y, Ha T, Wu B. DNA break induces rapid transcription repression mediated by proteasome-dependent RNAPII removal. Cell Rep. 2024 Jul 1;43(7):114420. doi: 10.1016/j.celrep.2024.114420. Epub ahead of print. PMID: 38954517.

Hernández-Martínez R, Covarrubias L. Interdigital cell death function and regulation: new insights on an old programmed cell death model. Dev Growth Differ. 2011 Feb;53(2):245-58. doi: 10.1111/j.1440-169X.2010.01246.x. PMID: 21338350.

Holoch D, Margueron R. Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity. Trends Biochem Sci. 2017 Jul;42(7):531-542. doi: 10.1016/j.tibs.2017.04.003. Epub 2017 May 5. PMID: 28483375.
Hu Q, Wu D, Chen W, Yan Z, Shi Y. Proteolytic processing of the caspase-9 zymogen is required for apoptosome-mediated activation of caspase-9. J Biol Chem. 2013 May 24;288(21):15142-7. doi: 10.1074/jbc.M112.441568. Epub 2013 Apr 9. PMID: 23572523; PMCID: PMC3663534.

Huang RX, Zhou PK. DNA damage response signaling pathways and targets for radiotherapy sensitization in cancer. Signal Transduct Target Ther. 2020 May 1;5(1):60. doi: 10.1038/s41392-020-0150-x. PMID: 32355263; PMCID: PMC7192953.

Hurlin PJ, Quéva C, Koskinen PJ, Steingrímsson E, Ayer DE, Copeland NG, Jenkins NA, Eisenman RN. Mad3 and Mad4: novel Max-interacting transcriptional repressors that suppress c-myc dependent transformation and are expressed during neural and epidermal differentiation. EMBO J. 1995 Nov 15;14(22):5646-59. doi: 10.1002/j.1460-2075.1995.tb00252.x. Erratum in: EMBO J. 1996 Apr 15;15(8):2030. PMID: 8521822; PMCID: PMC394680.

Ichim G, Lopez J, Ahmed SU, Muthalagu N, Giampazolias E, Delgado ME, Haller M, Riley JS, Mason SM, Athineos D, Parsons MJ, van de Kooij B, Bouchier-Hayes L, Chalmers AJ, Rooswinkel RW, Oberst A, Blyth K, Rehm M, Murphy DJ, Tait SWG. Limited mitochondrial permeabilization causes DNA damage and genomic instability in the absence of cell death. Mol Cell. 2015 Mar 5;57(5):860-872. doi: 10.1016/j.molcel.2015.01.018. Epub 2015 Feb 19. Erratum in: Mol Cell. 2015 Jun 4;58(5):900. PMID: 25702873; PMCID: PMC4352766.

Ishizaki Y, Jacobson MD, Raff MC. A role for caspases in lens fiber differentiation. J Cell Biol. 1998 Jan 12;140(1):153-8. doi: 10.1083/jcb.140.1.153. PMID: 9425163; PMCID: PMC2132591.

Janzen V, Fleming HE, Riedt T, Karlsson G, Riese MJ, Lo Celso C, Reynolds G, Milne CD, Paige CJ, Karlsson S, Woo M, Scadden DT. Hematopoietic stem cell responsiveness to exogenous signals is limited by caspase-3. Cell Stem Cell. 2008 Jun 5;2(6):584-94. doi: 10.1016/j.stem.2008.03.012. PMID: 18522851; PMCID: PMC2991117.

Jenkinson S, Koo K, Mansour MR, Goulden N, Vora A, Mitchell C, Wade R, Richards S, Hancock J, Moorman AV, Linch DC, Gale RE. Impact of NOTCH1/FBXW7 mutations on outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on the MRC UKALL 2003 trial. Leukemia. 2013 Jan;27(1):41-7. doi: 10.1038/leu.2012.176. Epub 2012 Jul 3. PMID: 22814294.

Jordan D, Hindocha S, Dhital M, Saleh M, Khan W. The epidemiology, genetics and future management of syndactyly. Open Orthop J. 2012;6:14-27. doi: 10.2174/1874325001206010014. Epub 2012 Mar 23. PMID: 22448207; PMCID: PMC3308320.

Ju BG, Lunyak VV, Perissi V, Garcia-Bassets I, Rose DW, Glass CK, Rosenfeld MG. A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. Science. 2006 Jun 23;312(5781):1798-802. doi: 10.1126/science.1127196. Erratum in: Science. 2011 May 6;332(6030):664. Erratum in: Science. 2014 Feb 21;343(6173):839. PMID: 16794079.

Juan AH, Kumar RM, Marx JG, Young RA, Sartorelli V. Mir-214-dependent regulation of the polycomb protein Ezh2 in skeletal muscle and embryonic stem cells. Mol Cell. 2009 Oct 9;36(1):61-74. doi: 10.1016/j.molcel.2009.08.008. PMID: 19818710; PMCID: PMC2761245

Justin N, Zhang Y, Tarricone C, Martin SR, Chen S, Underwood E, De Marco V, Haire LF, Walker PA, Reinberg D, Wilson JR, Gamblin SJ. Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. Nat Commun. 2016 Apr 28;7:11316. doi: 10.1038/ncomms11316. PMID: 27121947; PMCID: PMC4853476.

Kaluzhny Y, Yu G, Sun S, Toselli PA, Nieswandt B, Jackson CW, Ravid K. BclxL overexpression in megakaryocytes leads to impaired platelet fragmentation. Blood. 2002 Sep 1;100(5):1670-8. doi: 10.1182/blood-2001-12-0263. PMID: 12176887.

Kanuka H, Kuranaga E, Takemoto K, Hiratou T, Okano H, Miura M. Drosophila caspase transduces Shaggy/GSK-3beta kinase activity in neural precursor development. EMBO J. 2005 Nov 2;24(21):3793-806. doi: 10.1038/sj.emboj.7600822. Epub 2005 Oct 13. PMID: 16222340; PMCID: PMC1276714.

Kanuka H, Kuranaga E, Takemoto K, Hiratou T, Okano H, Miura M. Drosophila caspase transduces Shaggy/GSK-3beta kinase activity in neural precursor development. EMBO J. 2005 Nov 2;24(21):3793-806. doi: 10.1038/sj.emboj.7600822. Epub 2005 Oct 13. PMID: 16222340; PMCID: PMC1276714.

Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019 Apr 29;10(1):1930. doi: 10.1038/s41467-019-09982-5. PMID: 31036827; PMCID: PMC6488672.

Keeney S. Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. Genome Dyn Stab. 2008 Jan 1;2:81-123. doi: 10.1007/7050_2007_026. PMID: 21927624; PMCID: PMC3172816.

Kennedy NJ, Kataoka T, Tschopp J, Budd RC. Caspase activation is required for T cell proliferation. J Exp Med. 1999 Dec 20;190(12):1891-6. doi: 10.1084/jem.190.12.1891. PMID: 10601363; PMCID: PMC2195711.

Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Br J Cancer. 1972 Aug;26(4):239-57. doi: 10.1038/bjc.1972.33. PMID: 4561027; PMCID: PMC2008650.

Keyel PA. DNases in health and disease. Dev Biol. 2017 Sep 1;429(1):1-11. doi: 10.1016/j.ydbio.2017.06.028. Epub 2017 Jun 28. PMID: 28666955.

Kim Y, Kim B, Seong MW, Lee DS, Hong KT, Kang HJ, Yun J, Chang YH. Cryptic KMT2A/MLLT10 fusion detected by next-generation sequencing in a case of pediatric acute megakaryoblastic leukemia. Cancer Genet. 2023 Aug;276-277:36-39. doi: 10.1016/j.cancergen.2023.07.003. Epub 2023 Jul 11. PMID: 37478796.

King B, Ntziachristos P, Aifantis I. Hijacking T cell differentiation: new insights in TLX function in T-ALL. Cancer Cell. 2012 Apr 17;21(4):453-5. doi: 10.1016/j.ccr.2012.03.026. PMID: 22516255; PMCID: PMC3902167.

Kirk R. Pediatric oncology: Fusion-gene status crucial in RMS. Nat Rev Clin Oncol. 2010 Jul;7(7):356. doi: 10.1038/nrclinonc.2010.88. PMID: 20614557.

Konopka A, Atkin JD. The Role of DNA Damage in Neural Plasticity in Physiology and Neurodegeneration. Front Cell Neurosci. 2022 Jun 23;16:836885. doi: 10.3389/fncel.2022.836885. PMID: 35813507; PMCID: PMC9259845.

Konopka A, Atkin JD. The Role of DNA Damage in Neural Plasticity in Physiology and Neurodegeneration. Front Cell Neurosci. 2022 Jun 23;16:836885. doi: 10.3389/fncel.2022.836885. PMID: 35813507; PMCID: PMC9259845.

Krammer PH, Arnold R, Lavrik IN. Life and death in peripheral T cells. Nat Rev Immunol. 2007 Jul;7(7):532-42. doi: 10.1038/nri2115. PMID: 17589543.

Kulbay M, Bernier-Parker N, Bernier J. The role of the DFF40/CAD endonuclease in genomic stability. Apoptosis. 2021 Feb;26(1-2):9-23. doi: 10.1007/s10495-020-01649-7. Epub 2021 Jan 2. PMID: 33387146.

Kumar R, Vadlamudi RK, Adam L. Apoptosis in mammary gland and cancer. Endocr Relat Cancer. 2000 Dec;7(4):257-69. doi: 10.1677/erc.0.0070257. PMID: 11174847.

Larsen BD, Benada J, Yung PYK, Bell RAV, Pappas G, Urban V, Ahlskog JK, Kuo TT, Janscak P, Megeney LA, Elsässer SJ, Bartek J, Sørensen CS. Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress. Science. 2022 Apr 29;376(6592):476-483. doi: 10.1126/science.abi6378. Epub 2022 Apr 28. PMID: 35482866.

Larsen BD, Megeney LA. Parole terms for a killer: directing caspase3/CAD induced DNA strand breaks to coordinate changes in gene expression. Cell Cycle. 2010 Aug 1;9(15):2940-5. doi: 10.4161/cc.9.15.12335. PMID: 20714221; PMCID: PMC3040922.

Larsen BD, Rampalli S, Burns LE, Brunette S, Dilworth FJ, Megeney LA. Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. Proc Natl Acad Sci U S A. 2010 Mar 2;107(9):4230-5. doi: 10.1073/pnas.0913089107. Epub 2010 Feb 16. PMID: 20160104; PMCID: PMC2840077.

Li LY, Luo X, Wang X. Endonuclease G is an apoptotic DNase when released from mitochondria. Nature. 2001 Jul 5;412(6842):95-9. doi: 10.1038/35083620. PMID: 11452314.

Linqing Z, Guohua J, Haoming L, Xuelei T, Jianbing Q, Meiling T. Runx1t1 regulates the neuronal differentiation of radial glial cells from the rat hippocampus. Stem Cells Transl Med. 2015 Jan;4(1):110-6. doi: 10.5966/sctm.2014-0158. Epub 2014 Dec 3. PMID: 25473084; PMCID: PMC4275013.

Liu X, He Y, Li F, Huang Q, Kato TA, Hall RP, Li CY. Caspase-3 promotes genetic instability and carcinogenesis. Mol Cell. 2015 Apr 16;58(2):284-96. doi: 10.1016/j.molcel.2015.03.003. Epub 2015 Apr 9. PMID: 25866249; PMCID: PMC4408780.

Liu X, He Y, Li F, Huang Q, Kato TA, Hall RP, Li CY. Caspase-3 promotes genetic instability and carcinogenesis. Mol Cell. 2015 Apr 16;58(2):284-96. doi: 10.1016/j.molcel.2015.03.003. Epub 2015 Apr 9. PMID: 25866249; PMCID: PMC4408780.

Liu X, Zou H, Slaughter C, Wang X. DFF, a heterodimeric protein that functions downstream of caspase-3 to trigger DNA fragmentation during apoptosis. Cell. 1997 Apr 18;89(2):175-84. doi: 10.1016/s0092-8674(00)80197-x. PMID: 9108473.

Liu Y, Hernandez AM, Shibata D, Cortopassi GA. BCL2 translocation frequency rises with age in humans. Proc Natl Acad Sci U S A. 1994 Sep 13;91(19):8910-4. doi: 10.1073/pnas.91.19.8910. PMID: 8090743; PMCID: PMC44716.

Manceau L, Richard Albert J, Lollini PL, Greenberg MVC, Gilardi-Hebenstreit P, Ribes V. Divergent transcriptional and transforming properties of PAX3-FOXO1 and PAX7-FOXO1 paralogs. PLoS Genet. 2022 May 23;18(5):e1009782. doi: 10.1371/journal.pgen.1009782. PMID: 35604932; PMCID: PMC9166405

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ 3rd, Voigt P, Martin SR, Taylor WR, De Marco V, Pirrotta V, Reinberg D, Gamblin SJ. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature. 2009 Oct 8;461(7265):762-7. doi: 10.1038/nature08398. Epub 2009 Sep 20. PMID: 19767730; PMCID: PMC3772642.

Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, Dynlacht BD, Reinberg D. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol Cell. 2008 Nov 21;32(4):503-18. doi: 10.1016/j.molcel.2008.11.004. PMID: 19026781; PMCID: PMC3641558.

Martínez Valle F, Balada E, Ordi-Ros J, Vilardell-Tarres M. DNase 1 and systemic lupus erythematosus. Autoimmun Rev. 2008 May;7(5):359-63. doi: 10.1016/j.autrev.2008.02.002. Epub 2008 Mar 10. PMID: 18486922.

Maurya D, Rai G, Mandal D, Mondal BC. Transient caspase-mediated activation of caspase-activated DNase causes DNA damage required for phagocytic macrophage differentiation. Cell Rep. 2024 May 28;43(5):114251. doi: 10.1016/j.celrep.2024.114251. Epub 2024 May 17. PMID: 38761374.

Meers MP, Tenenbaum D, Henikoff S. Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. Epigenetics Chromatin. 2019 Jul 12;12(1):42. doi: 10.1186/s13072-019-0287-4. PMID: 31300027; PMCID: PMC6624997.

Miller KM, Jackson SP. Histone marks: repairing DNA breaks within the context of chromatin. Biochem Soc Trans. 2012 Apr;40(2):370-6. doi: 10.1042/BST20110747. PMID: 22435814.

Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer. 2007 Apr;7(4):233-45. doi: 10.1038/nrc2091. Epub 2007 Mar 15. PMID: 17361217.
Moeed A, Thilmany N, Beck F, Puthussery BK, Ortmann N, Haimovici A, Badr MT, Haghighi EB, Boerries M, Öllinger R, Rad R, Kirschnek S, Gentle IE, Donakonda S, Petric PP, Hummel JF, Pfaffendorf E, Zanetta P, Schell C, Schwemmle M, Weber A, Häcker G. The Caspase-Activated DNase drives inflammation and contributes to defense against viral infection. Cell Death Differ. 2024 Jul;31(7):924-937. doi: 10.1038/s41418-024-01320-7. Epub 2024 Jun 7. PMID: 38849575; PMCID: PMC11239672.

Mori H, Colman SM, Xiao Z, Ford AM, Healy LE, Donaldson C, Hows JM, Navarrete C, Greaves M. Chromosome translocations and covert leukemic clones are generated during normal fetal development. Proc Natl Acad Sci U S A. 2002 Jun 11;99(12):8242-7. doi: 10.1073/pnas.112218799. Epub 2002 Jun 4. PMID: 12048236; PMCID: PMC123052.

Mosesso P, Piane M, Palitti F, Pepe G, Penna S, Chessa L. The novel human gene aprataxin is directly involved in DNA single-strand-break repair. Cell Mol Life Sci. 2005 Feb;62(4):485-91. doi: 10.1007/s00018-004-4441-0. PMID: 15719174.

Murray TV, McMahon JM, Howley BA, Stanley A, Ritter T, Mohr A, Zwacka R, Fearnhead HO. A non-apoptotic role for caspase-9 in muscle differentiation. J Cell Sci. 2008 Nov 15;121(Pt 22):3786-93. doi: 10.1242/jcs.024547. Epub 2008 Oct 28. PMID: 18957517.

Nekrasov M, Wild B, Müller J. Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. EMBO Rep. 2005 Apr;6(4):348-53. doi: 10.1038/sj.embor.7400376. PMID: 15776017; PMCID: PMC1299286.
Nguyen TH, Barr FG. Therapeutic Approaches Targeting PAX3-FOXO1 and Its Regulatory and Transcriptional Pathways in Rhabdomyosarcoma. Molecules. 2018 Oct 28;23(11):2798. doi: 10.3390/molecules23112798. PMID: 30373318; PMCID: PMC6278278.

Nowell PC. Discovery of the Philadelphia chromosome: a personal perspective. J Clin Invest. 2007 Aug;117(8):2033-5. doi: 10.1172/JCI31771. PMID: 17671636; PMCID: PMC1934591.

Ohsawa S, Hamada S, Kuida K, Yoshida H, Igaki T, Miura M. Maturation of the olfactory sensory neurons by Apaf-1/caspase-9-mediated caspase activity. Proc Natl Acad Sci U S A. 2010 Jul 27;107(30):13366-71. doi: 10.1073/pnas.0910488107. Epub 2010 Jul 12. PMID: 20624980; PMCID: PMC2922127.

Olguin HC, Olwin BB. Pax-7 up-regulation inhibits myogenesis and cell cycle progression in satellite cells: a potential mechanism for self-renewal. Dev Biol. 2004 Nov 15;275(2):375-88. doi: 10.1016/j.ydbio.2004.08.015. PMID: 15501225; PMCID: PMC3322464.

Ottone T, Hasan SK, Montefusco E, Curzi P, Mays AN, Chessa L, Ferrari A, Conte E, Noguera NI, Lavorgna S, Ammatuna E, Divona M, Bovetti K, Amadori S, Grimwade D, Lo-Coco F. Identification of a potential "hotspot" DNA region in the RUNX1 gene targeted by mitoxantrone in therapy-related acute myeloid leukemia with t(16;21) translocation. Genes Chromosomes Cancer. 2009 Mar;48(3):213-21. doi: 10.1002/gcc.20633. PMID: 19023877.

Panizza S, Mendoza MA, Berlinger M, Huang L, Nicolas A, Shirahige K, Klein F. Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. Cell. 2011 Aug 5;146(3):372-83. doi: 10.1016/j.cell.2011.07.003. PMID: 21816273.

Park HH. Structural insight into CIDE domains: the Janus face of CIDEs. Apoptosis. 2015 Feb;20(2):240-9. doi: 10.1007/s10495-014-1067-z. PMID: 25488366.

Parrish AB, Freel CD, Kornbluth S. Cellular mechanisms controlling caspase activation and function. Cold Spring Harb Perspect Biol. 2013 Jun 1;5(6):a008672. doi: 10.1101/cshperspect.a008672. PMID: 23732469; PMCID: PMC3660825.

Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E, Helin K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. 2004 Oct 13;23(20):4061-71. doi: 10.1038/sj.emboj.7600402. Epub 2004 Sep 23. PMID: 15385962; PMCID: PMC524339.

Pommier Y, Sun Y, Huang SN, Nitiss JL. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. Nat Rev Mol Cell Biol. 2016 Nov;17(11):703-721. doi: 10.1038/nrm.2016.111. Epub 2016 Sep 21. PMID: 27649880; PMCID: PMC9248348.

Putinski C, Abdul-Ghani M, Brunette S, Burgon PG, Megeney LA. Caspase Cleavage of Gelsolin Is an Inductive Cue for Pathologic Cardiac Hypertrophy. J Am Heart Assoc. 2018 Dec 4;7(23):e010404. doi: 10.1161/JAHA.118.010404. PMID: 30486716; PMCID: PMC6405540.

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016 Jul 8;44(W1):W160-5. doi: 10.1093/nar/gkw257. Epub 2016 Apr 13. PMID: 27079975; PMCID: PMC4987876.

Ramsden DA, Nussenzweig A. Mechanisms driving chromosomal translocations: lost in time and space. Oncogene. 2021 Jun;40(25):4263-4270. doi: 10.1038/s41388-021-01856-9. Epub 2021 Jun 8. PMID: 34103687; PMCID: PMC8238880.

Relaix F, Rocancourt D, Mansouri A, Buckingham M. Divergent functions of murine Pax3 and Pax7 in limb muscle development. Genes Dev. 2004 May 1;18(9):1088-105. doi: 10.1101/gad.301004. PMID: 15132998; PMCID: PMC406297.

Ribeil JA, Zermati Y, Vandekerckhove J, Cathelin S, Kersual J, Dussiot M, Coulon S, Moura IC, Zeuner A, Kirkegaard-Sørensen T, Varet B, Solary E, Garrido C, Hermine O. Hsp70 regulates erythropoiesis by preventing caspase-3-mediated cleavage of GATA-1. Nature. 2007 Jan 4;445(7123):102-5. doi: 10.1038/nature05378. Epub 2006 Dec 10. PMID: 17167422.

Richly H, Aloia L, Di Croce L. Roles of the Polycomb group proteins in stem cells and cancer. Cell Death Dis. 2011 Sep 1;2(9):e204. doi: 10.1038/cddis.2011.84. PMID: 21881606; PMCID: PMC3186902.
Richly H, Aloia L, Di Croce L. Roles of the Polycomb group proteins in stem cells and cancer. Cell Death Dis. 2011 Sep 1;2(9):e204. doi: 10.1038/cddis.2011.84. PMID: 21881606; PMCID: PMC3186902.

Ringrose L, Paro R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet. 2004;38:413-43. doi: 10.1146/annurev.genet.38.072902.091907. PMID: 15568982.
Rode A, Maass KK, Willmund KV, Lichter P, Ernst A. Chromothripsis in cancer cells: An update. Int J Cancer. 2016 May 15;138(10):2322-33. doi: 10.1002/ijc.29888. Epub 2015 Oct 30. PMID: 26455580.

Rothlin CV, Ghosh S. Cracking the Cell Death Code. Cold Spring Harb Perspect Biol. 2020 May 1;12(5):a036343. doi: 10.1101/cshperspect.a036343. PMID: 31548182; PMCID: PMC7197433.
Rowley JD. Mapping of human chromosomal regions related to neoplasia: evidence from chromosomes 1 and 17. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5729-33. doi: 10.1073/pnas.74.12.5729. PMID: 271998; PMCID: PMC431867.

Rowley JD. Mapping of human chromosomal regions related to neoplasia: evidence from chromosomes 1 and 17. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5729-33. doi: 10.1073/pnas.74.12.5729. PMID: 271998; PMCID: PMC431867.

Sakahira H, Enari M, Nagata S. Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. Nature. 1998 Jan 1;391(6662):96-9. doi: 10.1038/34214. Erratum in: Nature. 2015 Oct 29;526(7575):728. doi: 10.1038/nature15532. PMID: 9422513.

Samejima K, Earnshaw WC. Differential localization of ICAD-L and ICAD-S in cells due to removal of a C-terminal NLS from ICAD-L by alternative splicing. Exp Cell Res. 2000 Mar 15;255(2):314-20. doi: 10.1006/excr.2000.4801. PMID: 10694446.

Samejima K, Earnshaw WC. Trashing the genome: the role of nucleases during apoptosis. Nat Rev Mol Cell Biol. 2005 Sep;6(9):677-88. doi: 10.1038/nrm1715. PMID: 16103871.

Schüler F, Dölken L, Hirt C, Kiefer T, Berg T, Fusch G, Weitmann K, Hoffmann W, Fusch C, Janz S, Rabkin CS, Dölken G. Prevalence and frequency of circulating t(14;18)-MBR translocation carrying cells in healthy individuals. Int J Cancer. 2009 Feb 15;124(4):958-63. doi: 10.1002/ijc.23958. PMID: 19030176; PMCID: PMC4216731.
Seale P, Sabourin LA, Girgis-Gabardo A, Mansouri A, Gruss P, Rudnicki MA. Pax7 is required for the specification of myogenic satellite cells. Cell. 2000 Sep 15;102(6):777-86. doi: 10.1016/s0092-8674(00)00066-0. PMID: 11030621.

Seenundun S, Rampalli S, Liu QC, Aziz A, Palii C, Hong S, Blais A, Brand M, Ge K, Dilworth FJ. UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. EMBO J. 2010 Apr 21;29(8):1401-11. doi: 10.1038/emboj.2010.37. Epub 2010 Mar 18. PMID: 20300060; PMCID: PMC2868576.

Shalini S, Dorstyn L, Dawar S, Kumar S. Old, new and emerging functions of caspases. Cell Death Differ. 2015 Apr;22(4):526-39. doi: 10.1038/cdd.2014.216. Epub 2014 Dec 19. PMID: 25526085; PMCID: PMC4356345.
Shkreta L, Chabot B. The RNA Splicing Response to DNA Damage. Biomolecules. 2015 Oct 29;5(4):2935-77. doi: 10.3390/biom5042935. PMID: 26529031; PMCID: PMC4693264.

Slaymaker IM, Fu Y, Toso DB, Ranatunga N, Brewster A, Forsburg SL, Zhou ZH, Chen XS. Mini-chromosome maintenance complexes form a filament to remodel DNA structure and topology. Nucleic Acids Res. 2013 Mar 1;41(5):3446-56. doi: 10.1093/nar/gkt022. Epub 2013 Jan 29. PMID: 23361460; PMCID: PMC3597688.

Solier S, Fontenay M, Vainchenker W, Droin N, Solary E. Non-apoptotic functions of caspases in myeloid cell differentiation. Cell Death Differ. 2017 Aug;24(8):1337-1347. doi: 10.1038/cdd.2017.19. Epub 2017 Feb 17. PMID: 28211870; PMCID: PMC5520450.

Sordet O, Rébé C, Plenchette S, Zermati Y, Hermine O, Vainchenker W, Garrido C, Solary E, Dubrez-Daloz L. Specific involvement of caspases in the differentiation of monocytes into macrophages. Blood. 2002 Dec 15;100(13):4446-53. doi: 10.1182/blood-2002-06-1778. Epub 2002 Aug 8. PMID: 12393560.

Steinauer N, Guo C, Zhang J. The transcriptional corepressor CBFA2T3 inhibits all-trans-retinoic acid-induced myeloid gene expression and differentiation in acute myeloid leukemia. J Biol Chem. 2020 Jul 3;295(27):8887-8900. doi: 10.1074/jbc.RA120.013042. Epub 2020 May 20. PMID: 32434928; PMCID: PMC7335779.

Steller H. Regulation of apoptosis in Drosophila. Cell Death Differ. 2008 Jul;15(7):1132-8. doi: 10.1038/cdd.2008.50. Epub 2008 Apr 25. PMID: 18437164.

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, McLaren S, Lin ML, McBride DJ, Varela I, Nik-Zainal S, Leroy C, Jia M, Menzies A, Butler AP, Teague JW, Quail MA, Burton J, Swerdlow H, Carter NP, Morsberger LA, Iacobuzio-Donahue C, Follows GA, Green AR, Flanagan AM, Stratton MR, Futreal PA, Campbell PJ. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 2011 Jan 7;144(1):27-40. doi: 10.1016/j.cell.2010.11.055. PMID: 21215367; PMCID: PMC3065307.

Stojic L, Jasencakova Z, Prezioso C, Stützer A, Bodega B, Pasini D, Klingberg R, Mozzetta C, Margueron R, Puri PL, Schwarzer D, Helin K, Fischle W, Orlando V. Chromatin regulated interchange between polycomb repressive complex 2 (PRC2)-Ezh2 and PRC2-Ezh1 complexes controls myogenin activation in skeletal muscle cells. Epigenetics Chromatin. 2011 Sep 5;4:16. doi: 10.1186/1756-8935-4-16. PMID: 21892963; PMCID: PMC3180244.
Stott RT, Kritsky O, Tsai LH. Profiling DNA break sites and transcriptional changes in response to contextual fear learning. PLoS One. 2021 Jul 1;16(7):e0249691. doi: 10.1371/journal.pone.0249691. PMID: 34197463; PMCID: PMC8248687.

Suberbielle E, Sanchez PE, Kravitz AV, Wang X, Ho K, Eilertson K, Devidze N, Kreitzer AC, Mucke L. Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-β. Nat Neurosci. 2013 May;16(5):613-21. doi: 10.1038/nn.3356. Epub 2013 Mar 24. PMID: 23525040; PMCID: PMC3637871.

Summers KE, Goff LK, Wilson AG, Gupta RK, Lister TA, Fitzgibbon J. Frequency of the Bcl-2/IgH rearrangement in normal individuals: implications for the monitoring of disease in patients with follicular lymphoma. J Clin Oncol. 2001 Jan 15;19(2):420-4. doi: 10.1200/JCO.2001.19.2.420. PMID: 11208834.
Tan JZ, Yan Y, Wang XX, Jiang Y, Xu HE. EZH2: biology, disease, and structure-based drug discovery. Acta Pharmacol Sin. 2014 Feb;35(2):161-74. doi: 10.1038/aps.2013.161. Epub 2013 Dec 23. PMID: 24362326; PMCID: PMC3914023.

Thalhammer V, Lopez-Garcia LA, Herrero-Martin D, Hecker R, Laubscher D, Gierisch ME, Wachtel M, Bode P, Nanni P, Blank B, Koscielniak E, Schäfer BW. PLK1 phosphorylates PAX3-FOXO1, the inhibition of which triggers regression of alveolar Rhabdomyosarcoma. Cancer Res. 2015 Jan 1;75(1):98-110. doi: 10.1158/0008-5472.CAN-14-1246. Epub 2014 Nov 14. PMID: 25398439.

Thompson PS, Cortez D. New insights into abasic site repair and tolerance. DNA Repair (Amst). 2020 Jun;90:102866. doi: 10.1016/j.dnarep.2020.102866. Epub 2020 Apr 30. PMID: 32417669; PMCID: PMC7299775.

Tian J, Wen M, Gao P, Feng M, Wei G. RUVBL1 ubiquitination by DTL promotes RUVBL1/2-β-catenin-mediated transcriptional regulation of NHEJ pathway and enhances radiation resistance in breast cancer. Cell Death Dis. 2024 Apr 12;15(4):259. doi: 10.1038/s41419-024-06651-4. PMID: 38609375; PMCID: PMC11015013.

Timson DJ, Singleton MR, Wigley DB. DNA ligases in the repair and replication of DNA. Mutat Res. 2000 Aug 30;460(3-4):301-18. doi: 10.1016/s0921-8777(00)00033-1. PMID: 10946235.
Tubbs A, Nussenzweig A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell. 2017 Feb 9;168(4):644-656. doi: 10.1016/j.cell.2017.01.002. PMID: 28187286; PMCID: PMC6591730.

Vier J, Fürmann C, Häcker G. Baculovirus P35 protein does not inhibit caspase-9 in a cell-free system of apoptosis. Biochem Biophys Res Commun. 2000 Oct 5;276(3):855-61. doi: 10.1006/bbrc.2000.3560. PMID: 11027559.
Walsh JG, Cullen SP, Sheridan C, Lüthi AU, Gerner C, Martin SJ. Executioner caspase-3 and caspase-7 are functionally distinct proteases. Proc Natl Acad Sci U S A. 2008 Sep 2;105(35):12815-9. doi: 10.1073/pnas.0707715105. Epub 2008 Aug 22. PMID: 18723680; PMCID: PMC2529079.

Wei L, Zhao X. A new MCM modification cycle regulates DNA replication initiation. Nat Struct Mol Biol. 2016 Mar;23(3):209-16. doi: 10.1038/nsmb.3173. Epub 2016 Feb 8. PMID: 26854664; PMCID: PMC4823995.

Whitcomb SJ, Basu A, Allis CD, Bernstein E. Polycomb Group proteins: an evolutionary perspective. Trends Genet. 2007 Oct;23(10):494-502. doi: 10.1016/j.tig.2007.08.006. Epub 2007 Sep 7. PMID: 17825942.

Widlak P, Li P, Wang X, Garrard WT. Cleavage preferences of the apoptotic endonuclease DFF40 (caspase-activated DNase or nuclease) on naked DNA and chromatin substrates. J Biol Chem. 2000 Mar 17;275(11):8226-32. doi: 10.1074/jbc.275.11.8226. PMID: 10713148.

Wiles ET, Selker EU. H3K27 methylation: a promiscuous repressive chromatin mark. Curr Opin Genet Dev. 2017 Apr;43:31-37. doi: 10.1016/j.gde.2016.11.001. Epub 2016 Dec 8. PMID: 27940208; PMCID: PMC5447479.

Wilson A, Murphy MJ, Oskarsson T, Kaloulis K, Bettess MD, Oser GM, Pasche AC, Knabenhans C, Macdonald HR, Trumpp A. c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. Genes Dev. 2004 Nov 15;18(22):2747-63. doi: 10.1101/gad.313104. PMID: 15545632; PMCID: PMC528895.

Wong CF, Tellam RL. MicroRNA-26a targets the histone methyltransferase Enhancer of Zeste homolog 2 during myogenesis. J Biol Chem. 2008 Apr 11;283(15):9836-43. doi: 10.1074/jbc.M709614200. Epub 2008 Feb 15. PMID: 18281287.

Wong CF, Tellam RL. MicroRNA-26a targets the histone methyltransferase Enhancer of Zeste homolog 2 during myogenesis. J Biol Chem. 2008 Apr 11;283(15):9836-43. doi: 10.1074/jbc.M709614200. Epub 2008 Feb 15. PMID: 18281287.

Woo EJ, Kim YG, Kim MS, Han WD, Shin S, Robinson H, Park SY, Oh BH. Structural mechanism for inactivation and activation of CAD/DFF40 in the apoptotic pathway. Mol Cell. 2004 May 21;14(4):531-9. doi: 10.1016/s1097-2765(04)00258-8. PMID: 15149602.

Wyllie AH, Kerr JF, Currie AR. Cell death: the significance of apoptosis. Int Rev Cytol. 1980;68:251-306. doi: 10.1016/s0074-7696(08)62312-8. PMID: 7014501.

Yadav AK, Polasek-Sedlackova H. Quantity and quality of minichromosome maintenance protein complexes couple replication licensing to genome integrity. Commun Biol. 2024 Feb 9;7(1):167. doi: 10.1038/s42003-024-05855-w. PMID: 38336851; PMCID: PMC10858283.

Yadav VK, Claeys Bouuaert C. Mechanism and Control of Meiotic DNA Double-Strand Break Formation in S. cerevisiae. Front Cell Dev Biol. 2021 Mar 2;9:642737. doi: 10.3389/fcell.2021.642737. PMID: 33748134; PMCID: PMC7968521.

Yamamoto K, Matsumoto H, Matsumoto S, Sakai R, Kitao A, Watanabe M, Goto H, Sugimoto T, Yano Y, Yakushijin K, Minami H. Unexpected appearance of KMT2A::MLLT10 fusion transcript in acute myeloid leukemia with t(5;11)(q31;q23.3). Cancer Genet. 2023 Apr;272-273:41-46. doi: 10.1016/j.cancergen.2023.02.002. Epub 2023 Feb 4. PMID: 36774707.

Yang JL, Tadokoro T, Keijzers G, Mattson MP, Bohr VA. Neurons efficiently repair glutamate-induced oxidative DNA damage by a process involving CREB-mediated up-regulation of apurinic endonuclease 1. J Biol Chem. 2010 Sep 3;285(36):28191-9. doi: 10.1074/jbc.M109.082883. Epub 2010 Jun 23. PMID: 20573957; PMCID: PMC2934684.

Yao L, Swartz P, Hamilton PT, Clark AC. Remodeling hydrogen bond interactions results in relaxed specificity of Caspase-3. Biosci Rep. 2021 Jan 29;41(1):BSR20203495. doi: 10.1042/BSR20203495. PMID: 33448281; PMCID: PMC7846959.

Yashar WM, Kong G, VanCampen J, Curtiss BM, Coleman DJ, Carbone L, Yardimci GG, Maxson JE, Braun TP. GoPeaks: histone modification peak calling for CUT&Tag. Genome Biol. 2022 Jul 4;23(1):144. doi: 10.1186/s13059-022-02707-w. PMID: 35788238; PMCID: PMC9252088.

Yokoyama H, Mukae N, Sakahira H, Okawa K, Iwamatsu A, Nagata S. A novel activation mechanism of caspase-activated DNase from Drosophila melanogaster. J Biol Chem. 2000 Apr 28;275(17):12978-86. doi: 10.1074/jbc.275.17.12978. PMID: 10777599.

Yu JR, Lee CH, Oksuz O, Stafford JM, Reinberg D. PRC2 is high maintenance. Genes Dev. 2019 Aug 1;33(15-16):903-935. doi: 10.1101/gad.325050.119. Epub 2019 May 23. PMID: 31123062; PMCID: PMC6672058.
Yu JR, Lee CH, Oksuz O, Stafford JM, Reinberg D. PRC2 is high maintenance. Genes Dev. 2019 Aug 1;33(15-16):903-935. doi: 10.1101/gad.325050.119. Epub 2019 May 23. PMID: 31123062; PMCID: PMC6672058.

Yuan JY, Horvitz HR. The Caenorhabditis elegans genes ced-3 and ced-4 act cell autonomously to cause programmed cell death. Dev Biol. 1990 Mar;138(1):33-41. doi: 10.1016/0012-1606(90)90174-h. PMID: 2307287.
Yuan JY, Horvitz HR. The Caenorhabditis elegans genes ced-3 and ced-4 act cell autonomously to cause programmed cell death. Dev Biol. 1990 Mar;138(1):33-41. doi: 10.1016/0012-1606(90)90174-h. PMID: 2307287.

Yuan W, Wu T, Fu H, Dai C, Wu H, Liu N, Li X, Xu M, Zhang Z, Niu T, Han Z, Chai J, Zhou XJ, Gao S, Zhu B. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. Science. 2012 Aug 24;337(6097):971-5. doi: 10.1126/science.1225237. PMID: 22923582.1974

Zermati Y, Garrido C, Amsellem S, Fishelson S, Bouscary D, Valensi F, Varet B, Solary E, Hermine O. Caspase activation is required for terminal erythroid differentiation. J Exp Med. 2001 Jan 15;193(2):247-54. doi: 10.1084/jem.193.2.247. PMID: 11208865; PMCID: PMC2193347.

Zhao B, Yang J, Ji P. Chromatin condensation during terminal erythropoiesis. Nucleus. 2016 Sep 2;7(5):425-429. doi: 10.1080/19491034.2016.1226717. Epub 2016 Aug 31. Erratum in: doi: 10.1016/j.devcel.2016.02.001. PMID: 27579498; PMCID: PMC5120594.