
Electronic Thesis and Dissertation Repository

8-13-2019 2:00 PM

Functional investigation of the role of the retinoblastoma protein in genome stability

Aren E. Marshall
The University of Western Ontario

Supervisor
Dick, Fred A.
The University of Western Ontario

Graduate Program in Biochemistry
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Aren E. Marshall 2019

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Molecular Biology Commons](#)

Recommended Citation

Marshall, Aren E., "Functional investigation of the role of the retinoblastoma protein in genome stability" (2019). *Electronic Thesis and Dissertation Repository*. 6342.
<https://ir.lib.uwo.ca/etd/6342>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Genome instability is an enabling characteristic of cancerous cells. It has recently been discovered that the retinoblastoma protein (pRB), typically known for its role in cell cycle regulation, also aids in the maintenance of genome stability. Intriguingly, mutations to the pRB gene, *RBI*, can arise late in tumorigenesis in cancer cells whose cell cycle regulation is already compromised by another mutation. This suggests that pRB's functions in genome stability could underlie cancer relevant characteristics that are independent of its ability to negatively regulate proliferation. The overall aim of this thesis is to characterize the different means through which pRB contributes to the preservation of genome integrity. Using CRISPR/Cas9, isogenic *RBI* mutant genotypes were created in a number of cancer cell lines. Cells with at least one mutant copy of *RBI* have increased basal levels of DNA damage and increased mitotic errors. When the underlying origins of these phenotypes were investigated further, I discovered elevated levels of reactive oxygen species as well as impaired homologous recombination repair in cells with *RBI* mutations. When xenografted into immune compromised mice, *RBI* mutation also results in an increased capacity to seed new tumors in the lungs. This thesis also investigates the functions of the pRB-condensin II complex in maintaining genome stability, specifically in interphase cells. Using a gene-targeted mouse model that disrupts the ability of pRB to recruit condensin II, *Rbl^L*, locations of pRB-dependent condensin II recruitment were investigated. I found that both condensin II and another architectural protein complex, TFIIC, are recruited to promoters between bidirectional genes by a mechanism that is reliant on pRB. Recruitment of these architectural proteins at bidirectional promoters is required to establish long-range chromosome interactions and transcriptional insulation between gene pairs. In addition, pRB deletion in cancer causes similar misregulation at divergent promoters, indicating that loss of insulation impacts the transcriptome of cancer cells. Overall, this work demonstrates that beyond altered proliferative control, loss of pRB can also contribute to cancer progression through enhanced DNA damage and altered chromosome topology.

Keywords

Retinoblastoma protein, genome instability, chemotherapy, DNA damage, cancer, chromosome architecture, condensin II, TFIIC, gene expression

Lay Abstract

In human cells, maintaining the integrity of DNA is critical to preserve proper function and health. If genomic instability does occur, it can lead to cancer. It has recently been discovered that the retinoblastoma protein (pRB), typically known for its role in regulating cellular growth rates to provide protection against cancer, also aids in the maintenance of genome stability. Intriguingly, mutations to pRB can arise in cancer cells where cellular growth is already compromised. This suggests that pRB's functions in genome stability could underlie cancer relevant characteristics that are independent of its ability to regulate cellular growth rates. The overall aim of this thesis is to characterize the different means through which pRB contributes to the preservation of genome integrity. To investigate this, pRB mutations were created in a number of cancer cell lines. Cells with pRB mutations have increased basal levels of DNA damage and increased errors when cells are dividing. When transplanted into immune compromised mice, cells with pRB mutation have an increased capacity to form new tumors in the lungs. This thesis also uses a mouse model with a targeted mutation to investigate the genome stability functions of pRB when it is in a complex with another protein, condensin II, which is known to fold DNA to package it tightly. I found that condensin II and another protein known to alter DNA packaging, TFIIC, are both reliant on pRB to be recruited to the beginning of genes facing opposite directions. Recruitment of condensin II and TFIIC at these locations is required for proper DNA packaging and gene expression. In addition, pRB deletion in cancer causes similar misexpression of genes at these locations, indicating that loss of pRB and hence condensin II and TFIIC localization impacts gene expression in cancer cells. Overall, this work demonstrates that beyond altered cellular growth rates, loss of pRB can also contribute to cancer progression through enhanced DNA damage and altered DNA packaging. In the future, these recently discovered characteristics could be used to select the best therapeutic tools for patients with pRB loss.

Co-Authorship Statement

All chapters were written by Aren Marshall and, with the exception of chapter 4, all chapters were edited by Dr. Fred Dick.

All experiments in chapter 2 were executed by Aren Marshall with assistance as follows: Michael Roes performed the IF staining and imaging for Figure 2.1D&E. Megan DeWeerd assisted in working out the conditions for the NHEJ and HR repair assays which eventually led to the final experiments in Figure 2.4A-F. Andrea Chaikovsky and Dr. Julien Sage provided clones of control and *RBI* knockout U2OS, NCI-H460 and NCI-H1792 cells used for Figures 2.1F-H, 2.2E&F, and 2.5C-F. Dr. Daniel Passos performed subcutaneous and tail vein injections and provided assistance with monitoring mice and harvesting tissue for Figures 2.7 and 2.8A-C. Dr. Passos also performed the genotyping for Figure 2.8C. Dr. Christopher Howlett conducted pathological analysis for Figure 2.7B.

All experiments in chapter 3 were performed by Aren Marshall with the exception of Figure 3.14D, which was previously performed by Dr. Charles Ishak and is unpublished work that also appears in his thesis.

Acknowledgments

I would like to thank my supervisor Dr. Fred Dick for the wonderful training environment provided. I was given the opportunity to pursue experiments that, at the beginning of my time here in the lab, I never would have envisioned. I also have to thank you for teaching me how to think critically, which is the most important ability I have gained from this experience. All of this is to be credited to your unwavering mentorship and encouragement throughout the years; I really could not ask for anything more from a supervisor.

Thank you to my advisory committee, Dr. Joe Torchia and Dr. Nathalie Bérubé, for your advice and insight throughout the years. I am grateful for all the discussions we have had and your investment in my development throughout my time as a graduate student.

I would also like to thank all past and present members of the Dick lab for making my time here so enjoyable. I am thankful for all of the discussions and the assistance over the years, scientific and otherwise, as well as the friendships that were formed. My time in graduate school would not have been the same without you all.

Finally, I would like to thank all of my family and friends. Without all of your love and support, I certainly would not be in the position I am today. I am incredibly grateful to have such amazing people encourage me daily in every aspect of my life. In particular, to my parents and my sister, thank you for always being there for me. Lastly, to my husband, thank you for vowing to encourage and inspire me, “especially when science seems to suck”; so far, you have done a great job!

Abbreviations

γ H2AX: phosphorylation of H2AX at Serine 139

γ IR: gamma irradiation

3C: chromosome conformation capture

4C: circularized chromosome conformation capture

8-oxoG: 8-oxoguanine

53BP1: p53-binding protein 1

AFM: atomic force microscopy

ATP: adenosine triphosphate

bp: base pair

BETA: Binding and Expression Target Analysis

BLM: Bloom syndrome protein

BrdU: 5-bromo-2'-deoxyuridine

BSA: bovine serum albumin

CA-DCF-DA: 5(6)-carboxy-2',7'-dichlorodihydrofluorescein diacetate

CAP-D2, D3, G, G2, H, H2: chromosome associated protein D2, D3, G, G2, H, H2 respectively

Cas9: clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein 9

CDK: cyclin-dependent kinase

CdLS: Cornelia de Lange syndrome

CEAS: Cis-regulatory Element Annotation System

CFS: common fragile sites

cGAS: cyclic GMP-AMP synthase protein

ChIP: chromatin immunoprecipitation

CIN: chromosome instability

CKI: cyclin-dependent kinase inhibitor

CNA: copy number alteration

CRISPR: clustered regularly interspaced short palindromic repeats

CTCF: CCCTC-binding factor

C-terminal: carboxy-terminal

DAPI: 4',6-diamidino-2-phenylindole

DAVID: Database for Annotation, Visualization and Integrated Discovery

dCAP-D3: *Drosophila* CAP-D3

DHFR: dihydrofolate reductase

DMEM: Dulbecco's Modified Eagle Medium

DNMT: DNA methyltransferase

DP: differentiation-regulated transcription factor-1 polypeptide

Drosophila: *Drosophila melanogaster*

DRTF1: differentiation-regulated transcription factor-1

E2F: E2 promoter binding factor

EGFR: epidermal growth factor receptor (gene)

EGS: ethylene glycol bis(succinimidyl succinate)

EOC: epithelial ovarian cancer

eRNA⁺: enhancer RNA-positive

ETC: extra transcription factor IIIc

EZH2: enhancer-of-zeste-homologue 2

EZH2i: enhancer-of-zeste-homologue 2 inhibitor

FBS: fetal bovine serum

FFPE: formalin-fixed paraffin-embedded

G1: Gap 1 phase of the cell division cycle

G2: Gap 2 phase of the cell division cycle

GAPDH: glyceraldehyde 3-phosphate dehydrogenase

GFP: green fluorescent protein

Gy: gray, unit referring to ionizing radiation

H&E: hematoxylin and eosin

H₂O₂: hydrogen peroxide

HDAC: histone deacetylase

HEAT: huntingtin, elongation factor 3, the PR65/A subunit of protein phosphatase 2A, TOR1; four proteins that contain the HEAT repeat domain

HGSOC: high-grade serous ovarian carcinoma

HR: homologous recombination

IC₅₀: half maximal inhibitory concentration

IF: immunofluorescence

IgG: immunoglobulin G

kb: kilobase

kDa: kilodalton

LINE: long interspersed nuclear element

LTR: long terminal repeat

LXCXE: leucine-any amino acid-cysteine-any amino acid-glutamate

M: mitosis

MACS: Model-based Analysis for ChIP-Seq

Mb: megabase

MBD: marked box domain

MDa: megadalton

MEF: mouse embryonic fibroblast

mESC: mouse embryonic stem cell

MIN: micro- and minisatellite instability

MMR: mismatch repair

MO: morpholino oligonucleotide

N: haploid ploidy (when referring to ploidy and DNA content)

NCSLC: non-small cell lung cancer

NGS: next generation sequencing

NHEJ: non-homologous end joining

p107: protein encoded by *RBL1*

p130: protein encoded by *RBL2*

PARPi: poly-ADP ribose polymerase (PARP) inhibitor

PBS: phosphate buffered saline

PFA: Paraformaldehyde

PI: propidium iodide

Pol: polymerase

PP1: protein phosphatase 1

PP2A: protein phosphatase 2A

pRB: retinoblastoma tumor suppressor protein

RBI: human retinoblastoma susceptibility gene

Rb1: mouse retinoblastoma susceptibility gene

Rb1^G: mouse retinoblastoma susceptibility gene containing E2F general binding site mutations (K461E and R542E)

Rb1^L: mouse retinoblastoma susceptibility gene containing the LXCXE binding cleft mutations (I746A, N750A, M754A)

RBL1: retinoblastoma like 1

RBL2: retinoblastoma like 2

RBF1: *Drosophila* pRB family homolog (protein)

RBLP: pRB large pocket

RBS: Robert syndrome

ROS: reactive oxygen species

RPA: replication protein A

S: DNA synthesis phase of cell division cycle

SA: stromal antigen

SCLC: small cell lung cancer

SEC: size exclusion chromatography

Seq: sequencing

sgRNA: single guide RNA

SINE: short interspersed nuclear element

SMC: structural maintenance of chromosomes

Suv39h1: suppressor of variegation 3-9 homolog 1

TAD: topologically associating domain (when referring to chromosome conformation)

TAD: transactivation domain (when referring to interaction surface on E2Fs)

T-ALL: T-cell acute lymphoblastic leukemia

TCGA: The Cancer Genome Atlas

TFIIIC: transcription factor IIIC

TSS: transcriptional start site

UFB: ultra-fine bridge

ZGA: zygotic genome activation

Table of Contents

Abstract.....	ii
Keywords.....	iii
Lay Abstract.....	iv
Co-Authorship Statement.....	v
Acknowledgments.....	vi
Abbreviations.....	vii
Table of Contents.....	xi
List of Tables.....	xvi
List of Figures.....	xvii
List of Appendices.....	xx
Chapter 1.....	1
1 Introduction.....	1
1.1 Retinoblastoma onset and the discovery of the retinoblastoma gene.....	1
1.2 The retinoblastoma protein is a cell cycle regulator.....	2
1.3 The pocket domain defines the pRB family.....	3
1.4 The pRB family is recruited to DNA by E2F transcription factors.....	6
1.5 Pocket proteins have overlapping and unique roles.....	7
1.6 E2F1 and pRB have a unique interaction interface.....	8
1.7 Evidence for a separate role of pRB in tumor suppression outside of cell cycle control.....	9
1.8 Genome instability is an enabling characteristic of cancer.....	11
1.9 Condensins are central components of mitotic chromosome dynamics.....	12
1.10 Condensin I and condensin II are differentially loaded onto mitotic chromosomes.....	15

1.11	Condensins can translocate DNA and extrude DNA loops	17
1.12	Cohesins are important for organizing the interphase genome	20
1.13	Condensin II binding has been associated with increased boundary strength....	25
1.14	Condensin II has functional roles outside of mitosis.....	27
1.15	Mutations in condensin components have been linked to human disease.....	30
1.16	The <i>Rb1</i> ^{L/L} mouse model reveals importance of condensin II interactions with the LXCXE binding cleft.....	32
1.17	pRB interacts with condensin II to mediate genome stability	34
1.18	Objectives of the present study.....	35
1.19	References.....	36
Chapter 2.....		55
2	<i>Rb1</i> deletion in pRB-pathway disrupted cells results in DNA damage and cancer progression.	55
2.1	Abstract.....	55
2.2	Introduction.....	55
2.3	Materials and Methods	57
2.3.1	Cell culture.....	57
2.3.2	Generation of <i>Rb1</i> deletions using CRISPR.....	58
2.3.3	Fluorescence microscopy.....	60
2.3.4	Gamma irradiation of cells.....	61
2.3.5	NHEJ and HR repair assays.....	61
2.3.6	Expression of HR factors	63
2.3.7	Determination of IC50 concentrations.....	64
2.3.8	ChIP-Sequencing	64
2.3.9	Flow cytometry	66
2.3.10	Nucleoside supplementation	66

2.3.11	Measurement of reactive oxygen species	67
2.3.12	Mouse xenografts.....	67
2.3.13	Data extraction from cBioPortal	69
2.4	Results.....	69
2.4.1	Spontaneous DNA damage in <i>RBI</i> deficient cancer cells	69
2.4.2	<i>RBI</i> mutant cells have randomly distributed DNA damage	76
2.4.3	Homologous recombination repair defects in <i>RBI</i> deficient cancer cells	80
2.4.4	Increased lung metastases in <i>RBI</i> mutant xenografts.....	88
2.5	Discussion.....	93
2.6	References.....	95
Chapter 3	101
3	An pRB-condensin II complex mediates long-range chromosome interactions and regulates expression at divergently paired genes	101
3.1	Abstract.....	101
3.2	Introduction.....	101
3.3	Materials and Methods	104
3.3.1	Cell culture.....	104
3.3.2	ChIP	104
3.3.3	ChIP-Seq libraries and alignment	105
3.3.4	Peak calling and annotation	106
3.3.5	RNA-Seq.....	107
3.3.6	Comparison of transcriptome analysis and ChIP-Seq data.....	108
3.3.7	qRT-PCR analysis of expression	109
3.3.8	Flow Cytometry	109
3.3.9	Nocodazole treatment	109
3.3.10	Chromosome conformation capture (3C) analysis	110

3.3.11	Circularized chromosome conformation capture (4C)-Seq	112
3.3.12	Gene expression in lung adenocarcinoma.....	113
3.3.13	EZH2 inhibitor treatment.....	114
3.4	Results.....	114
3.4.1	Condensin II and TFIIC occupy promoters in an pRB-dependent manner.....	114
3.4.2	Differential regulation of gene expression by condensin II at bidirectional promoters	125
3.4.3	pRB-condensin II complexes mediate long-range chromosome interactions.....	132
3.4.4	Misregulated gene expression at bidirectional promoters in <i>RBI</i> -deleted lung adenocarcinoma.....	139
3.4.5	Bidirectional promoters insulate epigenetic effects on gene regulation .	142
3.5	Discussion.....	147
3.6	References.....	150
Chapter 4	157
4	Discussion	157
4.1	Summary of findings	157
4.2	Impact on cancer therapy selection.....	158
4.3	<i>RBI</i> hemizygozity in cancer	159
4.4	<i>Rb1^L</i> and haploinsufficiency.....	160
4.5	Chromosome topology and cancer	161
4.6	Lack of cancer progression in <i>Rb1^{L/L}</i> mice.....	163
4.7	Further investigation of chromosome topology in <i>Rb1^{L/L}</i> cells	165
4.8	Multiple possibilities for long-range chromosome interactions facilitated by condensin II	166
4.9	Formation and recruitment of the pRB-condensin II complex.....	167

4.10 Summary of pRB functions in genome stability.....	170
4.11 References.....	170
Appendices.....	176
Curriculum Vitae	181

List of Tables

Table 2.1: Characterization of the U2OS <i>RBI</i> mutant clones generated by targeting exon 22.....	72
--	----

List of Figures

Figure 1.1: Interaction surfaces of pocket proteins and the unique interaction between pRB and E2F1.....	4
Figure 1.2: Depiction of the structural maintenance of chromosomes (SMC) family members cohesin, condensin I and condensin II.	14
Figure 1.3: Chromatin is arranged in the nucleus using hierarchical architecture.....	22
Figure 1.4: Chromatin dynamics affect cellular processes within a nucleus.....	29
Figure 2.1: CRISPR/Cas9 induced mutations in <i>RB1</i> cause DNA damage.....	70
Figure 2.2: Cancer cells with <i>RB1</i> mutations have elevated reactive oxygen species and sensitivity to cisplatin.	74
Figure 2.3: γ H2AX is randomly distributed in the genomes of <i>RB1</i> mutant cells.....	77
Figure 2.4: Defective homology directed repair of DNA breaks in <i>RB1</i> mutant cells.	81
Figure 2.5: Defective repair of γ IR induced DNA damage in <i>RB1</i> knockout cells.	84
Figure 2.6: Increased mitotic errors in <i>RB1</i> mutant cells.....	86
Figure 2.7: Increased lung metastases with <i>RB1</i> mutant cells in xenograft experiments. .	90
Figure 2.8: Single copy loss of <i>RB1</i> can create cancer enabling phenotypes.	92
Figure 3.1: CAP-D3 is enriched at proximal promoters.	115
Figure 3.2: CAP-D3 binds at some promoters in an pRB-dependent manner.....	118
Figure 3.3: CAP-D3 is localized to active promoters and changes in binding are correlated with activation and repression of genes.	121

Figure 3.4: Changes in condensin II and TFIIC binding are correlated with downregulation of <i>Tubb2a</i>	124
Figure 3.5: Condensin II binding is enriched at bidirectional promoters and is correlated with changes in expression.	126
Figure 3.6: Bidirectional promoters with differential expression changes upon loss of CAP-D3 binding in <i>Rb1^{L/L}</i> MEFs.....	128
Figure 3.7: Gene expression at bidirectional promoters is not altered by microtubule inhibitor induced aneuploidy.	131
Figure 3.8: Reduced condensin II and TFIIC localization at the <i>Hist1</i> gene cluster in <i>Rb1^{L/L}</i> MEFs are correlated with upregulation of some histone genes.	133
Figure 3.9: Preservation of 3C-detected chromatin loops in <i>Rb1^{L/L}</i> cells.	135
Figure 3.10: Genome wide significant interacting regions determined for <i>Cdca3/Usp5</i> and <i>Pole/Pxmp2</i> bidirectional promoters.....	137
Figure 3.11: Changes in pRB-condensin II complex binding correlated with altered interaction between the <i>Cdca3/Usp5</i> bidirectional promoter and the <i>Hoxa</i> locus.....	138
Figure 3.12: Altered long-range chromosome contacts in <i>Rb1^{L/L}</i> MEFs.	140
Figure 3.13: <i>RBI</i> deletion in lung adenocarcinoma is associated with increased expression of genes at bidirectional promoters.....	143
Figure 3.14: H3K27me3 localization is not altered as a result of changes in condensin II binding but pRB-condensin II complexes do insulate transcriptional environments at bidirectional promoters.	145
Figure 3.15: pRB-TFIIC-condensin II complexes organize transcriptional environments at bidirectional promoters.	148

Figure 4.1: Additional models of long-range chromosome contacts mediated by
condensin II at bidirectional promoters.168

List of Appendices

Appendix A: Permission for publication by Molecular and Cellular Biology	176
Appendix B: List of antibodies used.....	178
Appendix C: List of plasmids used.....	179

Chapter 1

1 Introduction

1.1 Retinoblastoma onset and the discovery of the retinoblastoma gene

Retinoblastoma, or childhood cancer of the retina, can be inherited due to germline mutation or can occur sporadically due to somatic mutation (Falls and Neel, 1951; Schappert-Kimmijser et al., 1966; Smith and Sorsby, 1958). Because many of the collected pedigrees showed affected sibships, and some of the affected sibs themselves transmitted the disease to their offspring, retinoblastoma was originally theorized to be caused by a single dominant mutant gene that is not completely penetrant (Neel and Falls, 1951; Smith and Sorsby, 1958). Retinoblastoma can occur either as a unilateral disease, affecting only one eye of the child, or as a bilateral disease, affecting both eyes. In 1971, Dr. Alfred Knudson used statistics to analyze clinical data and determined that retinoblastomas arise through a minimum of two mutational events, which later became known as the “two hit hypothesis” (Knudson, 1971). In his study, Knudson realized that bilateral retinoblastoma is generally found in children who have a family history of the disease. It was also noted that children with bilateral cases have an earlier mean age at diagnosis than those with unilateral cases. These observations, therefore, could be explained by his hypothesis; affected children with familial retinoblastoma inherit one mutational event in their genome, making them susceptible to retinoblastoma, but it is not until they acquire a second mutation, or “hit”, that they succumb to the disease. Patients who do not inherit the first mutation, however, need to have two independent mutational events occur to develop retinoblastoma. This hypothesis, therefore, is built on the assumption that there is a gene whose protein product is necessary to suppress retinoblastoma incidence.

Early genetic studies of retinoblastomas revealed that there were occasionally tumors with deletions in chromosome 13, more specifically at region 13q14 (Francke and Kung, 1976; Lele et al., 1963; Sparkes et al., 1980; Yunis and Ramsay, 1978). It was then postulated that the alleles found at this “retinoblastoma locus” were tumor suppressors

and loss of function of these alleles is associated with malignancy (Murphree and Benedict, 1984). This became the prototype to which other proposed tumor suppressor genes were compared. Eventually, in 1986, two independent groups cloned a single gene from this q14 segment of chromosome 13 which was shown not only to be deleted in retinoblastomas, but osteosarcomas as well (Friend et al., 1986; Lee et al., 1987). This gene became known as the retinoblastoma susceptibility gene, abbreviated as *RBI*.

1.2 The retinoblastoma protein is a cell cycle regulator

The first biochemical evidence that the retinoblastoma gene could be important for tumor suppression in other tissues outside of retina and bone came from studies of viral oncoproteins. Adenovirus E1A, simian virus 40 T antigen, and human papilloma virus E7 oncoproteins all require inactivation of the retinoblastoma protein (pRB) to transform cells, which suggested a broader tumor suppressive function of pRB (DeCaprio et al., 1988; Dyson et al., 1989; Whyte et al., 1988; Whyte et al., 1989). Shortly after these discoveries, it was also seen that loss of pRB function leads to loss of cellular proliferation control, as demonstrated in various cancer samples (Bookstein et al., 1990a; Huang et al., 1988; Takahashi et al., 1991). By utilizing human cancer cells with mutant versions of *RBI* and subsequently expressing exogenous wild type *RBI*, cells were shown to have reduced proliferation, colony formation in agar, and tumorigenicity in nude mice, further demonstrating the proliferative control of pRB (Bookstein et al., 1990b; Huang et al., 1988; Takahashi et al., 1991). pRB was also shown to be phosphorylated in late G1 phase just before cells enter S phase, which further suggested pRB may be acting as a cell cycle regulatory element (DeCaprio et al., 1989). Ultimately, pRB was shown to inhibit E2F transcription factors in G1 of the cell cycle, but pRB is unable to do so when it becomes hyperphosphorylated or when it is deregulated by viral oncoproteins (Chellappan et al., 1991; Helin et al., 1992; Hiebert et al., 1992; Nevins, 1992). These studies led to the identification of the G1 checkpoint of the cell division cycle which is controlled by pRB through binding to and inhibiting E2Fs and their target gene transcription (Dyson, 1998).

1.3 The pocket domain defines the pRB family

Two additional proteins were discovered that also have the ability to bind to viral oncoproteins and have sequence homology with pRB (Cobrinik et al., 1993; Ewen et al., 1991; Hannon et al., 1993; Li et al., 1993; Mayol et al., 1993; Zhu et al., 1993). These proteins are p107 and p130, encoded by *RBL1* (retinoblastoma like 1) and *RBL2* (retinoblastoma like 2), respectively. pRB, p107 and p130 are collectively referred to as the “pocket proteins” because of the pocket domain that is found in all three proteins (Figure 1.1A). This pocket domain itself contains a small pocket and a large pocket. The large pocket consists of an unstructured C-terminal domain along with the small pocket. This large pocket is necessary for binding E2F *in vivo*, inhibiting E2F target gene transcription and is also the minimal growth suppressing domain (Figure 1.1B) (Bremner et al., 1995; Hiebert et al., 1992; Qin et al., 1992; Yang et al., 2002). The small pocket can also be further divided into the A and the B subdomains which are separated by an unstructured, flexible linker region (Chow and Dean, 1996). These A and B subdomains each form a single cyclin fold which interact with each other to form the globular small pocket domain (Gibson et al., 1994; Lee et al., 1998). This small pocket is the minimal domain capable of interacting with viral oncoproteins and is sufficient to repress transcription (Chow and Dean, 1996; Chow et al., 1996; Hu et al., 1990; Kaelin et al., 1990; Sellers et al., 1995). Crystal structures revealed that the LXCXE motif found in viral oncoproteins is what binds a shallow groove in the B subdomain, which is now commonly referred to as the LXCXE binding cleft (Figure 1.1B) (Lee et al., 1998). Since then, it has been discovered that a number of cellular proteins bind to the LXCXE binding cleft, many of which are able to alter chromatin structure and act as co-repressors of transcription. Examples of proteins binding to pRB through the LXCXE binding cleft include histone deacetylases (HDAC1 and HDAC2), DNA methyl transferases (DNMT1), histone methyl transferases (Suv39h1), histone binding proteins (HP1), and condensin II (Brehm et al., 1998; Longworth et al., 2008; Magnaghi-Jaulin et al., 1998; Nielsen et al., 2001; Robertson et al., 2000).

Despite all of the structural similarities between the pocket proteins, there are also some key differences. Both p107 and p130 contain an insertion within their B

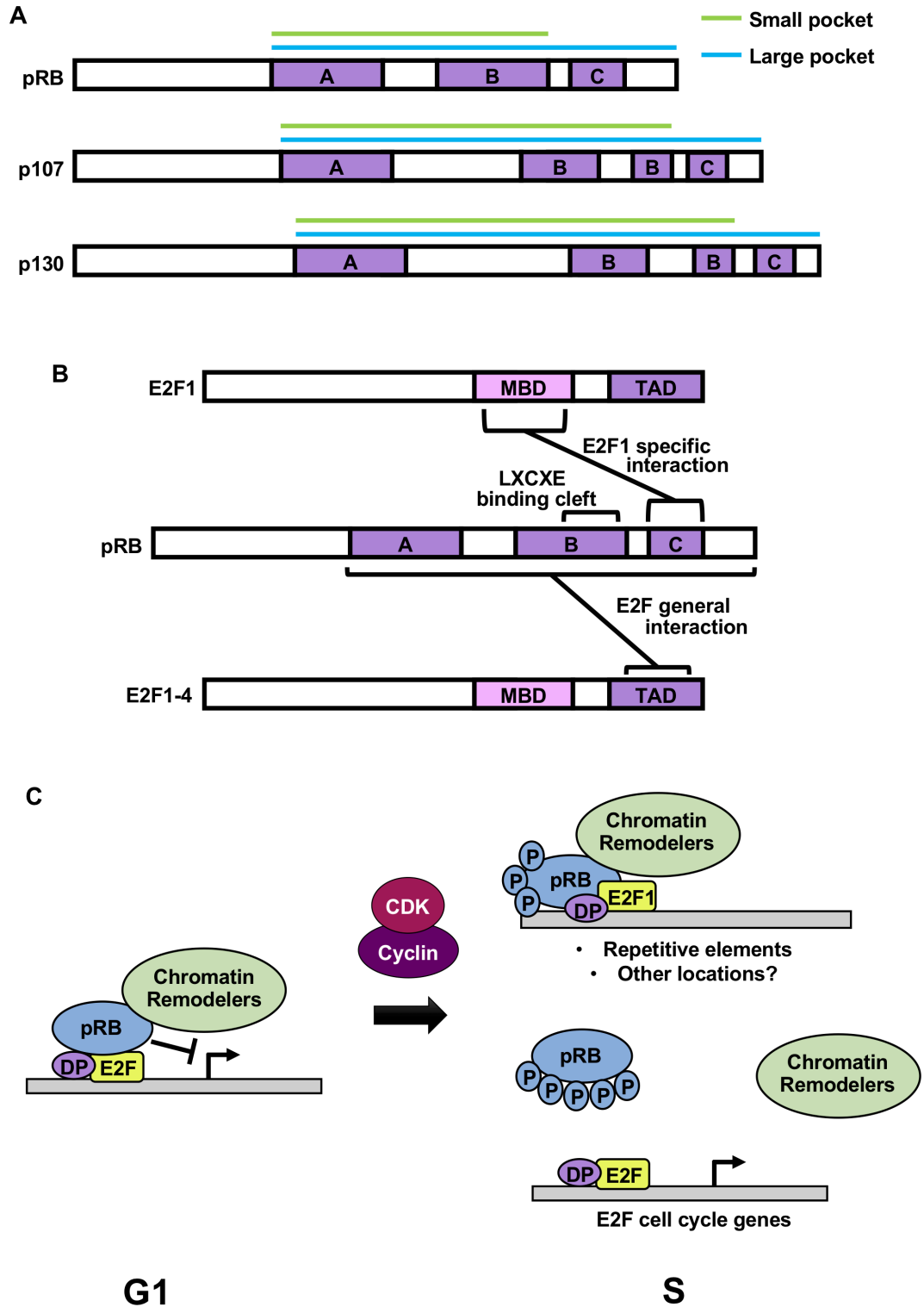


Figure 1.1: Interaction surfaces of pocket proteins and the unique interaction between pRB and E2F1.

Figure 1.1: Interaction surfaces of pocket proteins and the unique interaction between pRB and E2F1.

(A) The A, B and C-terminal subdomains are shown for all three pocket proteins. The large and small pockets are denoted by blue and green lines, respectively. (B) The three different interaction sites within the large pocket of pRB are shown. The E2F1 specific interaction is formed between the E2F1 marked box domain (MBD) and a minimal interaction site in the C-terminus of pRB, while the E2F general interaction is facilitated by the transactivation domains (TADs) of E2F1-4 and the large pocket domain of pRB. The third interaction site, the LXCXE binding cleft, is located in the B subdomain of the pocket. (C) The retinoblastoma protein binds to E2F-DP heterodimeric transcription factors to repress transcription of cell cycle genes during the G1 phase of the cell cycle. pRB also recruits chromatin remodeling proteins to these sites to assist in the formation of repressive heterochromatin. Upon mitogen stimulation during the G1-S phase transition, cyclin and cyclin-dependent kinase (CDK) complexes hyperphosphorylate pRB causing it to dissociate from E2F transcription factors that were bound through the E2F general interaction, leading to the activation of E2F cell cycle genes. However, the E2F1 specific interaction with pRB exhibits resistance to CDK phosphorylation and allows these proteins to continue to associate. This CDK-resistant complex of E2F1 and pRB has been seen to bind at repetitive elements, but other locations in the genome are largely unknown.

subdomains, have longer spacer regions linking subdomains A and B which allows them to interact with cyclin-cyclin dependent kinase (CDK) complexes, and they have a unique N-terminal domain that can inhibit CDKs (Classon and Dyson, 2001). pRB, on the other hand, has a specific E2F1 binding site within the pRB carboxy-terminal domain (Figure 1.1B) (Dick and Dyson, 2003; Julian et al., 2008). Also contained within the C-terminal domain of pRB are binding sites for cyclin-CDK complexes and protein phosphatase 1 (PP1) (Adams et al., 1999; Durfee et al., 1993; Tamrakar and Ludlow, 2000; Vietri et al., 2006).

1.4 The pRB family is recruited to DNA by E2F transcription factors

The first identification of any E2F transcription factor came from studying cell factors that could mediate transcriptional activation of the adenovirus E2 promoter, leading to the name E2 promoter binding factor (E2F) (Kovesdi et al., 1987). Further studies of transcriptional control showed that there were human genes regulated in not only an E2F-dependent manner but also a cell cycle-dependent manner; the first of which were *MYC*, which is involved in cellular proliferation, and the dihydrofolate reductase gene (*DHFR*), which is involved in metabolism of nucleic acids (Blake and Azizkhan, 1989; Hiebert et al., 1989; Thalmeier et al., 1989). This helped form the basis that the growth suppressive role of pRB was mediated through inhibiting E2F target gene transcription, which could become deregulated by viral oncoproteins (Nevins, 1992).

Since the initial discovery, it has become clear that the E2F transcription factors are a family of proteins essential for eukaryotic cell proliferation (Chen et al., 2009; Johnson and Degregori, 2006). The E2Fs can be divided based on their cellular activities; E2F1, E2F2 and E2F3 are generally described as “activator E2Fs” due to their ability to strongly activate E2F transcriptional targets to promote cell cycle progression, while E2Fs 4-8 are classified as “repressor E2Fs” and are believed to be required for cell cycle exit and differentiation (Chen et al., 2009).

All of the E2F members have a highly conserved DNA binding domain, while only E2Fs 1-6 have a highly conserved dimerization domain that is required for

interaction with DP family members (Girling et al., 1993; Morgunova et al., 2015; Wu et al., 1995; Zheng et al., 1999). The DP family of proteins consists of DP1, DP2/3 and DP4, each of which has preferred E2F associations. Notably, DP1 forms heterodimers with the activator E2Fs and together they bind to promoters of genes involved in DNA synthesis and cell cycle progression (DeGregori and Johnson, 2006). E2F7 and E2F8, however, have two DNA binding domains and interact with DNA as homodimers or heterodimers independently of DP family members (Morgunova et al., 2015). E2F1-5 each contain a transactivation domain that allows them to activate transcription of their gene targets, and also contain a pocket protein binding domain (Dimova and Dyson, 2005). Since pocket proteins lack intrinsic DNA binding activity, they rely on the E2Fs that they are able to interact with to be recruited to DNA. The pocket of pRB binds to E2Fs 1-4, while p107 and p130 can both interact with E2F4 and E2F5, and E2F6-8 do not have pocket protein binding domains (Dimova and Dyson, 2005; van den Heuvel and Dyson, 2008).

The pocket of pRB creates a docking site which is responsible for binding to the C-terminal transactivation domain of E2Fs, physically inhibiting transcriptional activation (Figure 1.1B) (Dick and Rubin, 2013). However, because pRB contains distinct binding sites for E2Fs and transcriptional repressor complexes, it can also enable the silencing of gene expression through recruiting chromatin remodeling factors (Giacinti and Giordano, 2006).

1.5 Pocket proteins have overlapping and unique roles

In theory, pRB, p107 and p130 are all capable of forming inhibitory complexes with E2Fs and recruiting co-repressors through their LXCXE binding clefts. However, the pocket proteins work in a concerted effort to assist in maintaining proliferative control as certain pocket protein-E2F complexes predominate in different phases of the cell cycle.

Not only do the pocket proteins have preferred E2F binding partners, but they also have differential expression throughout the cell cycle which also helps to explain their differential activity (Classon and Dyson, 2001). p107 is an E2F target gene; expression of p107 is low in quiescent cells and is most highly expressed in S phase (Xiao et al., 1996;

Zhu et al., 1995). The expression of p130 also fluctuates during the cell cycle, but it is expressed at high levels in quiescent and terminally differentiated cells and its expression decreases as cells progress through G1 and S phases. pRB, on the other hand, is expressed in both proliferating and non-cycling cells and for the most part, expression of pRB is relatively stable throughout the cell cycle (Buchkovich et al., 1989; Classon and Dyson, 2001). This relative stability of pRB expression during the cell cycle indicates that instead of its function being controlled by transcription, pRB is largely regulated by post-translational modifications, chiefly phosphorylation (Dick and Rubin, 2013). Cyclin-CDK complexes are largely responsible for the phosphorylation of pRB, as well as the other pocket proteins, leading to conformational changes which ultimately disrupt various protein interactions by masking binding surfaces (Dick and Rubin, 2013; Giacinti and Giordano, 2006). For pRB specifically, cyclin D-CDK4/6, followed by cyclin E-CDK2 complexes phosphorylate pRB at the G1-S phase transition leading to alleviation of E2F transcriptional repression through pRB dissociation (Figure 1.1C) (Burke et al., 2012; Calbó et al., 2002).

1.6 E2F1 and pRB have a unique interaction interface

Although phosphorylation by cyclin-CDK complexes produces conformational changes in pRB leading to dissociation from most of the activator E2Fs, a unique interaction between pRB and E2F1 is still able to form (Figure 1.1B). This second E2F1 binding site is outside of the A and B subdomains within the large pocket of pRB (Dick and Dyson, 2003). Within E2F1, a region termed the “marked box” domain (MBD), which excludes the transactivation domain, is required for this interaction (Dick and Dyson, 2003). This interaction is unique between pRB and E2F1, as it does not occur between pRB and any of the other E2Fs or E2F1 and any of the other pocket proteins (Cecchini and Dick, 2011; Dick and Dyson, 2003; Julian et al., 2008).

This specific interaction actually shows reduced affinity for a probe with the “typical” E2F consensus sequence and is also not able to repress transcription from luciferase reporters under the control of E2F promoters (Dick and Dyson, 2003; Julian et al., 2008). This alternate pRB-E2F1 conformation is also resistant to classic pRB-E2F dissociation signals (Figure 1.1C). Using this E2F1-specific binding site in pRB,

interactions with E2F1 are resistant to E1A-mediated displacement (Seifried et al., 2008). This pRB-E2F1 complex has also been noted to be resistant to CDK phosphorylation (Cecchini and Dick, 2011).

The pRB-E2F1 specific interaction, therefore, has been noted to play important regulatory roles outside of cell division. One function of this interaction is to allow pRB to regulate E2F1-induced apoptosis (Carnevale et al., 2012; Cecchini and Dick, 2011; Dick and Dyson, 2003; Julian et al., 2008; Seifried et al., 2008). This pRB-E2F1 complex is also important for recruiting condensin II to the genome to stabilize chromatin structure and help facilitate replication of certain repetitive regions of the genome (Coschi et al., 2014; Ishak et al., 2017). Another recently discovered function of this unique pRB and E2F1 interaction is to inhibit repeat element expression in the genome, notably endogenous retroviruses (Figure 1.1C) (Ishak et al., 2016).

1.7 Evidence for a separate role of pRB in tumor suppression outside of cell cycle control

The central defining feature of cancer is uncontrolled cell division. In multicellular organisms like humans, tissue homeostasis requires appropriate control of the cell division cycle. Deregulation of proliferative control is one of the major hallmarks of cancer therefore (Hanahan and Weinberg, 2011). Not surprisingly then, the activities of the pocket proteins are disrupted in many human cancers through alterations of upstream regulators (Burkhart and Sage, 2008; Sherr, 1996). While deregulated CDKs and CDK inhibitors (CKIs) can inactivate all the pocket proteins, only pRB is also seen to be commonly mutated in cancers while the other two pocket proteins, p107 and p130, are rarely directly inactivated (Burkhart and Sage, 2008; Dick and Rubin, 2013; McNair et al., 2018; Sherr, 1996). In addition, there are also some cancers that tend to be enriched for pRB-pathway alterations, while other cancer types appear to have this pathway relatively intact. For example, recent analyses using data from The Cancer Genome Atlas demonstrates that the majority of lung and breast cancers have alterations within the pRB-pathway, resulting from loss-of-function alterations in *RBI*, or gain-of-function alterations in genes encoding D-type cyclins and their associated CDKs, or loss-of-function alterations in the upstream CKI which is encoded by the *CDKN2A* gene (Dick et

al., 2018). Alternatively, there are some cancers like colorectal and prostate carcinomas, where alterations within the pRB-pathway are relatively uncommon (Dick et al., 2018).

Because of its role in regulating the G1-S transition, it is not surprising that deletion of the *RBI* gene prevents cell cycle arrest in response to a variety of signals (Knudsen and Knudsen, 2008). However, there have also been some interesting observations regarding *RBI* status and patient outcomes that have not been observed with mutations of other components of the pRB-pathway. Studies in patients with high-grade serous ovarian carcinoma (HGSOC) and lung adenocarcinoma have demonstrated that *RBI* loss is predictive of longer patient survival in response to chemotherapy (Cecchini et al., 2015; Garsed et al., 2018; Ludovini et al., 2004; Zhao et al., 2012). In these studies where proliferation was also investigated, loss of *RBI* was not correlated with increased proliferation (Cecchini et al., 2015; Garsed et al., 2018). In addition, studies that have specifically looked at more advanced cancers, such as those that have metastasized or are resistant to certain therapies, have observed that *RBI* gene loss may be more prevalent than in less advanced forms of the disease (Beltran et al., 2016; McNair et al., 2018; Robinson et al., 2017; Thangavel et al., 2017). Along these same lines, mutation in *RBI* has also been linked to transdifferentiation as a response to targeted therapeutics. In a study of non-small cell lung cancer (NSCLC) samples with epidermal growth factor receptor gene (*EGFR*) mutations where resistance to tyrosine kinase inhibitors occurred, the lung cancers that transformed to small cell lung cancer (SCLC) invariably lost pRB expression (Niederst et al., 2015). Similarly, prostate cancers that develop resistance to the antiandrogen enzalutamide may do so by switching from an epithelial to a neuroendocrine phenotype, which is enabled in part by loss of pRB expression (Ku et al., 2017; Mu et al., 2017). Overall, these examples all suggest that *RBI* mutation facilitates more than alterations to proliferative control, and loss of *RBI* may confer other cancer relevant characteristics that are independent of classical pRB-pathway functions of proliferative control in cancer.

1.8 Genome instability is an enabling characteristic of cancer

The failure of a cell to have efficient and error-free duplication of chromosomes before appropriately segregating them to daughter cells broadly defines genome instability. The genetic alterations that occur range from point mutations to chromosome rearrangements, so the types of genomic instability can be separated into categories depending on the event that has occurred. Some examples are chromosome instability (CIN), micro- and minisatellite instability (MIN), and instability leading to mutations as well as that leading to rearrangements.

Failures in either mitotic chromosome transmission or the spindle mitotic checkpoint can lead to CIN, defined by gain and/or loss of whole chromosomes or fractions of chromosomes (Draviam et al., 2004; Geigl et al., 2008). MIN results in repetitive DNA expansions and contractions and can materialize due to replication slippage, homologous recombination (HR), or impairment of mismatch repair (MMR) or nuclear excision repair (Aguilera and Gomez-Gonzalez, 2008; Draviam et al., 2004; Yao and Dai, 2014). Genome instability can also lead to mutations like base substitutions, micro-insertions and micro-deletions and these mutations are generally linked to errors in replication, impairment of base excision repair or MMR, or translesion DNA synthesis which is inherently error prone. Gross chromosomal rearrangements, such as translocations, duplications, inversions and deletions, are all generated by DNA breaks which are most commonly the result of replication stress (Aguilera and Gomez-Gonzalez, 2008).

Because cancer cells acquire mutations in the genome to evolve progressively to a neoplastic state, it should come as no surprise that genome instability is an enabling characteristic of cancer (Hanahan and Weinberg, 2000, 2011). In healthy cells, genome maintenance systems continuously monitor the genome to detect and resolve defects in the DNA, which ensures that the rates of spontaneous mutations are low (Salk et al., 2010). However, cancer cells usually have an increased rate of mutations, which is generally due to defects affecting the “caretakers” of the genome, which are genes whose products help to maintain the integrity of the genome (Kinzler and Vogelstein, 1997; Salk

et al., 2010). These caretakers are involved in inactivating or diverting mutagenic molecules before they lead to DNA damage, detecting DNA damage and stimulating the appropriate repair machinery, and directly repairing DNA damage (Ciccia and Elledge, 2010; Harper and Elledge, 2007; Jackson and Bartek, 2009; Negrini et al., 2010).

As previously mentioned, genome instability can be initiated by replication stress (Hanahan and Weinberg, 2011). Replication stress largely occurs during S phase of the cell cycle, which can affect chromosome segregation and lead to unfaithful transmission of genetic information to daughter cells. Endogenous or exogenous sources, including polymerase inhibition, limiting replicative factors, oxidative damage, DNA secondary structures, and RNA/DNA hybrids, can be sources of replication stress (Gelot et al., 2015). Phenotypes resulting from replication stress can be detected in G1 and include p53-binding protein 1 (53BP1) nuclear bodies, micronuclei and aneuploidy. Through use of gene targeted mice, it has been discovered that pRB is part of a complex with condensin II and E2F1 that aids in the maintenance of genome stability (Coschi et al., 2014; Ishak et al., 2017).

1.9 Condensins are central components of mitotic chromosome dynamics

The multisubunit protein condensin II is a member of the structural maintenance of chromosomes (SMC) family. The SMC family also consists of the SMC5/6 complex, cohesin, and condensin I (Losada and Hirano, 2005). These complexes are formed in eukaryotic organisms using six different SMC family members that establish three heterodimers. SMC1 and SMC3 are the core members of cohesin complexes, SMC2 and SMC4 are the fundamental units of the condensin complexes (Figure 1.2), and SMC5 and SMC6 are members of a complex associated with DNA repair and checkpoint responses (De Piccoli et al., 2009; Losada and Hirano, 2005). SMC proteins are central components needed for mitotic chromosome dynamics, regulation of gene expression and more, and as such are evolutionarily conserved in prokaryotes and eukaryotes (Cobbe and Heck, 2004). These SMC subunits are approximately 100 nm long before folding at a “hinge” domain in the middle of the protein, allowing the N- and C-termini to interact and form an ATPase “head” domain (Figure 1.2). These two domains in the protein are separated

by coiled coil domains that form the “arm” of the protein. For the most part, SMC proteins dimerize through their hinge-hinge interactions, although they also interact through head-head binding. While these complexes are chiefly defined based on their SMC subunits, non-SMC subunits are also unique to each type of SMC complex (Skibbens, 2019).

As previously mentioned, both condensin II and condensin I share the SMC subunits SMC2 and SMC4, meaning condensin II is distinguished from the other condensin complex by its unique, non-SMC subunits (Figure 1.2). Two of these subunits, CAP-D3 and CAP-G2, contain HEAT repeats, which are highly degenerate repeating motifs involved in protein-protein interactions and may also form a domain for binding double-stranded DNA (Neuwald and Hirano, 2000; Piazza et al., 2014). Condensin complexes lacking a single HEAT repeat subunit are unable to associate with chromosomes (Piazza et al., 2014). A kleisin family member, CAP-H2, is also found exclusively in condensin II. Kleisin proteins interact with SMC proteins at both N- and C-terminal domains to form a ring-like structure (Schleiffer et al., 2003). Condensin I, like condensin II, also has three unique subunits, which are the HEAT repeat containing proteins CAP-D2 and CAP-G, and the kleisin protein CAP-H.

In terms of their mitotic functions, in general, the SMC complex cohesin is important for proper resolution of sister chromatids while condensins I and II are important for mitotic chromosome condensation. In metazoans, cohesin is found on chromosomes from telophase of one cycle until anaphase onset of the next (Gerlich et al., 2006b). From telophase until replication commencement, cohesin binds DNA more dynamically, which is thought to mediate the role of cohesin in transcription. However, during DNA replication, a pool of cohesin binds DNA more stably where it establishes and maintains sister chromatid cohesion (Gerlich et al., 2006b). During mitosis, cell cycle regulated kinases control the removal of cohesin from chromosomes in a stepwise process, where only some arm and most pericentric cohesin is initially protected from removal (reviewed in (Nasmyth, 2011)). The cohesin that remains primarily at the centromeres during pre-metaphase and metaphase is responsible for holding sister chromatids together; it is this centromeric cohesion that is able to oppose the pulling

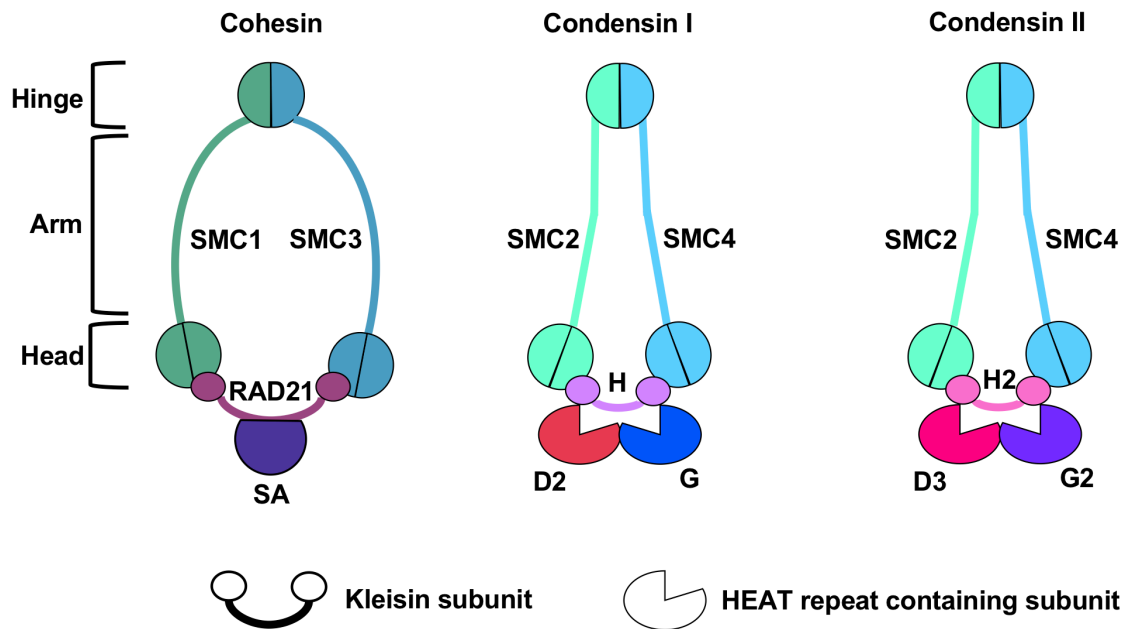


Figure 1.2: Depiction of the structural maintenance of chromosomes (SMC) family members cohesin, condensin I and condensin II.

Cohesin contains the SMC subunits SMC1 and SMC3, while both condensin proteins contain SMC2 and SMC4. Each SMC protein has a head domain which contains the N- and C-terminus of the protein, as well as a hinge domain, which is globular and binds to the hinge domain of another SMC protein. These two domains are separated by the arms of the SMC proteins, which are coiled coils. Non-SMC subunits that comprise cohesin are the kleisin RAD21 and stromal antigen (SA). Condensin I contains the kleisin CAP-H and HEAT repeat containing proteins CAP-D2 and CAP-G, and the counterparts to these proteins in condensin II are CAP-H2, CAP-D3 and CAP-G2, respectively.

forces of the mitotic spindle, which allows for correct chromosome alignment in metaphase. Meanwhile, in prophase, condensin II becomes more stably associated with chromosomes than it was in interphase and facilitates chromosome condensation (Gerlich et al., 2006a). After nuclear envelope breakdown, condensin I gains increased access to chromosomes and assists in further chromosome condensation. The different timing of condensin complexes binding to DNA is very much due to their different cellular locations throughout the cell cycle. Historically, condensin I was thought to be cytoplasmic during interphase and is only able to access chromosomes after nuclear envelope breakdown in prometaphase (Ono et al., 2004). However, some recent evidence has suggested that a small amount of condensin I is present in the nuclei of interphase cells, although the majority does remain cytoplasmic (Li et al., 2015; Zhang et al., 2016). Condensin II, on the other hand, is mostly nuclear throughout interphase and is loaded, to some extent, on DNA throughout the cell cycle (Gerlich et al., 2006a; Ono et al., 2004). At anaphase onset, after cohesin has been completely removed from chromatin, mitotic chromosomes are pulled to opposite poles of the cell (Gruber et al., 2003; Losada and Hirano, 2005). It is only after chromosomes have reached opposite poles that condensin I is unloaded from chromatin and condensin II has a weaker chromatin interaction (reviewed in (Jeppsson et al., 2014) and (Losada and Hirano, 2005)).

1.10 Condensin I and condensin II are differentially loaded onto mitotic chromosomes

Condensins have been thought to mediate linkages between chromatin fibers to organize them into higher order structures during chromosome segregation in mitosis (Cuylen and Haering, 2011). Through the use of global mapping of chromosomal interactions within mitotic chromosomes, it has been observed that when condensin is inactivated, loop formation and compaction are drastically disrupted (Gibcus et al., 2018; Kakui et al., 2017; Schalbetter et al., 2017). In human tissue culture cells, siRNA knockdowns of components of condensins I and II demonstrated that these complexes have different functions during mitosis (Hirota et al., 2004; Ono et al., 2004; Ono et al., 2003). Condensin I is primarily involved with lateral compaction of chromatin, whereas condensin II is chiefly responsible for its axial shortening (Shintomi and Hirano, 2011).

How this is accomplished has been revealed by further study of the individual condensin complexes; condensin II facilitates longer loops and lays the foundation for the axis of mitotic chromosomes, while condensin I facilitates shorter loops which thickens mitotic chromosomes (Green et al., 2012; Walther et al., 2018). Much of their action during mitosis can be attributed to their localization, where condensin II is centrally confined at the chromatid axis at the bases of chromatin loops and condensin I binds to the central chromatid axis but also reaches more peripheral sites (Gibcus et al., 2018; Walther et al., 2018). Condensin complexes have ATP-dependent DNA supercoiling activity and can also assist in the reannealing of complementary DNA strands (Kimura et al., 1998; Sakai et al., 2003; St-Pierre et al., 2009). It has been inferred that both these actions of condensin promote the assembly of mitotic chromosomes (Akai et al., 2011; Kimura and Hirano, 2000; St-Pierre et al., 2009; Sutani et al., 2015).

Phosphorylation is used to regulate condensin functions, and this mark is increased on condensins in mitosis (Abe et al., 2011; Piazza et al., 2013; Takemoto et al., 2004). Many phosphorylation sites have been identified on all of the condensin subunits in human cells through the use of phosphoproteomic studies, and some of these have been identified as crucial for proper condensin loading onto mitotic chromosomes (reviewed in (Kagami and Yoshida, 2016)). In addition, some studies have identified “recruiter” proteins that regulate mitotic chromosomal localization. One of these recruiter proteins for condensin II is protein phosphatase 2A (PP2A), which performs this task independent of its phosphatase activity (Takemoto et al., 2009). In addition to recruiter proteins, there also appear to be “receptor” proteins that may also be needed for condensin loading. CAP-G2 and CAP-D3, the HEAT repeat containing subunits of condensin II, for example, can interact with H4K20me1, a mark which accumulates during early mitosis (Liu et al., 2010). Altogether, these studies suggest that loading of condensin onto mitotic chromosomes requires both recruiter and receptor factors, and that phosphorylation of various condensin subunits may promote their association with these recruiters and/or the receptors. This also hints that condensin II binding to interphase chromosomes similarly requires a set of recruiter and receptor factors, some of which have been investigated in this thesis.

1.11 Condensins can translocate DNA and extrude DNA loops

Unlike cohesin which has open V or ring-like structures, early analysis of condensin structure generated images of lumen-less rod-like structures (Anderson et al., 2002; Melby et al., 1998; Yoshimura et al., 2002). Both yeast and chicken condensins have been subjected to chemical crosslinking followed with X-ray crystallography or mass spectroscopy, respectively, which have demonstrated a number of intermolecular crosslinks along the entire length of SMC2 and SMC4 coiled coil domains and is further evidence of a closed rod-like structure (Barysz et al., 2015; Soh et al., 2015). Interestingly, crosslinks between SMC ATPase heads have only been detected in ATPase mutant versions of the proteins, suggesting that during cycles of ATP binding and hydrolysis, SMC heads may adopt different conformations or may separate and re-close (Barysz et al., 2015; Diebold-Durand et al., 2017). However, another study examining only SMC2/SMC4 dimers using high-speed atomic force microscopy (AFM) in liquid never observed rod-like structures (Eeftens et al., 2016). Instead, this analysis showed that the SMC2/SMC4 coiled coils are highly flexible and the configuration of the dimers is dynamic over time. In fact, not only were these researchers able to see the ATPase heads of the dimers engaging and disengaging over time in a dynamic manner, but they also saw the heads dynamically interacting with the hinge (Eeftens et al., 2016). It is possible that a more rigid structure was not observed because the non-SMC subunits may impact coiled coil interactions within the complex. Resolving the structure and dynamics of the fully assembled and functional condensin complex, therefore, is important and will help resolve how condensin interacts with DNA.

Additional insights into how condensins may compact DNA have come from the finding that they are DNA translocases. In yeast, single-molecule imaging of condensin revealed that it is an ATP hydrolysis-dependent molecular motor (Terakawa et al., 2017). This assay used double-tethered DNA and showed condensin movement is largely unidirectional, although there was no initial preference in direction. Further examination also demonstrated that translocating condensins were able to bind and transport a separate DNA molecule, which confirmed a key expectation of the loop extrusion model, a

hypothesis that SMC protein complexes processively increase the size of loops in the genome (Terakawa et al., 2017).

Through directly visualizing DNA compaction by yeast condensin in real time, the loop extrusion model was able to be investigated more directly. In this study by Gangi et al., a double stranded DNA molecule had both its ends tethered to a substrate, close enough together that the DNA was not tauged and initially flexible. This DNA was visualized through staining and when fluorescently labelled condensin and ATP were introduced, DNA loops formed that stemmed from the site where condensin was bound (Ganji et al., 2018). These experiments also demonstrated that loop extrusion is asymmetric; it was seen that the distance between the loop region and one side of the DNA outside of the loop progressively got smaller as the loop increased in size, while the DNA content on the other side of the loop remained fixed. This suggests that one domain of condensin remains bound to DNA whereas another domain of condensin is able to actively drive DNA migration, which ultimately leads to loop extrusion.

To fully amalgamate data from condensin structure and its translocation properties into a functional model, step size analysis also needs to be accounted for. One set of experiments, where a single end of a DNA molecule was tethered to a glass substrate and the other end of DNA was attached to a streptavidin-coated magnetic bead, demonstrated that when condensin and ATP were introduced, bead migration, and thus DNA compaction, had average step sizes of between 177 and 210 nm (Keenholtz et al., 2017). This step size is approximately four times the length of the 50 nm long condensin complex. Another study also found that condensin can induce steps of hundreds of nanometers in size, and this data also peaks at approximately 200 nm (Eeftens et al., 2017). One explanation for larger step sizes is that condensin might translocate along DNA through a myosin- or kinesin-like “walking” mechanism where condensin would use cycles of ATP hydrolysis to yield head-over-head step sizes of up to 100 nm. Another possibility is that condensin may oligomerize to produce these larger step sizes, and although condensins exist in both monomeric and oligomeric complexes, condensin monomers are typically studied for detailed descriptions (examples are (Barysz et al., 2015; St-Pierre et al., 2009)).

However, one study that used size exclusion chromatography (SEC) to purify condensin studied both the monomeric form of the protein, which is approximately 600 kDa, as well as an earlier elution peak of approximately 2 MDa, which corresponds in mass to multimeric forms of condensin (Keenholtz et al., 2017). After validating all subunits of yeast condensin were present in both SEC peaks, further studies were conducted to compare and contrast the two states of condensin. Of note, condensins purified from the multimeric position were added to the flow cells described above, where a single end of a DNA molecule was tethered so extension could be tracked in real time from the other DNA end. When multimeric condensin was added to flow cells with ATP the fraction of the DNA compacted was approximately 70%, whereas this was reduced to 40-50% when monomeric condensin was used. This demonstrated, therefore, that the multimeric fraction of condensin leads to more robust DNA compaction. These differences could not be explained by differential abilities to bind ATP, different kinetics of the reactions, or different dwell times between addition of protein and ATP and the onset of compaction (Keenholtz et al., 2017).

Further evidence that condensins may oligomerize comes from the bidirectional translocation of *Bacillus subtilis*. In *B. subtilis*, the SMC-condensin complex is necessary to resolve and segregate newly replicated sister origins (Gruber et al., 2014; Wang et al., 2014). To do so, condensin is recruited to the origin by ParB, the widely conserved partitioning protein, which itself binds at centromeric *parS* sites adjacent to the replication origin (Breier and Grossman, 2007; Gruber and Errington, 2009; Sullivan et al., 2009). After showing that condensin stimulates DNA juxtaposition, or “zip-up”, that initiates at *parS* and progressively accumulates down the flanking DNA, the rate of DNA incorporated into the zip-up from ectopic *parS* sites was measured on either side of the chromosome (Wang et al., 2017). At one of these *parS* sites, one side of the DNA contained many highly transcribed genes, including an operon encoding for abundant ribosomal proteins and translation factors, while the sequence flanking the other side of the *parS* site was largely devoid of active transcription. Movement of condensin rings through the operon was impaired compared to movement in the direction that was free of active transcription, leading to asymmetric migration of condensin from the *parS* site. In other words, convergent transcription on one DNA duplex impedes condensin movement

along that DNA track, but on the partner DNA duplex, the rate of condensin movement is not affected. Along with the observation of asymmetric enrichment of SMC on either side of the ectopic *parS* sites, this is evidence for a model in which two condensin rings each encircle a single DNA duplex on either side of *parS* sites such that these partner DNA duplexes are tethered together by condensin “handcuffing”, or oligomerization.

Together, these studies have greatly advanced our knowledge of how condensin complexes are interacting with DNA. However, it is still not clear how condensins translocate DNA and extrude DNA loops within a dynamic DNA landscape. This makes it exceptionally difficult to determine exactly how condensin II interacts with its binding sites within the genome; is the same condensin II molecule able to interact simultaneously with two distant sites in the genome, or do condensin II molecules interact as oligomers between these sites, or do both scenarios occur, for instance.

1.12 Cohesins are important for organizing the interphase genome

Cohesins used to be chiefly thought of in terms of their role in the trans-tethering needed for sister chromatid segregation during mitosis (reviewed in (Jeppsson et al., 2014; Marston, 2014; Morales and Losada, 2018; Rudra and Skibbens, 2013)), however, two additional roles for cohesin were recognized based on research originally performed in yeast and *Drosophila*. First, mutations in cohesin not only result in defects in sister chromatid cohesion but also in drastic chromosome condensation defects, indicating that cohesins are important for cis-based DNA looping (Guacci et al., 1997). The second added role for cohesins is in transcriptional regulation; through DNA looping, cohesins facilitate communication between distal DNA regulatory elements, like enhancers and promoters (Rollins et al., 1999). Together, these studies have prepared the groundwork for current models of how cohesins generate higher-order chromatin structures which are indispensable for accurate transcription regulation. Because of the important role cohesins play in chromatin structure, cohesin pathways are essential for human development and mutations can have detrimental outcomes. Two cohesin based developmental disorders, for example, Robert syndrome (RBS) and Cornelia de Lange syndrome (CdLS), have common developmental defects including cleft palate, microcephaly, limb reduction

abnormalities, syndactyly and acute cognitive impairment (Skibbens, 2019). It is anticipated that the number of cohesin-related disorders, or “cohesinopathies”, which are more recently and generally classified with other “transcriptomopathies”, will increase considerably over time due to the range of tissues impacted by cohesin mutation as well as the genome-wide effect of cohesins on transcription (Banerji et al., 2017b; Yuan et al., 2015).

To impart transcriptional effects, cohesins organize the interphase genome through balancing the dynamic formation of two states: topologically associating domains (TADs) and compartments (Figure 1.3). For the most part, TADs are described as DNA loops of up to one megabase of DNA which usually have CCCTC-binding factor (CTCF) and cohesin binding at the loop base (Figure 1.3C&D). One general theory is that TADs insulate looped and non-looped DNA from transcriptional and chromatin-remodeling machinery as these factors migrate along DNA. Compartments, on the other hand, range in size from approximately 5 to 50 megabases of DNA and are untethered but self-interacting domains of transcriptionally active (open or “A” domains) or repressed (closed or “B” domains) chromatin states (Figure. 1.3B) (Haarhuis et al., 2017; Lieberman-Aiden et al., 2009; Rao et al., 2017; Schwarzer et al., 2017). At the moment, one theory of how genomic DNA is assimilated into compartments is that DNA-binding proteins like transcription factors and coactivators condense into high-concentration clusters in the nucleus. The condensation of these factors, mediated by low-complexity disordered regions in these proteins, sometimes leads to condensates forming droplets, producing a transient liquid phase separated from the rest of the nucleus (Plys and Kingston, 2018). Another model is that clustering may also emerge due to the interaction of histone modifications and transcription factors that possess similar characteristics (Haarhuis et al., 2017; Rao et al., 2017; Schwarzer et al., 2017). Compartments are normally dynamic and mobile, although they can become anchored. For example, repressed B compartments tend to be localized more to the periphery of the nucleus and may interact with the nuclear lamina (Hansen et al., 2018).

Although mice with mutations in cohesin and its regulators show some similarities to human cohesinopathies, the phenotypes can be mild as a result of extensive

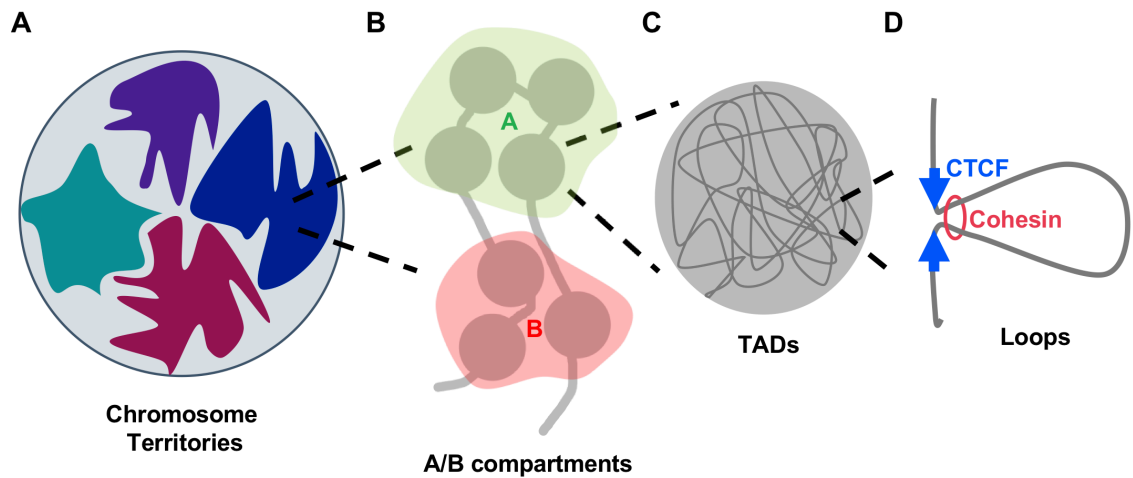


Figure 1.3: Chromatin is arranged in the nucleus using hierarchical architecture.

(A) The largest domain of chromatin architecture within the nucleus is chromosome territories, in which each chromosome occupies a distinct nuclear space. (B) Chromosome territories can be split into A and B compartments, which are transcriptionally more active or inactive, respectively. A compartments make more contacts than B compartments, and interchromosomal contacts between domains from the same compartments (A/A and B/B) are more frequent than those between different compartments (A/B). (C) Within compartments, there are topologically associated domains (TADs) which are largely defined by their interaction frequencies; two regions within a TAD associate on average more frequently with each other than regions outside of the TAD. (D) Inside TADs, loop formation occurs, particularly between enhancers and promoters. The most well-defined boundary factor and loop extrusion factor, respectively, are CTCF and cohesin, although other proteins have been and continue to be investigated.

upregulation of the wild type copy in heterozygotes; an example is the mutant mouse of one of cohesin's interacting partners, *Nipbl* (Kawauchi et al., 2009). One model organism that has been useful in studying cohesin's effects on transcription and development is zebrafish, largely through the use of antisense morpholino oligonucleotides (MOs), which allow not only for a greater range of gene knockdown but also the ability to delete multiple cohesin associated proteins simultaneously (Muto and Schilling, 2017). For example, the effects of cohesin on an important developmental process, the transition from maternal to zygotic control of transcription, or zygotic genome activation (ZGA), was successfully investigated in zebrafish. It was discovered that cohesins may be important for maternal mRNA turnover and ensuing expression of zygotic mRNAs, as embryos that have reduced levels of cohesin's kleisin protein, Rad21, have higher levels of maternal mRNAs (Meier et al., 2018). At post-ZGA stages, chromatin immunoprecipitation followed by sequencing (ChIP-Seq) revealed thousands of new genomic locations that recruit Rad21, many of which are promoters and transcriptional start sites (TSSs) also occupied by H3K4me1, H3K4me3 and/or H3K27ac histone modifications, indicative of active promoters and enhancers. However, transcripts from many genes in post-ZGA are differentially expressed upon Rad21 depletion, many more than those that are directly bound by Rad21 (Meier et al., 2018). Although it is likely that some of the bound genes may be regulated directly, as seen in other studies (Banerji et al., 2017a; Tsai et al., 2018; Yan et al., 2013), a larger fraction of genes in post-ZGA appears to be regulated indirectly (Meier et al., 2018). Therefore, this observation indicates that cohesin influences transcription through multiple mechanisms, including gene-specific regulation through local chromatin structures as well as more global effects, likely established by TADs and compartments.

Early ChIP studies in yeast showed that the density and distribution of cohesin along DNA correlate with the size of DNA loops (Blat and Kleckner, 1999; Glynn et al., 2004; Guacci et al., 1997). Since then, many studies in a variety of model organisms have shown cohesin plays an active role in TAD formation. In mouse zygotes, TADs are prevalent even at the one-cell stage, and when the gene that encodes Rad21 is knocked-out, TADs are largely absent (Gassler et al., 2017). TADs were also greatly reduced in both HeLa cells as well as in HCT-116 human colorectal carcinoma cells when RAD21

was degraded (Rao et al., 2017; Wutz et al., 2017). Re-expression of RAD21 lead to rapid reassembly of TADs, in particular near super enhancers, which are regions of the genome that contain a very high density of enhancer elements and elevated levels of H3K27ac (Rao et al., 2017).

Like with condensins, ATP is also important for cohesin-based looping and it is generally thought that cohesins also use loop extrusion to compact DNA (Ganji et al., 2018; Vian et al., 2018). The length of DNA loops seems to be defined based on cohesin binding and colocalizing with CTCF at the base of TADs (Dixon et al., 2012; Rao et al., 2014; Rao et al., 2017; Wutz et al., 2017). It is thought that when cohesin encounters CTCF, cohesin pauses or stops loop extrusion and this is what defines loop lengths. Further evidence of this is the fact that cohesin cannot translocate *in vitro* past DNA-bound CTCF (Davidson et al., 2016). In addition, degradation of CTCF does not abolish TADs but instead alters the length of TAD loops, further supporting the model that CTCF acts as a boundary for cohesin and in doing so defines DNA loop length (Wutz et al., 2017).

While cohesins are critical for the formation of TADs, they antagonize the formation of compartments. Depletion of the cohesin component Rad21 in both mouse zygotes and human colorectal carcinoma cells led to increased compartmentalization of active and inactive chromatin (Gassler et al., 2017; Rao et al., 2017). The opposite experiment can also be performed, where effects on compartmentalization can be assayed after depleting Wapl, a protein required for the release of cohesin from interphase chromatin (Kueng et al., 2006). Indeed, through increasing cohesin residency on chromatin through Wapl depletion, genomic compartmentalization was diminished (Gassler et al., 2017; Haarhuis et al., 2017; Wutz et al., 2017).

Altogether, these experiments demonstrate that cohesin is a key determinant of chromatin structure in interphase cells. At regions bound by cohesins, DNA loops are formed; the sizes of which seem to be regulated largely by CTCF. In regions devoid of cohesin and CTCF, however, DNA melds into untethered, epigenetically defined compartments of either transcriptionally active or repressed chromatin states.

1.13 Condensin II binding has been associated with increased boundary strength

Unlike compartments, whose interactions can have noticeable variability from cell-to-cell, TAD structure is largely consistent between various cell types and are also highly conserved across species (Dekker and Mirny, 2013; Dixon et al., 2012). This suggests the sub-megabase scale organization of chromosomes into TADs may be a conserved, bottom-up arrangement of chromatin organization and subsequently genome function. As a result, identifying the complete mechanism of how TADs are established and maintained has become an important area of investigation.

Despite the fact that most TAD boundaries have an enrichment of CTCF, only 15% of CTCF binding sites are located within boundary regions, suggesting that most CTCF binding sites are unrelated to the formation of TAD borders (Dixon et al., 2012). This also indicates that additional architectural proteins may complement CTCF binding to help with the formation of TAD borders, and this may help differentiate border-associated from non-border-associated CTCF binding sites.

A study performed in *D. melanogaster* discovered that the number of architectural proteins present at a TAD border directly correlates with the strength of the topological domain structure (Van Bortle et al., 2014). One of these architectural proteins investigated was the multisubunit RNA polymerase (Pol) III transcription factor TFIIC, consisting of TFIIC220, 110, 102, 90 63 and 35, and was previously shown to also bind at Pol III-independent regions, called extra TFIIC (ETC) loci (Teichmann et al., 2010; Van Bortle et al., 2014). In mice, ETC sites were seen to be close in proximity to CTCF and also have an enrichment of cohesin binding (Carriere et al., 2012), and in budding yeast, ETCs were observed to underlie condensin binding to chromosomes (D'Ambrosio et al., 2008). Therefore, the binding of cohesin, condensin, TFIIC and CTCF and their relation to TADs were investigated further (Van Bortle et al., 2014). Cohesin and condensin II complexes were seen to be enriched at the borders of TADs, along with TFIIC and CTCF (Van Bortle et al., 2014). In mouse embryonic stem cells (mESCs), a strong overlap in CTCF, cohesin, TFIIC and condensin II binding was observed, as well as a strong correlation between clustering of these architectural proteins and chromatin

organization. From this work, it was proposed that the range of TAD border strengths, which is associated with differences in protein occupancy, indicate the role of architectural proteins in long-range interactions. For example, combined binding of the architectural proteins TFIIC and CTCF, along with SMC family proteins cohesin and condensin II, may increase the propensity of all of these proteins to interact with one another and also increase the strength of interactions with other regulatory elements, supported by synergistic protein-protein and protein-DNA interactions. Fewer or less stable interactions of these architectural proteins at TAD boundaries, therefore, can lead to inter-TAD interactions resulting in weaker TAD boundaries (Van Bortle et al., 2014).

As previously mentioned, without cohesin, topological domains largely dissolve (Gassler et al., 2017), suggesting that cohesin is one of the main protein complexes responsible for this degree of chromosome organization. What role, then, is condensin II playing at these boundary sites? It was recently suggested that TFIIC serves as a binding factor for condensin II within the mammalian genome, and that the more abundant condensin II-TFIIC binding sites are within a boundary region, the stronger the insulation at this TAD boundary tends to be (Yuen et al., 2017). However, binding of condensin II at these boundaries does not appear to affect the architecture of individual TADs, but rather, the interaction between TAD boundaries to form compartments (Yuen and Gerton, 2018; Yuen et al., 2017). This idea is further strengthened by the fact that compartments are not lost upon cohesin loss, suggesting that additional proteins, like condensin II, could regulate this higher level of chromosome organization (Rao et al., 2017; Wutz et al., 2017).

Another recent study used Oligopaint FISH technology, which uses oligo libraries as a renewable source of FISH probes, to investigate compartment, or chromosome territory, organization in *Drosophila* cells (Beliveau et al., 2012; Rosin et al., 2018). Depletion of individual condensin II components in these cells greatly increased the amount of contact and intermixing between every tested pair of chromosome territories (Rosin et al., 2018). Conversely, increasing condensin II levels increased interphase compartment partitioning and chromosome compaction, leading to smaller “super territories” that are further separated in three-dimensional space. Overall, this study found

condensin II is essential for long-range chromatin interactions and offered a model in which these contacts are established by condensin II early in interphase to act as an “organizational bookmark” by giving precedence to intra-chromosomal folding directly after mitotic exit (Rosin et al., 2018).

Overall, although a defined role for cohesins in the establishment and maintenance of TADs in interphase cells has been recognized, further investigation into how condensin II affects topology is necessary. The molecular basis of compartment formation is not currently well-defined, but research is indicating that condensin II may play a role. To better understand how SMC complexes shape and preserve genomes, it is important to continue investigating how SMC complexes influence the three-dimensional topology of the genome, gene expression, and chromosome transmission in various situations. Because functions within the nucleus are spatially organized, altered chromatin dynamics can lead to modified transcription and genome instability (Figure 1.4). In this thesis, I investigate the role of condensin II at bidirectional promoters and its effects on chromosome contacts and gene expression.

1.14 Condensin II has functional roles outside of mitosis

Beyond the more recent studies linking condensin II to global chromatin topology, research in various model organisms have proposed condensin II mediates a variety of processes within the cell, again not related to mitosis. For example, polytene chromosomes, or maternal and paternal chromatids all aligned in register, are disassembled in *Drosophila* ovarian nurse cells during mid-oogenesis (Dej and Spradling, 1999). Condensin II is required for this disassembly, and was also found to be a negative regulator of transvection, a process in which certain mutant alleles are transcriptionally influenced through association with their allelic counterparts (Hartl et al., 2008). A newer study examined homologous chromosome pairing at high resolution using HiC, a technique which uses high throughput sequencing after chromosome conformation capture (3C) (Rowley et al., 2019). After examining homologous chromosome pairing after CAP-H2 knockdown, a general increase in pairing signal was detected, in particular within A compartment, or open chromatin, domains. Overall, this study verified that

condensin II either inhibits pairing of homologous chromosomes, or may be actively unpairing chromatin within A domains, at bound genomic regions (Rowley et al., 2019).

Another study in *Arabidopsis thaliana* investigated increased boron toxicity in plants with mutations in either the CAP-G2 and or the CAP-H2 subunit of condensin II (Sakamoto et al., 2011). The major cause of B toxicity in *Arabidopsis* is DNA damage, and condensin II is required for tolerance to the induction of DSBs, as well as to replication blocks induced by zeocin and aphidicolin. This implies that condensin II maintains genomic stability through reducing DNA damage and two possible mechanisms for how this occurs were suggested. The first was that condensin II might physically protect chromatin from genotoxic stresses, inferred from studies of the sole yeast condensin, which responds to nutrient starvation by compacting the genome in the nucleolus to stabilize it (Tsang et al., 2007a; Tsang et al., 2007b). The other possibility raised was that condensin II might be involved in the repair of DSBs and damaged replication forks, pointing towards a role of condensin II in homologous recombination repair (Sakamoto et al., 2011). Another piece of evidence that condensin II is important for homologous recombination came from a study in human cells, where a complex of condensin II and microcephalin/MCPH1 was proposed to mediate homologous recombination repair (Wood et al., 2008).

Prior to its canonical role in mitosis, condensin II has also been implicated in the resolution of sister chromatids in S phase, as seen in HeLa and a lymphoblastoid cell line (Ono et al., 2013). When condensin II is depleted from cells treated with low doses of aphidicolin to induce mild replication stress, chromatid axes have a very fragile appearance and sister chromatids are not distinct from each other. This ultimately leads to failure of the chromosomes to segregate. It is thought that under conditions of replication stress, the activation of dormant origins leads to the need to organize and resolve smaller chromatin loops, which could be a more challenging task for condensin II during S phase and ultimately manifests as structurally distorted axes in metaphase (Ono et al., 2013).

Recent research has also shown both condensin I and condensin II are simultaneously loaded at estrogen-responsive, enhancer RNA-positive (eRNA⁺) enhancers to promote eRNA transcription, enhancer-promoter looping, and full activation

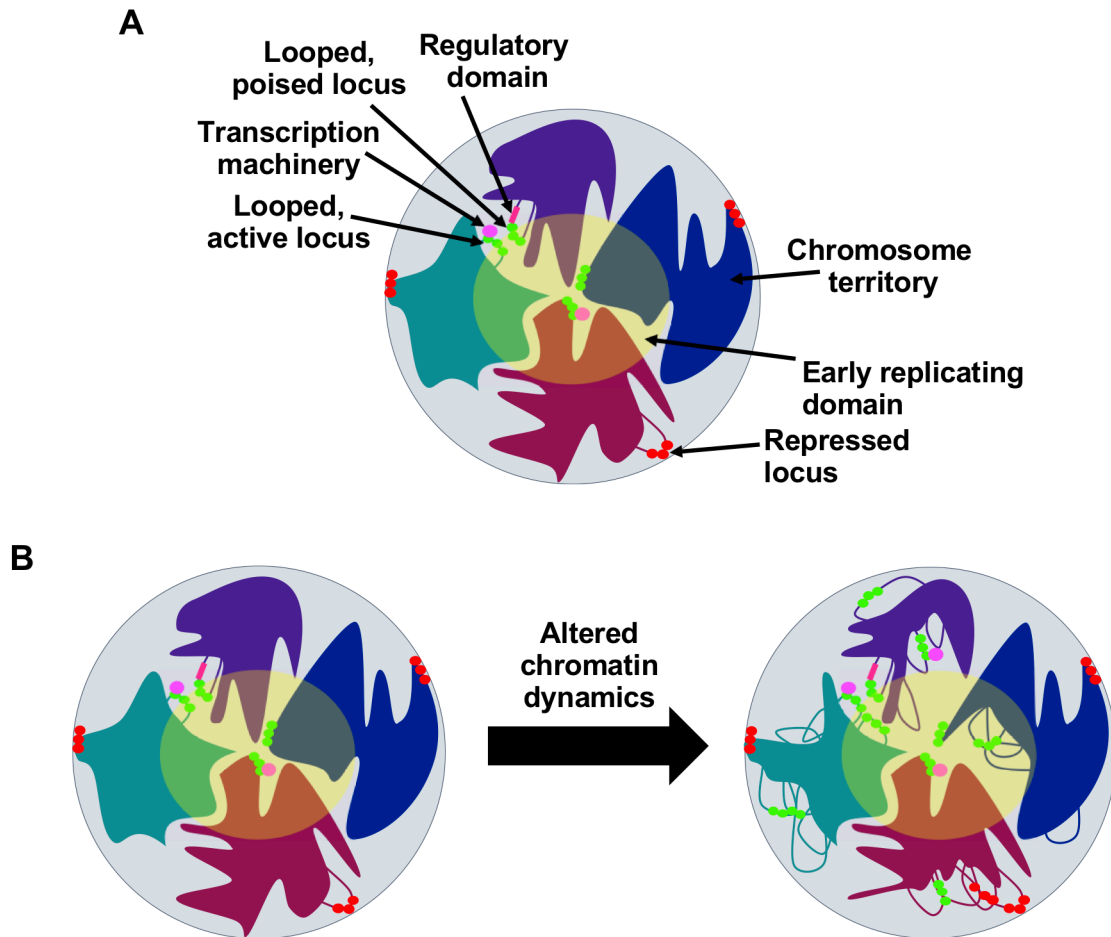


Figure 1.4: Chromatin dynamics affect cellular processes within a nucleus.

(A) Schematic diagram of the locations of replication and transcription domains. Looped poised and active loci tend to be close in three-dimensional space so transcriptional machinery can be shared. Similarly, early replicating regions also tend to cluster together and more inactive loci tend to be found near the nuclear lamina. **(B)** Altered chromatin dynamics can lead to altered transcription and genome instability. The nucleus on the right has many more looped poised and active loci, which can lead to an increase in transcription. Also, less DNA found within the early replicating domain can lead to delayed replication of certain parts of the genome, which can result in replication stress.

of the enhancers themselves (Li et al., 2015). Condensins are able to control the activation of ER- α -bound enhancers through maintaining a fine balance of co-activators and co-repressors, which in turn license RNA Pol II binding, eRNA transcription and ultimately complete activation of coding target genes. When condensin subunits are knocked down, the intensity/stability of enhancer-promoter loopings are reduced. This could be completely or partially due to reduced eRNA levels, as there is also the possibility that condensins could directly control higher-order chromatin architecture as well (Li et al., 2015).

Altogether, these studies highlight that condensin II has functional roles outside of its well-known role in mitosis. Many times, however, it is difficult to determine exactly how condensin II contributes to each process because of the complex inner-workings of the nucleus; it can be challenging to parse out exactly which alterations are leading to changes in phenotype, and if they are acting independently, synergistically, or additively, for example. In addition, some of the paradigms that condensin II is important for in one species may not translate to other organisms. This lack of clarity regarding more universal condensin II functions from model organisms with specialized processes is a strong motivator to study the roles of condensin II in mammalian organisms further.

1.15 Mutations in condensin components have been linked to human disease

Condensins have also recently been linked to some human diseases. One study performed in patients with extreme microcephaly discovered biallelic mutations in *NCAPD2* and *NCAPG*, genes for the CAP-D2 and CAP-G components of condensin I, respectively, as well as *NCAPD3*, the condensin II CAP-D3 subunit (Martin et al., 2016). All of these discovered mutations, however, are functionally hypomorphic, consistent with the essentiality of condensins in cell division (Hagstrom et al., 2002; Hudson et al., 2003; Oliveira et al., 2005; Ono et al., 2003; Siddiqui et al., 2006; Strunnikov et al., 1995). Parallel studies performed in a mutant mouse with hypomorphic loss-of-function *NcapH2*, the gene encoding the CAP-H2 subunit of condensin II, confirmed that mutation of condensin components also causes microcephaly in mice. In both the mouse model and affected patient fibroblasts, anaphase bridges were observed which led to the formation of

micronuclei and aneuploidy in these cells. Interestingly, one patient which had heterozygous variants of *NCAPD3* died of malignant anaplastic medulloblastoma at 11 years.

This was not the first time that a link between condensins and cancer was established, however. A study from a year prior used an algorithm they developed in a pan-cancer analysis to investigate mutated networks in 12 different cancer types from The Cancer Genome Atlas (TCGA) (Leiserson et al., 2015). From this analysis, a subnetwork of condensin I members, *NCAPD2*, *SMC2*, and *SMC4*, was found to be significantly mutated in bladder cancer. Additionally, a subnetwork of condensin II members, *NCAPD3*, *NCAPG2*, and *NCAPH2*, was significantly mutated in both lung adenocarcinoma and lung squamous cell carcinoma (Leiserson et al., 2015). Although these mutations were rare, this does suggest that condensins have a tumor suppressive role. Further evidence of this comes from a gene expression characterization study in prostate tumors, where *NCAPD3* expression was associated with decreased tumor recurrence after surgery (Lapointe et al., 2008). In addition, in the same hypomorphic loss-of-function *Ncaph2* mouse model utilized in the previously mentioned microcephaly study, defects in T cell development and subsequent thymic hypoplasia occur (Gosling et al., 2007). When this phenotype was investigated further, it was determined that thymocytes from these mice have a cell cycle block impairing progression into successful anaphase, which itself has an elevated frequency of abnormalities, such as anaphase bridges and lagging chromosomes (Woodward et al., 2016). The abnormal mitoses seen in these *Ncaph2* mutant cells subsequently leads to an accumulation of cells containing 4N DNA content or greater. Interestingly, this phenomenon is also seen in T cells of neonatal mutant mice, indicating these mitotic abnormalities precede malignant transformation.

Although these analyses demonstrate that mutations in condensin II can contribute to disease, the overall causal role of these mutations to diseases is unknown. With this data, it is difficult, if not impossible, to determine if all of these effects are due solely to mitotic abnormalities, or if other contributions of condensin II to processes outside of mitosis also contribute to disease progression. One way in which I have been able to

investigate the role of condensin II further within interphase cells is through the use of the *Rb1^{L/L}* mouse model.

1.16 The *Rb1^{L/L}* mouse model reveals importance of condensin II interactions with the LXCXE binding cleft

As previously described, the LXCXE binding cleft of the pocket proteins is located in a shallow groove in the B subdomain of the small pocket and this is the binding site used by many proteins that alter chromatin structure to bind pRB and act as co-repressors of transcription (Dick, 2007; Lee et al., 1998). Consequently, binding of these proteins to pRB assists in inhibiting cell cycle progression through inhibition of E2F target gene transcription. Viral oncoproteins, on the other hand, also use this LXCXE binding cleft to bind to and inactivate the pocket proteins (Dyson et al., 1992; Ewen et al., 1989; Lee et al., 1998; Munger et al., 1989; Whyte et al., 1989). Despite the fact that the LXCXE binding cleft on the pocket proteins leaves these proteins vulnerable to viral oncoproteins, this region is the most highly conserved among the pocket proteins and across several species of pRB (Lee et al., 1998). Therefore, interactions between cellular proteins and the LXCXE binding cleft of pRB must perform central tasks for this site to have remained so conserved over evolutionary time, and this may not be entirely explained by contributions to cell cycle regulation.

To begin examining what the cellular interactions with the LXCXE binding cleft are important for, point mutants that specifically disrupt this binding surface within pRB were generated. These initial studies demonstrated that when LXCXE interactions are disturbed, transcription from E2F-containing promoters can still be repressed and cell cycle arrest can still be induced but cannot be permanently maintained (Chan et al., 2001; Chen and Wang, 2000; Dahiya et al., 2000; Dick et al., 2000). To investigate the role of interactions with the LXCXE binding cleft of pRB *in vivo*, our lab generated a targeted mouse model where general E2F binding to the large pocket and specific E2F1 binding within the pRB C-terminal domain is not disrupted, but binding to the LXCXE binding cleft is abrogated. The *Rb1^{LXCXE}* allele (*Rb1^L* for short) encodes three alanine substitutions at amino acids within the LXCXE binding cleft (I746A, N750A and M754A), which disrupt protein interactions at this binding interface (Isaac et al., 2006).

The *Rb1^{L/L}* mouse model is viable, fertile, born at nearly the expected Mendelian ratios and does not succumb to spontaneous tumors (Isaac et al., 2006). *Rb1^{-/-}* mice, in contrast, die embryonically due to defects in proliferation and differentiation of placental cells (Wu et al., 2003). This implies LXCXE binding cleft interactions are not necessary for early embryonic development. One overt phenotype that has been seen with these mice, though, is a nursing defect in *Rb1^{L/L}* females which is linked to hyperplasia of the mammary ductal epithelium (Francis et al., 2009).

Studies in mouse embryonic fibroblasts (MEFs) from the *Rb1^{L/L}* genetic background revealed the level of mutant pRB^L protein is equivalent to that in wild type MEFs, and the levels of the other two pocket proteins are also not changed between genotypes implying compensatory effects are unlikely. Importantly, interactions mediated by the LXCXE binding cleft are defective in the pRB^L mutant, while those with E2F transcription factors are not altered. When these *Rb1^{L/L}* MEFs are cultured, similar growth rates to wild type are obtained, and these cells are also able to exit from the cell cycle in response to confluence arrest as well as serum starvation. Transcriptional repression of E2F target genes also remains largely intact in proliferating *Rb1^{L/L}* MEFs (Isaac et al., 2006).

However, there are also some apparent defects in *Rb1^{L/L}* MEFs. Failure to create a repressive chromatin environment at E2F target genes in senescent *Rb1^{L/L}* MEFs leads to defective inhibition of DNA replication so these cells fail to remain permanently arrested in senescence (Isaac et al., 2006; Talluri et al., 2010). In proliferating *Rb1^{L/L}* MEFs, there is an increase in >4N DNA compared to wild type, which is indicative of aneuploidy. This could be due to the fact that *Rb1^{L/L}* cells have abnormal pericentromeric heterochromatin that can lead to fusions of centromeres and ultimately errors in mitosis (Isaac et al., 2006). There is also a general hypocondensation of chromatin within *Rb1^{L/L}* cells; they take longer to progress from chromosome condensation to a fully aligned metaphase plate, which itself is also less compact than in wild type cells (Coschi et al., 2010). When these phenotypes were investigated further, it was discovered that this hypocondensation occurs prior to metaphase and this function of pRB is distinct from its G1-S phase cell cycle regulatory abilities. It was also seen that pRB interacts with

condensin II to maintain appropriate chromosomal structure, which led to the conclusion that many of the mitotic defects in *Rb1^{L/L}* cells could be caused by loss of this interaction (Coschi et al., 2014; Coschi et al., 2010).

1.17 pRB interacts with condensin II to mediate genome stability

Studies performed in *Drosophila* were the first to discover that pRB (named RBF1 in *Drosophila*) is important for chromatin structure, largely due to its interaction with the dCAP-D3 subunit of condensin II (Longworth et al., 2008). This study noted that RBF1 promotes the chromatin association of dCAP-D3, a finding which was further extended to human cells (Longworth et al., 2008). Longworth et al. also established the dependency of pRB on its LXCXE binding cleft for binding to CAP-D3 (Longworth et al., 2008).

Further investigation in *Rb1^{L/L}* MEFs revealed that only condensin II levels are reduced on chromatin, while the overall levels of chromatin bound condensin I and cohesin remain unchanged, indicating that the LXCXE binding cleft of pRB likely only affects the levels of condensin II binding (Coschi et al., 2010). However, in human RPE-1 cells depleted for pRB there is a reduction in both centromeric cohesin and condensin II, leading to an increase in intercentromeric distance and deformed centromere structure (Manning et al., 2010). These defects ultimately lead to merotelic attachments to chromosomes, which is when one or both sister kinetochores are attached to microtubules emanating from both spindle poles, and subsequently an increase in mitotic delay and in lagging chromosomes, similar to what has been seen in cells from the *Rb1^{L/L}* mouse (Coschi et al., 2010; Manning et al., 2010).

Further investigation of pRB and condensin II has revealed that these proteins form a complex along with E2F1 which localizes extensively to major satellite sequences within pericentromeric heterochromatin to facilitate replication (Coschi et al., 2014). It was also found that loss of even one wild type copy of *RB1* reduces recruitment of condensin II to pericentromeres and this also leads to aberrant replication and the ensuing defects in mitotic chromosome segregation (Coschi et al., 2014). Additionally, this study

demonstrated that the specific interaction between pRB and E2F1 is required for CAP-D3 to localization at pericentromeric heterochromatin (Coschi et al., 2014). In line with this, in cells with an *Rb1* mutant allele that disrupts the specific interaction between pRB and E2F1, condensin II localization is decreased at genomic locations where DNA damage accumulates and these cells also have replication abnormalities and aneuploidy (Ishak et al., 2017). However, a thorough investigation of pRB-dependent condensin II binding outside of repetitive regions of the genome has not been performed to investigate other consequences due to loss of this interaction.

1.18 Objectives of the present study

Altogether, there have been some interesting studies associating pRB with genome stability. First of all, there have been the paradoxical observations that the pRB pathway is misregulated in the vast majority of cancers, yet pRB loss specifically has been associated with better treatment outcome (Cecchini et al., 2015; Garsed et al., 2018; Ludovini et al., 2004; Zhao et al., 2012). These findings suggest that pRB is not solely a growth regulator and that loss of *Rb1* may lead to other cancer relevant characteristics that are independent of classical pRB-pathway functions. Secondly, studies where binding at the LXCXE binding cleft of pRB is abrogated has revealed defects in chromosome structure, particularly at the centromeres of mitotic chromosomes (Coschi et al., 2014; Coschi et al., 2010; Manning et al., 2010). Many mitotic defects in *Rb1^{L/L}* cells have been attributed to loss of interaction with condensin II. However, contributions of this pRB-condensin II complex to interphase chromosome dynamics and cellular processes outside of the pericentromere have not been investigated. Therefore, the overall aim of this study is to characterize the different ways in which pRB contributes to the maintenance of genome stability.

In chapter 2, CRISPR/Cas9 was used to create non-functional *Rb1* alleles in cancer cells that already have mutations affecting the pRB-pathway to investigate the specific role of *Rb1* deficiency. I hypothesized that *Rb1* mutations in pRB-pathway disrupted cells would lead to a decrease in genome stability. To test this hypothesis, I examined the levels of DNA damage in cells using γ H2AX foci. To investigate sources of DNA damage, I assessed their susceptibility to a number of chemical agents and explored

these vulnerabilities further. Locations of DNA damage were determined using γ H2AX ChIP-Seq to investigate if damage occurred at specific locations within the genome. Mouse xenograft experiments were also performed to assess if new cancer relevant phenotypes arise in *RBI* mutant cells.

In chapter 3, I investigate pRB-dependent condensin II localization genome-wide. I hypothesized that condensin II localizes extensively throughout the genome in an pRB-dependent manner within interphase nuclei and this has implications on transcriptional regulation and chromatin dynamics. In order to test this hypothesis, condensin II localization was resolved using ChIP-Seq in both wild type and *Rb1^{L/L}* MEFs and data were compared. To determine whether occupancy was functional, I performed RNA-Seq analysis in *Rb1^{+/+}* and *Rb1^{L/L}* MEF pairs. The influence of condensin II on local chromatin contacts was investigated using 3C at locations of interest. Chromatin interactions at these same locations were interrogated further using circularized chromosome conformation capture (4C)-Seq. Finally, we examined whether functional paradigms observed in *Rb1^{L/L}* MEFs could be extended to humans by comparing expression data from *RBI* deleted to *RBI* diploid lung adenocarcinoma samples.

1.19 References

Abe, S., Nagasaka, K., Hirayama, Y., Kozuka-Hata, H., Oyama, M., Aoyagi, Y., Obuse, C., and Hirota, T. (2011). The initial phase of chromosome condensation requires Cdk1-mediated phosphorylation of the CAP-D3 subunit of condensin II. *Genes Dev* 25, 863-874.

Adams, P.D., Li, X., Sellers, W.R., Baker, K.B., Leng, X., Harper, J.W., Taya, Y., and Kaelin, W.G., Jr. (1999). Retinoblastoma protein contains a C-terminal motif that targets it for phosphorylation by cyclin-cdk complexes. *Mol Cell Biol* 19, 1068-1080.

Aguilera, A., and Gomez-Gonzalez, B. (2008). Genome instability: a mechanistic view of its causes and consequences. *Nat Rev Genet* 9, 204-217.

Akai, Y., Kurokawa, Y., Nakazawa, N., Tonami-Murakami, Y., Suzuki, Y., Yoshimura, S.H., Iwasaki, H., Shiroiwa, Y., Nakamura, T., Shibata, E., *et al.* (2011). Opposing role of condensin hinge against replication protein A in mitosis and interphase through promoting DNA annealing. *Open biology* 1, 110023.

Anderson, D.E., Losada, A., Erickson, H.P., and Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. *The Journal of cell biology* *156*, 419-424.

Banerji, R., Skibbens, R.V., and Iovine, M.K. (2017a). Cohesin mediates Esco2-dependent transcriptional regulation in a zebrafish regenerating fin model of Roberts Syndrome. *Biol Open* *6*, 1802-1813.

Banerji, R., Skibbens, R.V., and Iovine, M.K. (2017b). How many roads lead to cohesinopathies? *Dev Dyn* *246*, 881-888.

Barysz, H., Kim, J.H., Chen, Z.A., Hudson, D.F., Rappsilber, J., Gerloff, D.L., and Earnshaw, W.C. (2015). Three-dimensional topology of the SMC2/SMC4 subcomplex from chicken condensin I revealed by cross-linking and molecular modelling. *Open biology* *5*, 150005.

Beliveau, B.J., Joyce, E.F., Apostolopoulos, N., Yilmaz, F., Fonseka, C.Y., McCole, R.B., Chang, Y., Li, J.B., Senaratne, T.N., Williams, B.R., *et al.* (2012). Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc Natl Acad Sci U S A* *109*, 21301-21306.

Beltran, H., Prandi, D., Mosquera, J.M., Benelli, M., Puca, L., Cyrta, J., Marotz, C., Giannopoulou, E., Chakravarthi, B.V., Varambally, S., *et al.* (2016). Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat Med* *22*, 298-305.

Blake, M.C., and Azizkhan, J.C. (1989). Transcription factor E2F is required for efficient expression of the hamster dihydrofolate reductase gene in vitro and in vivo. *Mol Cell Biol* *9*, 4994-5002.

Blat, Y., and Kleckner, N. (1999). Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. *Cell* *98*, 249-259.

Bookstein, R., Rio, P., Madreperla, S.A., Hong, F., Allred, C., Grizzle, W.E., and Lee, W.H. (1990a). Promoter deletion and loss of retinoblastoma gene expression in human prostate carcinoma. *Proc Natl Acad Sci U S A* *87*, 7762-7766.

Bookstein, R., Shew, J.Y., Chen, P.L., Scully, P., and Lee, W.H. (1990b). Suppression of tumorigenicity of human prostate carcinoma cells by replacing a mutated RB gene. *Science* *247*, 712-715.

Brehm, A., Miska, E.A., McCance, D.J., Reid, J.L., Bannister, A.J., and Kouzarides, T. (1998). Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* *391*, 597-601.

Breier, A.M., and Grossman, A.D. (2007). Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol Microbiol* *64*, 703-718.

Bremner, R., Cohen, B.L., Sopta, M., Hamel, P.A., Ingles, C.J., Gallie, B.L., and Phillips, R.A. (1995). Direct transcriptional repression by pRB and its reversal by specific cyclins. *Mol Cell Biol* 15, 3256-3265.

Buchkovich, K., Duffy, L.A., and Harlow, E. (1989). The retinoblastoma protein is phosphorylated during specific phases of the cell cycle. *Cell* 58, 1097-1105.

Burke, J.R., Hura, G.L., and Rubin, S.M. (2012). Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control. *Genes Dev* 26, 1156-1166.

Burkhardt, D.L., and Sage, J. (2008). Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat Rev Cancer* 8, 671-682.

Calbó, J., Parreño, M., Sotillo, E., Yong, T., Mazo, A., Garriga, J., and Graña, X. (2002). G1 Cyclin/Cyclin-dependent Kinase-coordinated Phosphorylation of Endogenous Pocket Proteins Differentially Regulates Their Interactions with E2F4 and E2F1 and Gene Expression. *Journal of Biological Chemistry* 277, 50263-50274.

Carnevale, J., Palander, O., Seifried, L.A., and Dick, F.A. (2012). DNA damage signals through differentially modified E2F1 molecules to induce apoptosis. *Mol Cell Biol* 32, 900-912.

Carriere, L., Graziani, S., Alibert, O., Ghavi-Helm, Y., Boussouar, F., Humbertclaude, H., Jounier, S., Aude, J.C., Keime, C., Murvai, J., *et al.* (2012). Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells. *Nucleic acids research* 40, 270-283.

Cecchini, M.J., and Dick, F.A. (2011). The biochemical basis of CDK phosphorylation-independent regulation of E2F1 by the retinoblastoma protein. *Biochem J*.

Cecchini, M.J., Ishak, C.A., Passos, D.T., Warner, A., Palma, D.A., Howlett, C.J., Driman, D.K., and Dick, F.A. (2015). Loss of the retinoblastoma tumor suppressor correlates with improved outcome in patients with lung adenocarcinoma treated with surgery and chemotherapy. *Human pathology* 46, 1922-1934.

Chan, H.M., Smith, L., and La Thangue, N.B. (2001). Role of LXCXE motif-dependent interactions in the activity of the retinoblastoma protein. *Oncogene* 20, 6152-6163.

Chellappan, S.P., Hiebert, S., Mudryj, M., Horowitz, J.M., and Nevins, J.R. (1991). The E2F transcription factor is a cellular target for the RB protein. *Cell* 65, 1053-1061.

Chen, H.Z., Tsai, S.Y., and Leone, G. (2009). Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer* 9, 785-797.

Chen, T.T., and Wang, J.Y. (2000). Establishment of irreversible growth arrest in myogenic differentiation requires the RB LXCXE-binding function. *Mol Cell Biol* 20, 5571-5580.

Chow, K.N., and Dean, D.C. (1996). Domains A and B in the Rb pocket interact to form a transcriptional repressor motif. *Mol Cell Biol* *16*, 4862-4868.

Chow, K.N., Starostik, P., and Dean, D.C. (1996). The Rb family contains a conserved cyclin-dependent-kinase-regulated transcriptional repressor motif. *Mol Cell Biol* *16*, 7173-7181.

Ciccia, A., and Elledge, S.J. (2010). The DNA damage response: making it safe to play with knives. *Mol Cell* *40*, 179-204.

Classon, M., and Dyson, N. (2001). p107 and p130: versatile proteins with interesting pockets. *Experimental cell research* *264*, 135-147.

Cobbe, N., and Heck, M.M. (2004). The evolution of SMC proteins: phylogenetic analysis and structural implications. *Molecular biology and evolution* *21*, 332-347.

Cobrinik, D., Whyte, P., Peeper, D.S., Jacks, T., and Weinberg, R.A. (1993). Cell cycle-specific association of E2F with the p130 E1A-binding protein. *Genes Dev* *7*, 2392-2404.

Coschi, C.H., Ishak, C.A., Gallo, D., Marshall, A., Talluri, S., Wang, J., Cecchini, M.J., Martens, A.L., Percy, V., Welch, I., *et al.* (2014). Haploinsufficiency of an RB-E2F1-Condensin II Complex Leads to Aberrant Replication and Aneuploidy. *Cancer discovery*.

Coschi, C.H., Martens, A.L., Ritchie, K., Francis, S.M., Chakrabarti, S., Berube, N.G., and Dick, F.A. (2010). Mitotic chromosome condensation mediated by the retinoblastoma protein is tumor-suppressive. *Genes & Development* *24*, 1351-1363.

Cuylen, S., and Haering, C.H. (2011). Deciphering condensin action during chromosome segregation. *Trends Cell Biol* *21*, 552-559.

D'Ambrosio, C., Schmidt, C.K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., and Uhlmann, F. (2008). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes Dev* *22*, 2215-2227.

Dahiya, A., Gavin, M.R., Luo, R.X., and Dean, D.C. (2000). Role of the LXCXE binding site in Rb function. *Mol Cell Biol* *20*, 6799-6805.

Davidson, I.F., Goetz, D., Zaczek, M.P., Molodtsov, M.I., Huis In 't Veld, P.J., Weissmann, F., Litos, G., Cisneros, D.A., Ocampo-Hafalla, M., Ladurner, R., *et al.* (2016). Rapid movement and transcriptional re-localization of human cohesin on DNA. *EMBO J* *35*, 2671-2685.

De Piccoli, G., Torres-Rosell, J., and Aragon, L. (2009). The unnamed complex: what do we know about Smc5-Smc6? *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* *17*, 251-263.

- DeCaprio, J.A., Ludlow, J.W., Figge, J., Shew, J.Y., Huang, C.M., Lee, W.H., Marsilio, E., Paucha, E., and Livingston, D.M. (1988). SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. *Cell* 54, 275-283.
- DeCaprio, J.A., Ludlow, J.W., Lynch, D., Furukawa, Y., Griffin, J., Piwnica-Worms, H., Huang, C.M., and Livingston, D.M. (1989). The product of the retinoblastoma susceptibility gene has properties of a cell cycle regulatory element. *Cell* 58, 1085-1095.
- DeGregori, J., and Johnson, D.G. (2006). Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. *Curr Mol Med* 6, 739-748.
- Dej, K.J., and Spradling, A.C. (1999). The endocycle controls nurse cell polytene chromosome structure during *Drosophila* oogenesis. *Development* 126, 293-303.
- Dekker, J., and Mirny, L. (2013). Biological techniques: Chromosomes captured one by one. *Nature* 502, 45-46.
- Dick, F.A. (2007). Structure-function analysis of the retinoblastoma tumor suppressor protein - is the whole a sum of its parts? *Cell Div* 2, 26.
- Dick, F.A., and Dyson, N. (2003). pRB Contains an E2F1 Specific Binding Domain that Allows E2F1 Induced Apoptosis to be Regulated Separately from other E2F Activities. *Mol Cell* 12, 639-649.
- Dick, F.A., Goodrich, D.W., Sage, J., and Dyson, N.J. (2018). Non-canonical functions of the RB protein in cancer. *Nat Rev Cancer* 18, 442-451.
- Dick, F.A., and Rubin, S.M. (2013). Molecular mechanisms underlying RB protein function. *Nat Rev Mol Cell Biol* 14, 297-306.
- Dick, F.A., Sailhamer, E., and Dyson, N.J. (2000). Mutagenesis of the pRB pocket reveals that cell cycle arrest functions are separable from binding to viral oncoproteins. *Mol Cell Biol* 20, 3715-3727.
- Diebold-Durand, M.L., Lee, H., Ruiz Avila, L.B., Noh, H., Shin, H.C., Im, H., Bock, F.P., Burmann, F., Durand, A., Basfeld, A., *et al.* (2017). Structure of Full-Length SMC and Rearrangements Required for Chromosome Organization. *Mol Cell* 67, 334-347 e335.
- Dimova, D.K., and Dyson, N.J. (2005). The E2F transcriptional network: old acquaintances with new faces. *Oncogene* 24, 2810-2826.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376-380.
- Draviam, V.M., Xie, S., and Sorger, P.K. (2004). Chromosome segregation and genomic stability. *Curr Opin Genet Dev* 14, 120-125.

- Durfee, T., Becherer, K., Chen, P.L., Yeh, S.H., Yang, Y., Kilburn, A.E., Lee, W.H., and Elledge, S.J. (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev* 7, 555-569.
- Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes Dev* 12, 2245-2262.
- Dyson, N., Guida, P., McCall, C., and Harlow, E. (1992). Adenovirus E1A makes two distinct contacts with the retinoblastoma protein. *J Virol* 66, 4606-4611.
- Dyson, N., Howley, P.M., Munger, K., and Harlow, E. (1989). The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 243, 934-937.
- Eeftens, J.M., Bisht, S., Kerssemakers, J., Kschonsak, M., Haering, C.H., and Dekker, C. (2017). Real-time detection of condensin-driven DNA compaction reveals a multistep binding mechanism. *EMBO J* 36, 3448-3457.
- Eeftens, J.M., Katan, A.J., Kschonsak, M., Hassler, M., de Wilde, L., Dief, E.M., Haering, C.H., and Dekker, C. (2016). Condensin Smc2-Smc4 Dimers Are Flexible and Dynamic. *Cell Rep* 14, 1813-1818.
- Ewen, M.E., Ludlow, J.W., Marsilio, E., DeCaprio, J.A., Millikan, R.C., Cheng, S.H., Paucha, E., and Livingston, D.M. (1989). An N-terminal transformation-governing sequence of SV40 large T antigen contributes to the binding of both p110Rb and a second cellular protein, p120. *Cell* 58, 257-267.
- Ewen, M.E., Xing, Y.G., Lawrence, J.B., and Livingston, D.M. (1991). Molecular cloning, chromosomal mapping, and expression of the cDNA for p107, a retinoblastoma gene product-related protein. *Cell* 66, 1155-1164.
- Falls, H.F., and Neel, J.V. (1951). Genetics of retinoblastoma. *AMA Arch Ophthalmol* 46, 367-389.
- Francis, S.M., Bergsied, J., Isaac, C.E., Coschi, C.H., Martens, A.L., Hojilla, C.V., Chakrabarti, S., Dimattia, G.E., Khoka, R., Wang, J.Y., *et al.* (2009). A functional connection between pRB and transforming growth factor beta in growth inhibition and mammary gland development. *Mol Cell Biol* 29, 4455-4466.
- Francke, U., and Kung, F. (1976). Sporadic bilateral retinoblastoma and 13q-chromosomal deletion. *Med Pediatr Oncol* 2, 379-385.
- Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 323, 643-646.
- Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., and Dekker, C. (2018). Real-time imaging of DNA loop extrusion by condensin. *Science* 360, 102-105.

Garsed, D.W., Alsop, K., Fereday, S., Emmanuel, C., Kennedy, C.J., Etemadmoghadam, D., Gao, B., GebSKI, V., Gares, V., Christie, E.L., *et al.* (2018). Homologous Recombination DNA Repair Pathway Disruption and Retinoblastoma Protein Loss Are Associated with Exceptional Survival in High-Grade Serous Ovarian Cancer. *Clin Cancer Res* 24, 569-580.

Gassler, J., Brandao, H.B., Imakaev, M., Flyamer, I.M., Ladstatter, S., Bickmore, W.A., Peters, J.M., Mirny, L.A., and Tachibana, K. (2017). A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J* 36, 3600-3618.

Geigl, J.B., Obenauf, A.C., Schwarzbraun, T., and Speicher, M.R. (2008). Defining 'chromosomal instability'. *Trends in genetics : TIG* 24, 64-69.

Gelot, C., Magdalou, I., and Lopez, B.S. (2015). Replication stress in Mammalian cells and its consequences for mitosis. *Genes (Basel)* 6, 267-298.

Gerlich, D., Hirota, T., Koch, B., Peters, J.M., and Ellenberg, J. (2006a). Condensin I stabilizes chromosomes mechanically through a dynamic interaction in live cells. *Current biology : CB* 16, 333-344.

Gerlich, D., Koch, B., Dupeux, F., Peters, J.M., and Ellenberg, J. (2006b). Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. *Current biology : CB* 16, 1571-1578.

Giacinti, C., and Giordano, A. (2006). RB and cell cycle progression. *Oncogene* 25, 5220-5227.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., *et al.* (2018). A pathway for mitotic chromosome formation. *Science* 359.

Gibson, T.J., Thompson, J.D., Blocker, A., and Kouzarides, T. (1994). Evidence for a protein domain superfamily shared by the cyclins, TFIIB and RB/p107. *Nucleic acids research* 22, 946-952.

Girling, R., Partridge, J.F., Bandara, L.R., Burden, N., Totty, N.F., Hsuan, J.J., and La Thangue, N.B. (1993). A new component of the transcription factor DRTF1/E2F. *Nature* 362, 83-87.

Glynn, E.F., Megee, P.C., Yu, H.G., Mistrot, C., Unal, E., Koshland, D.E., DeRisi, J.L., and Gerton, J.L. (2004). Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*. *PLoS Biol* 2, E259.

Gosling, K.M., Makaroff, L.E., Theodoratos, A., Kim, Y.H., Whittle, B., Rui, L., Wu, H., Hong, N.A., Kennedy, G.C., Fritz, J.A., *et al.* (2007). A mutation in a chromosome condensin II subunit, kleisin beta, specifically disrupts T cell development. *Proc Natl Acad Sci U S A* 104, 12445-12450.

Green, L.C., Kalitsis, P., Chang, T.M., Cipetic, M., Kim, J.H., Marshall, O., Turnbull, L., Whitchurch, C.B., Vagnarelli, P., Samejima, K., *et al.* (2012). Contrasting roles of condensin I and condensin II in mitotic chromosome formation. *J Cell Sci* *125*, 1591-1604.

Gruber, S., and Errington, J. (2009). Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in *B. subtilis*. *Cell* *137*, 685-696.

Gruber, S., Haering, C.H., and Nasmyth, K. (2003). Chromosomal cohesin forms a ring. *Cell* *112*, 765-777.

Gruber, S., Veening, J.W., Bach, J., Blettinger, M., Bramkamp, M., and Errington, J. (2014). Interlinked sister chromosomes arise in the absence of condensin during fast replication in *B. subtilis*. *Current biology : CB* *24*, 293-298.

Guacci, V., Koshland, D., and Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in *S. cerevisiae*. *Cell* *91*, 47-57.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B., *et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* *169*, 693-707 e614.

Hagstrom, K.A., Holmes, V.F., Cozzarelli, N.R., and Meyer, B.J. (2002). *C. elegans* condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid segregation during mitosis and meiosis. *Genes Dev* *16*, 729-742.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. *Cell* *100*, 57-70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* *144*, 646-674.

Hannon, G.J., Demetrick, D., and Beach, D. (1993). Isolation of the Rb-related p130 through its interaction with CDK2 and cyclins. *Genes Dev* *7*, 2378-2391.

Hansen, A.S., Cattoglio, C., Darzacq, X., and Tjian, R. (2018). Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* *9*, 20-32.

Harper, J.W., and Elledge, S.J. (2007). The DNA damage response: ten years after. *Mol Cell* *28*, 739-745.

Hartl, T.A., Smith, H.F., and Bosco, G. (2008). Chromosome alignment and transvection are antagonized by condensin II. *Science* *322*, 1384-1387.

- Helin, K., Lees, J.A., Vidal, M., Dyson, N., Harlow, E., and Fattaey, A. (1992). A cDNA encoding a pRB-binding protein with properties of the transcription factor E2F. *Cell* *70*, 337-350.
- Hiebert, S.W., Chellappan, S.P., Horowitz, J.M., and Nevins, J.R. (1992). The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. *Genes Dev* *6*, 177-185.
- Hiebert, S.W., Lipp, M., and Nevins, J.R. (1989). E1A-dependent trans-activation of the human MYC promoter is mediated by the E2F factor. *Proc Natl Acad Sci U S A* *86*, 3594-3598.
- Hirota, T., Gerlich, D., Koch, B., Ellenberg, J., and Peters, J.M. (2004). Distinct functions of condensin I and II in mitotic chromosome assembly. *J Cell Sci* *117*, 6435-6445.
- Hu, Q.J., Dyson, N., and Harlow, E. (1990). The regions of the retinoblastoma protein needed for binding to adenovirus E1A or SV40 large T antigen are common sites for mutations. *EMBO J* *9*, 1147-1155.
- Huang, H.J., Yee, J.K., Shew, J.Y., Chen, P.L., Bookstein, R., Friedmann, T., Lee, E.Y., and Lee, W.H. (1988). Suppression of the neoplastic phenotype by replacement of the RB gene in human cancer cells. *Science* *242*, 1563-1566.
- Hudson, D.F., Vagnarelli, P., Gassmann, R., and Earnshaw, W.C. (2003). Condensin is required for nonhistone protein assembly and structural integrity of vertebrate mitotic chromosomes. *Dev Cell* *5*, 323-336.
- Isaac, C.E., Francis, S.M., Martens, A.L., Julian, L.M., Seifried, L.A., Erdmann, N., Binne, U.K., Harrington, L., Sicinski, P., Berube, N.G., *et al.* (2006). The retinoblastoma protein regulates pericentric heterochromatin. *Mol Cell Biol* *26*, 3659-3671.
- Ishak, C.A., Coschi, C.H., Roes, M.V., and Dick, F.A. (2017). Disruption of CDK-resistant chromatin association by pRB causes DNA damage, mitotic errors, and reduces Condensin II recruitment. *Cell Cycle* *16*, 1430-1439.
- Ishak, C.A., Marshall, A.E., Passos, D.T., White, C.R., Kim, S.J., Cecchini, M.J., Ferwati, S., MacDonald, W.A., Howlett, C.J., Welch, I.D., *et al.* (2016). An RB-EZH2 Complex Mediates Silencing of Repetitive DNA Sequences. *Mol Cell*.
- Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature* *461*, 1071-1078.
- Jeppsson, K., Kanno, T., Shirahige, K., and Sjogren, C. (2014). The maintenance of chromosome structure: positioning and functioning of SMC complexes. *Nat Rev Mol Cell Biol* *15*, 601-614.
- Johnson, D.G., and Degregori, J. (2006). Putting the Oncogenic and Tumor Suppressive Activities of E2F into Context. *Curr Mol Med* *6*, 731-738.

- Julian, L.M., Palander, O., Seifried, L.A., Foster, J.E., and Dick, F.A. (2008). Characterization of an E2F1-specific binding domain in pRB and its implications for apoptotic regulation. *Oncogene* 27, 1572-1579.
- Kaelin, W.G., Jr., Ewen, M.E., and Livingston, D.M. (1990). Definition of the minimal simian virus 40 large T antigen- and adenovirus E1A-binding domain in the retinoblastoma gene product. *Mol Cell Biol* 10, 3761-3769.
- Kagami, Y., and Yoshida, K. (2016). The functional role for condensin in the regulation of chromosomal organization during the cell cycle. *Cellular and molecular life sciences : CMLS* 73, 4591-4598.
- Kakui, Y., Rabinowitz, A., Barry, D.J., and Uhlmann, F. (2017). Condensin-mediated remodeling of the mitotic chromatin landscape in fission yeast. *Nat Genet* 49, 1553-1557.
- Kawauchi, S., Calof, A.L., Santos, R., Lopez-Burks, M.E., Young, C.M., Hoang, M.P., Chua, A., Lao, T., Lechner, M.S., Daniel, J.A., *et al.* (2009). Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome. *PLoS genetics* 5, e1000650.
- Keenholtz, R.A., Dhanaraman, T., Palou, R., Yu, J., D'Amours, D., and Marko, J.F. (2017). Oligomerization and ATP stimulate condensin-mediated DNA compaction. *Sci Rep* 7, 14279.
- Kimura, K., Hirano, M., Kobayashi, R., and Hirano, T. (1998). Phosphorylation and activation of 13S condensin by Cdc2 in vitro. *Science* 282, 487-490.
- Kimura, K., and Hirano, T. (2000). Dual roles of the 11S regulatory subcomplex in condensin functions. *Proc Natl Acad Sci U S A* 97, 11972-11977.
- Kinzler, K.W., and Vogelstein, B. (1997). Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* 386, 761, 763.
- Knudsen, E.S., and Knudsen, K.E. (2008). Tailoring to RB: tumour suppressor status and therapeutic response. *Nat Rev Cancer* 8, 714-724.
- Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 68, 820-823.
- Kovesdi, I., Reichel, R., and Nevins, J.R. (1987). Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc Natl Acad Sci U S A* 84, 2180-2184.
- Ku, S.Y., Rosario, S., Wang, Y., Mu, P., Seshadri, M., Goodrich, Z.W., Goodrich, M.M., Labbe, D.P., Gomez, E.C., Wang, J., *et al.* (2017). Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science* 355, 78-83.

- Kueng, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K., and Peters, J.M. (2006). Wapl controls the dynamic association of cohesin with chromatin. *Cell* 127, 955-967.
- Lapointe, J., Malhotra, S., Higgins, J.P., Bair, E., Thompson, M., Salari, K., Giacomini, C.P., Ferrari, M., Montgomery, K., Tibshirani, R., *et al.* (2008). hCAP-D3 expression marks a prostate cancer subtype with favorable clinical behavior and androgen signaling signature. *Am J Surg Pathol* 32, 205-209.
- Lee, J.O., Russo, A.A., and Pavletich, N.P. (1998). Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. *Nature* 391, 859-865.
- Lee, W.H., Bookstein, R., Hong, F., Young, L.J., Shew, J.Y., and Lee, E.Y. (1987). Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science* 235, 1394-1399.
- Leiserson, M.D., Vandin, F., Wu, H.T., Dobson, J.R., Eldridge, J.V., Thomas, J.L., Papoutsaki, A., Kim, Y., Niu, B., McLellan, M., *et al.* (2015). Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet* 47, 106-114.
- Lele, K.P., Penrose, L.S., and Stallard, H.B. (1963). Chromosome Deletion in a Case of Retinoblastoma. *Ann Hum Genet* 27, 171-174.
- Li, W., Hu, Y., Oh, S., Ma, Q., Merkurjev, D., Song, X., Zhou, X., Liu, Z., Tanasa, B., He, X., *et al.* (2015). Condensin I and II Complexes License Full Estrogen Receptor alpha-Dependent Enhancer Activation. *Mol Cell* 59, 188-202.
- Li, Y., Graham, C., Lacy, S., Duncan, A.M., and Whyte, P. (1993). The adenovirus E1A-associated 130-kD protein is encoded by a member of the retinoblastoma gene family and physically interacts with cyclins A and E. *Genes Dev* 7, 2366-2377.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-293.
- Liu, W., Tanasa, B., Tyurina, O.V., Zhou, T.Y., Gassmann, R., Liu, W.T., Ohgi, K.A., Benner, C., Garcia-Bassets, I., Aggarwal, A.K., *et al.* (2010). PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. *Nature* 466, 508-512.
- Longworth, M.S., Herr, A., Ji, J.Y., and Dyson, N.J. (2008). RBF1 promotes chromatin condensation through a conserved interaction with the Condensin II protein dCAP-D3. *Genes Dev* 22, 1011-1024.
- Losada, A., and Hirano, T. (2005). Dynamic molecular linkers of the genome: the first decade of SMC proteins. *Genes Dev* 19, 1269-1287.

Ludovini, V., Gregorc, V., Pistola, L., Mihaylova, Z., Floriani, I., Darwish, S., Stracci, F., Tofanetti, F.R., Ferraldeschi, M., Di Carlo, L., *et al.* (2004). Vascular endothelial growth factor, p53, Rb, Bcl-2 expression and response to chemotherapy in advanced non-small cell lung cancer. *Lung Cancer* 46, 77-85.

Magnaghi-Jaulin, L., Groisman, R., Naguibneva, I., Robin, P., Lorain, S., Le Villain, J.P., Troalen, F., Trouche, D., and Harel-Bellan, A. (1998). Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* 391, 601-605.

Manning, A.L., Longworth, M.S., and Dyson, N.J. (2010). Loss of pRB causes centromere dysfunction and chromosomal instability. *Genes Dev* 24, 1364-1376.

Marston, A.L. (2014). Chromosome segregation in budding yeast: sister chromatid cohesion and related mechanisms. *Genetics* 196, 31-63.

Martin, C.A., Murray, J.E., Carroll, P., Leitch, A., Mackenzie, K.J., Halachev, M., Fetit, A.E., Keith, C., Bicknell, L.S., Fluteau, A., *et al.* (2016). Mutations in genes encoding condensin complex proteins cause microcephaly through decatenation failure at mitosis. *Genes Dev* 30, 2158-2172.

Mayol, X., Grana, X., Baldi, A., Sang, N., Hu, Q., and Giordano, A. (1993). Cloning of a new member of the retinoblastoma gene family (pRb2) which binds to the E1A transforming domain. *Oncogene* 8, 2561-2566.

McNair, C., Xu, K., Mandigo, A.C., Benelli, M., Leiby, B., Rodrigues, D., Lindberg, J., Gronberg, H., Crespo, M., De Laere, B., *et al.* (2018). Differential impact of RB status on E2F1 reprogramming in human cancer. *J Clin Invest* 128, 341-358.

Meier, M., Grant, J., Dowdle, A., Thomas, A., Gerton, J., Collas, P., O'Sullivan, J.M., and Horsfield, J.A. (2018). Cohesin facilitates zygotic genome activation in zebrafish. *Development* 145.

Melby, T.E., Ciampaglio, C.N., Briscoe, G., and Erickson, H.P. (1998). The symmetrical structure of structural maintenance of chromosomes (SMC) and MukB proteins: long, antiparallel coiled coils, folded at a flexible hinge. *The Journal of cell biology* 142, 1595-1604.

Morales, C., and Losada, A. (2018). Establishing and dissolving cohesion during the vertebrate cell cycle. *Curr Opin Cell Biol* 52, 51-57.

Morgunova, E., Yin, Y., Jolma, A., Dave, K., Schmierer, B., Popov, A., Eremina, N., Nilsson, L., and Taipale, J. (2015). Structural insights into the DNA-binding specificity of E2F family transcription factors. *Nature communications* 6, 10050.

Mu, P., Zhang, Z., Benelli, M., Karthaus, W.R., Hoover, E., Chen, C.C., Wongvipat, J., Ku, S.Y., Gao, D., Cao, Z., *et al.* (2017). SOX2 promotes lineage plasticity and antiandrogen resistance in TP53- and RB1-deficient prostate cancer. *Science* 355, 84-88.

- Munger, K., Werness, B.A., Dyson, N., Phelps, W.C., Harlow, E., and Howley, P.M. (1989). Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *EMBO J* 8, 4099-4105.
- Murphree, A.L., and Benedict, W.F. (1984). Retinoblastoma: clues to human oncogenesis. *Science* 223, 1028-1033.
- Muto, A., and Schilling, T.F. (2017). Zebrafish as a Model to Study Cohesin and Cohesinopathies. *Methods Mol Biol* 1515, 177-196.
- Nasmyth, K. (2011). Cohesin: a catenase with separate entry and exit gates? *Nat Cell Biol* 13, 1170-1177.
- Neel, J.V., and Falls, H.F. (1951). The rate of mutation of the gene responsible for retinoblastoma in man. *Science* 114, 419-422.
- Negrini, S., Gorgoulis, V.G., and Halazonetis, T.D. (2010). Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* 11, 220-228.
- Neuwald, A.F., and Hirano, T. (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. *Genome research* 10, 1445-1452.
- Nevins, J.R. (1992). E2F: a link between the Rb tumor suppressor protein and viral oncoproteins. *Science* 258, 424-429.
- Niederst, M.J., Sequist, L.V., Poirier, J.T., Mermel, C.H., Lockerman, E.L., Garcia, A.R., Katayama, R., Costa, C., Ross, K.N., Moran, T., *et al.* (2015). RB loss in resistant EGFR mutant lung adenocarcinomas that transform to small-cell lung cancer. *Nature communications* 6, 6377.
- Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E., *et al.* (2001). Rb targets histone H3 methylation and HP1 to promoters. *Nature* 412, 561-565.
- Oliveira, R.A., Coelho, P.A., and Sunkel, C.E. (2005). The condensin I subunit Barren/CAP-H is essential for the structural integrity of centromeric heterochromatin during mitosis. *Mol Cell Biol* 25, 8971-8984.
- Ono, T., Fang, Y., Spector, D.L., and Hirano, T. (2004). Spatial and temporal regulation of Condensins I and II in mitotic chromosome assembly in human cells. *Mol Biol Cell* 15, 3296-3308.
- Ono, T., Losada, A., Hirano, M., Myers, M.P., Neuwald, A.F., and Hirano, T. (2003). Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell* 115, 109-121.

Ono, T., Yamashita, D., and Hirano, T. (2013). Condensin II initiates sister chromatid resolution during S phase. *The Journal of cell biology* *200*, 429-441.

Piazza, I., Haering, C.H., and Rutkowska, A. (2013). Condensin: crafting the chromosome landscape. *Chromosoma* *122*, 175-190.

Piazza, I., Rutkowska, A., Ori, A., Walczak, M., Metz, J., Pelechano, V., Beck, M., and Haering, C.H. (2014). Association of condensin with chromosomes depends on DNA binding by its HEAT-repeat subunits. *Nature structural & molecular biology* *21*, 560-568.

Plys, A.J., and Kingston, R.E. (2018). Dynamic condensates activate transcription. *Science* *361*, 329-330.

Qin, X.Q., Chittenden, T., Livingston, D.M., and Kaelin, W.G., Jr. (1992). Identification of a growth suppression domain within the retinoblastoma gene product. *Genes Dev* *6*, 953-964.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., *et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* *159*, 1665-1680.

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., *et al.* (2017). Cohesin Loss Eliminates All Loop Domains. *Cell* *171*, 305-320 e324.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat Genet* *25*, 338-342.

Robinson, D.R., Wu, Y.M., Lonigro, R.J., Vats, P., Cobain, E., Everett, J., Cao, X., Rabban, E., Kumar-Sinha, C., Raymond, V., *et al.* (2017). Integrative clinical genomics of metastatic cancer. *Nature* *548*, 297-303.

Rollins, R.A., Morcillo, P., and Dorsett, D. (1999). Nipped-B, a *Drosophila* homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics* *152*, 577-593.

Rosin, L.F., Nguyen, S.C., and Joyce, E.F. (2018). Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in *Drosophila* nuclei. *PLoS genetics* *14*, e1007393.

Rowley, M.J., Lyu, X., Rana, V., Ando-Kuri, M., Karns, R., Bosco, G., and Corces, V.G. (2019). Condensin II Counteracts Cohesin and RNA Polymerase II in the Establishment of 3D Chromatin Organization. *Cell Rep* *26*, 2890-2903 e2893.

Rudra, S., and Skibbens, R.V. (2013). Cohesin codes - interpreting chromatin architecture and the many facets of cohesin function. *J Cell Sci* *126*, 31-41.

- Sakai, A., Hizume, K., Sutani, T., Takeyasu, K., and Yanagida, M. (2003). Condensin but not cohesin SMC heterodimer induces DNA reannealing through protein-protein assembly. *EMBO J* 22, 2764-2775.
- Sakamoto, T., Inui, Y.T., Uruguchi, S., Yoshizumi, T., Matsunaga, S., Mastui, M., Umeda, M., Fukui, K., and Fujiwara, T. (2011). Condensin II alleviates DNA damage and is essential for tolerance of boron overload stress in Arabidopsis. *The Plant cell* 23, 3533-3546.
- Salk, J.J., Fox, E.J., and Loeb, L.A. (2010). Mutational heterogeneity in human cancers: origin and consequences. *Annu Rev Pathol* 5, 51-75.
- Schalbetter, S.A., Goloborodko, A., Fudenberg, G., Belton, J.M., Miles, C., Yu, M., Dekker, J., Mirny, L., and Baxter, J. (2017). SMC complexes differentially compact mitotic chromosomes according to genomic context. *Nat Cell Biol* 19, 1071-1080.
- Schappert-Kimmijser, J., Hemmes, G.D., and Nijland, R. (1966). The heredity of retinoblastoma. *Ophthalmologica* 151, 197-213.
- Schleiffer, A., Kaitna, S., Maurer-Stroh, S., Glotzer, M., Nasmyth, K., and Eisenhaber, F. (2003). Kleisins: a superfamily of bacterial and eukaryotic SMC protein partners. *Mol Cell* 11, 571-575.
- Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., C, H.H., Mirny, L., *et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 551, 51-56.
- Seifried, L.A., Talluri, S., Cecchini, M., Julian, L.M., Mymryk, J.S., and Dick, F.A. (2008). pRB-E2F1 complexes are resistant to adenovirus E1A-mediated disruption. *J Virol* 82, 4511-4520.
- Sellers, W.R., Rodgers, J.W., and Kaelin, W.G., Jr. (1995). A potent transrepression domain in the retinoblastoma protein induces a cell cycle arrest when bound to E2F sites. *Proc Natl Acad Sci U S A* 92, 11544-11548.
- Sherr, C.J. (1996). Cancer cell cycles. *Science* 274, 1672-1677.
- Shintomi, K., and Hirano, T. (2011). The relative ratio of condensin I to II determines chromosome shapes. *Genes Dev* 25, 1464-1469.
- Siddiqui, N.U., Rusyniak, S., Hasenkampf, C.A., and Riggs, C.D. (2006). Disruption of the Arabidopsis SMC4 gene, AtCAP-C, compromises gametogenesis and embryogenesis. *Planta* 223, 990-997.
- Skibbens, R.V. (2019). Condensins and cohesins - one of these things is not like the other! *J Cell Sci* 132.

- Smith, S.M., and Sorsby, A. (1958). Retinoblastoma: some genetic aspects. *Ann Hum Genet* 23, 50-58.
- Soh, Y.M., Burmann, F., Shin, H.C., Oda, T., Jin, K.S., Toseland, C.P., Kim, C., Lee, H., Kim, S.J., Kong, M.S., *et al.* (2015). Molecular basis for SMC rod formation and its dissolution upon DNA binding. *Mol Cell* 57, 290-303.
- Sparkes, R.S., Sparkes, M.C., Wilson, M.G., Towner, J.W., Benedict, W., Murphree, A.L., and Yunis, J.J. (1980). Regional assignment of genes for human esterase D and retinoblastoma to chromosome band 13q14. *Science* 208, 1042-1044.
- St-Pierre, J., Douziech, M., Bazile, F., Pascariu, M., Bonneil, E., Sauve, V., Ratsima, H., and D'Amours, D. (2009). Polo kinase regulates mitotic chromosome condensation by hyperactivation of condensin DNA supercoiling activity. *Mol Cell* 34, 416-426.
- Strunnikov, A.V., Hogan, E., and Koshland, D. (1995). SMC2, a *Saccharomyces cerevisiae* gene essential for chromosome segregation and condensation, defines a subgroup within the SMC family. *Genes Dev* 9, 587-599.
- Sullivan, N.L., Marquis, K.A., and Rudner, D.Z. (2009). Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell* 137, 697-707.
- Sutani, T., Sakata, T., Nakato, R., Masuda, K., Ishibashi, M., Yamashita, D., Suzuki, Y., Hirano, T., Bando, M., and Shirahige, K. (2015). Condensin targets and reduces unwound DNA structures associated with transcription in mitotic chromosome condensation. *Nature communications* 6, 7815.
- Takahashi, R., Hashimoto, T., Xu, H.J., Hu, S.X., Matsui, T., Miki, T., Bigo-Marshall, H., Aaronson, S.A., and Benedict, W.F. (1991). The retinoblastoma gene functions as a growth and tumor suppressor in human bladder carcinoma cells. *Proc Natl Acad Sci U S A* 88, 5257-5261.
- Takemoto, A., Kimura, K., Yokoyama, S., and Hanaoka, F. (2004). Cell cycle-dependent phosphorylation, nuclear localization, and activation of human condensin. *J Biol Chem* 279, 4551-4559.
- Takemoto, A., Maeshima, K., Ikehara, T., Yamaguchi, K., Murayama, A., Imamura, S., Imamoto, N., Yokoyama, S., Hirano, T., Watanabe, Y., *et al.* (2009). The chromosomal association of condensin II is regulated by a noncatalytic function of PP2A. *Nature structural & molecular biology* 16, 1302-1308.
- Talluri, S., Isaac, C.E., Ahmad, M., Henley, S.A., Francis, S.M., Martens, A.L., Bremner, R., and Dick, F.A. (2010). A G1 checkpoint mediated by the retinoblastoma protein that is dispensable in terminal differentiation but essential for senescence. *Mol Cell Biol* 30, 948-960.

- Tamrakar, S., and Ludlow, J.W. (2000). The carboxyl-terminal region of the retinoblastoma protein binds non-competitively to protein phosphatase type 1alpha and inhibits catalytic activity. *J Biol Chem* *275*, 27784-27789.
- Teichmann, M., Dieci, G., Pascali, C., and Boldina, G. (2010). General transcription factors and subunits of RNA polymerase III: Paralogs for promoter- and cell type-specific transcription in multicellular eukaryotes. *Transcription* *1*, 130-135.
- Terakawa, T., Bisht, S., Eeftens, J.M., Dekker, C., Haering, C.H., and Greene, E.C. (2017). The condensin complex is a mechanochemical motor that translocates along DNA. *Science* *358*, 672-676.
- Thalmeier, K., Synovzik, H., Mertz, R., Winnacker, E.L., and Lipp, M. (1989). Nuclear factor E2F mediates basic transcription and trans-activation by E1a of the human MYC promoter. *Genes Dev* *3*, 527-536.
- Thangavel, C., Boopathi, E., Liu, Y., Haber, A., Ertel, A., Bhardwaj, A., Addya, S., Williams, N., Ciment, S.J., Cotzia, P., *et al.* (2017). RB Loss Promotes Prostate Cancer Metastasis. *Cancer Res* *77*, 982-995.
- Tsai, P.F., Dell'Orso, S., Rodriguez, J., Vivanco, K.O., Ko, K.D., Jiang, K., Juan, A.H., Sarshad, A.A., Vian, L., Tran, M., *et al.* (2018). A Muscle-Specific Enhancer RNA Mediates Cohesin Recruitment and Regulates Transcription In trans. *Mol Cell* *71*, 129-141 e128.
- Tsang, C.K., Li, H., and Zheng, X.S. (2007a). Nutrient starvation promotes condensin loading to maintain rDNA stability. *EMBO J* *26*, 448-458.
- Tsang, C.K., Wei, Y., and Zheng, X.F. (2007b). Compacting DNA during the interphase: condensin maintains rDNA integrity. *Cell Cycle* *6*, 2213-2218.
- Van Bortle, K., Nichols, M.H., Li, L., Ong, C.T., Takenaka, N., Qin, Z.S., and Corces, V.G. (2014). Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome biology* *15*, R82.
- van den Heuvel, S., and Dyson, N.J. (2008). Conserved functions of the pRB and E2F families. *Nat Rev Mol Cell Biol* *9*, 713-724.
- Vian, L., Pekowska, A., Rao, S.S.P., Kieffer-Kwon, K.R., Jung, S., Baranello, L., Huang, S.C., El Khattabi, L., Dose, M., Pruett, N., *et al.* (2018). The Energetics and Physiological Impact of Cohesin Extrusion. *Cell* *173*, 1165-1178 e1120.
- Vietri, M., Bianchi, M., Ludlow, J.W., Mittnacht, S., and Villa-Moruzzi, E. (2006). Direct interaction between the catalytic subunit of Protein Phosphatase 1 and pRb. *Cancer cell international* *6*, 3.
- Walther, N., Hossain, M.J., Politi, A.Z., Koch, B., Kueblbeck, M., Odegard-Fougner, O., Lampe, M., and Ellenberg, J. (2018). A quantitative map of human Condensins provides

new insights into mitotic chromosome architecture. *The Journal of cell biology* 217, 2309-2328.

Wang, X., Brandao, H.B., Le, T.B., Laub, M.T., and Rudner, D.Z. (2017). *Bacillus subtilis* SMC complexes juxtapose chromosome arms as they travel from origin to terminus. *Science* 355, 524-527.

Wang, X., Tang, O.W., Riley, E.P., and Rudner, D.Z. (2014). The SMC condensin complex is required for origin segregation in *Bacillus subtilis*. *Current biology : CB* 24, 287-292.

Whyte, P., Buchkovich, K.J., Horowitz, J.M., Friend, S.H., Raybuck, M., Weinberg, R.A., and Harlow, E. (1988). Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. *Nature* 334, 124-129.

Whyte, P., Williamson, N.M., and Harlow, E. (1989). Cellular targets for transformation by the adenovirus E1A proteins. *Cell* 56, 67-75.

Wood, J.L., Liang, Y., Li, K., and Chen, J. (2008). Microcephalin/MCPH1 associates with the Condensin II complex to function in homologous recombination repair. *J Biol Chem* 283, 29586-29592.

Woodward, J., Taylor, G.C., Soares, D.C., Boyle, S., Sie, D., Read, D., Chathoth, K., Vukovic, M., Tarrats, N., Jamieson, D., *et al.* (2016). Condensin II mutation causes T-cell lymphoma through tissue-specific genome instability. *Genes Dev* 30, 2173-2186.

Wu, C.L., Zukerberg, L.R., Ngwu, C., Harlow, E., and Lees, J.A. (1995). In vivo association of E2F and DP family proteins. *Mol Cell Biol* 15, 2536-2546.

Wu, L., de Bruin, A., Saavedra, H.I., Starovic, M., Trimboli, A., Yang, Y., Opavska, J., Wilson, P., Thompson, J.C., Ostrowski, M.C., *et al.* (2003). Extra-embryonic function of Rb is essential for embryonic development and viability. *Nature* 421, 942-947.

Wutz, G., Varnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J., *et al.* (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J* 36, 3573-3599.

Xiao, Z.X., Ginsberg, D., Ewen, M., and Livingston, D.M. (1996). Regulation of the retinoblastoma protein-related protein p107 by G1 cyclin-associated kinases. *Proc Natl Acad Sci U S A* 93, 4633-4637.

Yan, J., Enge, M., Whittington, T., Dave, K., Liu, J., Sur, I., Schmierer, B., Jolma, A., Kivioja, T., Taipale, M., *et al.* (2013). Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* 154, 801-813.

- Yang, H., Williams, B.O., Hinds, P.W., Shih, T.S., Jacks, T., Bronson, R.T., and Livingston, D.M. (2002). Tumor suppression by a severely truncated species of retinoblastoma protein. *Mol Cell Biol* 22, 3103-3110.
- Yao, Y., and Dai, W. (2014). Genomic Instability and Cancer. *J Carcinog Mutagen* 5.
- Yoshimura, S.H., Hizume, K., Murakami, A., Sutani, T., Takeyasu, K., and Yanagida, M. (2002). Condensin architecture and interaction with DNA: regulatory non-SMC subunits bind to the head of SMC heterodimer. *Current biology : CB* 12, 508-513.
- Yuan, B., Pehlivan, D., Karaca, E., Patel, N., Charng, W.L., Gambin, T., Gonzaga-Jauregui, C., Sutton, V.R., Yesil, G., Bozdogan, S.T., *et al.* (2015). Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. *J Clin Invest* 125, 636-651.
- Yuen, K.C., and Gerton, J.L. (2018). Taking cohesin and condensin in context. *PLoS genetics* 14, e1007118.
- Yuen, K.C., Slaughter, B.D., and Gerton, J.L. (2017). Condensin II is anchored by TFIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. *Sci Adv* 3, e1700191.
- Yunis, J.J., and Ramsay, N. (1978). Retinoblastoma and subband deletion of chromosome 13. *Am J Dis Child* 132, 161-163.
- Zhang, T., Paulson, J.R., Bakhrebah, M., Kim, J.H., Nowell, C., Kalitsis, P., and Hudson, D.F. (2016). Condensin I and II behaviour in interphase nuclei and cells undergoing premature chromosome condensation. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 24, 243-269.
- Zhao, W., Huang, C.C., Otterson, G.A., Leon, M.E., Tang, Y., Shilo, K., and Villalona, M.A. (2012). Altered p16(INK4) and RB1 Expressions Are Associated with Poor Prognosis in Patients with Non-small Cell Lung Cancer. *J Oncol* 2012, 957437.
- Zheng, N., Fraenkel, E., Pabo, C.O., and Pavletich, N.P. (1999). Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP. *Genes Dev* 13, 666-674.
- Zhu, L., van den Heuvel, S., Helin, K., Fattaey, A., Ewen, M., Livingston, D., Dyson, N., and Harlow, E. (1993). Inhibition of cell proliferation by p107, a relative of the retinoblastoma protein. *Genes Dev* 7, 1111-1125.
- Zhu, L., Zhu, L., Xie, E., and Chang, L.S. (1995). Differential roles of two tandem E2F sites in repression of the human p107 promoter by retinoblastoma and p107 proteins. *Mol Cell Biol* 15, 3552-3562.

Chapter 2

2 *RB1* deletion in pRB-pathway disrupted cells results in DNA damage and cancer progression.

2.1 Abstract

Proliferative control in cancer cells is frequently disrupted by mutations in the pRB-pathway. Intriguingly, *RB1* mutations can arise late in tumorigenesis in cancer cells whose pRB-pathway is already compromised by another mutation. In this study, we present evidence for increased DNA damage and instability in cancer cells with pRB-pathway defects when *RB1* mutations are induced. We generated isogenic *RB1* mutant genotypes with CRISPR/Cas9 in a number of cell lines. Cells with even one mutant copy of *RB1* have increased basal levels of DNA damage and increased mitotic errors. Elevated levels of reactive oxygen species as well as impaired homologous recombination repair underlie this DNA damage. When xenografted into immune compromised mice *RB1* mutant cells exhibit an elevated propensity to seed new tumors in recipient lungs. This study offers evidence that late arising *RB1* mutations can facilitate genome instability and cancer progression that are beyond the pre-existing proliferative control deficit.

2.2 Introduction

Loss of proliferative control is a defining feature of human cancer. Most cancer cells develop cell intrinsic mechanisms of supplying growth stimulatory signals as well as disrupting the response to cell cycle arrest cues (Hanahan and Weinberg, 2011). To this end, mutations in the retinoblastoma protein (pRB)-pathway are central to disrupting proliferative control in tumorigenesis (Burkhardt and Sage, 2008; Knudsen and Knudsen, 2008; Sherr and McCormick, 2002). Deletion of the pRB gene, *RB1*, prevents cell cycle arrest in response to a broad range of signals (Knudsen and Knudsen, 2008). Similarly, overexpression or hyperactivation of D-type cyclins and their associated cyclin-dependent kinases (CDKs) can lead to constitutive pRB phosphorylation and cell cycle entry. Lastly, deletion or promoter methylation of *CDKN2A* that encodes the CDK inhibitor p16 serves to deregulate kinase activity, causing constitutive phosphorylation of pRB. Cancer cell

genomes that sustain a single mutation in this pathway are considered to have disrupted pRB-pathway function and are deficient for cell cycle control (Dyson, 2016; Knudsen and Knudsen, 2008; Sherr, 1996). Historically, this concept of pRB-pathway inactivation suggested that mutations in different components of the pathway are relatively equivalent and additional mutations provide no subsequent advantage to cancer progression (Dick et al., 2018; Knudsen and Knudsen, 2008; Sherr, 1996; Sherr and McCormick, 2002).

A number of recent clinical observations challenge the logic of single pRB-pathway mutations in cancer. First, multiple studies have shown that *RBI* loss is specifically predictive of a favourable response to chemotherapy (Cecchini et al., 2015; Garsed et al., 2018; Ludovini et al., 2004; Zhao et al., 2012), whereas p16 expression or overall proliferative rates are not (Cecchini et al., 2015; Garsed et al., 2018; Zhao et al., 2012). This suggests that pRB-pathway mutations are not necessarily equivalent. Second, a number of studies have suggested that *RBI* gene loss is more prevalent in advanced cancers, or mechanistically contribute to progression or dissemination (Beltran et al., 2016; McNair et al., 2017; Robinson et al., 2017; Thangavel et al., 2017), a stage where cell autonomous proliferative control is presumably already deregulated. Collectively, these examples suggest that *RBI* mutation contributes more to tumor progression than just alterations to proliferative control and that *RBI* loss may confer other cancer relevant characteristics. Remarkably, some studies even highlight that single copy loss of *RBI* may be functionally significant (Coschi et al., 2014; Gonzalez-Vasconcellos et al., 2013; McNair et al., 2017; Zheng et al., 2002).

Beyond pRB's role in cell cycle control through E2F transcriptional regulation, it has been reported to participate in a host of functions that contribute to genome stability (Velez-Cruz and Johnson, 2017). These include chromosome condensation through pRB-dependent recruitment of condensin II and cohesin (Longworth et al., 2008; Manning et al., 2014). The pRB protein also influences repair of DNA breaks through both non-homologous end joining (NHEJ) (Cook et al., 2015), and homologous recombination (HR) (Velez-Cruz et al., 2016), and induction of mitochondrial biogenesis that impacts cell metabolism (Benevolenskaya and Frolov, 2015; Jones et al., 2016; Nicolay et al., 2015). Some of these functions, such as repair of DNA breaks by HR, are obligatorily

outside of pRB's role in G1-S phase regulation. In addition, other roles, such as effects on mitochondrial biogenesis and metabolism take place in proliferating populations of cells further suggesting that this is independent of G1-S regulation and the pRB-pathway. It is noteworthy that some atypical pRB functions in genome stability, or late stage cancer progression, may be sensitive to single copy loss (Coschi et al., 2014; Gonzalez-Vasconcellos et al., 2013; Sharma et al., 2010; Zheng et al., 2002). Thus, the existence of shallow *RBI* deletions may indicate that pRB's less well appreciated functions in genome stability could underlie cancer relevant characteristics that are independent of classical pRB-pathway function in cancer (Dick et al., 2018).

In order to test if *RBI* loss is relevant to cancer cells that already possess pRB-pathway disruption, we induced mutations in *RBI* using CRISPR/Cas9 in U2OS and NCI-H460 cell lines, that are reported to be defective for p16, as well as in the NCI-H1792 cell line, which possesses a gain-of-function alteration affecting CDK4. These cells displayed spontaneous DNA damage as evidenced by γ H2AX foci and elevated levels of reactive oxygen species. We also determined that *RBI* mutations decreased the ability to repair DNA breaks by homologous recombination, and this is supported by elevated levels of anaphase bridges in mitosis. *RBI* mutant cells were xenografted into immune compromised mice and this revealed similar growth kinetics in subcutaneous implantation, with *RBI* null showing greater propensity to colonize lungs. These experiments underscore the discovery that *RBI* mutation in cells that already possess pRB-pathway disruption creates DNA damage and fuels cancer progression.

2.3 Materials and Methods

2.3.1 Cell culture

U2OS cells and the resulting clones were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 50 U/mL penicillin and 50 μ g/mL streptomycin. H460 and H1792 cells and the resulting clones were grown in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% FBS, 2 mM L-glutamine, 50 U/mL penicillin and 50 μ g/mL streptomycin. Cells were grown at 37°C in humidified air containing 5% CO₂.

2.3.2 Generation of *RB1* deletions using CRISPR

For creation of *RB1* deletions, single guide RNAs (sgRNAs) targeting exon 22 of *RB1* were selected by using the CRISPR Design tool at <http://crispr.mit.edu/> (Cong et al., 2013). The sgRNA sequences were as follows: 5'-CACCGTATTATAGTATTCTATAACT-3' (X22B-top), 5'-AAACAGTTATAGAATACTATAATAC-3' (X22B-bottom), 5'-CACCGAGGATACTTTTGACCTACCC-3' (X22C-top), 5'-AAACGGGTAGGTCAAAAGTATCCTC (X22C-bottom). The X22B and X22C guides were each cloned into the pX459 plasmid (with wild type Cas9; Addgene #48139) and the pX462 plasmid (with the D10A mutant version of Cas9; Addgene #48141) Both plasmids contain a puromycin resistance cassette. Cells were seeded at a density of 10⁵ cells per well in a 6-well dish, and the next day a total of 1 µg per well of a 1:1 mix of X22B and X22C CRISPR plasmids (either pX459 or pX462) was transfected by use of X-tremeGENE HP DNA transfection reagent (Roche). The next day, each well was replated onto a 15 cm dish, and the day after that cells were cultured in selection medium with 2 µg/mL of puromycin for 2 days. Following that, cells were grown in normal cell culture media for approximately 12 days, following which single colonies were picked from the 15 cm plates using mechanical detachment with a pipette tip and placed into wells of a 48-well dish and allowed to grow. Cells were further passaged onto larger plates, and were genotyped using the following primers: X22F primer, 5'-TTACTGTTCTTCCTCAGACATTCAA-3'; and X22R primer, 5'-GGATCAAATAATCCCCCTCTCAT-3'. PCR products (445 bp for the wild type band) were run in an agarose gel, individual bands were gel purified using a Sigma GenElute Gel Extraction kit and sent for Sanger sequencing using the X22F and X22R primers shown above. For clones with multiple PCR products, bands were purified and sent for sequencing separately to determine individual *RB1* alleles in each clone. For alleles that could not easily be resolved by gel electrophoresis, PCR products were cloned into vectors using either the TOPO TA Cloning Kit for Sequencing (Invitrogen) or the CloneJET PCR Cloning Kit (Thermo Scientific).

The top scoring off-target intragenic locations determined for each sgRNA using the CRISPR Design tool were also sequenced to probe for mutations. sgRNA X22B had a potential off-target site in *ZNF699* (X22B_OT_ZNF699_F, 5'-GTGCCCTAAAACACTGAGGGA-3'; and X22B_OT_ZNF699_R, 5'-TTTATGATCAACAAGGACCAGAGC-3') while X22C has a potential off-target site *ALDH1L1* (X22C_OT_ALDH1L1_F, 5'-GCCACGCTATGCTTGTGATG-3'; and X22C_OT_ALDH1L1_R, 5'-CACCCCAGAGAAGGGAACAC -3'). PCR products were gel purified as above and sent for Sanger sequencing using their respective primers.

Nuclear extracts were prepared from U2OS CRISPR clones of interest and western blotting was carried out using previously described protocols (Cecchini and Dick, 2011). Antibodies raised against pRB (clone G3-245, BD Pharmingen; C-15, Santa Cruz) and Sp1 (H-225, Santa Cruz) were used for western blotting. Samples were western blotted using standard techniques.

To generate additional *RBI* knockout and control cell lines, sgRNAs targeting either exon 2 of *RBI* (5'-GGAGAAAGTTTCATCTG-3') or a gene desert region of the genome (5'-TGAGCCTATATTAATTGG-3') were utilized. The sgRNAs were cloned into the lentiCRISPR v2 vector (Addgene #52961), which also encodes Cas9. To generate lentivirus, 293T cells were transfected with the sgRNA vector and a 1:1:1 mixture of lentiviral packaging constructs (Addgene #12251, #12253, #8454) using polyethylenimine transfection reagent. Twenty-four hours after transfection, the 293T media was replaced, and recipient cells (U2OS, NCI-H1792, NCI-H460) were seeded for infection. The following day, media on the recipient cells was replaced with lentiviral media, and polybrene was added at a final concentration of 8 µg/mL. A second infection was performed the next day. Infected cells were then selected with 2 µg/mL puromycin for 3 days. To generate isogenic clones, populations of knockout (or control) cells were FACS sorted as single cells in 96-well plates (BD FACSAria II) and allowed to grow for approximately 2 weeks. Colonies were then expanded and screened for loss of pRB by immunoassay using the Simple Western™ system according to the manufacturer's instructions. Successful knockout clones were also genotyped to confirm clonogenic origin. Genomic DNA was extracted using the PureLink Genomic DNA Mini Kit

(Invitrogen), and the region surrounding the cut site was amplified by PCR using the following primers: X2F, 5'-TCACAGAAGTGTTTTGCTGCTT-3'; X2R, 5'-TTTGGTGGGAGGCATTTATGGA-3'. PCR products were purified using the DNA Clean and Concentrator Kit (Zymo Research) and sent for Sanger sequencing.

2.3.3 Fluorescence microscopy

Cells grown either on glass coverslips or in glass bottom plates were fixed in phosphate-buffered saline (PBS) containing 4% paraformaldehyde for 10 min and then permeabilized with PBS-0.3% Triton X-100 for 10 min at room temperature. The fixed cells were blocked in blocking buffer (PBS-0.3% Triton X-100 with either 5% donkey or goat serum depending on the species in which the secondary antibodies were raised) for at least 1 hr at room temperature. Cells were then incubated with primary antibody in blocking buffer at room temperature for 1 hr or at 4°C overnight. Antibodies raised against pRB (clone G3-245, BD Pharmingen), γ H2AX (clone JBW301, EMD Millipore), 53BP1 (H-300, Santa Cruz) and BLM (C-18, Santa Cruz) were used for IF. After 3 washes with PBS-0.3% Triton X-100, cells were incubated with secondary antibody diluted in blocking buffer for 1 hr at room temperature. Cells were washed twice with PBS-0.3% Triton X-100, incubated with 100 ng/mL 4',6-diamidino-2-phenylindole (DAPI) in PBS-0.3% Triton X-100 for 5 min, washed twice more with PBS-0.3% Triton X-100 and then washed once with PBS before mounting with Slowfade Gold Antifade mountant (S36936, ThermoFisher Scientific).

For 8-oxoguanine (8-oxoG) staining, cells were fixed and blocked as above, then washed with PBS-0.3% Triton X-100 and incubated in RNase solution (0.2 mg/mL RNase A, 10 mM Tris-HCl (pH 7.5), 15 mM NaCl, 0.1% Triton X-100 in 1X PBS) for 1 hr at room temperature. Cells were washed with PBS-0.3% Triton X-100 and then incubated in 2 M HCl for 10 min at room temperature, followed by a rinse with 50 mM Tris-HCl (pH 8.0). Cells were washed with PBS-0.3% Triton X-100, and then primary antibody incubation, using α -DNA/RNA Damage antibody raised against 8-oxoG (clone 15A3, ab62623, Abcam) and all subsequent steps were completed as above.

For confocal microscopy of pRB, γ H2AX and BLM, cells were examined on an Olympus Fluoview FV1000 confocal microscope system. For confocal microscopy of 53BP1, a Nikon A1R confocal microscope was used. For non-confocal microscopy, images were acquired using a Zeiss Axioskop 40 microscope and Spot flex camera. Foci were quantified using the Focinator (Oeck et al., 2015), while overall staining intensity in cells was quantified by ImageJ (Schneider et al., 2012).

2.3.4 Gamma irradiation of cells

Cells subjected to γ IR were plated at 100,000 cells per well in 6 well dishes with glass coverslips on the bottom. The next day, cells were exposed to a cobalt 60 source until a dose of either 2 Gy or 1 Gy was received. Cells were placed back in the cell culture incubator until the appropriate time point after treatment to fix cells for IF.

2.3.5 NHEJ and HR repair assays

For the HR repair assay, pDRGFP was used which was a gift from Maria Jasin (Addgene plasmid #26475; <http://n2t.net/addgene:26475>; RRID:Addgene_26475) and for the NHEJ assay, pimEJ5GFP was used which was a gift from Jeremy Stark (Addgene plasmid #44026; <http://n2t.net/addgene:44026>; RRID:Addgene_44026). pDRGFP was linearized using EcoRV and pimEJ5GFP was linearized using XhoI. These linearized fragments were then individually used for transfection using Lipofectamine 3000 transfection reagent (Invitrogen) into U2OS cells. The next day, each well was replated onto a 10 cm dish, and a day later, cells were cultured in selection medium with 2 μ g/mL of puromycin for 3 days. To isolate single cell colonies, limiting dilutions were then used to seed cells into 96 well plates. After approximately 3 weeks, wells with growth from single cell isolates were transferred to single wells of 12 well plates and after a few days were treated with puromycin again to ensure the selected clones did still contain either the NHEJ or HR constructs.

To determine the reporter efficiency in the isolated clones, 2 sets of transfections were performed per clone, again using Lipofectamine 3000 transfection reagent (Invitrogen). For the first set of transfections, each clone was transfected with a plasmid expressing the I-SceI endonuclease, pCBASceI, which was a gift from Maria Jasin

(Addgene plasmid #26477; <http://n2t.net/addgene:26477>; RRID:Addgene_26477), and a blasticidin marker, pMSCV-Blasticidin, which was a gift from David Mu (Addgene plasmid #75085 ; <http://n2t.net/addgene:75085>; RRID:Addgene_75085). The second set of transfections was with an empty backbone plasmid, pCAG-FALSE, which was a gift from Wilson Wong (Addgene plasmid #89689; <http://n2t.net/addgene:89689>; RRID:Addgene_89689) and pMSCV-Blasticidin. For both of these transfections, the blasticidin resistance plasmid was used in a 1:3 ratio with the complementary plasmid. The next day, each well was replated onto a 10 cm dish, and the day after that cells were cultured in selection medium with 10 µg/mL of blasticidin for 1 week. GFP positive cells were then quantified by flow cytometric analysis (FACS). To prepare cells for FACS, they were washed with PBS, trypsinized, resuspended in culture media, and washed twice with PBS. Cell pellets were then resuspended in 0.5 mL of flow cytometry staining buffer with propidium iodide (0.05% sodium azide and 0.5% BSA in 1X PBS with 0.01 mg/mL propidium iodide). For each reporter construct, the clone with the highest ratio of GFP signal when transfected with pCBASceI to GFP signal when transfected with pCAG-FALSE was selected for future studies.

To introduce CRISPR constructs into selected clones for each repair reporter, lentivirus particles were generated in HEK293T cells. Lentivirus was created for both lentiCRISPR v2 with no guide RNA inserted, and for lentiCRISPR v2 with the X22B sgRNA sequences for *RBI* (from above) inserted. lentiCRISPR v2 was a gift from Feng Zhang (Addgene plasmid #52961; <http://n2t.net/addgene:52961>; RRID:Addgene_52961). The X22B *RBI* guide sequences were inserted into the lentiCRISPR v2 plasmid as previously described (Sanjana et al., 2014; Shalem et al., 2014). Culture media containing lentiviral particles were transferred to appropriate U2OS HR and NHEJ reporter clones for 48 hours, followed by selection with 4 µg/mL puromycin for at least 5 days.

The population of U2OS HR and NHEJ reporter clones that were infected with lentiCRISPR v2 plasmids, either with or without the *RBI* guide, were then transfected with both pCBASceI and pMSCV-Blasticidin or pCAG-FALSE and pMSCV-Blasticidin and selected, as above, and analyzed by FACS to determine repair efficiency. These transfections were performed in the same transduced population of cells, but for

experimental replicates, the cells were seeded at different times and then subsequently transfected with plasmids. Cells grown in parallel to the transfected cells were used to prepare nuclear extracts for western blotting.

2.3.6 Expression of HR factors

Total RNA from cells was isolated using TRIzol reagent according to the standard protocol (Invitrogen). First-strand cDNA synthesis was performed using iScript cDNA Synthesis Kit (Bio-Rad). Isolated cDNA was used in qRT-PCR reactions with iQ SYBR Green Supermix (Bio-Rad) using the following primers: RPA2_F, 5'-CGAAAGCTATGGCAGCTCCT-3'; RPA2_R, 5'-GGCTCGGGCTCTTGATTTCT-3'; RAD54B_F, 5'-GCGAGGGGATAGCTGGTTAC-3'; RAD54B_R, 5'-AGTCGTGACCGGCGAAAAT-3'; BLM_F, 5'-GAGTCTGCGTGCGAGGATTA-3'; BLM_R, 5'-AGTGTCTGGCTGAGTGACG-3'; RAD51_F, 5'-AGCTGGGAAGTCAACTCAT-3'; RAD51_R, 5'-CCACACTGCTCTAACCGTGA-3'; RECQL_F, 5'-AGAGAAAGCCTATGAAGCAAGGA-3'; RECQL_R, 5'-GGCTTCTGCCGAACCTCATA-3'; BRCA1_F, 5'-CTGAAGACTGCTCAGGGCTATC-3'; BRCA1_R, 5'-AGGGTAGCTGTTAGAAGGCTGG-3'; XRCC2_F, 5'-GCGATGTGTAGTGCCCTCCA-3'; XRCC2_R, 5'-TTCAAGAATATCACCATGCACAGG-3'; BRCA2_F, 5'-AAGCACTCCAGATGGCACAAT-3'; BRCA2_R, 5'-GGGTACACAGGTAATCGGCT-3'; RAD52_F, 5'-ATGCTTTGGACAGTGCCAGT-3'; RAD52_R, 5'-ACATTCTGCTGCGTGATGGA-3'; GAPDH_F, 5'-ATGACCACAGTCCATGCCAT-3'; and GAPDH_R, 5'-TTGAAGTCAGAGGAGACCAC-3'. Resulting target Cq values were normalized to GAPDH, then expressed as fold change relative to the global wild type mean.

Nuclear extracts were prepared from clones for western blots, and the following antibodies were used: RPA32 (A300-244A, Bethyl), RAD54B (ab83311, Abcam), BLM (C-18, Santa Cruz), and RAD51 (ab63801, Abcam).

2.3.7 Determination of IC50 concentrations

For IC50 assays, U2OS cells were seeded at a density of 1,200 cells per well and H460 and H1792 cells were seeded at a density of 1,500 cells per well in 96 well dishes. Twenty-four hours after plating cells, media was replaced with media containing the drugs of interest at the appropriate concentrations. Technical triplicates were analyzed for each biological replicate. Serial dilutions of stock solutions of aphidicolin (APH), hydrogen peroxide (H₂O₂), etoposide, hydroxyurea (HU) and cisplatin were created so that a constant amount of drug was added to the media for each drug concentration used. After 72 hr, alamarBlue was added to an amount equal to 10% of the volume in the well (i.e. 10 μ L per well with 100 μ L of media and drug). After 4 hr of incubation, cytotoxicity was measured using a Synergy H4 Hybrid Reader (BioTek, USA) using excitation/emission wavelengths of 560 nm/590 nm. Values were corrected using a blank of media and alamarBlue only. The amount of fluorescence of alamarBlue for each well of drug treated cells was then normalized to the fluorescence value obtained for the untreated cells of the same technical replicate. These normalized fluorescence values relative to untreated cells were then analyzed using Prism. The drug concentrations were log transformed and the data were subsequently fit to a curve using nonlinear regression (log(inhibitor) vs. response (three parameters)). IC50 values were obtained from the best fit values, and IC50 values from three biological replicates were compared using Ordinary one-way ANOVA and Tukey's multiple comparisons test or paired *t*-test.

2.3.8 ChIP-Sequencing

ChIP was conducted according to protocols adapted from Cecchini *et al.* (Cecchini et al., 2014). Briefly, cross-linked chromatin was sonicated so most chromatin was \leq 400 bp. Sheared chromatin was then normalized between experimental groups and pre-cleared with protein G Dynabeads and IgG. Pre-cleared chromatin was then incubated with protein G Dynabeads and ChIP antibodies to immunoprecipitate proteins. Antibodies raised against γ H2AX (clone JBW301, EMD Millipore) and H4 (clone 62-141-13, EMD Millipore) were used for ChIP. Cross-links were reversed at 65°C, and samples were treated with RNase and proteinase K. DNA was isolated for library preparation, and 20 replicates per genotype for γ H2AX ChIP-Seq were pooled to achieve DNA yield required

for library preparation (NEBNext Ultra II DNA Library Prep Kit). ChIP libraries were sequenced using an Illumina NextSeq (High output 75 cycle kit), and processed reads are deposited in GEO (GSE125379).

Resulting FASTQ reads were aligned to the human genome build hg19 using Bowtie2 version 2.3.0 (Langmead and Salzberg, 2012). The following command was used: `bowtie2 -t -p 4 -D 15 -R 2 -L 32 -i S,1,0.75 -x hg19 -U <reads>.fastq -S <output>.sam`. Peaks were identified using MACS2 version `macs2 2.1.1.20160309` according to parameters stated below and the options to detect broad peak distributions for histone marks (Zhang et al., 2008). For H4 ChIP-Seq, the corresponding inputs were used as the control and for γ H2AX ChIP-Seq, the first input replicate was used as the control. The following command was used: `macs2 callpeak -t <ChIP>.bam -c <input>.bam -n <output> --outdir ./macs2/ -g hs --broad --broad-cutoff 0.1`.

To find abundance of ChIP-Seq reads in common fragile sites (CFS), the cytogenetically determined locations of CFS, as determined previously by Lukusa and Fryns (2008), were converted to human genomic coordinates (hg19) using the UCSC Genome Browser (Lukusa and Fryns, 2008; Tyner et al., 2017). Bedtools coverage was then used to find the number of alignments for each ChIP-Seq sample within the individual CFS (Quinlan, 2014). The abundance of reads that mapped to CFS were then converted to proportions by dividing by the total number of mapped reads. The proportion data were further normalized against input control and then ratios were made comparing the mutant proportions to the wild type proportions. A two-tailed one-sample t-test was performed to test if the normalized mean read count proportions of the *RBI*^{+/-} and the *RBI*^{-/-} ChIP-Seq assays at each of the CFS is equal to the normalized read count proportion of the corresponding CFS from the wild type control. A multi-test correction was applied to the calculated *P*-values (using "fdr" method from "p.adjust" function in R). Statistical analysis of sequence data was performed using R (version 3.4.2) and the plotting function used was lattice (v0.20-35).

For repeat analysis, another set of alignments were performed. For this analysis, reads were mapped using Bowtie version 1.2.1.1 with high stringency to the hg19

genome (Langmead et al., 2009). The following command was used: bowtie -S --best -m 1 --chunkmbs 500 -p 4 -t --un <not_aligned_unique> --max <multiple_reads_unique> hg19 <reads>.fastq <output>.sam. All remaining reads were mapped to repeat containing indexes using previously reported methods (Day et al., 2010), with indices also being derived from Repbase and Tandem Repeats Databases (Bao et al., 2015; Gelfand et al., 2007). For these remaining repeat alignments, the -m 1 parameter of the Bowtie mapping was changed to -k 1. Finally, all remaining reads were re-mapped to hg19 at low stringency to exhaustively match sequence tags to the mouse genome. The abundance of sequence tags that mapped to non-unique regions of the genome were compared by using log₂ ratios of γ H2AX precipitable tags per million mapped reads in mutant versus wild type and converted into heat maps using matrix2png (Pavlidis and Noble, 2003). To test for significance of enrichment of reads mapped to various repeat categories, the same analysis to test for significance within CFS was used (see above).

2.3.9 Flow cytometry

Cells were plated on 6 cm plates at a density of 100,000 cells per plate. Approximately 24 hr after, cells were pulsed with BrdU for a duration of 30 min. Cell cycle analysis was then carried out as previously described (Cecchini et al., 2012).

2.3.10 Nucleoside supplementation

For nucleoside complementation, cells were seeded at a density of 50,000 cells per 6 cm plate for flow cytometry and at 25,000 cells per 6 well plate with glass coverslips for immunofluorescence. Approximately 24 hr after seeding cells, media was replaced, either with or without the addition of nucleosides. To prepare nucleosides, uridine (Sigma) and cytidine (Sigma) were dissolved in autoclaved Milli-Q water to make 10 mM stocks, while adenosine (Sigma) and guanosine (Sigma) were dissolved to make 2 mM stocks. The suspensions were briefly boiled, filter sterilized, and added to complete medium at a final concentration of either 50 μ M or 10 μ M. 48 hr after nucleoside addition, cells were either fixed for γ H2AX IF using non-confocal microscopy and the Focinator (as above), or for flow cytometry. For analysis of DNA content by flow cytometry, propidium iodide stained DNA content was analyzed.

2.3.11 Measurement of reactive oxygen species

Cells were plated in 96 well plates at a density of 1,200 cells per well in DMEM without phenol red (31053-028, ThermoFisher Scientific). H₂O₂ was added 24 hr after seeding cells. Seventy-two hours later, 5(6)-carboxy-2',7'-dichlorodihydrofluorescein diacetate (carboxy-H2DCFDA; CA-DCF-DA; (C400, ThermoFisher Scientific)) at a stock concentration of 20 mM in DMSO was diluted in DMEM without phenol red to a concentration of 40 µM. This master mix of media and CA-DCF-DA was added directly to the wells already containing media and the drug of interest, to obtain a final concentration of 20 µM CA-DCF-DA. The cells were then put in the incubator for 45 min and readings were obtained using a Synergy H4 Hybrid Reader (BioTek, USA) using excitation/emission wavelengths of 492 nm/525 nm. Technical triplicates were analyzed for each biological replicate and the average background readings (cells treated with the highest concentration of the drug of interest for 72 hr and DMSO in place of CA-DCF-DA) from each cell line were subtracted from the average of each treatment reading for analysis of fluorescence.

2.3.12 Mouse xenografts

U2OS clones were grown in cell culture to approximately 80% confluence. Cells were washed with PBS, trypsinized, centrifuged and washed 3 times with Hanks' Balanced Salt Solution (HBSS, 1X). Cells were then resuspended in HBSS at a concentration of 5x10⁶ cells/mL so that 200 µL contained the 1x10⁶ cells required for each injection.

For subcutaneous injections, mice were approximately 8 weeks old and for the tail vein injections mice were approximately 13 weeks old when injected. All mice were given at least 3 days to acclimatize. All mice were female NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (stock number 005557, The Jackson Laboratory) and were housed and handled as approved by the Canadian Council on Animal Care, under an approved protocol (2016-068).

For subcutaneous injections, 1x10⁶ *RBI*^{+/+} cells were injected into the left flank of all mice, and 1x10⁶ *RBI*^{+/-} or *RBI*^{-/-} cells were injected into the right flank. The mice used

for the subcutaneous injections were euthanized approximately 8 weeks after injection. Necropsies were performed and tumor mass was determined. Tumors were then fixed in formalin for 48 hr and processed for histological assessment.

For tail vein injections, 1×10^6 cells were injected into the lateral tail vein. Mice were euthanized 8 weeks after tail vein injection. All animals were subjected to a thorough necropsy and all lungs as well as any abnormal tissues or organs were fixed in formalin for 48 hr and processed for histological assessment.

All tissues of interest from both studies were embedded, sectioned and stained with hematoxylin and eosin according to standard methods. Slides were imaged using an Aperio ScanScope slide scanner (Leica Biosystems).

For quantitative pathology of lungs from tail vein injected mice, images were analyzed using QuPath (Bankhead et al., 2017). Briefly, annotations were drawn around each individual lung. Within these annotated lungs, cells were detected using the cell detection command. The features within these cells were then smoothed by using the add smoothed features command (using $25 \mu\text{m}$ as the radius). Within the lungs, regions containing different cell types were annotated and these annotations were used to train a cell classifier. All possible 67 cell features were used to build the random trees classifier, using default parameters. A script was then made to determine the total cell area of all cell types called by the classifier within each lung, and the percentages of tumor cell area was calculated from these values. Tumor cell nodules were manually counted using the cell types determined by the classifier; anything thought to have derived from a single cell seeding event was considered a tumor cell nodule.

To determine the *RBI* genotype of seeded U2OS cells of interest, embedded mouse lung tissue was deparaffinized, lysed, formalin crosslinks were reversed, and DNA was isolated according to manufacturer's instructions (QIAamp DNA FFPE Tissue Kit, Qiagen). DNA was genotyped as above using PCR using genotyping primers (X22F and X22R).

2.3.13 Data extraction from cBioPortal

Only TCGA studies used for the Pan-Cancer Atlas with 150 samples or more on cBioPortal were selected to query (Cerami et al., 2012; Gao et al., 2013). Data from cBioPortal was obtained in January 2019. Mutation and CNA data was analyzed, with the gene set user-defined list being entered as “RB1: AMP HOMDEL HETLOSS mut”.

2.4 Results

2.4.1 Spontaneous DNA damage in *RB1* deficient cancer cells

To investigate *RB1* deficiency in pRB-pathway disrupted cells, we utilized a number of cell lines that are reported to be defective for p16, the product of *CDKN2A* gene, or that possess activation of cyclin D/CDK4. Initially we used p16 deficient U2OS cells (Forbes et al., 2017), and Cas9 with single guide RNA (sgRNA) pairs that target exon 22 of *RB1* because loss of this exon creates null alleles in cancer (Horowitz et al., 1990). Cells were transfected with plasmids to deliver pairs of sgRNAs and Cas9 (wild type or the D10A mutant). Following transient drug selection, colonies were isolated, expanded, and genotyped by PCR to search for *RB1* deletions (Figure 2.1A). Candidates were rigorously selected by checking pRB protein expression by western blotting (Figure 2.1B), ensuring heterozygous clones were not mixtures of wild type and knockout cells using fluorescent pRB staining (Figure 2.1C), and confirming that the most likely off targets were not mutated (Table 2.1). Using this approach, we selected four clones each for wild type and knockout *RB1* genotypes, and three clones for the *RB1*^{+/-} genotype that were used in subsequent experiments.

To determine if *RB1* mutation status affects genome stability in these engineered cell lines, DNA damage was assessed in untreated, proliferating cells by staining for γ H2AX. Foci were visualized by immunofluorescence microscopy and images were captured using confocal microscopy (Figure 2.1D). The quantity of foci per nucleus was determined and this revealed a significant increase in γ H2AX in the knockout and heterozygous lines compared to those that are wild type for *RB1* (Figure 2.1E).

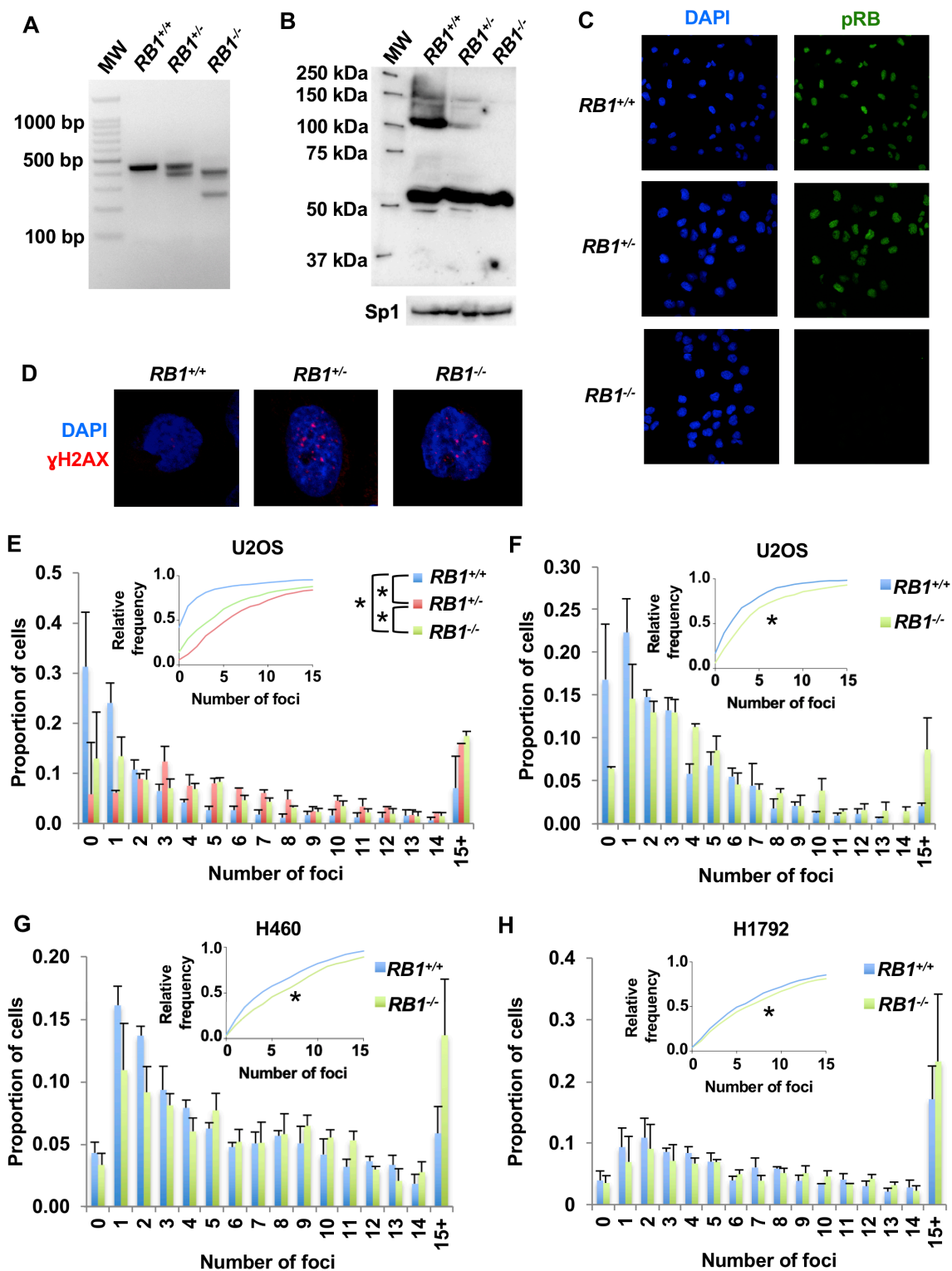


Figure 2.1: CRISPR/Cas9 induced mutations in *RB1* cause DNA damage.

Figure 2.1: CRISPR/Cas9 induced mutations in *RB1* cause DNA damage.

(A) An ethidium bromide stained, agarose gel shows examples of wild type, heterozygous, and homozygous mutant *RB1* genotypes that are detected by PCR amplification of exon 22 sequences. (B) A representative western blot showing pRB expression in control, heterozygous, and homozygous mutant cells is shown on top. Sp1 loading control is shown on the bottom. (C) Immunofluorescence microscopy was used to detect pRB expression (green) in cultures of control, heterozygous, or homozygous mutants. Cells were counterstained with DAPI to visualize nuclei (blue). (D) Representative confocal microscopy images of γ H2AX foci (red) in control, heterozygous, and homozygous *RB1* mutant cells. Cells were counterstained with DAPI to visualize nuclei (blue). (E) γ H2AX foci counts for each of the U2OS *RB1* genotypes. The average proportion of cells with discrete numbers of foci are shown as histograms, while the cumulative frequency of foci for each genotype is shown as an inset. The average distribution of foci for *RB1* wild type (4 different clones), heterozygous (3 different clones) and knockouts (4 different clones) were compared using the Kolmogorov-Smirnov test. (F) U2OS cells were transfected with CRISPR/Cas9 constructs targeting either a safe harbor site in the genome or exon 2 of the *RB1* gene. Three clones were selected for both control and knockout conditions and γ H2AX foci were quantified by fluorescence microscopy. The average proportion of γ H2AX foci for both *RB1* wild type and knockout genotypes are shown as histograms, while the cumulative relative frequency of foci is shown as an inset. Foci distributions were again compared by Kolmogorov-Smirnov test. (G) H460 lung cancer cells were stained for γ H2AX foci and analyzed as in F. (H) H1792 non-small cell lung cancer cells were analyzed as in F. All error bars are +1 SEM. * $P < 0.05$.

Table 2.1: Characterization of the U2OS *RB1* mutant clones generated by targeting exon 22.

Clone	Cas9	<i>RB1</i> Genotype ^A	Predicted protein ^B	Off-Target Genotypes ^C	
				<i>ALDH1L1</i>	<i>ZNF699</i>
U2OS	N/A	wild type	WT WT	N/A	N/A
21B2	D10A	wild type	WT WT	ND	ND
21B4	D10A	wild type	WT WT	WT	WT
11B5	D10A	wild type	WT WT	ND	ND
4C1	D10A	heterozygous	WT p.Y771fsX8	WT	WT
21D5	D10A	heterozygous	WT p.Q770HdelX10	WT	WT
21A1	D10A	heterozygous	WT p.I753delX12	WT	WT
12C3	D10A	null	p.I752IfsX9 p.E748CdelX11	WT	WT
5A5	WT	null	p.N757TX31 p.P776VdelX6	WT	WT
6B1	WT	null	p.Y756YfsX14 p.S751SfsX10	WT	WT
5C4	WT	null	p.S773FfsX14 p.Y756delX	WT	WT

^A Genotypes were determined by PCR and sequencing.

^B Amino acid coding changes were predicted based on nucleotide sequences.

^C The top scoring off-target intragenic locations determined for each sgRNA, *ZNF699* for sgRNA X22B and *ALDH1L1* for sgRNA X22C, were also sequenced to probe for unwanted mutations. ND=not determined.

We quantified γ H2AX foci abundance in additional pRB-pathway mutant cell lines to determine if increased DNA damage is a common consequence of pRB loss. We used cells that were modified with control non-targeting sgRNAs or that were produced by targeting exon 2 of *RBI*. pRB deficiency in these cells was confirmed by western blotting (data not shown). Using *RBI* null U2OS cells produced in this manner we again observed increased quantities of γ H2AX foci in pRB deficient cells compared to their controls (Figure 2.1F). In addition, clones of control and *RBI* knockout NCI-H460 (H460; *CDKN2A* deleted) and NCI-H1792 (H1792; *CDK4* amplified) lung cancer cells also displayed differences in spontaneous DNA damage (Figure 2.1G&H). This analysis demonstrated an elevation in the number of γ H2AX foci in each of the *RBI* deleted lines.

To further investigate the source of DNA damage in *RBI* mutant cells, we assessed their sensitivity to a number of chemical agents to determine if specific stresses could amplify defects that cause increased DNA damage. We tested aphidicolin, a DNA polymerase inhibitor that causes replication stress and etoposide, a topoisomerase inhibitor that creates DNA double stranded breaks. Hydrogen peroxide (H_2O_2) was used to induce oxidative damage, and cisplatin was used to create interstrand cross links, among other damaging effects. Representative U2OS clones of each genotype were treated for 72 hr with a range of chemical concentrations, after which alamarBlue was used to quantitate the cytotoxicity of each agent. These assays revealed that both heterozygous and homozygous *RBI* mutations sensitize cells to hydrogen peroxide and cisplatin, but not aphidicolin or etoposide (Figure 2.2A-D). Because platinum therapeutics preferentially benefit patients with pRB deficient cancers (Cecchini et al., 2015; Garsed et al., 2018), we also tested control and *RBI*^{-/-} H460 and H1792 cells for their sensitivity to cisplatin. This revealed increased sensitivity to cisplatin in H460 and H1792 cells upon *RBI* deletion (Figure 2.2E&F).

Overall, these drug sensitivities suggest that oxidative damage may underlie some aspects of the DNA damage phenotype in *RBI* mutant cells. We compared reactive oxygen species (ROS) levels in wild type and *RBI* mutant U2OS cells with and without H_2O_2 using a ROS indicator, 5(6)-carboxy-2',7'-dichlorodihydrofluorescein diacetate (CA-DCF-DA). For both untreated and H_2O_2 treated cells, there was more fluorescence

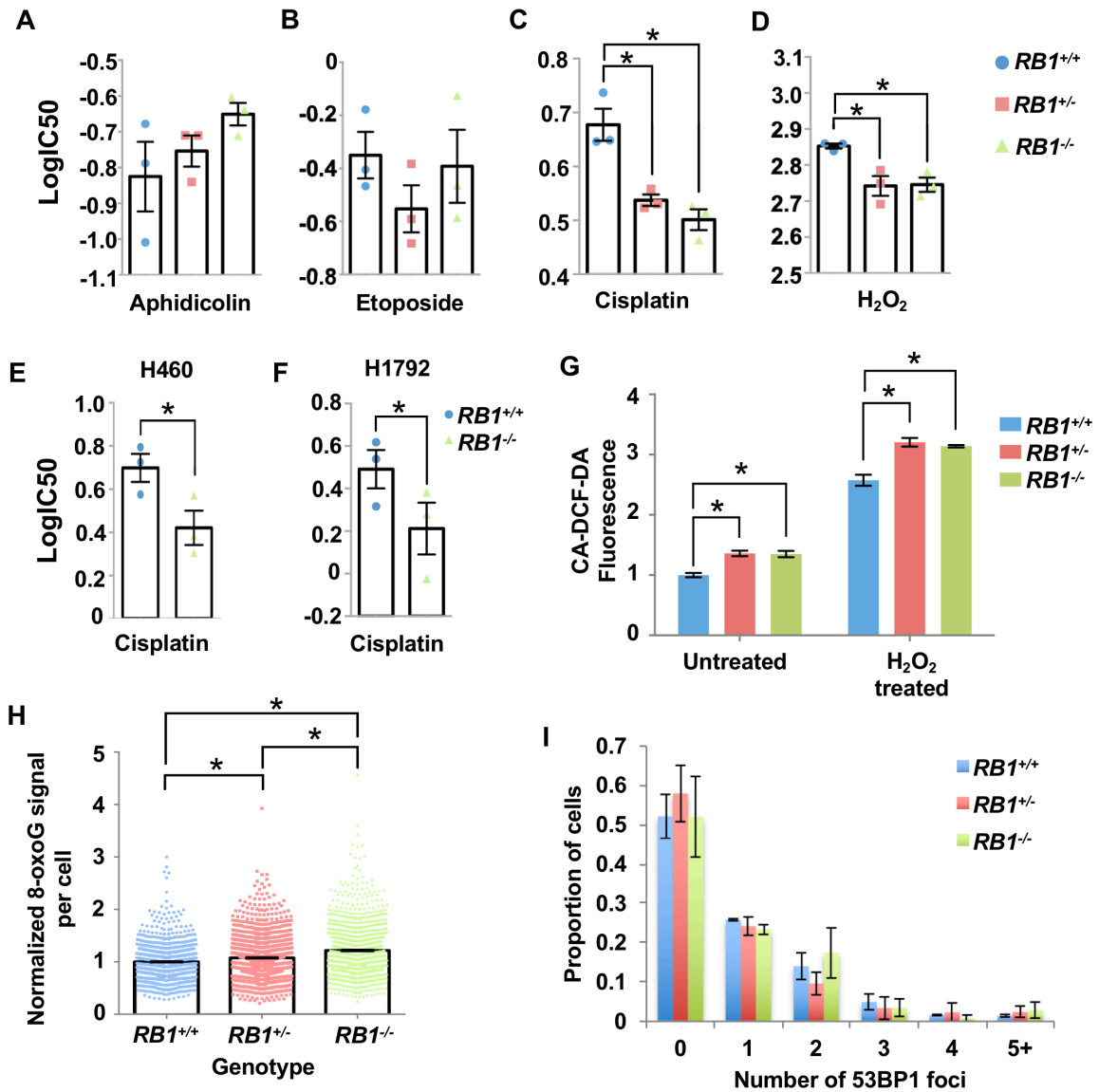


Figure 2.2: Cancer cells with *RB1* mutations have elevated reactive oxygen species and sensitivity to cisplatin.

Figure 2.2: Cancer cells with *RBI* mutations have elevated reactive oxygen species and sensitivity to cisplatin.

(A) Aphidicolin, (B) etoposide, (C) cisplatin and (D) hydrogen peroxide were added to cultures of the indicated genotypes of U2OS cells. Viability was assessed after 72 hr using alamarBlue and dose response curves were used to calculate half maximal inhibitory concentration (IC₅₀) values for each genotype. Both *RBI* mutant genotypes have significantly lower IC₅₀ values in response to cisplatin and hydrogen peroxide than control U2OS cells (as determined by one-way ANOVA). (E and F) H460 and H1792 cancer cells with *RBI* mutations were treated with cisplatin as in C. Differences in IC₅₀ values were determined using a paired *t*-test. (G) To detect reactive oxygen species (ROS), CA-DCF-DA was added to culture media at the end of 72 hr of mock or hydrogen peroxide treatment. Normalized fluorescence was averaged for four clones of *RBI* wild type and knockout genotypes, and three clones for the heterozygous genotype. Mean values were compared by two-way ANOVA. (H) Cells were fixed and stained for 8-oxoG and DAPI and visualized by fluorescence microscopy. The average 8-oxoG signal per nucleus was determined using ImageJ with DAPI staining defining nuclear area. Three clones per genotype were used and data were normalized to the mean signal from *RBI* wild type in duplicate experiments. Statistical significance in staining intensity was determined by Kruskal-Wallis one-way analysis of variance and Dunn's multiple comparisons test. (I) 53BP1 foci were quantitated for each *RBI* genotype using the Focinator as with γ H2AX. No significant differences were observed as determined by the Kolmogorov-Smirnov test. All error bars are ± 1 SEM. **P* < 0.05.

of the ROS indicator in *RB1* mutant cells, and *RB1*^{-/-} and *RB1*^{+/-} were equivalent (Figure 2.2G). We also fixed and stained cells for 8-oxoguanine (8-oxoG), one of the most abundant lesions resulting from oxidative modification of DNA (Furtado et al., 2012), and quantified the staining in DAPI-defined nuclear area using ImageJ (Schneider et al., 2012). Control U2OS values were used to normalize the 8-oxoG signal from *RB1* mutants. Again, both *RB1*^{+/-} and *RB1*^{-/-} U2OS cells had more 8-oxoG staining than the *RB1*^{+/+} cells (Figure 2.2H).

These experiments indicate that loss of *RB1* in cells with pre-existing pRB-pathway defects increases basal levels of DNA damage. Reactive oxygen species appear to be one source of this damage. A chemical agent that directly induces breaks (etoposide) did not selectively affect *RB1* mutant cells, however, sensitivity to a DNA cross linking agent (cisplatin) suggests a potential inability to repair DNA damage by homologous recombination. Both observations are consistent with a lack of 53BP1 foci, a marker of non-homologous end joining (NHEJ), in *RB1* mutant U2OS cells compared to controls (Figure 2.2I). This suggests the nature of DNA damage marked by γ H2AX in these *RB1* mutant cells is not necessarily double stranded DNA breaks, and that NHEJ is not the dominant pathway to repair damage in these cells. Overall, these experiments indicate that *RB1* loss contributes to an unstable genome, regardless of the proliferative control status of the cell.

2.4.2 *RB1* mutant cells have randomly distributed DNA damage

To further understand spontaneous DNA damage in *RB1* mutant U2OS cells, we sought to determine if damage occurred at specific locations within the genome. We performed ChIP-sequencing to identify DNA sequences associated with γ H2AX, as well as histone H4 as a control. Because spontaneous damage in untreated cell cultures is relatively rare, we pooled chromatin from 20 separate γ H2AX ChIP experiments per genotype to create each sequencing library (Figure 2.3A-E). We determined peak locations and number using Model-based Analysis for ChIP-Seq (MACS) (Zhang et al., 2008), and the quantity of γ H2AX and H4 peaks were similar between genotypes (Figure 2.3A). Looking at a large region of chromosome 4 as a representative view of the

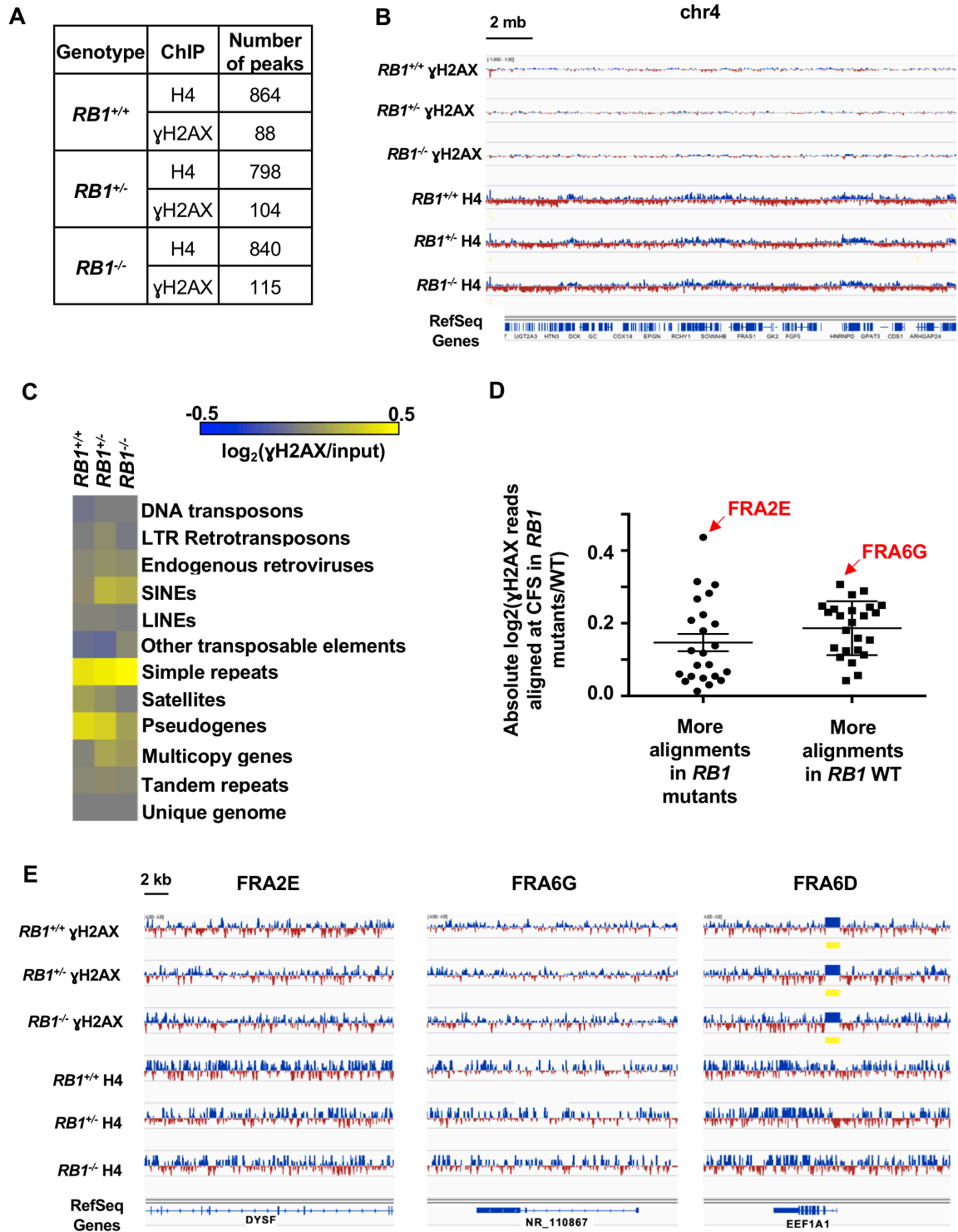


Figure 2.3: γ H2AX is randomly distributed in the genomes of *RB1* mutant cells.

Figure 2.3: γ H2AX is randomly distributed in the genomes of *RB1* mutant cells.

(A) Total number of MACS peaks found for H4 control and γ H2AX ChIP-Seq reads for the indicated genotypes. (B) A 20 Mb region of chromosome 4 is shown with ChIP-Seq read alignments for γ H2AX and H4. Tracks were normalized by subtracting input reads. Blue indicates more reads in the ChIP versus input, red indicates fewer reads. (C) The number of ChIP-Seq reads mapping to repetitive sequences, as well as unique genome regions, was determined. The heatmap shows the \log_2 ratios of the abundance of γ H2AX precipitable reads per million mapped reads versus input for each of the respective genotypes at each element analyzed. (D) Aligned γ H2AX ChIP-Seq read proportions within common fragile sites (CFS) were first normalized to their respective inputs, and then *RB1*^{+/-} and *RB1*^{-/-} were normalized to wild type and \log_2 transformed. A two-tailed one-sample *t*-test was performed to determine if the normalized mean read count proportions of the *RB1* mutants at the various CFS is equal to the normalized read count proportion of the corresponding CFS in the wild type. CFS where the false discovery rate was less than 0.1 were grouped according to whether there were significantly more alignments in the *RB1* mutants, or significantly more alignments in the *RB1*^{+/+} sample. There is no significant difference between these two categories (determined by unpaired *t*-test). FRA2E had the most reads in *RB1*^{+/-} and *RB1*^{-/-} compared to control while FRA6G had the most reads in the wild type compared to the mutants. Error bars are ± 1 SEM. (E) ChIP-Seq tracks for γ H2AX and H4 at representative CFS are displayed. FRA2E and FRA6G from D are shown, while FRA6D had no change in the proportion of γ H2AX reads that aligned between the genotypes. Regions of significant enrichment (MACS peaks) are denoted by yellow bars.

genome, we did not observe consequential differences between genotypes for γ H2AX or H4 peaks (Figure 2.3B). Since peak-finding at a genome scale failed to indicate obvious locations of DNA damage enrichment, we investigated individual genome sequence categories in search of elevated γ H2AX deposition. The proportion of aligned γ H2AX ChIP-Seq reads per million mapped reads versus input reads for each of the genotypes were compared within each repetitive element category and \log_2 transformed for display as a heat map (Figure 2.3C). Some categories, such as short interspersed nuclear elements (SINEs) and multicopy genes appear to have slight enrichment of γ H2AX localization in $RBI^{+/-}$ and $RBI^{-/-}$ compared to $RBI^{+/+}$ based on color, but a two-tailed one-sample t -test with FDR multi-test correction did not score these as significant. Therefore, even within repetitive sequences in the genome, there does not seem to be an enrichment of γ H2AX within RBI mutant cells compared to control.

Previous studies suggest that γ H2AX levels can be elevated at common fragile sites (CFS) of cancer cells under standard cell culture conditions because of replication stress (Harrigan et al., 2011). To investigate these locations, we quantified the number of γ H2AX ChIP-Seq reads and scaled them to the proportions of total aligned reads and normalized them to input levels. $RBI^{+/-}$ and $RBI^{-/-}$ were compared to wild type using a two-tailed one-sample t -test. This analysis revealed 23 CFS that had more γ H2AX alignments than the wild type and 24 CFS that had significantly less in RBI mutant cells compared to controls (Figure 2.3D). Figure 2.3E shows examples of CFS locations with the greatest increase in γ H2AX in RBI mutants (FRA2E), the greatest reduction in γ H2AX in RBI mutants (FRA6G), and unchanged γ H2AX levels (FRA6D). These examples appear highly similar between genotypes. Overall, it is possible that the distribution of γ H2AX within each CFS may be shifting slightly between the mutants and the wild type. However, there does not appear to be more of a bias in general for γ H2AX elevation at CFS in RBI mutant cells.

Collectively, our analysis of γ H2AX distribution across the genome suggests there is no particular chromosome location or sequence category that is preferentially enriched for this mark of DNA damage. These data suggest that the increase in γ H2AX foci

observed in *RBI*^{+/-} and *RBI*^{-/-} cells is likely due to an overall increase in DNA damage, and not newly arising locations, or “hotspots” of damage. This is in contrast to primary cells with a normal pRB-pathway that experience pRB loss and preferentially damage centromeric repeats (Coschi et al., 2014). The elevated sensitivity to peroxide and cisplatin, and increased ROS and 8-oxoG are consistent with DNA damage being randomly located in *RBI* mutant U2OS cells.

2.4.3 Homologous recombination repair defects in *RB1* deficient cancer cells

Another potential source of intrinsic DNA damage is defective repair. For this reason, we investigated the efficiency of HR and NHEJ repair using fluorescent reporters (Bennardo et al., 2008; Pierce et al., 1999). In this assay a promoterless, but functional, GFP is used to repair an adjacent break induced in a mutant, expressed, form of the GFP gene (Figure 2.4A). We generated clonal U2OS lines bearing this reporter and created a population of cells deleted for *RBI* with lentiviral delivery of Cas9 and an *RBI* specific sgRNA (Figure 2.4B). Introduction of the restriction enzyme I-SceI into these cells induced breaks, and *RBI* deficient cells were defective for their repair (Figure 2.4C). Similarly, we generated U2OS clones that stably maintain an NHEJ reporter for repair of induced breaks that links a constitutive promoter with a GFP gene. Loss of pRB expression was again confirmed by western blotting (Figure 2.4E), and induction of breaks was used to test repair in an *RBI* deficient background (Figure 2.4F). This failed to reveal a defect in repair, suggesting that *RBI* loss in U2OS cells specifically reduces HR repair.

To investigate if this reduction in HR repair is related to changes in expression of HR factors due to *RBI* loss, we examined DNA damage repair genes that are known to be regulated by E2Fs (Ren et al., 2002; Xu et al., 2007). Of these candidates, only one gene, *RPA2*, was seen to have significantly lower transcript levels in *RBI*^{-/-} U2OS cells (Figure 2.4G). However, when protein levels of a subset of these factors, including RPA2, were investigated, no changes in expression between genotypes were evident (Figure 2.4H). This suggests that the expression levels of key HR factors are not changed in these cells as a result of *RBI* loss.

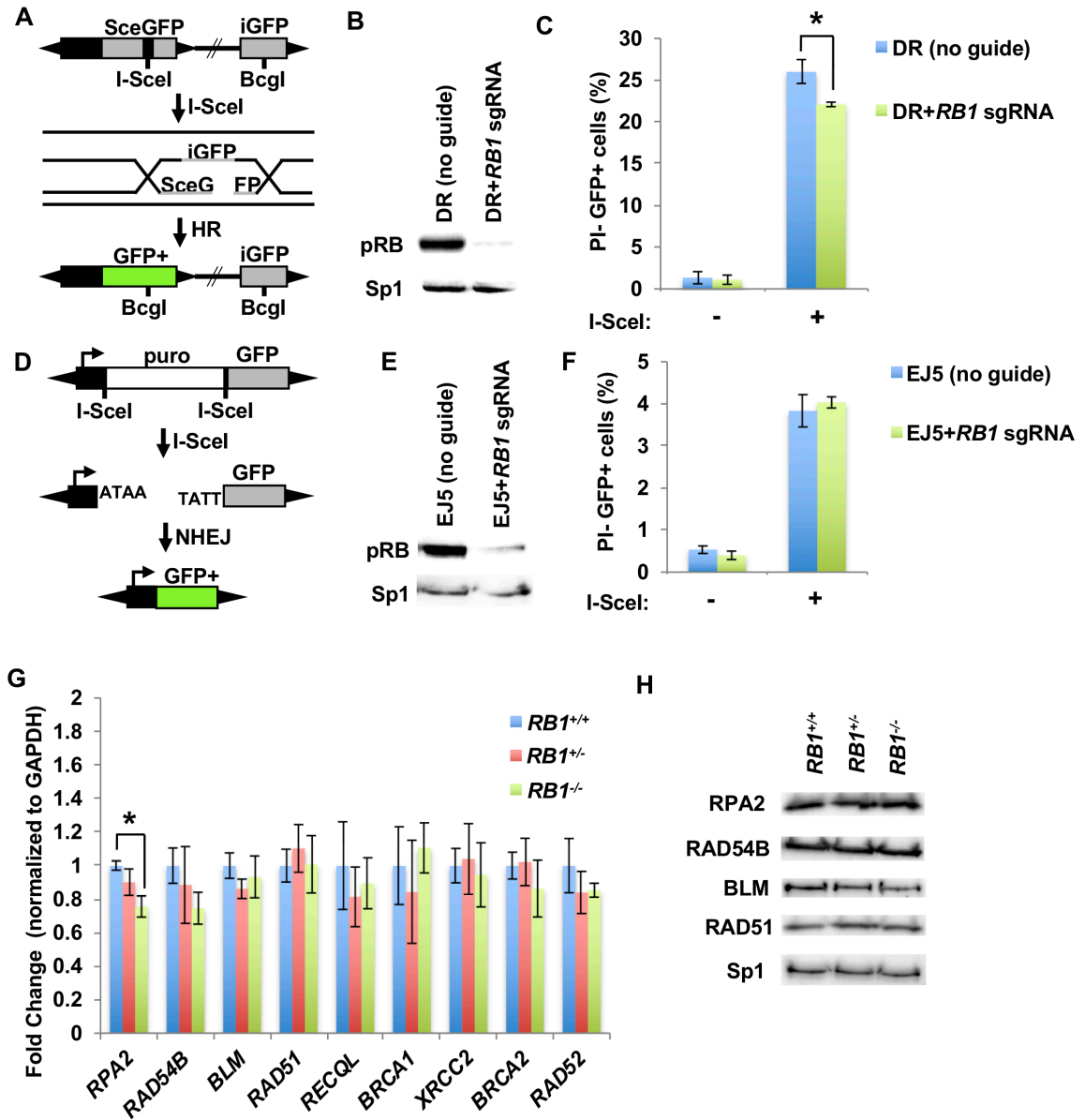


Figure 2.4: Defective homology directed repair of DNA breaks in *RB1* mutant cells.

Figure 2.4: Defective homology directed repair of DNA breaks in *RB1* mutant cells.

(A) Schematic of the DR homology directed repair construct used. Cleavage of an I-SceI site integrated into an expressed, but mutant, GFP gene (SceGFP), can be repaired from a downstream internal GFP fragment (iGFP). **(B)** U2OS cells with clonal integration of the HR reporter construct were ablated for *RB1* expression with lentiviral delivery of Cas9 and an *RB1* specific sgRNA. Relative expression of pRB in the population of cells was determined by western blotting and Sp1 serves as a loading control. **(C)** HR repair efficiency of *RB1* mutant U2OS cells was determined by transfecting an expression vector for I-SceI endonuclease (I-SceI +), or relevant negative control expression vector, and quantitating PI- and GFP+ cells by flow cytometry (n=5). All error bars are ± 1 SEM. $*P < 0.05$. **(D)** Schematic of the EJ5 NHEJ reporter system. DNA breaks at tandem I-SceI sites release the puromycin resistance gene, allowing NHEJ repair to join a promoter to GFP expressing sequence. **(E)** After generation of a stable U2OS clone containing the NHEJ reporter construct, *RB1* was deleted as above and confirmed by western blotting. **(F)** NHEJ repair efficiency was determined by transfecting an I-SceI endonuclease expression vector (I-SceI +), or negative control, and PI- and GFP+ cells were quantitated by flow cytometry (n=3). **(G)** RT-qPCR was performed to assess the transcript levels of various HR factors in *RB1* wild type (4 different clones), heterozygous (3 different clones) and knockout cells (4 different clones). Statistical differences in means were determined by one-way ANOVA. All error bars are ± 1 SEM. $*P < 0.05$. **(H)** Western blots showing expression of various HR factors in wild type, heterozygous, and homozygous *RB1* mutant clones are shown. Sp1 loading control is shown on the bottom.

To assess how *RBI* mutant U2OS cells respond to DNA breaks, each genotype was exposed to 2 Gy of gamma radiation (γ IR). Cells were fixed and stained for γ H2AX and DAPI to measure the amount of DNA damage. One hour after γ IR, despite pronounced γ H2AX foci in all genotypes, the amount of DNA damage was significantly greater in *RBI*^{-/-} clones compared to heterozygous or wild type cells (Figure 2.5A). After 24 hr, most of the DNA damage was repaired, and the *RBI*^{-/-} cells retained more γ H2AX foci than the other two genotypes (Figure 2.5B). This same increase in DNA damage in the *RBI* null background was seen in H460 cells (Figure 2.5C&D) and in H1792 cells (Figure 2.5E&F), both 1 hr and 24 hr after exposure to 1 Gy of γ IR. Overall, this indicates that cells completely lacking pRB are more sensitive to γ IR, likely because they are not able to repair DNA breaks as efficiently by HR repair.

Next, we investigated the fidelity of mitosis to determine if the elevated levels of DNA damage and impaired HR repair impacted chromosome segregation and aneuploidy (Gelot et al., 2015). Flow cytometry was performed on *RBI* deficient cells that were labeled and stained with BrdU and propidium iodide (Cecchini et al., 2012). This analysis failed to show statistically different changes in cell cycle phases between the different genotypes (Figure 2.6A). However, when DNA content greater than 4N was analyzed, there was a significant difference between wild type and *RBI* knockout cells, with *RBI*^{+/-} cells showing an intermediate value (Figure 2.6B). This suggests that mitotic errors in these cells may lead to aneuploidy.

To further investigate mitotic defects and their relationship with DNA damage and replication stress, cells were stained with DAPI and antibodies to BLM to visualize chromosome bridges, and mitotic figures were imaged using confocal microscopy (Figure 2.6C). We observed abundant chromosome bridges in *RBI*^{-/-} and *RBI*^{+/-} mutant cells (Figure 2.6D). In the *RBI* mutants there were some ultra-fine bridges (UFBs), which are “thread-like” DNA structures that stain only with BLM (Chan and Hickson, 2011). In *RBI*^{+/-} anaphase cells, 5% (2/43) were seen to have UFBs and in *RBI*^{-/-} cells, this was increased to 8% (3/39), whereas in wild type cells, UFBs were not observed (0/30). However, the majority of BLM bridges stained with DAPI, indicating that anaphase

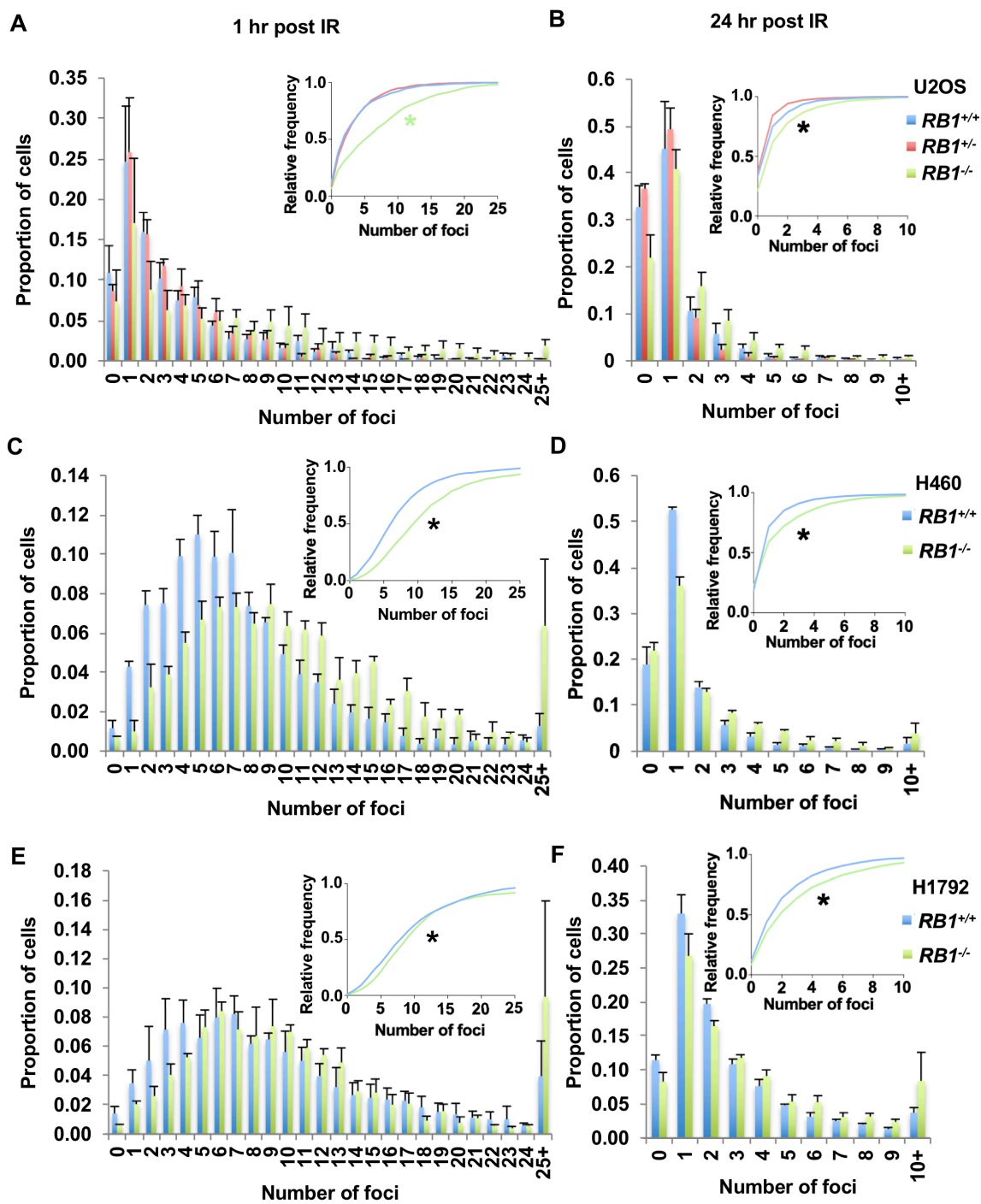


Figure 2.5: Defective repair of γ IR induced DNA damage in *RB1* knockout cells.

Figure 2.5: Defective repair of γ IR induced DNA damage in *RBI* knockout cells.

(A) U2OS cells were treated with 2 Gy γ IR and fixed 1 hr or (B) 24 hr after treatment and stained for γ H2AX. Three clones per genotype were used and γ H2AX foci were quantified. The proportion of cells with discrete numbers of foci are shown as histograms, while the cumulative frequency of foci is shown as an inset. Differences in foci distribution were determined using the Kolmogorov-Smirnov test. A green asterisk indicates *RBI*^{-/-} is statistically different than the other genotypes, while a black asterisk indicates all genotypes are statistically significantly different from each other. (C) H460 cells were treated with 1 Gy γ IR and fixed 1 hr or (D) 24 hr after treatment and stained for γ H2AX. (E) H1792 cells were treated with 1 Gy γ IR and fixed 1 hr or (F) 24 hr after treatment and stained for γ H2AX. All error bars are +1 SEM. **P* < 0.05.

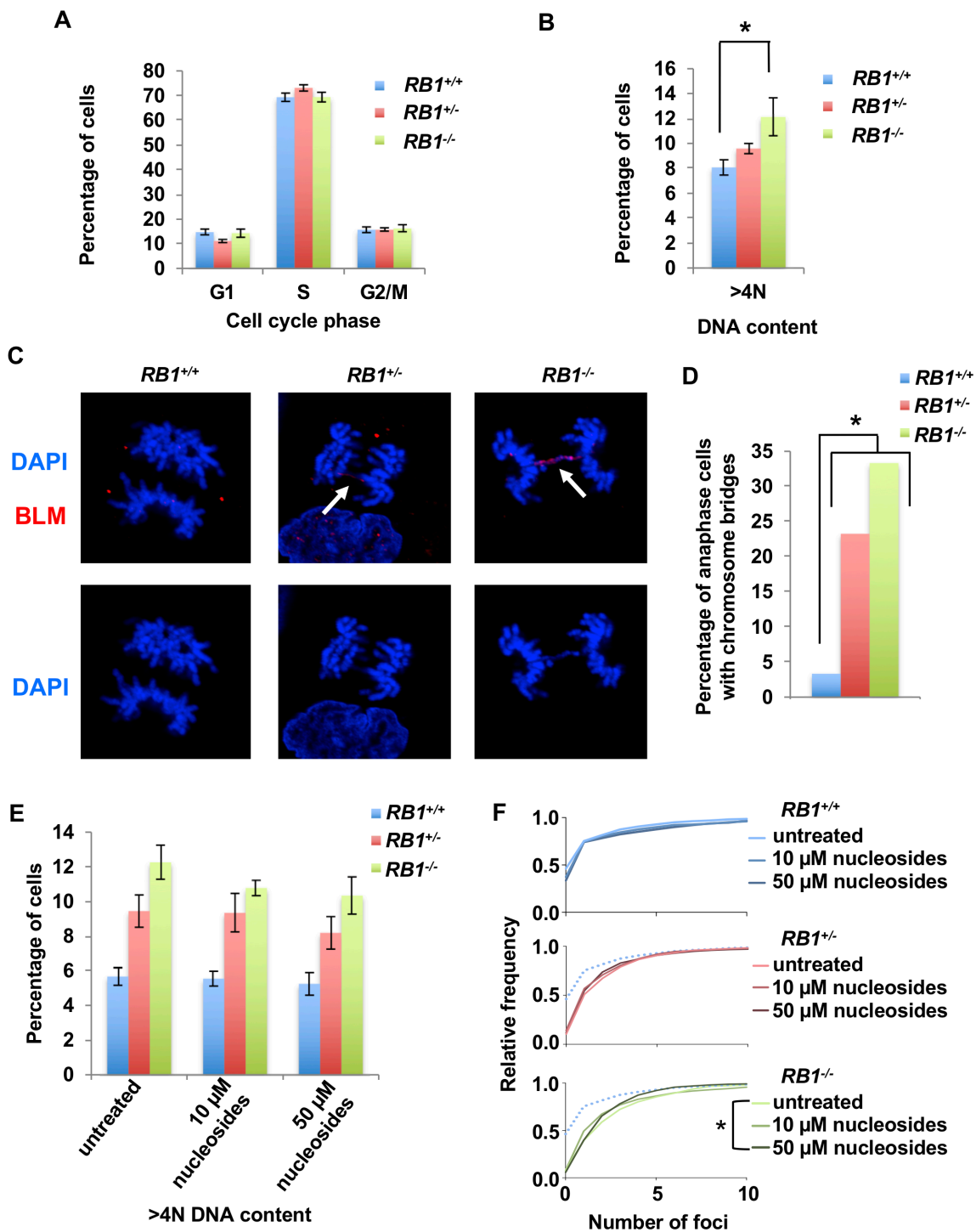


Figure 2.6: Increased mitotic errors in *RBI* mutant cells.

(A) BrdU and propidium iodide staining followed by flow cytometry were used to determine cell cycle phase distribution of asynchronous cultures of U2OS cells. Four clones for the wild type and knockout genotypes and three clones for the heterozygous genotype were analyzed. (B) Flow cytometry analysis shows the proportion of cells with greater than 4N DNA content. Means were compared using a one-way ANOVA. (C) Cells in anaphase were imaged by fluorescence microscopy using DAPI (blue) and BLM (red) in cells from each *RBI* genotype. Arrows indicate anaphase bridges that are stained by both DAPI and BLM. (D) The number of anaphase cells with DAPI stained chromosome bridges were quantitated. The proportion of cells with bridges is significantly higher in the knockout and heterozygous clone compared to the wild type clone as determined by the χ^2 -test. (E) Representative clones from each *RBI* genotype were either left untreated or treated with 10 μ M or 50 μ M nucleosides for 48 hr. Flow cytometry analysis of propidium iodide stained cells shows the proportion of cells with greater than 4N DNA content. Mean differences were compared by one-way ANOVA (n=3). (F) Nucleoside treated cells were fixed and stained for γ H2AX after 48 hr of culture. DNA damage is summarized in frequency plots. The blue dotted line in the *RBI*^{+/+} and *RBI*^{-/-} cumulative frequency plots represents the wild type untreated cells. Statistical significance between genotypes was determined by the Kolmogorov-Smirnov test. All error bars are ± 1 SEM. * $P < 0.05$.

bridges were most common. Anaphase bridges are known to occur in HR-defective cells, while UFBs can be induced by replication stress (Gelot et al., 2015).

Lastly, to investigate the link between UFBs in the *RB1* mutant cells and a possible increase in replication stress, we performed a nucleoside supplementation assay and investigated changes in DNA content and DNA damage as surrogate markers for nucleoside suppression of replication stress. Flow cytometry of cells stained with propidium iodide was performed 48 hr after media replacement, either with or without the addition of nucleosides. This analysis failed to show statistically different changes in greater than 4N DNA content within genotypes when comparing the different nucleoside treatment groups, although there was a trend towards decreased aneuploidy after nucleoside supplementation in all genotypes (Figure 2.6E). However, when we looked at DNA damage of cells grown in parallel, *RB1*^{-/-} cells treated with 50 μM nucleosides did have a significant reduction in γH2AX foci compared to untreated knockout cells (Figure 2.6F). Nevertheless, this reduction in DNA damage after nucleoside supplementation is modest and did not restore the levels of γH2AX foci to those seen in the *RB1*^{+/+} controls.

Taken together with experiments earlier in this report, defects in mitosis are best characterized as anaphase bridges that cause aneuploidy. These types of errors are consistent with HR repair deficiency and sensitivity to a DNA cross linking agent such as cisplatin. It appears that γH2AX foci can be suppressed slightly in *RB1*^{-/-} cells with nucleosides and *RB1*^{-/-} cells exhibit some UFBs, implying that they experience modest DNA replication stress. However, the lack of sensitivity to aphidicolin and the lack of increased γH2AX deposition at repetitive regions of the genome argues that replication stress in these cells is quite modest. Overall, our analysis of mitotic errors is consistent with a defect in HR repair being the main source of chromosome bridges in anaphase of these *RB1* mutant cells.

2.4.4 Increased lung metastases in *RB1* mutant xenografts

To further characterize the effects of induced *RB1* mutations in cells that already possess pRB-pathway defects, we performed xenograft experiments to determine if new cancer relevant properties arise upon loss of pRB. We injected cells subcutaneously into

immune compromised mice and allowed tumors to form for an eight-week period before analyzing growth by mass and histology (Figure 2.7A). This analysis revealed highly cellular tumors with abundant mitotic activity and foci of necrosis. Cells appeared epithelioid with no definite features of osteoid differentiation and they had small areas suggesting glandular differentiation. This phenotype was consistent among all genotypes (Figure 2.7B). Tumor masses were determined at end point and were not statistically different between genotypes, with *RBI*^{-/-} even trending towards a smaller size (Figure 2.7C). Mice were also tail vein injected and cell dissemination and proliferation were allowed to proceed for eight weeks at which time lungs were harvested, fixed, and sectioned. Hematoxylin and eosin-stained sections were digitally analyzed to quantitate cellular infiltration of U2OS cells (Figure 2.7D). This revealed a striking increase in *RBI*^{-/-} U2OS cells in the lungs of these mice compared to control and *RBI*^{+/-} cells as determined by the percent of lung section area that is occupied by malignant cells (Figure 2.7E). There were significantly more individual nodules of *RBI*^{-/-} cells per lung than the other genotypes (Figure 2.7F), further suggesting that *RBI* loss increased the efficiency of dissemination or establishment in the lung. Lastly, the area occupied by tumor cells per nodule is lower in both *RBI* mutant genotypes indicating that control cells form rarer, larger clusters of cells, whereas *RBI* mutants tend to seed more efficiently and perhaps proliferate more slowly (Figure 2.7G). We note that one mouse injected with *RBI*^{+/-} cells showed extensive dissemination and tumor burden that was highly reminiscent of mice injected with *RBI*^{-/-} cells (Figure 2.8A&B). To investigate this further we extracted DNA from paraffin embedded tumor material from this mouse. PCR analysis was used to genotype *RBI* exon 22 in these cells (Figure 2.8C). It confirmed that these cells maintained their wild type *RBI* allele, suggesting they had not undergone loss of heterozygosity as a means to acquire an *RBI*^{-/-} phenotype.

These xenograft experiments indicate that *RBI* mutations in U2OS cells have little effect on growth rate of primary tumors. Interestingly, *RBI*^{-/-} cells are much more efficient in colonizing recipient mouse lungs, suggesting that in addition to the DNA damage and genome instability phenotypes described earlier, *RBI* loss in pRB-pathway deficient cells imparts characteristics that enable cancer progression.

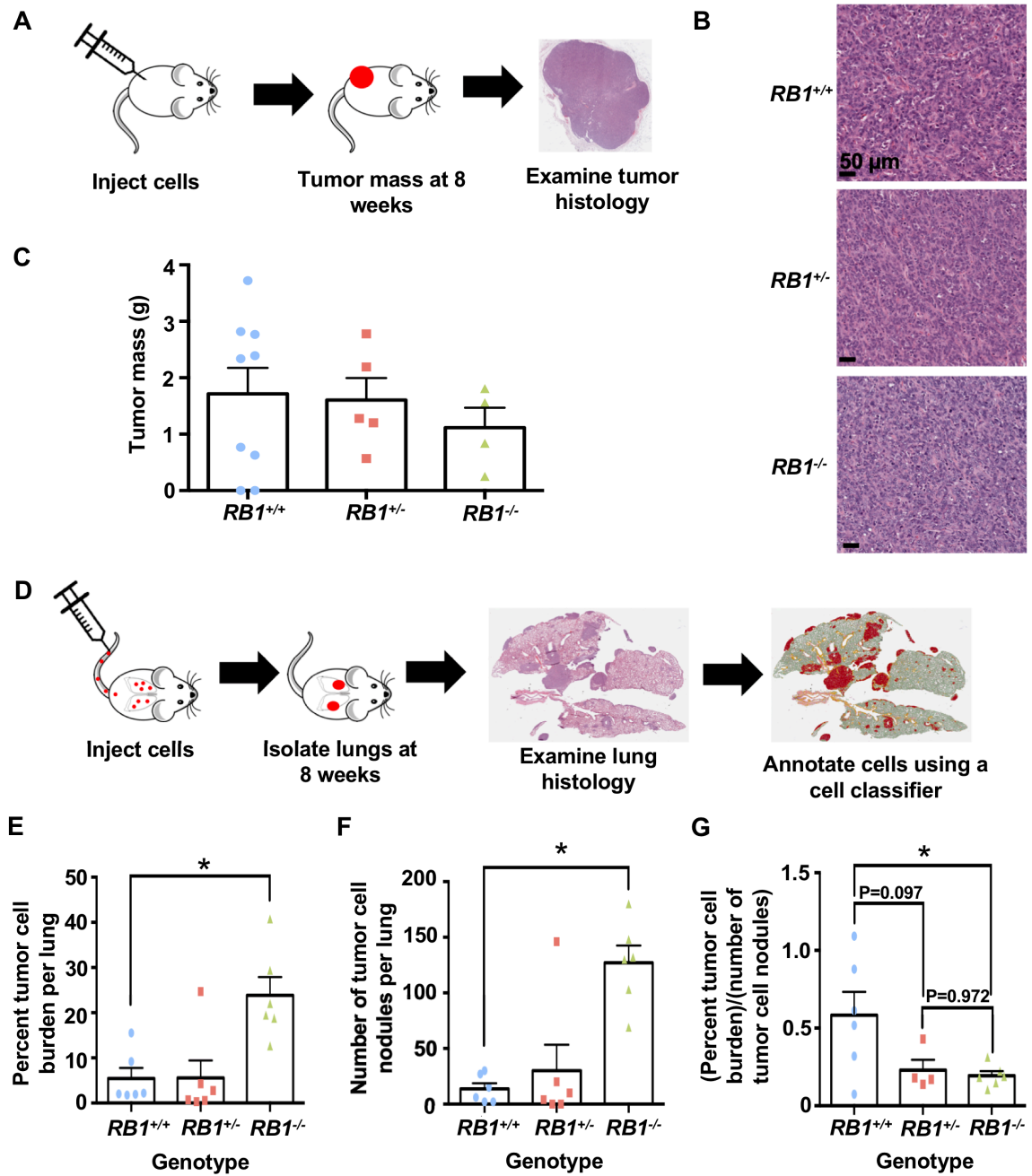


Figure 2.7: Increased lung metastases with *RB1* mutant cells in xenograft experiments.

Figure 2.7: Increased lung metastases with *RB1* mutant cells in xenograft experiments.

(A) Illustration of analysis of subcutaneous injections of *RB1* mutant U2OS cells. Tumors were allowed to form for eight weeks before analyzing tumor mass and histology. (B) Representative H&E stained tissue sections from each genotype of tumor. (C) Tumor masses from subcutaneous injected cells are shown. The means are not statistically different. (D) Schematic of tail vein injections to study dissemination to the lungs. Mice were injected and cell dissemination and proliferation were allowed to proceed for eight weeks. Lungs were then isolated, sectioned and stained with H&E, and analyzed using QuPath. (E) The percentage of lung area occupied by cancer cells was calculated from tissue sections and averaged between mice. (F) Tumor cell nodules were counted using the assistance of QuPath and averaged among recipient mice. (G) Percent tumor cell burden was divided by the number of tumor nodules to determine the average tumor cell area per nodule. Statistical significance between genotypes was determined with a *t*-test. All error bars are +1 SEM. * $P < 0.05$.

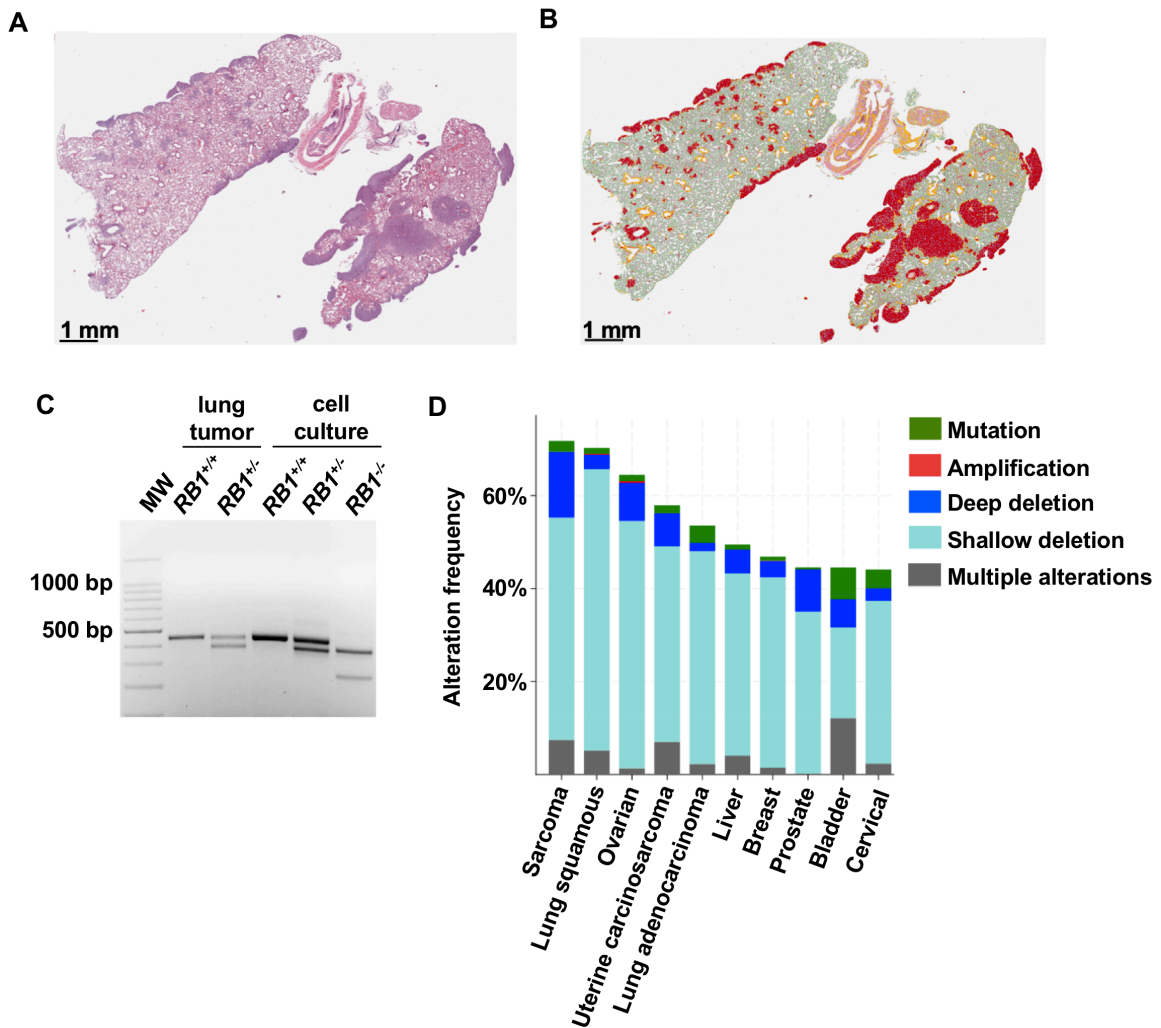


Figure 2.8: Single copy loss of *RB1* can create cancer enabling phenotypes.

(A) H&E staining of lung tissue with the highest number of *RB1*^{+/-} seeding events. (B) QuPath coloring to denote tumor tissue within these lungs (red). (C) DNA was extracted from nodules of tumor cells in paraffin embedded tissue containing these *RB1*^{+/-} cells, or a *RB1*^{+/+} control. PCR was performed to amplify exon 22 from recovered DNA and controls isolated from cell culture to verify the final genotype of cells in this sample. (D) The 10 most prevalent cancers were analyzed for *RB1* gene alterations using TCGA and Pan-Cancer Atlas data using cBioPortal. A deep deletion is consistent with biallelic loss of *RB1*, whereas a shallow deletion is suggestive of heterozygous *RB1* deletion.

2.5 Discussion

To test if *RB1* loss impacts cancer cells that already possess mutations disrupting the pRB-pathway, we used CRISPR/Cas9 to create non-functional *RB1* alleles. Using γ H2AX staining of cells, we found that even within cells deficient for one copy of *RB1*, there is an increase in spontaneous DNA damage. Experiments that probe drug sensitivity of *RB1* mutant cells, along with ChIP-Seq analysis of γ H2AX foci and analysis of DNA damage repair pathways, suggest a complex picture of cellular defects. We did not detect specific hotspots of DNA damage, but sensitivity to peroxide and cisplatin suggest that oxidative damage may cause sporadic DNA damage as reactive oxygen species are elevated in *RB1* deficient cells. Mutant *RB1* cells also have more abnormal mitoses characterized by anaphase bridges, reporter assays detect defects in HR repair, and nucleoside supplementation in culture media suppresses γ H2AX foci, suggesting additional endogenous sources of damage. Overall, this study reveals that *RB1* mutations lead to increased DNA damage that may enhance cancer progression.

The DNA damage phenotype caused by either single copy or homozygous mutation to *RB1* is unlikely to be attributable to a single root cause, and we expect that *RB1* deficiency may affect DNA damage in different cancer cells in varied ways. We report that *RB1* deletion compromises HR repair but not NHEJ and this is consistent with one previous report (Velez-Cruz et al., 2016), but contradicts another (Cook et al., 2015). Given that we observe more than 70% of U2OS cells incorporating BrdU in a brief pulse, they are likely biased towards HR repair pathways because of their cell cycle position. This is consistent with γ H2AX foci not being accompanied by 53BP1, and the activity levels of NHEJ reporters being almost an order of magnitude less than HR values, suggesting that U2OS cells are primed to use the HR repair pathway and thus phenotypes in these cells reflect this reality. We expect that DNA damage and defective repair described in this report are relevant to cancer progression phenotypes because graded differences between *RB1* wild type, heterozygous, and homozygous genotypes are reflected in 8-oxoguanine abundance, aneuploidy, and anaphase bridges. This stepwise trend in severity of phenotype in the molecular alterations resulting from *RB1* loss is similarly evident in the behaviour of *RB1*^{+/+}, *RB1*^{+/-}, and *RB1*^{-/-} mutant tail vein xenograft

experiments, suggesting that this increase in metastatic burden is related to a stepwise increase in DNA damage experienced by these different genotypes.

Single copy loss of *RBI* may contribute to cancer in a number of ways. Primary *RBI*^{+/-} cells from a number of sources are prone to mitotic errors (Coschi et al., 2014; Gonzalez-Vasconcellos et al., 2013; Zheng et al., 2002), and precursor lesions to retinoblastoma are characterized by aneuploidy (Dimaras et al., 2008), suggesting this is an early step in this disease. Therefore, partially defective *RBI* can contribute to the early stages of cancer through a distinct set of effects. However, a number of studies have highlighted that *RBI* loss is statistically enriched in advanced stages of cancer progression (Beltran et al., 2016; McNair et al., 2017; Robinson et al., 2017). In addition, analysis of the landscape of cancer alterations in TCGA consortia reveals so called “shallow deletions” of *RBI* as relatively commonplace (Figure 2.8D), and these events are unlikely to be explained by random, unselected events (Beroukhim et al., 2010). From this perspective, our study offers critical proof of concept that late stage loss of one copy of *RBI* can create cancer enabling phenotypes in the host cell, even if it already possesses pRB-pathway mutations. Complete elimination of *RBI* has a stronger effect on DNA damage phenotypes and dissemination to the lungs, as demonstrated in this study, however, highly abundant single copy *RBI* loss may represent a mutational compromise in which advantageous phenotypes are acquired with an economy of mutational changes (Davoli et al., 2013).

A number of studies correlate absence of pRB expression with improved patient survival following treatment that includes platinum-based chemotherapy (Cecchini et al., 2015; Garsed et al., 2018). Loss of pRB correlated with improved survival of lung adenocarcinomas treated by resection and adjuvant cisplatin or carboplatin, and a vinca alkaloid (Cecchini et al., 2015). More recently patients with high-grade serous ovarian carcinoma (HGSOC) that experienced exceptionally good clinical outcomes were studied (Garsed et al., 2018). These patients were treated with platinum-based agents, and loss of pRB was associated with long-term survival. From this perspective, U2OS cells engineered to be deficient for *RBI* demonstrate that pRB loss increases sensitivity to cisplatin. Given that *RBI* mutant cells were not more sensitive to another agent that

induces DNA breaks, etoposide, we interpret *RBI* deficiency to create a unique sensitivity to cisplatin that may relate to defective HR and higher endogenous ROS levels that create a highly specific sensitivity to this class of chemotherapy.

In conclusion, although there are many cancers that have mutations in the pRB-pathway that spare the *RBI* gene itself, further mutations to *RBI* surprisingly create cancer relevant characteristics that may influence disease progression. The frequency of shallow deletions of *RBI* across many cancers suggests that disease progression may select for these characteristics.

2.6 References

Bankhead, P., Loughrey, M.B., Fernandez, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., *et al.* (2017). QuPath: Open source software for digital pathology image analysis. *Sci Rep* 7, 16878.

Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* 6, 11.

Beltran, H., Prandi, D., Mosquera, J.M., Benelli, M., Puca, L., Cyrta, J., Marotz, C., Giannopoulou, E., Chakravarthi, B.V., Varambally, S., *et al.* (2016). Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat Med* 22, 298-305.

Benevolenskaya, E.V., and Frolov, M.V. (2015). Emerging links between E2F control and mitochondrial function. *Cancer Res* 75, 619-623.

Bennardo, N., Cheng, A., Huang, N., and Stark, J.M. (2008). Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS genetics* 4, e1000110.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., *et al.* (2010). The landscape of somatic copy-number alteration across human cancers. *Nature* 463, 899-905.

Burkhardt, D.L., and Sage, J. (2008). Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat Rev Cancer* 8, 671-682.

Cecchini, M.J., Amiri, M., and Dick, F.A. (2012). Analysis of cell cycle position in mammalian cells. *J Vis Exp* 59, e3491.

Cecchini, M.J., and Dick, F.A. (2011). The biochemical basis of CDK phosphorylation-independent regulation of E2F1 by the retinoblastoma protein. *Biochem J* 434, 297-308.

- Cecchini, M.J., Ishak, C.A., Passos, D.T., Warner, A., Palma, D.A., Howlett, C.J., Driman, D.K., and Dick, F.A. (2015). Loss of the retinoblastoma tumor suppressor correlates with improved outcome in patients with lung adenocarcinoma treated with surgery and chemotherapy. *Human pathology* 46, 1922-1934.
- Cecchini, M.J., Thwaites, M., Talluri, S., Macdonald, J.I., Passos, D.T., Chong, J.L., Cantalupo, P., Stafford, P., Saenz-Robles, M.T., Francis, S.M., *et al.* (2014). A retinoblastoma allele that is mutated at its common E2F interaction site inhibits cell proliferation in gene targeted mice. *Molecular and cellular biology* 34, 2029-2045.
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., *et al.* (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* 2, 401-404.
- Chan, K.L., and Hickson, I.D. (2011). New insights into the formation and resolution of ultra-fine anaphase bridges. *Seminars in cell & developmental biology* 22, 906-912.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., *et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823.
- Cook, R., Zoumpoulidou, G., Luczynski, M.T., Rieger, S., Moquet, J., Spanswick, V.J., Hartley, J.A., Rothkamm, K., Huang, P.H., and Mitnacht, S. (2015). Direct involvement of retinoblastoma family proteins in DNA repair by non-homologous end-joining. *Cell Rep* 10, 2006-2018.
- Coschi, C.H., Ishak, C.A., Gallo, D., Marshall, A., Talluri, S., Wang, J., Cecchini, M.J., Martens, A.L., Percy, V., Welch, I., *et al.* (2014). Haploinsufficiency of an RB-E2F1-Condensin II complex leads to aberrant replication and aneuploidy. *Cancer Discov* 4, 840-853.
- Davoli, T., Xu, A.W., Mengwasser, K.E., Sack, L.M., Yoon, J.C., Park, P.J., and Elledge, S.J. (2013). Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* 155, 948-962.
- Day, D.S., Luquette, L.J., Park, P.J., and Kharchenko, P.V. (2010). Estimating enrichment of repetitive elements from high-throughput sequence data. *Genome Biol* 11, R69.
- Dick, F.A., Goodrich, D.W., Sage, J., and Dyson, N.J. (2018). Non-canonical functions of the RB protein in cancer. *Nat Rev Cancer* 18, 442-451.
- Dimaras, H., Khetan, V., Halliday, W., Orlic, M., Prigoda, N.L., Piovesan, B., Marrano, P., Corson, T.W., Eagle, R.C., Jr., Squire, J.A., *et al.* (2008). Loss of RB1 induces non-proliferative retinoma: increasing genomic instability correlates with progression to retinoblastoma. *Hum Mol Genet* 17, 1363-1372.

- Dyson, N.J. (2016). RB1: a prototype tumor suppressor and an enigma. *Genes Dev* 30, 1492-1502.
- Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., Ponting, L., *et al.* (2017). COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* 45, D777-D783.
- Furtado, C., Kunrath-Lima, M., Rajao, M.A., Mendes, I.C., de Moura, M.B., Campos, P.C., Macedo, A.M., Franco, G.R., Pena, S.D., Teixeira, S.M., *et al.* (2012). Functional characterization of 8-oxoguanine DNA glycosylase of *Trypanosoma cruzi*. *PLoS One* 7, e42484.
- Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., *et al.* (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6, pii.
- Garsed, D.W., Alsop, K., Fereday, S., Emmanuel, C., Kennedy, C.J., Etemadmoghadam, D., Gao, B., GebSKI, V., Gars, V., Christie, E.L., *et al.* (2018). Homologous Recombination DNA Repair Pathway Disruption and Retinoblastoma Protein Loss Are Associated with Exceptional Survival in High-Grade Serous Ovarian Cancer. *Clin Cancer Res* 24, 569-580.
- Gelfand, Y., Rodriguez, A., and Benson, G. (2007). TRDB--the Tandem Repeats Database. *Nucleic acids research* 35, D80-87.
- Gelot, C., Magdalou, I., and Lopez, B.S. (2015). Replication stress in Mammalian cells and its consequences for mitosis. *Genes (Basel)* 6, 267-298.
- Gonzalez-Vasconcellos, I., Anastasov, N., Sanli-Bonazzi, B., Klymenko, O., Atkinson, M.J., and Rosemann, M. (2013). Rb1 haploinsufficiency promotes telomere attrition and radiation-induced genomic instability. *Cancer Res* 73, 4247-4255.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646-674.
- Harrigan, J.A., Belotserkovskaya, R., Coates, J., Dimitrova, D.S., Polo, S.E., Bradshaw, C.R., Fraser, P., and Jackson, S.P. (2011). Replication stress induces 53BP1-containing OPT domains in G1 cells. *J Cell Biol* 193, 97-108.
- Horowitz, J.M., Park, S.H., Bogenmann, E., Cheng, J.C., Yandell, D.W., Kaye, F.J., Minna, J.D., Dryja, T.P., and Weinberg, R.A. (1990). Frequent inactivation of the retinoblastoma anti-oncogene is restricted to a subset of human tumor cells. *Proc Natl Acad Sci USA* 87, 2775-2779.
- Jones, R.A., Robinson, T.J., Liu, J.C., Shrestha, M., Voisin, V., Ju, Y., Chung, P.E., Pellecchia, G., Fell, V.L., Bae, S., *et al.* (2016). RB1 deficiency in triple-negative breast cancer induces mitochondrial protein translation. *J Clin Invest* 126, 3739-3757.

- Knudsen, E.S., and Knudsen, K.E. (2008). Tailoring to RB: tumour suppressor status and therapeutic response. *Nat Rev Cancer* 8, 714-724.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10, R25.
- Longworth, M.S., Herr, A., Ji, J.Y., and Dyson, N.J. (2008). RBF1 promotes chromatin condensation through a conserved interaction with the Condensin II protein dCAP-D3. *Genes Dev* 22, 1011-1024.
- Ludovini, V., Gregorc, V., Pistola, L., Mihaylova, Z., Floriani, I., Darwish, S., Stracci, F., Tofanetti, F.R., Ferraldeschi, M., Di Carlo, L., *et al.* (2004). Vascular endothelial growth factor, p53, Rb, Bcl-2 expression and response to chemotherapy in advanced non-small cell lung cancer. *Lung Cancer* 46, 77-85.
- Lukusa, T., and Fryns, J.P. (2008). Human chromosome fragility. *Biochim Biophys Acta* 1779, 3-16.
- Manning, A.L., Yazinski, S.A., Nicolay, B., Bryll, A., Zou, L., and Dyson, N.J. (2014). Suppression of Genome Instability in pRB-Deficient Cells by Enhancement of Chromosome Cohesion. *Molecular cell* 53, 993-1004.
- McNair, C., Xu, K., Mandigo, A.C., Benelli, M., Leiby, B., Rodrigues, D., Lindberg, J., Gronberg, H., Crespo, M., De Laere, B., *et al.* (2017). Differential impact of RB status on E2F1 reprogramming in human cancer. *J Clin Invest*.
- Nicolay, B.N., Danielian, P.S., Kottakis, F., Lapek, J.D., Jr., Sanidas, I., Miles, W.O., Dehnad, M., Tschop, K., Gierut, J.J., Manning, A.L., *et al.* (2015). Proteomic analysis of pRb loss highlights a signature of decreased mitochondrial oxidative phosphorylation. *Genes Dev* 29, 1875-1889.
- Oeck, S., Malewicz, N.M., Hurst, S., Rudner, J., and Jendrossek, V. (2015). The Focinator - a new open-source tool for high-throughput foci evaluation of DNA damage. *Radiat Oncol* 10, 163.
- Pavlidis, P., and Noble, W.S. (2003). Matrix2png: a utility for visualizing matrix data. *Bioinformatics* 19, 295-296.
- Pierce, A.J., Johnson, R.D., Thompson, L.H., and Jasin, M. (1999). XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev* 13, 2633-2638.
- Quinlan, A.R. (2014). BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr Protoc Bioinformatics* 47, 11.12.1-11.12.34.

- Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R.A., and Dynlacht, B.D. (2002). E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev* 16, 245-256.
- Robinson, D.R., Wu, Y.M., Lonigro, R.J., Vats, P., Cobain, E., Everett, J., Cao, X., Rabban, E., Kumar-Sinha, C., Raymond, V., *et al.* (2017). Integrative clinical genomics of metastatic cancer. *Nature* 548, 297-303.
- Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11, 783-784.
- Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9, 671-675.
- Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., *et al.* (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84-87.
- Sharma, A., Yeow, W.S., Ertel, A., Coleman, I., Clegg, N., Thangavel, C., Morrissey, C., Zhang, X., Comstock, C.E., Witkiewicz, A.K., *et al.* (2010). The retinoblastoma tumor suppressor controls androgen signaling and human prostate cancer progression. *J Clin Invest* 120, 4478-4492.
- Sherr, C.J. (1996). Cancer cell cycles. *Science* 274, 1672-1677.
- Sherr, C.J., and McCormick, F. (2002). The RB and p53 pathways in cancer. *Cancer Cell* 2, 103-112.
- Thangavel, C., Boopathi, E., Liu, Y., Haber, A., Ertel, A., Bhardwaj, A., Addya, S., Williams, N., Ciment, S.J., Cotzia, P., *et al.* (2017). RB Loss Promotes Prostate Cancer Metastasis. *Cancer Res* 77, 982-995.
- Tyner, C., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Eisenhart, C., Fischer, C.M., Gibson, D., Gonzalez, J.N., Guruvadoo, L., *et al.* (2017). The UCSC Genome Browser database: 2017 update. *Nucleic acids research* 45, D626-D634.
- Velez-Cruz, R., and Johnson, D.G. (2017). The Retinoblastoma (RB) Tumor Suppressor: Pushing Back against Genome Instability on Multiple Fronts. *Int J Mol Sci* 18.
- Velez-Cruz, R., Manickavinayaham, S., Biswas, A.K., Clary, R.W., Premkumar, T., Cole, F., and Johnson, D.G. (2016). RB localizes to DNA double-strand breaks and promotes DNA end resection and homologous recombination through the recruitment of BRG1. *Genes Dev* 30, 2500-2512.
- Xu, X., Bieda, M., Jin, V.X., Rabinovich, A., Oberley, M.J., Green, R., and Farnham, P.J. (2007). A comprehensive ChIP-chip analysis of E2F1, E2F4, and E2F6 in normal and tumor cells reveals interchangeable roles of E2F family members. *Genome research* 17, 1550-1561.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137.

Zhao, W., Huang, C.C., Otterson, G.A., Leon, M.E., Tang, Y., Shilo, K., and Villalona, M.A. (2012). Altered p16(INK4) and RB1 Expressions Are Associated with Poor Prognosis in Patients with Nonsmall Cell Lung Cancer. *J Oncol* 2012, 957437.

Zheng, L., Flesken-Nikitin, A., Chen, P.L., and Lee, W.H. (2002). Deficiency of Retinoblastoma gene in mouse embryonic stem cells leads to genetic instability. *Cancer Res* 62, 2498-2502.

Chapter 3

3 An pRB-condensin II complex mediates long-range chromosome interactions and regulates expression at divergently paired genes

3.1 Abstract

Interphase chromosomes are organized into topologically associated domains by cohesins. Condensin II is analogous to cohesin, but its requirement in higher order chromatin organization in interphase mammalian nuclei is unknown because of limited loss of function approaches. Here we show that condensin II is recruited to divergently transcribed promoters by a mechanism that is dependent on the retinoblastoma protein (pRB). Long-range chromosome contacts are disrupted by loss of condensin II loading which leads to misexpression at these gene pairs. This study demonstrates that mammalian condensin II functions to organize long-range chromosome contacts and regulate transcription at specific genes. In addition, pRB dependence of condensin II suggests that widespread misregulation of chromosome contacts and transcriptional alterations are a consequence of pRB loss in cancer.

3.2 Introduction

Genomes are more complex than their one-dimensional, linear DNA sequences. Three-dimensional organization of interphase chromosomes imparts regulatory information that contributes to development and disease. In particular, functions are arranged spatially in the nucleus such that transcription, replication, and other processes are compartmentalized in discrete locations (Misteli, 2007). Understanding the organizational principles of interphase chromosomes has been aided by chromosome conformation based analyses (Rowley and Corces, 2018). In addition, investigation into the genetic requirement for specific architectural proteins is now revealing chromatin folding and compartmentalization on a whole genome scale with their underlying molecular mechanisms (Rowley and Corces, 2018). We are also beginning to appreciate that cell-fate decisions are driven by changes in environmental cues, triggering signal

transduction into the nucleus, which ultimately converge on altered activity of transcription factors and affect the recruitment of transcriptional and chromatin remodeling machinery (Stadhouders et al., 2019). Therefore, cell identity can be regarded as an emergent property resulting from the crosstalk between transcription factors, epigenetic modifications, chromatin-associated proteins, and the three-dimensional organization of the genome. Thus, it is not surprising that cell type specific chromosome territories have been observed, and that chromosomal rearrangements disrupting topologically associating domains (TADs) have been linked to developmental defects and cancer, further suggesting functional relevance of genome conformation (Fraser et al., 2015; Stadhouders et al., 2019).

Higher order chromosome organization is extensively supported by the structural maintenance of chromosomes (SMC) family of multisubunit protein complexes (van Ruiten and Rowland, 2018). In mammals, SMC family complexes include cohesin and condensins I and II (Hirano, 2016). All contain two SMC subunits as well as unique non-SMC proteins. Specifically, cohesin contains SMC1 and 3, and condensins utilize SMC2 and 4. Condensins are further differentiated by their non-SMC subunits, with CAP-D3, CAP-G2, and CAP-H2 being unique to condensin II (Hirano, 2016). All SMC containing protein complexes form ring like shapes that can entrap chromatin fibers to organize them into higher order structures (van Ruiten and Rowland, 2018). In S phase, cohesins ensure sister chromatid alignment following replication (Morales and Losada, 2018), and condensin II ensures chromatid separation (Ono et al., 2013). In mitosis condensins orchestrate compaction of chromatin during prophase (Gibcus et al., 2018). Condensin I is largely cytoplasmic during interphase and its access to chromosomes increases after nuclear envelope breakdown in prometaphase (Ono et al., 2003). Cohesins and condensin II, on the other hand, are nuclear during interphase. It is in this context that cohesins have been determined to play an essential role in organizing the three-dimensional architecture of chromosomes (Haarhuis et al., 2017; Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017; Wutz et al., 2017). Interestingly, these studies show that some long-range chromosome contacts are preserved in the absence of cohesins, and condensin II co-occurs at TAD boundaries (Van Bortle et al., 2014; Yuen et al., 2017), suggesting it may also contribute to chromosome organization in interphase. However, the effect of

condensin II loss of function in mammalian cells is relatively unexplored compared to cohesin.

Condensins have been proposed to mediate chromosome structure and transcriptional control in model organisms. For example, condensin molecules participate in dosage compensation of hermaphroditic worms to silence one X chromosome (Crane et al., 2015). Furthermore, condensin II in *Drosophila* influences gene transvection effects caused by interphase chromosome pairing (Rowley et al., 2019). However, these are species specific paradigms that may not be conserved in mammalian biology. Condensin mutations have effects on transcription that in fission yeast are largely the consequences of aberrant mitosis and aneuploidy (Hocquet et al., 2018). Consequently, separation of indirect cell cycle effects from direct roles for mammalian condensin II on chromosome structure during interphase are rare. In response to estrogen stimulation, both condensin I and II are loaded at hormone-responsive, RNA-positive (eRNA⁺) enhancers to promote eRNA transcription, enhancer-promoter looping, and full activation (Li et al., 2015b). However, investigation of basal transcriptional effects of condensin II loss through deletion of CAP-H2 in post-mitotic hepatocytes failed to reveal effects on transcription (Abdennur et al., 2018). From this perspective, our knowledge of condensin II function in interphase nuclei and its effects on transcriptional control remains limited.

The retinoblastoma protein (pRB) is typically known for its role in cell cycle entry regulation, however, pRB deficient primary fibroblasts are known to possess relaxed chromatin (Herrera et al., 1996). In addition, pRB functions to recruit condensin II to mitotic chromosomes in *Drosophila* and mammalian cells (Longworth et al., 2008). Defective recruitment leads to mitotic abnormalities characterized by decondensed chromosomes, misshapen centromeres, and aberrant spindle microtubule attachments (Coschi et al., 2010; Manning et al., 2010). Two different partial loss of function murine *Rb1* gene mutations have been shown to compromise the recruitment of condensin II to pericentromeric heterochromatin (Coschi et al., 2014; Ishak et al., 2017). One is a point mutant allele that disrupts a CDK insensitive pRB-E2F1 interaction that allows pRB to associate with chromatin in S phase (Ishak et al., 2016). The other is a three amino acid substitution that disrupts the LXCXE binding site used by viral oncogenes called *Rb1^L*

(Isaac et al., 2006). Prior studies have shown that when pRB-LXCXE interactions are disturbed, transcription from E2F-containing promoters can still be regulated during proliferation and in differentiation (Isaac et al., 2006; Talluri et al., 2010). However, defective condensin II recruitment in *Rb1^{L/L}* cells leads to DNA replication dependent damage at pericentromeric sequences (Coschi et al., 2014), suggesting that pRB may play a role in recruiting condensin II to interphase chromosomes for non-mitotic functions.

To investigate condensin II regulation of chromosome structure in interphase cells, we utilized primary *Rb1^{L/L}* fibroblasts to disrupt condensin II recruitment. Using a combination of ChIP-Seq, RNA-Seq and chromosome conformation capture approaches we demonstrate that pRB recruits chromosome architectural protein complexes TFIIC and condensin II to bidirectional promoters. In the absence of their recruitment, reciprocal misexpression was observed at these gene pairs, suggesting condensin II recruitment has direct effects on local transcription through an insulator type mechanism. Defective recruitment of condensin II diminished long-range chromosome contacts between bidirectional promoters and distant loci. Loss of human *RBI* in lung adenocarcinoma reveals misexpression of genes at these bidirectional promoters. These experiments highlight transcriptional and architectural functions of an pRB-TFIIC-condensin II complex that may represent a transcriptional insulator function that is inactivated upon pRB loss in cancer.

3.3 Materials and Methods

3.3.1 Cell culture

MEFs were prepared and cultured from E13.5 embryos according to standard methods (Thwaites et al., 2016). *Rb1^{+/+}*, *Rb1^{-/-}*, *Rb1^{L/L}*, *E2f1^{+/+}*, *E2f1^{-/-}* MEFs were genotyped as previously described and used by passage 4 (Isaac et al., 2006).

3.3.2 ChIP

ChIP was conducted according to protocols adapted from Cecchini *et al.* (Cecchini et al., 2014). Briefly, asynchronously cycling cells were fixed in either 1% formaldehyde in 1X PBS or in 2mM ethylene glycol bis(succinimidyl succinate (EGS) in

1X PBS followed by 1% formaldehyde. Both fixing reactions were neutralized with 0.125M glycine. Cross-linked chromatin was sonicated so most chromatin was ≤ 400 bp. Sheared chromatin was then normalized between experimental groups and pre-cleared with protein G Dynabeads and IgG. Pre-cleared chromatin was then incubated with protein G Dynabeads and ChIP antibodies to immunoprecipitate proteins. The following antibodies were used to precipitate proteins: anti-CAP-D3 (Coschi et al., 2010), anti-CAP-H2 (A302-275A, Bethyl), anti-H3K27me3 (07-449, EMD Millipore), anti-pRb (C-15, Santa Cruz), anti-pRb (M-153, Santa Cruz), anti-pRb (M-136, (Cecchini et al., 2014)), anti-pRb (s855, (Cecchini et al., 2014)), anti-pRb (Hyb4.1, Developmental Studies Hybridoma Bank), and anti-TFIIC-220 (A301-291A, Bethyl). Cross-links were reversed at 65°C, and samples were treated with RNase and proteinase K. DNA was isolated for qPCR and/or library preparation.

For ChIP-qPCR the following primer pairs were used: Lmnb2-F, 5'-TCGGAGGCTCTATGGGAAAC-3'; Lmnb2-R, 5'-AAGGACAGTGCTTAGGGACG-3'; Cdca3-F, 5'-TCTCTCGCATCCAATGAGCG-3'; Cdca3-R, 5'-TACCCGCGGCGCTTTTTATT-3'; Pole-F, 5'-TCATTGGCCGAAGCCGTAG-3'; Pole-R, 5'-TTCCTCAGGACCATTGCGAC-3'; Mcm3-F, 5'-ATCCAGGAAGTCCAAGTAGTCTCTC-3'; Mcm3-R, 5'-TTGAAGTGGTTAGCCAATCATAACG-3'; Mcm3_-2kb-F, 5'-GCCAAGGCAAAACAACAATTTCTAC-3'; Mcm3_-2kb-R, 5'-CTATCTCTTTGATTTTGGGTGGCTG-3'; Hist1h3b_F, 5'-GTCTGTTTGAGGACACCAACCT-3'; Hist1h3b_R, 5'-TTTGGGTTCCAGTTTGCACCTTG-3'; Hoxa10_F, 5'-ACTGGGGATCTCGGTCCTAC-3'; Hoxa10_R, 5'-CAGATACTGGGCGGTGGTC-3'.

3.3.3 ChIP-Seq libraries and alignment

For ChIP-Seq, some previously published data were used for analysis. FASTQ reads were downloaded from the following datasets: CAP-D3 (GSE55040: GSM1328449, GSM1328450 and GSM1328448) Pol2 (GSE36027: GSM918740 and GSM918761), H3K4me1 (GSE31039: GSM769028 and GSM769030), H3K4me3 (GSE31039: GSM769029 and GSM769030), H3K27ac (GSE29218: GSM851277 and

GSM723008), H3K27me3 (GSE77993: GSM2064311 and GSM2064301), H3K36me3 (GSE53939: GSM1303764 and GSM1303761), and H3K9me3 (GSE53939: GSM1303762 and GSM1303761) (Coschi et al., 2014; Ishak et al., 2016; Pedersen et al., 2014; Shen et al., 2012). ChIP-Seq libraries created for this study occasionally used DNA from multiple ChIP replicates per genotype. To obtain enough DNA for library preparation, for CAP-H2 and CAP-D3 ChIPs, 5 IPs were combined and for H3K27me3 ChIPs, 2 IPs per library were used. ChIP libraries were sequenced using an Illumina NextSeq (High output 75 cycle kit).

All FASTQ reads (both from ChIP-Seq data published for other studies and from ChIP-Seq data obtained for this study) were aligned to mouse genome build mm9 using Bowtie2 version 2.3.0 (Langmead and Salzberg, 2012). Reads aligning to multiple locations of a particular element were distributed randomly to these positions, while multiple reads mapping to the same location were retained as previously described (Bulut-Karslioglu et al., 2014). The following command was used: `bowtie2 -t -p 4 -D 15 -R 2 -L 32 -i S,1,0.75 -x mm9 -U <reads>.fastq -S <output>.sam`.

3.3.4 Peak calling and annotation

Peaks were identified using MACS version 1.4.0rc2 according to the parameters stated below (Zhang et al., 2008). The following command was used: `macs14 -t <ChIP>.sam -c <input>.sam -n <output> -g mm -B -S`. For analysis of CAP-D3 localization, only peak locations that had 10 or more tags, a *P*-value of at most 10^{-5} , and a fold enrichment of at least 4 were used for subsequent analyses. Reads and peak locations were visualized using Integrative Genomics Viewer (IGV) (Robinson et al., 2011; Thorvaldsdottir et al., 2013).

Peak enrichment per genomic region was determined using CEAS and the mm9 RefSeq table supplied with the program (Shin et al., 2009). To determine the number and significance of peaks overlapping with other peaks or genomic elements, BEDTools was used (Quinlan, 2014; Quinlan and Hall, 2010).

To create heatmaps of reads, deepTools bamCompare was first used to generate bigWig files of ChIP reads normalized to input (Ramirez et al., 2014; Ramirez et al.,

2016). `computeMatrix` was then used to calculate read enrichment scores at promoters with wild type CAP-D3 peaks and `plotHeatmap` was used to plot those scores. To perform Pearson correlation analysis and create clustered heatmaps, `mutliBigwigSummary` computed the average scores for each of the files in every genomic region, and `plotCorrelation` was used to plot the data.

Functional annotation of genes of interest was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis et al., 2003). Genes were submitted to DAVID and analyzed using the functional annotation tool with the following categories selected: `GOTERM_BP_DIRECT`, `GOTERM_CC_DIRECT`, `GOTERM_MF_DIRECT`, `UP_KEYWORDS`, `UP_SEQ_FEATURE`, `KEGG_PATHWAY`, `BIOCARTA`, `INTERPRO`, `PIR_SUPERFAMILY`, and `SMART`. Functional annotation clustering was then used to cluster all of the gene enrichment results, with default parameters. The enrichment score for the annotation clusters represents the geometric mean, in $-\log$ scale, of the members' p-values in each corresponding annotation cluster. Functional annotation clusters were manually described based on their over-arching theme.

3.3.5 RNA-Seq

Total RNA from proliferating MEFs was isolated using TRIzol reagent according to the standard protocol (Invitrogen). Three *Rb1*^{+/+} and *Rb1*^{L/L} MEF pairs were used for this analysis. All samples were further processed and sequenced at the London Regional Genomics Centre (Robarts Research Institute, London, Ontario, Canada; <http://www.lrgc.ca>). Total RNA samples were quantified using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific) and quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies Inc.) and the RNA 6000 Nano kit (Caliper Life Sciences). They were then processed using the ScriptSeq Complete Gold Kit (H/M/R) (Illumina Inc.) which includes Ribo-Zero rRNA removal. Samples were fragmented, cDNA was synthesized, tagged, cleaned-up and subjected to PCR with indexed reverse primers (ScriptSeq Index PCR Primers) to permit equimolar pooling of samples into one library. Samples were sequenced on an Illumina NextSeq 500 (Mid output 150 cycle kit). FASTQ data files were then downloaded from BaseSpace.

For alignment of reads, a STAR v2.5.0a was used (Dobin et al., 2013). To generate a genome index, the mm9 chromFa.tar.gz file was downloaded from UCSC to create mm9_full_genome.fa, and the following command was used: STAR --runThreadN 4 --runMode genomeGenerate --genomeDir <STAR_genome_indices_directory> --genomeFastaFiles mm9_full_genome.fa --genomeSAsparseD 2 --limitGenomeGenerateRAM 20000000000 (Mouse Genome Sequencing et al., 2002). For the alignments, the following STAR command was used: STAR --runMode alignReads --runThreadN 8 --genomeDir <STAR_genome_indices_directory> --readFilesIn <reads>.fastq --sjdbGTFfile mm9_UCSC_knownGene.gtf --sjdbOverhang 149 --sjdbInsertSave All --outSAMtype BAM Unsorted SortedByCoordinate --outFileNamePrefix <output> --outReadsUnmapped Fastx --outMultimapperOrder Random --outSAMattributes NH HI AS NM XS --outWigType wiggle. StringTie v1.3.2 was then used to assemble RNA-Seq alignments into potential transcripts with the following command: stringtie <STAR_aligned_sample>.sortedByCoord.out.bam -o <output>.gtf -p 4 -G mm9_UCSC_knownGene.gtf -A <output>_gene_abund.tab -B -e (Pertea et al., 2015). The prepDE.py script provided with StringTie was then utilized to extract read count information from the generated files and DESeq2 was used to determine differential transcript abundance between *Rb1*^{+/+} and *Rb1*^{L/L} genotypes. When analyzing overall gene expression, all normalized transcript read counts from DESeq2 were summed for each gene and each of the three *Rb1*^{L/L} MEF preparations were compared to the average of three wild type MEFs. Heatmaps of data were created using matrix2png (Pavlidis and Noble, 2003).

3.3.6 Comparison of transcriptome analysis and ChIP-Seq data

To compare CAP-D3 ChIP-Seq binding sites to changes in transcription, BETA was used (Wang et al., 2013). First, RNA-Seq data were converted to BETA specific format (BSF) with the fold-change values being reported as wild type versus *Rb1*^{L/L} so expression changes when CAP-D3 could bind were being highlighted. The expression data were then compared to the filtered list of high confidence CAP-D3 peaks (see above) using BETA, taking into account CTCF binding sites using the following command:

BETA plus -p <CAP-D3_peaks>.bed -e <expression_changes>.txt -k BSF --bl -g mm9 --gs mm9.fa -n <output>.

3.3.7 qRT-PCR analysis of expression

Total RNA from proliferating MEFs was isolated using TRIzol reagent according to the standard protocol (Invitrogen). First-strand cDNA synthesis was performed using random primers, RNaseOUT and SuperScript III Reverse Transcriptase according to manufacturer's instructions (Invitrogen). Isolated cDNA was used in qRT-PCR reactions with iQ SYBR Green Supermix (Bio-Rad) using the following primers: Cdca3_exp_F, 5'-GTAGCAGACCCTCGTTCACC-3'; Cdca3_exp_R, 5'-ATTCCGACGCTTCTGTCTCC-3'; Usp5_exp_F, 5'-ATGGCGGAGCTGAGTGAAGA-3'; Usp5_exp_R, 5'-ATAGAGGCCACCTCAGACT-3'; Pole_exp_F, 5'-GAGAAGGTGCCTGTGGAACA-3'; Pole_exp_R, 5'-GCTGTAGGCGGTTGGTAAGA-3'; Pxmp2_exp_F, 5'-GACTGCCTAGCTGTTGGGTG-3'; Pxmp2_exp_R, 5'-CCAAGGGCTGACAAAATGCC-3'; GAPDH_exp_F, 5'-GCACAGTCAAGGCCGAGAAT-3'; and GAPDH_exp_R, 5'-GCCTTCTCCATGGTGGTGAA-3'. Resulting target Cq values were normalized to GAPDH, then expressed as fold change relative to the global wild type mean.

3.3.8 Flow Cytometry

Cells were plated on 10 cm plates at a density of 900,000 cells per plate. Approximately 24 hours after, cells were pulsed with BrdU for a duration of 2 hours. Cell cycle analysis was then carried out as previously described (Cecchini et al., 2012).

3.3.9 Nocodazole treatment

Cells were plated on 10 cm plates at a density of 900,000 cells per plate or on 6 cm plates at a density of 300,000 cells per plate. Approximately 24 hours after seeding, cell media was replaced with standard culture media containing either DMSO only (control) or 20 ng/mL nocodazole. One day after treatment with nocodazole, cells were processed for flow cytometry and qRT-PCR expression analysis (see above).

3.3.10 Chromosome conformation capture (3C) analysis

3C was conducted according to protocols adapted from Hagege *et al.* (Hagege *et al.*, 2007). Briefly, cells were trypsinized, centrifuged and resuspended to make a single-cell suspension followed with crosslinking in 1% formaldehyde/10% FCS/PBS. The reactions were quenched with glycine and cells were then pelleted and lysed in 5 mL cold lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1% Triton X-100 and protease inhibitors). Samples were digested with HindIII, ligated using T4 DNA ligase, samples were treated with proteinase K and cross-links were reversed at 65°C. Samples were then treated with RNase A followed by phenol-chloroform purification. 3C samples were digested with a second restriction enzyme, EcoRI, to help minimize potential PCR biases resulting from limited template accessibility. The resulting final 3C products were quantified in triplicates by quantitative TaqMan real-time PCR after 3C DNA was normalized to a final concentration of 50 ng/μL. Bacterial artificial chromosomes (BACs), clones RP23-333M22 and RP23-55K5, containing the *Pxmp2/Pole* and *Cdca3/Usf5* bidirectional promoters of interest, respectively, were used as control templates. To generate random ligation products of HindIII fragments at the regions of interest, BACs were individually digested with HindIII, ligated, and digested again with EcoRI. This ligated BAC DNA was serially diluted and used to generate standard curves in each qPCR runs for each primer pair to which all 3C products were normalized. The 3C signals at each locus were further normalized to those from a control locus, ERCC3, with primers and a probe described previously (Splinter *et al.*, 2006). Probe and primer sequences are as follows (primer names correspond to approximate position (in kb) relative to bidirectional promoters): *Cdca3*_const, 5'-TAGAGCAAAGCTACACCGGG-3'; *Cdca3*_Probe, 5'-FAM-AGAGAGATCTATCCAGGTCTCACAGGCC-TAMRA-3'; *Cdca3*_93, 5'-GACCACTGCGAGACGGAAG-3'; *Cdca3*_73, 5'-CCCCAGATACTCAATCCCTG-3'; *Cdca3*_57, 5'-CCTCTCCCCTCCTTTCTTCC-3'; *Cdca3*_45, 5'-TACAGATGGTTGCGAGGCAC-3'; *Cdca3*_33, 5'-GCTGGGAGGATGAGAAAATGAC-3'; *Cdca3*_13, 5'-ATCCAGAGATTCACGCTTGCT-3'; *Cdca3*_+1, 5'-CCTGGAGGAGGCCATTCAAG-3'; *Cdca3*_+8, 5'-CCAAGTCTACCATCTCGGGG-3'; *Cdca3*_+9, 5'-

TTGGGTGACTGAGATAGCCC-3'; Cdca3_+37, 5'-CCTGGCACACAGCAAGC-3';
 Cdca3_+48, 5'-ACCAGAACAGACATCTCAAGTAACA-3'; Cdca3_+57, 5'-
 TGCATAGCGATGGGTTTCACA-3'; Cdca3_+59, 5'-
 TGGCCAGAATGACTGCAAAGAA-3'; Cdca3_+63, 5'-
 CCCTCTCCCAGGGTAAAGC-3'; Cdca3_+64, 5'-
 GTTCCAGAACTTTCCTCTCTTTGC-3'; Cdca3_+69, 5'-
 GAGGCATGCCACGTAAGC-3'; Cdca3_+70, 5'-CCC GCAGGAGAGGGACTAAA-3';
 Cdca3_+71, 5'-TCGGAATTTGCTCCCAGAGTC-3'; Cdca3_+79, 5'-
 GTGGGGACAGTAGAGGAAGC-3'; Cdca3_+80, 5'-GCTGAGCCCAGTGGAAAG-3';
 Cdca3_+88, 5'-GGCTCTGTTTACCACA ACTGTC-3'; Cdca3_+96, 5'-
 AGAAGCTGACTTGGGCTACAT-3'; Cdca3_+123, 5'-
 GTCTAACGGGTTGGTAGAGTGG-3'; Cdca3_+127, 5'-
 ACATAATAACACCAGGGGCCG-3'; Cdca3_+129, 5'-
 CACTGACTGACGCACAGAAGAA-3'; Pole_const, 5'-GGTGTCTGCTGCCAAGG-
 3'; Pole_Probe, 5'-FAM-TGCTCCGCGCGGCTAC-TAMRA-3'; Pole_-30, 5'-
 ACAGCCAGA ACTACACAGAGG-3'; Pole_-25, 5'-
 ACTGTTGTGGACTGATGCTTAGAA-3'; Pole_-24, 5'-
 TGCCCTTGTCTGAAGTCTGC-3'; Pole_-23, 5'-CTCAGGAGACAGGCAAGCTAA-
 3'; Pole_-6, 5'-ATGGCAATCACAGGCACAAG-3'; Pole_-5, 5'-
 CTGTCCTCATCTGTGCCCTC-3'; Pole_-4, 5'-TGCCAAAGTATGGGGGATGTG-3';
 Pole_+3, 5'-TTCCTCGGTGGGCATTCTTC-3'; Pole_+5, 5'-
 AGAAGATTGTGATCAGTGTTTCTGA-3'; Pole_+11, 5'-
 CTTGTTGAACTCCTGCTCTTGC-3'; Pole_+18, 5'-CTAATGGAACCGAGGAGCCG-
 3'; Pole_+28, 5'-GCAATACACAAACGCTCTTGGTC-3'; Pole_+32, 5'-
 AGCCAGACCAACCACTCTTC-3'; Pole_+36, 5'-
 TGAGAGTTGTCCTGTATGGAACG-3'; Pole_+50, 5'-
 AAAATGCCACCTTGCTGTG-3'; Pole_+62, 5'-
 TTCAGGGTTCTCTCTTTGGGAGTG-3'; Pole_+66, 5'-
 GTGTTCTTCCTGTGTATATACTTGC-3'; Pole_+77, 5'-
 CTCCCTTAAGTTTGTCGGGCT-3'; Pole_+82, 5'-TGGAGGAATGTGACTGGGGA-
 3'; Pole_+84, 5'-CTGGGCGCTTGGAGGTTTAC-3'; Pole_+101, 5'-

TTGAAGCCTAGGTGGGAGTCTT-3'; Pole_+104, 5'-
 ACAGGAGAGAGGCAGGTATGTC-3'; ERCC3_1, 5'-
 GCCCTCCCTGAAAATAAGGA-3'; ERCC3_2, 5'-GACTTCTCACCTGGGCCTACA-
 3'; ERCC3_Probe, 5'-FAM-AAAGCTTGCACCCTGCTTTAGTGGCC-TAMRA-3'.

3.3.11 Circularized chromosome conformation capture (4C)-Seq

Samples were fixed, digested with HindIII, ligated and crosslinks were reversed as detailed in the 3C protocol above. To generate a 4C products, 3C products were processed according to protocols adapted from Splinter *et al.* (Splinter et al., 2012). Briefly, 3C products were digested with NlaIII, ligated with T4 DNA ligase, and purified. Primers were designed for the bidirectional promoters of interest for inverse PCR reactions. For Illumina sequencing, all reading primers started with an Illumina read adapter sequence, 5'-
 AATGATACGGCGACCACCGAGATCTACACACTCTTCCCTACACGACGCTC
 TTCCGATCT-3'; followed by a unique tag 0-3 nucleotides in length. These tags were unique to each sample and were used both to ensure that nucleotide content for every cycle of sequencing is not the same (despite using the same primer sequence for all samples at the same viewpoint) and to pool multiple samples together. After the tag, the reading primers contained sequence unique to the viewpoint of interest, as close as possible to the HindIII digestion site. For the *Pole/Pxmp2* bidirectional promoter, this sequence was 5'-GATTCACTCCAAACTCCACAAA-3', and for the *Cdca3/Usp5* bidirectional promoter, it was 5'-AGCAAGAGAGTGTAGCTAAG-3'. For the reverse sequencing primers for these viewpoints, the primers again started with an Illumina adapter sequence, 5'-CAAGCAGAAGACGGCATAACGAGAT-3'; followed by sequence unique to the viewpoint of interest, close to the NlaIII restriction sites. For the *Pole/Pxmp2* bidirectional promoter, this sequence was 5'-
 TCCAAAGGATATATGAGGTTTCG-3', and for the *Cdca3/Usp5* bidirectional promoter, it was 5'-GTCTGACTTGCAGTTTTTCAG-3'. 4C product for sequencing was then prepared using Expand Long Template Polymerase (Roche) and 3.2 µg of 4C template split between 16 PCR reactions. PCR reactions were pooled after completion and

purified. 4C PCR products were then pooled equimolarly into one library and sequenced on an Illumina NextSeq (Mid Output 150 cycle kit).

Sequence reads were first trimmed to 20 bp total after the primer and bait sequence (excluding the HindIII site) using the FASTQ trimmer from the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). w4CSeq was then used to identify 4C sites, statistically significant regions, and look at intra- and inter-chromosomal interactions (Cai et al., 2016). For this analysis, the mm9 genome was used, 500 enzyme sites was selected as the bin size for trans chromosomes (size_inter), 100 enzyme sites was selected as the bin size for cis chromosomes (size_intra), 3000 enzyme sites was selected as the background window size for the cis chromosome (window_intra), and 0.05 was used as the FDR threshold. To compare the similarity of significant interacting regions, BEDTools Jaccard was used (Quinlan, 2014; Quinlan and Hall, 2010). BEDTools merge was used to combine the significant interacting regions together from each replicate for each viewpoint, and BEDTools intersect to determine the merged regions that were the same between genotypes.

3.3.12 Gene expression in lung adenocarcinoma

To determine the expression of genes in lung adenocarcinoma samples, the TCGA PanCancer Atlas dataset on cBioPortal was used. Only the “complete tumors” were analyzed, with the gene set user-defined list being entered as “RB1: AMP HOMDEL HETLOSS mut” as well as only the gene names for the other genes we were interested in. The plot function was then used to obtain data comparing *RBI* copy number to the mRNA expression *Z*-scores of genes of interest. All genes that contained an *Rbi*^{+/+} CAP-D3 peak within their bidirectional promoter region that is ≤5000 bp and was the higher expressing of the two genes from each bidirectional promoter in *Rbi*^{L/L} compared to *Rbi*^{+/+} MEFs were attempted to be analyzed in these tumor samples. However, not all mouse genes were able to be analyzed in human, and this was particularly true for the RIKEN genes. *Z*-scores from *RBI* “deep deletion” and “shallow deletion” samples were then manually grouped to obtain the deletion category, and were compared to *Z*-scores from the *RBI* diploid population for each gene of interest using t-tests. Data from genes that had significantly increased *Z*-scores in *RBI* deleted tumors (deep and shallow

deletion combined) compared to *RBI* diploid tumors were then submitted to matrix2png to create a heatmap where the data were normalized such that each row had a mean of zero and a variance of 1 (Pavlidis and Noble, 2003).

3.3.13 EZH2 inhibitor treatment

To inhibit EZH2 in MEFs, asynchronously growing P4 MEFs were seeded at a density of 4×10^5 cells per 10 cm plate. The next day, when cells were at approximately 50% confluency, treatment with GSK343 or vehicle (DMSO) with treatment volumes of 0.1% of total media volume were started. Treatments of 2.5 μ M or 5 μ M GSK343 were administered once every 24 hr for 48 hr. GSK343 was stored at a stock concentration of 1000x treatment dose in DMSO at -20°C protected from light. After 48 hr of drug treatment, total RNA from proliferating MEFs was isolated, converted to cDNA and expression was analyzed as above. Whole cell extracts were also collected after 48 hr of drug treatment and H3K27me3 (07-449, EMD Millipore) blots were performed.

3.4 Results

3.4.1 Condensin II and TFIIIC occupy promoters in an pRB-dependent manner

To investigate genome wide localization of condensin II during interphase, we performed chromatin immunoprecipitation followed by high throughput sequencing (ChIP-Seq). We precipitated CAP-D3, a subunit that is unique to condensin II, from chromatin isolated from primary mouse embryonic fibroblasts (MEFs). Regions of local enrichment were determined using Model-based Analysis for ChIP-Seq (MACS) (Zhang et al., 2008). Using Cis-regulatory Element Annotation System (CEAS) (Shin et al., 2009), the distribution of CAP-D3 ChIP-Seq peaks within genomic features was determined. This data demonstrated that 40% of 1547 total CAP-D3 peaks were within 1 kb of a transcriptional start site (TSS), which we refer to as proximal promoters (Figure 3.1A). We carried out ChIP-Seq for both CAP-D3 and CAP-H2 and used a combination of ethylene glycol bis(succinimidyl succinate) (EGS) and formaldehyde fixation, to preserve protein-protein interactions that may recruit condensin II to chromatin and capture these longer distance interactions (Figure 3.1B). These ChIP-Seq experiments

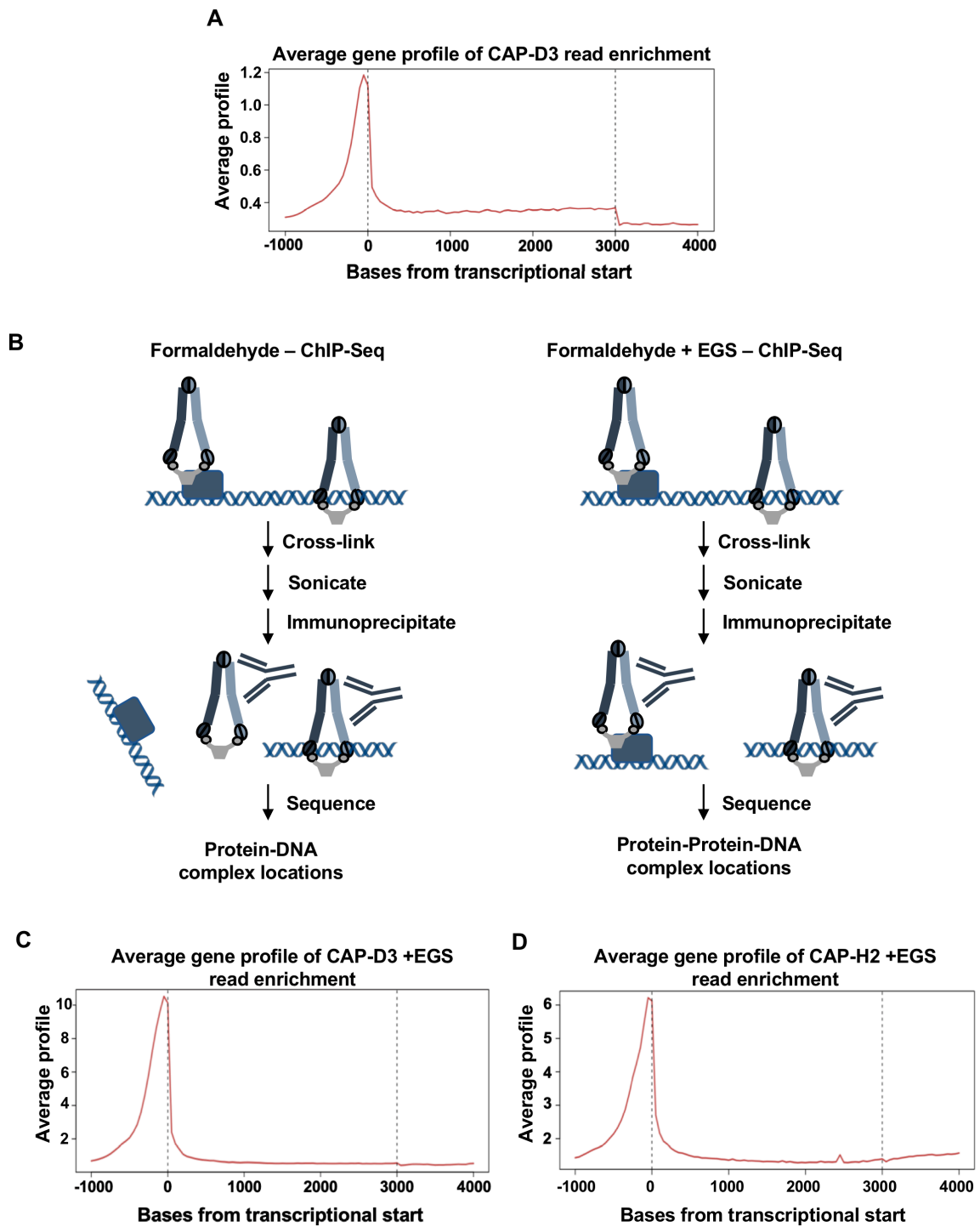


Figure 3.1: CAP-D3 is enriched at proximal promoters.

Figure 3.1: CAP-D3 is enriched at proximal promoters.

(A) Average distribution of *Rb1*^{+/+} CAP-D3 ChIP-Seq reads 1000 bp upstream and downstream of genes, as well as average ChIP signals on a meta-gene, where every gene is normalized to a length of 3000 bp. **(B)** Schematic of captured ChIP fragments with formaldehyde versus formaldehyde and ethylene glycol bis(succinimidyl succinate) (EGS) fixation. Formaldehyde uses methylene bridges to crosslink and is primarily a protein-DNA crosslinker, whereas EGS is a protein-protein crosslinker, using N-hydroxysuccinimide (NHS) esters to yield stable amide bonds (Hoffman et al., 2015; Tian et al., 2012). Since condensin II may not be directly binding to DNA at many locations it interacts with, EGS+formaldehyde dual crosslinking was used to have a more in-depth look at chromatin contacts. **(C)** Average distribution of EGS and formaldehyde fixed CAP-D3 ChIP-Seq reads 1000 bp upstream and downstream of genes, as well as average ChIP signals on a meta-gene, where every gene is normalized to a length of 3000 bp. **(D)** Average distribution of EGS and formaldehyde fixed CAP-H2 ChIP-Seq reads 1000 bp upstream and downstream of genes, as well as average ChIP signals on a meta-gene.

revealed an order of magnitude more peaks, as there were now 17576 total CAP-D3 peaks for example, but peaks still showed similar promoter occupancy (Figure 3.1C&D), suggesting promoters are the principal condensin II binding site. Since CAP-H2 peaks intersect with CAP-D3 peaks at proximal promoters with 72% frequency, we conclude these represent condensin II complexes and not free subunits.

Since pRB recruits condensin II to pericentromeric heterochromatin (Coschi et al., 2014), we compared CAP-D3 occupancy locations from ChIP-Seq data between wild type and *Rb1^{L/L}* murine fibroblasts to learn where pRB-condensin II localizes in interphase nuclei. This *Rb1* mutant encodes a three amino acid substitution (I746A, N750A, M754A) in the pRB pocket domain designed to disrupt interactions with viral oncogenes such as HPV E7, while preserving interactions with E2F transcription factors (Isaac et al., 2006). Condensin II recruitment to pericentromeric heterochromatin is reduced in *Rb1^{L/L}* cells (Coschi et al., 2014), but chromatin loading of cohesin and condensin I is normal (Coschi et al., 2010). Using MACS, 66% of all CAP-D3 peaks found in wild type MEFs are lost in the *Rb1^{L/L}* mutant. Figure 3.2A shows CAP-D3 ChIP-Seq read enrichment at 52 well-known E2F regulated cell cycle promoters for both genotypes. CAP-D3 peaks were found in 9 of these promoters in wild type and 7 of these promoters lost peaks in *Rb1^{L/L}* cells. This suggests that the large quantity of promoters occupied by pRB and condensin II are unrelated to pRB's canonical function in G1-S regulation of E2F transcription. An example of one of these locations is the *Timm13* promoter, which is best known as the *Lmnb2* origin of replication (Figure 3.2B&C). Previous studies have identified pRB and E2F1 at this location by ChIP, despite its lack of an E2F DNA sequence element (Avni et al., 2003; Mendoza-Maldonado et al., 2010). Furthermore, MACS determined CAP-D3 reads contributing to the peak at this location was significant in both *Rb1^{+/+}* and *Rb1^{L/L}* ChIP-Seq experiments (Figure 3.2B), although read build ups are clearly reduced in *Rb1^{L/L}*. ChIP-qPCR was used to quantitatively compare CAP-D3 binding at *Timm13/Lmnb2* and it showed that CAP-D3 binding was only significantly enriched over the IgG background control in *Rb1^{+/+}* MEFs (Figure 3.2D). This indicates that condensin II recruitment is pRB-dependent, but the *Rb1^L* mutation does not completely eliminate it in ChIP-Seq experiments. To examine pRB-dependency of CAP-D3 localization at promoters, CAP-D3 ChIP-Seq reads from *Rb1^{+/+}*

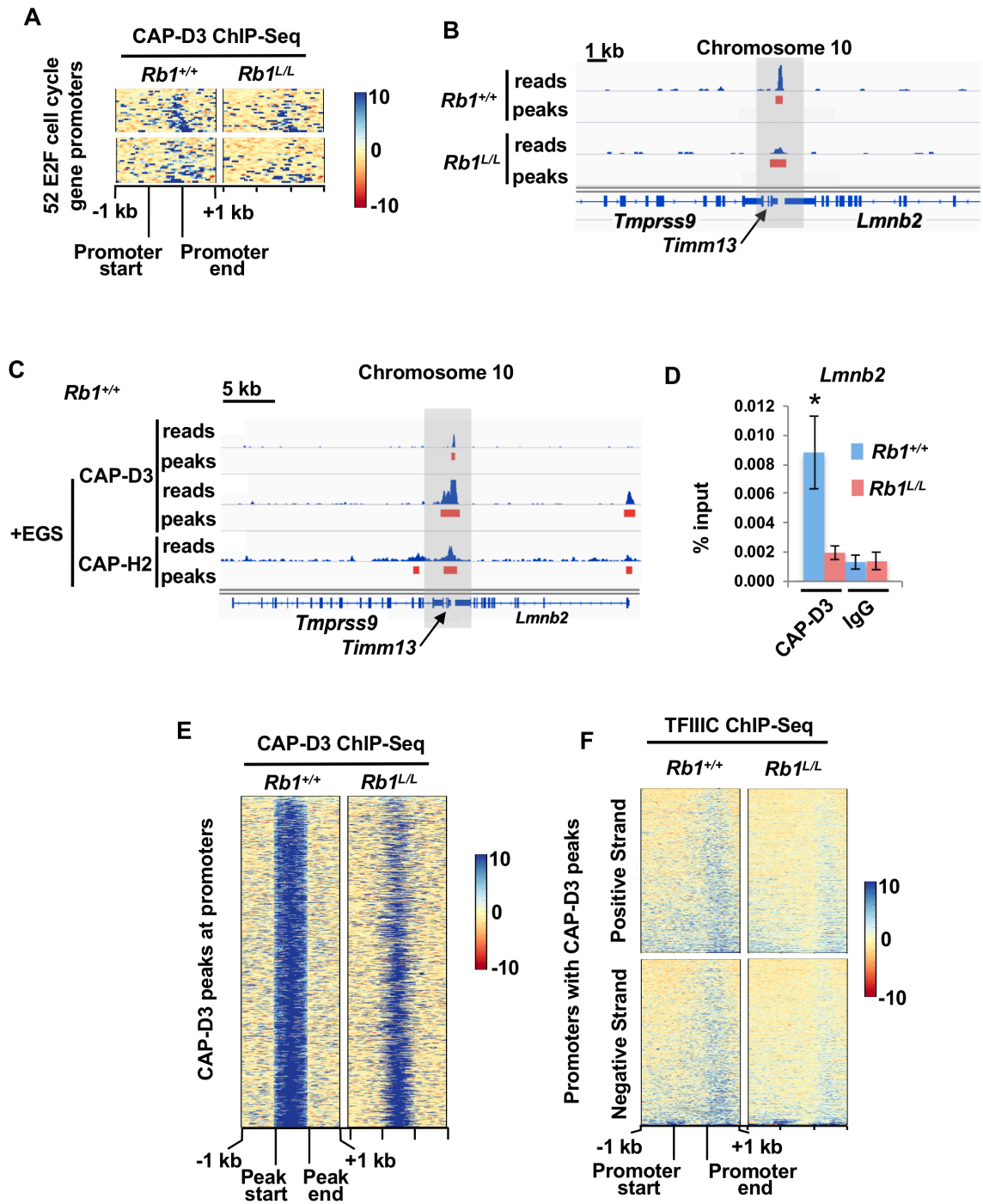


Figure 3.2: CAP-D3 binds at some promoters in an pRB-dependent manner.

Figure 3.2: CAP-D3 binds at some promoters in an pRB-dependent manner.

(A) Heatmaps of CAP-D3 ChIP-Seq read abundance relative to input controls in *Rb1*^{+/+} and *Rb1*^{L/L} MEFs at 1 kb proximal promoter regions of 52 pRB regulated E2F cell cycle target genes. Heatmaps are divided into those promoters on the forward strand (top) and those on the reverse strand (bottom). (B) ChIP-Seq tracks for CAP-D3 from *Rb1*^{+/+} and *Rb1*^{L/L} MEFs at the *Timm13* promoter/*Lmnb2* origin of replication where there is a CAP-D3 peak in both genotypes. Scales of read build-ups are set to 25. (C) ChIP-Seq tracks for CAP-D3 and CAP-H2 from *Rb1*^{+/+} MEFs at the *Lmnb2* origin of replication. Cells were fixed either with formaldehyde alone or formaldehyde in combination with EGS (+EGS) prior to ChIP. All scales of read build-ups are set to 50 to show the relative differences. (D) ChIP-qPCR for CAP-D3 and IgG controls at the *Lmnb2* origin of replication, seen in B&C. (n=7), error bars are ± 1 SEM, * $P < 0.05$; determined by *t*-test. (E) Heatmaps of CAP-D3 ChIP-Seq read abundance at 1271 promoters that contain an *Rb1*^{+/+} CAP-D3 peak within 1 kb of their transcriptional start sites. The heatmap shows the location of the *Rb1*^{+/+} CAP-D3 peaks in the center, as well as 1 kb upstream and downstream of these regions. (F) Heatmaps of TFIIC ChIP-Seq read abundances in *Rb1*^{+/+} and *Rb1*^{L/L} MEFs at the 1271 1 kb promoter regions that contain an *Rb1*^{+/+} CAP-D3 peak. Heatmaps are divided into promoters on the forward strand (top) and reverse strand (bottom) and are centered on the promoter regions.

and *Rb1^{L/L}* MEFs were compared at 1271 promoters that were most enriched for CAP-D3 (Figure 3.2E). At these promoter locations in *Rb1^{L/L}* cells, CAP-D3 peaks are lost 30% of the time, or are shifted or diminished similar to *Timm13/Lmnb2*, further indicating that condensin II binding is extensively reduced in the *Rb1^{L/L}* genotype at these locations. Therefore, we sought to use comparisons between wild type and *Rb1^{L/L}* MEFs to study the effects of condensin II loss on interphase chromosomes.

Condensin II recruitment to promoters has also been shown to co-occur with the RNA polymerase III transcription factor and chromosome architectural protein complex TFIIC (Yuen et al., 2017). To reconcile roles for pRB and TFIIC in condensin II recruitment, we carried out ChIP-Seq for subunit 1 of TFIIC in *Rb1^{+/+}* and *Rb1^{L/L}* MEFs. This data revealed that TFIIC is abundant just 3' of the promoter regions containing CAP-D3 peaks (Figure 3.2F), and TFIIC read build-ups in this 3 kb window are reduced in *Rb1^{L/L}* MEFs. Using MACS, duplicate ChIP-Seq experiments revealed 67% of these promoter views contain TFIIC peaks in *Rb1^{+/+}* cells, and this is decreased modestly to 53% in *Rb1^{L/L}* cells as locations adjacent to the TSS are unaffected by the *Rb1^L* mutation. This indicates that recruitment of TFIIC and condensin II, two known components of topological boundaries, is pRB dependent specifically at these promoter locations.

To gain insight into condensin II's role at mammalian promoters, we investigated their chromatin state from publicly available ChIP-Seq datasets generated with murine fibroblasts. Extensive simultaneous occupancy of these promoters with both Pol2 and H3K4me3 is a strong indication of active transcription (Barski et al., 2007) (Figure 3.3A), and there was 98% peak overlap between these modifications and CAP-D3 containing promoters. H3K4me1, a mark of inactive/poised enhancers, was de-enriched at the TSS of these promoters (Figure 3.3A), and overlaps infrequently with these promoters (<20%). H3K27ac, a mark of active enhancers and promoters, is also highly abundant at these promoters with 85% overlap (Barski et al., 2007; Creyghton et al., 2010; Wang et al., 2008). Conversely, repressive chromatin marks such as H3K27me3 and H3K9me3 were almost completely excluded from these promoters (Figure 3.3A), with 2% and 1% occupancy respectively. Lastly, H3K36me3, which preferentially marks exons in gene bodies, was largely devoid from these promoter regions of interest (<6%, Figure 3.3A)

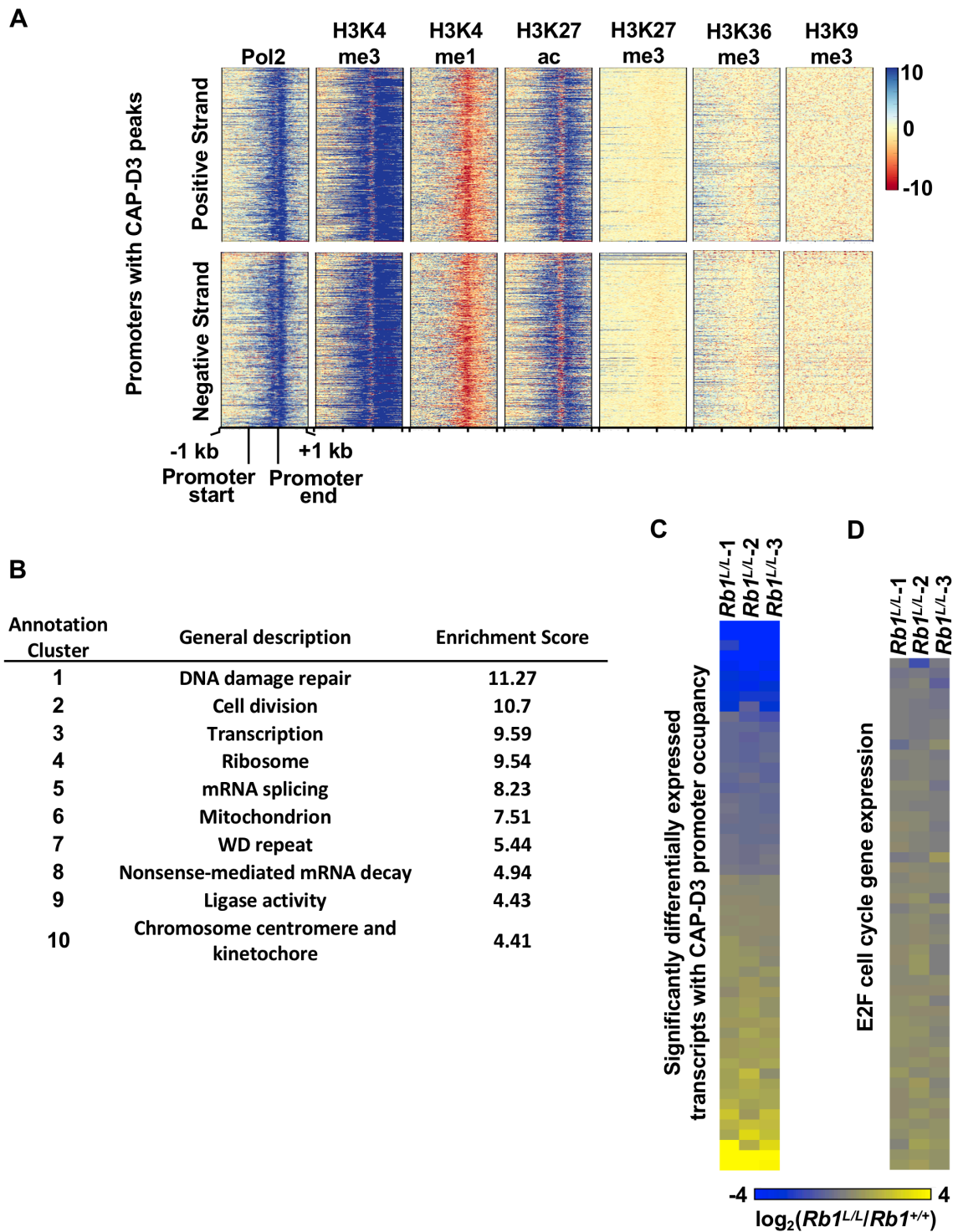


Figure 3.3: CAP-D3 is localized to active promoters and changes in binding are correlated with activation and repression of genes.

Figure 3.3: CAP-D3 is localized to active promoters and changes in binding are correlated with activation and repression of genes.

(A) Heatmaps of ChIP-Seq read abundances for histone tail modifications and RNA Pol2 in wild type MEFs at the 1271 1 kb promoter regions that contain a CAP-D3 peak. Heatmaps are divided into promoters on the forward strand (top) and reverse strand (bottom). (B) Top 10 most enriched annotation clusters from the genes that contain an *Rb1*^{+/+} CAP-D3 peak within 1 kb of their TSS. Analysis was performed using DAVID and the enrichment score for each annotation cluster represents the geometric mean, in – log scale, of the members' *P*-values in each cluster. (C) Expression from three *Rb1*^{L/L} MEF preparations relative to the average of three wild type MEFs, as determined by RNA-sequencing (RNA-Seq). The heatmap is of transcripts with significantly different expression that also had an *Rb1*^{+/+} CAP-D3 peak within their 1 kb proximal promoter. (D) Heatmap of the overall change in gene expression for pRB regulated E2F cell cycle target genes determined by RNA-Seq as in C.

(Barski et al., 2007; Kolasinska-Zwierz et al., 2009). The most obvious conclusion from this work is that pRB-condensin II complexes are localized to active promoters. Importantly, this pattern of histone modifications is highly divergent from known regulatory roles for pRB in gene silencing that are characterized by establishment of H3K9me3 in senescence (Narita et al., 2003), H3K27me3 in terminal differentiation during development (Blais et al., 2007), and deacetylation of histones in quiescence (Narita et al., 2003). This further emphasizes that pRB recruitment of condensin II to these promoters represents a different functional paradigm.

The genes bound by condensin II within their proximal promoters were subject to functional annotation clustering using the Database for Annotation, Visualization and Integrated Discovery (DAVID) to determine the biological paradigms that could be regulated by the pRB-condensin II complex (Dennis et al., 2003). This analysis demonstrates that this complex could be responsible for regulating several important processes within a cell, such as DNA damage repair, cell division, transcription and translation (Figure 3.3B). RNA-Seq was then used to investigate if changes in condensin II localization in *Rb1^{+/+}* compared to *Rb1^{L/L}* MEFs are linked to changes in expression of these genes. This shows that 54 genes that have condensin II binding within their proximal promoters have significantly differentially expressed transcripts in *Rb1^{L/L}* compared to *Rb1^{+/+}* MEFs, 25 of which are downregulated and 29 of which are upregulated (Figure 3.3C). This is in contrast to E2F regulated cell cycle genes, from Figure 3.2A, in which 6 genes were significantly upregulated in *Rb1^{L/L}* MEFs (Figure 3.3D).

The presence of H3K27ac at condensin II occupied promoters (Figure 3.3A), may be indicative of enhancer interactions, or may demonstrate that these occupied locations can function as enhancers themselves. For this reason, it is not surprising that some genes exhibited reduced expression in *Rb1^{L/L}* cells at locations where condensin II recruitment was reduced. The *Tubb2a* locus illustrates this effect with CAP-D3 and TFIIC loss near the TSS in *Rb1^{L/L}* cells and approximately 40% reduction in transcript abundance for this gene (Figure 3.4A-C). A more intriguing result is the apparent upregulation of gene expression at condensin II occupied promoters when it is diminished from these locations

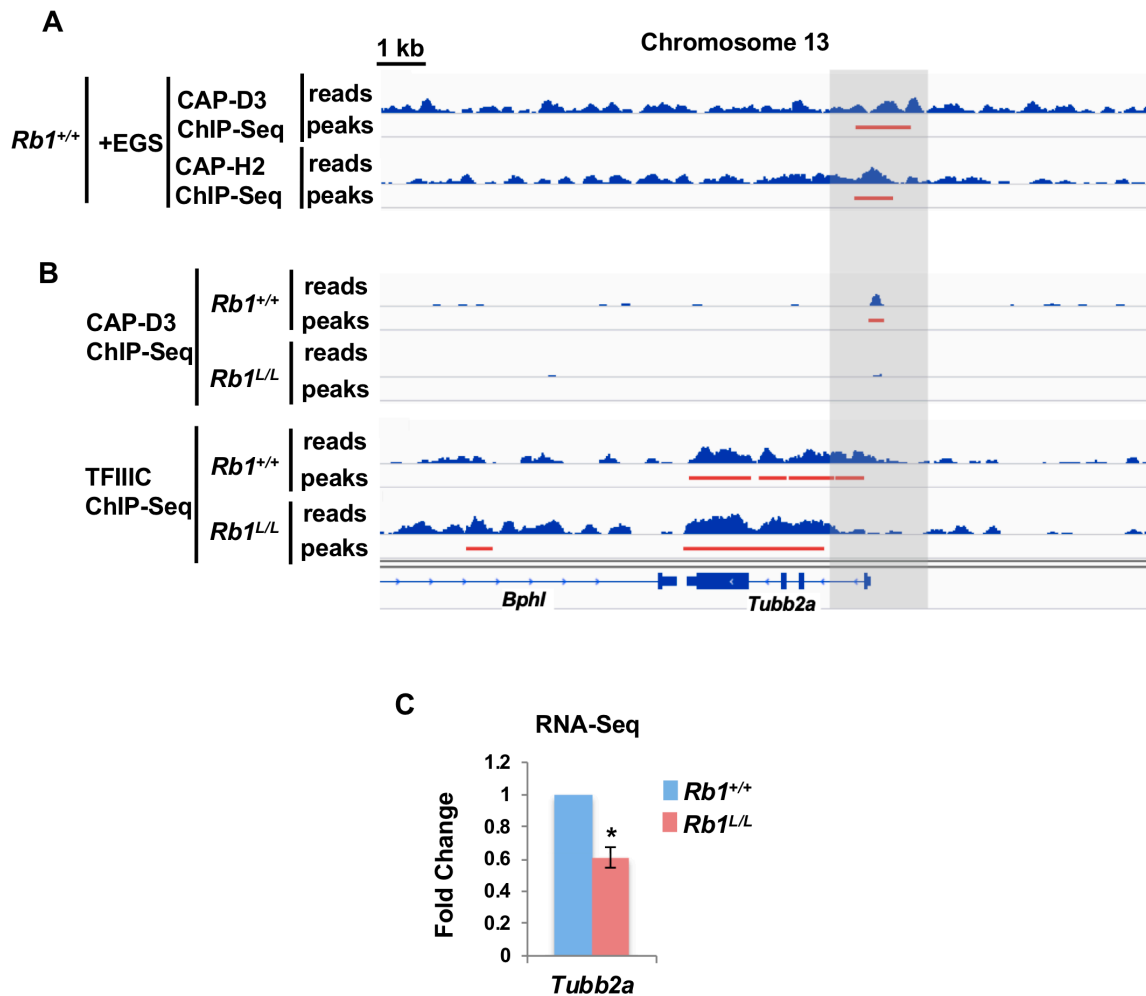


Figure 3.4: Changes in condensin II and TFIIC binding are correlated with downregulation of *Tubb2a*.

(A) Genome browser view of the *Tubb2a* gene is shown. ChIP-Seq for CAP-D3 and CAP-H2 from wild type MEFs fixed with formaldehyde in combination with EGS (+EGS) are displayed at the top. (B) CAP-D3 and TFIIC ChIP-Seq tracks for both *Rb1*^{+/+} and *Rb1*^{L/L} formaldehyde fixed MEFs are shown at the bottom. All scales of read build-ups in A and B are set to 35 to show the relative differences. (C) Relative expression of the *Tubb2a* gene in *Rb1*^{L/L} MEFs, as determined by RNA-Seq. Error bar is ± 1 SEM, * $P < 0.05$.

in *Rb1^{L/L}* cells. One possible explanation may be that condensin II functions as an insulator between closely spaced promoters and reduction of occupancy enhances transcription of an already active promoter.

Further investigation of pRB-condensin II occupied promoter locations demonstrates that it is enriched at bidirectional promoters in *Rb1^{+/+}* MEFs compared to their expected frequency in the genome (Figure 3.5A). Furthermore, condensin II presence at bidirectional promoters is more than expected by chance, given the relative abundance of bidirectional promoters (Figure 3.5B). Analysis of expression of genes on either side of bidirectional promoters bound by CAP-D3 in *Rb1^{+/+}* MEFs demonstrates that upregulation and/or downregulation occurs at many of these locations in *Rb1^{L/L}* cells (Figure 3.5C). The contribution of condensin II to the differential gene expression in *Rb1^{L/L}* cells was further investigated using Binding and Expression Target Analysis (BETA), which was able to relate CAP-D3 peak locations with expression of nearby genes (Figure 3.5D) (Wang et al., 2013). This further suggested that condensin II may have repressive and activating influences at neighboring genes. To validate this correlation, the locations of CAP-D3 peaks were computationally redistributed throughout the genome, while maintaining the same number and length of peaks. These “shuffled peaks” are not correlated with transcript changes in *Rb1^{L/L}* compared to *Rb1^{+/+}* MEFs, further suggesting a functional connection between condensin II occupancy and nearby gene expression. Consequently, we focused our investigation on condensin II’s role at bidirectional promoters.

3.4.2 Differential regulation of gene expression by condensin II at bidirectional promoters

We used data from our BETA analysis to determine locations of altered gene expression between *Rb1^{L/L}* and *Rb1^{+/+}* MEFs and found bidirectional promoters where expression of one of the two genes was correlated to CAP-D3 binding. We focused on two bidirectional promoter locations, *Usp5/Cdca3* and the *Pxmp2/Pole* to determine condensin II’s role in gene regulation at these locations. In both cases, CAP-D3 peaks were found in the intergenic region between the two divergently transcribed genes in wild type MEFs (Figure 3.6A). This peak was lost in *Rb1^{L/L}* MEFs at the *Pxmp2/Pole* location,

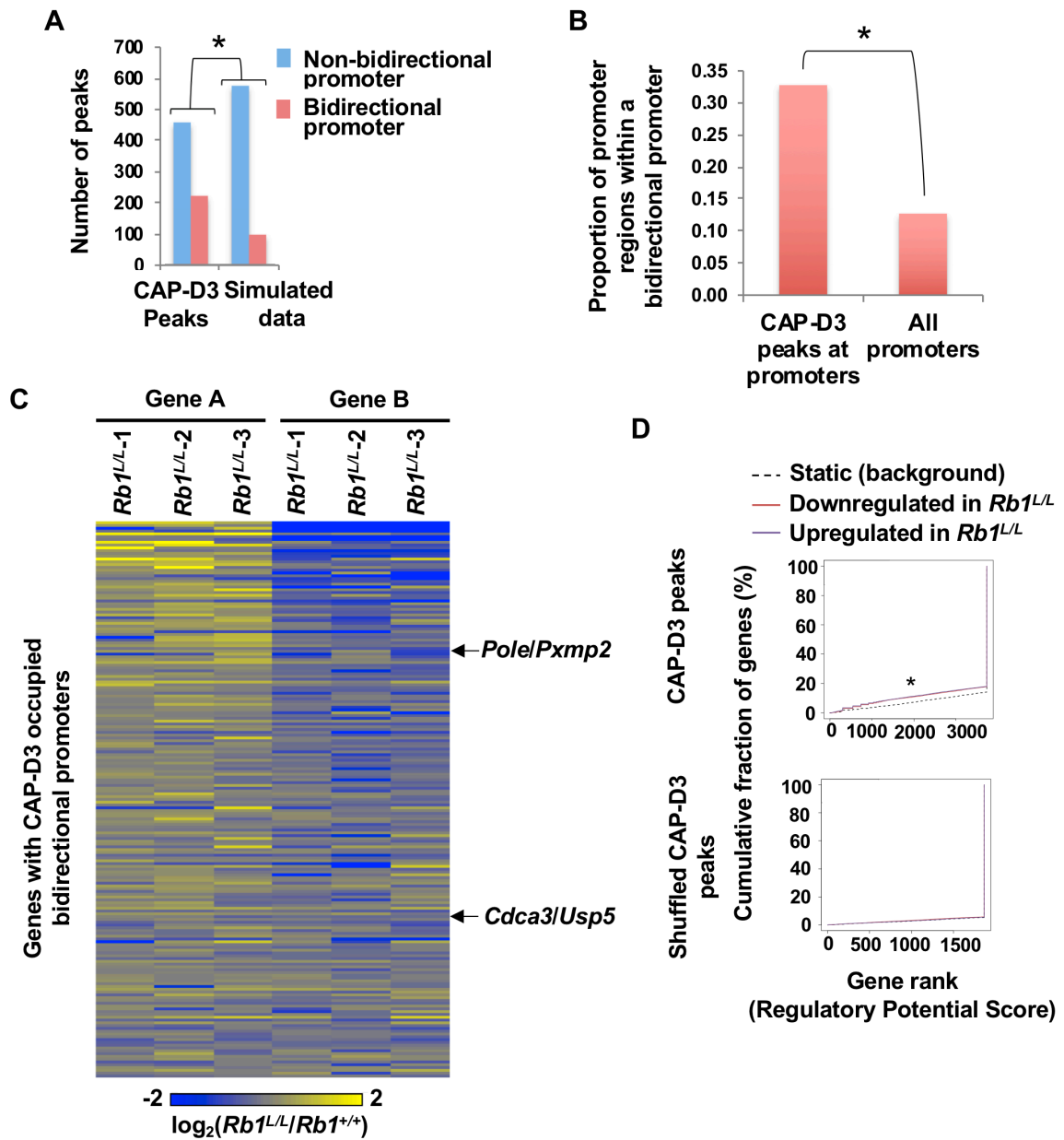


Figure 3.5: Condensin II binding is enriched at bidirectional promoters and is correlated with changes in expression.

Figure 3.5: Condensin II binding is enriched at bidirectional promoters and is correlated with changes in expression.

(A) Peaks that are within 3 kb upstream from the TSS of genes were grouped according to whether they are found at bidirectional promoters or not. Compared with simulated data, where peaks are shuffled to promoters at random, there are more CAP-D3 peaks at bidirectional promoters as determined by χ^2 -test. (B) Peaks that are within 3 kb upstream from the TSS of genes were grouped according to whether they are found at bidirectional promoters or not. Compared with the proportion of promoters genome wide (“All promoters”) that reside within a bidirectional promoter region, there are more promoters containing CAP-D3 peaks that are part of bidirectional promoters, as determined by χ^2 -test. (C) Expression changes from RNA-Seq of genes that contain an *Rb1*^{+/+} CAP-D3 peak within their bidirectional promoter region that is ≤ 5000 bp. Each row represents a bidirectional promoter, with the higher expressing of the two genes in *Rb1*^{L/L} compared to *Rb1*^{+/+} MEFs on the left (labelled Gene A). (D) Binding and Expressing Target Analysis (BETA) was used to establish if condensin II localization in asynchronous *Rb1*^{+/+} MEFs, as determined by CAP-D3 ChIP-Seq, was correlated with changes of expression between *Rb1*^{L/L} and *Rb1*^{+/+} MEFs, as determined by RNA-Seq. Genes that had either an increase or a decrease in expression in the *Rb1*^{L/L} background (when CAP-D3 cannot bind to pRB) also tended to have an enrichment of CAP-D3 binding sites in wild type MEFs. When peak locations were redistributed randomly throughout the genome, these “shuffled peaks” are not correlated with transcriptional changes in *Rb1*^{L/L} compared to *Rb1*^{+/+} MEFs. * $P < 0.05$.

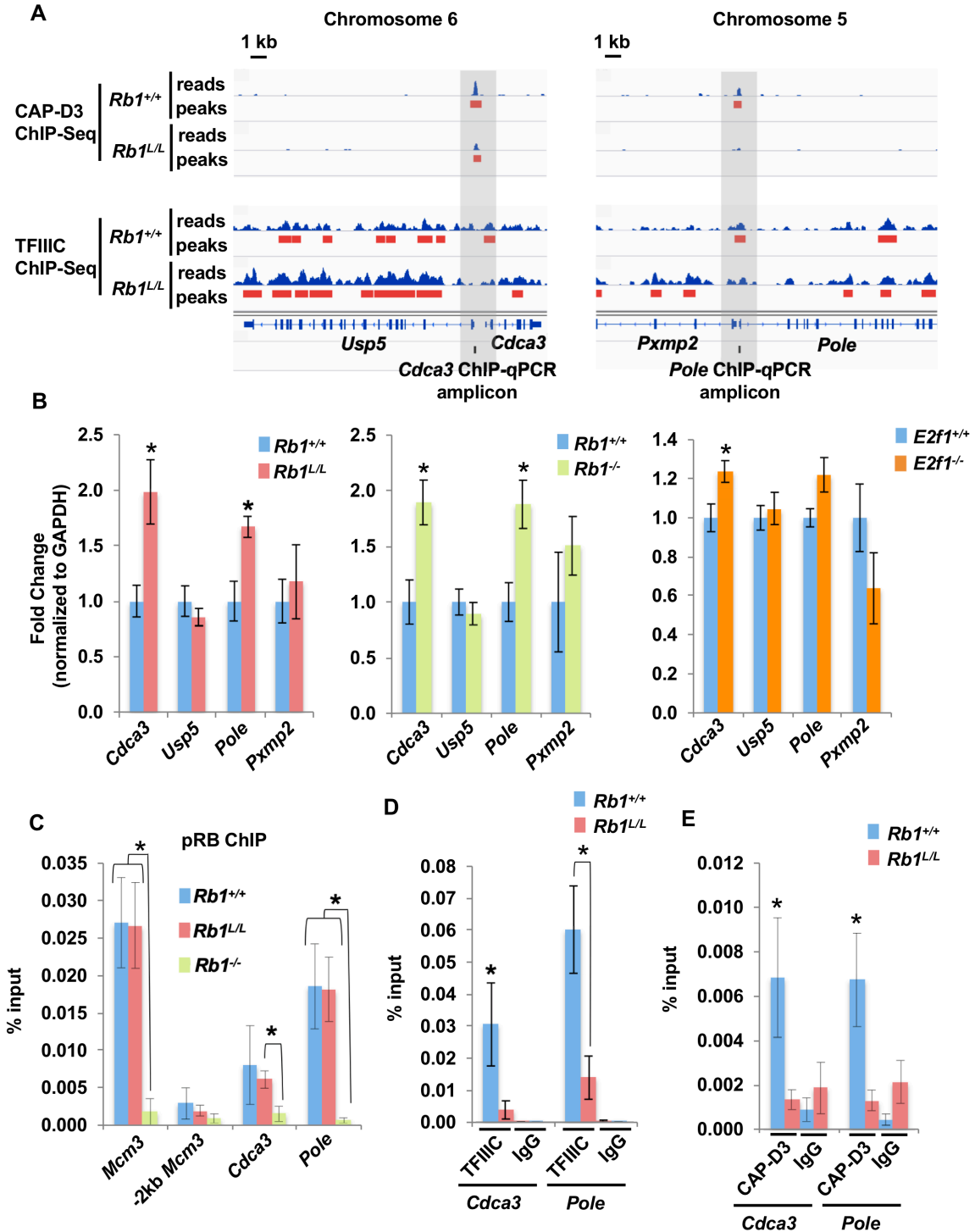


Figure 3.6: Bidirectional promoters with differential expression changes upon loss of CAP-D3 binding in *Rb1*^{L/L} MEFs.

Figure 3.6: Bidirectional promoters with differential expression changes upon loss of CAP-D3 binding in *Rb1^{L/L}* MEFs.

(A) *Cdca3/Usp5* and *Pole/Pxmp2* bidirectional promoter locations and their respective CAP-D3 and TFIIC ChIP-Seq reads and peaks in both *Rb1^{+/+}* and *Rb1^{L/L}* MEFs. All scales of read build-ups are set to 35. (B) RT-qPCR of the two sets of bidirectional genes in *Rb1^{+/+}* and *Rb1^{L/L}*, *Rb1^{+/+}* and *Rb1^{-/-}*, and wild type and *E2f1^{-/-}* MEF littermate pairs. (n=5). (C) pRB ChIP-qPCR results from asynchronously growing *Rb1^{+/+}*, *Rb1^{L/L}* and *Rb1^{-/-}* MEFs at the bidirectional promoter locations indicated in A. The *Mcm3* promoter and -2 kb upstream location are used as positive and negative controls for pRB ChIPs, respectively (n=5). (D) TFIIC ChIP-qPCR results from *Rb1^{+/+}* and *Rb1^{L/L}* MEFs at the same bidirectional promoter locations as above (n=6). (E) CAP-D3 ChIP-qPCR results from asynchronously growing *Rb1^{+/+}* and *Rb1^{L/L}* MEFs (n=5). All error bars are ± 1 SEM. * $P < 0.05$; determined by *t*-test.

and diminished at the *Usp5/Cdca3* in *Rb1^{L/L}*. TFIIC binding at these locations also appeared to require wild type pRB (Figure 3.6A).

We used qPCR to quantitate mRNA levels for each of these four transcripts. Consistent with condensin II regulating expression at both of these bidirectional promoters, expression of one of the genes was increased in the *Rb1^{L/L}* MEFs; *Cdca3* at the *Usp5/Cdca3* bidirectional promoter and *Pole* at the *Pxmp2/Pole* promoter (Figure 3.6B). This change was also seen at both locations in *Rb1^{-/-}* cells compared to wild type, as well as in *E2f1^{-/-}* MEFs at the *Usp5/Cdca3* bidirectional promoter (Figure 3.6B). This indicates that loss of an pRB-condensin II complex may play a role in the differential expression of these divergently paired genes.

To understand the architecture of protein complexes at these locations, ChIP-qPCR was used to quantitate the binding of pRB, TFIIC, and condensin II components. First, pRB localization is not changed in *Rb1^{L/L}* MEFs compared to *Rb1^{+/+}* MEFs at these promoters (Figure 3.6C). Only *Rb1^{-/-}* cells are devoid of pRB at these locations and this pattern matches occupancy at *Mcm3*, a classical pRB-E2F transcriptional target (Figure 3.6C). In addition, TFIIC localization is significantly reduced between *Rb1^{+/+}* and *Rb1^{L/L}* MEFs at both bidirectional promoters (Figure 3.6D). Lastly, CAP-D3 is present at these locations in *Rb1^{+/+}* cells, but reduced to background levels in *Rb1^{L/L}* MEFs (Figure 3.6E). Collectively, these experiments demonstrate pRB dependent recruitment of both TFIIC and condensin II to these locations. When the *Rb1^L* encoded protein is present, but unable to assemble this complex, expression of one gene in each pair is elevated, implicating TFIIC and condensin II in this regulatory event.

Loss of condensin II function affects progression through mitosis (Nishide and Hirano, 2014), and *Rb1^{L/L}* MEFs similarly display misshapen chromosomes, mitotic delay, and aneuploidy in daughter cells (Coschi et al., 2010; Isaac et al., 2006). Studies of condensin mutants in fission yeast suggest that gene misexpression may be a consequence of mitotic errors and aneuploidy (Hocquet et al., 2018). Therefore, to increase the quantity of *Rb1^{+/+}* MEF cells with >4N DNA and to increase G2/M phase levels to a similar quantity as in *Rb1^{L/L}* fibroblasts, nocodazole was used to inhibit mitotic progression (Figure 3.7A&B). This induced mitotic errors and delayed progression, but

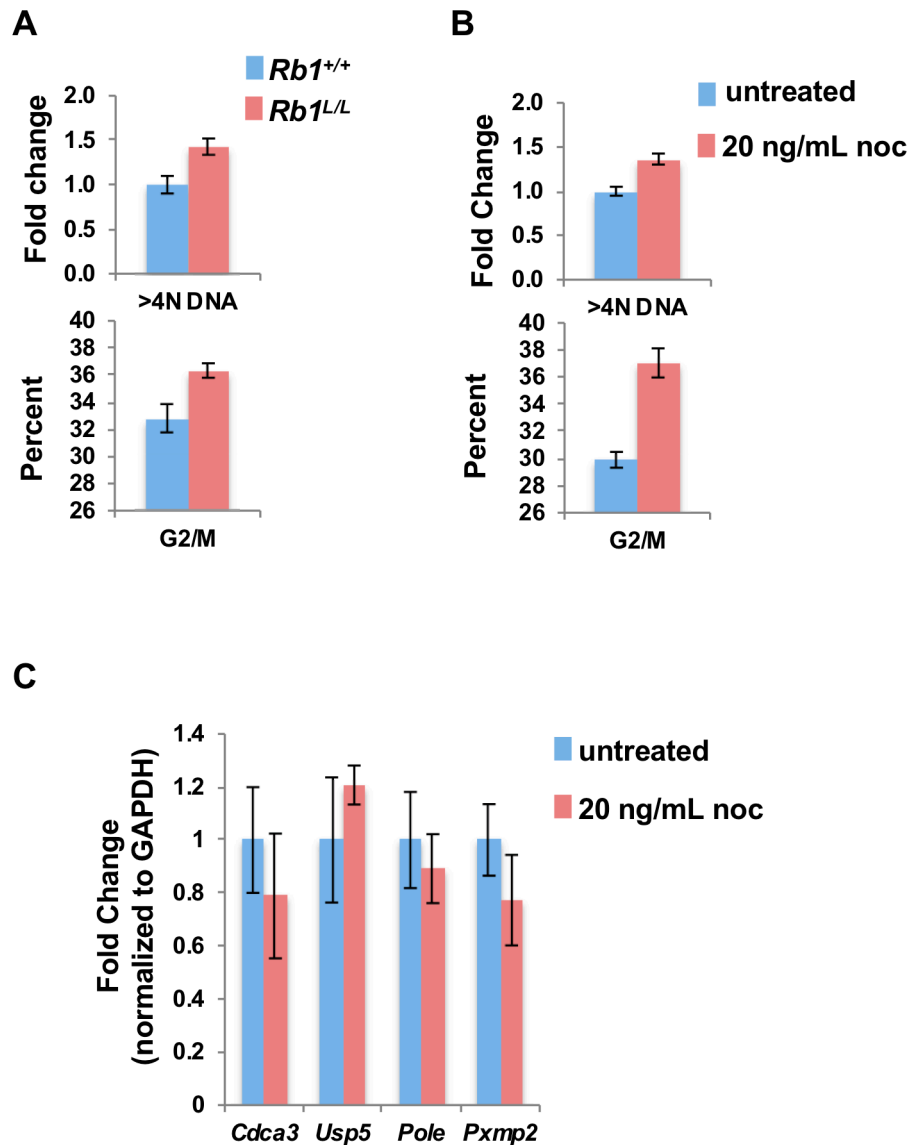


Figure 3.7: Gene expression at bidirectional promoters is not altered by microtubule inhibitor induced aneuploidy.

(A) Differences in >4N DNA content between *Rb1*^{+/+} and *Rb1*^{L/L} MEFs and the percent of cells in G2/M as determined by PI staining and flow cytometry (n=6). (B) *Rb1*^{+/+} MEFs were nocodazole (noc) treated for 24 hours to increase the >4N DNA and the G2/M content in wild type MEFs similarly to the *Rb1*^{L/L} MEFs and the percent of cells in G2/M as determined by PI staining and flow cytometry (n=4). (C) RT-qPCR was used to quantitate expression of genes at the bidirectional promoters in untreated and nocodazole treated *Rb1*^{+/+} cultures. All error bars are ± 1 SEM.

did not cause expression changes at the *Usp5/Cdca3* and *Pxmp2/Pole* bidirectional promoters (Figure 3.7C). This suggests that indirect effects on transcription from defective condensin II function in mitosis in *Rb1^{L/L}* cells is unlikely to explain misregulation at these bidirectional promoters. Additionally, depletion of condensin II components restricts progression through mitosis in mammalian cells and potentially causes indirect transcriptional effects through synchronization. For example, transcription of histone gene clusters is tightly linked to S phase to ensure chromatinization is coordinated with DNA replication, and accumulation of cells in mitosis may lower transcript levels for these genes simply because the number of cells in S phase is reduced. Asynchronously proliferating *Rb1^{L/L}* cell cultures have similar quantities of S phase as *Rb1^{+/+}* (Isaac et al., 2006). Our analysis shows that the histone cluster on chromosome 13 displays pRB dependent occupancy of TFIIC and condensin II (Figure 3.8A-C) and when this is reduced in *Rb1^{L/L}* cells, expression of this gene cluster is actually elevated (Figure 3.8D). Overall, these analyses argue against potential cell cycle and mitotic fidelity effects on transcription in *Rb1^{L/L}* fibroblasts and further emphasize that loss of condensin II occupation likely has direct effects on gene expression at bidirectional promoters.

3.4.3 pRB-condensin II complexes mediate long-range chromosome interactions

Because of the transcriptional changes seen at bidirectional promoters and the ability of condensin II to affect chromatin architecture and looping, conformation dependent interactions were investigated. Genome views of regions surrounding *Usp5/Cdca3* and *Pxmp2/Pole* revealed regular peak locations for condensin II (Figure 3.9A&D). One possibility is that condensin II organizes loops between these locations to influence gene expression within these territories. For this reason, local chromatin folding was interrogated using chromosome conformation capture (3C) analysis. We performed 3C-qPCR on cross-linked, HindIII cut, and ligated chromatin from four wild type and *Rb1^{L/L}* MEF pairs using *Usp5/Cdca3* and *Pxmp2/Pole* bidirectional promoters as the reference points (Figure 3.9B&E). This revealed sites of local interaction within these chromosomal regions. Most notably, some loops appeared in locations that are unrelated to condensin II peaks (Figure 3.9C, ~60 kb from the *Cdca3* bait). Furthermore, there do

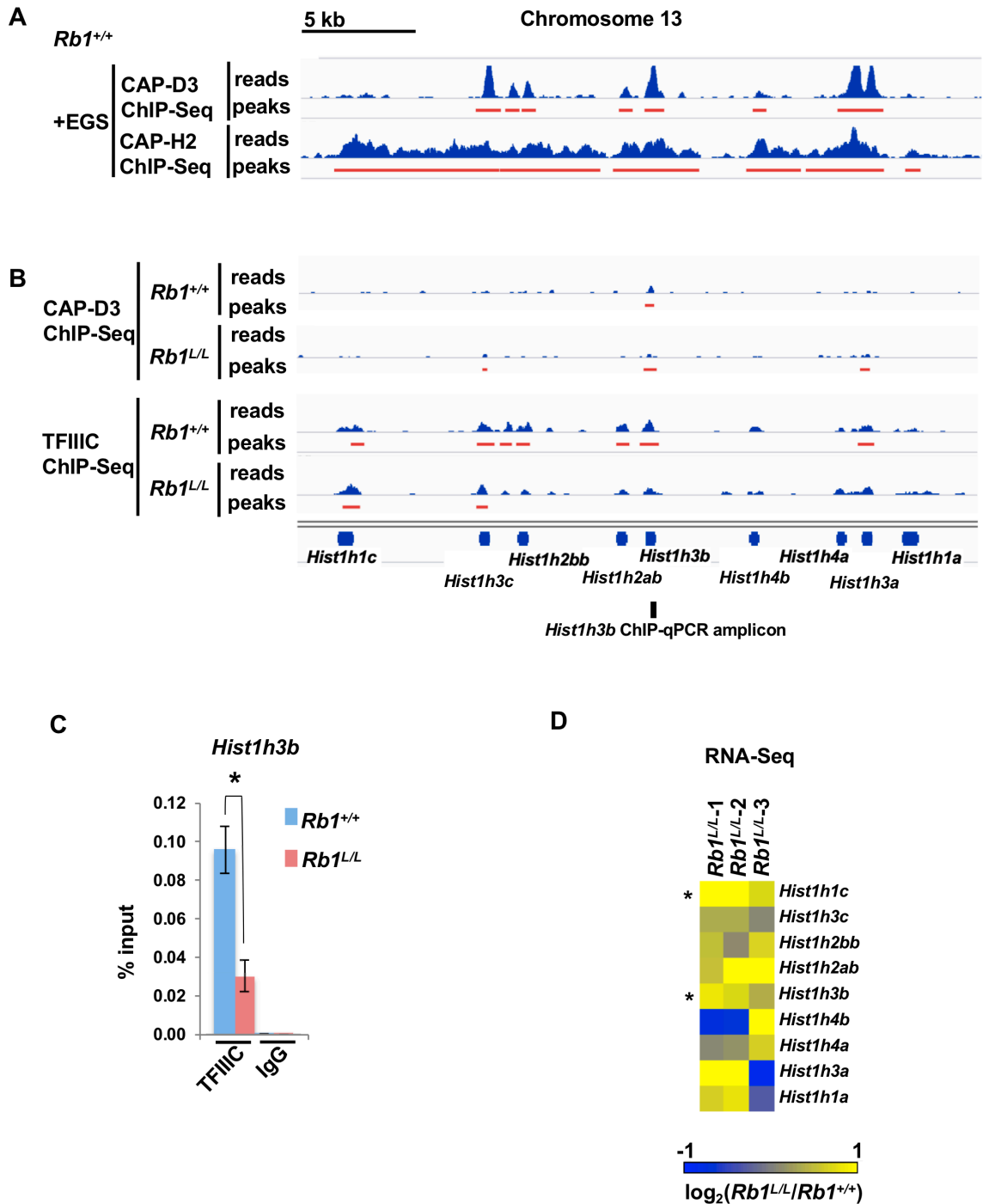


Figure 3.8: Reduced condensin II and TFIIIC localization at the *Hist1* gene cluster in *Rb1*^{L/L} MEFs are correlated with upregulation of some histone genes.

Figure 3.8: Reduced condensin II and TFIIC localization at the *Hist1* gene cluster in *Rb1^{L/L}* MEFs are correlated with upregulation of some histone genes.

(A) Genome browser view of the *Hist1* gene cluster on chromosome 13. ChIP-Seq for CAP-D3 and CAP-H2 from wild type MEFs fixed with formaldehyde in combination with EGS (+EGS) prior to ChIP are displayed at the top. **(B)** CAP-D3 and TFIIC ChIP-Seq tracks for both *Rb1^{+/+}* and *Rb1^{L/L}* formaldehyde fixed MEFs are shown. All scales of read build-ups in A and B are set to 50 to show the relative differences. **(C)** TFIIC ChIP-qPCR results from *Rb1^{+/+}* and *Rb1^{L/L}* MEFs at the *Hist1h3b* gene location indicated in B. (n=6), error bars are ± 1 SEM. * $P < 0.05$; determined by *t*-test. **(D)** Expression changes of individual genes within the *Hist1* cluster as determined by RNA-Seq. The expression from three *Rb1^{L/L}* MEF preparations is displayed relative to the average of three wild type MEFs. Differentially expressed genes between genotypes with an adjusted *P*-value of < 0.1 are indicated with an asterisk.

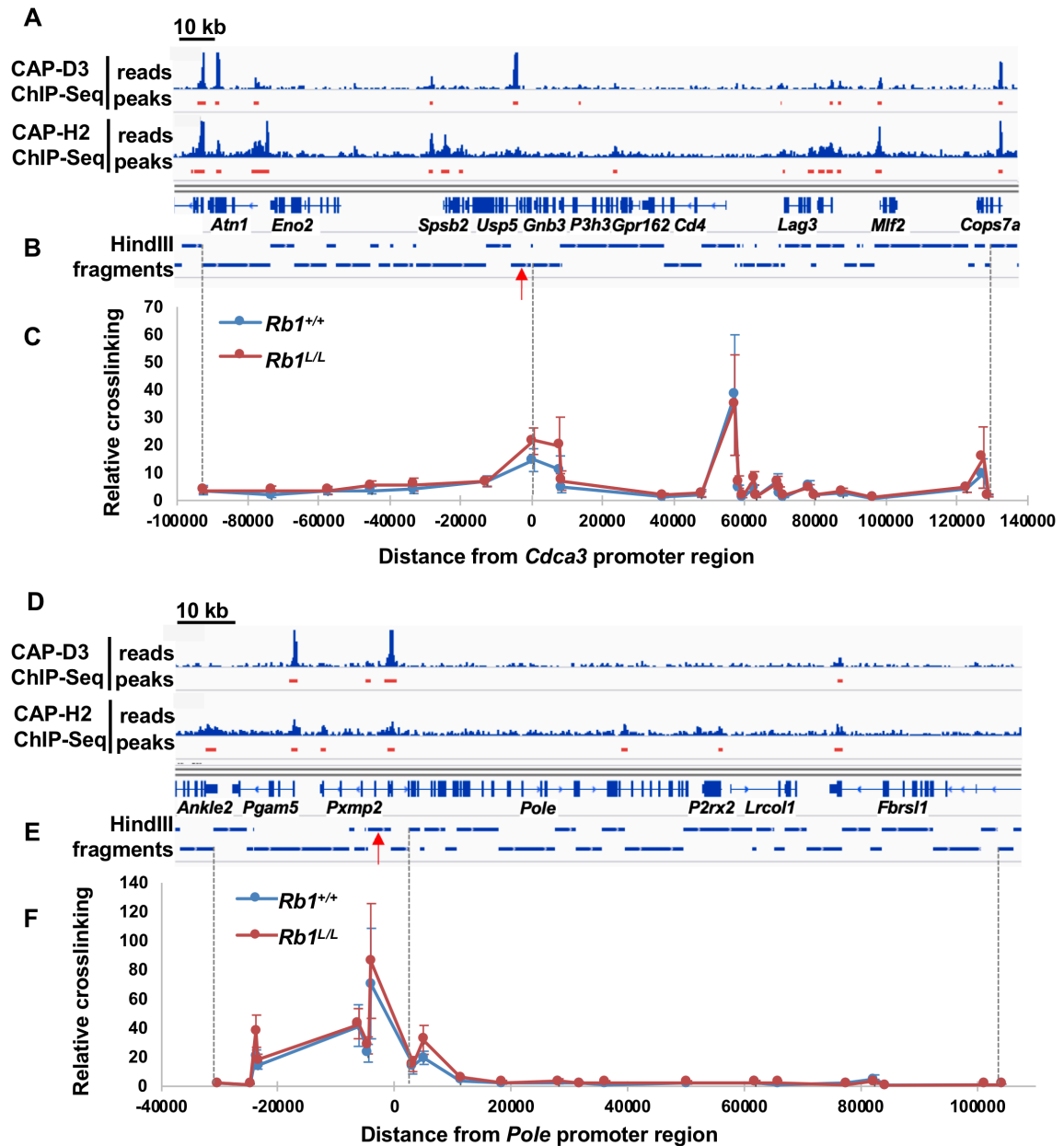


Figure 3.9: Preservation of 3C-detected chromatin loops in *Rb1^{L/L}* cells.

(A&D) CAP-D3 and CAP-H2 ChIP-Seq read tracks (EGS and formaldehyde fixed), surrounding the *Cdca3/Usp5* bidirectional promoter in A and the *Pole/Pxmp2* locus in D. Read build-up scale is 50. (B&E) HindIII restriction enzyme fragments are shown as tracks in relation to the gene structure of these genomic locations. (C&F) 3C interaction frequencies for selected HindIII fragments (indicated as dots) in the *Cdca3/Usp5* and *Pole/Pxmp2* regions of interest are displayed in graphical format. The grey dotted vertical lines relate restriction enzyme fragments to specific data points. The bait fragments are indicated by red arrows. 3C crosslinking frequencies were normalized to the ERCC3 locus ($n=4$). All error bars are ± 1 SEM.

not appear to be any interactions that are gained or lost between genotypes (Figure 3.9C&F). Since local chromatin conformation does not appear to be different between genotypes, despite loss of condensin II at these locations, short-range loops seem unlikely to be responsible for the changes in gene expression of *Cdca3* and *Pole* in *Rb1^{L/L}* MEFs.

To broaden our search for condensin II dependent chromosome interactions, we used *Usp5/Cdca3* and *Pxmp2/Pole* promoter viewpoints as the bait regions for circularized chromosome conformation capture (4C)-Seq. Chromatin were processed from four wild type and *Rb1^{L/L}* littermate pairs and alignment and analysis of data was performed using w4CSeq (Cai et al., 2016). From this data, significant interacting regions were produced for each individual experiment (Figure 3.10A) and these were merged together such that each region shown was found in two or more replicates for each genotype (Figure 3.10B).

Initially we focused on one region that was common to *Rb1^{+/+}* and *Rb1^{L/L}* 4C-Seq data, but that showed different boundaries. The *Hoxa10* locus, approximately 70 Mb away from *Cdca3* on chromosome 6, contains ChIP-Seq peaks for CAP-D3 and extensive contacts with CAP-H2 suggesting condensin II occupies this location (Figure 3.11A). Based on 4C-Seq data, *Hoxa10* interactions were detected in all four *Rb1^{+/+}* 4C-Seq replicates using *Usp5/Cdca3* as the bait, but only one in *Rb1^{L/L}* cells (Figure 3.11A). Consistent with previous experiments, ChIP-qPCR reveals that pRB localization is not changed between *Rb1^{+/+}* and *Rb1^{L/L}* MEFs at *Hoxa10* (Figure 3.11B), but CAP-D3 localization at *Hoxa10* is lost in the *Rb1^{L/L}* genetic background (Figure 3.11C). This suggests defective condensin II recruitment by pRB can impact long-range chromosome interactions. Furthermore, a previous 4C-Seq report demonstrated that *Hoxa10*, used as a bait, interacted with a region of DNA that contains the *Usp5/Cdca3* bidirectional promoter in MEFs (Sridharan et al., 2009), and this supports the validity of this chromosome contact (Figure 3.11D).

The genome wide representation of locations that interact with each respective bait in each genotype reveals that most long-range contacts are retained in *Rb1^{L/L}* cells and that some contacts were lost and new contacts appeared in *Rb1^{L/L}* mutant cells compared to wild type (Figure 3.10B). However, when interactions that were lost in

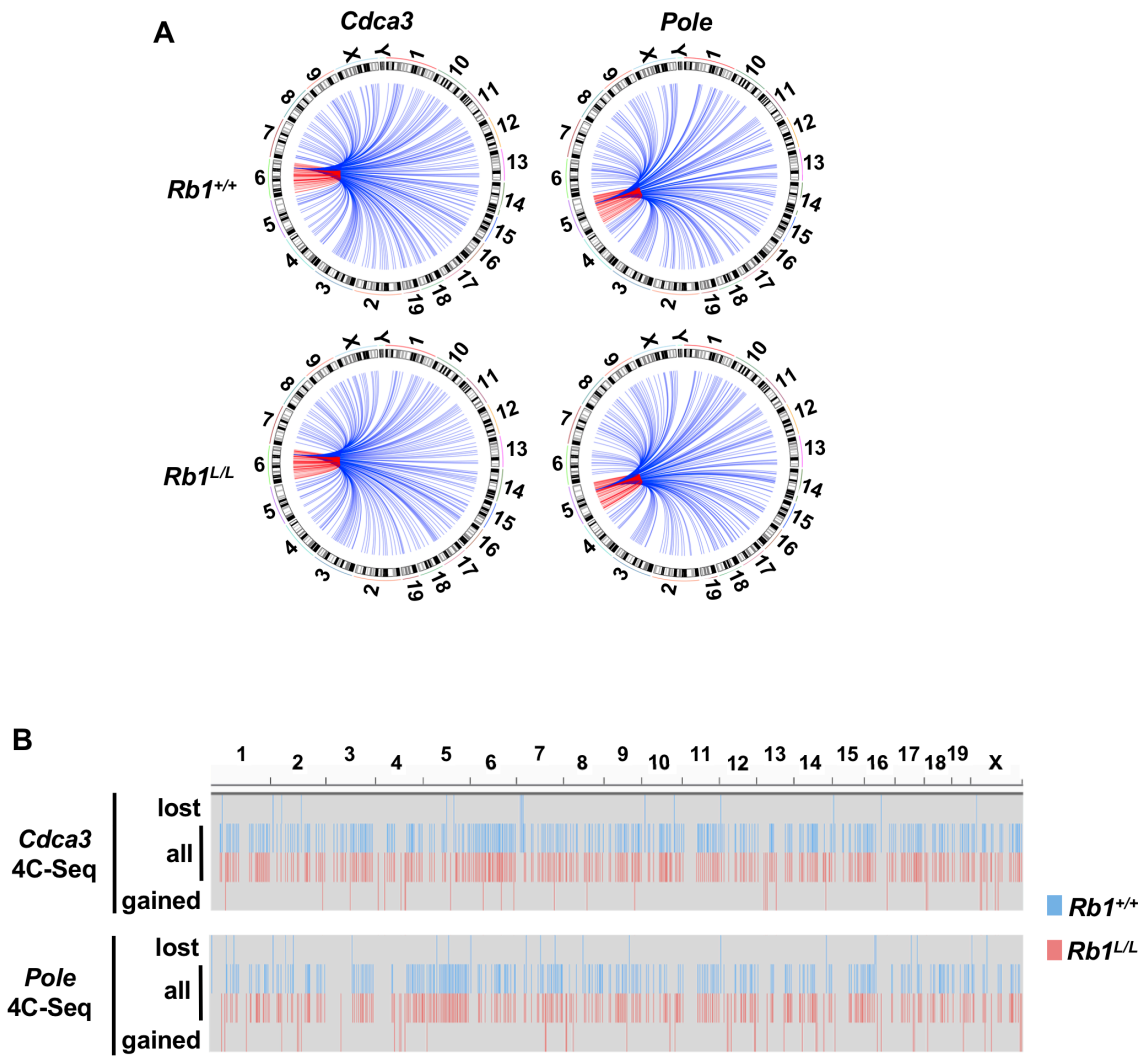


Figure 3.10: Genome wide significant interacting regions determined for *Cdca3/Usps* and *Pole/Pxmp2* bidirectional promoters.

(A) 4C-Seq was used to determine chromosomal regions that interact with the *Cdca3/Usps* and *Pole/Pxmp2* promoters. Circos plots showing all significant interacting regions for one of the four MEF pairs analyzed from both 4C-Seq viewpoints. Red indicates intrachromosomal interactions while blue shows all interchromosomal interactions. (B) Statistically significant interacting regions were determined using w4CSeq for each of four different wild type and *Rb1*^{L/L} replicates. Regions that were identified in more than one replicate are shown as vertical red or blue bars and chromosome position is shown across the top. The complete set of regions identified in *Rb1*^{+/+} and in *Rb1*^{L/L} cells are labeled as ‘all’, while ‘lost’ regions represent contacts identified in wild type MEFs but absent in *Rb1*^{L/L} cells and ‘gained’ regions represent novel contacts found only in *Rb1*^{L/L} cells from that viewpoint.

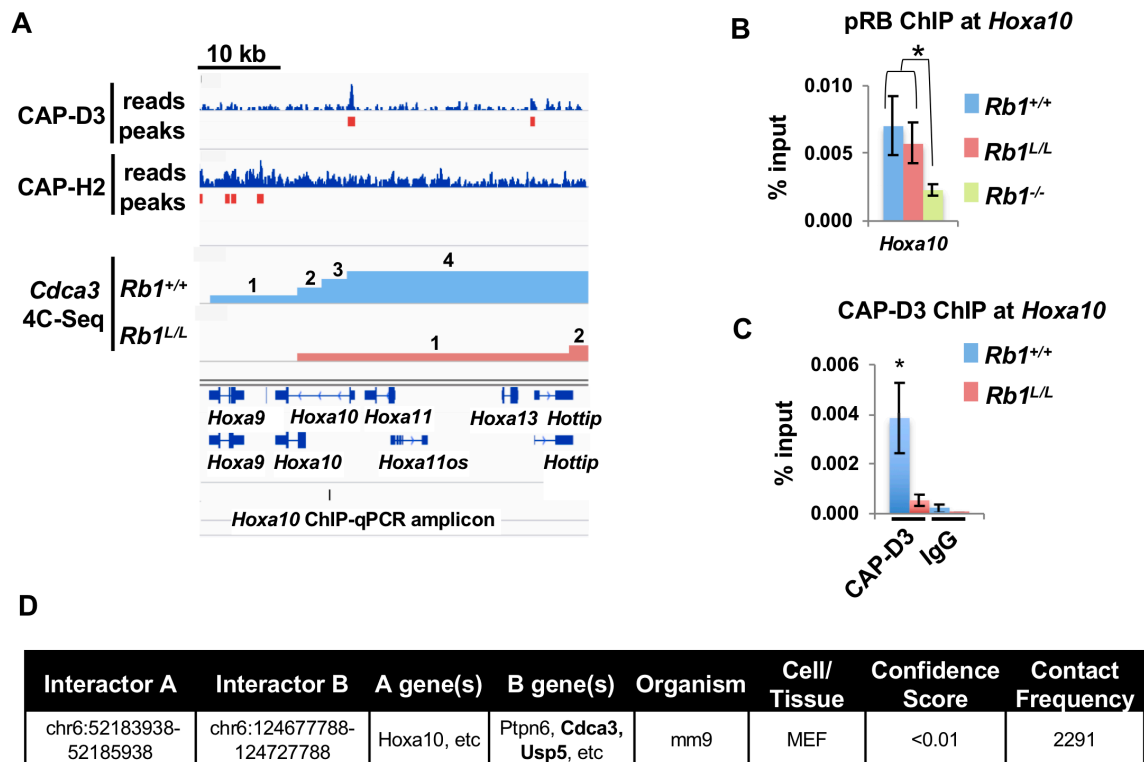


Figure 3.11: Changes in pRB-condensin II complex binding correlated with altered interaction between the *Cdca3/Usf5* bidirectional promoter and the *Hoxa* locus.

(A) Genome browser view of the *Hoxa* locus on chromosome 6 showing CAP-D3 and CAP-H2 ChIP-Seq reads and peaks from EGS and formaldehyde fixed wild type MEFs displayed at the top (with read build-ups scale set to 20). 4C-Seq significant interacting regions from individual replicates of each genotype from the *Cdca3* 4C-Seq experiments are shown. Numbers above the bars indicate the number of replicates that had significant interaction with each region of the *Hoxa* locus. (B) pRB ChIP-qPCR results from asynchronously growing *Rb1*^{+/+}, *Rb1*^{L/L} and *Rb1*^{-/-} MEFs at the indicated location in *Hoxa10* (n=4). (C) CAP-D3 ChIP-qPCR results from *Rb1*^{+/+} and *Rb1*^{L/L} MEFs at *Hoxa10* (n=4). All error bars are ± 1 SEM. * $P < 0.05$; determined by *t*-test. (D) 4C-Seq data from a study by Denholtz *et al.* indicating that the *Hoxa10* region on chromosome 6 interacts with a region on chromosome 6 containing the *Cdca3/Usf5* bidirectional promoter in MEFs (Denholtz *et al.*, 2013). Contact frequency is the normalized count of sequencing reads from each region seen to interact with the *Hoxa10* locus. Data were obtained from 4DGenome (Teng *et al.*, 2016).

Rb1^{L/L} cells were compared with the interactions that were gained, we discovered that the new regions that interact with the *Cdca3/Usp5* viewpoint in *Rb1^{L/L}* cells were more variable in size, while new interactions with *Pole/Pxmp2* promoter were larger in *Rb1^{L/L}* cells (Figure 3.12A). From the *Cdca3/Usp5* viewpoint, 15 interactions were found in the wild type but not *Rb1^{L/L}* cells, and this was increased to 22 interactions at the *Pole/Pxmp2* viewpoint, averaging to approximately 18 interactions that are lost at each region in *Rb1^{L/L}* compared to *Rb1^{+/+}* cells (Figure 3.12B). One of these locations, which is shared between the two different viewpoints, is on chromosome 12 (Figure 3.12C). Further downstream of this location, though, chromosome contacts with each of the viewpoints between genotypes are consistent. Despite loss of wild type chromosome contacts, there were approximately 30 interactions that are gained with each region in *Rb1^{L/L}* compared to *Rb1^{+/+}* cells (25 from the *Cdca3/Usp5* viewpoint and 36 from the *Pole/Pxmp2* viewpoint; Figure 3.12D). One interesting location where we see novel interactions in *Rb1^{L/L}* MEFs, from both viewpoints, is on chromosome 4, which again is close in proximity to undisturbed chromosome contacts between genotypes (Figure 3.12E). Taken together, this data suggests subtle alterations in chromosome conformation and contacts in *Rb1^{L/L}* MEFs.

Overall, the chromosome interactions are similar between the two genotypes at both bidirectional promoter viewpoints in wild type and *Rb1^{L/L}* cells. However, chromatin conformations in *Rb1^{L/L}* MEFs appear more dynamic, as seen by more variable and larger unique interacting regions, and with more interactions gained than lost. Furthermore, our analysis recapitulates a previously observed long-range interaction between *Hoxa10* and the *Usp5/Cdca3* promoter region and demonstrates that it is dependent on pRB-condensin II recruitment to these locations. This confirms a role for pRB and condensin II in connecting insulated bidirectional promoters with distant genomic contacts.

3.4.4 Misregulated gene expression at bidirectional promoters in *RB1*-deleted lung adenocarcinoma

To determine if bidirectional promoters could also be regulated by the pRB-condensin II complex in human cells and if this could be related to disease, we utilized the lung adenocarcinoma TCGA PanCancer Atlas dataset on cBioPortal. At the

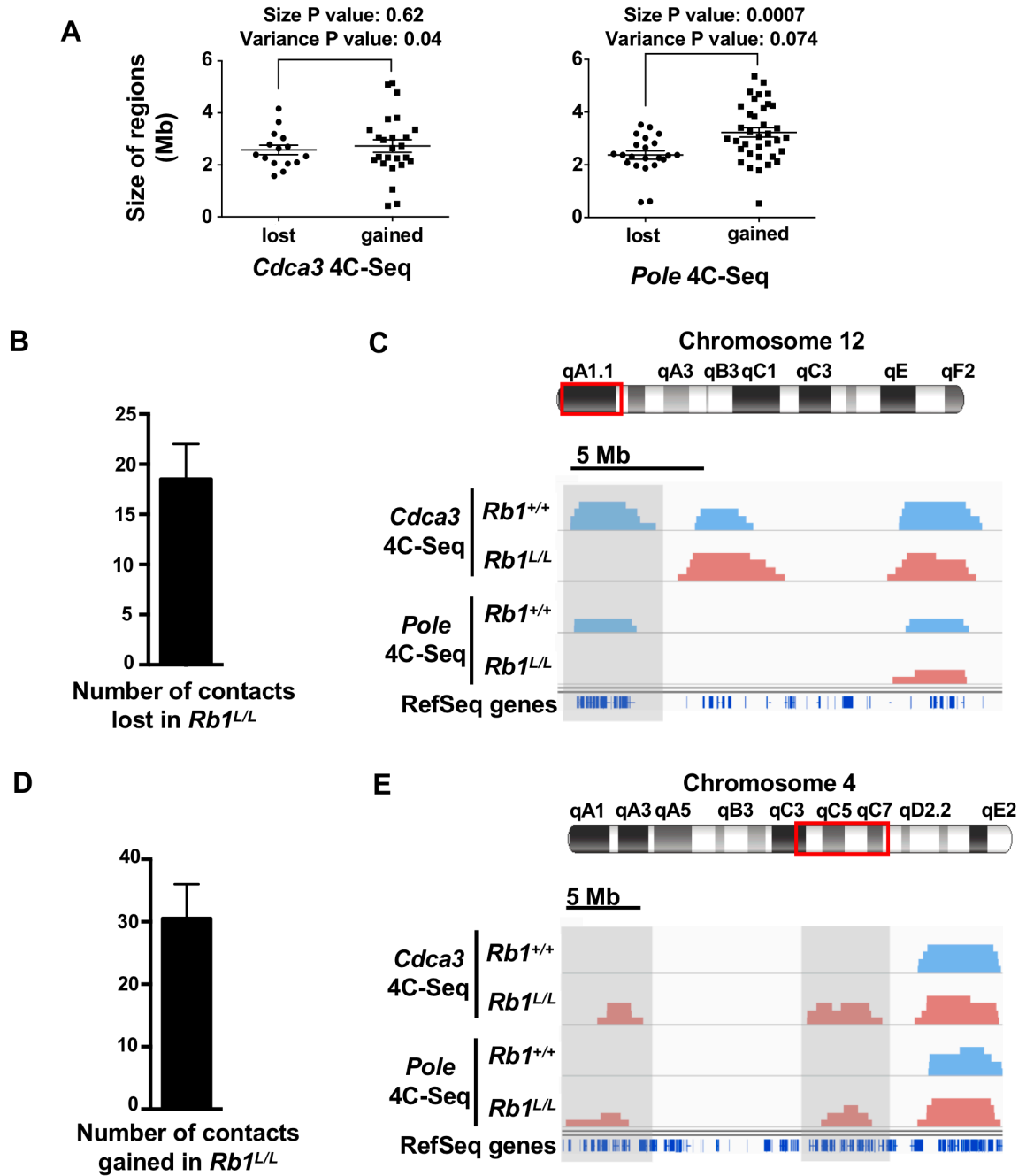


Figure 3.12: Altered long-range chromosome contacts in *Rb1^{L/L}* MEFs.

Figure 3.12: Altered long-range chromosome contacts in *Rb1^{L/L}* MEFs.

(A) The overall size of the significantly interacting regions that were either lost or gained in *Rb1^{L/L}* MEFs compared to wild type with the *Cdca3/Usp5* and *Pole/Pxmp2* viewpoints, from Figure 3.10B. Each data point represents an interacting region and mean sizes and the variance in sizes of these regions were compared between the indicated genotypes using a *t*-test and an F test, respectively. (B) The number of 4C-Seq significant interacting regions, from A, found in wild type MEFs and lost in *Rb1^{L/L}* MEFs at the *Cdca3* or the *Pole* viewpoints were averaged. (C) Genome browser view of an area of chromosome 12 containing a 4C-Seq significant interacting region with both the *Cdca3/Usp5* and the *Pole/Pxmp2* viewpoints in *Rb1^{+/+}* MEFs, but not in *Rb1^{L/L}* MEFs, highlighted in grey. Also shown is a region where only the *Cdca3/Usp5* viewpoint makes significant contact, and another region where both viewpoints significantly interact, regardless of genotype. The height of the 4C-Seq bars indicates the number of replicates that had significant interaction, and the red box on the ideogram illustrates the region shown. (D) The number of 4C-Seq significant interacting regions, from A, found in *Rb1^{L/L}* MEFs, but not in *Rb1^{+/+}* MEFs, at the *Cdca3* or the *Pole* viewpoints were averaged. (E) Two different 4C-Seq significant interacting regions found on chromosome 4 in *Rb1^{L/L}* MEFs, but not wild type MEFs, from both the *Cdca3/Usp5* and the *Pole/Pxmp2* viewpoints, as highlighted in grey. Same data display as in C. All error bars are ± 1 SEM.

POLE/PXMP2 bidirectional promoter, there is a much more significant increase in the expression of *POLE* than *PXMP2* in tumors that have *RBI* deletion compared to those that are diploid for *RBI* (Figure 3.13A). Likewise, at the *CDCA3/USP5* bidirectional promoter, there is only a significant increase in *USP5* expression in tumors that have *RBI* deletion (Figure 3.13A). These results parallel the increase in expression of *Pole* and *Cdca3* that was seen in and *Rb1^{L/L}* and *Rb1^{-/-}* cells compared to wild type (Figure 3.6B). To determine if this change in expression between *RBI* deleted and diploid tumors occurs at other bidirectional promoters where condensin II binding was observed in MEFs, genes that contained an *Rb1^{+/+}* CAP-D3 peak within their bidirectional promoter region and were the higher expressing of the two genes from each bidirectional promoter in mouse *Rb1^{L/L}* compared to *Rb1^{+/+}* cells (“Gene A” from Figure 3.5C) were analyzed within the same lung adenocarcinoma dataset. In total, approximately one quarter of these genes (39 of the 159 genes with expression data in these tumors) had increased expression when *RBI* is deleted compared to when both copies are retained (Figure 3.13B), indicating loss of pRB-condensin II insulating effects are evident in cancer cell gene expression patterns.

3.4.5 Bidirectional promoters insulate epigenetic effects on gene regulation

We sought to test pRB-condensin II complexes as insulators through functional experiments that explored whether pRB-condensin II could maintain divergent transcriptional regulation in the presence of perturbed local heterochromatin. We carried out H3K27me3 ChIP-Seq on *Rb1^{+/+}* and *Rb1^{L/L}* MEFs and noted that at both bidirectional promoters, there are regions of high H3K27me3 that are found on the 3’ side of the CAP-D3 peak in these genome views (Figure 3.14A). Interestingly, this is also the side of each bidirectional promoter where there is increased expression (*Cdca3* and *Pole*) in *Rb1^{L/L}* MEFs. However, we did not see changes in localization of this mark surrounding bidirectional promoters as a result of altered condensin II binding (Figure 3.14A&B). To confirm that localization of H3K27me3 was not changing genome wide, H3K27me3 ChIP-Seq alignments were compared using the Pearson correlation method, which showed extremely high similarity between alignments, regardless of genotype (Figure 3.14C). We treated wild type MEFs with modest concentrations of an EZH2 inhibitor (EZH2i; GSK343) that were insufficient to diminish global H3K27me3 levels and instead

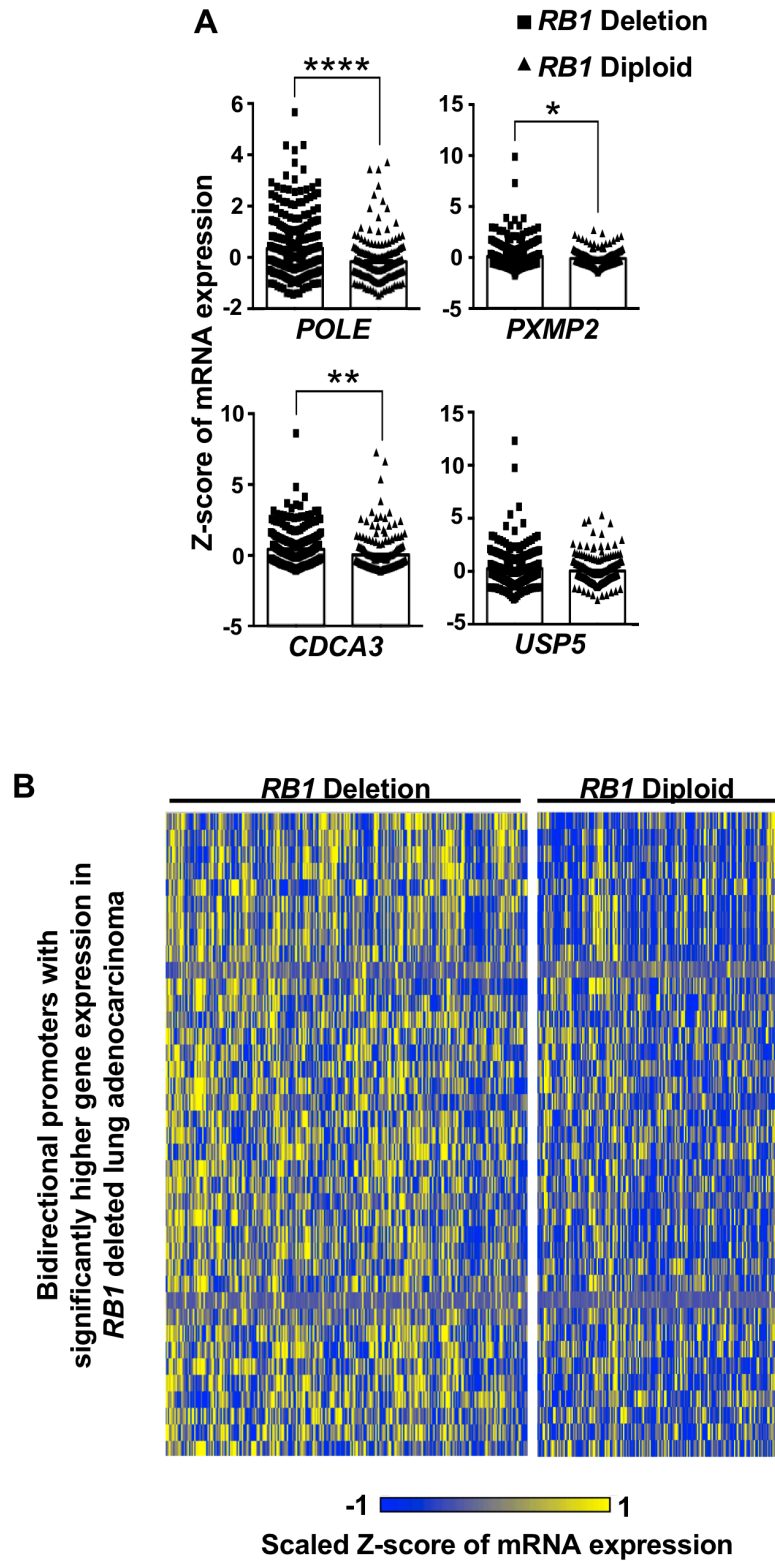


Figure 3.13: *RB1* deletion in lung adenocarcinoma is associated with increased expression of genes at bidirectional promoters.

Figure 3.13: *RB1* deletion in lung adenocarcinoma is associated with increased expression of genes at bidirectional promoters.

(A) Comparison of mRNA expression Z-scores from cBioPortal's TCGA PanCancer Atlas lung adenocarcinoma samples based on *RB1* copy number. Z-scores from samples with *RB1* "deep deletion" and "shallow deletion" were combined for the *RB1* deletion category, and were compared to Z-scores from the *RB1* diploid population for each gene of interest using *t*-tests. * $P < 0.05$, ** $P < 0.005$, **** $P < 0.0001$. **(B)** Genes that contained an *Rb1*^{+/+} CAP-D3 peak within their bidirectional promoter region that is ≤ 5000 bp and was the higher expressing of the two genes from each bidirectional promoter in *Rb1*^{L/L} compared to *Rb1*^{+/+} MEFs ("Gene A" from Figure 3.5C) were analyzed in cBioPortal's TCGA PanCancer Atlas lung adenocarcinoma samples. mRNA expression Z-scores were compared as in A. Only those genes that had significantly higher gene expression in *RB1* deleted tumors (the left half of the heatmap) compared to *RB1* diploid tumors (on the right half of the heatmap) are displayed. Each row was normalized to have a mean of 0 and a variance of 1.

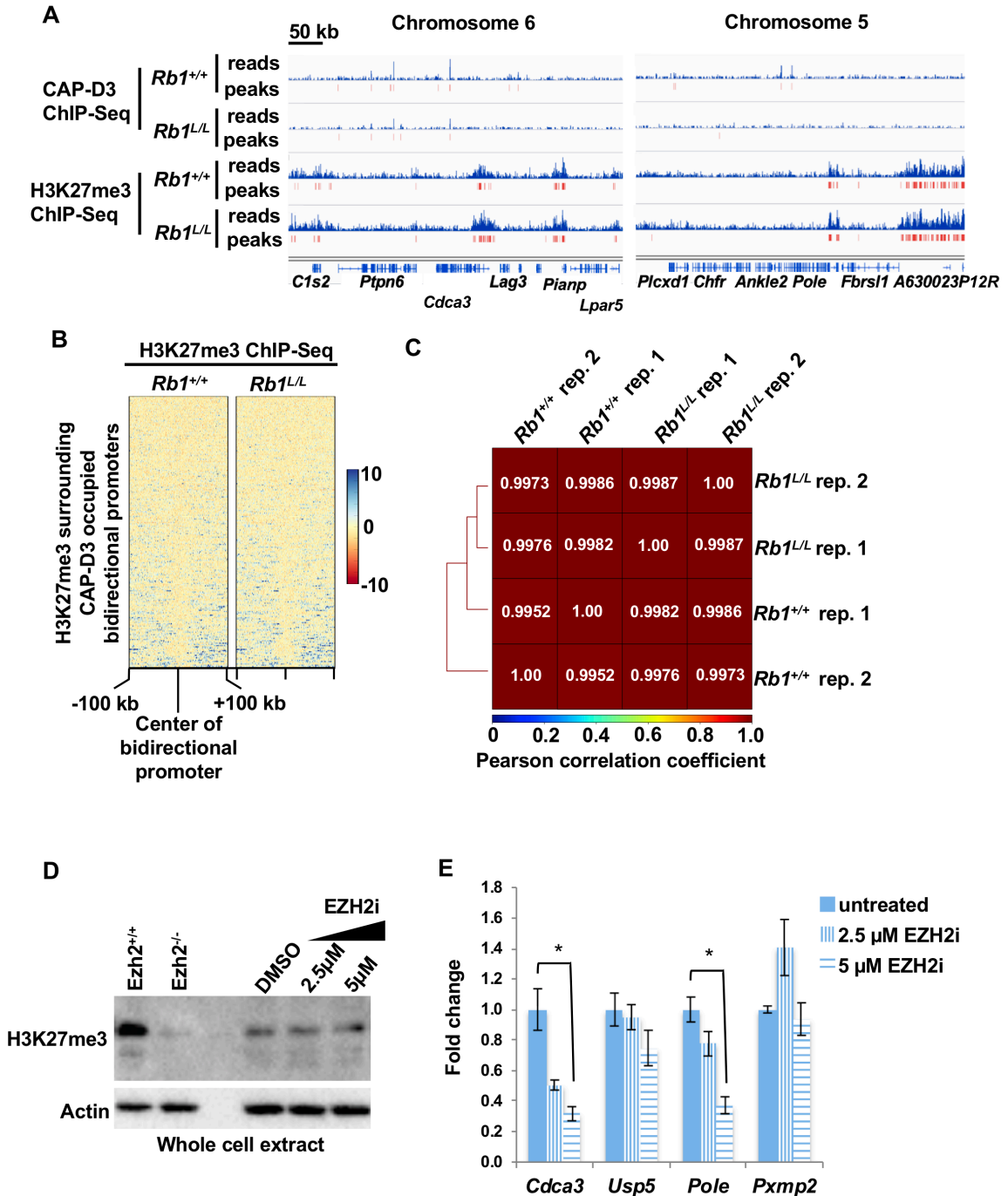


Figure 3.14: H3K27me3 localization is not altered as a result of changes in condensin II binding but pRB-condensin II complexes do insulate transcriptional environments at bidirectional promoters.

Figure 3.14: H3K27me3 localization is not altered as a result of changes in condensin II binding but pRB-condensin II complexes do insulate transcriptional environments at bidirectional promoters.

(A) Genome browser view of 500 kb regions surrounding the *Cdca3/Usp5* and *Pole/Pxmp2* bidirectional promoter locations with ChIP-Seq tracks for CAP-D3 and H3K27me3 in these regions. (B) Heatmaps of H3K27me3 ChIP-Seq read abundance surrounding the ≤ 5000 bp bidirectional promoters that contain an *Rb1*^{+/+} CAP-D3 peak. The heatmaps are centered on the bidirectional promoters and are showing 100 kb upstream and downstream of these regions. (C) Clustered heatmap of Pearson correlations for H3K27me3 ChIP-Seq alignments for every genomic region. The two different ChIP-Seq replicates (denoted as “rep.” in the figure) for each of the genotypes were all compared with each other. (D) Wild type MEFs were treated every 24 hours with the indicated concentrations of GSK343, an EZH2 inhibitor (EZH2i), or DMSO vehicle, for 48 hours total. Extracts were blotted for H3K27me3 and actin. Extract from *Ezh2*^{-/-} MEFs serve as a control for antibody specificity. (E) MEFs treated as in D had RNA isolated to determine the relative expression at *Cdca3/Usp5* and *Pole/Pxmp2* bidirectional promoters. All error bars are ± 1 SEM. * $P < 0.05$; determined by *t*-test.

perturb the epigenetic environment (Figure 3.14D). Analysis of gene expression levels by RT-qPCR of *Cdca3/Usp5* and *Pole/Pxmp2* demonstrates that *Cdca3* and *Pole* are repressed, but their bidirectional promoter mate is unchanged (Figure 3.14E). This demonstrates that pRB-condensin II complexes contribute to insulation of transcriptional effects on opposing sides of occupied bidirectional promoters.

3.5 Discussion

In order to study loss of function effects caused by defective condensin II chromatin loading in interphase, we utilized a targeted murine mutant *Rb1^L* where pRB-condensin II interactions are impaired (Coschi et al., 2014). Using ChIP-Seq and chromosome conformation capture approaches we demonstrated that pRB organizes chromosome architectural protein complexes TFIIC and condensin II in proximal promoter regions of genes. In particular, bidirectional promoters are occupied by these proteins and they may mediate long-range chromosome interactions. These data suggest that pRB-TFIIC-condensin II occupancy at these genome locations can act as architectural organizers and insulators that partition expression potential for closely spaced genes (Figure 3.15A). Reduction of TFIIC and condensin II recruitment caused by the *Rb1^L* allele alters long-range contacts mediated from these locations and misregulates one of the gene pairs (Figure 3.15B). Our data also indicates that loss of condensin II recruitment does not alter the boundaries or abundance of repressive chromatin surrounding these regions. This insight into mammalian genome organization and transcriptional control raises new questions.

Studies from mammals and model organisms indicate that condensins localize to TAD boundaries along with TFIIC, CTCF, and cohesin complexes (Crane et al., 2015; Li et al., 2015a; Van Bortle et al., 2014; Yuen et al., 2017). Data from *Drosophila* suggests that TAD boundaries containing condensin II are the strongest insulators between domains (Van Bortle et al., 2014). While this suggests a role in chromosome organization, it also suggests redundancy of function between different components at these locations, or that RNA polymerase movement may simply direct condensin II to these locations without a functional purpose. Our study looks at condensin II function in chromosome topology in isolation, as cohesin and condensin I chromatin loading are

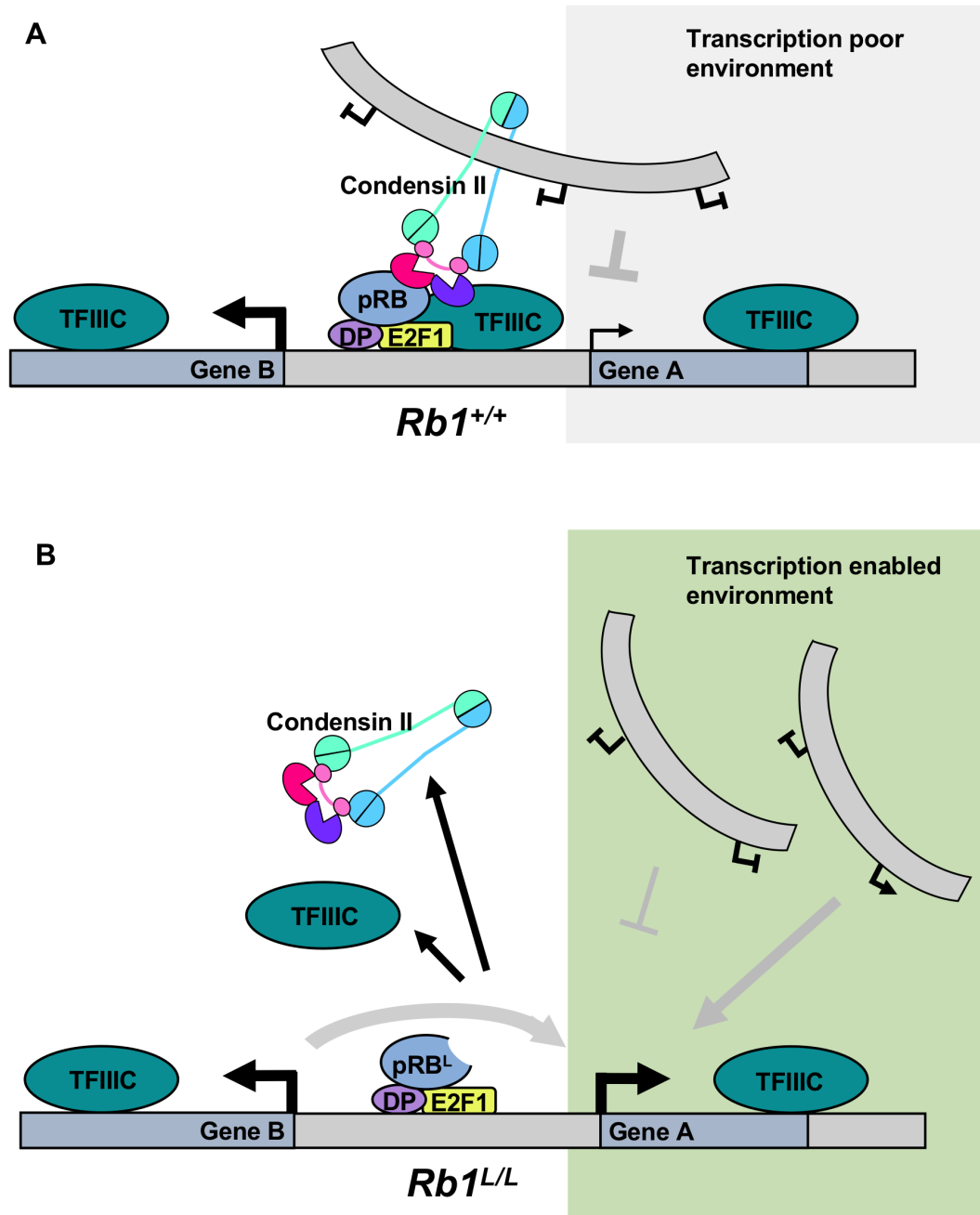


Figure 3.15: pRB-TFIIC-condensin II complexes organize transcriptional environments at bidirectional promoters.

(A) In wild type cells pRB-TFIIC-condensin II is responsible for making long-range chromosome interactions at bidirectional promoters where one of the two genes (Gene A) is sequestered in an environment that reduces its expression. **(B)** In *Rb1^{L/L}* cells pRB is present, but is defective for recruiting TFIIC and condensin II to bidirectional promoters. Loss of this complex reduces contact with certain genomic locations, allowing novel and potentially more transcriptionally active regions to influence the environment around Gene A facilitating higher level expression.

normal in *Rb1^{L/L}* cells (Coschi et al., 2010; Manning et al., 2010). One conclusion from this work is that most distant chromosome contacts involving *Cdca3* and *Pole* are retained in *Rb1^{L/L}* cells compared to *Rb1^{+/+}* in our interaction maps. This supports the concept of redundancy of three-dimensional architecture in genome organization. Our data also reveal that *Rb1^L* causes the consistent loss of chromosome contacts with specific regions seen in wild type cells when TFIIC and condensin II recruitment are diminished. In addition, the overall effect of the *Rb1^L* mutant on *Pole/Pxmp2* promoter contacts creates broader unique interacting chromosome territories, and more variable sizes of unique interacting chromosome territories for *Cdca3/Usp5* promoter contacts. Thus, mammalian condensin II-dependent contributions to three-dimensional chromosome architecture are measurable and their loss has consequences in isolation from cohesin and its regulators.

Our data indicates that loss of pRB recruitment of condensin II and TFIIC has local effects on transcription. The viewpoints we investigated in detail demonstrate that condensin II has little role in organizing short-range loops. For this reason, local effects on transcript levels are best explained by the altered long-range chromosome interactions. We note that our model of condensin II mediated chromosome contacts (Figure 3.15) is consistent with gene transcription territories in nuclear organization that have been proposed in other studies (Denholtz et al., 2013; Imakaev et al., 2012). It is possible that our work reveals subtle reorganization, or loss of these territories. Furthermore, recent studies of chromosome topology changes induced by degradation of CTCF or cohesin components indicate that functional compartments created by some long-range contacts are maintained (Haarhuis et al., 2017; Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017; Wutz et al., 2017). Our data suggests that condensin II may be an unidentified factor that mediates these contacts, although this will need to await further investigation using Hi-C approaches.

The involvement of pRB in TFIIC and condensin II recruitment to gene promoters in interphase may seem surprising. We note that some prior studies of mammalian condensin II have utilized cell lines that are transformed by viral oncogenes that target pRB (Liu et al., 2010; Yuen et al., 2017). pRB's role in organizing long-range

contacts would likely be disrupted in these cells, or would rely on compensatory mechanisms to load condensin II in pRB's absence. A number of previous studies have demonstrated that pRB interacts with condensin II and this interaction is conserved in *Drosophila* (Coschi et al., 2010; Longworth et al., 2008). Furthermore, TFIIC has also previously been shown to bind to pRB (Chu et al., 1997), indicating that these protein complexes have been known to function together, but their role in long-range chromosome interactions and transcriptional control at divergent promoters was unknown.

Cancer genomic studies have revealed that *RBI* deletions often occur late in cancer progression and are preferentially enriched in metastatic disease (Robinson et al., 2017), or in acquired resistance to targeted therapeutic agents (Dick et al., 2018). These are scenarios where deregulated proliferation has already been established, and suggest pRB loss has additional contributions to cancer biology. Therefore, alterations to interphase chromosome structure and gene expression, mediated by pRB disruption in cancer, may be linked to misregulation of condensin II. Consistent with this, when we analyzed genes from bidirectional promoters where condensin II was seen to bind in our mouse model, we found that compared to *RBI*^{+/+} cancers, those with *RBI* deletion frequently upregulate these genes. This study demonstrates that organization of long-range chromosome interactions and divergent promoter insulation by condensin II is an unexpected, but obligatorily common target for disruption during cancer progression because of its dependence on pRB.

3.6 References

- Abdennur, N., Schwarzer, W., Pekowska, A., Shaltiel, I.A., Huber, W., Haering, C.H., Mirny, L., and Spitz, F. (2018). Condensin II inactivation in interphase does not affect chromatin folding or gene expression. bioArxiv <https://doi.org/10.1101/437459>.
- Avni, D., Yang, H., Martelli, F., Hofmann, F., ElShamy, W.M., Ganesan, S., Scully, R., and Livingston, D.M. (2003). Active localization of the retinoblastoma protein in chromatin and its response to S phase DNA damage. *Mol Cell* 12, 735-746.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837.

Blais, A., van Oevelen, C.J., Margueron, R., Acosta-Alvear, D., and Dynlacht, B.D. (2007). Retinoblastoma tumor suppressor protein-dependent methylation of histone H3 lysine 27 is associated with irreversible cell cycle exit. *J Cell Biol* 179, 1399-1412.

Bulut-Karslioglu, A., De La Rosa-Velazquez, I.A., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galan, C., Winter, G.E., Engist, B., Gerle, B., *et al.* (2014). Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol Cell* 55, 277-290.

Cai, M., Gao, F., Lu, W., and Wang, K. (2016). w4CSeq: software and web application to analyze 4C-seq data. *Bioinformatics* 32, 3333-3335.

Cecchini, M.J., Amiri, M., and Dick, F.A. (2012). Analysis of cell cycle position in mammalian cells. *J Vis Exp* 59, e3491.

Cecchini, M.J., Thwaites, M., Talluri, S., Macdonald, J.I., Passos, D.T., Chong, J.L., Cantalupo, P., Stafford, P., Saenz-Robles, M.T., Francis, S.M., *et al.* (2014). A retinoblastoma allele that is mutated at its common E2F interaction site inhibits cell proliferation in gene targeted mice. *Molecular and cellular biology* 34, 2029-2045.

Chu, W.M., Wang, Z., Roeder, R.G., and Schmid, C.W. (1997). RNA polymerase III transcription repressed by Rb through its interactions with TFIIB and TFIIC2. *J Biol Chem* 272, 14755-14761.

Coschi, C.H., Ishak, C.A., Gallo, D., Marshall, A., Talluri, S., Wang, J., Cecchini, M.J., Martens, A.L., Percy, V., Welch, I., *et al.* (2014). Haploinsufficiency of an RB-E2F1-Condensin II complex leads to aberrant replication and aneuploidy. *Cancer Discov* 4, 840-853.

Coschi, C.H., Martens, A.L., Ritchie, K., Francis, S.M., Chakrabarti, S., Berube, N.G., and Dick, F.A. (2010). Mitotic chromosome condensation mediated by the retinoblastoma protein is tumor-suppressive. *Genes Dev* 24, 1351-1363.

Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* 523, 240-244.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., *et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* 107, 21931-21936.

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M., and Plath, K. (2013). Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* 13, 602-616.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology* 4, P3.

Dick, F.A., Goodrich, D.W., Sage, J., and Dyson, N.J. (2018). Non-canonical functions of the RB protein in cancer. *Nat Rev Cancer* 18, 442-451.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21.

Fraser, J., Williamson, I., Bickmore, W.A., and Dostie, J. (2015). An Overview of Genome Organization and How We Got There: from FISH to Hi-C. *Microbiol Mol Biol Rev* 79, 347-372.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., *et al.* (2018). A pathway for mitotic chromosome formation. *Science* 359.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B., *et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* 169, 693-707 e614.

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

Herrera, R.E., Chen, F., and Weinberg, R.A. (1996). Increased histone H1 phosphorylation and relaxed chromatin structure in Rb-deficient fibroblasts. *Proc Natl Acad Sci USA* 93, 11510-11515.

Hirano, T. (2016). Condensin-Based Chromosome Organization from Bacteria to Vertebrates. *Cell* 164, 847-857.

Hocquet, C., Robellet, X., Modolo, L., Sun, X.M., Burny, C., Cuylen-Haering, S., Toselli, E., Clauder-Munster, S., Steinmetz, L., Haering, C.H., *et al.* (2018). Condensin controls cellular RNA levels through the accurate segregation of chromosomes instead of directly regulating transcription. *Elife* 7.

Hoffman, E.A., Frey, B.L., Smith, L.M., and Auble, D.T. (2015). Formaldehyde crosslinking: a tool for the study of chromatin complexes. *J Biol Chem* 290, 26404-26411.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods* 9, 999-1003.

- Isaac, C.E., Francis, S.M., Martens, A.L., Julian, L.M., Seifried, L.A., Erdmann, N., Binne, U.K., Harrington, L., Sicinski, P., Berube, N.G., *et al.* (2006). The retinoblastoma protein regulates pericentric heterochromatin. *Mol Cell Biol* 26, 3659-3671.
- Ishak, C.A., Coschi, C.H., Roes, M.V., and Dick, F.A. (2017). Disruption of CDK-resistant chromatin association by pRB causes DNA damage, mitotic errors, and reduces Condensin II recruitment. *Cell Cycle* 16, 1430-1439.
- Ishak, C.A., Marshall, A.E., Passos, D.T., White, C.R., Kim, S.J., Cecchini, M.J., Ferwati, S., MacDonald, W.A., Howlett, C.J., Welch, I.D., *et al.* (2016). An RB-EZH2 Complex Mediates Silencing of Repetitive DNA Sequences. *Mol Cell* 64, 1074-1087.
- Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* 41, 376-381.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359.
- Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.T., Cubenas-Potts, C., Hu, M., Lei, E.P., Bosco, G., *et al.* (2015a). Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* 58, 216-231.
- Li, W., Hu, Y., Oh, S., Ma, Q., Merkurjev, D., Song, X., Zhou, X., Liu, Z., Tanasa, B., He, X., *et al.* (2015b). Condensin I and II Complexes License Full Estrogen Receptor alpha-Dependent Enhancer Activation. *Mol Cell* 59, 188-202.
- Liu, W., Tanasa, B., Tyurina, O.V., Zhou, T.Y., Gassmann, R., Liu, W.T., Ohgi, K.A., Benner, C., Garcia-Bassets, I., Aggarwal, A.K., *et al.* (2010). PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. *Nature* 466, 508-512.
- Longworth, M.S., Herr, A., Ji, J.Y., and Dyson, N.J. (2008). RBF1 promotes chromatin condensation through a conserved interaction with the Condensin II protein dCAP-D3. *Genes Dev* 22, 1011-1024.
- Manning, A.L., Longworth, M.S., and Dyson, N.J. (2010). Loss of pRB causes centromere dysfunction and chromosomal instability. *Genes & development* 24, 1364-1376.
- Mendoza-Maldonado, R., Paolinelli, R., Galbiati, L., Giadrossi, S., and Giacca, M. (2010). Interaction of the retinoblastoma protein with Orc1 and its recruitment to human origins of DNA replication. *PLoS One* 5, e13720.
- Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell* 128, 787-800.

Morales, C., and Losada, A. (2018). Establishing and dissolving cohesion during the vertebrate cell cycle. *Curr Opin Cell Biol* 52, 51-57.

Mouse Genome Sequencing, C., Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., *et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520-562.

Narita, M., Nunez, S., Heard, E., Narita, M., Lin, A.W., Hearn, S.A., Spector, D.L., Hannon, G.J., and Lowe, S.W. (2003). Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* 113, 703-716.

Nishide, K., and Hirano, T. (2014). Overlapping and non-overlapping functions of condensins I and II in neural stem cell divisions. *PLoS genetics* 10, e1004847.

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* 169, 930-944 e922.

Ono, T., Losada, A., Hirano, M., Myers, M.P., Neuwald, A.F., and Hirano, T. (2003). Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell* 115, 109-121.

Ono, T., Yamashita, D., and Hirano, T. (2013). Condensin II initiates sister chromatid resolution during S phase. *The Journal of cell biology* 200, 429-441.

Pavlidis, P., and Noble, W.S. (2003). Matrix2png: a utility for visualizing matrix data. *Bioinformatics* 19, 295-296.

Pedersen, M.T., Agger, K., Laugesen, A., Johansen, J.V., Cloos, P.A., Christensen, J., and Helin, K. (2014). The demethylase JMJD2C localizes to H3K4me3-positive transcription start sites and is dispensable for embryonic development. *Mol Cell Biol* 34, 1031-1045.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33, 290-295.

Quinlan, A.R. (2014). BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr Protoc Bioinformatics* 47, 11 12 11-34.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841-842.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic acids research* 42, W187-191.

- Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* *44*, W160-165.
- Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., *et al.* (2017). Cohesin Loss Eliminates All Loop Domains. *Cell* *171*, 305-320 e324.
- Robinson, D.R., Wu, Y.M., Lonigro, R.J., Vats, P., Cobain, E., Everett, J., Cao, X., Rabban, E., Kumar-Sinha, C., Raymond, V., *et al.* (2017). Integrative clinical genomics of metastatic cancer. *Nature* *548*, 297-303.
- Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat Biotechnol* *29*, 24-26.
- Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. *Nat Rev Genet* *19*, 789-800.
- Rowley, M.J., Lyu, X., Rana, V., Ando-Kuri, M., Karns, R., Bosco, G., and Corces, V.G. (2019). Condensin II Counteracts Cohesin and RNA Polymerase II in the Establishment of 3D Chromatin Organization. *Cell Rep* *26*, 2890-2903 e2893.
- Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., C, H.H., Mirny, L., *et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature* *551*, 51-56.
- Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenko, V.V., *et al.* (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature* *488*, 116-120.
- Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. (2009). CEAS: cis-regulatory element annotation system. *Bioinformatics* *25*, 2605-2606.
- Splinter, E., de Wit, E., van de Werken, H.J., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods* *58*, 221-230.
- Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* *20*, 2349-2354.
- Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. *Cell* *136*, 364-377.
- Stadhouders, R., Fillion, G.J., and Graf, T. (2019). Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* *569*, 345-354.

- Talluri, S., Isaac, C.E., Ahmad, M., Henley, S.A., Francis, S.M., Martens, A.L., Bremner, R., and Dick, F.A. (2010). A G1 checkpoint mediated by the retinoblastoma protein that is dispensable in terminal differentiation but essential for senescence. *Mol Cell Biol* *30*, 948-960.
- Teng, L., He, B., Wang, J., and Tan, K. (2016). 4DGenome: a comprehensive database of chromatin interactions. *Bioinformatics* *32*, 2727.
- Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* *14*, 178-192.
- Thwaites, M.J., Coschi, C.H., Isaac, C.E., and Dick, F.A. (2016). Cell Synchronization of Mouse Embryonic Fibroblasts. *Methods Mol Biol* *1342*, 91-99.
- Tian, B., Yang, J., and Brasier, A.R. (2012). Two-step cross-linking for analysis of protein-chromatin interactions. *Methods Mol Biol* *809*, 105-120.
- Van Bortle, K., Nichols, M.H., Li, L., Ong, C.T., Takenaka, N., Qin, Z.S., and Corces, V.G. (2014). Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome biology* *15*, R82.
- van Ruiten, M.S., and Rowland, B.D. (2018). SMC Complexes: Universal DNA Looping Machines with Distinct Regulators. *Trends in genetics : TIG* *34*, 477-487.
- Wang, S., Sun, H., Ma, J., Zang, C., Wang, C., Wang, J., Tang, Q., Meyer, C.A., Zhang, Y., and Liu, X.S. (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nature protocols* *8*, 2502-2515.
- Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q., *et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* *40*, 897-903.
- Wutz, G., Varnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J., *et al.* (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J* *36*, 3573-3599.
- Yuen, K.C., Slaughter, B.D., and Gerton, J.L. (2017). Condensin II is anchored by TFIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. *Sci Adv* *3*, e1700191.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* *9*, R137.

Chapter 4

4 Discussion

4.1 Summary of findings

My thesis investigates different contributions of pRB to the maintenance of genome stability. The work here suggests *RBI* mutation affects the level of DNA damage through various mechanisms as well as the organization of long-range chromosome contacts through pRB's interaction with condensin II. Overall, this demonstrates that pRB has other previously unappreciated roles within the nucleus which assist in tumor suppression through preserving genome integrity.

My work in chapter 2 identifies increased genome instability in cancer cell lines when *RBI* mutations are induced, despite an already compromised pRB-pathway. Even cells with single copy loss of *RBI* have increased levels of basal DNA damage as well as mitotic errors. Increased quantities of reactive oxygen species as well as defects in homologous recombination appear to be the main contributors to the increase in spontaneous DNA damage. *RBI* mutation also results in a heightened capacity to seed new tumors in the lungs of xenografted immune compromised mice.

In chapter 3, the dependency of condensin II on pRB for its localization throughout the genome is investigated. This study demonstrates that pRB is a recruitment factor for both condensin II and TFIIC at many promoters throughout the genome. The complex of pRB-condensin II-TFIIC is further enriched at bidirectional promoters, where diminished localization of these proteins leads to reciprocal misexpression of these divergently paired genes. Chromosome contacts were investigated demonstrating that loss of condensin II binding leads to altered long-range chromatin interactions with between bidirectional promoters and distant loci. Data from lung adenocarcinoma patients demonstrates that a similar mechanism may occur in humans, as some genes at bidirectional promoters have significantly higher expression in patients with *RBI* deletion compared to those that have retained both copies of *RBI*.

Collectively, my work has added to our current model of pRB-mediated tumor suppression, highlighting the fact that its contributions to tumor suppression encapsulate more than simply maintenance of proliferative control through binding to and repressing E2F transcription factors (Dyson, 1998). This initial finding has caused cell cycle independent functions of pRB to be largely overlooked when assessing its contributions to tumor suppression. However, my work demonstrates that loss of *RBI* can also contribute to cancer progression through enhanced DNA damage and altered chromosome topology. My investigation into the various measures pRB employs to support genome stability may elucidate answers to questions regarding pRB status in relation to patient outcome, disease progression, and genome reorganization.

4.2 Impact on cancer therapy selection

My work in cancer cell lines with *RBI* loss in chapter 2 confirms increased vulnerability to cisplatin, agreeing with studies where loss of pRB correlated with improved survival in response to chemotherapy incorporating a platinum-based agent like cisplatin or carboplatin (Cecchini et al., 2015; Garsed et al., 2018; Ludovini et al., 2004). Thus, it appears as though these platinum-based chemotherapies are exploiting the genomic instability in these *RBI*-deficient tumors, which could be what is responsible for the improved outcomes in these patients. Since I have shown that cells with *RBI* loss have impaired homologous recombination, PARP inhibitors (PARPis) might be useful to potentially increase treatment success in patients with low levels of pRB. This is because PARPis exhibit synthetic lethal effects when applied to cells with defective HR (Konstantinopoulos et al., 2015). In some clinical trials of PARPis, in fact, platinum sensitivity has been used as a clinical surrogate of HR deficiency, and thus has been used as an eligibility criterion when selecting patients (examples are (Ledermann et al., 2012; Liu et al., 2014)).

Only recently have PARPis been used in combination with carboplatin, and because of the common *BRCA1* and *BRCA2* mutations in epithelial ovarian cancer (EOC) and breast cancer, this combination therapy has been tested chiefly in these settings and the results seem favourable. For example, in a recent phase 2 trial, the efficacy and tolerability of PARPi olaparib in combination with paclitaxel and carboplatin

chemotherapy, followed by olaparib maintenance monotherapy was assessed versus chemotherapy alone in patients with platinum-sensitive, recurrent, high-grade serous ovarian carcinoma (HGSOC) (Oza et al., 2015). This study found that olaparib plus chemotherapy followed by maintenance monotherapy significantly improved progression-free survival versus chemotherapy alone (Oza et al., 2015). However, despite the fact that an extension in progression free survival was observed in the cohort of all patients, the difference between treatment groups was larger in the subset of patients with *BRCA* mutations, so subsequent phase 3 trials have focused solely on patients with these mutations (Moore et al., 2018; Pujade-Lauraine et al., 2017). Therefore, it would be interesting to determine if some of the other patients that respond well to PARPis in combination with chemotherapy are those that have pRB loss.

4.3 *RB1* hemizygozity in cancer

As previously discussed in the introduction, Dr. Alfred Knudson proposed what became known as the “two hit hypothesis” when studying retinoblastoma (Knudson, 1971). As part of this study, Knudson proposed that loss of one *RB1* allele did not accelerate loss of the second *RB1* allele; in other words, it was thought that hemizygozity of *RB1* did not create haploinsufficiency, but rather that the wild type condition was recapitulated. However, some recent studies, including data from chapter 2 of this thesis, has demonstrated that single copy loss of *RB1* may contribute to cancer progression through multiple mechanisms.

Genomic instability used to be thought of as a by-product of tumorigenesis, however, through our increased knowledge of cancer progression, it is now thought of as an enabling characteristic (Hanahan and Weinberg, 2000, 2011). Data from chapter 2, therefore, demonstrating that single copy loss of *RB1* leads to increased DNA damage compared to otherwise isogenic cancer cells that are diploid for *RB1* suggests that pRB is haploinsufficient in terms of the maintenance of genome stability, and thus, tumor suppression. In line with this, there is evidence of genome instability through an increase in mitotic errors in primary *RB1*^{+/-} cells (Coschi et al., 2014; Gonzalez-Vasconcellos et al., 2013; Zheng et al., 2002), and patients with inherited retinoblastoma are more likely

to acquire second primary neoplasms than the general population (Abramson et al., 1976; Marees et al., 2008).

The most common non-ocular cancers patients with hereditary retinoblastoma develop are bone and soft tissue sarcomas, melanoma and brain tumors (Kleinerman et al., 2012). Of the bone sarcomas reported in retinoblastoma survivors, many are diagnosed between 10 and 20 years of age, and osteosarcoma is the most common type (Kleinerman et al., 2012).

The cancer cell line we used to investigate *RBI* heterozygosity, U2OS, was cultivated from an osteosarcoma patient. Because many of the second primary neoplasms arising in retinoblastoma patients are of mesenchymal cell origin, and I demonstrated increased genomic instability in osteosarcoma cells with single copy loss of *RBI*, it is tempting to consider that mesenchymal cells may have a special requirement for two functional copies of *RBI*. However, we also demonstrated in chapter 2 through investigation of cBioPortal data that a variety of already established cancers have “shallow deletion” of *RBI*. Although it may be true that *RBI* hemizyosity might contribute uniquely to initial cancer onset in mesenchymal cells, it appears as though *RBI* single copy loss is more ubiquitous than that and thus its contributions to genome stability are likely more universal, although this remains to be investigated.

4.4 *Rb1^L* and haploinsufficiency

Not only does complete loss of a single *RBI* allele lead to haploinsufficiency, previous work has also demonstrated a genome instability phenotype in *Rb1^{L/+}* cells (Coschi et al., 2014). It has been shown that pRB and condensin II localize and contribute to the integrity of pericentromeric heterochromatin, and even in *Rb1^{L/+}* cells, localization of condensin II to these regions is decreased (Coschi et al., 2014). Ultimately, this resulted in a similar frequency of mitotic errors as seen in the homozygous mutants, demonstrating haploinsufficiency. Given that a single *Rb1^L* allele leads to reduced recruitment of condensin II at pericentromeric heterochromatin compared to wild type, a remaining question from chapter 3 is if this same genotype can also lead to reduced recruitment of condensin II at promoter regions as well. Comparing the amount of

condensin II recruitment at bidirectional promoters to the long-range chromatin interactions at these sites in *Rb1*^{+/+}, *Rb1*^{L/+} and *Rb1*^{L/L} cells could help reveal the amount of condensin II localization required to maintain appropriate chromosome topology. Interestingly, when expression data from lung adenocarcinoma patients was analyzed at bidirectional promoters in chapter 3, due to the small number of patients with “deep deletion”, consistent with biallelic loss, the *RBI* deleted category also included patients with “shallow deletion”, which is suggestive of heterozygous deletion of this allele. Therefore, since the significantly higher expression of bidirectional promoter genes occurred in patients with both deep and shallow deletion of *RBI* combined compared to patients that were diploid for *RBI*, this suggests that localization of condensin II at bidirectional promoters is gene-dosage dependent as well.

4.5 Chromosome topology and cancer

With the advent of chromosome capture technology, we are beginning to appreciate that cell-fate decisions are driven by changes in environmental cues, activating signal transduction into the nucleus, which ultimately converge in the activation or silencing of DNA sequence-specific regulators, commonly transcription factors, and affect the recruitment of transcriptional and chromatin remodeling machinery (Stadhouders et al., 2019). The formation and continuation of cell-type-specific gene expression programs are therefore a consequence of the interaction between transcription factors and the chromatin landscape that they interact with. This means that cell identity can be regarded as an “emergent property”, meaning it results from the interaction between components at various levels of organization, namely between transcription factors, chromatin-associated proteins, epigenetic modifications, and the three-dimensional organization of the genome. Accordingly, genome conformation is partly cell-type specific; the position of chromosome territories and regions of intermingling between them has been seen to vary between cell types and evidence is suggesting that 10-40% of TAD boundaries are cell-type-specific and that the strength of boundary insulation is dynamic (Stadhouders et al., 2019).

One way in which chromosome conformation has been seen to be disrupted in cancer is through chromosomal rearrangements and mutations that disrupt TAD

boundaries, which can lead to incorrect insulated “neighborhoods”, or DNA loops, and ultimately incorrect promoter-enhancer connections. Recently, insulated neighborhoods in T-cell acute lymphoblastic leukemia (T-ALL) were mapped and frequent microdeletions were discovered that abolish boundary sites of insulated neighborhoods containing well known T-ALL proto-oncogenes (Hnisz et al., 2016). Through database mining, an enrichment of mutations at CTCF boundaries of constitutive neighborhoods in the cancer genomes of 31 different tumor types were also found, showing that somatic mutations of insulated neighborhood boundaries occur in the genomes of a variety of cancers (Hnisz et al., 2016). Another study found that the promoter for a long non-coding RNA gene, *PVT1*, competes with the *MYC* promoter for contact with a nearby enhancer element in cis and that mutations encompassing the *PVT1* promoter region are recurrent in human cancers (Cho et al., 2018). The promoter of the *PVT1* gene was termed a “tumor-suppressive DNA element” since it limits *MYC* oncogene expression by acting as a DNA boundary element between *MYC* and downstream enhancers (Cho et al., 2018). Additionally, cancer cells can also activate novel enhancers to influence the transcription profile. For example, the activation of super enhancers in breast cancer, T-ALL, and diffuse large B-cell lymphoma have been seen to increase the amount of CD47, a cell surface molecule that inhibits phagocytosis of cells that express it (Betancur et al., 2017). Finally, a subset of gliomas have gain-of-function *IDH* mutations as initiating events, which leads to hypermethylation throughout the tumor genome. This hypermethylation overlaps some CTCF sites and was seen to disrupt its binding, leading to changes in chromosome structure and altered oncogene expression (Flavahan et al., 2016).

Altogether, the previous examples demonstrate that changes in conformation have been associated with cancer cell types. However, changes in chromosome topology in all of these examples are linked to mutations or changes in DNA elements, and not from alterations in the actual architectural proteins responsible for maintaining proper chromosome topology in cells. Although there are some examples of mutations specifically affecting proteins previously implicated in interphase chromosome topology like CTCF and components of cohesin, overall, these are not mutations that are commonly associated with cancer (Hnisz et al., 2018).

My work in chapter 3 has revealed that loss of pRB can systemically alter chromosome topology through loss of the architectural functions of the pRB-condensin II complex, and this is likely happening on a routine basis in cancers when *RBI* mutations occur. When pRB is inactivated in cancer, reorganization or loss of chromosome territories at closely spaced bidirectional promoters is expected to occur, as evidenced by increased expression of particular bidirectional promoter genes in lung adenocarcinomas with *RBI* deletion. As previously mentioned, cancer genomic studies reveal that *RBI* deletions often appear late in cancer progression and are enriched in metastatic disease (Robinson et al., 2017), or in acquired resistance to targeted therapeutic agents (Dick et al., 2018). From my work, it can be inferred that one reason this may be happening is to create a state of epigenetic plasticity, where more plastic chromatin may sample alternative transcriptional programs or gene pathways, some of which may confer a growth advantage (Flavahan et al., 2017). To confirm loss of pRB in cancer cells leads to a more dynamic chromatin state with novel chromatin conformations, techniques like Hi-C, an “all-versus-all” chromosome conformation method, will be useful. Comparisons could be made with cell types of origin to gain insight into the extent of topology changes from non-cancerous to cancerous cells, and additional comparisons can then be made to cancers that have similar mutational spectrums but have different *RBI* status to determine how much loss of the pRB-condensin II complex is able to further contribute to altered chromosome conformations that will likely be observed in many cancers.

4.6 Lack of cancer progression in *Rb1^{L/L}* mice

Despite altered chromosome topology and an altered transcriptional profile in interphase cells from *Rb1^{L/L}* mice described in chapter 3, and the fact that these cells have hypocondensation of chromatin in mitotic cells and lagging chromosomes that lead to aneuploidy (Coschi et al., 2010; Isaac et al., 2006), these mice do not develop spontaneous tumors (Isaac et al., 2006). This may suggest that the maintenance of proper chromosome structure is not one of the more important aspects of tumor suppression by pRB. Conversely, altered chromosome topology in interphase and mitotic cells does create the opportunity for genetic change that can contribute to cancer pathogenesis, but the lack of an inappropriate growth-promoting signal is likely what is keeping these cells

from becoming cancerous. Evidence for this comes from *Rb1^{L/L}* mice that were crossed into the cancer-prone *Trp53^{-/-}* background. *Rb1^{L/L}; Trp53^{-/-}* mice succumb to cancer more rapidly than *Trp53^{-/-}* controls and *Rb1^{L/L}; Trp53^{-/-}* mice also have a trend toward more aggressive tumors, increased numbers of animals with multiple tumors, and more frequent metastases, indicating that the *Rb1^{L/L}* allele increases cancer susceptibility (Coschi et al., 2010).

Interestingly, *Rb1^{G/G}* mutant mice, which have a targeted mutation in the pRB pocket that disrupts the ability of pRB from associating with E2Fs through the pocket domain, and thus have increased expression of E2F target genes, are also viable and do not succumb to spontaneous tumors (Cecchini et al., 2014). Except for an accelerated entry into S phase in response to serum restimulation following serum starvation, cell cycle regulation in *Rb1^{G/G}* MEFs largely parallels what is seen in wild type cells (Cecchini et al., 2014). Overall, studies using this mouse model suggest that loss of E2F transcriptional repression is insufficient to cause tumor formation. Similar to what was seen with *Rb1^{L/L}* mice though, when *Rb1^{G/G}* were crossed with *Trp53^{-/-}* mice, *Rb1^{G/G}; Trp53^{-/-}* mutant mice had significantly shorter disease-free survival compared to *Trp53^{-/-}* controls. Therefore, it would be interesting to create an *Rb1^{G,L/G,L}* mouse line and assess phenotypes. Clearly both E2F transcriptional control and genome stability are facets of pRB tumor suppression, and if both of these are rendered non-functional in a compound mutant, tumor formation may occur. Additionally, similarly to what is seen in *Rb1^{-/-}* mice, there is always the possibility that *Rb1^{G,L/G,L}* mice may die embryonically (Wu et al., 2003).

Intriguingly, when thymocytes from the *Rb1^{L/L}* mouse were previously examined, no aneuploid or tetraploid cells were found (Isaac et al., 2006). This implies that endogenous proliferation rates may not be high enough to generate aneuploid cells faster than they can be eliminated, and thus, the mitotic defects observed in *Rb1^{L/L}* MEFs were not able to be detected *in vivo*. This eradication of cells may be due the cyclic GMP-AMP synthase (cGAS) protein, which is a cytosolic DNA sensor that activates innate immune responses, including the induction of interferons (Chen et al., 2016). Delivery of DNA to the host cytoplasm by microbial infection is pathogen-associated, while self-DNA is

danger-associated when it enters the cytoplasm from the nucleus, as it can be the result of DNA damage or the reverse transcription of retroelements, for example. Therefore, in *Rb1^{L/L}* mice, it is highly likely that cells with more severely altered chromosome topology are being cleared by the innate immune system through the cGAS pathway. This could be investigated further through a *Rb1^{L/L}; Mb21d1^{-/-}* compound mutant mouse, where *Mb21d1* is the gene that codes for cGAS. Without activation of immune defenses against cells that may be premalignant, the *Rb1^L* mutation might lead to a cancerous phenotype in this *Rb1^{L/L}; Mb21d1^{-/-}* background. It would also be interesting to probe if the cGAS pathway is only killing cells with mitotic problems leading to aneuploidy, or if it is able to somehow target cells that have non-mitotic phenotypes, like those with more plastic chromatin in interphase. It is possible that mitotic problems and dynamic interphase chromatin will go hand-in-hand in cells from these mice.

4.7 Further investigation of chromosome topology in *Rb1^{L/L}* cells

Although my work has thoroughly investigated chromosome conformation at specific loci in the genome using 3C and 4C-Seq techniques, differences in TADs as well as chromatin compartments, or regions of active or repressed chromatin states, remain to be investigated between wild type and *Rb1^{L/L}* cells. Changes in these types of chromosome conformations in *Rb1^{L/L}* cells could be studied further using an all-versus-all approach like Hi-C. Based on my 4C-Seq data, I predict that Hi-C data would display changes predominantly in compartments between the wild type and mutant cells, as I saw no difference in short-range loops at the bidirectional promoters I investigated but did find differences in long-range chromosome contacts. This agrees with previous predictions that condensin II is important for facilitating compartment formation based on the localization of condensin II at boundary-boundary interaction sites (Yuen and Gerton, 2018). Because of the partial loss of function mutation, use of *Rb1^{L/L}* cells would be one of the best ways to study the effects of diminished condensin II binding on chromosome conformation genome-wide in interphase cells. However, one obvious barrier to carry out this experiment is the amount of sequencing depth required. Generally speaking, there is a direct relationship between mapping resolution and sequencing depth for a Hi-C assay,

and the rule of thumb is to increase resolution by a factor of n , the number of reads needs to be increased by a factor of n^2 (Lieberman-Aiden et al., 2009). In genomes as large as those from human and mouse, generating contact profiles with resolution from 40 kb to 1 kb requires hundreds of millions to billions of paired-end reads (Han et al., 2018). Because many of the effects we saw in *Rb1^{L/L}* MEFs using 4C-Seq data were relatively subtle, higher resolution Hi-C contact maps might be required to observe global consequences of diminished condensin II binding in *Rb1^{L/L}* cells. Therefore, the amount of sequencing as well as the computational resources required makes Hi-C an expensive technique to pursue, meaning the decision to carry out this experiment merits a cost-benefits analysis. There are also alternative “many-to-all” methods of chromosome conformation capture, which enable the generation of high-resolution maps at a subfraction of the genome through enrichment of specific regions of interest out of the Hi-C library prior to sequencing (reviewed in (Sati and Cavalli, 2017)). Regions that could be enriched for based on my data, therefore, are bidirectional promoters where condensin II is seen to bind. This could reduce the cost of sequencing and the amount of computational resources required, but could still result in more global data, and thus, more global conclusions could be drawn.

4.8 Multiple possibilities for long-range chromosome interactions facilitated by condensin II

Although our initial model of how the pRB-TFIIC-condensin II complex could be organizing chromosome conformation at bidirectional promoters in chapter 3 showed a single condensin II molecule entrapping a single piece of duplex DNA (Figure 3.15A), there are several possible ways in which condensin II could be mediating this interaction. For example, since there has been evidence of condensin oligomerization (Keenholtz et al., 2017; Wang et al., 2017), it remains feasible that multiple condensin II complexes interact to bring chromatin close in three-dimensional space. How this oligomerization between condensin II complexes occurs is also largely unknown. One likely possibility is that the ATPase heads of different SMC dimers could interact and form a double ring structure, for example (Hirano, 2006). It has also been proposed that homotypic HEAT-HEAT interactions (e.g. between CAP-D2 and CAP-D2) might be important for

intercomplex crosstalk between condensins (Kinoshita et al., 2015). Therefore, some of the other ways in which the pRB-condensin II complex could be mediating long-range chromosome contacts at bidirectional promoters are depicted in Figure 4.1. The given scenarios are certainly not exhaustive however; how condensin complexes interact with DNA and form loops is still a mystery. One recent review actually listed some rather important outstanding questions in the SMC field, which include how do SMC complexes entrap DNA to form initial chromatin loops, how do these complexes processively expand chromatin loops, and do all SMC complexes form chromatin loops as monomers (van Ruiten and Rowland, 2018)?

4.9 Formation and recruitment of the pRB-condensin II complex

While the association of condensin II with pRB has been characterized as dependent on the LXCXE binding cleft, there are questions that still exist about the biochemical properties of this complex. For example, we do not know if there are post-translational modifications on either pRB or condensin II that are required for this complex to form, or whether certain modifications may prevent complex formation. Since pRB interacts with E2F1 using its specific interaction surface in the C-terminal domain when forming a complex with condensin II (Coschi et al., 2014; Ishak et al., 2017), we do know that hyperphosphorylation of pRB would not inhibit its interaction with E2F1 (Cecchini and Dick, 2011). Also, since binding of E2F1 to phosphorylated pRB has been suggested to alter the binding specificity of E2F1 such that it can bind at non-canonical E2F consensus sequences on DNA (Dick and Dyson, 2003; Tao et al., 1997), phosphorylated pRB in complex with condensin II remains an intriguing possibility. Further evidence of this possibility comes from chapter 3, where I demonstrate that condensin II binds at many promoters in an pRB-dependent manner, largely outside of those that are related to pRB's canonical function in G1-S regulation of E2F transcription. In addition, phosphorylation is also a common post translational modification that regulates the functions of condensin II (reviewed in (Kagami and Yoshida, 2016)), suggesting that it, too, could have specific modifications that may facilitate its interaction with pRB.

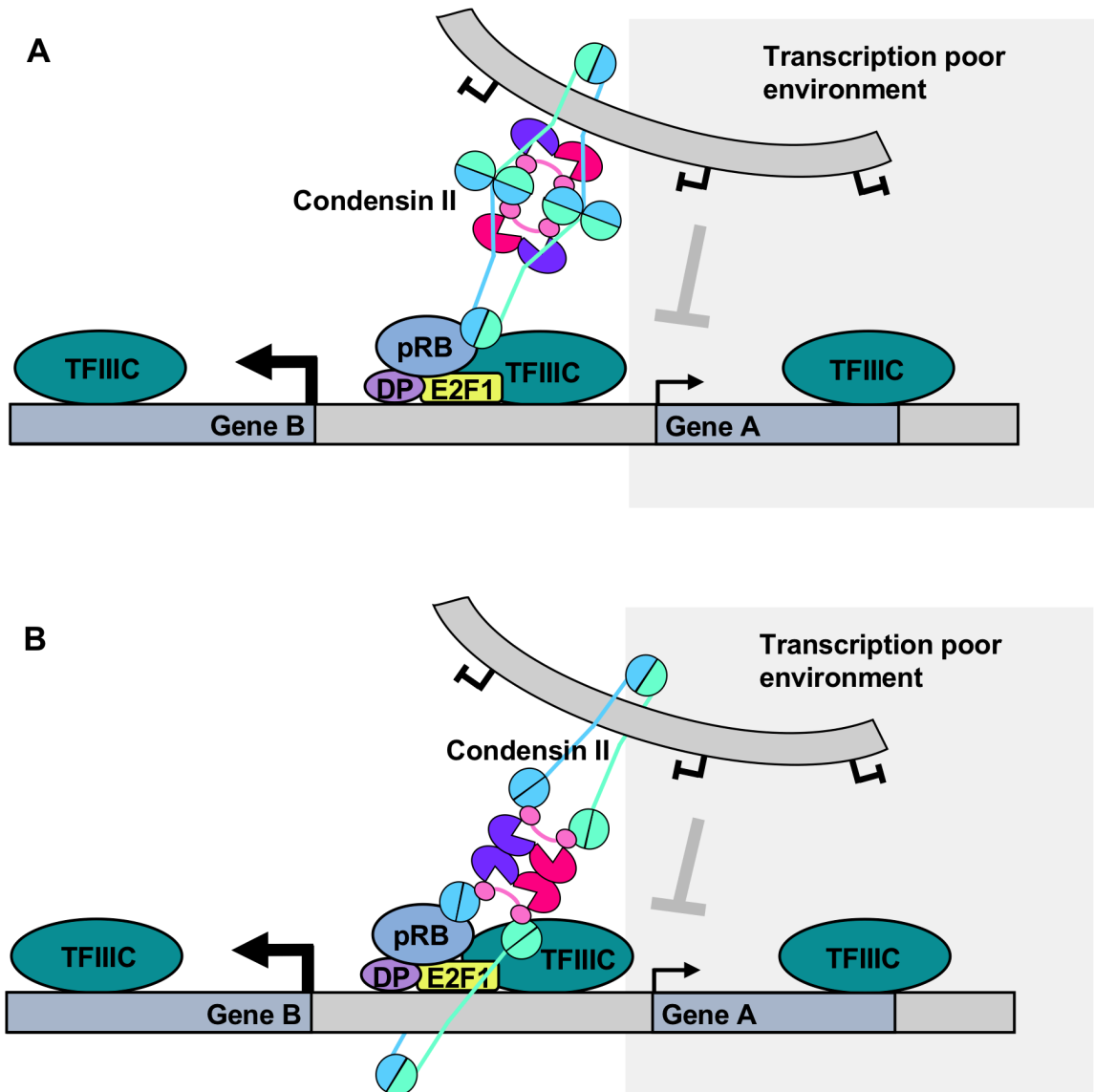


Figure 4.1: Additional models of long-range chromosome contacts mediated by condensin II at bidirectional promoters.

(A) Condensin II dimerization through the SMC ATPase heads could result in the formation of a double ring structure that is able to bring two pieces of chromatin close in proximity. In this diagram, the hinge domain of condensin II is binding to pRB and TFIIC, and only one condensin II complex is physically encircling DNA. (B) Homotypic HEAT-HEAT interactions could also be important for condensin II intercomplex crosstalk. In this diagram, the head domains of condensin II are interacting with pRB and TFIIC, and both condensin II subunits are physically encircling a DNA strand. Interactions in A and B between proteins and DNA can be mixed-and-matched, and these possibilities of how condensin II could be mediating chromosome contacts at bidirectional promoters are not exhaustive.

Moreover, we still do not know whether interactions between pRB and certain condensin II subunits are direct or indirect, and which subunits are essential for this interaction. Work in chapter 3 has shown that TFIIC is likely a part of the same complex, as reduced recruitment of TFIIC at promoters in *Rbl^{L/L}* cells also occurs. Further support of this comes from previous work demonstrating interactions between pRB and TFIIC (Chu et al., 1997), and TFIIC and condensin II (Yuen et al., 2017). Therefore, one possibility is that pRB binds to TFIIC, which in turn is able to recruit condensin II. However, further biochemical experiments are needed to determine how these proteins may all interact and whether additional co-factors are required for formation of this pRB-TFIIC-condensin II complex.

As previously mentioned, condensin II binds at many non-E2F target gene promoters in an pRB-dependent manner, begging the question of how these proteins are recruited to these sites. Again, it is possible that E2F1 is responsible for localization of this complex and is binding at non-canonical E2F consensus sequences on DNA (Dick and Dyson, 2003; Tao et al., 1997). Alternatively, the recruitment of pRB and E2F1 to sites of DNA damage suggests that these proteins may be recruited through sequence-independent means within the genome (Cook et al., 2015; Velez-Cruz et al., 2016). One possibility is that histone marks may mediate binding of the pRB and condensin II containing complex. Condensin II might be responsible for interactions with histones, as interactions with specific histone modifications have been previously demonstrated. For example, the CAP-G2 and CAP-D3 HEAT-repeat containing subunits can associate with H4K20me1 histone tails (Liu et al., 2010), and CAP-D3 can interact with H3K4me3 (Yuen et al., 2017). Another option is that TFIIC could be responsible for some degree of sequence-specific binding of these proteins, as sites of TFIIC-condensin II binding were previously demonstrated to be enriched for a motif similar to the consensus B box motif TFIIC is known to bind (Yuen et al., 2017). Overall, although my thesis has increased our knowledge of where the pRB-condensin II complex binds and its roles at target locations, there are still several questions about the properties of this complex that remain to be answered.

4.10 Summary of pRB functions in genome stability

In summary, my thesis establishes previously unappreciated roles for pRB which aid in tumor suppression by maintaining genome integrity. Through the use of cancer cell lines with already compromised pRB-pathways, I observed increased levels of basal DNA damage as well as mitotic errors when *RBI* was mutated. I also demonstrated increased genome instability even when cells have single copy loss of *RBI*. The main source of the increased spontaneous DNA damage in pRB mutant cells appears to be increased quantities of reactive oxygen species as well as defects in homologous recombination. There was also an enhanced capacity to seed new tumors in the lungs of xenografted immune compromised mice when *RBI* was mutated. Furthermore, this thesis also reveals that pRB is a recruitment factor for both condensin II and TFIIC at many promoters throughout the genome, particularly at bidirectional promoters, where diminished localization of these proteins leads to misexpression of many divergently paired genes. Loss of condensin II binding is also associated with altered long-range chromatin interactions with between bidirectional promoters and distant loci. A similar mechanism may occur in humans, as some genes at bidirectional promoters have significantly higher expression in lung adenocarcinoma patients with *RBI* deletion compared to those that are diploid for *RBI*.

Overall, my work has added to our current model of pRB-mediated tumor suppression, emphasising that its involvement in tumor suppression expands beyond solely maintaining proliferative control through repressing E2F transcription factors (Dyson, 1998). My thesis reveals that loss of *RBI* can also contribute to cancer progression through increased DNA damage and altered chromosome topology and these findings may shed light on associations between pRB status and patient outcome, disease progression, and genome reorganization.

4.11 References

Abramson, D.H., Ellsworth, R.M., and Zimmerman, L.E. (1976). Nonocular cancer in retinoblastoma survivors. *Trans Sect Ophthalmol Am Acad Ophthalmol Otolaryngol* 81, 454-457.

Betancur, P.A., Abraham, B.J., Yiu, Y.Y., Willingham, S.B., Khameneh, F., Zarnegar, M., Kuo, A.H., McKenna, K., Kojima, Y., Leeper, N.J., *et al.* (2017). A CD47-associated super-enhancer links pro-inflammatory signalling to CD47 upregulation in breast cancer. *Nature communications* 8, 14802.

Cecchini, M.J., and Dick, F.A. (2011). The biochemical basis of CDK phosphorylation-independent regulation of E2F1 by the retinoblastoma protein. *Biochem J.*

Cecchini, M.J., Ishak, C.A., Passos, D.T., Warner, A., Palma, D.A., Howlett, C.J., Driman, D.K., and Dick, F.A. (2015). Loss of the retinoblastoma tumor suppressor correlates with improved outcome in patients with lung adenocarcinoma treated with surgery and chemotherapy. *Human pathology* 46, 1922-1934.

Cecchini, M.J., Thwaites, M.J., Talluri, S., MacDonald, J.I., Passos, D.T., Chong, J.L., Cantalupo, P., Stafford, P.M., Saenz-Robles, M.T., Francis, S.M., *et al.* (2014). A retinoblastoma allele that is mutated at its common E2F interaction site inhibits cell proliferation in gene-targeted mice. *Mol Cell Biol* 34, 2029-2045.

Chen, Q., Sun, L., and Chen, Z.J. (2016). Regulation and function of the cGAS-STING pathway of cytosolic DNA sensing. *Nat Immunol* 17, 1142-1149.

Cho, S.W., Xu, J., Sun, R., Mumbach, M.R., Carter, A.C., Chen, Y.G., Yost, K.E., Kim, J., He, J., Nevins, S.A., *et al.* (2018). Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element. *Cell* 173, 1398-1412 e1322.

Chu, W.M., Wang, Z., Roeder, R.G., and Schmid, C.W. (1997). RNA polymerase III transcription repressed by Rb through its interactions with TFIIB and TFIIC2. *J Biol Chem* 272, 14755-14761.

Cook, R., Zoumpoulidou, G., Luczynski, M.T., Rieger, S., Moquet, J., Spanswick, V.J., Hartley, J.A., Rothkamm, K., Huang, P.H., and Mitnacht, S. (2015). Direct involvement of retinoblastoma family proteins in DNA repair by non-homologous end-joining. *Cell Rep* 10, 2006-2018.

Coschi, C.H., Ishak, C.A., Gallo, D., Marshall, A., Talluri, S., Wang, J., Cecchini, M.J., Martens, A.L., Percy, V., Welch, I., *et al.* (2014). Haploinsufficiency of an RB-E2F1-Condensin II Complex Leads to Aberrant Replication and Aneuploidy. *Cancer discovery.*

Coschi, C.H., Martens, A.L., Ritchie, K., Francis, S.M., Chakrabarti, S., Berube, N.G., and Dick, F.A. (2010). Mitotic chromosome condensation mediated by the retinoblastoma protein is tumor-suppressive. *Genes & Development* 24, 1351-1363.

Dick, F.A., and Dyson, N. (2003). pRB Contains an E2F1 Specific Binding Domain that Allows E2F1 Induced Apoptosis to be Regulated Separately from other E2F Activities. *Mol Cell* 12, 639-649.

Dick, F.A., Goodrich, D.W., Sage, J., and Dyson, N.J. (2018). Non-canonical functions of the RB protein in cancer. *Nat Rev Cancer.*

- Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes Dev* *12*, 2245-2262.
- Flavahan, W.A., Drier, Y., Liao, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suva, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* *529*, 110-114.
- Flavahan, W.A., Gaskell, E., and Bernstein, B.E. (2017). Epigenetic plasticity and the hallmarks of cancer. *Science* *357*.
- Garsed, D.W., Alsop, K., Fereday, S., Emmanuel, C., Kennedy, C.J., Etemadmoghadam, D., Gao, B., GebSKI, V., Gares, V., Christie, E.L., *et al.* (2018). Homologous Recombination DNA Repair Pathway Disruption and Retinoblastoma Protein Loss Are Associated with Exceptional Survival in High-Grade Serous Ovarian Cancer. *Clin Cancer Res* *24*, 569-580.
- Gonzalez-Vasconcellos, I., Anastasov, N., Sanli-Bonazzi, B., Klymenko, O., Atkinson, M.J., and Rosemann, M. (2013). Rb1 haploinsufficiency promotes telomere attrition and radiation-induced genomic instability. *Cancer Res* *73*, 4247-4255.
- Han, J., Zhang, Z., and Wang, K. (2018). 3C and 3C-based techniques: the powerful tools for spatial genome organization deciphering. *Mol Cytogenet* *11*, 21.
- Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. *Cell* *100*, 57-70.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* *144*, 646-674.
- Hirano, T. (2006). At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol* *7*, 311-322.
- Hnisz, D., Schuijers, J., Li, C.H., and Young, R.A. (2018). Regulation and Dysregulation of Chromosome Structure in Cancer. *Annual Review of Cancer Biology* *2*, 21-40.
- Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., *et al.* (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* *351*, 1454-1458.
- Isaac, C.E., Francis, S.M., Martens, A.L., Julian, L.M., Seifried, L.A., Erdmann, N., Binne, U.K., Harrington, L., Sicinski, P., Berube, N.G., *et al.* (2006). The retinoblastoma protein regulates pericentric heterochromatin. *Mol Cell Biol* *26*, 3659-3671.
- Ishak, C.A., Coschi, C.H., Roes, M.V., and Dick, F.A. (2017). Disruption of CDK-resistant chromatin association by pRB causes DNA damage, mitotic errors, and reduces Condensin II recruitment. *Cell Cycle* *16*, 1430-1439.

- Kagami, Y., and Yoshida, K. (2016). The functional role for condensin in the regulation of chromosomal organization during the cell cycle. *Cellular and molecular life sciences : CMLS* 73, 4591-4598.
- Keenholtz, R.A., Dhanaraman, T., Palou, R., Yu, J., D'Amours, D., and Marko, J.F. (2017). Oligomerization and ATP stimulate condensin-mediated DNA compaction. *Sci Rep* 7, 14279.
- Kinoshita, K., Kobayashi, T.J., and Hirano, T. (2015). Balancing acts of two HEAT subunits of condensin I support dynamic assembly of chromosome axes. *Dev Cell* 33, 94-106.
- Kleinerman, R.A., Schonfeld, S.J., and Tucker, M.A. (2012). Sarcomas in hereditary retinoblastoma. *Clin Sarcoma Res* 2, 15.
- Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 68, 820-823.
- Konstantinopoulos, P.A., Ceccaldi, R., Shapiro, G.I., and D'Andrea, A.D. (2015). Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. *Cancer discovery* 5, 1137-1154.
- Ledermann, J., Harter, P., Gourley, C., Friedlander, M., Vergote, I., Rustin, G., Scott, C., Meier, W., Shapira-Frommer, R., Safra, T., *et al.* (2012). Olaparib maintenance therapy in platinum-sensitive relapsed ovarian cancer. *The New England journal of medicine* 366, 1382-1392.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-293.
- Liu, J.F., Barry, W.T., Birrer, M., Lee, J.M., Buckanovich, R.J., Fleming, G.F., Rimel, B., Buss, M.K., Nattam, S., Hurteau, J., *et al.* (2014). Combination cediranib and olaparib versus olaparib alone for women with recurrent platinum-sensitive ovarian cancer: a randomised phase 2 study. *Lancet Oncol* 15, 1207-1214.
- Liu, W., Tanasa, B., Tyurina, O.V., Zhou, T.Y., Gassmann, R., Liu, W.T., Ohgi, K.A., Benner, C., Garcia-Bassets, I., Aggarwal, A.K., *et al.* (2010). PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. *Nature* 466, 508-512.
- Ludovini, V., Gregorc, V., Pistola, L., Mihaylova, Z., Floriani, I., Darwish, S., Stracci, F., Tofanetti, F.R., Ferraldeschi, M., Di Carlo, L., *et al.* (2004). Vascular endothelial growth factor, p53, Rb, Bcl-2 expression and response to chemotherapy in advanced non-small cell lung cancer. *Lung Cancer* 46, 77-85.

- Marees, T., Moll, A.C., Imhof, S.M., de Boer, M.R., Ringens, P.J., and van Leeuwen, F.E. (2008). Risk of second malignancies in survivors of retinoblastoma: more than 40 years of follow-up. *Journal of the National Cancer Institute* *100*, 1771-1779.
- Moore, K., Colombo, N., Scambia, G., Kim, B.G., Oaknin, A., Friedlander, M., Lisysanskaya, A., Floquet, A., Leary, A., Sonke, G.S., *et al.* (2018). Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *The New England journal of medicine* *379*, 2495-2505.
- Oza, A.M., Cibula, D., Benzaquen, A.O., Poole, C., Mathijssen, R.H., Sonke, G.S., Colombo, N., Spacek, J., Vuylsteke, P., Hirte, H., *et al.* (2015). Olaparib combined with chemotherapy for recurrent platinum-sensitive ovarian cancer: a randomised phase 2 trial. *Lancet Oncol* *16*, 87-97.
- Pujade-Lauraine, E., Ledermann, J.A., Selle, F., GebSKI, V., Penson, R.T., Oza, A.M., Korach, J., Huzarski, T., Poveda, A., Pignata, S., *et al.* (2017). Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Oncol* *18*, 1274-1284.
- Robinson, D.R., Wu, Y.M., Lonigro, R.J., Vats, P., Cobain, E., Everett, J., Cao, X., Rabban, E., Kumar-Sinha, C., Raymond, V., *et al.* (2017). Integrative clinical genomics of metastatic cancer. *Nature* *548*, 297-303.
- Sati, S., and Cavalli, G. (2017). Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* *126*, 33-44.
- Stadhouders, R., Fillion, G.J., and Graf, T. (2019). Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* *569*, 345-354.
- Tao, Y., Kassatly, R.F., Cress, W.D., and Horowitz, J.M. (1997). Subunit composition determines E2F DNA-binding site specificity. *Mol Cell Biol* *17*, 6994-7007.
- van Ruiten, M.S., and Rowland, B.D. (2018). SMC Complexes: Universal DNA Looping Machines with Distinct Regulators. *Trends in genetics : TIG* *34*, 477-487.
- Velez-Cruz, R., Manickavinayaham, S., Biswas, A.K., Clary, R.W., Premkumar, T., Cole, F., and Johnson, D.G. (2016). RB localizes to DNA double-strand breaks and promotes DNA end resection and homologous recombination through the recruitment of BRG1. *Genes Dev* *30*, 2500-2512.
- Wang, X., Brandao, H.B., Le, T.B., Laub, M.T., and Rudner, D.Z. (2017). *Bacillus subtilis* SMC complexes juxtapose chromosome arms as they travel from origin to terminus. *Science* *355*, 524-527.
- Wu, L., de Bruin, A., Saavedra, H.I., Starovic, M., Trimboli, A., Yang, Y., Opavska, J., Wilson, P., Thompson, J.C., Ostrowski, M.C., *et al.* (2003). Extra-embryonic function of Rb is essential for embryonic development and viability. *Nature* *421*, 942-947.

Yuen, K.C., and Gerton, J.L. (2018). Taking cohesin and condensin in context. *PLoS genetics* *14*, e1007118.

Yuen, K.C., Slaughter, B.D., and Gerton, J.L. (2017). Condensin II is anchored by TFIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. *Sci Adv* *3*, e1700191.

Zheng, L., Flesken-Nikitin, A., Chen, P.L., and Lee, W.H. (2002). Deficiency of Retinoblastoma gene in mouse embryonic stem cells leads to genetic instability. *Cancer Res* *62*, 2498-2502.

Appendices

Appendix A: Permission for publication by Molecular and Cellular Biology

Data presented in chapter 2 is published in Molecular and Cellular Biology, American Society for Microbiology (ASM) journals.

Copyright © American Society for Microbiology, Molecular and Cellular Biology, doi: 10.1128/MCB.00105-19

Marshall AE, Roes MV, Passos DT, DeWeerd MC, Chaikovsky AC, Sage J, Howelett CJ, Dick FA. *RBI* deletion in RB-pathway disrupted cells results in DNA damage and cancer progression. (2019). Mol Cell Biol. doi: 10.1128/MCB.00105-19.

See the following page for the ASM Journals Statement of Author Rights, which includes their policy on permissions for including published material in a thesis.

ASM Journals Statement of Author Rights

Authors may post their articles to their institutional repositories

ASM grants authors the right to post their accepted manuscripts in publicly accessible electronic repositories maintained by funding agencies, as well as appropriate institutional or subject-based open repositories established by a government or non-commercial entity.

In preparation for the REF 2021, ASM would like to remind Authors that the current author fee structure, along with the policy outlined above, allows authors to comply with the HEFCE deposition requirements. If authors have paid a fee to make their article "gold" open access then there is no need to worry about these deposition requirements (See section 38 of the "Policy for open access in Research Excellence Framework 2021" document).

Please note that ASM makes the final, typeset articles from its primary-research journals available free of charge on the ASM Journals and PMC websites 6 months after final publication.

Authors may post their articles in full on personal or employer websites

ASM grants the author the right to post his/her article (after publication by ASM) on the author's personal or university-hosted website, but not on any corporate, government, or similar website, without ASM's prior permission, provided that proper credit is given to the original ASM publication.

Once ASM grants permission to the author, ASM requests that the posting release date for the content be no earlier than 6 months after the final publication of the typeset article by ASM.

Authors may make copies of their articles in full

Corresponding authors are entitled to 10 free downloads of their papers. Additionally, all authors may make up to 99 copies of his/her own work for personal or professional use (including teaching packs that are distributed free of charge within your own institution). For orders of 100 or more copies, you should seek ASM's permission or purchase access through Highwire's Pay-Per-View option, available on the ASM online journal sites.

Authors may republish/adapt portions of their articles

ASM also grants the author the right to republish discrete portions of his/her article in any other publication (including print, CD-ROM, and other electronic formats) of which he or she is author or editor, provided that proper credit is given to the original ASM publication. An ASM author also retains the right to reuse the full article in his/her dissertation or thesis. "Proper credit" means either the copyright lines shown on the top of the first page of the PDF version, or "Copyright © American Society for Microbiology, [insert journal name, volume number, year, page numbers and DOI]" of the HTML version. For technical questions about using Rightslink, please contact Customer Support via phone at (877) 622-5543 (toll free) or (978) 777-9929, or e-mail Rightslink customer care at customercare@copyright.com.

Appendix B: List of antibodies used

Antibody	Target	Species	Source	ChIP (Ab:Chromatin)	WB	IF
C-15	pRB	Rabbit	Santa Cruz	5 ug : 80 ug	1:1000	
M-153	pRB	Rabbit	Santa Cruz	5 ug : 80 ug		
M-136	pRB	Sheep	(Cecchini et al., 2014)	5 ug : 80 ug		
S855	pRB	Rabbit	(Cecchini et al., 2014)	5 ug : 80 ug		
Hyb4.1	pRB	Mouse	Developmental Studies Hybridoma Bank	500 uL : 80 ug		
G3-245	pRB	Mouse	BD Pharmingen		1:250	1:400
H-225	Sp1	Rabbit	Santa Cruz		1:1000	
05-636	H2A.X pSer139	Mouse	EMD Millipore	4 ug: 200 ug		1:400
sc-7790	BLM	Goat	Santa Cruz		1:200	1:150
sc-22760	53BP1	Rabbit	Santa Cruz			1:400
ab62623	DNA/RNA damage (8-oxoG)	Mouse	Abcam			1:200
A300- 244A	RPA32	Rabbit	Bethyl		1:2000	
ab83311	RAD54B	Rabbit	Abcam		1:500	
ab63801	RAD51	Rabbit	Abcam		1:2000	
05-858	Histone H4	Rabbit	EMD Millipore	10 ug : 200 ug		
CAP-D3	CAP-D3	Rabbit	(Coschi et al., 2010)	10 ug : 200 ug		
A302- 275A	CAP-H2	Rabbit	Bethyl	10 ug : 200 ug		
07-449	H3K27me3	Rabbit	EMD Millipore	10 ug : 200 ug		
A301- 291A	TFIIIC-220	Rabbit	Bethyl	6 ug : 120 ug		

Appendix C: List of plasmids used

Name	Description	Obtained/ constructed	Bacterial Resistance	Selectable markers	Stock Number
pX459	Plasmid expressing WT Cas9	Addgene #48139	Ampicillin	Puromycin	0681
pX459+ X22B	Plasmid expressing WT Cas9 and sgRNA for exon 22 of <i>RBI</i> (X22B sequence)	Addgene #48139 and Fred Dick	Ampicillin	Puromycin	0689
pX459+ X22C	Plasmid expressing WT Cas9 and sgRNA for exon 22 of <i>RBI</i> (X22C sequence)	Addgene #48139 and Fred Dick	Ampicillin	Puromycin	0690
pX462	Plasmid expressing D10A mutant of Cas9	Addgene #48141	Ampicillin	Puromycin	0682
pX462+ X22B	Plasmid expressing D10A mutant of Cas9 and sgRNA for exon 22 of <i>RBI</i> (X22B sequence)	Addgene #48141 and Fred Dick	Ampicillin	Puromycin	0693
pX462+ X22C	Plasmid expressing D10A mutant of Cas9 and sgRNA for exon 22 of <i>RBI</i> (X22C sequence)	Addgene #48141 and Fred Dick	Ampicillin	Puromycin	0694
pDRGFP	<i>In vivo</i> homologous recombination substrate	N. Bérubé	Ampicillin	Puromycin	0784
pimEJ5GFP	<i>In vivo</i> non-homologous end joining substrate	Addgene #44026	Ampicillin	Puromycin	0815
pCBASceI	I-SceI endonuclease expression vector (pCAGGS backbone)	N. Bérubé	Ampicillin	N/A	0785
pMSCV-Blasticidin	Empty backbone with blasticidin resistance	Addgene #75085	Ampicillin	Blasticidin	0816

pCAG- FALSE	pCAG promoter driving no CDS (pCAGGS backbone)	Addgene #89689	Ampicillin	N/A	0817
lentiCRISPR v2	Lentiviral backbone expressing WT Cas9	Addgene #52961	Ampicillin	Puromycin	0818
lentiCRISPR v2+X22B	Lentiviral backbone expressing WT Cas9 and sgRNA for exon 22 of <i>RBI</i> (X22B sequence)	Addgene #52961 and Michael Roes	Ampicillin	Puromycin	0819
pMD2.G	VSV-G envelope expressing plasmid	Greg Fonseca/ Mymryk	Ampicillin	N/A	0656
psPAX2	Lentiviral packaging plasmid	Greg Fonseca/ Mymryk	Ampicillin	N/A	0655

Curriculum Vitae

Aren Elizabeth Marshall

Education

PhD Candidate 2013 – 2019

Biochemistry
The University of Western Ontario, London, ON
Supervisor: Dr. Fred Dick

Bachelor of Science; with distinction 2008 – 2013

Honors Specialization in Genetics and Biochemistry
The University of Western Ontario

Awards and Scholarships

Schulich Graduate Scholarship (\$2000/year)	2015 - 2018
Ontario Graduate Scholarships; PhD (\$15000/year)	2015 - 2018
Western Graduate Research Scholarship (\$4500/year)	2013 - 2018
Howard Hughes Medical Institute Stipend (\$750)	2016
CIHR Strategic Training Program in Cancer Research and Technology Transfer (\$3100)	2014 - 2015
Ontario Graduate Scholarship; MSc (\$15000)	2014 - 2015
Canada Graduate Scholarship (Master's), NSERC (\$17500)	2013 - 2014
Ontario Graduate Scholarship; MSc <i>declined</i> (\$15000)	2013 - 2014
Western Gold Medal- Honors Specialization in Genetics and Biochemistry	2013
Andrew and Sarah Hamilton Scholarship (\$1500/year)	2009 - 2013
Continuing Entrance Scholarship (\$2000/year)	2008 - 2013
NSERC Undergraduate Student Research Award (\$4500)	2012
Biochemistry 2280A Award (\$100)	2011
Governor General's Academic Medal	2008
Orillia Business Women's Association Lisa Brooking Young Woman of the Year	2008

Publications

Marshall AE, Ishak CA, Dick FA. In preparation. An RB-Condensin II complex mediates long range chromosome interactions and influences expression at divergently paired genes. (*Submitted*)

Marshall AE, Roes MV, Passos DT, DeWeerd MC, Chaikovsky AC, Sage J, Howelett CJ, Dick FA. *RBI* deletion in RB-pathway disrupted cells results in DNA damage and cancer progression. (2019). *Mol Cell Biol.* 39, e00105-00119.
Submitted image was chosen for the cover of Mol Cell Biol. vol. 39, issue 16

Ishak CA, **Marshall AE**, Passos DT, White CR, Kim SJ, Cecchini MJ, Ferwati S, MacDonald WA, Howlett CJ, Welch ID, Rubin SM, Mann MRW, Dick FA. An RB-EZH2 complex mediates silencing of repetitive DNA sequences. (2016) Mol Cell. 64, 1074-1087

Coschi CH, Ishak CA, Gallo D, **Marshall AE**, Talluri S, Wang J, Cecchini MJ, Martens AL, Percy V, Welch I, Boutros PC, Brown GW, Dick FA. Haploinsufficiency of an RB-E2F1-Condensin II complex leads to aberrant replication and aneuploidy. (2014) Cancer Discov. 4, 840-53

International and National Meetings (Presenter is underlined)

Marshall AE and Dick FA. An pRB-Condensin II complex mediates long range chromosome interactions and regulates expression at divergently paired genes. ICEGB DNA Tumour Virus Meeting, Trieste, Italy, July 2019. (Oral presentation)

Marshall AE, Coschi CH, Dick FA. Functional investigation of transcriptional control by the retinoblastoma protein and condensin II. The LXXXII Cold Spring Harbor Symposium on Quantitative Biology: Chromosome Segregation & Structure, Cold Spring Harbor, New York, June 2017. (Poster)

Ishak CA, **Marshall AE**, Passos DT, White CR, Cecchini MJ, Ferwati S, Macdonald WA, Mann MR, Dick FA. The E1A-insensitive pRB-E2F1 complex silences endogenous retroviruses. The DNA Tumor Virus Meeting, Trieste, Italy, July 2015. (Oral presentation)

Marshall AE, Coschi CH, Dick FA. The localization and properties of a pRB, E2F1, and condensin II complex. Second Canadian Conference on Epigenetics, London, Canada, June 2014. (Poster)

Ishak CA, Passos DT, Ferwati S, **Marshall AE**, Dick FA. Alternate pRB-E2F1 complexes silence repetitive elements. Second Canadian Conference on Epigenetics, London, Canada, June 2014. (Oral presentation)

Coschi CH, Ishak CA, Gallo D, **Marshall AE**, Talluri S, Wang J, Cecchini MJ, Martens AL, Percy V, Welch I, Boutros PC, Brown GW, Dick FA. Haploinsufficiency of an RB-E2F1-Condensin II complex causes replication stress and aneuploidy. Meeting on the Cell Cycle, Cold Spring Harbor, New York, May 2014. (Oral presentation)

Coschi CH, Ishak CA, Gallo D, **Marshall AE**, Talluri S, Wang J, Cecchini MJ, Martens AL, Percy V, Welch I, Boutros PC, Brown GW, Dick FA. Haploinsufficiency of an RB-E2F1-Condensin II complex Causes Under Replication and Contributes to Mesenchymal Cancers. Third International Retinoblastoma Meeting, Monterey, California, October 7, 2013. (Oral presentation)

Marshall AE, Eitutis SE, Osborne Locke ME, Daley M, Hill KA. An array-based genomic survey of copy number variation across wild caught and inbred mice implicates different environmental-responsive candidate genes. Environmental Mutagenesis and Genomics Society Annual Meeting, Monterey, California, September 2013. (Oral presentation)

Local Meetings (Presenter is underlined)

Marshall AE and Dick FA. An RB-condensin II complex mediates long range chromosome interactions and regulates expression at divergently paired genes. Department of Oncology Research and Education Day, London, Canada, June 14, 2019. (Selected for poster award)

Marshall AE and Dick FA. An RB-Condensin II complex mediates long range chromosome interactions and regulates expression at divergently paired genes. Dr. Maud L. Menten Memorial Spring Symposium, Department of Biochemistry, London, Canada, May 3, 2019. (Oral presentation)

Marshall AE, Roes MV, Dick FA. Identification of cancer vulnerabilities created by RB1 gene loss. Department of Oncology Research and Education Day, London, Canada, June 8, 2018. (Selected for oral presentation and awarded best talk)

Marshall AE, Roes MV, Dick FA. Investigation of reduced pRB expression in treatment response. Oncology Grand Rounds: Young Cancer Researcher Trainee Showcase, London Regional Cancer Program, London, Canada, December 5, 2017. (Oral presentation)

Marshall AE, Coschi CH, Dick FA. Functional investigation of transcriptional control by the retinoblastoma protein and condensin II. Department of Oncology Research and Education Day, London, Canada, June 16, 2017. (Poster)

Marshall AE, Roes MV, Dick FA. Investigation of pRB hemizyosity and its effect on genome instability in cancer cells. Department of Oncology Research and Education Day, London, Canada, June 17, 2016. (Poster)

Marshall AE, Roes MV, Dick FA. Investigation of pRB hemizyosity and its effect on genome instability in cancer cells. London Health Research Day, London, Canada, March 29, 2016. (Poster)

Marshall AE, Coschi CH, Dick FA. Investigation of replication stress in pRB haploinsufficient cells. Department of Oncology Research and Education Day, London, Canada, June 26, 2015. (Poster)

Research and Professional Experience

**Biochemistry Graduate Student Association
Representative**

September 2018 – Present

Biochemistry Department, The University of Western Ontario, London, ON

Maud L. Menten Lecture Student Representative **September 2016 – August 2018**
Biochemistry Department, The University of Western Ontario, London, ON

Co-founder of Victoria Research Labs **February 2016 – August 2018**
Bioinformatics Group
Victoria Hospital, London Health Sciences Centre, London, ON

Canadian Cancer Society's Research Information Outreach Team **2016 – 2018**
Research Information Outreach Team (RIOT); Let's Talk Science; London, ON

Graduate Student Representative for the **September 2015 – August 2017**
Biochemistry Undergraduate Program Committee
Biochemistry Department, The University of Western Ontario, London, ON

Member of the Mustang Bioinformatics Club **September 2015 – August 2016**
Graduate students, The University of Western Ontario, London, ON

Advanced Sequencing Technologies & Applications **November 10 – 22, 2015**
Cold Spring Harbor Laboratory course, Cold Spring Harbor, NY

Honors Thesis Student **September 2012 – April 2013**
The University of Western Ontario, London, ON
Supervisor: Dr. Kathleen Hill

Agroecology Research Intern **May 2011 – August 2012**
A&L Biologicals Inc., London ON
Supervisor: Dr. George Lazarovits