THE UNIVERSITY OF CHICAGO

DECIPHERING HOW GENOMIC INSTABILITY IN HEMATOPOIETIC CELLS DRIVES TUMORIGENESIS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

COMMITTEE ON CANCER BIOLOGY

BY

STEPHEN OREN ARNOVITZ

CHICAGO, ILLINOIS

JUNE 2022

Dedication

To my parents and grandparents, who supported me throughout my academic pursuits and inspired me to seek to understand the world around me by asking meaningful questions.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACMG	American College of Medical Genetics
ALC	Absolute lymphocyte count
alt-EJ	Alternative end joining
AML	Acute myeloid leukemia
ANC	Absolute neutrophil count
AT	Ataxia telangiectasia
BIR	Break induced replication
BM	Bone marrow
BMF	Bone marrow failure
BRCT	BRCA C Terminus
CBC	Complete blood count
CBF	Core binding factor
cDNA	Complementary DNA
CH	Clonal hematopoiesis
ChIP	Chromatin immunoprecipitation
CHK2	Checkpoint kinase 2 protein
CldU	5-Chloro-2'-deoxy-uridine
cRSS	Cryptic recombination signal sequences
CTCL	Cutaneous T-cell Leukemia
DDR	DNA damage response
DP	Double positive (CD4+CD8+)
DSB	Double strand break

ES	Embryonic stem
ETP	Etoposide
FA	Fanconi anemia
FISH	Fluorescence in situ hybridization
FPKM	Frequency per kilobase million
gDNA	Genomic DNA
GDSC	Genomics of Drug Sensitivity in Cancer
HM	Hematopoietic malignancy
HR	Homologous recombination
HRD	Homologous recombination deficiency
HSPC	Hematopoietic stem and progenitor cell
HU	Hydroxyurea
IC50	Half maximal inhibitory concentration
IdU	5-Iodo-2’deoxy-uridine
Ig	Immunoglobulin
IHC	Immunohistochemistry
IP	Intraperitoneal
LK	Lineage- Sca1- cKit+
LN	Lymph node
LSK	Lineage- Sca1+ cKit+
MDS	Myelodysplastic syndrome
MPN	Myeloproliferative neoplasm
MPO	Myeloperoxidase

NES	Normalized enrichment score
NHEJ	Non-homologous end joining
NHL	Non-Hodgkin's lymphoma
OR	Odds ratio
PB	Peripheral blood
PCA	Principal component analysis
PCR	Polymerase chain reaction
PheWAS	Phenome-wide association study
pI:pC	Polyinosinic polycytidylic acid
pLoF	Predicted loss of function
PTCL	Peripheral T-cell Leukemia
qPCR	Quantitative PCR
qRT-PCR	Quantitative reverse transcription PCR
RSS	Recombination signal sequences
SP	Single positive
T-ALL	T-cell acute lymphoblastic leukemia
TCGA	The Cancer Genome Atlas
TCR	T-cell receptor
TdT	Terminal deoxynucleotidyl transferase
TF	Transcription factor
t-MN	Therapy related myeloid neoplasm
UK	United Kingdom
WBC	White blood cell

Acknowledgments

I would first like to thank my team of mentors that supported me throughout my doctoral work, Dr. Lucy Godley, Dr. Fotini Gounari, and Dr. Jane Churpek. We took on a unique mentorship relationship across multiple groups and projects, which was facilitated by a collaborative environment and a true team spirit for driving these projects. I thank all of my mentors for their flexibility and cooperation throughout this process.

I would like to thank Dr. Lucy Godley for her invaluable contributions to my project design, scientific thinking, and professional development. I am very grateful for the clinical perspective and bedside to bench approaches pursued in the Godley lab and the opportunity to participate in projects with direct translational value. I also am endlessly grateful for the assistance in pursuit of my postdoctoral position and her encouragement for me to think big about my future.

I would also like to thank Dr. Fotini Gounari for significant contributions to my scientific thinking and confidence as an independent researcher. I greatly benefited from an open-door policy and endless opportunities to discuss the scientific outlook of my projects in exquisite mechanistic detail. I also thank Fotini for the opportunity to participate in a wide range of projects and add significant immunology training to my cancer background. I also need to thank Fotini for endless emotional support throughout this process and for her deep care about my wellbeing and future prospects.

I would also like to thank Dr. Jane Churpek for her guidance and support throughout my PhD career. I am grateful for her willingness to take me on as her first graduate trainee and continuing to support me even after her departure from the

University of Chicago. I benefited tremendously from our discussions about the nuanced mechanisms involved in my project and carefully designing a research program to address meaningful questions. I would also like to thank Jane for significant emotional support through some of the difficulty of balancing my varied projects.

I would also like to thank my thesis committee members, Dr. Megan McNerney, Dr. Barbara Kee, Dr. Panos Nitzchristos and Dr. Scott Oakes, for their time, thoughtful critiques, and advice on optimizing the value of my thesis work. I would particularly like to thank Dr. Megan McNerney for helping me to navigate the complexities of my projects and managing my trajectory throughout my PhD studies. Her leadership on the committee and openness for discussion on any topic was essential for me completing my graduate studies.

I am also deeply grateful to the Pinsky Family for their financial contribution to support my studies and further my career as a research scientist. Moreover, I am thankful for our yearly check in meetings and their clear interest in the content and progress of my studies. This was a unique opportunity to engage with the direct translational benefit of my work alongside some of the patients who will be helped from the results of my studies. They truly made me feel a part of the family.

I would also like to thank all of the exceptional people in the research groups I participated in. In the Godley and Churpek Labs, I would like to thank former and past members including: Dr. Sakshi Uppal, Dr. John Cao, Dr. Anastasia Hains, Dr. Michael Drazer, Dr. Lorraine Canham, Dr. Krisitina Bigelow, Dr. Kiran Tawana, Dr. Imo Akpan, Dr. Ryan Stubbins, Dr. Allison West, Shawn Albert, Janet Lepore, Art Wolin, Maya Lewinsohn, Matthew Pozsgai, Matthew Jones, Aelin Kim, Kelsey McNeely, Anase Asom

and Daniel Mendez. I would particularly like to thank Dr. John Cao, Dr. Anastasia Hains and Dr. Michael Drazer who have grown to be close friends and were extremely supportive to me in pursuing my PhD. I will always be grateful to the sanity they provided and the help and support on my research. In the Gounari lab, I would like to thank Dr. Priya Mathur, Dr. Jasmin Quandt, Dr. Akinola (Junior) Emmanuel, Dr. Osman Abu, Melissa Tracy, Azam Mohsin, Greer Gurewitz, Francis Apolinario and Soumi Mondal. I would particularly like to thank Dr. Priya Mathur and Dr. Jasmin Quandt, whose friendship and support was invaluable. Additionally, Priya provided endless amounts of training and assistance in project design and interpretation of results. Jasmin trained me in flow cytometry and other critical lab procedures that are an essential part of my toolkit today.

I also need to thank my friends, or rather my Chicago family, for their unwavering support through many years of my PhD studies. The encouragement, understanding, and love I received from this excellent network of human beings buoyed me throughout my studies. I would particularly like to thank James Kowalsky and Susie Bernero for their friendship, meals, and constant emotional support. My roommate, Ashley Heyer, was also essential for me completing my graduate work, and I thank her for her understanding, encouragement, and keeping calories in my body.

Finally, and most importantly, I would like to thank my family. I cannot put into words the importance of their involvement in my pursuits of a career in science.

Abstract

Mutations in DNA repair factors are increasingly recognized for having deleterious effects on hematopoiesis and increasing risk for developing hematopoietic malignancies (HMs). These genomically unstable cancers exhibit increases in somatic mutation rates as well as cases of large-scale chromosomal aberrations and translocations. In my thesis work, I investigated alterations in the homologous recombination (HR) DNA repair pathway, centered on the *ATM-CHEK2-BRCA1* response to double strand breaks (DSBs). My work addresses the mechanisms of genomic instability that leads to an HM due to alterations in HR repair and effects on the cellular response to DNA damage and replication stress. The findings from my thesis work highlight the dual role that HR-associated factors play at stressed replication forks and implicate replication-mediated DNA damage in the etiology of genomic instability in hematopoietic cells. I investigated the mechanisms of genomic instability leading to recurrent *Tcra*/*Myc-Pvt1* translocations in a T-cell leukemia model with aberrantly stabilized β-catenin. I show that DSBs in the *Tcra* site of the translocation are Rag-generated whereas the *Myc-Pvt1* DSBs are not. I find that aberrant activation of β-catenin in thymocytes leads to a Tcf-1 mediated downregulation of HR-pathway member that promotes the retention of replication-mediated DSBs, providing the conditions for translocations to form. I also investigated a mouse model with hematopoietic-specific knockout (KO) of *Brca1* that produces bone marrow failure and HMs with widespread chromosomal aberrations. My investigations into the mechanism of this genomic instability in the absence of this central HR factor implicate replication fork restart failures and the use of more error prone backup pathway for DSB repair,

including non-homologous end joining (NHEJ) and alternative end joining (alt-EJ). Importantly, I show that bone marrow that is heterozygous for *Brca1* also shows mild deficiencies at stressed replication forks and increased expression of NHEJ and alt-EJ factors. Finally, I investigated germline variants in the cell cycle regulator, *CHEK2*, for their contribution to increased risk for developing an HM. My work helps to identify two *CHEK2* alleles, c.470C>T/p.I200T and c.1283C>T/p.S428P, as increasing the odds of developing an HM for patient carriers. Furthermore, I use a mouse model of the *CHEK2* p.I200T allele and show that these mice develop leukocytosis, clonal hematopoiesis, and HMs at late stages. This suggest that variants in *CHEK2* can alter the proliferation rates and somatic mutation rates in hematopoietic cells, contributing to genomic instability and outgrowth of bone marrow clones. Taken together, my studies highlight the dual role that HR factors play in repairing DSBs and in managing replication stress. I show how altered function can lead to failures of both replication fork protection and DSB repair, which act synergistically to increase genomic instability in these cells. These findings contribute additional context both to our understanding of current risks for carrying a mutation in these genome maintenance genes, and also opening up new therapeutic targets for treatment of HR-deficient HMs.

CHAPTER I

Introduction

Genomic instability is a well characterized hallmark of cancer that drives both malignant transformation and cancer cell plasticity (Hanahan and Weinberg, 2011). As DNA is the template for all cellular machinery, including transcription and replication, maintaining the fidelity of DNA is essential for viability. There are many endogenous and exogenous threats to genome integrity, leading to upwards of 70,000 lesions per day that a cell must sense and respond to (Lindahl and Barnes, 2000). Therefore, mammalian cells have evolved a diverse array of context specific and sometimes overlapping mechanisms, collectively known as the DNA damage response (DDR), for fixing or eliminating damaged cells to prevent the acquisition of mutations.

Double strand breaks and the DNA damage response

One of the most dangerous forms of DNA damage is double strand breaks (DSBs), which are formed when both strands of the DNA duplex are broken simultaneously. Failures of DSB repair pathways contribute to a variety of developmental, immunological, and neurological disorders in addition to being a major source of mutations in cancer (Jackson and Bartek, 2009). Below, I will discuss the canonical DNA damage response to DSBs and the regulation of pathway choice in mediating repair. Importantly, I focus on how pathway selection contributes to error rates and likelihood for acquisition of mutations.

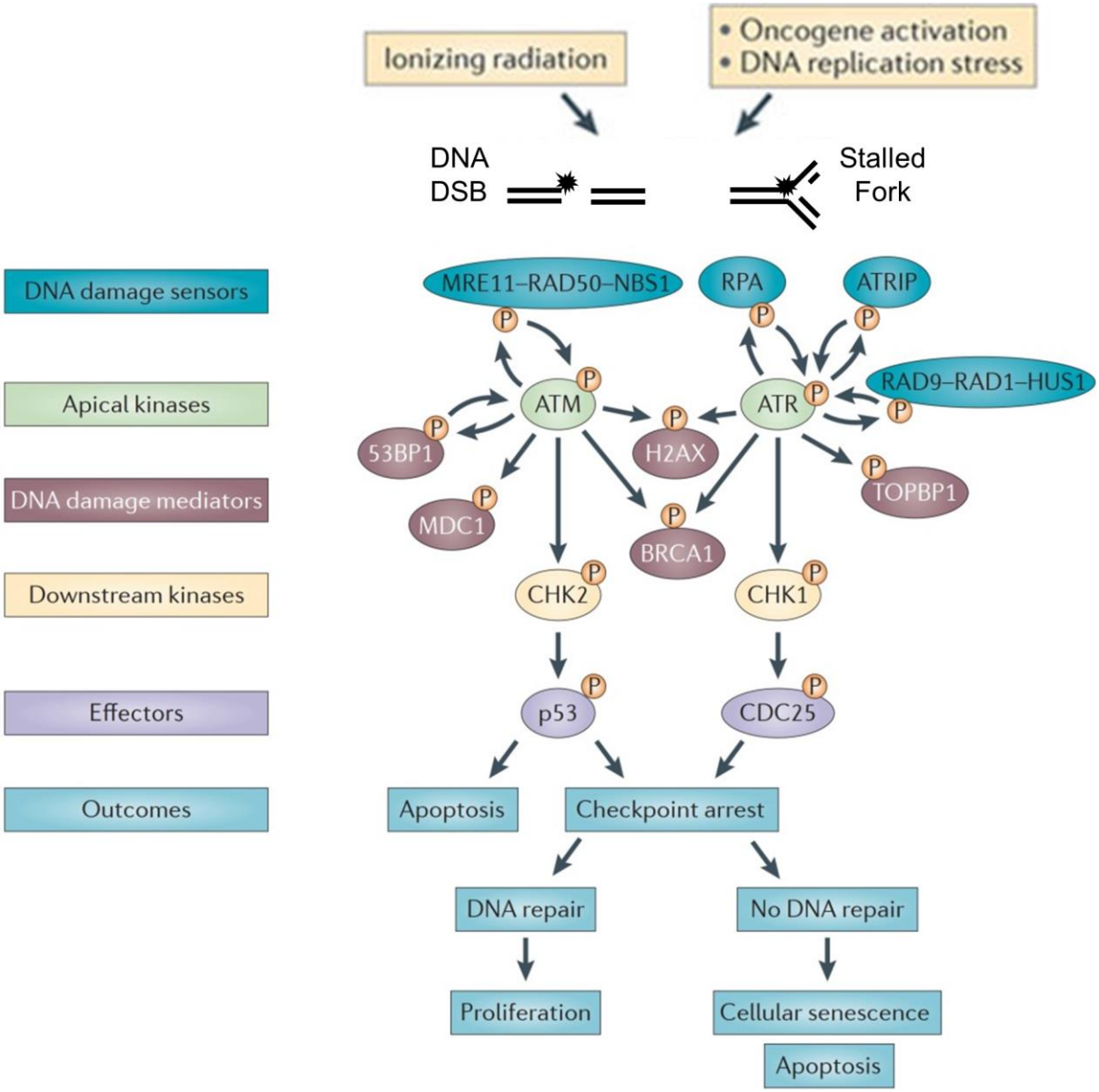
DNA damage response and replication stress signaling

The basic mechanics of the DDR response to DSBs involves sensing of the damage, transduction of a signaling response, and the engagement of cellular machinery for cell cycle arrest, DNA repair, or apoptosis (Figure 1.1). One of the earliest events after sensing DSBs is the activation of the ATM kinase, which initiates the DDR by phosphorylating histone H2AX (γH2ax) and other downstream targets. γH2ax often serves as a marker of DSBs in chromatin and is the site of recruitment for DDR proteins (Rogakou et al., 1998). ATM also phosphorylates and activates the effector protein CHEK2, which in turn phosphorylates many downstream targets to carry out the DDR (Matsuoka et al., 1998). A parallel and sometimes overlapping pathway is mediated by ATR and CHEK1 in response to replication stress and single stranded DNA (ssDNA) at stalled forks, however there can be some cross talk and overlap between these two pathways in response to DSBs (Sulli et al., 2012) (Figure 1.1). The checkpoint kinases CHEK1/2 are the transducers of DDR signaling and play a key role in arresting the cell cycle when damage is present to allow sufficient time for DNA to be repaired before continuing replication and cell division, and thus preventing the transmission of DNA DSBs. If the DDR cannot resolve the damage, the cells are then eliminated through apoptosis or senescence, often mediated by TP53 (Figure 1.1).

DNA repair pathways for double strand breaks

There are two main DDR pathways that are used to repair DSBs, known as homologous recombination (HR) and non-homologous end joining (NHEJ) (Ceccaldi et al., 2016a). Importantly, these two pathways have different error rates, and they

Figure 1.1 ATM and ATR signaling in response to DSBs and replication stress



Schematic of the DNA damage response (DDR) to double strand breaks (DSBs). There are two main sensors including the MRN (MRE11-RAD50-NBS1) complex for DSBs and the 9-1-1 (RAD9-RAD1-HUS1) complex and RPA for replication stress and ssDNA. These sensors active the kinases ATM and ATR, which phosphorylate a number of downstream targets to mediate the DDR, including BRCA1 for HR and 53BP1 for NHEJ. Additionally, ATM and ATR activate downstream kinases CHEK1 and CHEK2, which control cell cycle and cell survival by activating effectors such as p53 and CDC25A. DDR-mediated cellular outcomes may be cell death by apoptosis; transient cell cycle arrest followed by repair of DNA damage and resumption of proliferation; or cellular senescence caused by the persistence of unrepaired DNA damage. (Adapted from Sulli et al., 2012)

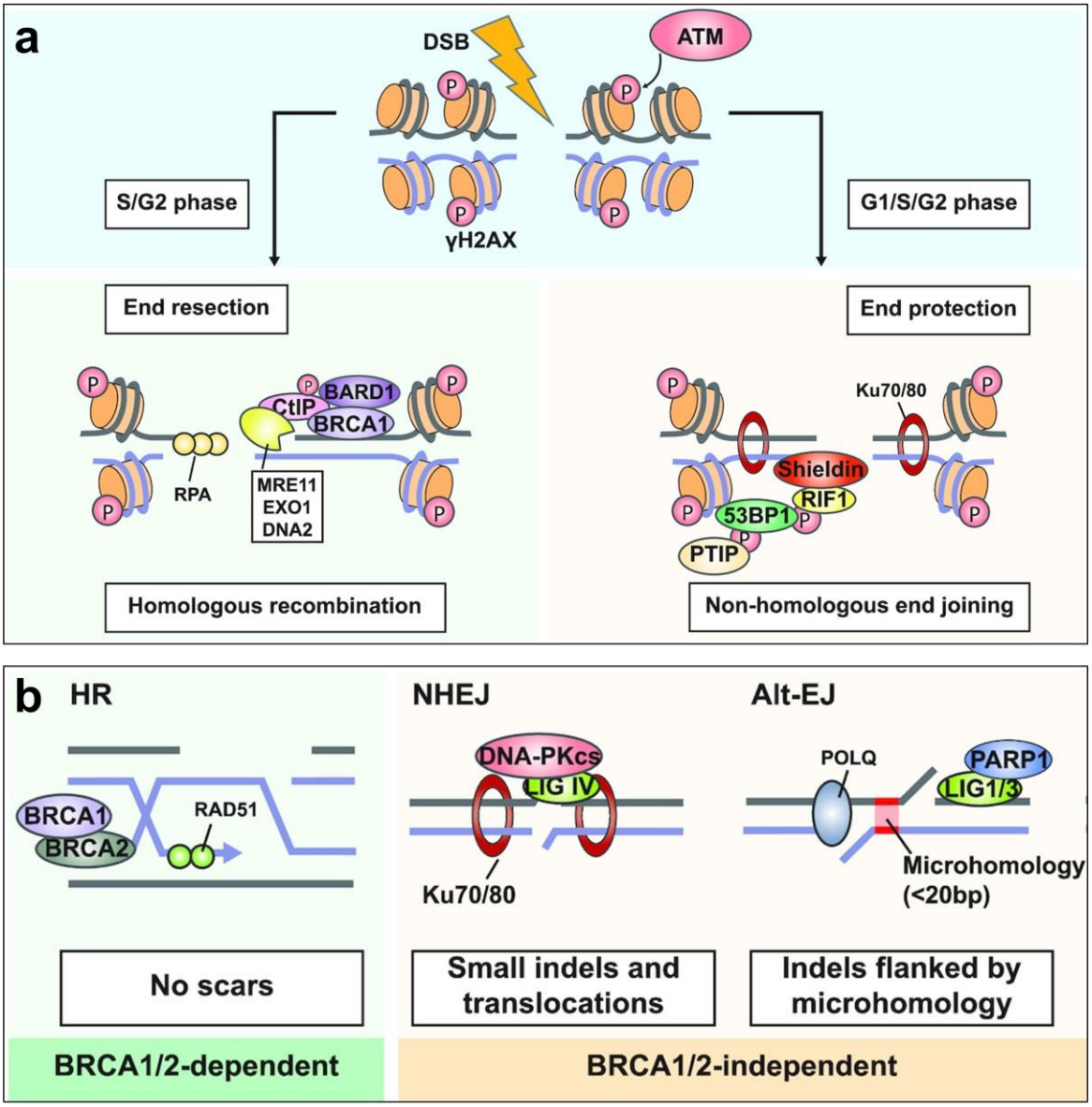
compete for DNA ends in a cell-cycle regulated manner. HR is considered to be error free, as it is only active in S-phase when a sister chromatid can serve as a template to mediate repair. In contrast, NHEJ acts at all phases of the cell cycle and is considered to be error prone, as it directly ligates DSB ends that are in proximity leading to small insertions, deletions, and substitutions at the break sites (Figure 1.2). Chromosomal translocations can arise from failures of HR during strand invasion when complex DNA structures are formed as well as from aberrant NHEJ if DSBs from disparate genomic loci are ligated together. Accordingly, loss of HR factors and the increased use of NHEJ increases baseline mutations rates, which contributes to genomic instability.

Alternative end joining (alt-EJ) is a third DSB repair pathway that has been described more recently, and is thought to serve mainly as a back-up pathway to HR, and to a lesser extent, NHEJ (Yousefzadeh and Wood, 2013). Whether alt-EJ plays normal roles in DNA repair in cells with a full complement of DDR programs remains unclear. The alt-EJ pathway is mediated by POLQ (also known as POLθ) and the use of internal microhomologies (<20bp) in the process of ligating DSB ends (Roerink et al., 2014; Wood and Doublié, 2016) (Figure 1.2b). This process also involves both PARP1 and Ligase 1 or 3, and is considered to be highly error prone due to homology search, flap trimming after ligation, and template insertions (van Schendel et al., 2016).

DNA damage pathway choice is mediated by an end-resection step

The decision regarding DNA repair pathway use at DSBs is regulated by an end-resection step. For NHEJ activity, the KU70/80 heterodimer rapidly binds blunt DSB ends and protects them from nucleolytic processing in cooperation with 53BP1 and the

Figure 1.2 DNA repair pathway choice for double strand breaks is mediated by cell cycle and end resection



a. DSBs are first recognized by ATM, which phosphorylates atypical histone H2AX, making γH2ax. DSB repair pathways are recruited to γH2ax in chromatin to mediate repair in a cell-cycle specific manner. HR and NHEJ compete for DNA ends and pathway choice is determined by an end-resection step. In NHEJ, the Ku70/80 heterodimer rapidly binds DSB ends and prevents nucleolytic processing in conjunction with the Shieldin complex. During S-phase, HR-machinery evicts NHEJ proteins and recruits the MRN complex to facilitate end resection, which is required before RAD51-mediated strand invasion to access the template strand for HR. **b.** Core components of the 3 known DSB repair pathways and the relative error-rates associated with repair activity. (Adapted from Stok et al., 2021)

Shieldin complex (Dev et al., 2018; Shao et al., 2012) (Figure 1.2a). During S-phase, the NHEJ machinery is evicted from DSB ends, and the MRN/CtIP complex is recruited in conjunction with BRCA1 (Escribano-Díaz et al., 2013). MRE11 then initiates 3' to 5' end resection, which is required to produce ssDNA to serve as the invading strand for HR activity (Chapman et al., 2012). This prevents further activity by the NHEJ machinery, which has low affinity for ssDNA (Dynan and Yoo, 1998; Scully et al., 2019). Importantly, alt-EJ can also act on resected ends, which may have important consequence in HR-deficient cells (*e.g.,* those with germline *ATM*, *CHEK2*, and *BRCA1/2* mutations) on DDR pathway choice and mutagenesis rates (Mateos-Gomez et al., 2017). Given the different error rates of repair programs, pathway selection is related directly to mutagenesis rates and contributions to genomic instability. Therefore, failure of DNA repair pathways or their regulation by altered cell cycle signaling is a common precursor to cancer development.

Dysregulation of DNA repair genes leads to genomic instability and cancer

Germline mutations in DNA repair genes increase risk for cancer

The role of DDR pathway members in health and disease is exemplified by inherited germline mutations in key genome maintenance genes. Errors in DDR pathways are common underlying events at both the germline and somatic levels that contribute to cancer development and other forms of disease (Ciccia and Elledge, 2010). Germline mutations in the ATM-CHEK2-BRCA1 axis are frequently associated

with elevated rates of cancer development, and hematopoietic malignancies (HMs) are often among the tumor spectrum in these patients.

For example, Ataxia telangiectasia (AT) is an autosomal recessive developmental disorder that occurs in individuals who inherit two pathogenic variants in *ATM* (Hall et al., 2021; Scott et al., 2002). This disorder is characterized by neurological and immune disfunction as well as a 37-fold increase in susceptibility to cancer, the most frequent of which are leukemias and lymphomas (Olsen et al., 2001). Additionally, individuals who inherit a single pathogenic variant in *ATM* have a 6.8-fold higher risk for developing malignancies, which also include cases of leukemia and lymphoma (Swift et al., 1990). However, it remains unclear whether HMs in these patients are causatively associated with AT risk alleles or are independent events in aging individuals, and additional studies will be needed to clarify this relationship. Furthermore, heterozygous germline pathogenic variants in *TP53*, perhaps the most well described tumor suppressor in human cancer, cause Li Fraumeni syndrome, an autosomal dominant cancer predisposition syndrome. These patients experience high cancer rates as part of the phenotype, including leukemias and lymphomas (Varley, 2003). *TP53* is also the gene most commonly somatically mutated in human cancer (Kandoth et al., 2013).

Pathogenic germline variants in *CHEK2* are also inherited in an autosomal dominant pattern and confer risk to a variety of cancers. For example, deleterious alleles such as del1100C and I200T are associated with development of colon, prostate, and lobular breast cancers (Filippini and Vega, 2013; Liang et al., 2018; Liu et al., 2012; Smith et al., 2010; Wang et al., 2015). Although the association with HMs has been less clear, the *CHEK2* p.I200T allele has been linked to HMs, particularly in Eastern

European populations (Cybulski et al., 2004; Janiszewska et al., 2012, 2018). The association of *CHEK2* mutations with HMs is addressed in more detail in Chapter IV.

Finally, *BRCA1* is a well-known tumor suppressor that has been causatively linked to genomic instability in a variety of human cancers. Pathogenic germline variants in *BRCA1* are associated with greatly elevated lifetime risk for breast (50-70%) and ovarian (40-60%) cancer as well as lesser risk for stomach, pancreas, prostate, colon, and other cancers (Hu et al., 2021; King et al., 2003; Thompson et al., 2002). Risk for HM in carriers of germline *BRCA1* variants has been more controversial (Friedenson, 2007, 2016; Iqbal et al., 2016). Nevertheless, work from our laboratory and others has now clearly shown a role for *BRCA1* in restraining HMs, as hematopoietic knock-out produces bone marrow failure (BMF) and HMs in mouse models (Mgbemena et al., 2017; Vasanthakumar et al., 2015). *BRCA1* is now recognized as a Fanconi Anemia (FA) like gene and is increasingly implicated in HMs, which is discussed in more detail specifically in Chapter V.

Therefore, the entire signaling axis for HR responses to DSBs has been implicated in human cancers, including HMs, where mutations in genome maintenance pathways contribute to genomic instability and tumorigenesis. DDR pathways may also be lost or silenced due to the selective pressure from high rates of proliferation and DNA synthesis in cancer cells, further contributing to genomic instability during cancer progression (Burrell et al., 2013; Cahill et al., 1999). Importantly, decades of investigation have associated a wide array of mutations in DDR genes with risk for various cancer, including some cases where different pathogenic variants in the same gene can confer risk for transformation in different tissues. (Chapman et al., 2012). The

nuanced differences between these deficits can also be exploited to understand DDR pathway members in their normal roles and functions further.

Clonal hematopoiesis and the contribution of DNA repair genes

Clonal hematopoiesis (CH) is characterized by the outgrowth of somatic clones due to the acquisition of mutations in hematopoietic stem or progenitor cells that provide a competitive growth advantage. The set of mutated genes that confer such a proliferative expansion have been identified in several large-scale studies, the most frequent of which are *DNMT3A, TET2, ASXL1, PPM1D, JAK2, SF3B1, SRSF2, and TP53* (Bick et al., 2020). Interestingly, this gene set includes those encoding epigenetic modifiers, splicing factors, and DNA repair genes. Although it is possible that CH represents a similar degree of somatic mosaicism that is seen across many tissues, the presence of these particular CH clones is associated with increased risk for developing an HM as well as cardiovascular diseases, such as strokes and heart attacks (Jaiswal and Ebert, 2019). Nevertheless, not all CH will progress to an HM, and the conditions contributing to transformation remain unclear.

Extensive work has focused on the mechanisms by which these different CH-associated mutations promote clonal expansions, which has produced different mechanistic paradigms depending on the type of mutated gene. For epigenetic modifiers such as *DNMT3A* and *TET2*, recent studies suggest that mutations in these genes lead to enhanced self-renewal capacity and maintenance of more stem-like characteristics in aging hematopoietic stem and progenitor cells (HSPCs) (Köhnke and Majeti, 2021). Additionally, inflammation appears to promote the expansion of clones

with these mutations (Bowman et al., 2018). In contrast, mutations in DNA repair genes such as *PPM1D* and *TP53* are thought to provide a survival advantage leading to clonal expansions. Interestingly, CH that is detected after prior cytotoxic treatments is enriched for *PPM1D* and *TP53* clones (Coombs et al., 2017). Importantly, although initial hypotheses posited that these mutations in CH genes may be caused by DNA-damaging therapeutics themselves, the mutational profiles of expanded clones in this setting suggested that these somatically mutated CH clones were present prior to treatment and were clonally selected (Wong et al., 2015). Indeed, CH clones with mutations in DNA repair genes *PPM1D* and *TP53* were more likely to expand after cytotoxic exposures, whereas *DNMT3A* clones showed expansion after autologous transplantation (Wong et al., 2018). This study also identified additional DDR genes with somatic mutations in patients previously treated with cytotoxic therapy who subsequently developed CH. Taken together, these data suggest that DNA repair associated CH has a different mechanism and kinetics than other CH clones, which may inform risk for progression to an HM, or the subtype of HM that is likely to develop.

DNA repair genes may also play a role in the predisposition for developing a CH clone in the first place. Recent large-scale studies have attempted to identify factors that are associated with the acquisition of CH, the most well characterized of which is age, with more CH detected in older individuals (Jaiswal et al., 2014). Two recent studies with >90,000 individuals sought to identify germline variants that are associated with CH. Interestingly, some of the most frequently occurring variants were in the telomere maintenance gene, *TERT*, and variants in DDR genes were also found, including *ATM* and *CHEK2* (Bick et al., 2020; Kar et al., 2022; Silver et al., 2021). Therefore, DNA

repair genes likely play an important role in the etiology of CH and potentially the progression to HMs, with DDR-associated CH clones potentially serving as biomarkers of previous cytotoxic therapeutic exposure and differential risk for developing HMs.

Chromosomal aberrations are common features of hematopoietic malignancies

There are many forms of genomic instability that can arise when DNA repair pathways fail. Although failures in DSB repair are more likely to produce translocations, the vast majority of DNA damage in cellular genomes are single base changes and single stranded breaks (Nussenzweig and Nussenzweig, 2010). Importantly, the types of DNA damage can be used to classify cancer types and treatment plans. Furthermore, signatures of DNA damage are now being used to correlate mutation profiles with particular DDR pathway failures, providing additional mechanistic insight to their normal use and function (Alexandrov et al., 2013, 2020). Therefore, the presence of a particular type of damage repair product, such as a chromosomal translocation or mutation signature, can provide insight into the type of DNA damage that occurs in those cells.

Chromosomal translocations, which are the products of failed or aberrant DDR responses, are a common feature of genomic instability in human cancers. Frequently, balanced chromosomal translocations can lead to tumorigenesis by placing oncogenes under the control of highly active promoters, or by the production of fusion proteins that escape normal regulatory controls on expression (Nussenzweig and Nussenzweig, 2010). For example, the Philadelphia chromosome was the first described balanced chromosomal translocation as a cause for cancer, which was discovered in part at the University of Chicago by Dr. Janet Rowley (Rowley, 1973). In patients with chronic

myeloid leukemia (CML), a translocation between chromosomes 9 and 22 (t(9;22)(q34;q11)) produces a fusion product known as BCR-ABL, which leads to constitutive Abelson kinase activity that contributes to transformation (Salesse and Verfaillie, 2002). Although chromosomal translocations were thought to be overrepresented in HMs, the improvement of sequencing technologies has now identified many translocations in solid tumors as well (Mertens et al., 2015). Nevertheless, it remains unclear how many of these NGS-detected translocations are truly pathogenic or chance events in established malignancies. Furthermore, although there has been extensive work on the etiology of translocations, particularly in HMs, the complete mechanistic picture remains to be defined.

Importantly, chromosomal aberrations are frequently pathogenic in HMs (Küppers, 2005; Mitelman et al., 2007) (Table 1.1). Additionally, the presence of particular translocations and cytogenetic abnormalities has long been used to classify leukemias and design treatment plans, particularly in acute myeloid leukemias (AML) (Grimwade and Hills, 2009; Mrózek et al., 2004). Although extensive work has identified many recurrent translocations in these patients, the precise mechanisms that create the conditions for translocations to form remains incomplete. Importantly, the presentation of larger scale chromosomal aberrations in leukemias may provide insight into the type of DNA damage that occurs in hematopoietic cells. Additionally, continued investigation into the etiology of translocations can provide important diagnostic and treatment value as well as provide mechanistic insight into basic biological processes for maintaining genomic stability.

Table 1.1 Frequencies of balanced chromosome aberrations and gene fusions in cancer

Diagnosis	Number of balanced aberrations	Number of recurrent balanced aberrations	Number of gene fusions	Estimated proportion with gene fusions
Haematological disorders				
Acute myeloid leukaemia	1,785	267	109	20%
Myelodysplastic syndromes	498	54	28	<1%
Chronic myeloid leukaemia	750	152	15	100%
Chronic myeloproliferative disorders	194	17	19	<1%
Acute lymphoblastic leukaemia	1,139	155	82	30%
Mature B-cell neoplasms	1,713	227	69	30%
Mature T-cell neoplasms	425	21	20	15%
Hodgkin lymphoma	63	2	5	<1%
Malignant solid tumours				
Respiratory system	282	3	2	<1%
Digestive system	435	11	2	<1%
Breast	343	13	3	<1%
Female genital organs	176	9	3	<1%
Male genital organs	41	0	4	80%
Urinary tract	225	7	7	<1%
Endocrine system	39	3	15	35%
Nervous system	412	9	1	<1%
Skin	238	5	0	<1%
Bone	216	3	7	15%
Soft tissues	342	16	33	20%

The numbers of balanced aberrations, recurrent balanced aberrations and gene fusions based on published neoplasia-associated karyotypes and their corresponding gene fusions. The estimates in the last column take into consideration the relative frequencies of all morphological tumor entities within each organ system. (Mitelman et al., 2007)

Sources of double strand breaks involved in translocations in human cancer

Sources of DSBs that serve as partner loci in translocations include both exogenous and endogenous sources. Exogenous sources involve both environmental exposures and therapeutic interventions such as ionizing radiation (Chapman et al., 2012). Endogenous sources include programmed activity, as with RAG-mediated recombination of *BCR* and *TCR* loci and AID-mediated somatic hypermutation of *BCR*

loci during lymphocyte development (Nussenzweig and Nussenzweig, 2010). Alternatively, spontaneous damage can occur during normal cellular activities (Tubbs and Nussenzweig, 2017). One reason for the frequency of translocations in HMs is that lymphocytes are unique among somatic cells because their development requires the controlled generation of DSBs for rearranging receptor loci for adaptive immunity. In both B and T cell development, RAG recombinases generate controlled DSBs to recombine the T cell receptor (TCR) and immunoglobulin (Ig) loci. Below, I will detail the programmed and spontaneous sources of endogenous DSBs that can serve as substrates for translocations to form, with a focus on T-cell development as an example of RAG-mediated activities and cell cycle regulation.

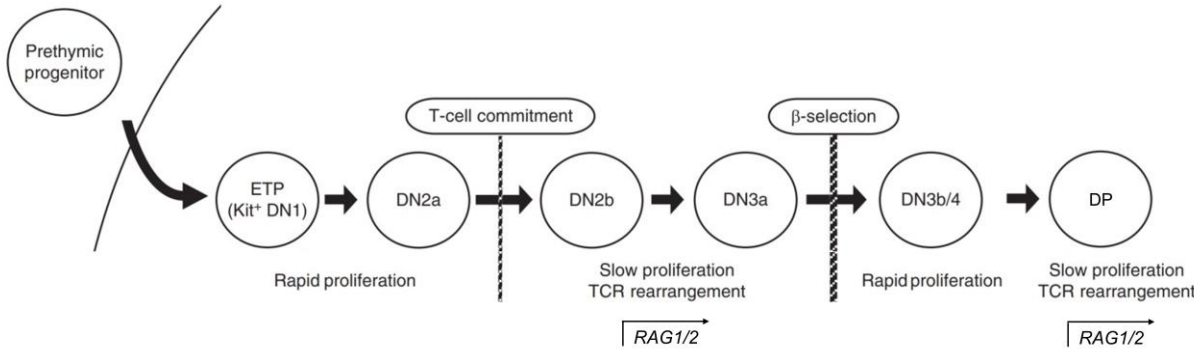
Thymocyte development requires coordination of proliferation and receptor rearrangements

T-cell development occurs in the thymus and is a highly regulated progression through multiple precursor stages as cells progressively acquire T-cell characteristics and lose multipotency. These phenotypically distinct precursor stages can be defined by the surface expression of CD4 and CD8, with the earlier stages lacking these markers (double-negative, DN). Relatively small numbers of early thymic progenitors (ETP/DN1, KIT^+ $CD44^+$ $CD25^-$) enter the thymus, which is followed by a proliferative burst as these cells receive Notch signaling from the thymic cortical environment (Rothenberg, 2000). This provides a larger number of DN2 cells (KIT^+ $CD44^+$ $CD25^-$) to enter the next phase of thymic development where lineage commitment to T-cell fate begins, which is controlled by the expression of *BCL11B* in addition to other lineage-specifying

transcription factors (Ikawa et al., 2010). Importantly, recombination of *TCR* loci is initiated during the DN3a stage (KIT^- $CD44^-$ $CD25^+$), which is marked by low proliferation rates to facilitate the expression of RAG-recombinases in the G1 cell cycle phase (Hosokawa and Rothenberg, 2018). The mature αβTCR is made up of the TCRα (encoded by *TCRA*) and TCRβ (encoded by *TCRB*) chains, which both undergo VDJ-recombination for diversification of the receptors (discussed in more detail below). At the DN3a stage, only the TCRβ chain is recombined, which is then assembled with an invariant pre-TCRα chain to make the pre-TCR. Cells that productively rearrange the TCRβ chain and assemble a pre-TCR on their surface initiate progression to the DN3b/4 stages (KIT^- $CD44^-$ $CD25^-$)(Hosokawa and Rothenberg, 2018). Importantly, this is followed by a burst of rapid proliferation that is critical for the full phenotypic differentiation at later stages and increasing potential diversity in TCRα/TCRβ chain pairs (Kreslavsky et al., 2012). Cells then progress to the double positive (DP, $CD4^+CD8^+$) stage where once again cell cycle must be arrested to facilitate RAG expression and the recombination of the TCRα chain. DP cells can then assemble the mature αβTCR on their surface. This is followed by testing the αβTCR through the processes of positive and negative selection and eventual maturation to CD4 or CD8 single positive cells (Shah and Zuniga-Pflucker, 2014). Although there are many additional features of the T-cell developmental program, here I have focused on the proliferative capacities of these developmental precursors to highlight the cell cycle regulation required in the context RAG-expression (Figure 1.3). The balance between proliferation and cell cycle arrest for VDJ recombination is critical for proper thymic developmental programs. The details of RAG-mediated VDJ-recombination and the

coordination with NHEJ DNA repair factors is discussed below.

Figure 1.3 Thymocyte developmental progression and proliferative dynamics

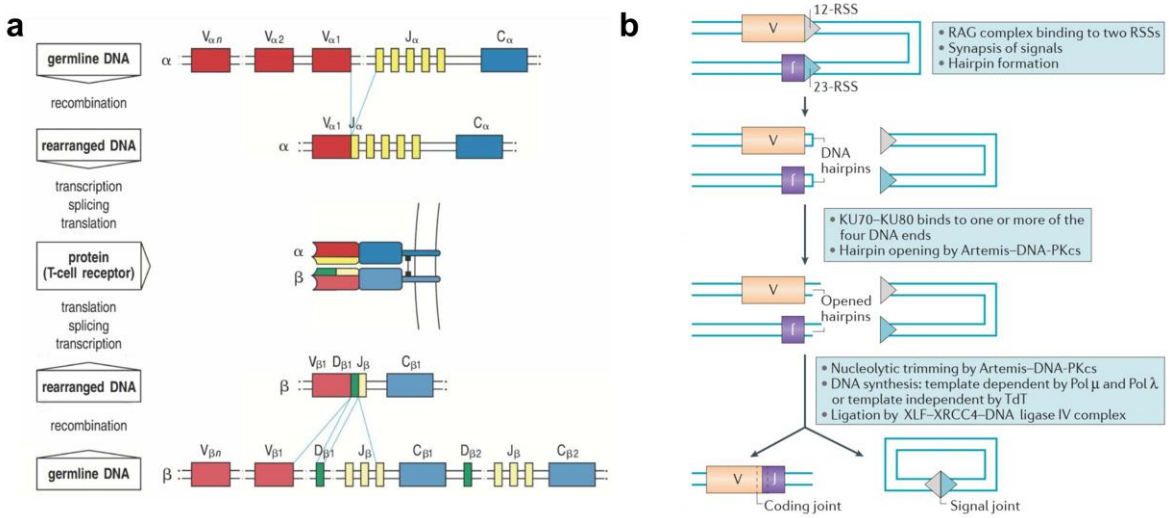


Schematic of thymocyte development focused on proliferation and receptor recombination. A small number of prethymic progenitor cells migrate to the thymus and become early thymic progenitors (ETP). A rapid burst of proliferation provides large numbers of DN2a cells (double negative, $CD4^{-}CD8^{-}$). After T-cell commitment, cell cycle slows to facilitate *RAG1/2* expression to begin rearranging the TCRβ chain in the DN2b-DN3a stages. Successfully recombined TCRβ chains are assembled into the pre-TCR and undergo β-selection where successful pre-TCR signaling leads to progression to the DN3b/4 stage. This is followed by a second burst of proliferation prior to progression to the DP (double positive, $CD4^{+}CD8^{+}$) stage. DP thymocytes again arrest cell cycle to facilitate *RAG1/2* expression for rearranging the TCRα chain. (Adapted from Hosokawa and Rothenbeg, 2018)

Programmed DNA double strand breaks during lymphocyte development.

The TCRα and TCRβ chains are made up of multiple gene segments that involve rearrangement and selection of a single variable (V), joining (J), and, in TCRβ chains, diversity (D) gene segments from a large number of possible gene segments in each region (Figure 1.4a). For example, there are 61 J gene segments in the *TCRα* locus and a single segment will be recombined with one of ~70 V segments. Each gene segment is flanked by recombination signal sequences (RSSs) that are made up of a conserved nonamer and heptamer sequence with either a 12 or 23 base pair (bp) spacer. RAG recombinases can bind these RSSs and bring a V and J region together, where

Figure 1.4 VDJ recombination in T-cells for *TCR* assembly



a. Schematic of VDJ recombination for assembly of the TCR. For the TCRα chain (top) a single V_α gene segment is recombined with a single J_α gene to produce a functional variable region exon. Transcription and splicing then joins the variable region with constant region (C_α) leading to expression of the TCRα chain. Similarly, the β-chain involves recombination of a V_β, D_β, and J_β gene segment to produce a variable region exon that is spliced with the C_β region to produce a TCRβ exon. The α- and β-chain proteins then pair to produce a fully assembled αβTCR. (Murphy, *Janeway's Immunobiology*, 2021). **b.** Schematic of RAG-mediated VDJ recombination. RAG binds a pair of 12-RSS and 23-RSS on V and J segments according to the 12/23 rule and mediates synapsis between gene segments. RAG then nicks RSS sites and produces DNA hairpins at each DSB end on the V and J segments, leaving blunt DSBs on the intervening DNA region. NHEJ machinery is activated at all four DSB ends. Hairpins on V and J segments are then cleaved by the DNA-PK/ARTEMIS complex. This is followed by NHEJ repair producing a coding join to complete a variable gene segment. (Adapted from Lieber, 2016).

directional orientation is maintained by always pairing a 12-bp RSS with a 23-bp RSS, known as the 12/23 rule. Once RAG binds, each RSS site is nicked and a trans-esterification step produces a hairpin on one side of the DNA duplex (coding end), and blunt DSB ends on the other (signal end) that is recognized and repaired by the NHEJ machinery. The coding end is also processed by NHEJ factors where the KU70/80 heterodimer binds, while the DNA-PK/ARTEMIS complex opens the hairpin producing DNA ends with ssDNA. The single strand overhangs of opened coding ends can

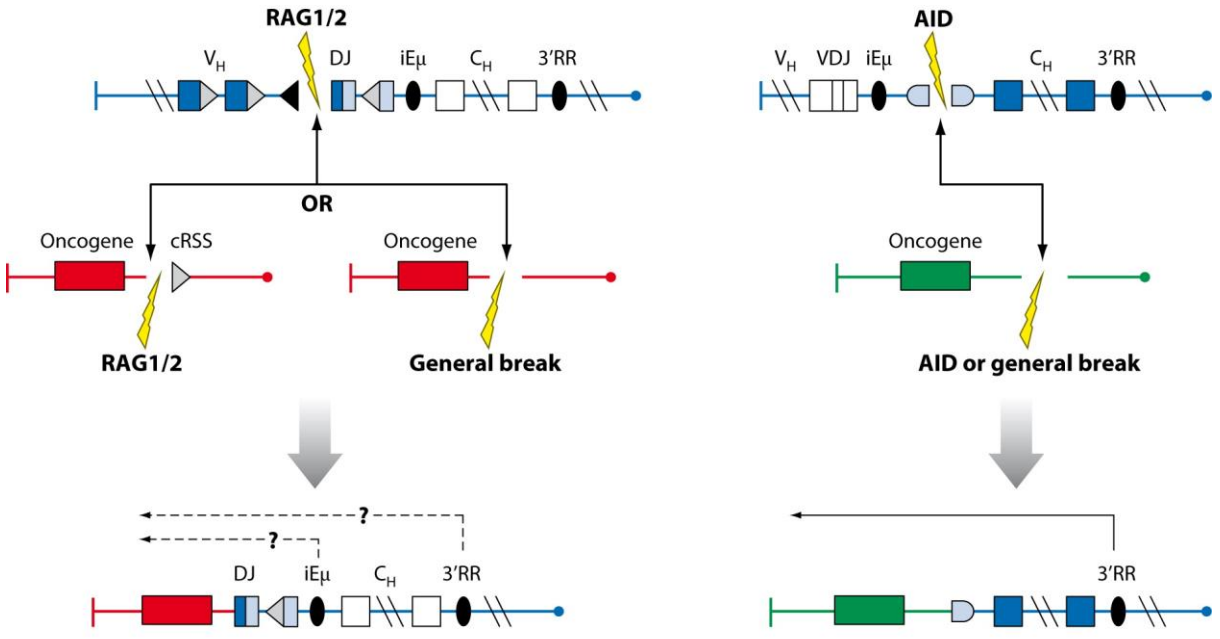
partially align between V and J segments, and Terminal deoxynucleotidyl Transferase (TdT) mediates gap filling with nontemplated (N) nucleotides, which greatly increases receptor diversity by the stochastic addition of bases in the join. Endonucleases also remove unpaired nucleotides, which also adds to receptor diversity. Finally, DNA ligase IV and XRCC4 cooperate to ligate the V and J segments together (Figure 1.4b). Therefore, as VDJ recombination produces DSB ends that require NHEJ mediated repair, the cell cycle must be strictly regulated to keep cells in G1 when NHEJ factors are active and RAG is expressed. Given the balance between proliferative bursts and VDJ-recombination during thymocyte development, failure of cell cycle regulators or use of alternative DNA repair pathways at RAG-mediated DSBs can contribute to translocation formation.

Although I have focused here on T-cell development, an analogous RAG-mediated process for VDJ recombination of the *Ig* light and heavy chains occurs in B-cells. Additionally, in B-cells, programmed genomic instability occurs through the activity of AID in somatic hypermutation and class switch recombination. Translocations in lymphocytes have been associated mechanistically with aberrant activity of RAG1/2 and/or AID (Gostissa et al., 2011; Lieber, 2016; Robbiani and Nussenzweig, 2013). Indeed, one of the two partner loci in chromosomal translocations in HMs frequently involves either an *Ig* or *TCR* region, suggesting that perturbations during lymphocyte development can produce DSB substrates for translocation formation.

Nevertheless, it is worth noting that two DSBs from disparate chromosomes are required for translocations to form. As noted above, off-target RAG or AID activity may induce DSBs elsewhere in the genome (Lieber, 2016; Nussenzweig and Nussenzweig,

2010) (Figure 1.5). For example, RAG activity at a cryptic RSS site in the *BCL2* locus has been implicated in the formation of *IgH-BCL2* translocations in follicular lymphoma (Vaandrager et al., 2000). Additionally, recurrent DSBs in non-*Ig* loci have been associated with AID activity, and DSBs in the *c-Myc* locus are AID-dependent in a model of Burkitt's lymphoma with *IgH-c-Myc* translocations (Qian et al., 2014; Robbiani et al., 2008). Nevertheless, there are also chromosomal translocations that appear to be entirely independent of RAG and AID. Additionally, translocations are also observed in

Figure 1.5 Contribution of RAG and AID activity to translocation formation



Schematic example of chromosomal translocations that lead to oncogene expression when fusions involve the *Ig* receptor loci. (Left) RAG generates DSBs during VDJ recombination; a second break near an oncogene occurs due to off-target RAG activity at a cryptic RSS (cRSS) or a spontaneous (general) DSB from another source. Oncogene expression is then activated by either the *Ig* intronic enhancer (iEμ) or the *Ig* regulatory region (3'RR) encompassing four longer range enhancer elements. (Right) AID-mediated DSBs during class switch recombination are joined with DSBs near oncogenes. Similarly, the second break may be spontaneous or due to off target AID activity at non-*Ig* loci. *Ig* enhancer elements then lead to aberrant expression of an oncogene. (Adapted from Gotissa et al, 2011)

high rates in myeloid cells, where *MLL* genes are frequently involved in oncogenic fusions. DSBs in the *MLL* loci have been associated with a variety of mechanisms, including the activity of topoisomerases. Indeed, a subset of therapy related myeloid neoplasms (t-MNs) is associated with prior treatment with topoisomerase II inhibitors, and these t-MNs correlate with the presence of MLL-involved translocations (McNerney et al., 2017). Whether topoisomerase activity is involved in *de novo* AMLs with MLL fusions remains unclear, however DSBs in *KMT2C* and *KMT2D* have also been associated with early replication fragile sites (Ray Chaudhuri et al., 2016), suggesting that these breaks could be derived from failures at the replication fork, where topoisomerases are active in relieving torsional strain from DNA unwinding. Therefore, although RAG and AID activity are sometimes responsible for non-*Ig* and non-*TCR* breaks, the mechanisms involved in DSB formation outside of the receptor loci that are involved in chromosomal translocations in HMs constitute an active area of research.

Spontaneous sources of endogenous DNA double strand breaks

Replication stress leading to fork collapse and breakage is the most common source of spontaneous endogenous DSBs (Nussenzweig and Nussenzweig, 2010). Replication forks slow or stall when encountering obstacles to processive replication. These include nicks, gaps, or DSBs in the template strand which stall the fork machinery (Zeman and Cimprich, 2014). Abasic sites or misincorporated nucleotides cause similar stalling (Dalgaard, 2012). Topological blocks to the replication machinery also include DNA-protein complexes, R-loops (DNA-RNA hybrids), torsional stress, and collisions of the replication machinery with a second converging fork or with

transcriptional machinery (Nussenzweig and Nussenzweig, 2010; Stok et al., 2021). Depletion of dNTP and histone pools, which can occur in highly proliferative cancer cells, also induce replication stress due to the lack of available biomaterials to synthesize new DNA (Anglana et al., 2003; Poli et al., 2012). Reactive oxygen species (ROS) have also been known to induce DSBs directly or alter bases in ssDNA leading to replication-stress (Yu and Anderson, 1997).

Interstrand crosslinks (ICLs) that prevent separation of DNA strands are a particularly toxic form of replication stress. Chemical exposures to compounds such as aldehydes and platinum-based chemotherapies also cause ICLs (Deans and West, 2011). ICLs must be removed or bypassed to complete replication, which involves different DNA repair machinery from that used for DSBs (Deans and West, 2011). Some ICLs are removed in G1 by the nuclear excision repair pathway (Deans and West, 2011). In S-phase, ICLs are engaged by the FA pathway for fork stabilization and unhooking of the damaged base (Deans and West, 2011). Importantly, this form of ICL removal may require an HR step for the final resolution and restart of replication (Ceccaldi et al., 2016b), which is discussed in more detail in later sections. Furthermore, the FA pathway is required for normal hematopoiesis, and germline mutations in FA genes are well known for predisposition to bone marrow failure, aplastic anemia, and HMs (Nalepa and Clapp, 2018). Moreover, prior work from our laboratory and others has established *BRCA1* as a member of the FA pathway (Ceccaldi et al., 2016b; Mgbemena et al., 2017; Vasanthakumar et al., 2015). Therefore, there is significant overlap with HR machinery and BRCA1 in the resolution of replication stress and replication-mediated DSBs, particularly at ICLs.

Replication stress response at stalled forks and potential for DNA breaks

Replication fork stalling and stabilization

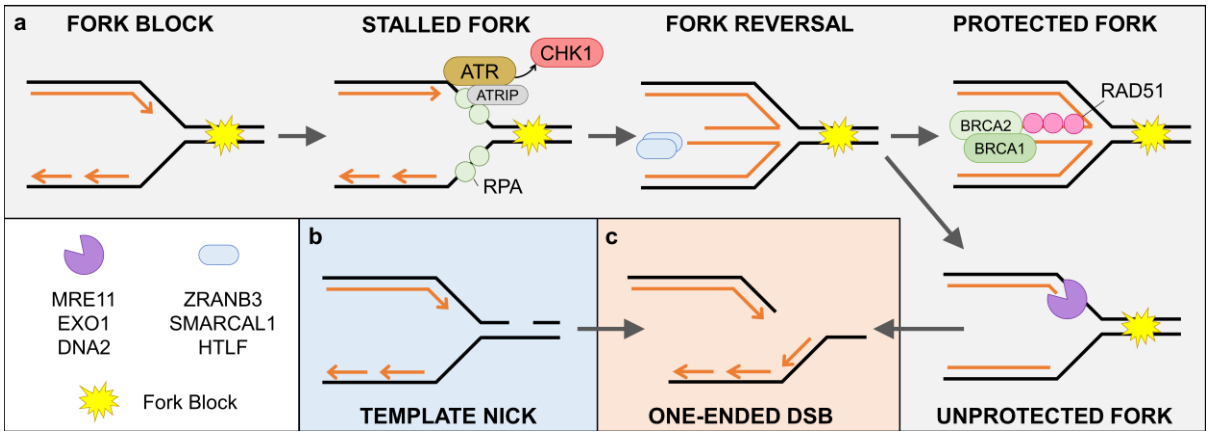
The need to replicate the genome requires opening the DNA helix which temporarily renders it more vulnerable to mutations and rearrangements. As replication forks progress, ssDNA and free DNA ends are produced, and nucleosomes are displaced, to provide access to the replication machinery. Therefore, DNA at replication forks is more accessible to nucleases and other potential forms of damage, which is exacerbated by replication blockages that stall forks in these open configurations. When replication forks stall, polymerases become uncoupled from helicases, which continue to unwind DNA. This leads to open stretches of ssDNA, which are bound quickly by RPA, thereby activating ATR (Zeman and Cimprich, 2014). ATR then mediates global slowing of S-phase, prevents new origins from firing, and helps regulate RPA pools (Neelsen et al., 2013; Syljuåsen et al., 2005; Toledo et al., 2013). One of the main targets for the ATR kinase is CHEK1, which mediates cell cycle arrest. Although these ATR and RPA activities can help stabilize a stalled fork, this is often not sufficient to rescue the fork and promote its restart.

Fork reversal is often an intermediate step that preserves nascent DNA strands at a stabilized fork by reversing them into a “chicken foot” structure that resembles a Holliday junction (Atkinson and McGlynn, 2009). This activity is performed by the helicases ZRANB3, SMARCAL1, and HTLF (Liu et al., 2020). Critically, this separates

the nascent strand from any potential damage in front of it on the template strand. Furthermore, this allows additional time for resolution of replication stress before restart.

Importantly, reversed forks form a one-ended DSB-like structure and are therefore subject to DDR responses and potential resection (Scully et al., 2019). Consequently, access to the nascent strand must be strictly regulated to prevent degradation by MRE11, DNA2, or EXO1 (Lemaçon et al., 2017). Over-activity of these nucleases at a regressed fork can lead to fork cleavage, generating a one-ended DSB (Figure 1.6a,c). Interestingly, loss of BRCA1 leads to increases in replication-mediated DSBs, which is rescued by deletion of ZRANB3, SMARCAL1, or HTLF (Quinet et al., 2017). Therefore, fork reversal can protect genome stability by preserving stalled forks, but they are also a danger to genomic integrity if not properly regulated. Additionally,

Figure 1.6 Replication stress can produce one-ended double strand breaks



Schematic of how replication stress can lead to a one-ended double strand break (DSB) **a.** A replication impediment leads to fork stalling, where open single stranded DNA (ssDNA) is bound by RPA. ATR and ATRIP then activate CHK1 for cell cycle arrest. Nascent strands at stalled forks are regressed and reversed into a “chicken foot” structure by ZRANB3, SMARCAL1 and HTLF. BRCA1/2 and RAD51 participate in protection of regressed forks. Excessive resection at unprotected forks by MRE11, EXO1, or DNA2 leads to strand breakage and a one-ended DSB. **b.** A ssDNA nick leads to a DSB when the replication machinery unwinds the template and hits the gap. **c.** Schematic of one-ended DSB from either gap or nuclease mechanisms.

nicks and ssDNA can turn into one ended DSBs when the replication machinery reaches the gap and falls off the template strand (So et al., 2017) (Figure 1.6b,c)

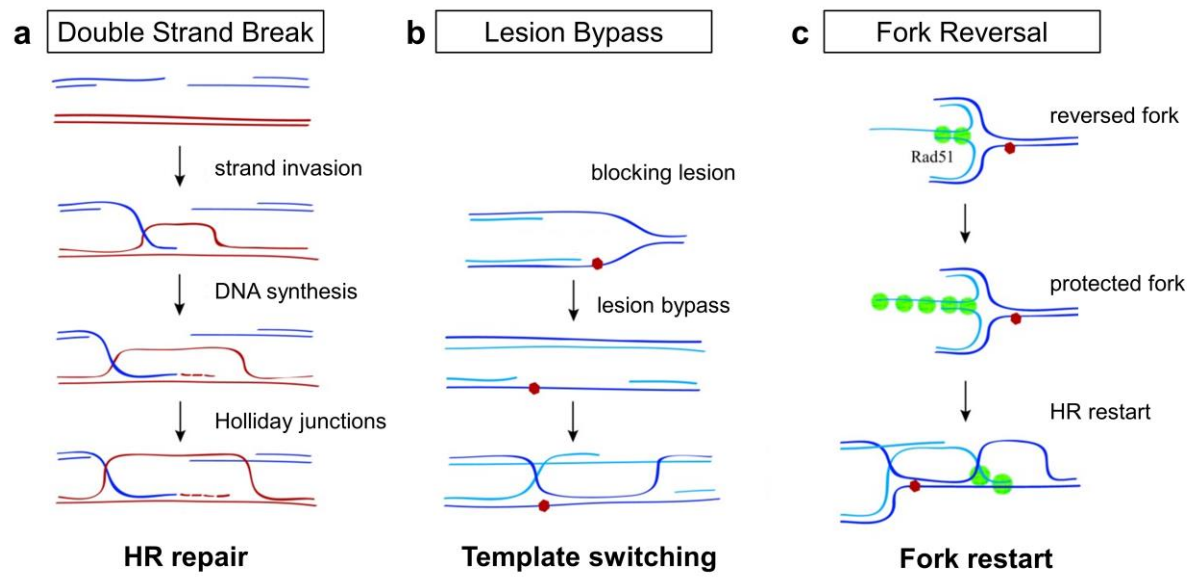
Replication fork collapse produces single ended DSBs

Replication fork collapse and breakage produce one-ended DSB structures, which serve as ideal substrates for translocation formation. Especially in HR deficiency when NHEJ predominates, the proximity of two DNA ends is important for re-ligation. Although two-ended breaks in a continuous DNA strand are usually topologically proximal, one-ended DSBs may be joined with another DSB anywhere in the genome. Especially in the context of widespread replication stress producing multiple one-ended DSBs, inter-chromosomal translocations may become more frequent.

Structural similarities to HR intermediates in replication stress responses

The recovery from stalled replication forks and DSBs shares additional structural features, providing some hints as to the replication functions for HR-associated proteins. Restarting replication after a blocking lesion is encountered often involves HR-proteins regulating the formation of a D-loop (displacement loop) leading to Holliday junctions. This includes replication forks forming HR-like recombination intermediates by utilizing a sister chromatid to circumvent fork blockages in a mechanism known as template switching (Figure 1.7b). Additionally, reversed forks may restart replication through strand invasion past the blocking lesion (Figure 1.7c). Therefore, in addition to one-ended DSB structures at stalled or collapsed forks, these bypass mechanisms also share structural features with HR-mediated DSB repair. In both cases, DSB end-like

Figure 1.7 Structural similarities to HR intermediates in replication fork stress and restart responses



Schematic of D-loop intermediates and Holliday junction formation. **a.** Traditional homologous recombination (HR) response to a DNA doubles strand break. Resected ends invade the homologous stand to restart DNA synthesis, ending in a double Holliday junction structure. **b.** Lesion bypass by template switching involving D-loop formation past blocking lesion. **c.** Blocking lesion leads to fork reversal into chicken foot structures, with Rad51 protecting nascent DNA end. Nascent DNA end strand invasion past the blocking lesion for HR restart past blocking lesion produces D-loop structure. (Adapted from Prado, 2018.)

structures are processed in two major steps that are necessary to initiate recombination-based repair that maintains genomic integrity. First, resection of the 5' strand at DSB ends by nucleases, including MRE11, EXO1, and DNA2, creates a ssDNA 3' overhang (Mijic et al., 2017; Thangavel et al., 2015). This simultaneously promotes HR-mediated repair and reduces NHEJ activity due to differential affinity for resected end structures (Chapman et al., 2012). In the second step, a nucleofilament is formed by loading RAD51 onto the 3' ssDNA, which is required for invasion of the homologous strand to allow for templated repair or fork restart. In both standard DSBs and replication fork restart processes, this formation of a D-loop intermediate is often

necessary to allow for error free repair and is recognized and processed by the HR-machinery or other DDR pathways (Daley et al., 2014; Prado, 2018).

D-loops can also be toxic structures if not properly regulated and resolved, which are frequently mediated by HR-associated proteins (Daley et al., 2014). Thus, DSBs and stalled forks share structural features that require the activity of HR proteins for resolution and maintenance of genome integrity. Consequently, mutations in genes involved in HR repair of DSBs also have deleterious effects on replication stress responses and are implicated increasingly in replication-mediated instability.

HR proteins have overlapping and separate functions in DNA damage and replication stress responses

In some cases, HR pathway members appear to perform the same function at both DSBs and stalled forks, but recent studies have also uncovered replication-specific functions for HR proteins. Excellent examples of this are BRCA1/2, which appear to have overlapping roles with HR-repair in loading RAD51 onto resected DNA ends to facilitate strand invasion, but they also play replication-specific functions in fork reversal and protection from over digestion by nucleases (Chen et al., 2018; Jones and Petermann, 2012; Prado, 2018). RAD51 also shares some of its known roles in HR-mediated DSB processing with replication fork stress responses in strand protection and invasion, but also appears to have functions in fork reversal that are independent of its enzymatic activity and strand exchange (Mason et al., 2019). Therefore, there is a need to resolve the shared and separate functions of HR proteins in DSB repair and

replication stress to understand risks to genomic instability when these pathways fail or when cells experience proliferative or damage stress.

Fanconi Anemia supports the overlap of DNA replication and damage stress programs in hematopoietic cells

FA is an autosomal recessive disorder that is characterized by BMF, congenital defects, cancer predisposition, and decreased ability to repair DNA ICLs (Brosh et al., 2017). FA results from biallelic germline mutations in any of 22 FANC genes, with a few exceptions, including X-linked *FANCB* mutations and dominant negative *FANCR*/*RAD51* mutations (Brosh et al., 2017). Some of the hematopoietic-related symptoms include anemia, leukopenia, and thrombocytopenia that can manifest as fatigue, easy bruising, and nose or gum bleeding. This can progress to BMF, as cellular deficiencies lead to the attrition of HSPCs. Although some of the clinical presentations of FA can vary based on the affected gene, one common feature is hypersensitivity to low dose mitomycin C (MMC) which induces chromone breakage and aberrant repair leading to multiradial chromosomes (Nalepa and Clapp, 2018). This is due to the central role of the FA pathway in repairing DNA ICLs that are induced by MMC or other crosslinking agents. Thus, hematopoietic cells harboring biallelic *FANC* gene mutations are especially sensitive to replicative stress induced by ICLs, which contributes to loss of HSPCs and hematopoietic dysfunction in these patients.

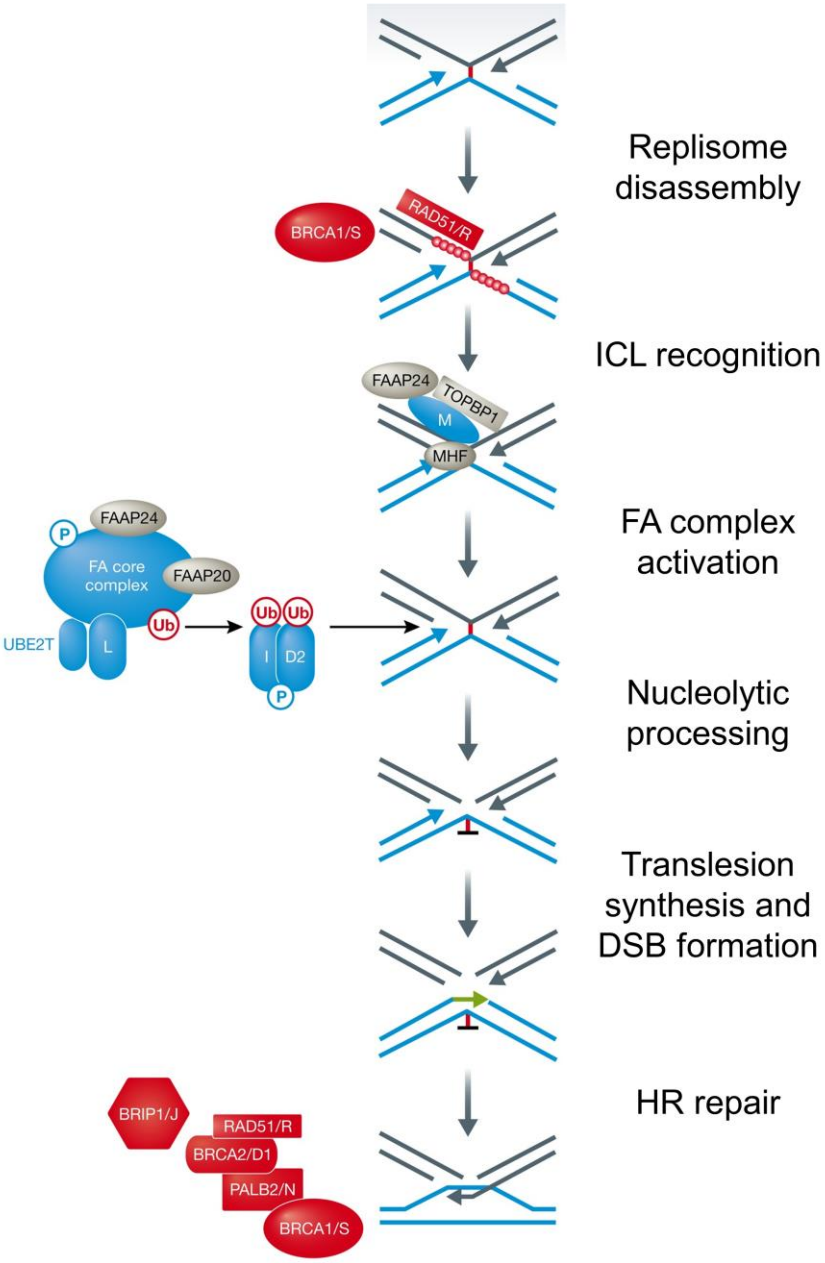
Inclusion of *BRCA1/FANCS* as a FA gene was more controversial as germline biallelic mutations in *BRCA1* are exceedingly rare. However, recent characterization of a woman with congenital abnormalities that harbored biallelic germline mutations in

BRCA1 exhibited clinical phenotypes associated with FA, including hypersensitivity to crosslinking agents that were reversed by expression of a *BRCA1* transgene (Sawyer et al., 2015). Furthermore, work from our laboratory and others has established that hematopoietic-specific knock-out of *Brca1* in mouse models leads to BMF and HMs (Mgbemena et al., 2017; Vasanthakumar et al., 2015). Furthermore, HSPCs lacking *Brca1* are hypersensitive to MMC, and even unperturbed cells show large scale chromosomal abnormalities (Vasanthakumar et al., 2015). These studies helped establish *BRCA1* as a *bona fide* FA gene.

Interestingly, some monoallelic mutations in FA genes are now appreciated for cancer predisposition in other tissues in non-FA patients, which overlap with known cancer risks for *BRCA1* and *BRCA2* mutation carriers. For example, germline mutations in *FANCJ/BRIP1* increase risks for developing ovarian cancer (Rafnar et al., 2011). Germline mutations in *FANCN/PALB2* and *FANCM* have also been associated with increased risk for developing breast cancer (Hofstatter et al., 2011; Kiiski et al., 2014). These findings provide further evidence for *BRCA1* playing a key role in genome maintenance in hematopoietic cells that are challenged with replication stressors in addition to potential roles in HR-mediated repair of DSBs.

There is significant functional overlap between the FA pathway and BRCA1-mediated DNA repair responses both in the removal of ICLs and at stressed replication forks more broadly. When two convergent replication forks stall at an ICL, BRCA1 is involved in replisome disassembly and RAD51 loading on to ssDNA to protect the replication fork (Figure 1.8). The ICL is then recognized by FANCM, which recruits the FA complex to the ICL. This complex is activated by ubiquitination of FANCI/D2, which

Figure 1.8 Removal of ICLs by the FA complex involves BRCA1 and an HR-repair step



Schematic of the FA pathway response to ICLs. Two converging replication forks stall at an ICL (red) leading to replisome disassembly and RAD51 loading onto ssDNA, which is facilitated by BRCA1. FANCM then recognizes the ICL leading to activation of the FA core complex through ubiquitination (Ub) of the FANCI and FANCD2 heterodimer. The activated FA complex then recruits multiple nucleases that unhook and process the ICL. This results in one strand (blue) with an unreplicated DNA gap that is filled in by translesion synthesis. This results in a DSB on the other strand (grey), which requires HR-mediated repair to complete replication of both strands. (Adapted from Michl et.al., 2016)

promotes nucleolytic processing and unhooking of the ICL (Michl et al., 2016). Importantly, the final stages of ICL removal involve HR-mediated repair of the DSB that results from this FA-mediated ICL processing (Figure 1.8). Furthermore, the FA pathway is activated in response to other replication stressors independent of ICLs, and both FANCI and FANCD2 have been shown to associate with MCM proteins in the replicative helicase under stress (Lossaint et al., 2013). There is also evidence for FA pathway activation in response to R-loops, G-quadruplexes, and the depletion of nucleotide pools by hydroxyurea, which are all ICL-independent forms of replication stress (Niraj et al., 2019). These features of the FA-pathway and the overlap with BRCA1-mediated genomic stability pathways highlight the importance of replicative stress in the etiology of hematopoietic dysfunction associated with mutations in *BRCA1* and other FA genes.

Further evidence for the importance of replicative stress in HSPC dysfunction due to mutations in *BRCA1* or other FA genes is exhibited by mouse models for FA. Interestingly, *Fanca*$^{-/-}$ mice do not spontaneously develop congenital or hematopoietic abnormalities (Cheng et al., 2000). Nevertheless, cells from these mice still exhibit hypersensitivity to MMC and HSPCs proliferate poorly *ex vivo* (Río et al., 2002). Similar phenotypes are observed in *Fanc*$^{-/-}$ and *Fancg*$^{-/-}$ mice which also fail to develop BMF or HMs, but HSPCs have reduced repopulating capacity and are also hypersensitive to crosslinking agents (Parmar et al., 2009). However, replicative stress induced by MMC exposure in *Fanc*$^{-/-}$ mice elicits BMF phenotypes (Bakker et al., 2013). Similarly, the Milsom group finds that induction of replicative stress in HSPCs through serial bleeding or induction of inflammation is sufficient to produce BMF in *Fanca*$^{-/-}$ mice (Walter et al.,

2015). Importantly, these replicative stressors are independent of ICLs and merely proliferation of HSCPs in these mice leads to replication-mediated DSBs and depletion of stem cell compartments. These replication-induced phenotypes are similar to the BMF observed in our mouse *Brca1*$^{-/-}$ mouse model, suggesting that replication stress is also a significant component of hematopoietic dysfunction in HSCPs in these animals. Taken together, these finding support a key role for FANC proteins, including Brca1, in restraining replication-mediated DNA damage in HSPCS and hematopoietic toxicities.

Synthetic lethality approaches for targeting DNA repair deficiencies

Excitingly, mechanistic understanding of DDR deficiency has led to a new type of therapeutic approach known as synthetic lethality (Lord and Ashworth, 2017). Cells that have lost one DNA repair program must rely on the remaining alternative pathways for managing DNA damage. This cellular state allows for selective targeting of cancer cells by exploiting the relative lack of redundancy in the DDR by hitting a back-up pathway that is now essential for tumor cell survival but can be tolerated by normal cells with a full complement of DDR programs (Chan and Giaccia, 2011). One of the best and currently most promising example of this approach is exhibited by the use of PARP inhibitors (PARPi) in BRCA1- or BRCA2-deficient breast and ovarian cancers. In these HR-deficient tumors, DSBs are instead managed by other DNA repair pathways including single strand annealing, base excision repair, and alt-EJ. Given that PARP1 plays a role in promoting these other forms of repair, inhibition of PARP1 in HR-deficient cells disrupts these alternative genome maintenance programs and sensitizes these tumors to therapeutic agents that induce DNA damage or replication stress (Farmer et

al., 2005; Jagtap and Szabó, 2005). Although the initial paradigm suggested that PARPi efficacy was due to the loss of other core DNA repair programs, recent findings also implicate the accumulation of ssDNA and trapping of PARP1 at replication forks as part of the mechanism of action for this synthetic lethal approach. Therefore, there is strong evidence that DDR pathway inhibitors are a successful strategy in cancer therapy, and efficacy may be linked to replication stress failures.

Following the success of PARPi use in BRCA-deficient cancers, many subsequent studies have attempted to identify additional synthetic lethal interactions in DNA repair pathways. This has included a wide variety of approaches, including CRISPR and RNAi screens targeting components of the DDR. For example, synthetic lethal screens have identified some NHEJ factors as being essential in HR-deficient conditions including DNA-PK (Mohiuddin and Kang, 2019). DNA-PK is made up of the Ku70/80 heterodimer and the catalytic subunit, DNA-PKcs, which are core factors in the NHEJ machinery for processing DSB ends. Inhibition of DNA-PK has shown success in BRCA1-deficent cells and ATM-deficient lymphoma models, and there are currently four different inhibitors being tested in clinical trials in a variety of settings (Topatana et al., 2020). Other synthetic lethality screens have identified DDR pathway components and cell cycle regulators as promising therapeutic targets, including WEE1, ATR, and CHK1 (Topatana et al., 2020). Furthermore, independent of screening approaches, molecular characterization of DNA repair responses in the context of HR-failures has identified additional synthetic lethal targets in BRCA1-deficient cells including POLQ and FANCM (Panday et al., 2021; Zatreanu et al., 2021; Zhou et al., 2021). Interestingly, these latter findings implicate the involvement of replication fork failures, suggesting that the overlap

of DDR proteins in managing stalled replication forks is involved in these observed synthetic lethal phenotypes. Therefore, deeper molecular understanding of how cells manage damage and replication stress in the context of defective genome maintenance programs will likely uncover new synthetic lethal targets for drug development.

The role of β-catenin in T-cell acute lymphoblastic leukemias

Tcf-1, the canonical DNA binding co-factor for β-catenin, plays essential roles at several stages of thymocyte development, including T cell specification, progression through both the pre-TCR checkpoint and the $CD4^+CD8^+$ DP stage, thymic selection, and the choice of post-selection T-cell lineages (Zhao et al., 2021). However, the relative involvement of β-catenin in these Tcf-1-mediated developmental processes remains unclear, and Tcf-1 activity during normal development may be independent of canonical WNT signaling. Nevertheless, there is substantial evidence for the involvement of β-catenin in T-cell malignancies. In these pathogenic conditions, aberrant stabilization of β-catenin could disrupt the normal functions of Tcf-1 in development and alter expression programs to promote transformation, which is addressed in greater detail in Chapter III.

In human T-cell acute lymphoblastic leukemia (T-ALL), the most frequent genetic perturbations can be categorized by mutations leading to activation of NOTCH signaling (>50%) and PTEN loss/AKT activation (20-45%) (Girardi et al., 2017; Gutierrez et al., 2009). As PTEN loss and AKT activation restrain the activity of GSK3β, these mutations lead to indirect stabilization of β-catenin as it is no longer phosphorylated and marked for degradation (Ashihara et al., 2015). Mouse models with hematopoietic specific

ablation of *Pten* by *Mx1-Cre* lead to T-ALLs with *Tcra-Myc* translocations and leukemia initiating cells that exhibit high levels of β-catenin (Guo et al., 2008). Importantly, deletion of a β-catenin allele in this *Pten* null model decreased the incidence and latency of T-ALL development (Guo et al., 2008). Similarly, the thymic specific activation of Akt using an *Lck-Cre* mouse model also produces DP T-cell leukemias with *Tcra-Myc* fusions. Furthermore, a separate study found that β-catenin stabilization synergizes with *Pten* loss in a Notch-independent T-ALL mouse model, and Wnt signaling was required for the initiation and maintenance of these tumors (Kaveri et al., 2013). Therefore, β-catenin activation likely plays an important role in the transformation of T-cells with mutations in the PTEN/AKT pathway.

There is also evidence for β-catenin activity in Notch-driven T-ALL models. For example, a T-ALL model driven by Notch signaling in murine fetal liver cells requires β-catenin both for the inhiation of leukemias and the transcription of Myc (Gekas et al., 2016). Similarly, a separate Notch-driven T-ALL model found that leukemia initiating cells required both Wnt and Hif1a signaling for their maintenance (Giambra et al., 2015). Therefore, although β-catenin signaling is more strongly associated with T-ALLs with PTEN/AKT mutations, aberrant stabilization of β-catenin may represent an underappreciated feature in the etiology of T-ALLs more broadly. This is corroborated by a study of pediatric T-ALL patients, unselected for genetic mutation status, in which nearly all patients exhibited elevated expression of *β-catenin* mRNA and nuclear localization of β-catenin (Ng et al., 2014). Taken together, this is substantial evidence for the involvement of β-catenin in the development and maintenance of T-ALL, which warrants further investigation.

Hypothesis and specific aims

In my thesis work, I investigated the mechanisms by which disrupted DNA repair and replication stress programs lead to genomic instability in hematopoietic cells, which drives tumorigenesis. I used model systems that are united by dysregulation of various nodes of the DDR response, centered on the ATM-CHEK2-BRCA1 axis for repair of DSBs. In all cases, preliminary evidence suggested that replication stress responses were involved in the etiology of genomic instability in these models. ***I hypothesized that dysfunctional replication stress responses are linked to failure of HR-mediated DNA repair, leading to replication-mediated DSBs that are repaired by more error-prone pathways that increase genomic instability in hematopoietic cells.*** This is due to the increasingly appreciated overlapping roles for HR factors in DNA damage and replication stress responses. This overall hypothesis was tested in three Specific Aims, each of which has a sub-hypothesis:

Specific Aim 1: Define the mechanisms of genomic instability in transforming T-cells. Here, I used a system in which stabilized β-catenin in late thymic development leads to recurrent *Tcra/Myc-Pvt1* translocations that mirror those seen in human T-cell leukemia. ***I hypothesized that aberrant β-catenin stabilization redirects Tcf1 to downregulate HR repair factors, which disables checkpoint signaling and allows for transmission of replication-mediated DSBs into the G1 phase where they are aberrantly joined with rearranging receptor loci.*** The results from this work can be found in Chapter II: Aberrant β-catenin activation guides Tcf-1 to promote genomic instability and thymocyte transformation.

Specific Aim 2: Determine whether germline variants in *CHEK2* predispose to hematopoietic malignancies. Here I used the Godley Laboratory's patient cohort focused on individuals and families with clustering of hematopoietic malignancies and solid tumors to identify recurrent germline mutations in *CHEK2* and evaluate their contribution to altered hematopoiesis and transformation. ***I hypothesized that variants in CHEK2 are associated with clonal hematopoiesis and hematopoietic malignancies due to deleterious effects on CHEK2 function in DNA repair and cell cycle regulation*.** The results from this work can be found in Chapter III: Germline *CHEK2* variants as risk alleles for clonal hematopoiesis and hematopoietic malignancies in humans and mice.

Specific Aim 3: Elucidate the mechanisms of genomic instability *Brca1*-deficient hematopoietic cells. Here I used a system in which complete loss of *Brca1* in murine bone marrow cells leads to BMF and HMs with widespread chromosomal aberrations. ***I hypothesized that complete loss of Brca1 leads to genomic instability in hematopoietic cells due to a combined deficiency in DNA repair and replication fork protection, and that replication stress can induce a similar phenotype in cells that are heterozygous for Brca1*.** The results from this work can be found in Chapter IV: Loss of *Brca1* in hemopoietic cells leads to replication-mediated genomic instability and large-scale chromosomal aberrations.

CHAPTER II

Methods

Animal husbandry and generation of genetically engineered mice

Cd4-Cre/Ctnnb1^{ex3fl} (CAT), *Cd4-Cre/ Ctnnb1*^{ex3fl}/*Tcf7*$^{fll/fl}$ (CAT-*Tcf7*$^{\Delta}$), *CD4-Cre/Ctnnb* ^{ex3fl}/*BclXL*$^{fl/fl}$(CAT-*BclXL*$^{\Delta}$), *Cd4-Cre/Tcf7*$^{fll/fl}$ (*Tcf7*$^{\Delta}$), *CD4-Cre/BclXL*$^{fl/fl}$(*BclXL*$^{\Delta}$) or littermate control *Cd4-Cre (*Cre*)* were used in all described experiments. The *Ctnnb*^{ex3fl} allele was reported previously(Harada et al., 1999), and mice carrying *CD4-Cre*, *Tcf7*^{fl} or *BclXl*^{fl} alleles and *Rag*$^{-/-}$ mice were procured from the Jackson Laboratory. Mice were maintained on a C57BL/6 background and housed at the University of Chicago animal facility in accordance with protocol #71880, approved by the University of Chicago Institutional Animal Care and Use Committee. Transplant recipient mice carrying the CD45.1 allele were procured from the Jackson Laboratory.

For CHEK2 p.I200T mouse model studies, a knock-in allele harboring the equivalent variant in the mouse *Chek2* sequence (T8677C/ p.I161T) and homology arms targeting exons 3-5 were designed. Transgenic mice were generated by injection into murine embryonic stem (ES) cells on a C57BL/6 background in collaboration with the Transgenics and ES Cell Facility at the University of Chicago. Transgenic mice carrying the I161T allele were confirmed by Southern blotting and long-range PCR across homology arms to the expected endogenous locus. The presence of T8677C allele and subsequent genotyping of the mouse cohort was completed by Sanger sequencing (Table 2.1).

Brca1 conditional knock-out mice were made by interbreeding *Mx1-Cre* and *Brca1*$^{F22\text{-}24}$, in which exons 22-24 that encode the second BRCA C Terminus (BRCT)

domain of Brca1 are floxed (Jackson Laboratory, strain #017835). Pups were injected intraperitoneally with 10ug Polyinosinic:polycytidylic acid (pI:pC) on days 14, 16, and 18 to generate an interferon response leading to *Mx1-Cre* expression and excision of floxed *Brca1* allele. Mice were maintained on a C57BL/6 background and housed at the University of Chicago animal facility in accordance with protocol #71370, approved by the University of Chicago Institutional Animal Care and Use Committee.

Table 2.1 Primers for genotyping and qRT-PCR

Primer Name	Purpose	Sequence (5'?3')	Project
Brca1-Geno_F	Genotyping	AGGGCCATGATTGTCAGTTC	Brca1
Brca1-Geno_R	Genotyping	GATGGAAGCTCCTTCACCAC	Brca1
Mx1-Cre_F	Genotyping	GCATTACCGGTCGATGCAACGAGTGA	Brca1
Mx1-Cre_R	Genotyping	GAGTGAACGAACCTGGTCGAAATCAG	Brca1
Polq-qPCR-F	qRT-PCR	GAAAAGCACCTGAACCCCCT	Brca1
Polq-qPCR-R	qRT-PCR	TTATCCGTCCTGTAGCCGTG	Brca1
Chek2_I161T_gDNA_Geno-F	Genotyping	CCTTGATTGTCTTCTTACTGCTG	Chek2
Chek2_I161T_gDNA_Geno-R	Genotyping	GCCTTTCCCAATAAGCTCGG	Chek2
Chek2_Total_qRT-PCR-F	qRT-PCR	GCTATGGGCTCTTCAGGATGG	Chek2
Chek2_Total_qRT-PCR-R	qRT-PCR	ACAGTGGTCCATCGAAGCAAT	Chek2
Chek2_I161T_cDNA_Geno-F	Sanger Sequencing	CCACTGTTGAGAAGGACGGACA	Chek2
Chek2_I161T_cDNA_Geno-R	Sanger Sequencing	GTTGTTACTCAGAGGACAGCGTT	Chek2
18S rRNA F	qRT-PCR	GAGGGAGCCTGAGAAACGG	
18S rRNA R	qRT-PCR	GTCGGGAGTGGGTAATTTGC	

Tissue Isolation

Thymii, spleens, and lymph nodes (LN) were resected in ice cold FACS buffer (2% FBS in PBS) and dissociated to single cell suspension using 0.70μ cell strainers. Bone marrow (BM) was extracted from resected femurs, fibulas, humeri, and hips by crushing with mortar and pestle in ice cold FACS buffer followed by passage through 0.70 or 0.40m strainers. Peripheral blood (PB) was collected by submandibular bleeding or heart stick at endpoint into EDTA-coated collection tubes. Red blood cells were lysed

with ACK buffer (150 mM NH_4Cl, 10 mM $KHCO_3$ and 0.1 mM EDTA (pH 7.4)) for 3 min on ice. The reaction was halted by flooding with ice cold 1x PBS. Cells were immediately spun down and resuspended in fresh FACS buffer. Thymii in CAT studies were extracted from 6-8 week old mice for pre-transformed tissues or from 2-3 month old mice for transformed tissue. All other tissue collections were obtained from mice between 2-4 months of age, unless otherwise noted.

Flow cytometry and fluorescence-activated cell sorting of murine lymphocytes

Lymphocytes were surface stained in FACS buffer for 30 min on ice containing antibody cocktails using the panels listed below. Samples were washed 2x in ice cold FACS buffer before acquisition. For intracellular staining, cells were then fixed for 30 minutes on ice and permeabilized for 15 minutes on ice using the Foxp3/Transcription Factor Fixation kit (eBioscience, 00-5521) according to manufacturer recommendations. Cells were then resuspended in permeabilization buffer containing intracellular antibodies and stained on ice for 30 minutes. Cells were then washed 2x in permeabilization buffer and finally resuspended in FACS buffer. Samples were acquired on a Fortessa 4-15 or X20 flow cytometer (Becton Dickinson) at the Cytometry and Antibody Technology Core at University of Chicago. The data were analyzed with FlowJo software (Becton Dickinson).

For CAT studies, surface antibodies were against CD4 (clone GK1.5, BD Biosciences), CD8 (clone 53-6.7, eBiosciences), Tcrβ (clone H57-597, eBiosciences), and Ccr7 (clone 4B12, eBiosciences). Cells were stained for viability using LIVE/DEAD

Aqua fluorescent reactive dye (Molecular Probes–Life Technologies, L34963). Intracellular staining was for γH2ax (Anti-H2AX (pS139), BD Biosciences, BDB562377).

For *CHEK2* and *BRCA1* hematopoietic stem and progenitor cell (HSPC) studies, lineage depleted (below) BM was first resuspended in FC block (BD, BDB553142) for 10 minutes on ice. Cells were then re-stained with the same biotinylated lineage panel to measure depletion efficiency. These cells were subsequently stained with a surface antibody panel including CD117/cKit (clone ACK2, Biolegend), Sca1 (Clone D7, Biolegend), CD34 (RAM34, BD) and streptavidin-FITC or APC (Biolegend 405201 or 405207). For stem cell panels, SLAM markers were targeted by the addition of antibodies against CD48 (clone HM48-1, BD) and CD150 (clone TC15-12F12.2, Biolegend) to the surface panel. Cell cycle panels were completed with the addition of Ki67 (clone SolA15, eBiosciences) to the surface panels. Additionally, after the final intracellular stain, these cells are also stained for DAPI (Sigma-Aldrich, 10236276001) at 1ug/mL for 5-10min at room temperature, followed by 2x wash in FACS buffer.

For *CHEK2* PB immunophenotyping panels, staining mixes contained antibodies targeting CD34 (RAM34, BD), CD117/cKit (clone ACK2, Biolegend), CD3 (clone 17A2, BD), CD4 (clone GK1.5, Biolegend), CD8 (clone 53-6.7, Biolegend), and Cd11b (clone M1/70, Biolegend).

Lineage depletion of BM

HSPCs were enriched by lineage depletion using magnetic LS columns (Miltenyi #130-042-401. Total BM (50-150x10^6 cells) was incubated for 30 minutes on ice in a cocktail of biotinylated antibodies targeting differentiated cell markers CD3 (clone 145-

2C11, Tonbo Biosciences), Gr1/Ly6G (clone R86-8C5, Biolegend), CD19 (clone eBio-1D3, eBiosciences), Cd11b (clone M1/70, eBiosciences), Ter119 (clone TER-119, Biolegend), and B220 (clone RA3-682, Biolegend). All antibodies were used at a 1:200 dilution in FACS buffer and were subsequently washed 2x in FACS buffer. Cells were then resuspended in the appropriate volume of MACS buffer (2%BSA in 1xPBS, filtered) containing streptavidin microbeads (Miltenyi, #130-048-101) according to manufacturer recommendations. Cells were incubated with microbeads for 15 minutes at 4°C, washed 2x with MACS buffer, and applied to LS columns. Samples sometimes required additional filtration through 0.45m strainers to prevent clogging in columns from cellular debris. Lineage negative cells were collected from the initial flow through and three 3mL washes in 15mL conical tubes positioned under columns in magnetic stands. Cells were then spun down, resuspended in FACS buffer, and used for subsequent studies. When required, lineage$^+$ cells were collected from LS columns using manufacturer supplied plungers and used for confirmation genotyping or other analysis.

Translocation Breakpoint Detection

Sequencing libraries from four independent CAT leukemia samples were created with the Nextera Mate Pair Preparation Kit (FC-132-1001), which allows for large insert-size (up to 12kb) and is ideal for the detection of structural variation. Each biological sample was replicated across two lanes and sequenced using HiSeq2500. Reads were then preprocessed and special adaptor clipping procedures were necessary before any alignment procedure NextClip (Leggett et al., 2014). Aligned MP reads were post-processed with Samtools v0.1.18 (Li et al., 2009) and Picard v1.70

(http://picard.sourceforge.net/). BreakDancer v1.3.6 (Chen et al., 2009) was used to detect inter-chromosomal translocations. Finally, the distribution of breakpoints across chromosomes was analyzed using the 5kb binned data. Primers were then designed for consensus translocations involving Myc-Pvt1 loci and precise clonal breakpoint sequences were determined with Sanger sequencing (Table 2.2).

Table 2.2 Primers for Sanger sequencing of translocation breakpoints

Break	Chromosome	Locus	Sequence (5'→3')
1	15	*Myc-Pvt1*	GAATTCTGAACCTGCAGAAGGAGC
1	14	*Tcra*	GATTGGGAGTCACAGCAACAGTTG
2	15	*Myc-Pvt1*	ATAAGGACCCATGCTTGACACCAC
2	14	*Tcra*	GTTGAAGTGTGAGCATGGCAGAAG
3	15	*Myc-Pvt1*	GCAGCTCTCAGAGTTTCAAAGCTG
3	14	*Tcra*	CTTCCAGGCACTTGGAAATGTTGG
4	15	*Myc-Pvt1*	AACCAGGCTATTGTCTTGCAGGTG
4	14	*Tcra*	CCTAGTGATCCAGTCGTGTTGAGT

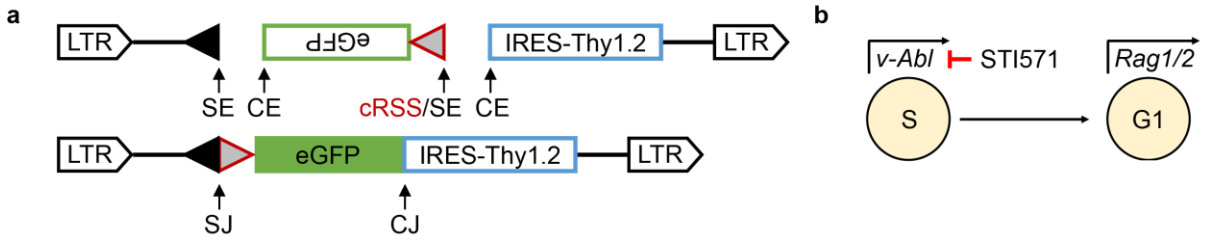
Identification of RSS and cryptic RSS sites

Genomic sequences from translocation breakpoints +/- 200 bp were investigated for RSS and cryptic RSS (cRSS) sites using the online reference database and predication tool RSSsite (Merelli et al., 2010). For *Myc-Pvt1* loci, the DnaGrab tool was used to score potential RSSs in this region. The top scoring potential cRSS site containing a 23-bp spacer was selected from each of four CAT lymphomas, representing the most likely substrate for fusion with 12-bp spacer RSS sites in *Tcra* loci according to the 12/23 rule.

Rag Recombination Assay

As previously described, a plasmid-based Rag-recombination assay(Bredemeyer et al., 2006) was adapted to evaluate potential RSS sites for Rag-cleavage. Briefly, a modified pMX-INV vector harbors an inverted eGFP sequence flanked by canonical RSS sites such that successful Rag-recombination flips the cassette in frame, leading to eGFP expression (pMX-INV-GFP, gift Dr. Barry Sleckman, UAB). An IRES-Thy1.2 marker is 3' to the eGFP cassette, to mark successful viral transduction. The top scoring potential cryptic RSS site from *Myc-Pvt1* break sites were synthesized by IDT, annealed, and cloned to replace the upstream RSS site (Figure 2.1a). Viral supernatants were produced by transfection into Plat-E cells using FuGENE reagents according to manufacturer recommendations (Promega E2311). Viral supernatants were collected, filtered, and used to transduce v-Abl transformed pro-B cells (gift Dr. Barry Sleckman, UAB These highly proliferative cells are forced into G1 by treatment with 3uM STI571 (imatinib, Selleck Chem, S2475) for three days, inducing Rag

Figure 2.1 Schematic of Rag-reporter system



a. Schematic of pMX-INV-GFP construct. An eGFP cassette is flanked by directional RSS or cRSS sites (triangles) where Rag recombinases act to generate hair-pins at coding ends (CE) and blunt signal ends (SE). Rag mediates inversion of the eGFP cassette to generate a coding join (CJ) and signal join (SJ) that puts eGFP in frame to be expressed. A flanking IRES-Thy1.2 reports on successful transduction of the construct **b**. Schematic of highly proliferative (v)-Abl transformed pre-B-cells (gift, Sleckman) in S-phase (S) that are driven into G1 by treatment with Abelson kinase inhibitor ST1571 (Imatinib) to initiate Rag expression.

expression and recombination (Figure 2.1b). After imatinib treatment, cells were collected, stained for Thy1.2 (anti-CD90.2/Th1.2-PE, Biolegend #553005) on ice for 30 minutes in FACS buffer, and evaluated for eGFP expression by flow cytometry using an LSRII (Becton Dickinson).

Nascent DNA Fiber Assays

For gentler selection conditions prior to culture, thymocytes were enriched for DPs (double positive, $CD4^+CD8^+$)] using a CD4+ T Cell Isolation kit (Miltenyi, 130-104-454) according to the manufacturer protocols. Cells were subsequently cultured in thymocyte media (IMDM, 10%FBS, 1%Pen/Strep, 50mM β-mercaptoethanol) for 1 hour at 5% CO_2 and 37°C to acclimate cells to culture conditions.

For untreated fibers, cells were pulse labeled first with 25 μM IdU (MP Biochemicals, 0210025701) for 25 minutes, washed once in 1xPBS, and subsequently resuspended in media with 250 μM CldU (Sigma, C6892) for an additional 25 minutes. For stressed replication assays, cells were co-cultured in 2mM Hydroxyurea (Sigma, H8627) during the second pulse. For fork recovery assays, cells were cultured in 1mM HU for 3 hours between the IdU and CldU pulses. For fork protection, cells were cultured in 2mM HU for 5 hours after the CldU pulse. After culture, cells were immediately collected on ice and diluted to $7.5x10^5$ cells/mL. Cells were then spotted onto the top of glass slides, briefly dried, and lysed dropwise (100mM Tris pH 7.5, 0.5% SDS, 50mM EDTA). After 2-3 minutes, slides were tilted at a 20-40° angle allowing droplet to roll down the slide length. Slides were then air dried for 45 minutes, fixed in a 3:1 methanol/ acetic acid solution for 10 minutes, and denatured in 2.5M HCl for 80

minutes. Fibers were stained at 4°C overnight in a humidified chamber with 1%BSA containing primary antibodies rat-α-BrdU (1:200, BD Biosciences, B44) and mouse-α-BrdU (1:25, Abcam, BU1/75(ICR1)), which label IdU and CldU, respectively. Slides were then washed in PBS, fixed for 10 minutes (3% PFA, 3.4% sucrose in PBS), and stained with α-rat Alexa Fluor 488 (1:500, Invitrogen, A-11006) and α-mouse Alexa Fluor 594 (1:400, Invitrogen, A-11005) for 1.5 hours at room temperature. Slides were then washed and mounted with ProLong Anti-fade mounting media (Invitrogen, P36930) and #1.5 cover glass. Images were acquired with an Olympus IX81 inverted microscope with a 100X, NA 1.45 objective at the Integrated Light Microscopy Core at the University of Chicago. Images were visualized and individual DNA fibers were hand selected using ImageJ. Individual fiber track lengths and CldU:IdU ratios were measured using custom Matlab scripts to avoid user bias.

Chromatin immunoprecipitation and sequencing

2×10^7 total thymocytes from 6- to 8-week old CAT mice were paraformaldehyde fixed at a final concentration of 1% (2106-01, J.T. Baker) for 15 min at room temperature (23 °C), quenched with glycine (0.125 M), and washed with ice-cold PBS containing protease inhibitors. Cells were resuspended in lysis buffer (10 mM Tris, pH 7.4, 1 mM EDTA, 1% TritonX-100, 0.1% sodium deoxycholate, 0.8 M NaCl, and 0.1% SDS) for 10 min at 4°C and sonicated to an average size of 300 bp. The chromatin was incubated overnight with Protein G Dynabeads (10004D, Invitrogen) coupled to 5μg of antibodies to H3K27Ac (Abcam, #4729), H3K27me3 (Millipore #07-449), or H3K9K14Ac (Millipore, #06-599). Beads were washed five times with lysis buffer (10 mM Tris, pH

7.4, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.5 M NaCl, and 0.1% SDS) and once with 1×TE. The chromatin was eluted with elution buffer (2% SDS and 20mM Tris-HCl, pH 6.8) and reverse cross-linked overnight at 65 °C. RNase A was added (50µg/mL) and incubated at 37°C for 1 h. Proteinase K was added to a final concentration of 240µg/mL and incubated at 56°C for 2h. DNA was ethanol precipitated and resuspended in elution buffer (Qiagen). ChIP material was prepared for sequencing in accordance with the Illumina/Solexa Genomic DNA protocol. Approximately 20ng of immunoprecipitated DNA was end repaired, polyadenylated, ligated to Illumina TruSeq indexed adaptors, and purified with AMPure XP Beads (A63880, Beckman Coulter). Adaptor-ligated DNA was PCR amplified with KAPA Hifi DNA Polymerase (KK2601, Kapa Biosystems). PCR products were separated on a 2% agarose gel, and DNA fragments between 200 and 500 bp were excised and purified (28706, Qiagen). Sequencing was performed at the University of Chicago Genomics Facility using a HiSeq 4000 sequencer. Additional ChIP data for Tcf-1 and H3K4me3 and some *CD4-Cre* profiles are publicly available from previous reports (GSE46662, GSE32311, SRP142342).

RNA Isolation and Sequencing

Total RNA was extracted using TRIzol (15596026, Invitrogen) according to the protocol described by the Immunological Genome Project (https://www.immgen.org**/)**. For CAT thymocyte studies, $1x10^7$ live $CD4^+CD8^+$ DP thymocytes were sorted from three *Cd4-Cre*, CAT, and CAT-*Tcf7*$^{\Delta}$ mice. RNA sequencing libraries were generated and sequenced by the University of Chicago Genomics Facility. For *Chek2* and *Brca1*

studies, RNA was extracted from 1-3x10^6 cells as above, and cDNA was generated using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies, 4368814). Quantitative PCR (qPCR) was performed with Power SYBR Green PCR Master Mix (Life Technologies 4368706). Primers for qRT-PCR can be found in Table 2.1

For *Brca1* studies, RNA-seq was performed on 50-100,000 sorted LSKs from murine BM. RNA was isolated using the RNeasy Micro Kit (Qiagen, #74004), and libraries were made with SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input (Takara, #634412). Sequencing was performed with Illumina Hi-Seq 4000 at the University of Chicago Genomics Facility.

Genome mapping and data analysis

Sequenced ChIP datasets were mapped with the Galaxy (https://usegalaxy.org/**)** suite of tools. Data were groomed and aligned to the mouse mm9 genome with Bowtie, allowing up to one mismatch and retaining only uniquely mapped reads, and unmapped reads were filtered. Peak calling was performed with MACS via HOMER (Heinz et al., 2010). Transcription-factor peak calling was performed relative to input controls with the requirement that peaks were at a minimum fivefold enriched over input and meet a P-value cutoff of 10×10^{-5}. Motif enrichment analysis was performed with the HOMER motif discovery algorithm (Heinz et al., 2010). Tcf-1 and histone ChIP profiles were represented as heatmaps and enrichment histograms using ngs.plot software (Shen et al., 2014). RNA-seq datasets were aligned to mouse mm9 as above ChIP-seq studies. Batch correction, when necessary, was performed with the RUVs function in RUVSeq (Risso et al., 2014). Unbiased clustering analysis of normalized RNA-seq reads was

performed with ExpressCluster (www.GenePattern.org) (Reich et al., 2006). Differential gene expression analysis was performed with EdgeR (Robinson et al., 2009). Only genes that exhibited a counts per million (CPM) value greater than 0.5 in at least three samples were kept, and TMM normalization was applied using the calcNormFactors function in EdgeR (Liao et al., 2018). Genes with transcript abundance differences below $P < 0.05$ were considered to be significantly differentially expressed. Pathway enrichment analysis was performed using Metascape. Gene set enrichment analysis (GSEA) (Subramanian et al., 2005) was run against all gene sets within KEGG, HALLMARK, Reactome and GO databases (Broad, MsigDB).

Spearman correlation density plots

Density plots were created with the stat_bin2d function in the ggplot2 package in R, with 30 bins in each dimension. For visualization purposes, the axis ranges of some density plots were limited to highlight the high-probability regions of the plot. Spearman correlation coefficients and P values were computed in R with the cor and cor.test functions. Genes were then filtered for the greatest relevance to the rescue condition by selecting those with the tightest fit to the Spearman correlation (i.e. distance ≤ 1 from the fit line) to perform pathway analysis on the ”restored” genes relative to CAT.

Olaparib treatments of CAT lymphoma *in vivo*

Suspensions of CAT lymphomas were isolated, and $2x10^6$ cells were transferred into sub-lethally irradiated (450 rads) $Rag^{-/-}$ mice by tail vein injection. Three days post-transplant, mice were treated 5x/week for three weeks with intraperitoneal injections of

50mg/kg olaparib (AstraZeneca, Lynparza). The progression of CAT lymphomas was enumerated by the fraction of $CD4^+CD8^+$ cells identified in weekly PB sampling (100ul) and flow cytometry after staining for CD4 and CD8, as described above.

Genomics of Drug Sensitivity in Cancer (GDSC) database analysis

Normalized IC50 data were extracted for all compounds and all cancer cell lines from the Genomics of Drug Sensitivity in Cancer database (Iorio et al., 2016). Cell line metadata were also extracted and used to define T-cell and leukemia/ lymphoma cell lines, which were integrated to IC50 data using custom R-scripts. Metadata for inhibitors were used to classify pathway targets and data were reduced to drugs of interest (DOIs) targeting WNT, PARP, or AKT pathways. Means of normalized IC50 data for DOIs in T-ALL or leukemia and lymphoma cell lines were compared to all other cancer lines using two-sided, unpaired t-tests. Linear correlations of IC50 data for olaparib versus other WNT or PARP inhibitors was performed in R using the lm() function and data were depicted using the ggplot2 package.

Patient cohorts for *CHEK2* germline mutations

Within the exploratory cohort, patients who underwent clinical or research testing for germline variants at the University of Chicago medical center from August 2015 through February 2021 were assessed retrospectively for the presence of a reported variant in the *CHEK2* gene. All probands had germline tissue assessed with DNA extracted from cultured fibroblasts; cascade testing of family members was via saliva or PB. Variants were identified with a target capture panel (University of Chicago Genetics

Services Laboratory) or by Sanger sequencing. Variants were curated and classified in accordance with American College of Medical Genetics (ACMG) guidelines. Clinical and pedigree details were collected by chart review. Population frequencies of germline *CHEK2* variants was extracted from the Exome Aggregation Consortium (ExAc) in gnomAD (Karczewski et al., 2020) for p.I200T, p.S428P, and p.T367fs variants. This study was approved by the University of Chicago Institutional Review Board (protocol #11-0014), and all study participants gave written informed consent for research participation.

Statistics for *CHEK2* study

Between-group differences were calculated by Chi-square, Fishers' exact, or t-test. Odds ratios were calculated using comparing germline *CHEK2* variant frequencies versus the gnomAD control population with between-group comparisons using a binomial test; $p<0.05$ was taken as significant.

UK Biobank PheWAS Analysis

PheWAS data from the United Kingdom (UK) Biobank (Sudlow et al., 2015) were obtained through the genebass portal (https://genebass.org, GRCh38, v.0.7.8-alpha) (Karczewski et al., 2021). Phenotypes were selected for all predicted loss of function (pLoF) gene burden associations with *CHEK2* as well as *CHEK2* p.I200T alone at a significance level of $1e^{-7}$ with a SKAT-O burden test. Single-variant associations were examined for the top phenome term (e.g., platelet crit) identified.

Monitoring hematopoiesis and health of mice with the CHEK2 p.I161T allele

A cohort of *Chek2*$^{wt/wt}$, *Chek2*$^{wt/I161T}$, and *Chek2*$^{I161T/I161T}$ mice was monitored by daily health assessments in addition to monthly weight checks and submandibular bleeds (50-150uL) to assess PB counts. Complete blood counts (CBCs) were performed using a Hemavet 950 (CDC Technologies, Oxford, CT), and results were binned into two-month groupings. Wright-Giemsa stained PB smears were assessed visually. Peripheral and mediastinal tumor monitoring was performed by palpitation and scruffing, respectively. Mice with rapidly elevating lymphocyte counts or changes in health appearance due to weight loss, ruffled fur, hunched posture, decreased mobility, or labored breathing were euthanized, and full necropsies were performed with a focus on hematopoietic tissues. Endpoint necropsies included terminal bleed for CBC by heart stick, blood smears, and histological examination of fixed tissues from BM (sternum), spleen, thymus, lymph nodes, and liver. Fixation was performed in 10% neutral-buffered formalin for 24-48 hours, with sternums being decalcified for 2 hours in Cal-rite (Richard-Allan Scientific, #5501) post-fixation. Tissues were stored in 70% ethanol until being embedded in paraffin and sectioned at 3-4µ. Single cell suspensions of BM, PB, splenocytes, lymph nodes, and any malignant tissues were prepared for flow cytometry staining panels. Immunophenotyping was performed by flow followed by confirmation with additional immunohistochemical (IHC) staining of sectioned, fixed tissues for relevant cell-type specific markers. Immunohistochemical staining was performed and optimized in collaboration with the Human Tissue Resource Center at University of Chicago.

Monitoring clonal hematopoiesis in mice with CHEK2 p.I161T

A cohort of *Chek2*$^{wt/wt}$, *Chek2*$^{wt/I161T}$, and *Chek2*$^{I161T/I161T}$ mice was bled (200-250µL) at 6-9 months of age. Two rounds of RBC lysis (ACK buffer on ice for 3 minutes) were preformed followed by genomic DNA (gDNA) extraction using standard phenol:cholorform protocols. DNA concentrations were determined using the Qubit dsDNA broad range assay kit (Thermo Fisher, Q32850). 1ug of DNA was sent for next generation sequencing using the MSK-IMPACT 585 gene panel, which has been developed for assessing clonal hematopoiesis in mice(Loberg et al., 2019). Matched tail and PB samples were used to eliminate variants unique to our colony. Variants present in PB but not tails from the same animal were considered CH clones.

Variant calling for clonal hematopoiesis in mouse peripheral blood was performed using the methods established by Memorial Sloan Kettering Cancer Center.

Protein extraction

Cell pellets were resuspended with wide bore tips in ice cold RIPA buffer (50mM Tris pH 8.8, 1% TritonX-100, 0.1% SDS, 150mM NaCl, 5mM EDTA, 0.5mM EGTA 0.5% sodium deoxycholate) containing protease (Calbiochem, #539134) and phosphatase (Thermo, #78420) inhibitors. Cells were vortexed briefly, incubated on ice for 10 minutes, and then vortexed for 30 min at 4°C. After spinning at 10,000g for 30 minutes in a chilled centrifuge, lysates were collected from the supernatants and stored at -80°C.

Western blotting

Proteins were separated using an SDS-PAGE system on 6-12% acrylamide gels depending on the protein of interest. Membranes were probed with primary antibodies diluted in 1%BSA in 1x PBS for 1 hour at room temperature or overnight at 4°C. Primary antibodies were against p21 (BD, # 556431), p53 (Santa Cruz, sc-126), γH2ax (Histone H2A.X (Ser139), Millipore, #05-636), 53Bp1 (Novus, NB100-904), Phospho-Chk1 (Ser345, Cell Signaling, #2348), CyclinA (Santa Cruz, sc-239), CyclinD2 (Santa Cruz, sc-53637), Gapdh (Cell Signaling, #2118), Lamin A (Santa Cruz, sc-71481), Ligase3 (Santa Cruz, sc-135883), Parp1 (Santa Cruz, sc-53643 or Cell Signaling, #9532), or Topoisomerase I (Abcam, ab109374). Membranes were washed 3x in TBS-T. Secondary antibody stains were performed at 1:5000 dilution in 1% milk with antibodies against rabbit IgG (Millipore 401393-2ML) or mouse IgG (Cell Signaling Technology 70765). Membranes were then incubated with Western-lightning Plus-ECL, enhanced chemiluminescence (PerkinElmer) and exposed using standard x-ray film.

Murine BM culture

Mouse BM cells ($1x10^6$ cells/mL) were cultured in HSPC-enriching conditions using Stem Span SFEM media (Stemcell Technologies, #09600) containing mIL-3 (10ng/mL), hIL-6 (10ng/mL), and SCF (50ng/uL).

DNA damage foci

12-mm round #1.5 glass coverslips were poly-l-lysine coated and placed in 12-well plates, centered in the bottom of each well. Cells were resuspended in FACS buffer at

10^7cells/mL, and 100uL of the cell mixture was pooled onto the top of each coverslip. Cells were allowed to settle and attach to coverslip by 15-minute incubation at room temperature. Coverslips were then gently washed with 0.5mL 1xPBS/well and aspirated. Cells were then fixed in 4% v/v paraformaldehyde (EMS, #50980488) in PBS for 10 minutes at room temperature, then washed 3x with PBS. Cells were then permeabilized by immersion in 0.1% Triton X-100 in PBS for 15 minutes at room temperature and washed 3x with PBS. Coverslips were then blocked for 1 hour at room temperature in PBS containing a mixture of 2.5% horse serum, 2.5% goat serum, and 2.5% rabbit serum. Blocking buffer was diluted to 1% serum and used to dilute primary antibodies against γH2ax (Histone H2A.X (Ser139), Millipore, #05-636) or 53Bp1 (Novus, NB100-904). The primary antibody mixture was spotted onto parafilm in a humidified chamber and coverslips were inverted onto the staining mix. Coverslips were incubated overnight at 4°C, protected from light. Coverslips were returned to 12-well plates and washed 3x in diluted blocking buffer (1% total serum in PBS). Secondary staining was performed at room temperature for 2 hours with goat anti-Mouse Alex Flour 594 (Invitrogen, A-11005) or goat anti-Rabbit Alex Flour 488 (Invitrogen, A-11034) prepared at 1:400. Coverslips were then washed 3x in PBS. Nuclear staining was then performed by incubation in 1ug/mL DAPI for 10 minutes at room temperature followed by 3x PBS washes. Finally, coverslips were mounted (Invitrogen, P36930) on glass slides and sealed with clear nail polish. Imaging was performed at the Integrated Light Microscopy Core at University of Chicago on an Olympus DSU Spinning Disk Confocal microscope. Images were processed using ImageJ and R.

Replication stress in murine BM

To induce replicative stress in mouse BM, animals were injected intraperitoneally with 0.5mg/kg Polyinosinic:polycytidylic acid (pI:pC). Animals were sacrificed 72 hours later, and BM was extracted.

CHAPTER III

Aberrant β-catenin activation guides Tcf-1 to promote genomic instability and thymocyte transformation

The data in this chapter are adapted from a manuscript currently submitted to *Nature Communications*: Arnovitz S, Mathur P, Tracy M, Mohsin A, Mondal S, Quandt J, Hernandez KM, Khazai K, Emmanuel AO, Gounari F. Aberrant β-catenin activation guides Tcf-1 to promote genomic instability and thymocyte transformation.

In this chapter, I designed and performed the experiments, analyzed the data, and wrote the manuscript. P.M. designed and performed experiments and analyzed data. M.T., S.M., and J.Q. performed experiments. A.M. assisted in DNA fiber analysis. K.M.H performed breakpoint detection analysis. A.O.E. participated in study design and ran experiments. F.G. conceived the study, analyzed the data, and edited the manuscript.

Introduction

Genomic instability is a well characterized hallmark of cancer that drives both malignant transformation and cancer cell plasticity, which underlies treatment escape. It is estimated that a human cell is subjected to ~70,000 DNA lesions a day and the integrity of its genome depends on a coordinated balance between the generation and faithful repair of these DNA lesions (Lindahl and Barnes, 2000). Although extensive research into DNA repair pathways has identified many cancer risks and treatment opportunities, the precise molecular mechanism generating large scale genomic alterations remains incompletely understood.

Chromosomal aberrations and gene fusions are particularly prevalent in hematological malignancies, including up to 80% of T-cell acute lymphoblastic leukemia (T-ALL) (Graux et al., 2006; Vermeer et al., 2008; Van Vlierberghe et al., 2008). Importantly, T-ALL is an aggressive hematologic malignancy with poor prognosis and limited treatment options that remain largely untargeted (Belver and Ferrando, 2016; Samra et al., 2020; Teachey and O'Connor, 2020). Although dose escalation and combinatorial strategies have improved pediatric cure rates to >90%, these regimens are not well tolerated by adult patients (Samra et al., 2020). Excitingly, molecular and genomic classifications have led to improved new therapies including antibody and CAR-T regimens that predominately target B-cell ALL (Samra et al., 2020). However, similar success in T-ALL has lagged.

Lymphocytes are unique among somatic cells because their development requires the controlled generation of DNA double strand breaks (DSBs) to recombine the T-cell receptor (TCR) and Immunoglobulin (Ig) loci. Rag recombinases catalyze

these DSBs at recombination signal sequences (RSSs) that flank the V, D, and J segments of these loci. Rag activity occurs during the G0/G1 phases of the cell cycle when the predominant DSB repair pathway is non-homologous end joining (NHEJ) (Fugmann et al., 2000; Tubbs and Nussenzweig, 2017; Tubbs et al., 2018). NHEJ is error prone as it directly ligates DNA ends leading to small insertions, deletions, and potentially translocations (Lieber, 2010). Homologous recombination (HR), the other major DSB repair pathway, is favored during replication when a sister chromatid is present, providing a template for error-free repair (Jasin and Rothstein, 2013). Genome stability during thymocyte development requires a refined coordination between cell cycle and repair mechanisms. This is because T-cell development is a highly regulated stepwise progression through precursor stages that require bursts of proliferation followed by G0/G1 arrest to facilitate Rag-mediated receptor rearrangements (Love and Bhandoola, 2011; Rothenberg, 2019). When DSBs are detected, checkpoint kinases Chek2 or Chek1 signal cell cycle arrest to allow for repair of DNA lesions and maintenance of genome stability (Kastan and Bartek, 2004). In thymocytes, the inability to resolve replication induced DSBs promptly due to failure of repair pathways and/or checkpoint signaling could allow their persistence into G0/G1. At this phase, which relies on the error-prone NHEJ repair mechanism, these breaks could serve as aberrant substrates during Rag-recombination, leading to translocations. Not surprisingly, chromosomal aberrations involving the recombining T cell receptor genes are frequently observed in T-ALL cases (Larmonie et al., 2013).

Stabilizing mutations in β-catenin and activation of the Wnt signaling pathway have been described in T-cell malignancies (Lento et al., 2013) including precursor (T-

ALL), peripheral (PTCL), cutaneous (CTCL) and adult T-cell leukemia (ATL) (Bellei et al., 2006; Groen et al., 2008; Ng et al., 2014; Ram-Wolff et al., 2010). A more recent study showed that β-catenin is required for the initiation of T-cell leukemia downstream of NOTCH signaling (Gekas et al., 2016), which is the most frequently activated pathway due to mutations in T-ALL patients. Additionally, both loss of PTEN and activation of PI3K/AKT signaling, which mark ~47% of T-ALL (Gutierrez et al., 2009), stabilize β-catenin by inactivating GSK3β and preventing the phosphorylation events that initiate its proteasomal degradation (Dose et al., 2014; Guo et al., 2007; Kaveri et al., 2013). Mouse models of thymocyte-specific β-catenin activation, via mutation of β-catenin itself, loss of Pten, or expression of an active Akt, produce genomically unstable T-cell leukemias that mirror human disease (Dose et al., 2014; Guo et al., 2008, 2007; Kaveri et al., 2013; Timakhov et al., 2009). Importantly, β-catenin signaling is required in PTEN-deficient models of T-cell transformation (Guo et al., 2008) highlighting the need to understand the mechanisms by which β-catenin activation drives T-cell transformation.

The T cell specific partners of β-catenin in the context of Wnt signaling are the HMG domain-containing DNA binding proteins Tcf-1 (encoded by Tcf7) and Lef-1. Upon Wnt activation, β-catenin is stabilized and transported to the nucleus where it interacts with DNA-bound Tcf-1 and Lef-1, promoting chromatin accessibility and target gene expression (Mosimann et al., 2009). Tcf-1 has essential roles at almost all stages of T-cell development and function (Zhao et al., 2021). We and others have shown that Tcf-1 is required for T cell specification (Germar et al., 2011; Weber et al., 2011), progression to the $CD4^+CD8^+$ DP stage (Emmanuel et al., 2018), thymic selection, and the choice of

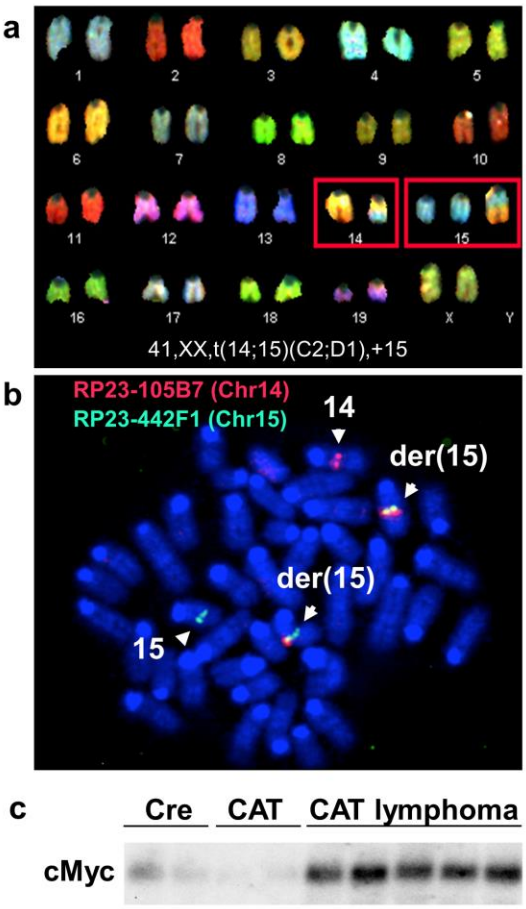
post-selection T-cell lineages (Steinke et al., 2014). These studies established that Tcf-1 modulates the chromatin landscape and transcription profiles of thymocytes (Johnson et al., 2018). Specifically, we showed that in DP thymocytes Tcf-1 acts both by directly binding its conserved DNA motif directly or through the indirect coordination with other regulatory proteins (Emmanuel et al., 2018). Tcf-1 binding can promote either up- or downregulation of target gene expression depending on sequence context and coordination with binding partners. It remains unclear how oncogenic β-catenin stabilization affects the natural functions of Tcf-1 to promote transformation.

Here I used our previously established mouse model of T-ALL, which relies on conditional stabilization of β-catenin in DP thymocytes, to test the mechanism by which β-catenin and Tcf-1 cooperate in transforming these cells (Dose et al., 2014; Guo et al., 2007). The leukemias in this model have chromosomal translocations that are identical to aberrations seen in T cell leukemia models resulting from PTEN ablation or constitutive AKT activation (Timakhov et al., 2009). I found that conditionally ablating Tcf-1 at the time of β-catenin stabilization abrogates leukemogenesis. This observation allowed me to define the specific mechanisms of transformation. My findings indicate that activation of β-catenin redirects Tcf-1 binding and promotes gene expression changes that compromise the replication process and checkpoint responses to replication stress. As a result, DNA damage generated during replication is aberrantly joined with Rag breaks induced during the subsequent G0/G1 phase. The HR impairments of these leukemias render them more vulnerable to Parp inhibitors and provide a therapeutic possibility.

Results

β-catenin induced translocations link DSBs from two distinct processes.

Figure 3.1 CAT mice have recurrent *Tcra/Myc-Pvt1* translocations leading to overexpression of Myc



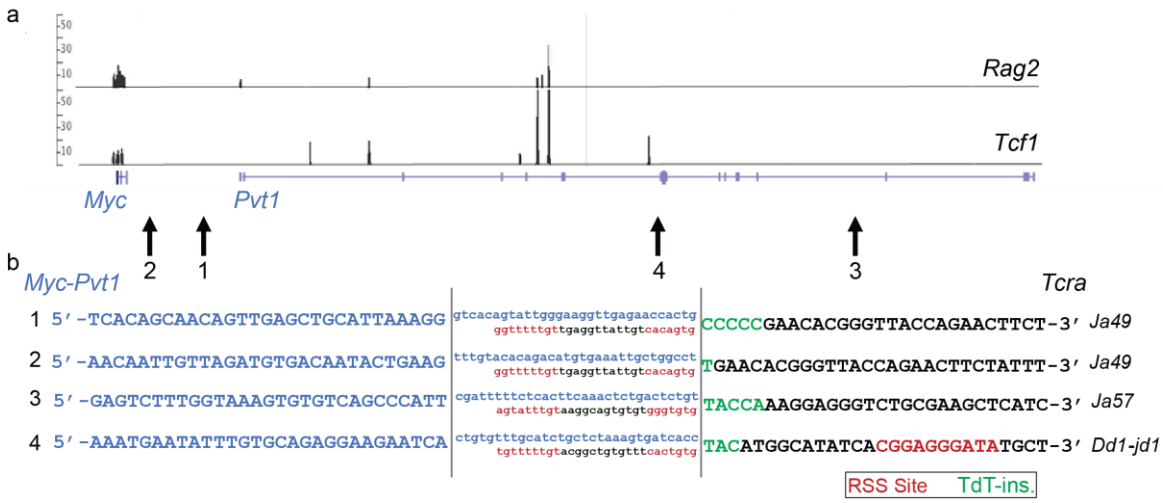
a. Representative spectral karyotyping analysis from transformed CAT thymocytes exhibit translocations between chromosomes 14 with *Tcra* and 15 with *Myc-Pvt1*. **b.** FISH analysis shows colocalization of probes for *Tcra* (pink) and *Myc-Pvt1* (green) (Dose et al., 2014) **c.** Western blot analysis exhibits overexpression of the *Myc* oncogene in transformed CAT thymocytes.

We have reported previously that *CD4-Cre* mediated stabilization of β-catenin (Harada et al., 1999) (CAT mice) produces DP T-cell lymphomas (Guo et al., 2007). These have recurrent translocations in which the *Tcra* locus on chromosome 14 is illegitimately linked to the *Myc-Pvt1* locus on chromosome 15 (Dose et al., 2014), driving elevated expression of the *Myc* oncogene and malignant transformation of DP thymocytes (Figure 3.1). As lymphomagenesis in CAT mice is dependent on Rag activity, I sought to elucidate the mechanisms underpinning the genomic instability that allows for aberrant *Tcra/Myc-Pvt1* fusions. I first evaluated publicly available ChIP data from DP thymocytes in addition to our own Tcf1 ChIP and showed that both Rag2 (Teng et al., 2015) and Tcf1 bind the Myc-Pvt1 locus (Figure 3.2a)

Therefore, I hypothesized that off-target Rag activity may generate DSBs in the *Myc-Pvt1* locus, which are then aberrantly joined with breaks generated during normal *Tcra* rearrangements. To address this postulate, I utilized data generated by Kyle Hernandez wherein we made whole genome mate pair libraries, which is a next generation sequencing approach optimized for the detection of structural variations through the identification discordant reads. To establish the precise translocation breakpoints and sequences, I designed PCR primers (Table 2.2) throughout the *Tcra* and *Myc-Pvt1* loci walking toward junctions established in *Tcra*/*Myc-Pvt1* hybrid reads. These analyses revealed that the *Tcra* breakpoint site involved early rearranging J fragments of the *Tcra* locus (Ja49, Ja57, Dd1-jd1). Normal thymocyte development contains inherent danger for genomic stability, as it involves generating DSBs and joining pairs of RSS sites with either 12 or 23 bp spacers (according to the 12/23 rule), which flank V- D- and J- fragments. Additionally, enzymes such as Terminal deoxynucleotidyl transferase (TdT) then mediate gap filling at coding joins that is critical for increasing receptor diversity for adaptive immunity. As expected, the *Tcra* break site of the translocations contained a proximal canonical recombination signal sequence (RSS), suggesting these breaks were Rag-generated. Furthermore, small insertions of N-nucleotides next to RSS sites are evidence of TdT activity (Figure 3.2b). In contrast, no canonical RSS sequences were identifiable within the *Myc-Pvt1* breakpoint sites.

As Rag can also generate DSBs at alternative sites known as cryptic RSS (cRSS) sites, I employed the reference database and prediction tool RSSsite (Merelli et al., 2010) to evaluate the 200bp region flanking the *Myc-Pvt1* breakpoints, accounting for potential resection from DSB processing enzymes prior to translocation formation.

Figure 3.2 Evidence of Rag activity of at translocation loci



a. ChIP seq tracks of Rag2(Teng et al., 2015) and Tcf-1 binding in the Myc-Pvt1 locus of WT DP thymocytes (top) and corresponding break sites from sequenced Tcr/Myc-Pvt1 translocations in four CAT lymphomas (arrows, 1-4). **b**. Tcra/Myc-Pvt1 breakpoint sequences from whole genome sequencing (upper case) and flanking endogenous region (lower case) are shown, highlighting heptamers and nonamers of RSS sites (red) with 12 bp spacers and TdT-inserted N-nucleotides at junctions (green).

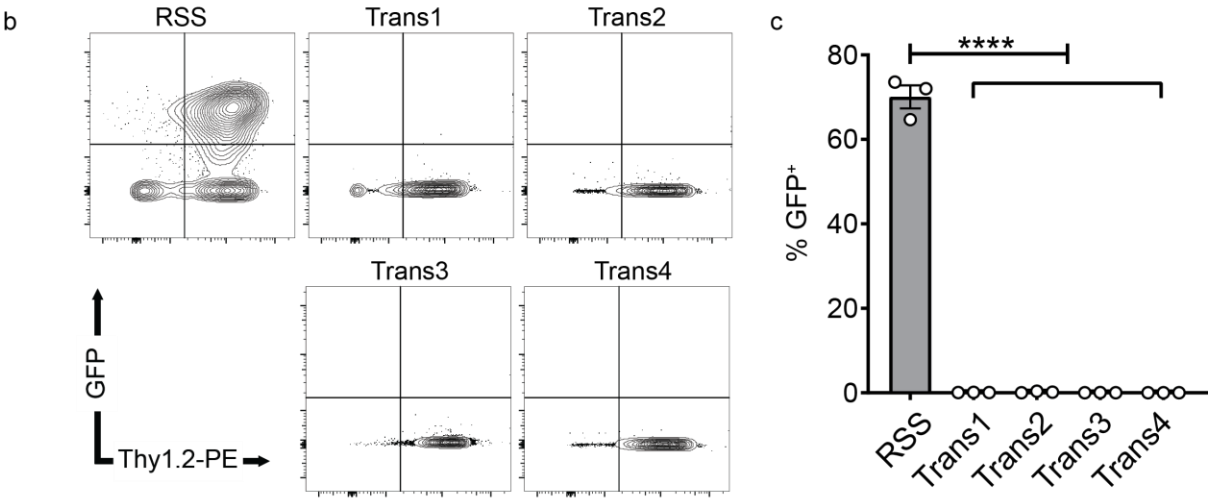
Although only one low quality cRSS site was identified in four separate lymphomas (Figure 3.3a), I also functionally tested the top scoring sequences from each lymphoma using a Rag-activity reporter assay (Figure 2.1a). I prioritized 23-bp spacer cRSSs that would pair with the 12-bp RSS sites identified in the *Tcra* site of the translocation (Figure 3.2b). The potential cRSS site from each lymphoma was synthesized and cloned into the pMX-INV-GFP Rag retroviral recombination reporter (Bredemeyer et al., 2006). In pMX-INV, Rag activity mediates inversion of an anti-sense green fluorescent protein (GFP) cDNA flanked with RSS sites and initiates GFP expression. These vectors were used to transduce (v)-Abl kinase-transformed pre-B-cell lines (gift of Dr. Sleckman). Treatment of these transduced and highly proliferative pre-B-cell lines with the Abl kinase inhibitor, STI571, blocks the G1-to-S transition and rapidly initiates Rag expression and activity (Figure 2.1b). In contrast to a control vector containing a bona

fide RSS site, I recovered no GFP expression from any of the potential *Myc-Pvt1* cRSSs in these experiments, demonstrating that the cloned sequences were not Rag substrates (Figure 3.3b-c). These findings indicate that the illegitimately repaired DSBs at the *Myc-Pvt1* site of the translocation breakpoint result from alternative, Rag-independent mechanism(s).

Figure 3.3 Identification and functional testing of potential cryptic RSSs

a

Translocation	Top Candidate cRSS **Site**	5'bps to BS	3'bps to BS	Strand	Score	
1	cacagtattgggaaggttgagaaccactgtcctagatca	96	134	+	-51.22	PASS
2	cacagacatgtgaaattgctggcctgaagtcacataagc	98	136	+	-63.30	FAIL
3	caaactctgactctgtagcctgctataggcgtagctgcc	109	147	+	-73.73	FAIL
4	caccttagctggggcaggggacccagacattgatgaaat	128	166	+	-72.81	FAIL



a. Highest scoring potential cryptic RSS (cRSS) site from *Myc-Pvt1* break points +/- 200bps using RSSsite (bps = base pairs, BS=break site). **b**. Representative flow cytometry dot plots of GFP$^+$ cells after 3-day STI571 treatment of v-Abl-transformed pre-B cells containing pMX-INV-GFP vectors with cloned cRSSs from Myc-Pvt1 sites (above) or a control, canonical RSS site (RSS). Vector also harbors an IRES-Thy1.2, marking successful viral transduction of Rag-reporter cells. **e**. Representative flow cytometric histograms of GFP+Thy1.2+ cells (left) and cumulative analysis from 3 independent experiments (right); data are represented as the mean ± s.e.m., and statistical testing is depicted as two-sided, unpaired t-tests; ****P ≤ 0.0001

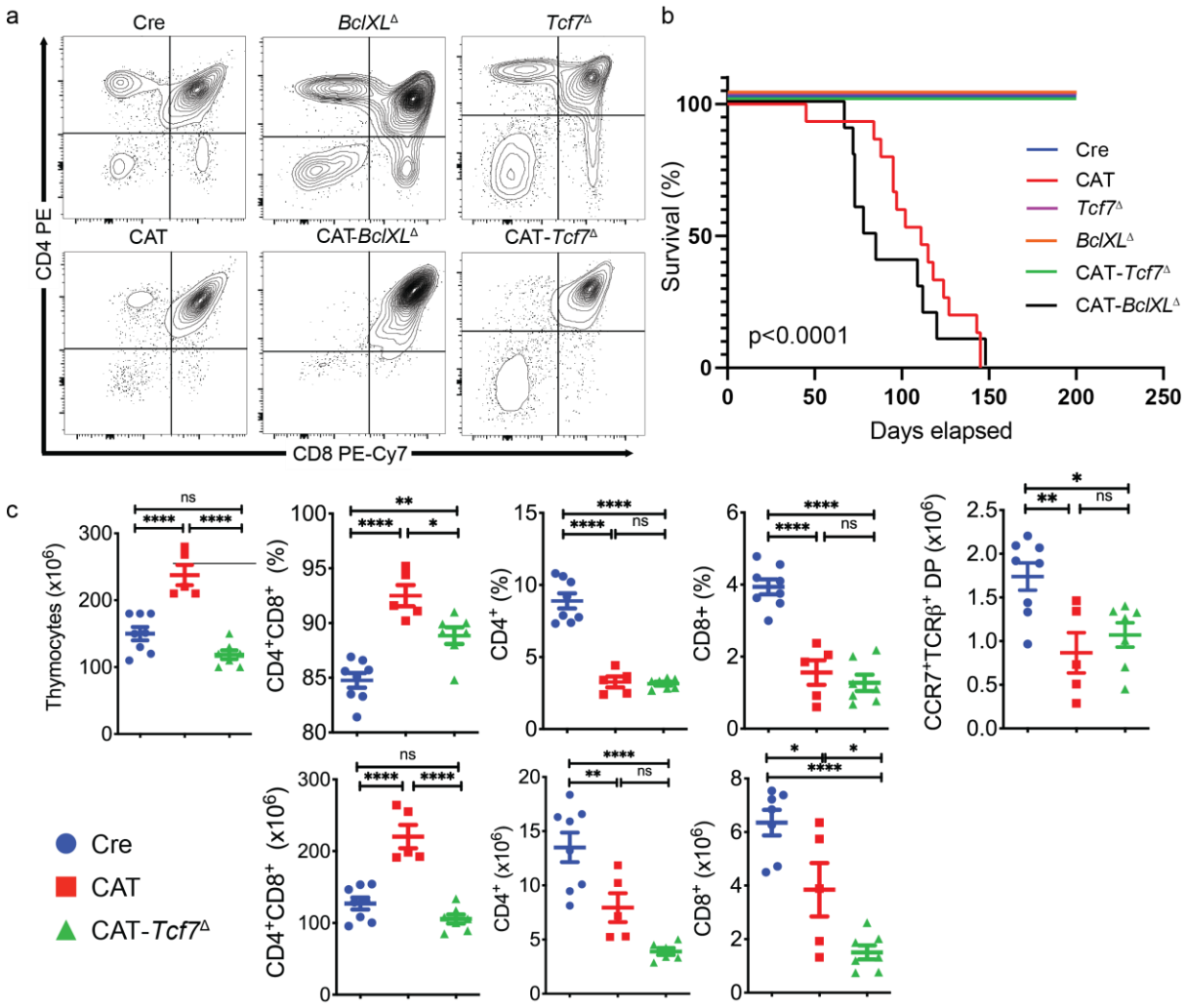
Tcf-1 is essential for β-catenin mediated transformation of DP thymocytes.

I next investigated processes that are altered by β-catenin activation that may contribute to transformation. Our earlier studies showed that thymocytes with activated

β-catenin are less sensitive to irradiation-induced cell death. This was linked to the overexpression of *BclXL*, the main antiapoptotic mediator in DP thymocytes (Ma et al., 1995), by showing that BclXL inhibitors abrogated the survival advantage of CAT thymocytes (Dose et al., 2014). Therefore, to determine if enhanced DP thymocyte survival contributes to their transformation, I conditionally ablated BclXL in DP thymocytes simultaneously with activation of β-catenin by crossing CAT mice with mice carrying a *BclXL*^{fl} allele (Walton et al., 2001). Compound *CD4-Cre/Ctnnb* ^{ex3fl}*/BclXL*$^{fl/fl}$*(*CAT-*BclXL*$^{\Delta}$*),* mice had the same developmental block in the transition of DP thymocytes to the single positive (SP) stages described in our previous reports(Guo et al., 2007) and succumbed to leukemia with the same frequency and latency as CAT mice (Figure 3.4a-b). Therefore, I concluded that the BclXL-mediated increase in survival of CAT DP thymocytes does not contribute to their developmental defect or their transformation *in vivo*.

I also hypothesized that the developmental block in CAT mice could be mediated by the DNA binding partner of β-catenin, Tcf-1, and may be contributing to transformation. I therefore produced compound *Cd4-Cre/ Ctnnb1*^{ex3fl}*/Tcf7*$^{fl/fl}$ (CAT-*Tcf7*$^{\Delta}$) mutant mice with conditional ablation of Tcf-1 in CAT DP thymocytes. Tcf-1 deletion did not resolve the developmental block and CAT-*Tcf7*$^{\Delta}$ thymocytes had reduced numbers of CCR7^{+}TCRβ^{+} post-selected DPs as well as fewer CD4^{+} and CD8^{+} SP thymocytes, as seen in CAT mice (Figure 3.4a,c). Nevertheless, ablation of Tcf-1 completely abrogated leukemogenesis (Figure 3.4b). This exciting finding indicates that Tcf-1 specifically mediates a fraction of the processes that are altered by β-catenin activation in DP thymocytes. Importantly, since these Tcf-1 mediated functions are critical for

Figure 3.4 Ablation of *Tcf7* but not *BclXl* rescues CAT lymphomas



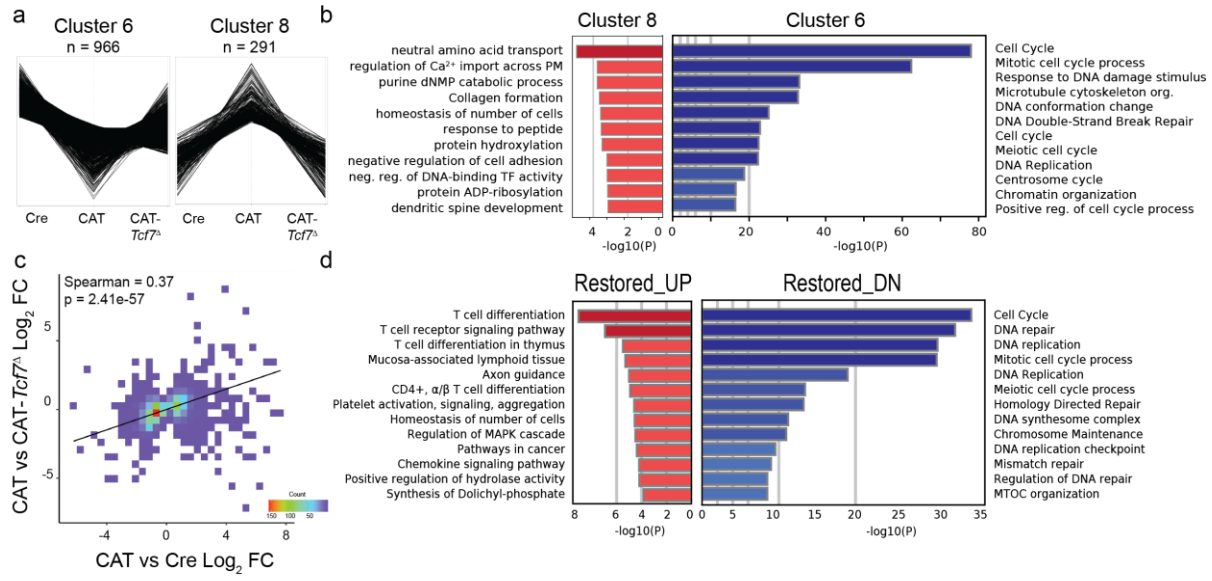
a. Representative flow cytometry contour plots of live thymocytes from 6–8-week-old mice stained for CD4 and CD8 to show developmental progression. b. Kaplan Meier survival curve analysis for *Cre* (n=10), CAT (n=15), *Tcf7*$^{\Delta}$ (n=10), *BclXL*$^{\Delta}$ (n=10), and those with co-deletions CAT-*Tcf7*$^{\Delta}$ (n=5) or CAT-*BclXL*$^{\Delta}$ (n=10). c. Flow cytometric histograms indicating the percentage (top) and total number (bottom) of thymocytes in the indicated late developmental stages in *Cre* (n=8), CAT (n=5), and CAT-*Tcf7*$^{\Delta}$ (n=7) mice; data are represented as the mean ± s.e.m., and statistical testing is depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.

leukemogenesis, the CAT-*Tcf7*$^{\Delta}$ mouse model offers a valuable tool to identify the mechanisms involved in transformation.

Aberrant β-catenin uses Tcf-1 to downregulate genome maintenance pathways.

To identify the Tcf-1 mediated processes involved in the transformation of CAT thymocytes by β-catenin, I performed RNA sequencing in pre-transformed (6 weeks) sorted DPs from CAT-*Tcf7*$^{\Delta}$ mice and compared to prior profiles of Cre and CAT. Unbiased clustering analysis (ExpressCluster, www.GenePattern.org) of expression levels (FPKM) across all genes in Cre, CAT, and CAT-*Tcf7*$^{\Delta}$ thymocytes highlighted two gene clusters that were upregulated (Cluster 8, n=291) or downregulated (Cluster 6, n=966) in CAT thymocytes and restored in the absence of Tcf-1 (Figure 3.5a). Interestingly, pathway analysis of the downregulated genes (http://www.metascape.org)

Figure 3.5 RNA-seq analysis confirms rescue and identifies Tcf1-controlled expression program

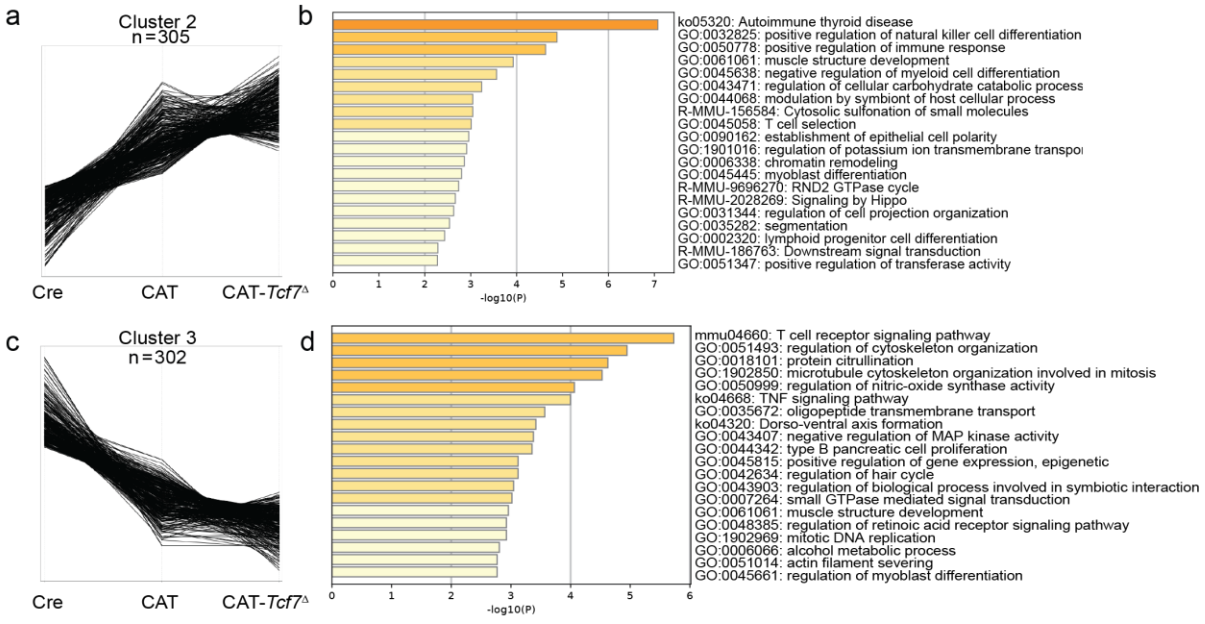


a. Unsupervised clustering of triplicate RNA-seq profiles (FPKMs) for Cre (n=3), CAT (n=3), and CAT-*Tcf7*$^{\Delta}$ (n=3) DP (CD4^{+}CD8^{+}) thymocytes; gene clusters associated with altered expression in CAT thymocytes that are restored to Cre levels in CAT-*Tcf7*$^{\Delta}$ are shown. **b**. Functional pathways enriched in Cluster 6 and Cluster 8 gene sets (Metascape, http://www.metascape.org). **c**. Spearman correlation of differentially expressed genes ($p < 0.05$) between CAT and either Cre or CAT-*Tcf7*$^{\Delta}$. **d.** Pathway analysis of restored DEGs that were upregulated (red) or downregulated (DN, blue) in CAT versus Cre and CAT-*Tcf7*$^{\Delta}$ (Metascape).

identified strong enrichment in cell cycle, DNA repair, and DNA replication pathways (Cluster 6, Figure 3.5b). In contrast, upregulated genes did not show strong enrichment for functionally relevant pathways (Cluster 8, Figure 3.5b). I also investigated two gene clusters that were altered in CAT thymocytes and remained at similar levels in CAT-Tcf7$^{\Delta}$ cells; these genes exhibited mild enrichment for basic T-cell and homeostatic functions (Figure 3.6). Taken together, this analysis suggests that downregulated genes in CAT thymocytes include those responsible for dysregulated genome maintenance pathways that may contribute to transformation. On the other hand, genes that are not restored could contribute to the continued developmental defects.

To determine the extent of rescue in CAT-*Tcf7*$^{\Delta}$ thymocytes at the transcriptional level further, I identified the differentially expressed genes (DEGs, $p < 0.05$) between

Figure 3.6 Unrestored gene clusters do not contribute to genomic instability



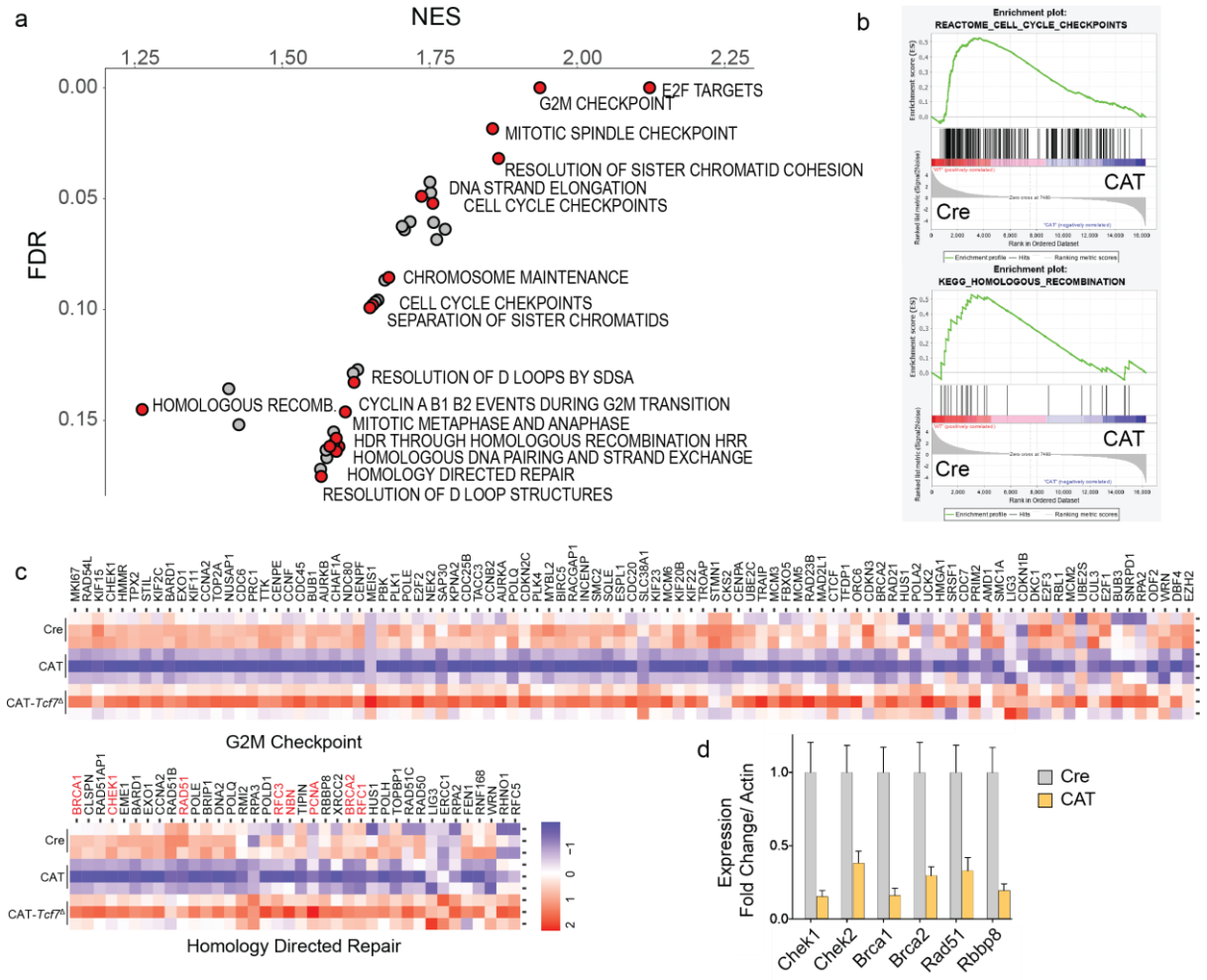
a,c. Profiles from unsupervised clustering of triplicate RNA-seq profiles (FPKMs) for Cre (n=3), CAT (n=3), and CAT-*Tcf7*$^{\Delta}$ (n=3) DP ($CD4^{+}CD8^{+}$) thymocytes; gene clusters associated with altered expression in CAT thymocytes that remain dysregulated in CAT-*Tcf7*$^{\Delta}$ are shown. Pathway analysis (Metascape) for **b**, Cluster 2 and **d**, Cluster 3.

CAT and either Cre or CAT-*Tcf7*$^{\Delta}$ cells, finding 1181 and 3735 DEGs, respectively. As expected, Spearman rank correlation of these two sets of DEGs established that deletion of *Tcf7* in CAT thymocytes reverts transcriptional programs to those in Cre thymocytes (Spearman = 0.37, P=2.41x10^{-57}; Figure 3.5c). Furthermore, the restored DEGs that were downregulated by β-catenin and returned by ablation of Tcf-1 also showed strong enrichment of genome maintenance pathways, specifically implicating homologous recombination, the error-free program for DSB repair (Figure 3.5d, blue). Interestingly, the upregulated DEGs showed enrichment in T-cell development pathways and are likely contributing to the DP block in CAT thymocytes (Figure 3.5d, red). Nevertheless, restoration of these genes upon the deletion of *Tcf-7* is insufficient to overcome the developmental stalling in β-catenin stabilized thymocytes.

Finally, I performed Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) on Cluster 6 genes, which accounts for magnitude and rank order of gene expression changes associated with specific cellular functions. The most significantly enriched pathways confirmed the Tcf-1 mediated dysregulation of cell cycle checkpoints and genome maintenance pathways in CAT thymocytes. (Figure 3.7a-b, red). I further identified the core enrichment genes at the leading edge of the GSEA profiles in several relevant pathway associations (Figure 3.7c). Importantly, core regulators of HR-mediated repair, including *Brca1*, *Brca2*, *Rad51*, and *Rbbp8,* and replication or cell cycle checkpoint engagement, including *Pcna, Rfc1, Chek1* and *Chek2*, were downregulated in CAT thymocytes. I confirmed the downregulation of these key genome maintenance genes using qRT-PCR. Pathways involving chromosome maintenance and processes necessary for mitosis were also overrepresented. Taken

together, these data provide strong transcriptional evidence that β-catenin contributes to the genomic instability of thymocytes through a Tcf-1 mediated downregulation of genes required for cell cycle regulation, faithful DNA replication, and DNA repair of DSBs, whereas a set of genes that are positively regulated by Tcf-1 are more likely involved in the developmental effects of β-catenin.

Figure 3.7 β-catenin uses Tcf-1 to downregulate replication, HR repair, and checkpoint control pathway genes



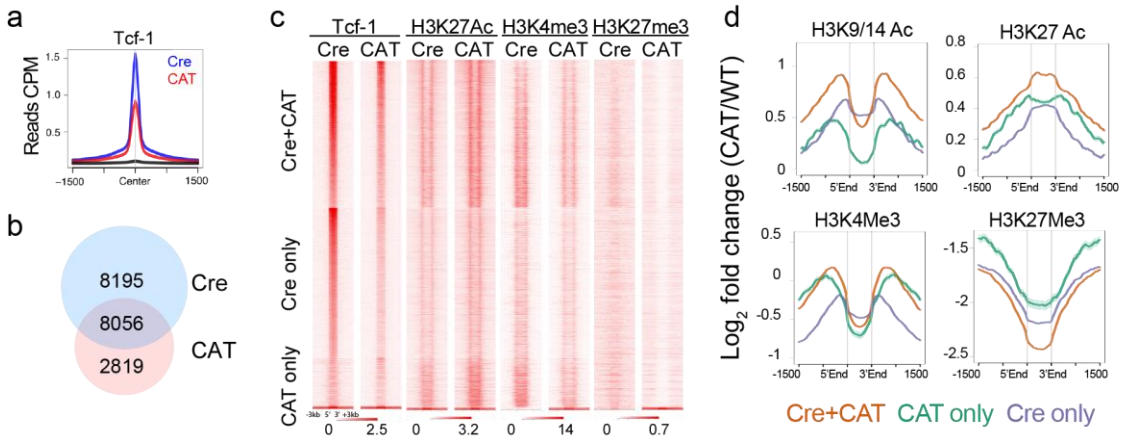
a. Gene set enrichment analysis (GSEA/ MSigDB) of Cluster 6 (Figure 3.4a) gene expression showing significant signatures (NES > 1) with relevant genome maintenance programs labeled and highlighted in red. **b.** Representative GSEA profiles and **c.** heatmaps of core enrichment genes of the indicated pathways; central replication and repair genes are highlighted (red). **d.** qRT-PCR confirmation of downregulation of core HR and checkpoint genes

Dominant active β-catenin directs novel Tcf-1 binding to HR repair and checkpoint control genes

I then hypothesized that expression changes in CAT thymocytes may reflect β-catenin mediated redirection of Tcf-1 binding and/or changes in chromatin accessibility due to its role in epigenetic regulation. Therefore, I mapped the distribution of Tcf-1 binding in CAT DP thymocytes using ChIP-seq. I identified 10,875 high confidence Tcf-1 peaks (p=10e-4) in these cells compared to 16,251 that had previously been detected in Cre DP thymocytes (Emmanuel et al., 2018). Tcf-1 peaks in CAT were less enriched overall compared to Cre thymocytes (Figure 3.8a). Although the majority of Tcf-1 peaks (8056) in CAT were shared with those in Cre DP thymocytes, there were 2819 new Tcf-1 peaks that are unique to CAT thymocytes (Figure 3.8b).

To assess the effect of β-catenin stabilization on the chromatin landscape, I determined the distribution of histone marks in CAT DP thymocytes by ChIP seq and compared to earlier data from Cre DP thymocytes (Emmanuel et al., 2018). I mapped histone 3 acetylation at lysines 9/14 (H3K9/14Ac), which marks active promoters, H3K27Ac, which marks active enhancers, H3K4me3, which is associated with activation of transcription, and H3K27me3, which marks poised/closed chromatin. Overall, near Tcf-1 peaks, enrichment for H3K9/14Ac and H3K27Ac marks increased whereas enrichment for H3K27me3 marks decreased in CAT compared to Cre DP thymocytes (log_2FC, Figure 3.8c-d). The degree of changes in histone mark enrichment varied depending on whether the Tcf-1 peaks were unique to Cre or CAT or shared between the two genotypes. The greatest increase of H3K9/14Ac and H3K27Ac and reduction in H3K27me3 marks was in shared Cre and CAT Tcf-1 sites compared to the novel CAT

Figure 3.8 Dominant active β-catenin alters the epigenetic landscape of DP thymocytes

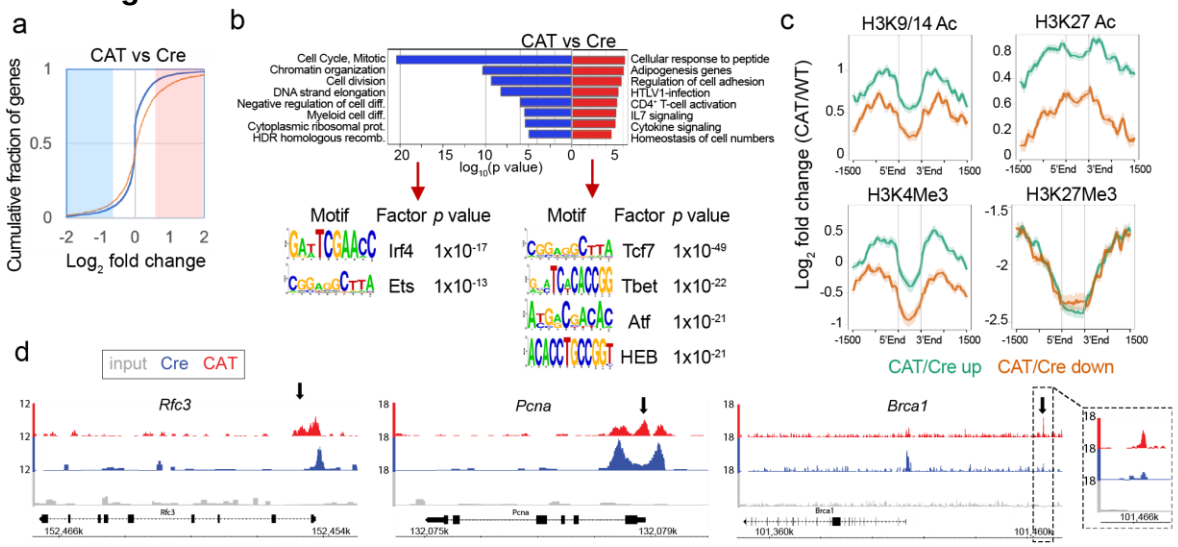


a. Histogram overlays of Tcf-1 binding centered on Tcf-1-bound regions in Cre DP thymocytes (±1.5 kb). **b.** Venn diagram of overlapping Tcf-1 peaks between Cre and CAT DP thymocytes. **c.** Heat maps of ChIP-seq peaks for Tcf-1 and indicated histone marks, centered on Tcf-1 binding (+/- 3kb) at shared or unique sites as indicated **d.** Comparative enrichment histograms of histone modifications associated with active promoters (H3K9/14Ac), active enhancers (H3K27ac), transcription (H3K4Me3), and closed/ poised chromatin (H3K27me3).

Tcf-1 sites, suggesting that accessibility increased more in the shared sites. In contrast to these changes in the H3K9/14Ac, H3K27Ac, and H3K27me3 marks, H3K4me3 showed little change in Cre versus CAT for shared or CAT only Tcf-1 peaks and showed reduced enrichment in Cre only Tcf-1 peaks. Thus, β-catenin stabilization enhances accessibility in Tcf-1 bound sites consistent with the classical understanding of its epigenetic functions.

The novel Tcf-1 binding sites in CAT cells and the corresponding chromatin changes suggest that these sites may affect the expression of the associated genes. After annotating, I compared expression changes of these genes in Cre vs CAT to all genes using cumulative distribution function (CDF) analysis. Consistent with the finding that stabilization of β-catenin results in an overall increase in chromatin accessibility, the

Figure 3.9 CAT-unique Tcf-1 binding is enriched at HR repair and checkpoint control genes

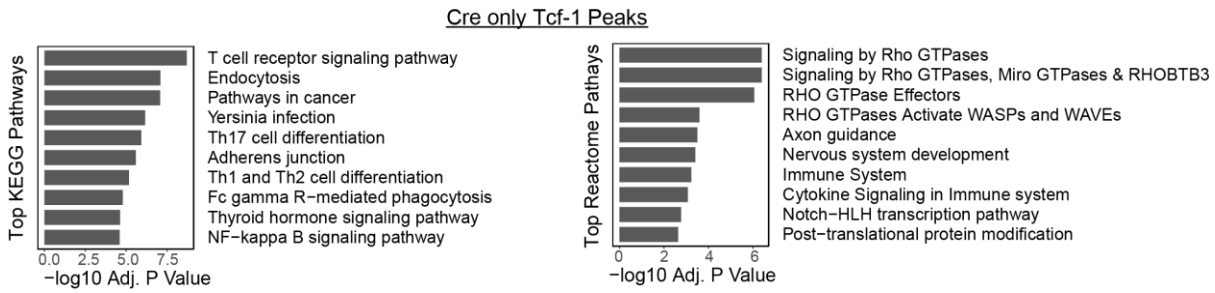


a. CDF plot of expression changes in genes uniquely bound by Tcf-1 in CAT DP thymocytes (n=2819, red line) compared to all genes (blue line). **b**. Pathway analysis of downregulated (blue) and upregulated (red) genes uniquely bound by Tcf-1 in CAT (top) and the most significantly enriched transcription-factor-binding motifs (HOMER, bottom). **c.** Comparative enrichment histograms for histone modifications (H3K9/14Ac, H3K27ac, H3K4Me3, and H3K27Me3) centered on CAT-unique Tcf-1 binding sites corresponding to up- or downregulated genes in CAT versus Cre as indicated. **d**. Representative Tcf-1 ChIP-seq enrichment tracks (Integrated Genome Browser) at key replication and repair genes that are differentially expressed and have novel Tcf-1 sites in CAT thymocytes (arrows).

genes associated with novel Tcf-1 binding were significantly more likely to be upregulated in CAT thymocytes (K&S $p<0.0001$, Figure 3.9a). I then identified two groups of significantly up ($Log_2FC>0.56$) or downregulated ($Log_2FC<-0.56$) genes with novel Tcf-1 peaks in CAT thymocytes. Pathway enrichment analysis (http://www.metascape.org/) revealed that the group of upregulated genes with new Tcf-1 peaks was enriched in genes involved in general T cell functions, whereas the group of downregulated genes was enriched in HR DNA repair and cell cycle pathways (Figure 3.9b). I also investigated the Tcf-1 peaks that were unique to Cre thymocytes, which did not appear to contribute to genome maintenance pathways and were also

enriched for homeostatic or basic T-cell functions (Figure 3.10).

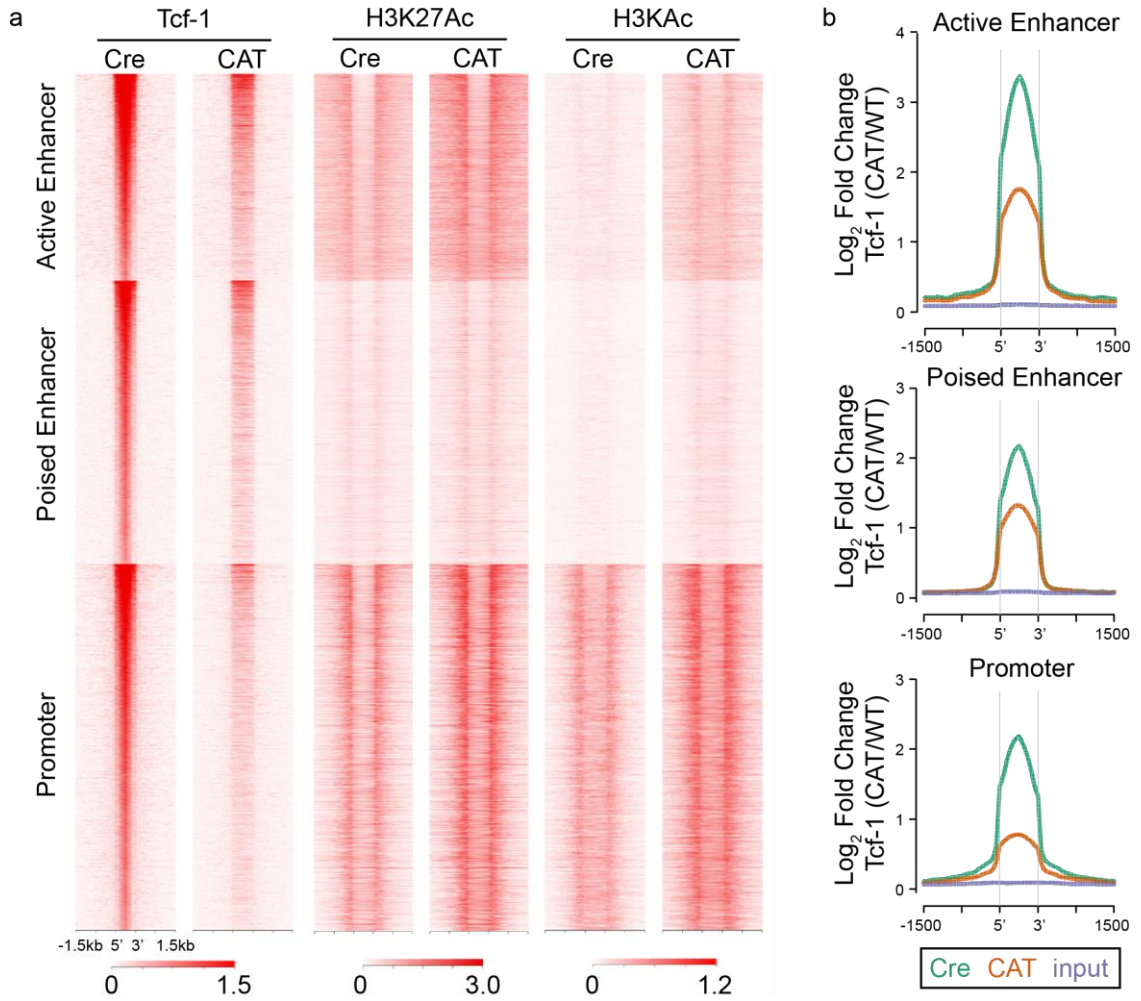
Figure 3.10 Cre-unique Tcf-1 binding is not enriched for genome maintenance



Pathway enrichment analysis (g:Profiler, biit.cs.ut.ee/gprofiler) for genes associated with Tcf-1 binding in Cre that are lost in CAT mice (n=8,195 peaks); top enriched KEGG and Reactome pathways are shown

I further characterized the chromatin profiles specifically at these CAT-unique, Tcf-1-bound DEGs and found that novel Tcf-1 binding led to greater increases in accessibility in upregulated compared to downregulated (Figure 3.9c) genes. Additionally, I evaluated Tcf-1 binding in regulatory elements, as were previously defined by histone profiling in Cre DP thymocytes (Emmanuel et al., 2018). I found some mild bias toward loss of Tcf-1 at promoters, but both poised and active enhancers also showed reduced enrichment for Tcf-1 (Figure 3.11). Similarly, representative tracks show novel Tcf-1 binding sites in both promoters and putative enhancers of important genome maintenance genes *Rfc3*, *Pcna* and *Brca1* (Figure 3.9d), which were identified in the transcriptional analysis (Figure 3.7). To understand how Tcf-1 coordinates these opposing expression programs further, I performed motif enrichment analysis on novel Tcf-1 binding sites associated with genes that were up- or downregulated in CAT thymocytes. Tcf-1 peaks in upregulated genes were enriched for the Tcf-1 binding motif (*Tcf7)* in addition to other key transcriptional regulators in T-cell development such as

Figure 3.11 Tcf-1 binding at promoters and enhancers



a. Heat maps of ChIP-seq peaks for Tcf-1 and indicated histone marks centered on Tcf-1 binding (+/- 1.5kb) at promoters, poised enhancers, or active enhancers, as previously defined by histone profiles in Cre DP thymocytes **b.** Corresponding enrichment histograms.

Tbet and HEB. In contrast, Tcf-1 peaks in downregulated genes were not enriched for their own consensus motif, suggesting Tcf-1 acts in distinct complexes with other transcriptional regulators to enforce these different outcomes on chromatin and transcription. Altogether these findings provide a mechanistic explanation for the downregulation of HR repair and cell cycle genes in CAT thymocytes and directly link these expression changes to Tcf-1 binding. This outcome also explains why ablation of

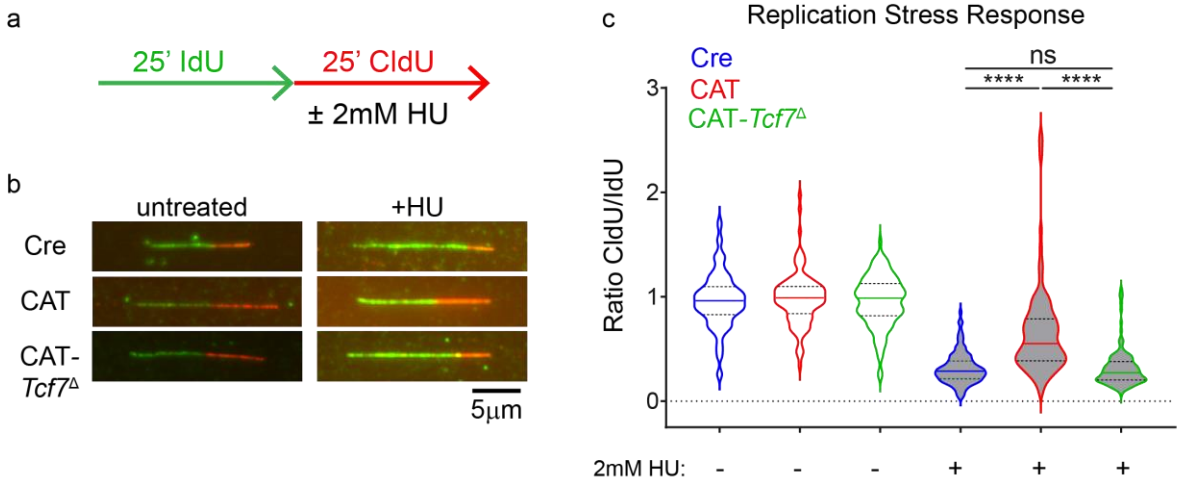
Tcf-1 eliminates these expression changes and highlights their direct mechanistic link to transformation.

Altered replication process and response to replication stress in CAT thymocytes.

My previous findings established that the DSBs in the *Myc-Pvt1* locus were Rag-independent (Figure 3.3-3). On the other hand, the transcriptional evidence suggested that replication checkpoints may be impaired in pre-transformed CAT thymocytes. Moreover, failures in replication and repair programs are increasingly appreciated as a source of genomic instability in cancer (Tubbs and Nussenzweig, 2017). Therefore, to test replication fork fidelity directly, I performed nascent DNA fiber assays in DP thymocytes cultured with the replication stressor, Hydroxyurea (HU). In these experiments, newly replicated DNA is pulse-labeled by two successive co-cultures with nucleotide analogs 5-Iodo-2'deoxy-uridine (IdU) and 5-Cloro-2'deoxy-uridine (CldU). The fidelity of replication fork arrest due to replication stress checkpoints is tested through the addition of 2mM HU during the CldU pulse (Figure 3.12a). Compared to the equal IdU and CldU track lengths in untreated thymocytes, both Cre and CAT-*Tcf7*$^{\Delta}$ DPs exhibit shortening of CldU tracks during HU co-culture conditions (Figure. 3.12b-c). CAT thymocytes, however, continue to replicate under HU stress leading to higher ratios of CldU to IdU track lengths, suggesting that replication fork arrest is impaired (Figure. 3.11b-c).

I then tested whether CAT thymocytes were able to repair DNA damage generated by etoposide (ETP), which freezes topoisomerase II after generating DSBs and prevents re-ligation, with strongest effects on early replicating genomic regions

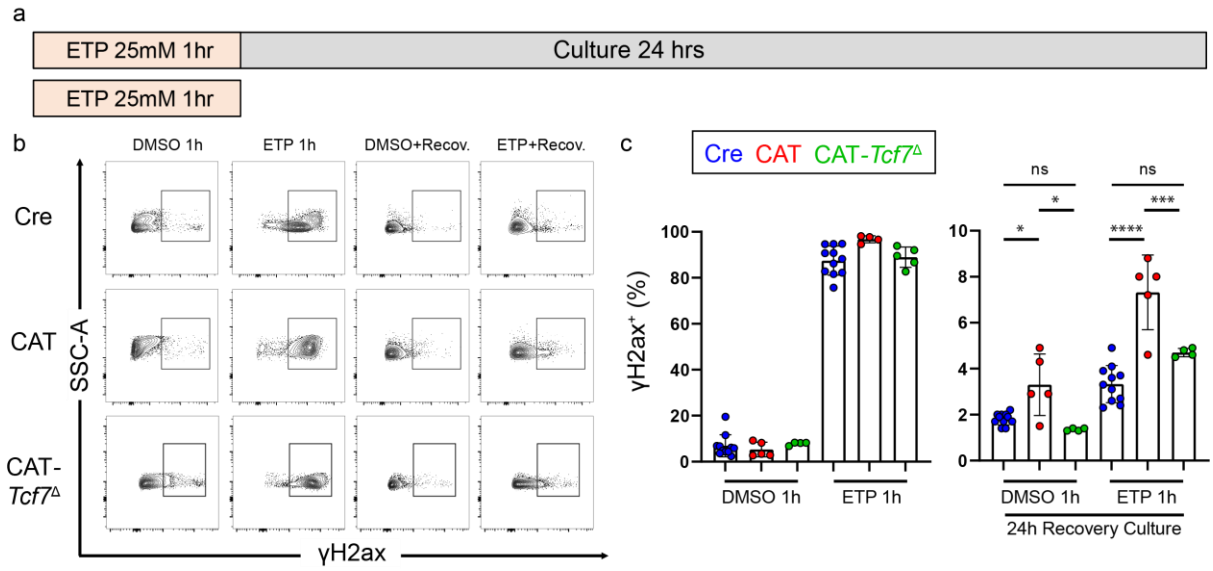
Figure 3.12 CAT thymocytes are less sensitive to replication stress



a. Schematic of assay and **b.** representative nascent DNA fibers from Cre, CAT, and CAT-*Tcf7*$^{\Delta}$ CD4^{+}CD8^{+} (DP) thymocytes cultured in successive 25 minute pulses of media containing nucleotide analog IdU (green) followed by CldU (red) ± 2mM Hydroxyurea (HU) during the CldU pulse. **b.** Ratios of CldU:IdU tract lengths from untreated (vehicle) and HU-treated thymocytes in the replication stress assay. Data are representative of three biological replicates per condition in two independent experiments (n >100 fibers/ condition); results depicted as violin plots with median (solid) and quartile (dotted) lines, with statistics from two-sided, unpaired t-tests; ****P ≤ 0.0001.

(Tubbs and Nussenzweig, 2017). After a brief one-hour culture in 25mM ETP, nearly all thymocytes were positive for the early DSB marker, phosphorylated H2ax (γH2ax), suggesting that the detection of DNA breaks and sensitivity to DNA damage induction was normal in CAT thymocytes. However, after ETP wash-out and 24 hours of recovery in fresh media, CAT thymocyte cultures retained more γH2ax^{+} cells (Figure 3.13b-c). This finding suggests that CAT thymocytes are either less proficient at repairing DSBs and/ or they fail to initiate apoptosis to eliminate unrepaired cells. Taken together, these findings support a model wherein CAT thymocytes fail to recognize replication stress and allow unrepaired DSBs to escape cell cycle checkpoints that typically restrict replication-mediated damage to the error-free, HR-repair in S-phase. This would permit *Myc-Pvt1* breaks due to impaired replication and/or repair programs to coexist with Rag-

Figure 3.13 CAT thymocytes retain replication-associated DNA damage



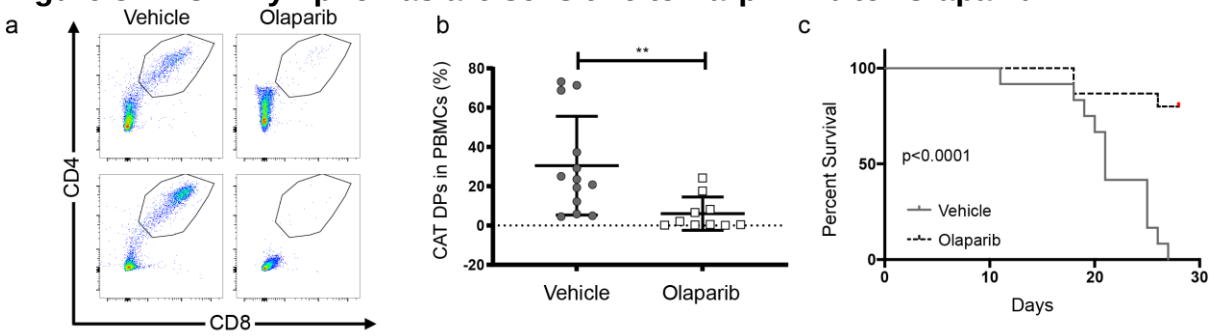
a. Schematic of assay and **b.** representative flow cytometric contour plots of intracellular staining for DNA damage marker γH2ax in live DP thymocytes cultured *ex vivo* for 1 hour in 25mM etoposide (ETP) or vehicle control (DMSO), followed by 24-hour recovery culture in fresh media. **c**. Cumulative analysis of Cre (n=11), CAT (n=5), and CAT-*Tcf7*$^{\Delta}$ (n=4) DP cultures showing the percent of live, γH2ax^{+} cells from two independent experiments in the indicated conditions. Data are represented as the mean ± s.e.m. and statistical testing depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.

mediated breaks in *Tcra* during G1, providing the conditions for error-prone DSB repair by NHEJ resulting in chromosomal translocations.

β-catenin induced leukemias are sensitive to PARP inhibitors.

I then hypothesized that this β-catenin and Tcf1 mediated downregulation of HR DNA repair genes could also render the resulting lymphomas sensitive to Parp inhibitors (PARPi). Olaparib is a PARPi that was first approved for use in cancer patients that have lost HR activity (most commonly loss of BRCA1/2) due to its synthetic lethal targeting of cancer cells that are now overly dependent on backup DNA repair pathways that require Parp1 (Lord and Ashworth, 2017). To test this hypothesis, I transferred CAT

Figure 3.14 CAT lymphomas are sensitive to Parp inhibitor Olaparib

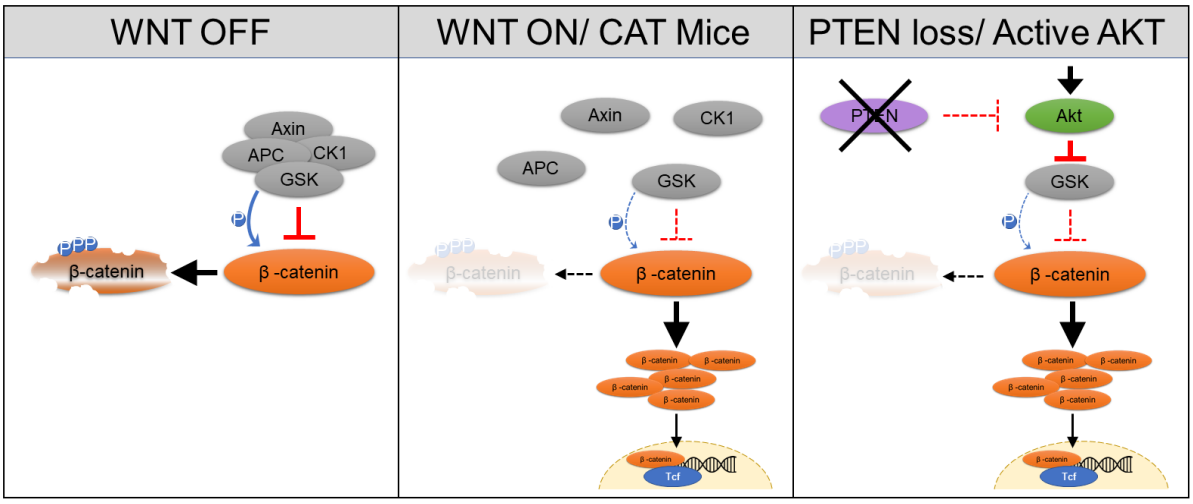


a. Representative flow cytometry pseudocolor dot plots and **b**. cumulative quantification of engraftment flow of live, $CD4^+CD8^+$ DPs in peripheral blood (PB) of $Rag^{-/-}$ mice (n=10/ donor) that received CAT lymphomas (n=3). PB was sampled 2-weeks post transfer during treatment course, with recipients split evenly between Olaparib (50mg/kg/day) or vehicle controls. Data represented as the mean ± s.e.m. and statistical testing depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01 **c.** Kaplan Meier survival curve of CAT lymphoma recipients in the indicated conditions.

lymphomas into the periphery of $Rag^{-/-}$ mice and treated lymphoma recipients with a three-week course of olaparib (50mg/kg daily for 5 days) or vehicle control. Excitingly, tracking of $CD4^+CD8^+$ DPs in peripheral blood showed that olaparib treatment controlled the expansion of CAT lymphomas and prolonged survival of transplant recipients (Figure 3.14).

To connect these findings to human leukemia directly, I next investigated whether PARPi sensitivity might be applicable more broadly in human T-ALL patients. Loss of PTEN, the second most frequent mutation (20-47%) in T-ALL patients (Gutierrez et al., 2009; Palomero et al., 2007), activates AKT and disrupts the GSK3B/CK1 destruction complex, leading to stabilization of β-catenin (Figure 3.15). Furthermore, elevated β-catenin has been reported in patients with mutations activating NOTCH (> 50%), making β-catenin a common feature of T-ALL pathology (Girardi et al., 2017). I therefore compared the transcriptional profiles derived from five CAT lymphomas to PTEN target genes identified in human T-ALL using GSEA

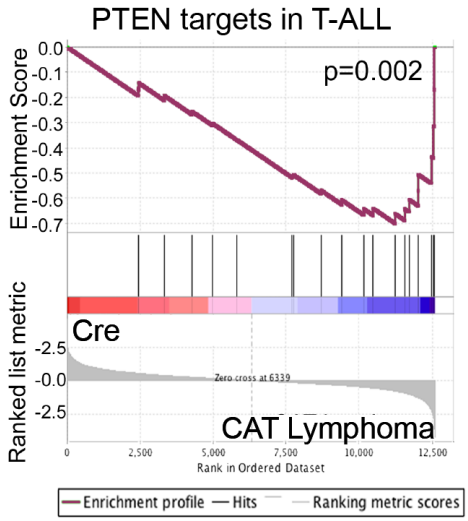
Figure 3.15 Schematic of Wnt signaling and PTEN loss on β-catenin activity



When no WNT signaling is present, the GSK3B(GSK)/CK1 destruction complex (grey) phosphorylates (blue) β-catenin (orange), initiating its degradation by the proteasome. In the presence of a WNT ligand, or when the phosphorylation sites are excised in CAT mice, β-catenin is stabilized, accumulates, and translocates to the nucleus where it interacts with Tcf-1 (blue) to alter chromatin and transcription. Similarly, mutations leading to loss of PTEN (purple) or activation of AKT (green) signaling inhibit GSK kinase activity, leading to indirect stabilization of β-catenin.

(Palomero et al., 2007). Like in PTEN deficient T-ALL, these genes were also downregulated in CAT lymphomas, supporting the relevance of the CAT mice to these human tumors (Figure 3.16). Furthermore, I extracted drug sensitivity data from the Genomics of Drug Sensitivity in Cancer (GDSC) database, which contains normalized IC_{50} data on hundreds of compounds in more than 900 cancer cell lines. In comparison to all lines, I found that T-ALL samples were more sensitive to three separate WNT inhibitors (CHIR-99021, SB-216763, XAV-939) as well as four compounds targeting AKT, which includes both *NOTCH1* and *PTEN* mutant lines (Figure 3.17a). Importantly, 3/4 PARPi compounds in this database (olaparib, ABT-88 (veliparib), and AG-014699 (rucaparib)) were also more sensitive in T-ALL lines (Figure 3.17a). Additionally, expanded analysis on all leukemia and lymphoma cell lines found similar sensitivity profiles (Figure 3.17b), suggesting relevance to other hematologic malignancies.

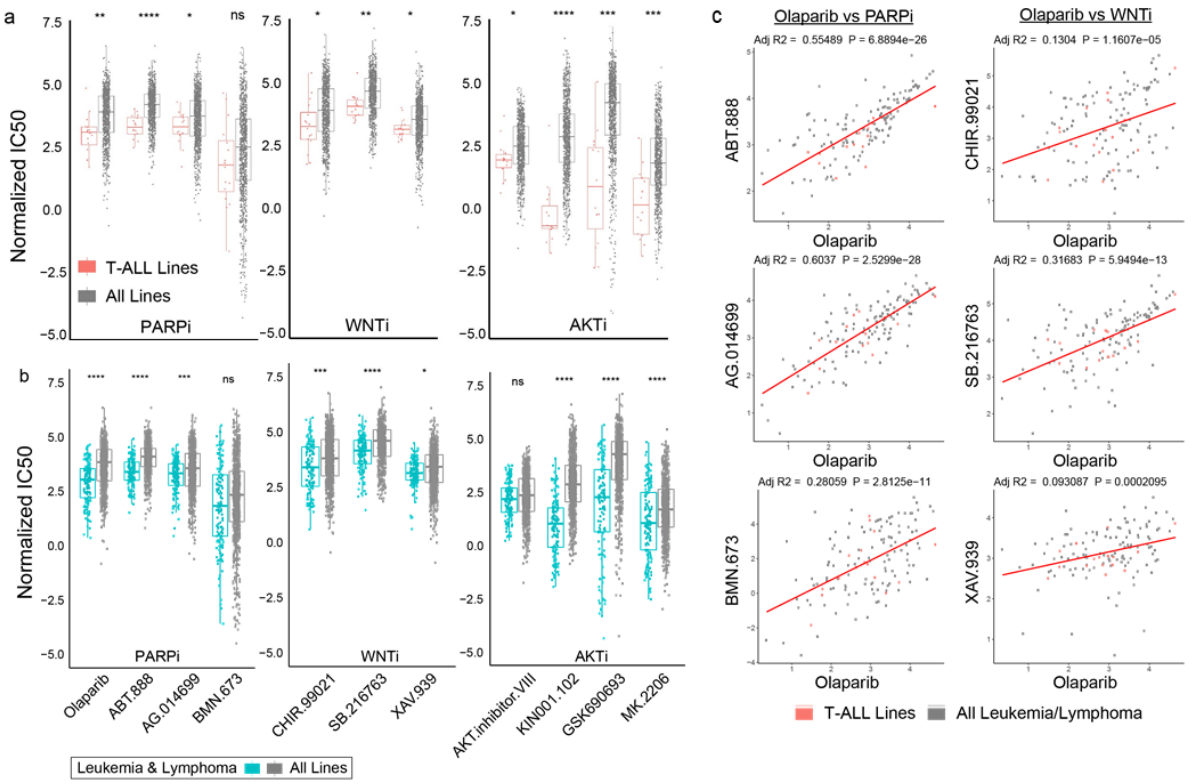
Figure 3.16 CAT lymphomas resemble transcriptional profiles of T-ALL samples with *PTEN* mutations



Gene set enrichment analysis comparing PTEN targets in T-ALL gene signature (Palomero et al, 2007) to RNA expression profile derived from five CAT lymphomas.

Moreover, the degree of drug sensitivity to PARP targeting compounds correlated with that of WNT targeted therapies (Figure 3.17c). Taken together, these findings support a paradigm where stabilization of β-catenin signaling leads to dysregulated HR in T-ALL, opening a therapeutic opportunity for synthetic lethal drugs in a patient population with urgent needs for targeted therapies.

Figure 3.17 T-ALL and leukemia and lymphoma cell lines are more sensitive to both WNT and PARP inhibitors



Normalized IC50 data from the Genomics of Drug Sensitivity in Cancer (GDSC) database comparing **a.** T-ALL (n=16) or **b.** leukemia/ lymphoma (n=150) cell lines to all other cancer cell lines (n=974, n = 840, respectively) for sensitivity to inhibitors targeting PARP (PARPi), WNT (WNTi), or AKT (AKTi), as indicated. Data represented as the mean ± s.e.m. and statistical testing depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. **c.** Correlations of olaparib sensitivity (IC50) to other PARPi drugs (left) or WNTi compounds (right) in all leukemia and lymphoma cell lines (T-ALL lines highlighted in red). Linear regressions analysis performed in R assuming Gaussian distribution and correlations represented as adjusted R-squared values (adj R2).

Discussion

Chromosomal translocations involving the immunoglobulin (Ig) loci in B-cells and the T-cell receptor (Tcr) loci in T-cells have been frequently reported in leukemia and lymphoma patients (Lieber, 2016). A common partner to Ig and Tcr inter-chromosomal fusions is the *Myc-Pvt1* locus whose abnormal expression has been causally implicated

in a variety of cancer types (Lieber, 2016; Tseng et al., 2014). In this study, I provide a mechanistic explanation for a recurrent, spontaneous *Tcra/Myc-Pvt1* translocation promoted by aberrantly stabilized β-catenin in late thymocyte development. I demonstrate that although DSBs at the *Tcra* breakpoint site of the translocation are generated by the Rag recombinase, DSBs at the *Myc-Pvt1* site result from Tcf-1 controlled impairment of the HR repair and cell cycle checkpoint processes. Specifically, stabilized β-catenin redirects Tcf-1 binding to novel sites associated with genes involved in replication-associated DNA damage resulting in their downregulation. I further demonstrate that this HR repair impairment implicated in the etiology of the translocations opens an exciting therapeutic opportunity for PARP inhibitors, with potential for use more widely in Wnt/β-catenin stabilized leukemias and lymphomas.

Tcf-1, a member of the Tcf/Lef family of transcriptional regulators, plays critical roles at multiple stages of T-cell development and function, including the DP stage investigated here (Escobar et al., 2020; Mielke et al., 2019; Zhao et al., 2021). Classically, Tcf-1 can both negatively regulate genes when in complex with factors like Groucho, while also positively regulating Wnt-target genes, acting as the canonical DNA-binding partner of β-catenin (Mosimann et al., 2009). Accordingly, here I show widespread and bidirectional alterations in Tcf-1 dependent transcriptional programs when β-catenin is aberrantly stabilized. I show that β-catenin can partially redirect Tcf-1 from its normal role and distribution on chromatin in DP-thymocytes. Interestingly, genes that are upregulated upon novel Tcf-1 binding in CAT thymocytes contain its consensus motif, while downregulated genes exhibit new Tcf-1 binding to alternative TF motifs. This may help explain how Tcf-1 can simultaneously coordinate different

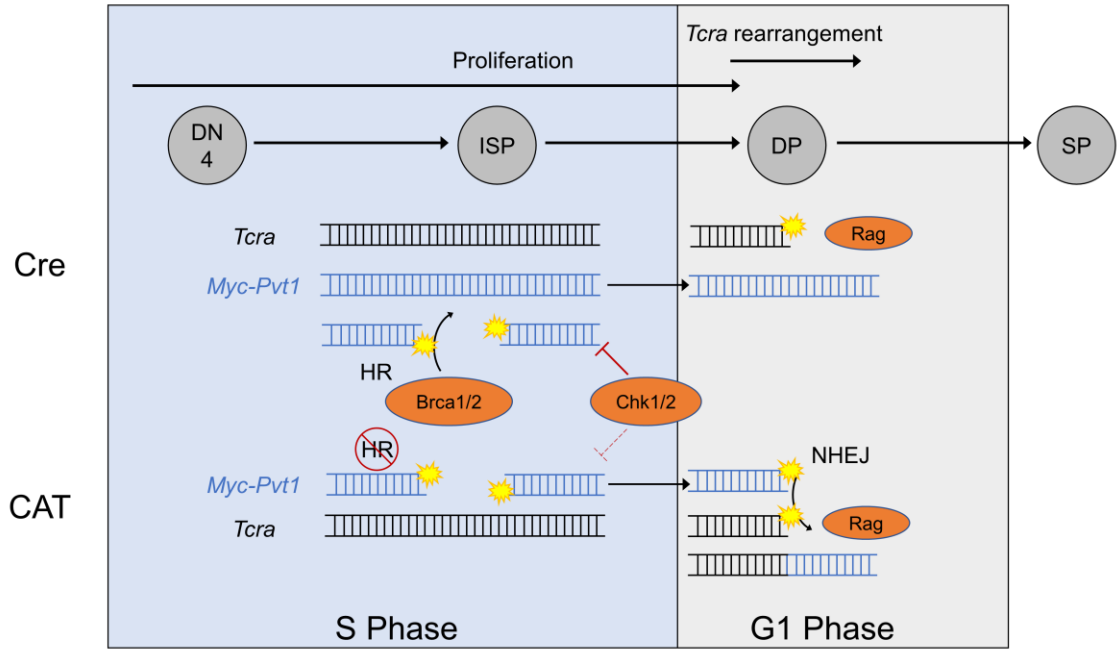
outcomes on gene regulation, which has been invoked by our group and others studying Tcf-1 and its different regulatory outcomes in T cells (Emmanuel et al., 2018; Zhao et al., 2021). At present, my analysis cannot define precise rules for secondary factors. Future studies will need to address TF-complex formation and coordination with Tcf-1 throughout thymocyte development.

Tcf-1 has been appreciated as an epigenetic regulator and has been predominately associated with enhancing chromatin accessibility (Johnson et al., 2018). Work from our lab and others has confirmed widespread alterations in chromatin accessibility mediated by Tcf-1 in both DP thymocytes and T-regulatory cells (Emmanuel et al., 2018; Quandt et al., 2021; van der Veeken et al., 2020). Accumulating evidence suggests that Tcf-1 acts in various molecular complexes with other transcriptional and epigenetic regulators to alter expression programs by modifying the chromatin landscape. More recently, Tcf-1 has also been found to have intrinsic HDAC activity (Xing et al., 2016). Here I show that stabilized β-catenin directs Tcf-1 to novel binding sites and significantly enhances chromatin accessibility at both existing and novel Tcf-1 binding sites. This fits within the general paradigm of Tcf-1 acting on top of a previously existent T-cell chromatin state to control stage-specific programs by fine-tuning the chromatin landscape during T cell development and activation (Emmanuel et al., 2018; Johnson et al., 2018). In agreement with this idea, novel Tcf-1 binding was frequently found proximal to previously existing Tcf-1 binding sites in Cre thymocytes (Fig 4g), which could be evidence of alterations in regulatory complexes at these loci. Thus, deregulated activation of β-catenin in DP thymocytes

engages Tcf-1 to predominately enhance chromatin accessibility and promote aberrant gene expression programs.

The present findings highlight a Tcf-1-controlled cluster of downregulated genes that is significantly enriched in genome maintenance pathways, which I implicate in DSB retention at the Myc-Pvt1 locus. Importantly, this includes core regulators of homologous recombination, such as *Brca1* and *Brca2*, indicating that an impaired error-free DSB repair underlies the translocation formation, as has been shown in other translocation studies (Scully and Livingston, 2000; Vasanthakumar et al., 2015). Furthermore, pathway analysis strongly implicated altered replication biology, which can also lead to DSBs through a variety of mechanisms. These include replication fork errors or under-replicated DNA leading to chromosome compaction and segregation errors that produce DSBs and translocations when the mitotic G2 checkpoint arrest fails to facilitate repair. Indeed, prior work in APC-mutant colon cancers has shown that overactive Wnt/β-catenin signaling contributes to anaphase bridges that lead to high levels of genomic instability in these tumors (Aoki et al., 2007). In this study, I provide functional evidence that CAT thymocytes are less efficient at resolving DSBs generated by the replication poison etoposide and are also insensitive to HU stress in nascent fiber assays, which can lead to under replicated DNA regions that are a danger to genomic stability. Taken together, I suggest a model wherein Tcf-1 mediated downregulation of the HR and repair machinery produces replication failures that lead to DSBs, which are also not properly recognized for repair and fail to initiate S/G2 arrest. This would allow the coexistence of breaks in the Myc-Pvt1 locus that traverse the cell cycle and are

Figure 3.18 Model of *Tcra/Myc-Pvt1* fusion formation in CAT thymocytes



During the later stages of T-cell development, proliferative DN4/ISP cells transition to the DP stage where cell cycle is arrested in G1 to facilitate Rag expression and *Tcra* rearrangement. In WT (Cre) cells, breaks that are sustained in the *Myc-Pvt1* locus are repaired by BRCA1 and HR during an S/G2 cell cycle arrest mediated by checkpoint kinases Chek1 and Chek2. In CAT cells, both HR and cell cycle checkpoints are downregulated, leading to the persistence of *Myc-Pvt1* breaks into G1. This allows for coexistence of replication-associated DSBs in *Myc-Pvt1* and Rag-mediated DSBs in *Tcra*, where NHEJ repair and ligation produces translocations. (DN = double negative ($CD4^-CD8^-$), ISP = immature single positive, DP = double positive ($CD4^+CD8^+$), SP = single positive (mature CD4+ or CD8+))

aberrantly joined to rearranging receptor loci in the following G1, placing *Myc* under the control of highly active regulatory elements (Figure 3.18).

The present mechanistic study of *Tcra/Myc-Pvt1* fusion conditions also uncovered an HR-deficiency that renders the subsequent lymphomas sensitive to the Parp inhibitor, olaparib. Human T-ALL patients, particularly adults and pediatric cases after relapse, have poor outcomes and few therapeutic options (Samra et al., 2020). Although mutations in β-catenin are not directly reported in T-ALL, stabilization and

elevated Wnt signaling is seen broadly including pediatric cases, NOTCH-driven transformation, PTEN loss with or without activating NOTCH, maintenance of leukemic stem cells, as well as several subtypes of adult differentiated T-ALL (Bellei et al., 2006; Groen et al., 2008; Lento et al., 2013; Ng et al., 2014; Ram-Wolff et al., 2010). Excitingly, I highlight the potential for a wider application of PARPi in T-ALL due to β-catenin associated HR deficiency as nearly all T-ALL cell lines exhibit sensitivity compared to other cancer lines. Importantly, sensitivity to WNT pathway inhibitors correlated with PARPi sensitivity within leukemia and lymphoma lines. This finding suggests T-ALL patients may see significant benefits with the addition of PARPi to therapeutic regimens. Future work should address whether β-catenin stabilization can also lead to dysregulation of homologous recombination in non-T cell cancers, particularly B-cell leukemias harboring *Ig-Myc* translocations, and with the potential for use as a biomarker to predict PARPi sensitivity.

CHAPTER IV

Germline *CHEK2* Variants as Risk Alleles for Clonal Hematopoiesis and Hematopoietic Malignancies in Humans and Mice

The data in this chapter are adapted from a manuscript currently in preparation: Stephen Arnovitz*, Ryan Stubbins*, Anase S. Asom*, Imo Akpan, Daniel Mendez, Maya Lewinsohn, Matthew Jones, Jason Cheng, Ashwin Koppayi, Michael Drazer, and Lucy A. Godley. Target journal: *Nature Medicine*. *These authors contributed equally to this work.

In this chapter, I designed and performed experiments, analyzed the data, and wrote and edited the manuscript. R.S. performed experiments, analyzed data, and helped write the manuscript. R.S. and A.S.A. analyzed clinical cohort data. D.M. helped track a mouse cohort. J.C. performed IHC immunophenotyping on mouse tumors. M.L. and M.J. helped establish the *Chek2* p.I161T mouse line. A.K. and R.S. adapted and ran variant calling pipeline for MSK-IMPACT data. M.D. performed UK Biobank analysis. L.A.G. conceived the study, analyzed the data, and edited the manuscript. I would also like to acknowledge Linda Degenstein and Xiu Chen in the Transgenics Core at University of Chicago for help in establishing the knock-in mouse line.

Introduction

The *CHEK2* gene encodes the checkpoint kinase 2 (CHK2) protein, an integral effector kinase in the ATM-CHK2-BRCA1 DNA damage response (DDR) pathway (Falck et al., 2001; Matsuoka et al., 1998). Detection of a DNA double-stranded break by the Mre11 complex and other proteins leads to activation of ATM, which subsequently phosphorylates several CHK2 residues, bringing the activation loops of the kinase domains into proximity and promoting multimerization and further autophosphorylation (Lee and Paull, 2007; Oliver et al., 2007; Stracker et al., 2009). Activated CHK2 subsequently phosphorylates several downstream targets, such as *TP53* and *BRCA1*, which then modulate cell cycle, apoptosis, autophagy, and DNA repair by homologous recombination (Bartek et al., 2007; Lee et al., 2000; Stolarova et al., 2020; Zhang et al., 2004).

Germline variants in *CHEK2* have been described in several populations, the most common being c.1100delC/p.T367fs (gnomAD allele frequency = 0.00181) , c.444+1G>A (gnomAD allele frequency = 0.00022), and c.599T>C/p.I200T, previously known as p.I157T under the hg19 nomenclature (gnomAD allele frequency = 0.00434) (Bell et al., 1999; Cybulski et al., 2004; Wu et al., 2001). Germline *CHEK2* variant carriers have been shown to be at higher risk for cancers of the breast (Couch et al., 2017; Cybulski et al., 2004; Desrichard et al., 2011; Girard et al., 2019), prostate(Cybulski et al., 2004; Seppälä et al., 2003), thyroid (Cybulski et al., 2004; Kaczmarek-Ryś et al., 2015), gastrointestinal system (Obazee et al., 2019; Zhunussova et al., 2019), and urinary tract (Kinnersley et al., 2016). However, there has been controversy around the degree of risk conferred by *CHEK2* variants given that the allele

frequency of many common variants can exceed 1% in some populations (*e.g.,* p.I200T allele frequency in gnomAD of 2.5% in Finnish and 4.6% in Estonians; p.T367fs and pI200T combined in up to 15% of Polish groups (Cybulski et al., 2004; Sutcliffe et al., 2020). The reasons for this are complex, but may be explained partly by the incomplete penetrance of many *CHEK2* variants among other factors (Meijers-Heijboer et al., 2002).

The risk that pathogenic (P) and likely pathogenic (LP) *CHEK2* variants impart on developing hematologic malignancies remains an active area of investigation. Two studies identified non-synonymous *CHEK2* variants as being moderate risk factors for non-Hodgkin lymphoma (NHL) development (Havranek et al., 2015; Rudd et al., 2006). Previous studies also suggest germline *CHEK2* variants are a risk factor for a variety of myeloid cancers: myeloproliferative neoplasms (MPNs) (Janiszewska et al., 2012), myelodysplastic syndromes (MDS) (Janiszewska et al., 2018), and therapy-related myeloid neoplasms (t-MNs) (Churpek et al., 2016; Singhal et al., 2021). One recent study of P/LP germline variants from the BEAT Acute Myeloid Leukemia (AML) cohort identified *CHEK2* variants in 2.05% of newly diagnosed AML patients (Yang et al., 2021). Nevertheless, the relatively high carrier rates for *CHEK2* variant alleles in normal populations have made their classification as risk alleles for hematopoietic malignancies challenging.

Clonal hematopoiesis (CH) is an acquired disorder wherein a subset of hematopoietic cells develop a single-nucleotide variant (SNV), insertion/deletion (indel), structural variant (SV), or copy number alteration (CNA) in the absence of other features of a hematologic malignancy (Saiki et al., 2021; Steensma et al., 2015). These

mutations typically result in a proliferative advantage for this subset of cells (Steensma et al., 2015). CH is associated with risks of myeloid(Jaiswal et al., 2014) and lymphoid (Jaiswal et al., 2014; Saiki et al., 2021). malignancies as well as cardiovascular diseases(Jaiswal et al., 2017). The strongest association with CH development is age, with more frequent and larger clones identified in older patients (Jaiswal et al., 2014). However, environmental exposures (Coombs et al., 2017; Wong et al., 2018) and chronic inflammatory states (Dharan et al., 2021; Savola et al., 2018) have also been implicated.

More recently, there has been an increasing awareness that germline variants contribute to the development of CH (Silver et al., 2021). One large scale study examined SNVs from individuals that correlated with the development of CH, and the gene with the strongest correlation with CH was *CHEK2* (Bick et al., 2020). Moreover, suppression of *CHEK2* by RNA interference promoted proliferation in primitive Lin^- $CD34^+$ cells in long-term culture (Bao et al., 2020). This is consistent with the role of CHEK2 in mediating cell cycle arrest such that loss of function provides a proliferative advantage that could contribute to clonal outgrowths. Furthermore, CHEK2 also mediates the DNA damage response to double stand breaks (DSBs), including through the phosphorylation of BRCA1 to promote repair by homologous recombination (HR). Of the two main pathways used for repair of DSBs, HR is considered to be error free as it utilizes a sister chromatid in S-phase. During the rest of the cell cycle or when HR fails, DSBs are repaired by the more error-prone pathway, non-homologous end joining (NHEJ), which directly ligates DBS ends. Therefore, loss of CHEK2 function could contribute to the development of CH both through increased somatic mutation rates due

to altered DNA repair signaling and increased proliferation as cell cycle arrest is impaired. Furthermore, a genome-wide association study (GWAS) identified variant rs555607708 in *CHEK2*, which encodes c.1100delC , as having an odds ratio (OR) of 4.07 for the development of *JAK2* p.V617F CH (Hinds et al., 2016). Although these findings provide circumstantial evidence for an association between germline *CHEK2* variants and CH, but the precise mechanism(s) by which CHEK2 defects lead to clonal outgrowth remain to be explored

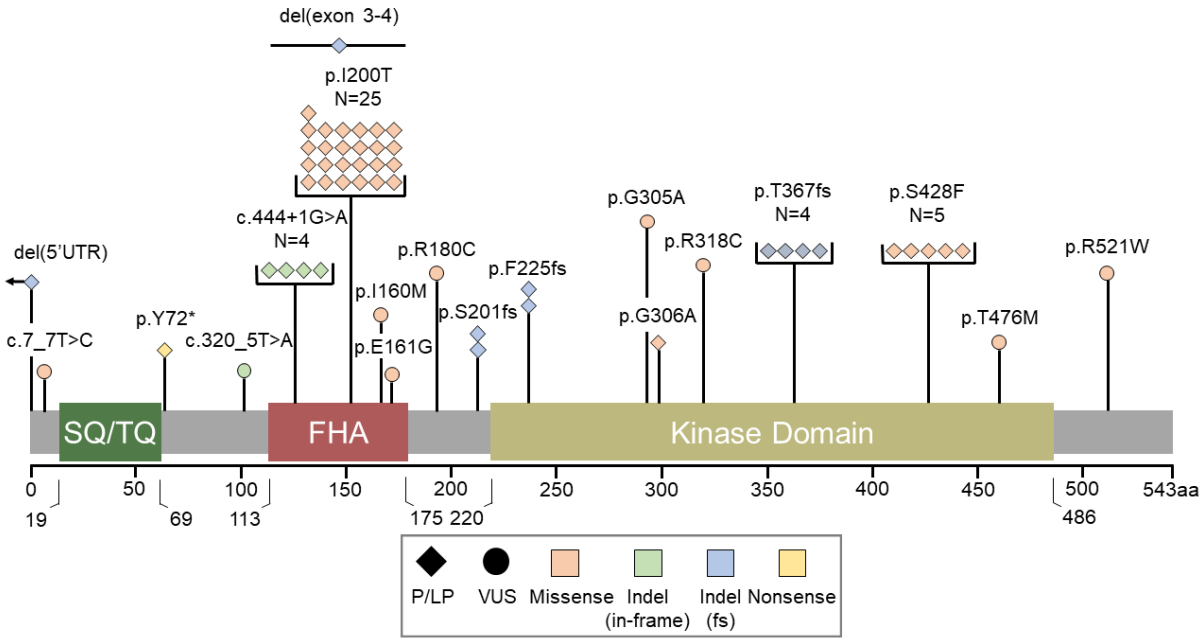
In this study, I worked with Ryan Stubbins and Anase Asom to analyze the Godley Laboratory's cohort of more than 1450 patients and families with clustering of hematopoietic malignancies (HMs) to identify the frequency and spectrum of P/LP germline variants in *CHEK2*. I describe the phenome-wide association study (PheWAS) features of *CHEK2* variant carriers and identify the spectrum of hematopoietic malignancies and other cancers in these patients, with particular focus on the p.I200T allele, representing the most frequently identified variant in the cohort. Finally, I developed a knock-in mouse model carrying the equivalent of the p.I200T variant (p.I161T) for functional testing of the impact on hematopoiesis and leukemogenesis. Excitingly, these mice exhibit mild leukocytosis and develop CH, consistent with the human phenotype, by 6-9 months and a variety of hematopoietic malignancies by 18-24 months, again consistent with our observations from people.

Results

Identification of germline *CHEK2* variants in the University of Chicago clinical cohort

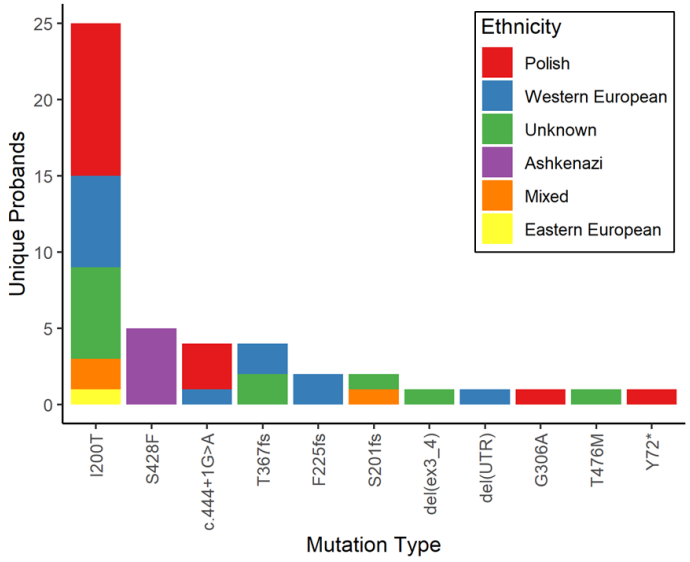
To examine the frequency of deleterious germline *CHEK2* variants in patients with hematopoietic malignancies, Ryan Stubbins and I selected patients from the Godley cohort who underwent clinical or research testing for deleterious germline cancer predisposition variants between 2015-2021, and identified all carriers of *CHEK2* variants. Within the exploratory cohort of 544 eligible patients, 9% (N=47/544) of unique probands carried P/LP *CHEK2* variants and 2% (N=9/544) carried variants of uncertain significance (VUSs). Among unique probands, the most common P/LP were

Figure 4.1 Schematic of CHEK2 and variants identified in patient cohorts



A lollipop plot demonstrates the distribution and location of pathogenic/ likely pathogenic (P/LP, diamonds) variants and variants of uncertain significance (VUS, circles) along the CHEK2 protein from unique probands. Functional domains are indicated as colored and labeled rectangles along protein structure with N-terminus on left and amino acid (AA) position listed underneath (FHA = forkhead-associated domain). Colors of proband shapes represent type of mutation as missense (red), in-frame insertion/deletion (indel, green), frame shift (fs) indel (blue) or nonsense (yellow). Protein change is labeled next to each lollipop variant.

Figure 4.2 Ethnicity of *CHEK2* variant carriers



Number of unique probands (y-axis) for each *CHEK2* variant (x-axis) in patient cohort, with bars colored by ethnicity: Polish (red), Western European (blue), unknown (green), Ashkenazi (purple), mixed (orange), and Easter European (yellow).

c.599T>C/p.I200T (53%, N=25/47), c.1283C>T/p.S428F (11%, N=5/47), c.444+1G>A (9%, N=4/47), and c.1100delC/p.T367fs (9%, N=4/47). The bulk of the P/LP variants affected either the forkhead-associated (FHA) region (62%, N=29/47) or the kinase domain (28%, N=13/47). The majority of P/LP variants in the Godley cohort are missense (68%, N=32/47) followed by indels (30%, N=14/47) and nonsense (2%, N=1/47) mutations (Figure 4.1). The self-reported ethnic distribution for probands was variable, although the most common ethnic origin for the c.599T>C/p.I200T variant was Polish (40%, N=10/25) and for c.1283C>T/p.S428F was Ashkenazi (100%, N=5/5) (Figure 4.2).

A total of 38 patients in our cohort are heterozygous for the *CHEK2* c.599T>C/p.I200T, including family members of unique probands identified by cascade testing. Clinical characteristics of these patients contrasted to other *CHEK2* variant

Table 4.1 Baseline characteristics of *CHEK2* p.I200T variant carriers contrasted to other CHEK2 variant carriers

	Exploratory Cohort (All Patients)			
	CHEK2 p.I200T (N=38)		Other P/LP *CHEK2* variant (N=27)	
	N	%	N	%
Status				
Proband	25	66%	22	81%
Sibling	3	8%	2	7%
Child	4	11%	1	4%
Parent	5	13%	1	4%
Secondary	1	3%	1	4%
Gender				
Male	19	50%	9	33%
Female	19	50%	18	67%
Ethnic Background				
Polish	17	45%	6	22%
Ashkenazi	0	0%	6	22%
Western European	8	21%	9	33%
Mixed	2	5%	0	0%
Eastern European	1	3%	1	4%
Unknown	10	26%	5	19%
Numer of Malignancies				
0	7	18%	3	11%
1	21	55%	16	59%
≥ 2	10	26%	8	30%
Median Age at 1st Diagnosis	49 years		59 years	
(Range)	(18 - 85 years)		(32 - 75 years)	
Malignancy Type	of N=31		of N=24	
Any Hematopoietic	28	90%	19	79%
Only Solid Organ	3	10%	5	21%
Hematopoietic Subtype	of N=29*		of N=20**	
MDS	4	14%	1	5%
MPN	6	21%	3	15%
AML	7	24%	4	20%
B-ALL	2	7%	1	5%
NHL	5	17%	5	25%
Myeloma	4	14%	6	30%
Hodgkin's	1	3%	0	0%

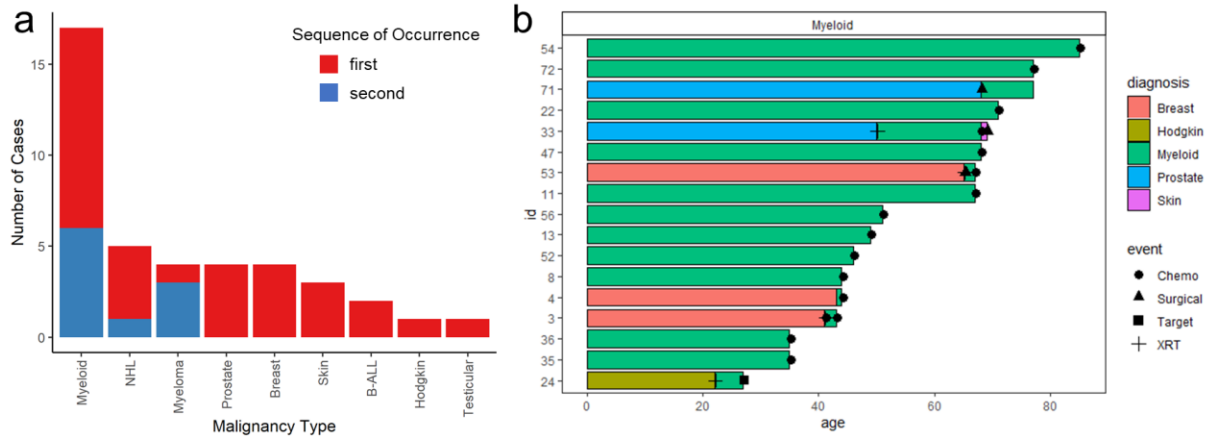
* One case of dual diagnosis of Hodgkin's and MPN

** One case with dual diagnoses of NHL and MDS

carriers are summarized in Table 4.1. Among the *CHEK2* p.I200T variant carriers, 81% (N=21/38) had developed a malignancy at last follow-up, with 26% (N=10/38) having developed ≥ 2 malignancies. This was comparable to carriers of other *CHEK2* P/LP variants with 89% (N=16/27) having developed malignancies and 30% (N=8/27) having developed ≥ 2 malignancies. The median age at first malignancy diagnosis for *CHEK2* p.I200T carriers was 49 years (range 18-85 years) and for other P/LP *CHEK2* variant carriers was 59 years (range 32-75 years, p = 0.22). Of these malignancies, the most common were myeloid (41%, N=17/41), non-Hodgkin lymphoma (NHL) (12%, N=5/41), myeloma (10%, N=4/41), prostate (10%, N=4/41), and breast (10%, N=4/41) (Figure 4.3a). Within the *CHEK2* p.I200T variant carriers, 74% (N=28/38) had developed any hematopoietic malignancy as well as 70% (N=19/27) in the other *CHEK2* P/LP variant carriers. A significant fraction of the hematopoietic malignancies developed after a prior solid organ malignancy, and subsequent myeloid neoplasms comprised 35% of total myeloid neoplasm diagnoses (N=6/17), myeloma 75% (N=3/4), and NHL 20% N=1/5) (Figure 4.3b). Within the myeloid neoplasms, the most common subtypes were AML (41%, N=7/17) and MPN (35%, N=6/17) followed by MDS (24%, N=4/17). (Table 4.1) In patients diagnosed with a myeloid malignancy, 24% (N=4/17) had received prior cytotoxic chemotherapy or radiation therapy for a solid organ malignancy (Figure 4.3b). Importantly, the majority of hematopoietic malignancies occurred first and were independent of any prior exposure to DNA damaging therapies, suggesting that carrying the *CHEK2* p.I200T can confer risk to spontaneous disease development.

We also investigated the spectrum of additional mutations arising in patients that carried the *CHEK2* p.I200T allele and developed a myeloid malignancy for evidence of

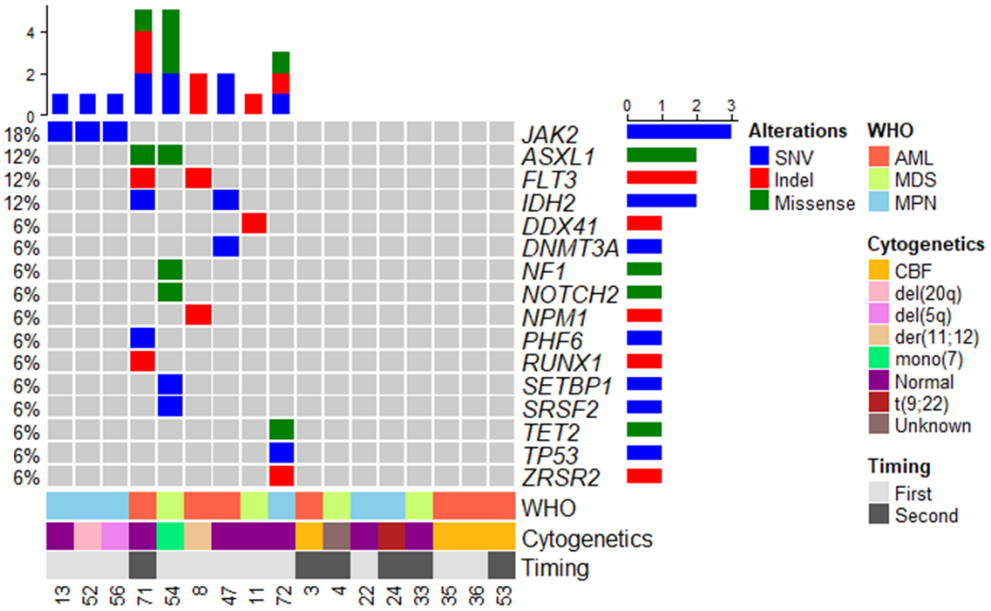
Figure 4.3 Order of occurrence and treatment profiles for malignancies in *CHEK2* p.I200T variant carriers



a. Number of cases (y-axis) for each cancer diagnoses (x-axis) in *CHEK2* p.I200T variant carriers with bar colors indicating if it was patient's first (red) or second (blue) cancer. Cancer types include from left to right: myeloid malignancy, non-Hodgkin lymphoma (NHL), myeloma, prostate, breast, skin, B-cell acute lymphoblastic leukemia (B-ALL), Hodgkin's lymphoma, and testicular. **b.** Order of occurrence of malignancies and treatment profiles in *CHEK2* p.I200T variant carriers. Waterfall plot depicting the age at diagnosis (x-axis) for each patient carrying a *CHEK2* p.I200T variant (y-axis, patient ID number) that experienced a myeloid malignancy. Order of occurrence is depicted moving left to right with bar colors representing cancer type: breast (red), Hodgkin lymphoma (gold), myeloid malignancy (green), prostate (blue), and skin (purple). Shapes at bar ends represent cancer treatment type including: chemotherapy (circle), surgery (triangle), targeted therapy (square), or radiation (XRT, cross).

cooperating mutations or common somatic variants associated with CH. There was an over-representation of core binding factor (CBF) cytogenetic abnormalities, at 57% (N=4/7), which is significantly enriched compared to a control population from The Cancer Genome Atlas (TCGA) (Cancer Genome Atlas Research Network et al., 2013) using a Fisher's exact test. (p=0.003). This finding is consistent with genomic instability leading to chromosomal translocations due to an altered DDR in the context of *CHEK2* p.I200T functionality. Furthermore, the most common molecular abnormality was *JAK2* p.V617F (18%, N=3/17), which is consistent with the same variant being identified in

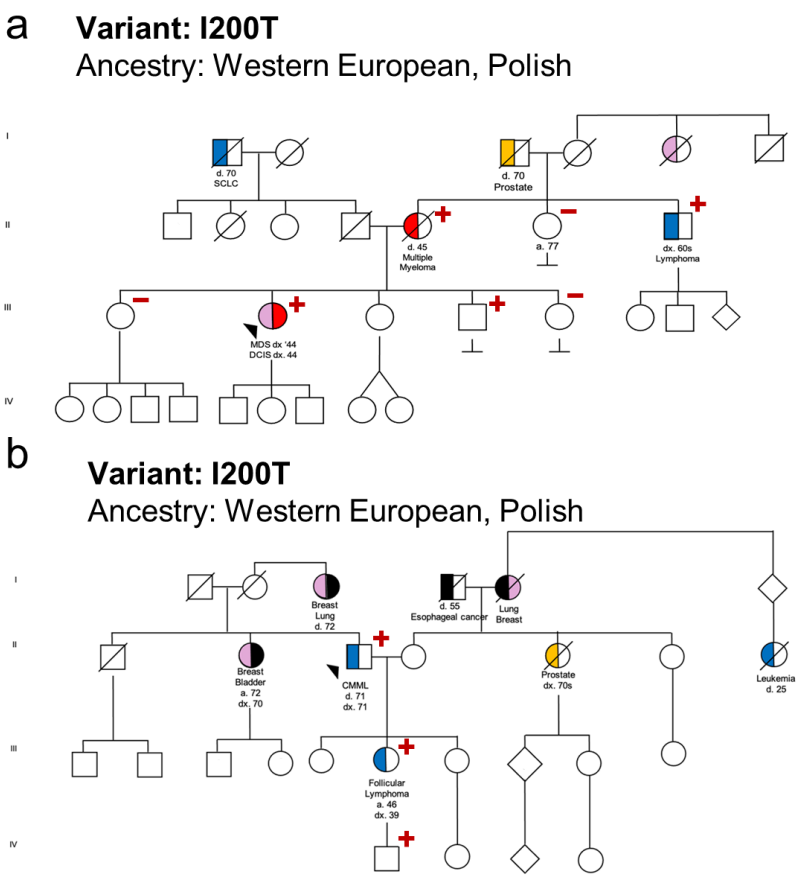
Figure 4.4 Mutation spectrum in myeloid malignancies from *CHEK2* p.I200T variant carriers



Summary of mutations and cytogenetics identified in OncoPlus panel performed for patients carrying the *CHEK2* p.I200T allele and with a myeloid malignancy diagnosis. Each column represents a single patient with patient IDs listed on x-axis. From bottom up, first row indicates whether myeloid malignancy was patient's first (light grey) or second (dark grey) cancer. Second row indicates cytogenetics of patient sample, with colors corresponding to translocations or karyotype type, listed on right of chart key. Third row indicates world health organization (WHO) classification of myeloid disease as acute myeloid leukemia (AML, orange), myelodysplastic syndrome (MDS, green) or a myeloproliferative neoplasm (MPN, blue). Remaining rows each indicate a gene (right, gene IDs), with the proportion of patients in the panel with that mutation listed on the left (% of total). Adjacent bar graph on right is the cumulative number of patients with mutations in that gene. Adjacent bar graph on top is cumulative number of mutations per patient. Colors in adjacent bar graphs and gene-patient grid correspond to mutation type as single nucleotide variants (SNV, blue), insertion/ deletion (indel, red), and missense (green).

previous analysis in the context of the *CHEK2* c.1100delC/p.I200T variant (Hinds et al., 2016) (Figure 4.4). Finally, within affected families, segregation of the *CHEK2* p.I200T variant with cancer development was established in 36% (N=9/25) of cases; representative pedigrees are demonstrated in Figure 4.5.

Figure 4.5 Representative pedigrees and segregation with disease



Representative pedigrees showing segregation of *CHEK2* p.I200T allele (red plus) with malignancy, represented by fill color (yellow = prostate, pink = breast, red = myeloid, blue = lymphoid, black = other). Family members that were tested and were WT for *CHEK2* are indicated with a red minus. Sex is indicated by shapes as squares (male), circles (female), or diamonds (unknown). Each row indicates a generation, siblings are connected by top bracketing, and marriage is indicated by horizontal connecting lines. Slashes indicate deceased individuals.

Within the patient subset of P/LP *CHEK2* variant carriers that had clinical testing for suspicion of a heritable cancer syndrome (70%, N=33/47) had a diagnosis of a hematopoietic malignancy. For the *CHEK2* p.I200T variant, the allele frequency was 0.026 (N=14 variants/544 total tests) in the clinical cohort versus 0.00489 (N=691 variants/141,208 total alleles) in the gnomAD control cohort, giving an OR for carrying a *CHEK2* P/LP variant of 5.37 (95% CI 3.14 – 9.18, $p < 0.0001$). For the *CHEK2* p.S428P

variant, the allele frequency was 0.006 (N=3 variants/544 total tests) in the clinical cohort versus 0.00025 (N=19 variants/76,097 total alleles) in the gnomAD control cohort, yielding an OR of 22.20 (95% CI 6.55 – 75.25, $p < 0.0001$). For the *CHEK2* p.T267fs variant, the allele frequency was 0.018 (N=1 variants/544 total tests) in the clinical cohort versus 0.00172 (N=131 variants/76,103 total alleles) in the gnomAD control cohort, giving an OR of 2.14 (95% CI 0.53 – 8.65, $p = 0.2877$). These results are summarized in Table 4.2. Of note, the Godley cohort was enriched based on hereditary cancer signals and hematological malignancies within families, which has the potential to inflate OR esitmates when comaring to the non-cancer gnomAD population. Future studies will require similar analysis in either an all encompassing heredetiray cancer population or an all encompassing hematological cohort for more accurate risk estimates.

Table 4.2 Variant frequencies in hematologic malignancy patients with clinical testing for *CHEK2* variants versus a non-cancer gnomAD control population

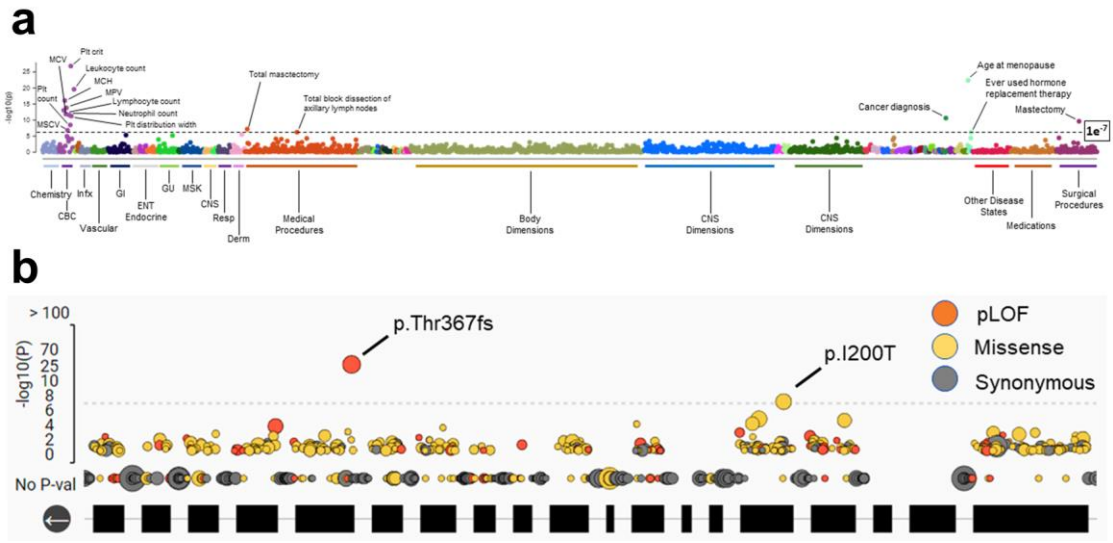
Hematologic Malignancy Patients with *CHEK2* Variant Total (n = 33), excluding research-only testing			Non-cancer ExAc Control Population (gnomAD)		Hematologic Malignancy vs gnomAD cohort	Significance
Variant	Proportion of Individuals with the Mutation	Variant Frequenecy	ExAc Allele Number (excluding homozygous)	Allele Frequency	OR (95% CI)	*p*
p.I200T (c.470T>C)	14 variant 544 total tests	0.026	691 variants 141,208 total alleles	0.00489	5.37 (3.14 to 9.18)	$p < 0.0001$
p.S428P (c.1283C>T)	3 variant 544 total tests	0.006	19 variants 76,097 total alleles	0.00025	22.20 (6.55 to 75.25)	$p < 0.0001$
p.T367fs (c.1100delC)	1 variant 544 total tests	0.002	131 variants 76,103 total alleles	0.00172	2.14 (0.53 to 8.65)	$p = 0.2877$
Total *CHEK2*	33 *CHEK2* 544 total tests	0.061				

Odds ratio calculation for likelihood of carrying a *CHEK2* P/LP variant in patients with a hematological malignancy selected for germline genetic testing based on familial cancer signals compared to a non-cancer control population (gnomAD)

PheWAS for *CHEK2* variants identifies hematological terms

To support the findings that these *CHEK2* variants can increase risk for hematopoietic malignancies, Michael Drazer and Ryan Stubbins analyzed the recently published UK Biobank PheWAS database to search for *CHEK2* variant associations with health phenotype data in 400,000 individuals; I helped review and interpret the results (Karczewski et al., 2021). PheWAS for pLoF variants in *CHEK2* identified 16 terms that were enriched at a $1e^{-7}$ significance level. Of these, 63% (N = 10/16) were related to complete blood count (CBC) phenotype terms. The remainder included age at menopause ($p = 4.05e^{-23}$), cancer diagnosed by doctor ($p = 2.28e^{-11}$), mastectomy ($p = 2.35e^{-10}$) and total mastectomy ($p = 5.53e^{-8}$), block dissection of axillary lymph nodes (*p*

Figure 4.6 UK Biobank data PheWAS for *CHEK2* variant carriers



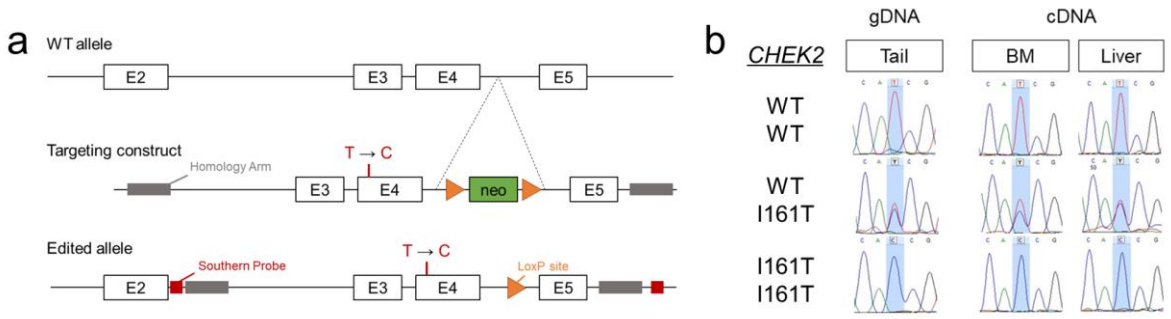
a. PheWAS data from the UK Biobank showing association of predicted loss of function variants in *CHEK2* with health phenotypes (x-axis). Terms above dotted line are significant ($\log_{10}$ p value > $1e^{-7}$) and labeled; complete blood count terms are purple on the left. **b.** Association of individual *CHEK2* variants with 'platelet crit.', the top term hit in above PheWAS. Each circle represents an individual variant, with color representing mutation type as predicted loss of function (pLOF, orange), missense (yellow) or synonymous (grey). Variants plotted by $\log_{10}$ p value (y-axis) association to term platelet crit, with variants above the dotted line being significant ($1e^{-7}$) and labeled.

= 5.74e^{-7}), and ever used hormone replacement therapy (p = 5.88e^{-7}), which are consistent with known associations of *CHEK2* variants to breast cancer (Figure 4.6a). Single variant associations with the top identified phenotype term (Platelet crit) identified significance for p.T367fs (*i.e.,* c.1100delC) (N=1293 variants/532,210 total alleles, allele frequency 2.42e-3, p = 5.15e^{-23}) and p.I200T (N=350 variants/537,488 total alleles, allele frequency 6.51e^{-4}, p = 2.06e^{-7}) (Figure 4.6d). PheWAS specifically for the *CHEK2* c.599T>C /p.I200T variant revealed a similar cluster of complete blood count phenotype terms, with the top being platelet crit (p = 2.06e^{-7}), cytomegaloviral disease (p = 3.81e^{-6}), and monocyte count (p = 2.69e^{-5}).

Establishing the *Chek2* p.I161T mouse model

Having identified the *CHEK2* p.I200T variant and its association to hematopoietic malignancies in patient cohorts, I collaborated with the Transgenics core at University of Chicago to help establish a mouse model to test the hypothesis that this *CHEK2* variant causes hematopoietic phenotypes, such as CH and cancer development. A targeting construct carrying the *Chek2* exon 4 variant that encodes p.I161T, the equivalent of the human p.I200T allele, was designed by Maya Lewinsohn and Matt Jones in collaboration with Xiu Chen in the Transgenics Core (Figure 4.7a). We opted for a simple knock-in animal rather than a conditional targeting of hematopoietic cells to appropriately model human germline carriers. We did not expect any viability issues with this targeted allele because human homozygous carriers of the *CHEK2* p.I200T allele have been reported. I performed PCR screening on genomic DNA (gDNA)

Figure 4.7 Establishing the *Chek2* p.I161T mouse model



a. Schematic of targeting construct used to generate the *Chek2*^{I161T} allele. Top represents the wild type (WT) allele with exons (E) 2-5 depicted as rectangles. Middle is targeting construct with T>C nucleotide change (that encodes p.I200T) labeled in red and homology arms in solid grey. Targeting allele harbors neomycin cassette (neo, green) for selection, which is flanked by loxP sites (orange triangle). Post neomycin excision allele in established animals is depicted on the bottom, with location of Southern blot probes for locus insertion confirmation labeled in red. **b.** Sanger sequencing chromatograms from genomic DNA (gDNA, left) and complementary DNA (cDNA, right) from the indicated tissues, with corresponding *Chek2* p.I161T T>C base change highlighted in blue.

isolated from 198 ES cell clones generated by Xiu Chen, and I selected 2 that were positive for the altered *Chek2* allele to be used for blastocyst injections, which were performed by Linda Degenstein in the Transgenics Core. I PCR screened chimeric pups for the presence of the *Chek2* p.I161T allele to set up new breeders. I established two independent mouse lines, one derived from each ES cell clone, in order to compare phenotypes in equivalently altered mice and reduce risk of off-target effects being interpreted as a phenotype. I performed confirmation of proper locus insertion of targeted *Chek2* p.I161T allele using a combination of Southern blotting and long-range PCR across homology arms, followed by Sanger sequencing. Once established, I continued to confirm the presence of the *Chek2*^{I161T} allele by Sanger sequencing of gDNA from mouse tails (Figure 4.7b). Furthermore, Sanger sequencing of cDNA from bone marrow and liver confirmed expression of the *Chek2*^{I161T} allele and at

approximately even ratios in *Chek2*$^{WT/I161T}$ samples (Figure 4.7b). (Similar results were obtained with cDNA from testes, brain, thymus, and spleen, data not shown). Confirmation of even expression by qRT-PCR and Western blotting is in progress.

After establishing two independent *Chek2* p.I161T mouse lines, Daniel Mendez and I genotyped at least 150 pups from heterozygous parents from each line. I tracked Mendelian ratios and confirmed expected inheritance patterns and no embryonic lethality was observed in either line. Daily observation of both *Chek2*$^{WT/I161T}$ and *Chek2*$^{I161T/I161T}$ mice through 8 weeks showed no developmental defects. At this point, lines were considered equivalent in gross phenotype. Therefore, Daniel Mendez and I established cohorts of *Chek2*$^{WT/WT}$, *Chek2*$^{WT/I161T}$, and *Chek2*$^{I161T/I161T}$ mice from one of the *Chek2* p.I200T lines to investigate CH development and propensity for altered hematopoiesis or malignancies; the other line was maintained at low numbers in case additional observation was necessary and to confirm subsequent phenotypes.

Clonal hematopoiesis in *Chek2*^{I161T} mice

Clonal hematopoiesis (CH) is characterized by the outgrowth of hematopoietic cells carrying somatic mutations and is often considered a precursor to the development of hematopoietic malignancies. Importantly, recent studies identified germline *CHEK2* variants, including I200T, as conferring the highest risk for CH within the general population (Bick et al., 2020). Therefore to detect somatic clones in mice expressing the *Chek2*^{I161T} allele, I used the murine IMPACT next generation sequencing panel developed at Memorial Sloan Kettering Cancer Center, which covers 585 cancer-driving and CH-associated genes (Loberg et al., 2019). I compared gDNA from peripheral

Table 4.3 CH is detected in PB of 6-9 month old *Chek2*^{I161T} mice

Genotype	CH	Mouse ID	Gene	Variant Reads (%)
Chek2$^{WT/I161T}$	50% (2/4)	3440	*Mgam*	1
		3508	*Ercc3*	1
Chek2$^{I161T/I161T}$	80% (4/5)	814	*Rbm10*	11
			Cic	1
		828	*Mef2b*	19
			Amer1	3
		848	*Esr1*	1
		966	*Cux1*	12
			Cux1	8
			Whsc1l1	8
			Rbm10	5
			Prkd1	2
			Smarca4	1
			Mgam	1
			Cd274	1

Somatic clones detected in the indicated genes in heterozygous (green, n=4) and homozygous (n=5, blue) *Chek2*^{I161T} expressing knock-in mice using the murine IMPACT panel, with clone sizes given in the right-most column as a fraction of total reads in that gene. Only mice with at least one clone are shown. No clones were detected in *Chek2*$^{WT/WT}$ mice (n=1).

blood (PB) to matched tail DNA to correct for any variants unique to mice from our colony. Excitingly, I readily identified CH in 6-9 months old *Chek2*$^{I161T/I161T}$ mice (4/5, 80%) as well as *Chek2*$^{WT/I161T}$ mice (2/4, 50%), with larger clone sizes seen in homozygous compared to heterozygous mice (Table 4.3). To date, I have not identified CH in WT mice (n=1), however data analysis for a larger cohort of animals, including and later timepoints, is ongoing.

There are several outstanding questions about the incidence and kinetics of CH in the *Chek2*^{I161T} mice that are currently being addressed by additional IMPACT panel testing. First, as only one *Chek2*$^{WT/WT}$ mouse has currently been analyzed, I have collected material from an additional 4 WT mice in the 6-9 month range for comparison to those currently represented in Table 4.3. I have also established a second set of older *Chek2*$^{WT/WT}$ (n= 5), *Chek2*$^{WT/I161T}$ (n= 5), and *Chek2*$^{I161T/I161T}$ (n= 5) mice that were sampled at 12-17 months old to test if CH expands with age, as is seen in human

Table 4.4 Samples being analyzed in latest MSK-IMPACT panel

Mouse ID	*Chek2* p.I161T	Age At Assay (M)	Cohort
3580	HET	17.2	Old CH
3591	HET	17.1	Old CH
3707	HET	15.6	Old CH
3708	HET	15.6	Old CH
4098	HET	12.4	Old CH
3366	HOM	19.3	Old CH
3575	HOM	17.2	Old CH
3576	HOM	17.2	Old CH
3701	HOM	15.6	Old CH
4087	HOM	12.7	Old CH
3709	WT	15.6	Old CH
3704	WT	15.6	Old CH
4090	WT	12.4	Old CH
4097	WT	12.4	Old CH
4201	WT	12.4	Old CH
4307	WT	9.5	Young CH
4309	WT	9.5	Young CH
4349	WT	8.6	Young CH
4348	WT	8.6	Young CH
966	HOM	23.0	CH+, Endpoint
848	HOM	24.0	CH+, Endpoint
828	HOM	24.0	CH+, Endpoint

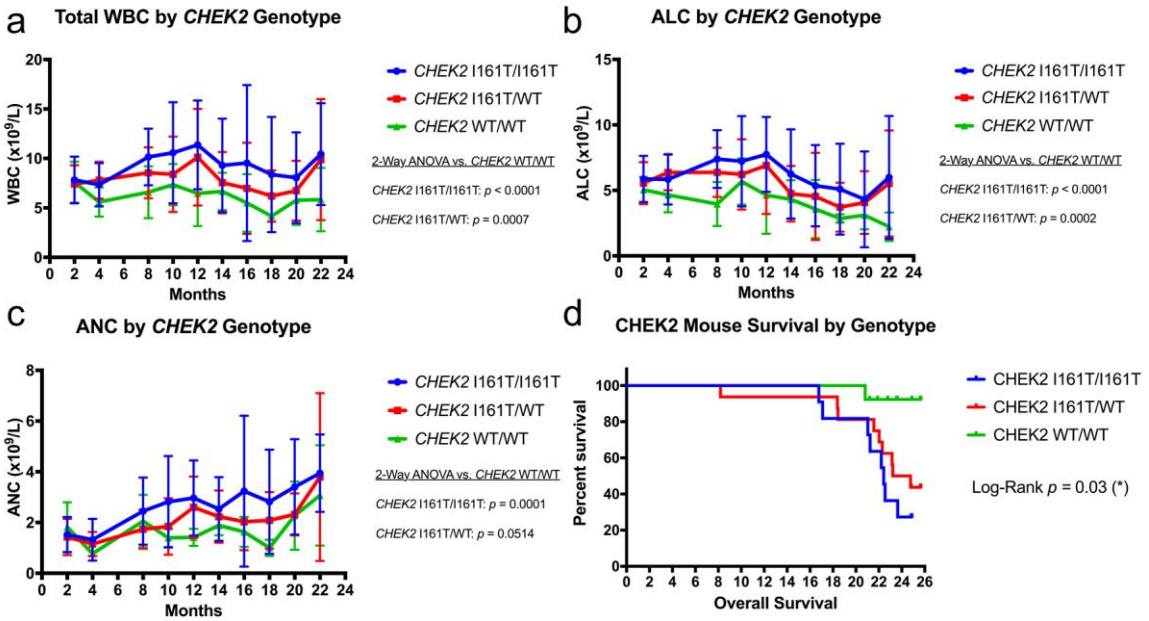
(M = months, DOB = date of birth)

patients. I have also collected PB at endpoint from three animals that were positive for CH in the initial cohort and have submitted these for IMPACT testing to identify persistent or additional clones from the same animals. All the requisite tail and PB gDNA samples have been collected, submitted to MSK-IMPACT, run through panel sequencing, and data are currently in our hands (Table 4.4). Ryan Stubbins and Ashwin Koppayi are working on adapting our variant calling pipeline from human to mouse, and we expect to be able to run analysis to identify clones shortly. These analyses will be critical for defining the onset and kinetics of CH in these mice.

Hematopoiesis and leukemogenesis in the *Chek2*^{I161T} mouse model

I tracked peripheral blood CBCs in a second cohort of *Chek2*$^{WT/WT}$ (n= 10), *Chek2*$^{WT/I161T}$ (n= 15), and *Chek2*$^{I161T/I161T}$ (n= 10) mice throughout lifetime, with assistance from Daniel Mendez and Anase Asom in running CBCs and performing blood smears. Both homozygous and heterozygous mice exhibited mild but significant leukocytosis, with the most pronounced effects on white blood cells (WBC) and neutrophils (Figure 4.8a-c). This lymphocyte expansion fits with the identification of CH

Figure 4.8 Complete blood counts and survival in *Chek2*^{I161T} mutant mice



Lifetime monthly peripheral blood CBCs showing **a.** total white blood cells (WBC, x10^9/L), **b.** absolute lymphocyte counts (ALC, x10^9/L), and **c**. absolute neutrophil counts (ANC, x10^9/L) for *Chek2*$^{WT/WT}$ (green), *Chek2*$^{WT/I161T}$ (red), and *Chek2*$^{I161T/I161T}$ (blue) mice. Age in months is depicted on x-axis. Statistical testing between genotypes performed as 2-way ANOVA over complete dataset. **d.** Kaplan-Meier for overall survival (y-axis) in *Chek2*^{I161T} allele carriers compared to WT controls. Time in months depicted on x-axis.

in 6–9-month-old mice, which could contribute to a proliferative advantage that is detected as an expansion in PB counts.

Furthermore, aged mice harboring a *Chek2*^{I161T} allele also show reduced overall survival as some mice go on to develop HMs (Figure 4.8d). To date, we have detected a range of HMs, including B- and T-cell leukemia/ lymphomas and at least one apparent myeloid malignancy with an expansion of Cd11b$^+$c-Kit$^+$ BM cells (Table 4.5). These HMs have involved a variety of hematopoietic tissues including spleen, lymph nodes, bone marrow, peripheral blood, and liver (Table 4.5). This is consistent with the wide range of HMs identified in patient cohorts. Nevertheless, the immunophenotyping remains incomplete and additional work is necessary to identify the full spectrum of

Table 4.5 Summary of current immunophenotyping of potential HMs in *Chek2*^{I161T} mutant mice

Mouse ID	Sex	Genotype	Survival (months)	Gross	H&E	Flow	IHC Target Plan	Diagnosis
3503	M	HET	23.21	Ocular mass	Circulating abn lymphs	CD19 expansion in PB, CD19+ mass	IHC- Tdt, CD20	B-Cell Leukemia
3426	M	HET	25.10	LN_Masses	TBD	CD19+ mass, no PB CD19+; ? Separate CD3+ clone	IHC - Tdt, CD20	B-Cell Lymphoma
3434	F	HET	23.15	Malocculision	PB Blasts	analysis ongoing		
3506	F	HET	22.03	Weight loss	PB - Normal	analysis ongoing		
3577	F	HET	21.57	High WBC	PB Blasts	analysis ongoing		
3512	F	HET	18.38	LN, derm	No slides?			
3505	F	HOM	22.55	Mesenteric Mass	PB Leukopenia	Mass CD19+, no PB CD19 expansion	IHC - Tdt, CD20	B-Cell Lymphoma
3510	M	HOM	22.22	High WBC	PB - Monocytosis?	CD34+CD11b+CD117+ BM Expansion	IHC - MPO, CD33	Myeloid malignancy
3502	M	HOM	19.44	LN_Masses	TBD	CD3+CD4+ lymphoid	IHC - Tdt, CD20	T-Cell Lymphoma
3501	M	HOM	22.42	High WBC	LN_MSC - lymphoid infiltrate	analysis ongoing		
3438	F	HOM	16.77	High WBC	No slides?	analysis ongoing		
3504	M	HOM	21.04	Missing	Missing			
3437	F	HOM	21.24	Spleen infiltrate?	SP - Infiltrate (?monocytes)			

Summary of current immunophenotyping results of HMs identified in heterozygous (HET) and homozygous (HOM) *Chek2* p.I161T allele carriers including both. Gross refers to observed organ pathology at necropsy. H&E are observations from hematoxylin and eosin stains of peripheral blood smears and spleen or liver touch preparations. Flow column contains results from initial immunophenotyping flow cytometry panel. IHC target plan refers to necessary stains for final diagnosis. Diagnosis contains current designation of HM subtype based on all relevant data. IHC will be used to finalize all HM diagnoses.

HMs that develop in these animals. With Ryan's assistance performing flow cytometry, we have completed an immunophenotyping panel on all mouse tumors where viable material is available. We are currently processing paraffin-fixed tissues to confirm flow findings by immunohistochemistry (IHC) with more specific cell-type markers. The summary of the current results and remaining analysis are indicated in Table 4.5. Conditions for IHC staining for CD3, CD4, CD8, and CD20 are already determined, and Ryan Stubbins is currently working with Terri Li in the Human Tissue Resource Center to optimize staining for additional murine targets CD33, myeloperoxidase (MPO), and Terminal deoxynucleotidyl transferase (TdT). IHC stains will be used to finalize immunophenotyping based on initial flow results, or will be used independently where viable material is not available. When paired with the flow cytometry panels, we will be able to more clearly define spectrum of HMs in these mice. Although it is certainly possible that the *CHEK2* p.I200T variant confers broad risk in HSPCs due to

generalized impact on genome stability, completing the above analysis will be important for defining the model system for future studies.

Discussion

It has remained unclear whether germline variants in *CHEK2* confer risk to HMs in addition to well described pathogenic variants known for association to other solid cancers. Additionally, although previous reports have implicated *CHEK2* variants in a variety of hematopoietic disorders including NHL, MPNs, MDS, t-MNs, and AML (Churpek et al., 2016; Havranek et al., 2015; Janiszewska et al., 2012, 2018; Rudd et al., 2006; Singhal et al., 2021; Yang et al., 2021), the apparent incomplete penetrance of some variants, including p.I200T, has prevented firm conclusions about risk for HM. This is due to the relatively high frequency of this variant in general populations, particularly the Finnish and Polish populations noted earlier, and the number of apparently unaffected carriers without hematopoietic symptoms. In this study, I participated in analysis done by Ryan Stubbins and Anase Asom on the Godley laboratory's patient cohort with clustering of hematopoietic malignancies. This analysis provides evidence that *CHEK2* variants, particularly p.I200T, segregate with disease. Patients with an HM in the Godley cohort were also more likely to carry the *CHEK2* p.I200T and p.S428P variant alleles (OR= 5.37 and 22.2, respectively) in comparison to non-cancer control population, whereas the Ashkenazi associated pT367fs (c.1100delC) is not. Even with the potential for increased carrier rates for some *CHEK2* variants in eastern European populations with the Godley cohort, the association with hematopoiesis is strongly supported by UK Biobank data on 400,000

individuals, which would include non-penetrant *CHEK2* variant carriers that do not contribute to altered CBC terms. Furthermore, I functionally validated a role for *CHEK2* p.I200T in hematopoietic phenotypes by generating a mouse model of this allele (p.I161T) that shows both mild leukocytosis and an overrepresentation of hematopoietic malignancies in aged animals. Taken together, this is strong evidence that the *CHEK2* p.I200T variant has deleterious effects on its function and confers increased risk for hematopoietic malignancies.

Understanding germline predisposition to cancer provides important information for genetic counseling in patients with a strong family history of cancer in addition to being useful for designing treatment and surveillance strategies. Furthermore, due to the incomplete penetrance of some of the *CHEK2* variants, it has been suggested that additional cooperating mutations or environmental factors may also play a role. In line with this, there are several features that are of note in our patient cohort. A substantial number of patients developed a HM after treatment for a prior cancer, suggesting that exposure to DNA damaging or cytotoxic agents may contribute to secondary cancers in these patients. Nevertheless, although this fits mechanistically with the role of *CHEK2* in DNA repair, the majority of HMs occurred as a primary cancer, suggesting risk in these patients is independent of treatment exposures. Furthermore, age to onset and subtype of HM was widely distributed in *CHEK2* variant carriers and does not appear to provide any predictive value for HM development. Future studies with larger patient numbers will be required to elucidate the contribution of environmental factors or the potential for cooperating mutations further. Pending mouse immunophenotyping and analysis of a second familial cancer patient cohort from our collaborators, here I provide supporting

evidence for the inclusion of *CHEK2* p.I200T as a pathogenic variant predisposing to HM. Therefore, genetic counseling would provide potential patient value at any age but may be particularly important for patients surviving a primary cancer.

Clonal hematopoiesis is the process by which hematopoietic stem and progenitor cells acquire somatic mutations over time that provide clonal growth advantages, which is often seen as a requisite precursor for the development of a HM. Although some of the driving mutations that confer clonal advantages are well described, a recent study with whole genome sequencing on 97,691 patients identified germline mutations in *CHEK2* as the most significantly associated with development of these CH clones (Bick et al., 2020). *CHEK2* was also identified in another large study as a potential germline risk allele in MPN (Bao et al., 2020), which can be seen as an intermediate state between CH and the development of a HM. These findings suggest that HSPCs with deleterious *CHEK2* variants are unable to detect and/or correct somatic mutations, giving rise to CH. In this study, I support this paradigm by finding evidence of somatic clones in mice carrying the *Chek2* p.I161T allele at 6-9 months of age, well before any HMs developed. This is likely the result of increased somatic mutation rates due to loss of CHEK2 in DNA damage responses. Mechanistically, failure to phosphorylate downstream targets like BRCA1 or the inability to arrest cells in S/G2 phase when a DSB is present could shift the balance away from error free HR-mediated repair toward more error-prone repair pathways. Furthermore, as CHEK2 can also initiate apoptosis to eliminate persistently damaged cells, deleterious effects on CHEK2 function can produce a survival advantage that contributes to the retention of mutated cells. Finally, in addition to acquired mutations that could provide a selective growth advantage, loss

of CHEK2-mediated cell cycle arrest could increase baseline proliferation rates and promote clonal outgrowths.

Nevertheless, it remains unclear whether any of the somatic mutations identified in the CH analysis here are driving clonal expansions or are merely passengers that mark an expanding clone. In the initial cohort, I did not identify mutations in any of the well-defined clonal driver genes established in human CH, such as *DNMT3A*, *TET2* or *ASXL1* (Bick et al., 2020; Jaiswal and Ebert, 2019). As noted, I have already begun experiments to use serial sampling from the same animals to detect expansions of particular clones over time and how early CH begins in these animals. I expect this analysis to select for the more functionally relevant somatic mutations that are driving clonal expansions and are retained throughout time. Additionally, inclusion of more animals and accumulating larger gene lists from the larger set of samples currently being analyzed (Table 4.4) may allow for identification of cellular pathways involved if individual gene mutations are too rare. Similar resampling at end point in mice that were positive for CH and later developed a HM will help identify additional mutations involved in transformation and which clonal mutations persisted or expanded throughout life. These approaches will more clearly define *Chek2*^{I161T} CH and allow for more direct comparison to human patient cases.

Additionally, it is worth noting that many HMs in the *Chek2* p.I161T mice do not develop until very late ages. It will be important to firmly clarify that these HMs are directly related to the altered allele and are not a consequence of normal aging, when even WT animals can develop HMs. My current analysis does not suggest that WT animals in our colony have significant rates of spontaneous HM development in this age

range, but the aforementioned immunophenotyping will also clarify these results. Additionally, it will be ideal if the exposure experiments can identify factors that accelerate CH and HM development in these animals. This may be necessary in order for them to serve as a more useful research tool, as two-year mouse studies are expensive and difficult to perform with proper controls.

Not all CH will go on to progress to a malignancy, and it is currently unclear what factors contribute to transformation, which is an active area of research. Due to the high level of accessibility of PB, it is possible that some CH that is detected in humans is representative of somatic mosaicism that is present in all pre-malignant tissues. Thus, there is an urgent clinical need to identify the risk factors for developing a HM once CH is detected. Excitingly, to my knowledge, this is the first spontaneous mouse model of CH to be described. This has important value for the field, as modeling CH in mice has been difficult without adding confounding conditions, such as inflammation from irradiation in transplant-based approaches. This will allow for future studies to address the impact of inflammation or therapeutics directly on the incidence and kinetics of CH and the conditions for transformation to a HM.

One hypothesis for the incomplete penetrance of *CHEK2* p.I200T allele is that baseline functions are not significantly impaired and additional DNA damage or replication stress is required for progression to a malignancy. This is particularly relevant to patients that develop a secondary HM after treatment for a primary cancer. In the analysis of the clinical cohorts, there are too few cases to determine whether particular treatment modalities contributed to HMs in these patients. Therefore, the *Chek2*^{I161T} mice could be treated with irradiation or chemotherapeutic agents that

induce DNA damage or target DNA replication to see if this exacerbates CH or decreases time to HM development. It will be interesting to test if different classes of therapeutic agents with different mechanisms of action produce different outcomes on CH or HM development. For example, I have designed future experiments to treat *Chek2* p.I161T mice with either a single sub-lethal dose of γ-irradiation (e.g., 450rad), an alkylating agent (ENU or cyclophosphamide), or a platinum agent (Cisplatin). Treated animals will be tracked by CBCs to identify earlier expansions of WBCs or neutrophils, and the incidence and severity of CH and HMs will also be evaluated. These studies will establish whether environmental agents exacerbate the phenotypes observed in *Chek2* p.I200T allele carriers, and whether DNA damaging agents

Another emerging model suggests that systemic inflammation may be required for CH progression to HM. For example, our group in collaboration with the Jabri laboratory has shown that MPNs that develop in the *Tet2*$^{-/-}$ CH mouse model depend on an IL-6-mediated systemic inflammation as a response to translocation of microbes across a defective gut immune barrier (Meisel et al., 2018). Therefore, it will be interesting to see if inducing inflammation (*i.e.,* irradiation, pulses of pI:pC injections, or infection models) can exacerbate CH clones. Conversely, inhibiting inflammation (*i.e.,* IL-6 blockade, treatment with NSAIDs, or crossing to mice lacking specific immune receptors) may inhibit clonal expansions or the progression to HM.

Finally, crossing to other genetically engineered mouse models will allow for direct testing of any cooperating mutations that are identified in CH analysis or in future human patient studies. For example, it would be interesting to cross the *Chek2* p.I200T line with a *Tet2*-deficient mouse to model the acquisition of a driving mutation identified

in human CH in the context of this germline variant. Similar approached could be undertaken if recurrent mutations are identified in the mouse IMPACT panel CH analysis.

Although I have supported a role for *CHEK2* variants contributing to the hematopoietic phenotypes in human patients by finding evidence of clonal outgrowths and HMs in the *Chek2* p.I161T mouse model, the precise molecular functions that are affected by this variant remain to be defined. As noted previously, the *CHEK2* I200T variant lies in the FHA domain, which includes the dimerization interface required for kinase activation. This is particularly interesting as most previously established P variants, defined by germline variants associated with solid tumors, affect the kinase domain, whereas these results suggest that variants in the FHA domain may bias toward the development of hematopoietic malignancies. Work from other laboratories has suggested that variants in the FHA domain can reduce binding to downstream targets such as BRCA1, p53, and CDC25A (Falck et al., 2001; Li et al., 2002; Wu et al., 2001). Future work should test binding and activation of this CHEK2 axis in hematopoietic cells from the *Chek2* p.I161T mouse model using co-immunoprecipitation and phosphorylation-specific Western blotting. Furthermore, the ability of the mutant Chek2 protein to arrest the cell cycle successfully, repair damaged DNA, or initiate apoptosis could also be investigated to provide mechanistic details into the conditions contributing to CH and altered hematopoiesis in *Chek2*^{I161T} mice. I have also established several human lymphoblastic cell lines from the patient cohorts, and similar analysis for the molecular deficiencies of the FHA domain variant could be tested in this

system. This characterization of an FHA domain variant may also help classify future VUSs identified in this region.

Additionally, as CH in humans is thought to arise in HSPCs leading to expansions of stem cell clones, I would expect to detect similar clonal expansions in murine HSPCs carrying a *Chek2*^{I161T} allele. As noted previously, one study using shRNA knock down of *CHEK2* in Lin-CD34+ cells showed expansions in a CFU assay. I have already established a murine HSPC flow panel and am currently testing for expansions of murine HSPCs. Similar functional analysis of stem cell health could be tested in future experiments using colony forming assays or serial bone marrow transplants.

Finally, the significant risk to HMs identified in the smaller number of patients carrying the *CHEK2* p.S428P variant warrants further study. It is likely not realistic to pursue a mouse model of this variant at this time, however some functional characterization could be performed in lymphoblastic cell lines derived from these patients or with CRISPR to introduce this variant into a cell line model. Ideally, additional corroborating clinical evidence from other groups implicating this variant in dysfunction could emerge after publication of this report. Excitingly, our collaborator Afaf Osman at the University of Utah is providing such a dataset with an additional cohort of families acquired on HMs; arrival of those data await final approval from the University of Utah Institutional Review Board. Interestingly, this patient cohort has a different ethnic distribution and is likely to have less Polish and Eastern European families, which should enable more precise risk analysis for individuals carrying a mutated *CHEK2* allele.

CHAPTER V

Loss of *Brca1* in hemopoietic cells leads to replication-mediated genomic instability and large-scale chromosomal aberrations

In this chapter, I describe work that will require additional experiments/replicates, as detailed in the chapter, before it can be submitted for publication. The contributors to the work presented in this chapter include:

Myself: I have designed and performed all of the experiments, analyzed the data, and anticipate writing the manuscript.

Arthur Wolin: maintained mice and performed delayed bleeding CBC tracking.

Anase S. Asom: assisted with Western blots and DNA fiber imaging.

Julian Lutz: performed some DNA damage foci experiments.

Jane E. Churpek: conceived the study and advised experimental designs and analysis

Fotini Gounari: advised in experimental design and analysis.

Steven Kron: supervised Julian Lutz.

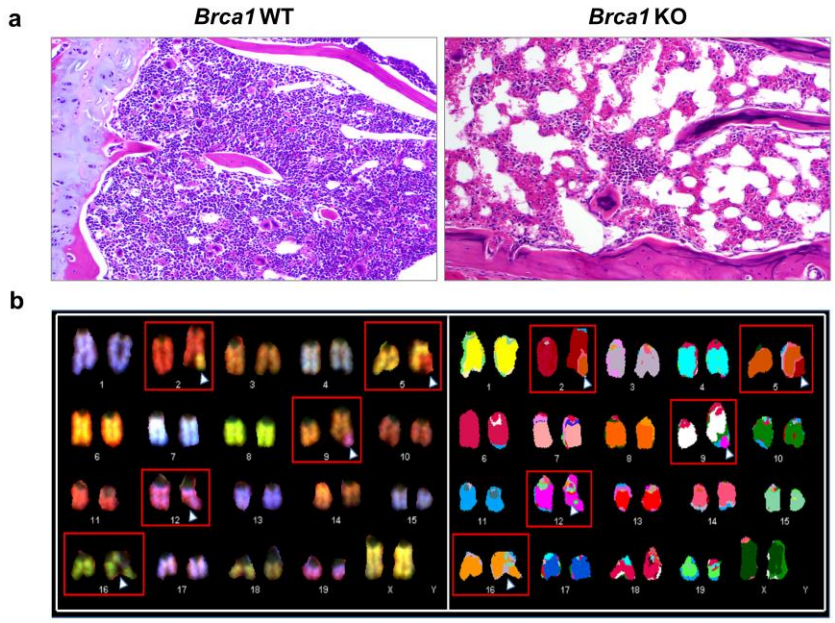
Lucy A. Godley: conceived the study and advised experimental designs.

Introduction

Genomic instability is a well-characterized hallmark of cancer that drives both malignant transformation and cancer cell plasticity. Nevertheless, our understanding of the precise molecular mechanisms involved in these critical cellular processes remains incomplete. BRCA1 is a well-known tumor suppressor that has been causatively linked to genomic instability in a variety of human cancers. Germline mutations in *BRCA1* are associated with greatly elevated lifetime risk for breast (50-70%) and ovarian (40-60%)(Hu et al., 2021; King et al., 2003) cancer as well as lesser risk for stomach, pancreas, prostate and colon cancers (Thompson et al., 2002).

Although *BRCA1* is less well described in the bone marrow (BM) compartment, an emerging body of literature suggests that it plays a key role in restraining the development of hematopoietic malignancies (HMs) and deficiency can result in leukemogenesis (Friedenson, 2007; Vasanthakumar et al., 2015). For example, cancer patients who develop therapy-related myeloid neoplasms (t-MNs) frequently carry germline mutations in a variety of cancer susceptibility genes, including *BRCA1* (Bannon and DiNardo, 2016; Churpek et al., 2016; Feurstein et al., 2016; McNerney et al., 2017). Furthermore, *BRCA1* and *BRCA2* are members of the Fanconi Anemia (FA) pathway that is essential for normal hematopoiesis (Mamrak et al., 2017). Previous work from our group and others have shown that hematopoietic-specific knock out of *Brca1* (KO, *Mx1-Cre*/*Brca1*$^{F22\text{-}24/F22\text{-}24}$) in mice produces FA-like bone marrow failure (BMF) and hematopoietic malignancies with widespread chromosomal aberrations (Mgbemena et al., 2017; Vasanthakumar et al., 2015). (Figure 5.1) Nevertheless, the cellular processes leading to genomic insanity in the absence of *Brca1* in hematopoietic

Figure 5.1 Deletion of *Brca1* in murine bone marrow leads to marrow failure and widespread chromosomal aberrations



a. Representative hematoxylin and eosin staining of fixed sternum sections from a *Brca1* WT (left) and a KO mouse that developed bone marrow failure (BMF). **b.** Representative spectral karyotype analysis from metaphase spreads of two *Brca1* KO mice that developed BMF. Red boxes indicate chromosomes with structural abnormalities. (Vasanthakumar et al., 2016)

cells remain to be determined.

The traditional model of hematopoietic output is a hierarchical organization of dormant and quiescent stem cells that give rise to a more proliferative subset of multipotent progenitors. This is thought to preserve the self-renewal and regenerative capacity of the stem cell pool while progenitor cells contribute the bulk of cells required for the daily output of new blood (Orkin and Zon, 2008). Excessive cycling of hematopoietic stem cells leads to increased acquisition of mutations and eventual exhaustion of the BM (Singh et al., 2020). A recent study showed that merely exiting quiescence and beginning to cycle is sufficient to increase DNA damage in hematopoietic stem cells (Walter et al., 2015). It is currently unclear whether *Brca1* loss

has outsized effects on subpopulations in hematopoietic stem and progenitor cell (HSPC) pools and how this contributes to the observed BMF or malignancies that develop.

Extensive work on the functional roles of BRCA1 has firmly established it as a central player in several genome maintenance pathways, most prominently in homologous recombination (HR) machinery for the DNA damage response (DDR) to double strand breaks (DSBs) (Chen et al., 2018). The type of DNA repair pathway used to resolve DNA DSBs can also contribute to genomic instability and may in part explain the larger number of chromosomal aberrations in some *BRCA1*-deficient cells, including murine BM (Kote-Jarai et al., 2006; Moynahan and Jasin, 2010; Vasanthakumar et al., 2015). Although error-free HR acts in S-phase using a homologous DNA strand as template, NHEJ is considered to be error prone as it directly ligates DSB ends. Genomic instability in *BRCA1*-deficient cells has been attributed to the use of NHEJ during S-phase, which would typically be replaced by HR machinery in normal cells (Chapman et al., 2012). Additionally, a more recently described third DNA repair pathway known as alternative end joining (alt-EJ) can also function in the absence of HR and uses small microhomology to ligate DSB ends, with increased rates of insertions and deletions, making it highly error-prone (Stok et al., 2021) (Figure 1.2) It is currently unclear which of these pathways are utilized by *BRCA1*-deficient hematopoietic cells and to what extent they contribute to genomic instability.

In addition to a role in HR, it is increasingly clear that BRCA1 and BRCA2 play key genome maintenance roles at stalled replication forks (Kolinjivadi et al., 2017). This includes HR-independent functions such as preventing nucleases from processing

stalled forks and as a scaffolds for recruiting other fork repair and remodeling proteins (Schlacher et al., 2011; Stok et al., 2021) (Figure 1.6a). Additionally, depending on the type of replication block, the resolution step may require an HR-like function, which is currently under active investigation (Prado, 2018). Importantly, faulty replication can produce one-ended DSB structures either from failed forks that are processed by nucleases or from stalled forks that are reversed into a chicken-foot structure until the block is resolved (Figure 1.6). Failures of BRCA1 function could both contribute to the production of these replication-associated DSBs as well as the inability to properly repair them (Stok et al., 2021).

In either case, aberrant DNA damage or replication stress responses could contribute to the genomic instability in *BRCA1*-deficient BM through a variety of non-mutually exclusive mechanisms. It is important to elucidate which cellular functions are contributing to genomic instability in *BRCA1*-deficient hematopoietic cells in order to define the risk factors contributing to tumorigenesis in these patients and tailor treatment approaches to avoid t-MNs arising in hematopoietic tissues.

Although much of our mechanistic understanding of BRCA1 deficiency derives from complete loss of function studies, heterozygosity is the more clinically relevant paradigm. Patients with germline mutations in *BRCA1* still have one functional copy in most cells, and homozygous germline pathogenic variants in human patients are exceedingly rare (Sawyer et al., 2015). Additionally, given that loss of heterozygosity (LOH) has been considered a rate limiting step in progression to breast and ovarian cancer for *BRCA1*-mutation carriers, cells with a single copy were thought to be largely normal (Fackenthal and Olopade, 2007; Martins et al., 2012). However, there is now

doubt about the requirement and timing of LOH events, and emerging evidence shows that cooperating mutations acquired during *BRCA1*-heterozygosity (*BRCA1*$^{+/-}$) are required for transformation (such as *TP53* truncations) (Martins et al., 2012; Sedic and Kuperwasser, 2016). Recent data also suggests that alternative pathways independent of an LOH event can lead to cancer in patients with a single pathogenic variant in *BRCA1*. Furthermore, *BRCA1*$^{+/-}$ cells can experience an induced haploinsufficiency under replication or damage stressors that is independent of DSB repair activity (Pathania et al., 2014; Sedic et al., 2015). This is particularly concerning given that most chemotherapeutic agents used in cancer patients carrying a *BRCA1* mutation target replication and repair pathways, including the newer PARP1 inhibitors (PARPi).

PARPis were approved first for use in breast and ovarian cancer patients with germline *BRCA1/2* mutations and have become a standard part of care regimens in these patients(Rose et al., 2020). PARPis work via synthetic lethality due to inhibition of DDR pathways that compensate for HR failures and are essential in *BRCA1*-deficient cancer cells, but are tolerated by cells that retain a functional copy of *BRCA1* (Ashworth and Lord, 2018). Even with great treatment outcomes, concerns exist about adverse events affecting hematopoietic tissues This includes a warning to monitor patients for hematopoietic toxicities due to Myelodysplastic Syndrome/Acute Myeloid Leukemia (MDS/AML) occurring in <1.5% of patients exposed to Lynparza (olaparib) monotherapy, with the majority of those events having a fatal outcome. Some of the most frequent major adverse events in PARPi trials include neutropenia, thrombocytopenia, and anemia (Wang and Li, 2021). Additionally, a recent meta-analysis of PARPi-treated patients for a variety of cancer types identified a significantly

increased risk (odds ratio = 2.63, p = 0.026) for developing MDS/AML compared to placebo treated controls arms (Morice et al., 2021). This suggests that patients with germline *BRCA1* mutations could have largely normal BM, but an induced haploinsufficiency and genomic instability when exposed to systemic cytotoxic therapies or PARPis for a primary cancer, potentially explaining the prevalence of t-MNs. As PARP1 is thought to act in back-up DDR pathways and has also been reported at stalled replication forks, understanding how replication and repair are affected in *BRCA1* heterozygous cells is critical for determining risk of PARPi use. This is particularly important as PARPi use is expanded to other cancer types and is increasingly considered for maintenance therapy, as is already the case in ovarian and pancreatic cancer patients post adjuvant therapy (Ledermann et al., 2014).

My project aims to address which genomic integrity functions are affected by *Brca1*-deficiency in BM, particularly in the context of heterozygosity. Elucidating the mechanisms involved will inform better treatment strategies for *BRAC1*-mutation carriers and simultaneously will provide information about BRCA1's role in normal hematopoiesis, which could also be leveraged for new treatment strategies.

Results

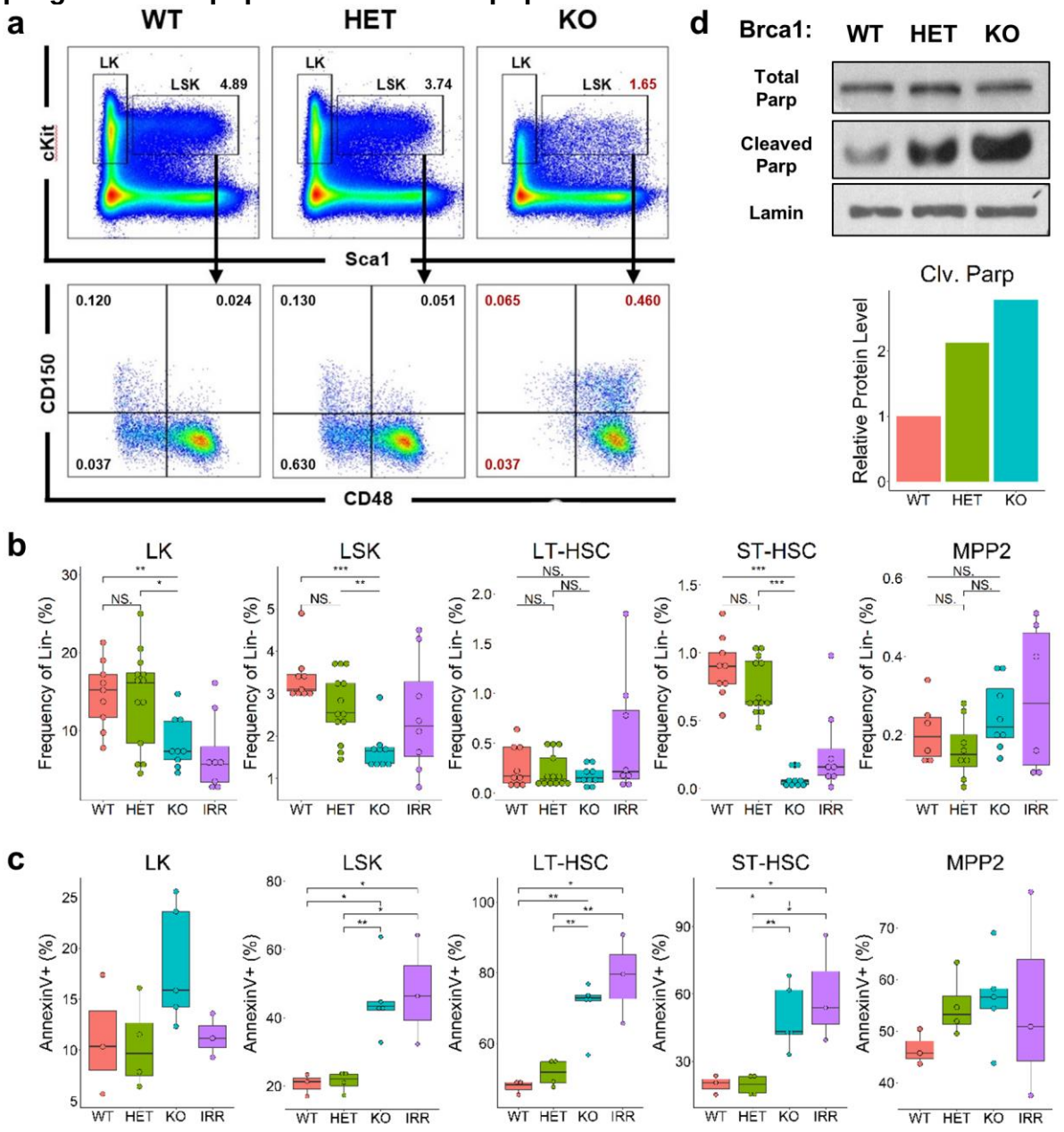
Hematopoietic stem and progenitor cells are lost in bone marrow lacking Brca1

In our previous report, we showed that mice lacking Brca1 in hematopoietic tissues experience altered hematopoiesis characterized by macrocytic anemia and decreased total white blood cell counts followed by BMF or a the development of a HM (Vasanthakumar et al., 2015). Therefore, I first characterized the HSPCs in BM from

Mx1-Cre/ *Brca1*$^{+/+}$ (WT), *Mx1-Cre*/ *Brca1*$^{F22\text{-}24/+}$ (HET) and *Mx1-Cre*/*Brca1*$^{F22\text{-}24/F22\text{-}24}$ (KO) mice using flow cytometry. First, more proliferative progenitor subsets were defined as LK (Lineage^{-}, Sca-1^{-}, cKit^{+}) or LSK (Lineage^{-}, Sca1^{+} cKit^{+}), and LSKs were subset further by SLAM markers (Oguro et al., 2013) to define both short-term hematopoietic stem cells (ST-HSC, LSK CD48^{-}CD150^{-}) and long-term hematopoietic stem cells (LT-HSC, LSK CD48^{-}CD150^{+}) (Figure 5.2a). KO marrow experiences loss of all major stem and progenitor cell compartments, consistent with the observed BMF and altered hematopoiesis in these animals (Figure 5.2b). Additionally, the MPP2 (LSK CD48^{+}CD150^{+}) subset, which is known to expand during emergency hematopoiesis after transplantation or irradiation (Pietras et al., 2015), is mildly expanded in KO animals as well. Interestingly, KO marrow retained some LT-HSCs, which may be a feature of younger 6–8-week-old animals before attrition can fully exhaust stem cell pools. Alternatively, incomplete deletion of the floxed *Brca1* allele could lead to the retention of comparatively healthy *Brca1*$^{+/-}$ HSPCs.

I next addressed whether the lack of HSPCs was due to slowed production or cell death. Flow cytometry for AnnexinV/PI revealed increases in apoptosis in KO marrow in all HSPC subsets, including LT-HSCs (Figure 5.2c), a finding that was confirmed in bulk BM by Western blot. In contrast, I observed increased Parp cleavage in both HET and KO BM (Figure 5.2d). Taken together, KO animals lose nearly all HSPCs due to apoptosis, whereas HET BM is largely unaffected.

Figure 5.2 Brca1-deficient hematopoietic cells exhibit loss of stem and progenitor cell populations due to apoptosis



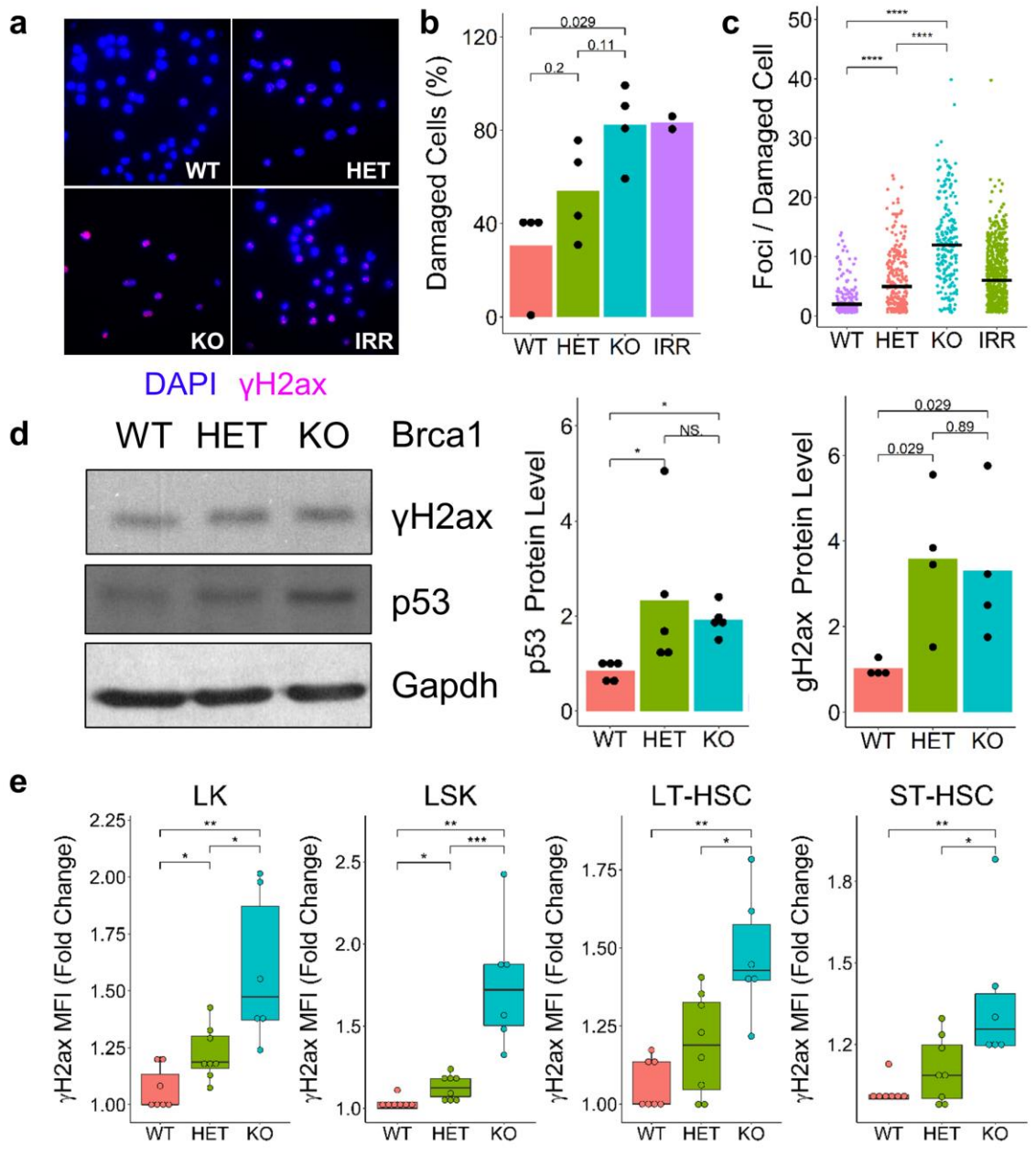
a. Representative flow cytometric pseudocolor dotplots and gating strategy. **b.** Quantification of live lineage negative (Lin^-) hematopoietic stem and progenitor cell (HSPC) populations depicting LK ($Lin^-Sca1^-cKit^+$), LSK ($Lin^-Sca1^+cKit^+$), LT-HSC (LSK $CD48^-CD150^+$), ST-HSC (LSK CD48-$CD150^-$) and MPP2 (LSK $CD48^+CD150^+$) **c.** Cell death depicted as fraction $AnnexinV^+$ from flow cytometry in the indicated HSPC subpopulations **d.** Western blot for apoptosis marker cleaved Parp1 in bulk bone marrow. (WT = $Brca1^{+/+}$, HET = $Brca1^{+/-}$, KO = $Brca1^{-/-}$, IRR = WT+400rads) Boxplots represented as the mean ± s.e.m., and statistical testing is depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.

Brca1-deficient mice have elevated DNA damage at baseline

Given the role of Brca1 in DNA repair, I next investigated the level of DNA damage using immunofluorescent staining for γH2ax, which is a histone that is rapidly phosphorylated upon detection of DSBs in DNA. *Brca1*-deficient BM has more damaged cells (≥2 γH2ax foci/ cell) than WT controls, with KO BM appearing similar to irradiated WT control mice 2 hours after exposure (IRR, 400rad), which induces widespread DSBs. (Figure 5.3a,b) Interestingly, HET BM showed an intermediate level of DNA damage compared to KO and WT animals and the amount of damage per cell correlated with the 'dosage' of *Brca1* (Figure 5.3b-c). Of note, DNA damage appears lower in irradiated (IRR) cells due to merging and the inability to discriminate individual foci. Damaged cells stabilize p53 to arrest the cell cycle and engage either DNA repair programs or apoptosis. Accordingly, both γH2ax and p53 levels are elevated in HET and KO BM by Western blot (Figure 5.3d).

Finally, I investigated whether there is any bias in the HSPC subpopulations experiencing elevated levels of DNA damage using intracellular flow cytometry for γH2ax. Unperturbed KO BM again showed higher levels of DSBs in all HSPCs (Figure 5.3e). Interestingly, I only observed increases in γH2ax in HETs at the LK and LSK progenitor level, whereas stem cells were less affected. These findings suggest that KO BM experiences high levels of DSBs leading to cell death and loss of HSPCs. In contrast, elevated DNA damage in unperturbed HET BM is insufficient to induce widespread apoptosis or lead to HSPC population losses.

Figure 5.3 Unperturbed Brca1-deficient mouse bone marrow has elevated DNA double strand breaks

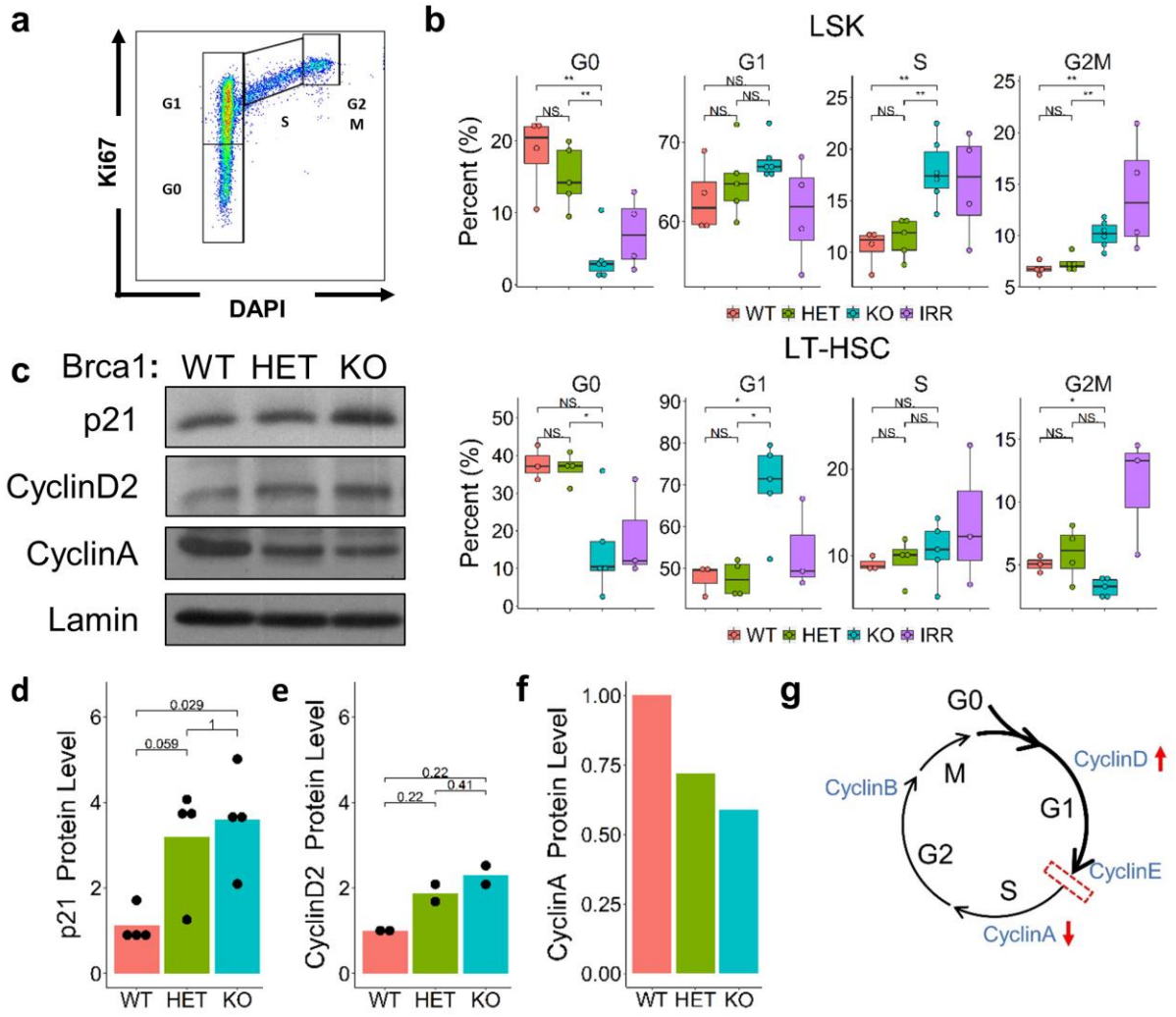


a. Representative images of immunofluorescent staining for DNA damage marker, γH2ax in BM cells **b.** Percentage of damaged cells (≥2 γH2ax foci/ cell) **c.** Total number of γH2ax foci per damaged cell. **d.** Representative Western blot (left) and cumulative quantification (right) for DNA damage markers γH2ax and p53. **d.** Flow cytometry for DNA damage by HSPC subpopulation as fold increase in γH2ax mean fluorescent intensity (MFI) over WT. (WT = Brca1$^{+/+}$, HET = Brca1$^{+/-}$, KO = Brca1$^{-/-}$, IRR = WT+400rads). Statistical testing is depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.

Brca1-deficient bone marrow exhibits an altered cell cycle and replication program

As DNA damage leads to cell cycle arrest to mitigate its effects, determining the cell cycle phase can provide mechanistic information as to the source and outcome of damage stress. Furthermore, DNA-repair pathways are regulated by cell cycle phase, with the Brca1-controlled error-free HR occurring only in S-phase when a sister chromatid is present to serve as a template. Therefore, I determined the cell cycle profile of the HSPCs in WT, HET, and KO BM cells and compared to irradiation-induced DSBs using flow cytometry for DNA content (DAPI) and the proliferation marker Ki67. As expected for WT cells, HSPC subpopulations exhibited relatively quiescent stem cell pools that maintain a greater number of non-cycling cells in G0, whereas LK and LSK progenitors are the more proliferative subsets, supplying the bulk of hematopoietic cells needed for homeostasis (Figure 5.4b) (Páral et al., 2018; Qiu et al., 2014). In contrast, when the BM is stressed or unable to produce sufficient differentiated blood cells, LT-HSCs can exit quiescence and proliferate to increase BM output (Qiu et al., 2014). *Brca1* KO BM exhibits a near complete loss of quiescent LT-HSCs and LSKs (Figure 5.4a-b), which contributes to the increase in cycling LSKs in S-phase (Figure 5.4b), a finding consistent with irradiated WT HSPCs that exit dormancy and begin cycling to meet the acute demands for new blood cell production. Importantly, whereas irradiated WT marrow produces predominantly a G2 block, *Brca1* KO BM is blocked in G1 at both the LT-HSC and LSK level (Figure 5.4b). However, a single copy of *Brca1* in HET BM is sufficient to protect LT-HSCs and LSKs from altered cycling and no significant changes are apparent even with the intermediate damage profiles identified previously.

Figure 5.4 Brca1-deficient mouse bone marrow exhibits altered cell cycle with loss of quiescence and a G1/S-block



a. Representative flow cytometry gating for cell cycle phase by Ki67 and DNA marker DAPI **b.** Quantification of cell cycle phase for HSPC subpopulations LSK (top) and LT-HSC (bottom).**c.** Representative Western blots and **d-f,** cumulative quantification by densitometry for cell cycle markers p21, CyclinD2 and CyclinA in whole bone marrow **g.** Schematic of cell cycle block in *Brca1*-deficient marrow.

Western blots from Brca1-deficient marrow also show elevated p21, which mediates cell cycle arrest downstream of p53 (Figure 5.4c-d). Furthermore, relative enrichment of CyclinD2 and depletion of CyclinA is consistent with a G1 block in *Brca1* KO marrow (Figure 5.4a,e-f). Intermediate but insignificant differences in HET BM in

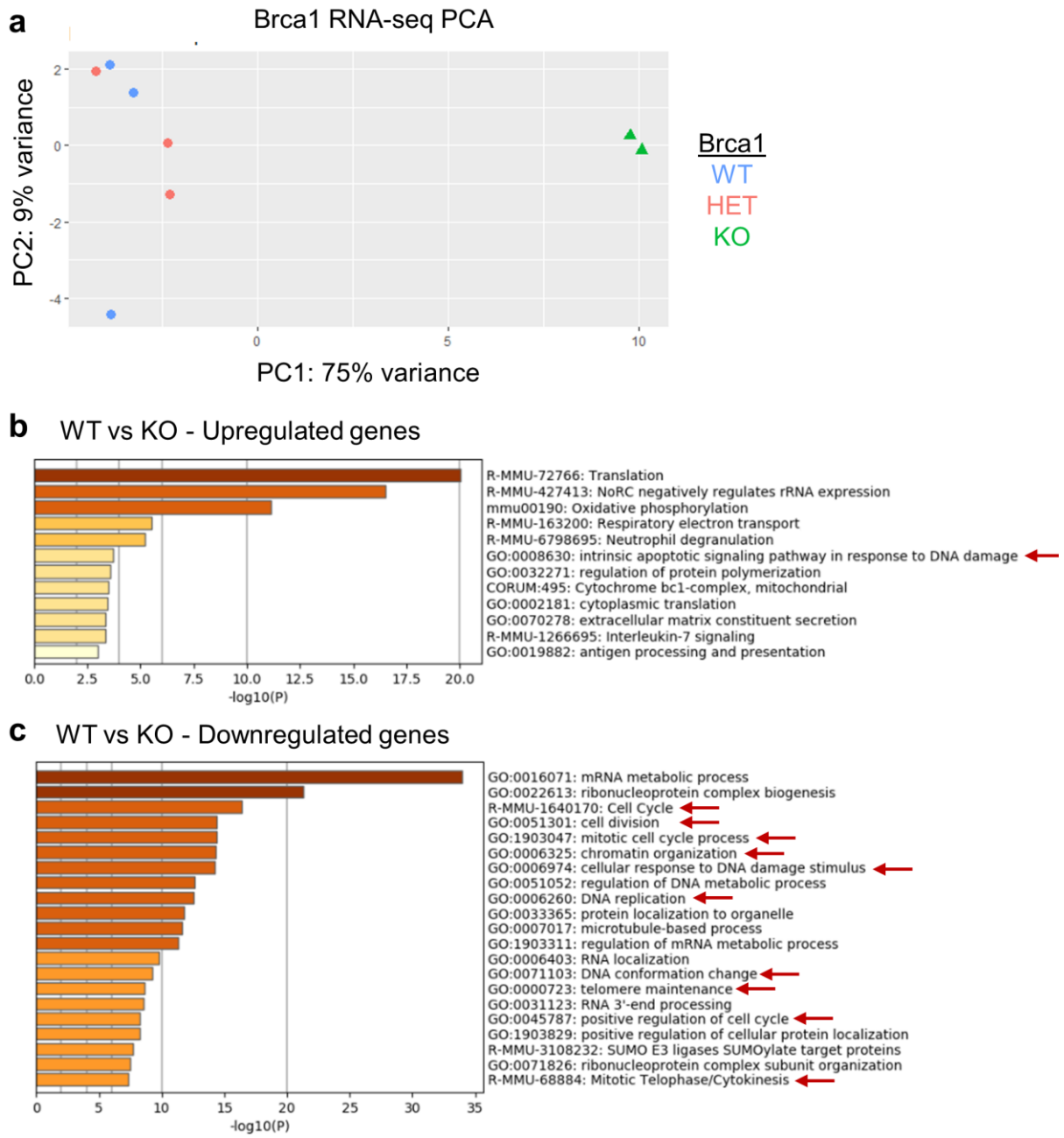
Westerns for the these same cell cycle targets could be due to the use of bulk BM with mixed populations or difference in specificity between Ki67 and cyclin proteins for determining cell cycle stage. Taken together, this evidence suggests that *Brca1* KO BM cells attempt to increase blood cell production by activating quiescent HSPCs, which then experience a G1 block due to elevated damage levels as they begin to proliferate.

To contextualize the failure of genome maintenance programs further and identify sources of the DNA damage levels observed in *Brca1*-deficient BM, I performed RNA sequencing (RNA-seq) on sorted LSKs from WT, HET, and KO mice. Interestingly, principal component analysis (PCA) of normalized counts from all genes shows that WT and HET samples cluster together, whereas KO marrow has a distinct expression program (Figure 5.5a). Moreover, I identified 334 up-regulated and 733 down-regulated genes ($p \leq 0.05$, fold change $\geq \pm 0.5$) in KO compared to WT marrow. Pathway enrichment analysis (metascape.org) of differentially expressed genes highlighted down-regulated genome maintenance programs including cell cycle regulation, DNA replication, DNA repair, and chromosome compaction related to mitosis (Figure 5.5b). Up-regulated pathways included apoptotic signaling in response to DNA damage, corroborating the mechanistic findings from flow cytometry and Western blot assays (Figure 5.5c).

Replication forks are more sensitive to stress in the absence of Brca1

Taken together, these results suggest that *Brca1* KO HSPCs are unable to cope with high levels of DSBs sustained during normal hematopoiesis, leading to apoptosis and cell death. This puts additional proliferative demands on KO HSPCs, which exit

Figure 5.5 Brca1 KO LSKs exhibit loss of cell cycle, replication, and DNA repair expression programs



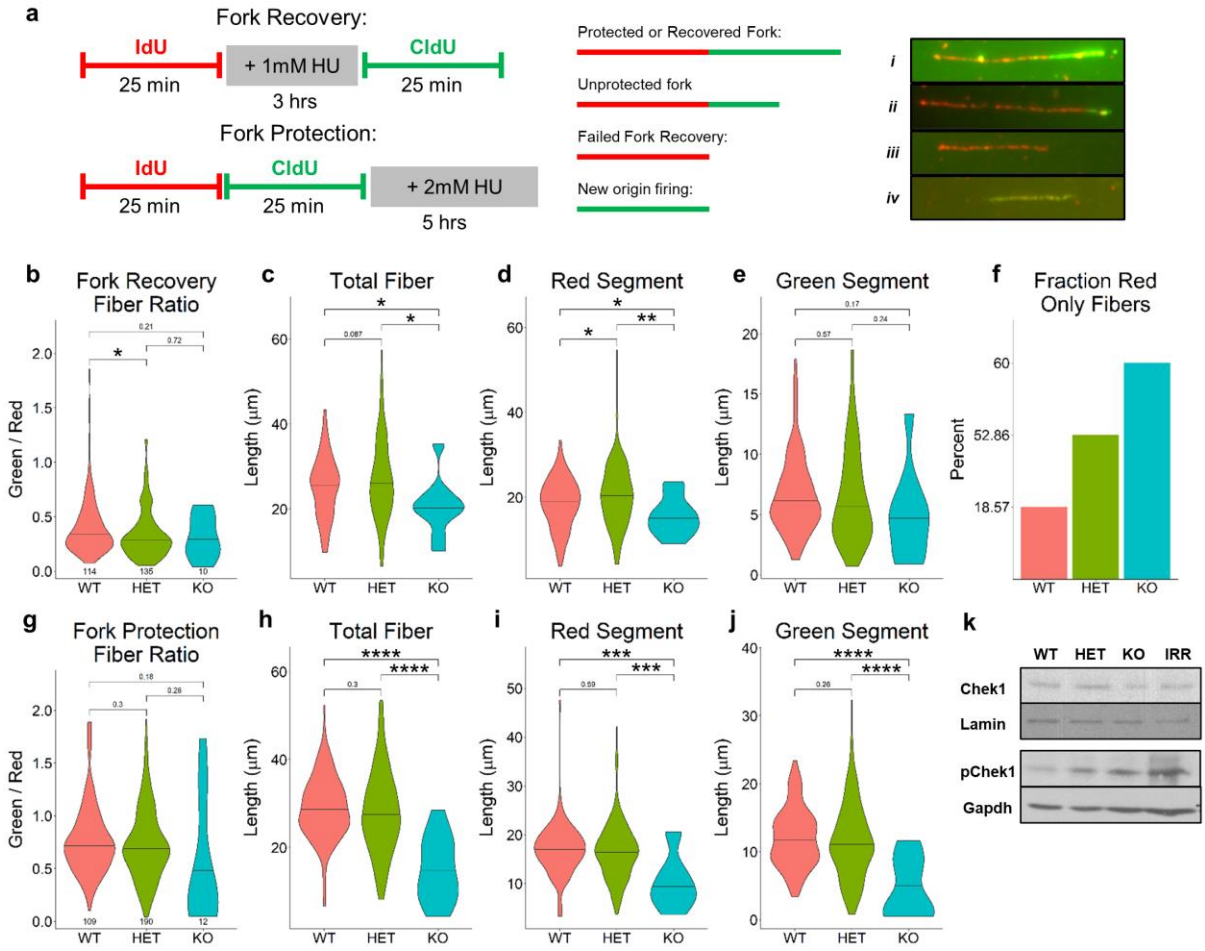
a. Principal component analysis of normalized counts from RNA-seq in Brac1 WT (n=3), HET (n=3), and KO (n=2) sorted LSKs. Pathway enrichment analysis (metascape.org) on differentially expressed genes that are **b,** upregulated or **c,** downregulated in KO compared to WT LSKs. Red arrows indicate genome maintenance pathways.

quiescence and attempt to restore normal blood production. Nevertheless, even newly cycling stem cells likely sustain additional damage as they become blocked in G1. This is consistent with failure early in DNA replication in the transition from G1 to S-phase, rather than a G2 block which is associated with problems later in replication and in preparing chromosomes for mitosis.

After having established that *Brca1*-deficient BM cells have higher levels of DSBs and the associated cell cycle and DDR changes, I next sought to identify the source of this DNA damage. Increased number of DSBs could be explained either by accumulation due to lack of repair or increased incidence related to loss of other Brca1 functions. Although not mutually exclusive, the apparent successful repair of the elevated damage in HET BM cells suggests that the incidence of DSBs was increased.

Due to Brca1's more recently described role in protecting stalled replications forks and the observed increase in proliferation of *Brca1* KO BM, I next investigated the fidelity of the replication fork in Lineage- BM from WT (n=2), HET (n=2), and KO (n=1) mice using variations of the nascent DNA fiber assay. In these assays, cells were cultured for one hour in an HSPC-promoting media and replicating cells were then treated with two 25-minute pulses of nucleotide analogs IdU and CldU. For fork recovery assays, the two nucleotide pulses were separated by a three-hour treatment with the replication stressor hydroxyurea (HU, 1mM). Normal cells will hold and protect stalled replication forks until restart after HU wash-out when they begin to incorporate CldU, producing two-color fibers (Figure 5.6a). For fork protection assays, cells were cultured in 2mM HU for five hours after the second CldU pulse. Normal fork protection will prevent nucleases from processing the nascent strand during the stronger HU

Figure 5.6 Nascent fiber assays suggest Brca1-deficient bone marrow has an impaired replication program with less efficient fork recovery



a. Schematic of nascent fiber assays and the expected outcomes. Representative fibers (*i-iv*) pictured on the right. **c-f.** Fork recovery assay read outs with **b,** 2-color fiber ratios of CldU/IdU segment lengths, **c,** total fiber lengths, **d,** IdU segment lengths (red), **e,** CldU segment lengths (green), and **f,** the proportion of total fibers counted that are IdU-only (red). **g-j.** Fork protection assay read out with **g,** 2-color fiber ratios, **h,** total fiber lengths, **i,** IdU segment lengths (red), **j,** CldU segment lengths (green). Data are depicted as violin plot with cross line representing the mean. Statistical testing is depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.**k.** Western blotting for Chek1 and phosphorylated-Chek1(pChek1, Ser345) in whole cell lysates from BM in the indicated genotypes; Gapdh and Lamin served as the loading controls.

conditions, whereas unprotected forks will experience shortening of the CldU track

(Figure 5.6a). Additionally, the total track lengths (particularly in the first IdU pulse)

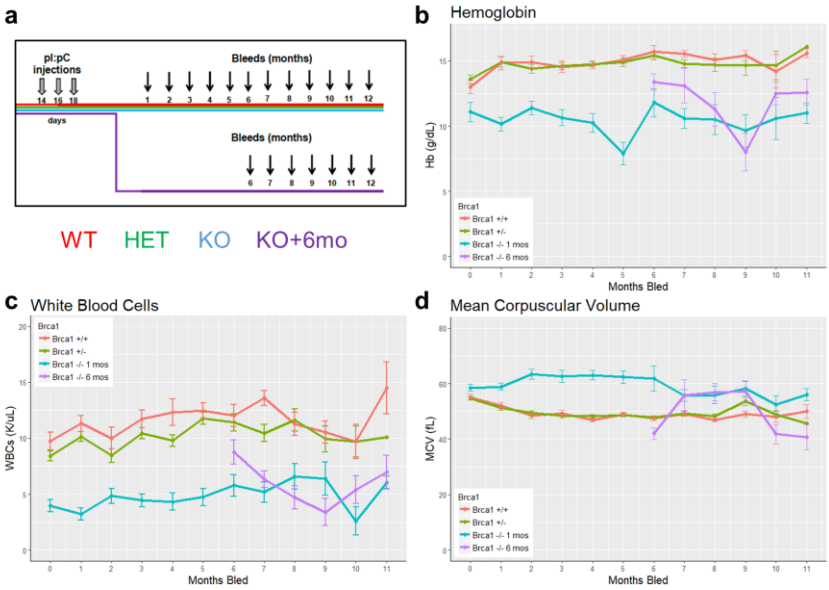
serve as a measure of replication speed and fidelity in both assays. Schematics of the assays and expected outcomes with representative fibers can be found in Figure 5.6a.

In both fiber experiments, KO cells appear to show a reduction in total fiber lengths (Figure 5.6c,h). This shortening was also observed in the IdU portion (red) of two-color fibers, which is before any replication stress is applied (Figure 5.6d,i). Therefore, baseline replication appears to be slower and experiencing kinetic issues in KO cells. Although difficult to interpret due to low numbers and slowed replication in KO conditions, these cells may have reduced fork recovery and protection ratios compared to WT cells. Importantly, there was a strong over-representation of IdU-only fibers from KO cells (Figure 5.6f), which would be expected from fork failures that cannot restart during the CldU culture. This further indicates that KO BM cells may have a faulty replication program, rendering them more sensitive to replication stress. HET BM cells did not show any signs of replication fidelity issues and had normal total fiber and IdU track lengths (Figure 5.6c-d,h-i). Nevertheless, HET BM cells appeared to have reduced fork recovery rates with a significant reduction of two-color fiber ratios (Figure 5.6b), whereas fork protection ratios were similar to WT cells (Figure 5.6g). Additionally, HET BM cells also produced a greater number of IdU-only fibers than WT cells in fork recovery assays (Figure 5.6f). Taken together, this evidence suggests that HET BM cells are normal under steady state conditions, but replication stress can lead to increased fork failure rates. Although these results support the overall hypothesis that loss of Brca1 can lead to faulty replication stress responses that contribute to genomic instability, at least one additional technical replicate and additional biological replicates (at least n=3) are required to make firm conclusions. Nevertheless, Western blotting for

phosphorylated Chek1 (pChek1), which is a marker of replication stress and is activated by Atr signaling, is elevated in both HET and KO total BM cells (Figure 5.6k). Although additional biological replicates (n=2) are required here as well, this is a second modality providing evidence of replication stress, even in HET BM cells.

In addition to decreased fork restart in fiber assays, several lines of evidence suggest problems with DNA replication may be the main source of DSBs leading to genomic instability Brca1-deficient HSPCs. First, recent functional studies in other cell types have shown that Brca1 plays a role in preventing DSB formation at stalled replication forks, both in protecting from nucleases and promoting reorganization and restart of nascent DNA ends (Stok et al., 2021). Secondly, my RNA-seq data strongly implicate DNA replication in addition to DNA repair cellular programs in *Brca1* KO LSKs. Third, my damage profiling of HET BM cells by flow cytometry indicates that the more proliferative LK and LSK subsets had elevated DSBs, but the more quiescent stem cells were unaffected (Figure 5.3e). Fourth, we had observed that the latency to BMF or HMs appeared longer in KO mice after our initial report was published. I hypothesized that the stress of monthly bleeding for characterization of the initial cohort had contributed to the poor health outcomes in the KO mice. Indeed, when KO animals are not bled until six months of age, the changes in peripheral blood counts are altered only mildly, but quickly converge with the counts in the original KO cohort as bleeding continues (Figure 5.7, purple). This suggests that proliferation and cycling can cause DNA damage in Brca1-deficient cells. Finally, the high incidence of chromosomal structural alterations identified in both BMF and HM in KO mice are more consistent with one-ended DSBs generated by replication fork failures. This is in contrast to the small mutations,

Figure 5.7 Bleeding stress contributes to altered hematopoiesis in *Brca1* KO mice.

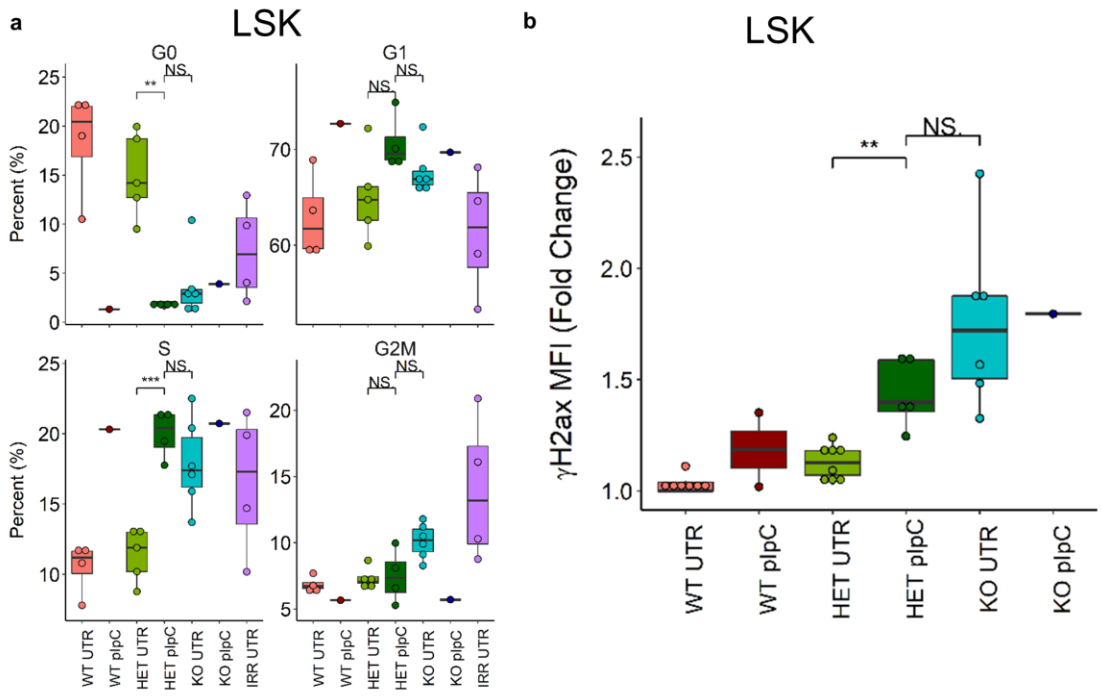


a. Schematic of monthly bleeding of mice of the indicated *Brca1* genotypes, with a second KO arm that is not bled until 6 months of age (purple). Complete blood counts from bleeding cohorts comparing levels of **b,** hemoglobin (Hb, g/dL), **c,** white blood cells (WBC, K/uL), and **d,** mean corpuscular volume (MCV, fL). Time in months on the x-axis.

insertions, and deletions that would be expected simply from alternative repair pathway use on 'normal' DSBs where both ends are proximal for re-ligation (Chapman et al., 2012). These findings suggest that the lack of Brca1 activity at replication forks leads to replication-mediated DSBs that are then repaired with more error prone mechanism due to HR-deficiency.

Further evidence for replication-mediated DSBs in HSPCs comes from a recent report by Walter *et al.* that exit from quiescence and cell cycle entry is sufficient to induce DSBs in normal LT-HSCs. Additionally, forcing HSPCs to cycle using repeated pI:pC injections as a chronic stressor led to BMF and attrition of HSPCs in an otherwise asymptomatic *Fanca*$^{-/-}$ mouse model (Walter et al., 2015). Similarly, I hypothesized that

Figure 5.8 Replication stress in HET marrow induces DNA damage and an altered cell cycle program similar to KOs



Flow cytometric profiling of BM LSKs extracted from mice 48-hours after injection with pI:pC in comparison to untreated controls for **a,** cell cycle phase and **b,** levels of DNA damage determined by fold change of γH2ax (MFI). HET pI:pc comparisons to UTR HET and KO conditions are depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001

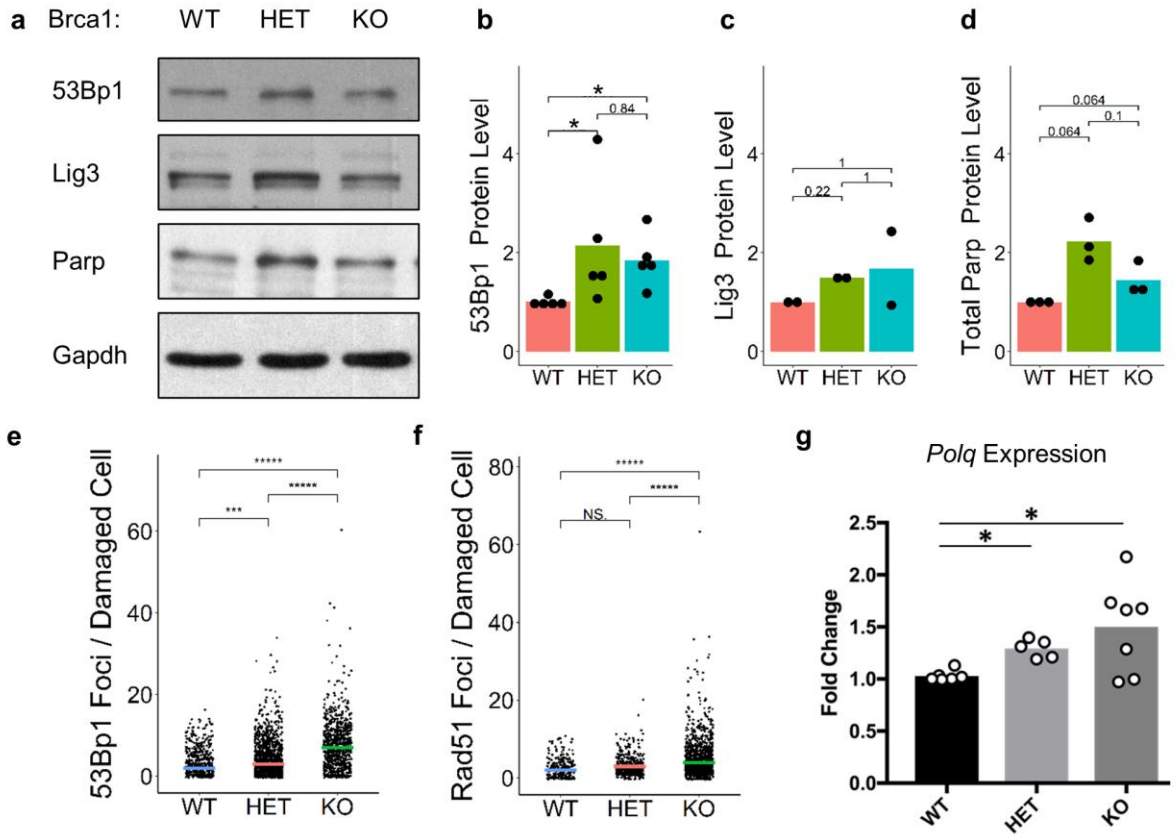
the milder phenotypes observed in the Brca1$^{+/-}$ BM cells under steady state conditions could be exacerbated by replicative stress, leading to higher levels of DSBs than WT cells due to the reduced pool of Brca1. Therefore, I treated two-month old HET mice with intraperitoneal injections 5mg/kg pI:pC and evaluated cell cycle profiles and levels of DSBs in HSPCs by flow cytometry 48 hours later. Both WT and HET LSKs exhibited loss of quiescence and increased cycling cells in S-phase (Figure 5.8a). Importantly, pI:pC treatment produced higher levels of damage in HET BM cells relative to WT mice and exhibited γH2ax staining on par with untreated KO BM (Figure 5.8b). Nevertheless, at least one more technical replicate and increased numbers in control groups (up to

n=5) would be required to finalize these results. Taken together, this experiment suggests that HET mice are largely normal under steady state conditions but experience an induced-haploinsufficiency under replicative stress.

Alternative DNA repair pathways are upregulated in Brca1-deficient mice

Brca1 plays a central role in HR repair of DSBs, which is backed up by more error prone pathways NHEJ and alt-EJ in the absence of functional HR activity. As such, I investigated whether these DNA repair pathways were upregulated in *Brca1* HET or KO BM by Western blot. Protein levels for 53Bp1, a core member of the NHEJ machinery, was significantly upregulated in both HET and KO marrow (Figure 5.9a-b). There was also a mild but insignificant increase in alt-EJ pathway members Ligase 3 (Lig3) and Parp1 (Figure 5.9a-d). Interestingly, HET BM cells consistently showed over-representation of Parp1, which may be evidence of a HET-specific phenotype with some functional HR working in these cells. Nevertheless, lower levels of total Parp1 in KO conditions may be due to increased cleavage in apoptotic cells. I further investigated DNA repair pathway choice by co-staining for NHEJ core factor 53Bp1 in γH2ax immunofluorescent DNA damage repair foci experiments. In both HET and KO BM, there was an increase in the total number of 53Bp1 foci in $\gamma H2ax^{+}$ cells. I also show that in KO, but not HET, BM there is an increase in the number of Rad51 foci, which has been shown to restore HR in the absence of Brca1 (Martin et al., 2007). Finally, although Ligase 3 and Parp1 play roles in other DNA repair pathways or cellular functions, *Polq* (Polθ) is an atypical polymerase that appears unique to alt-EJ function. In the absence of good primary antibodies, I performed qRT-PCR on RNA extracted

Figure 5.9 Brca1-deficient bone marrow upregulates alternative DNA repair pathways, including cNHEJ and alt-EJ.



a. Western blot from BM whole cell lysates probed for the indicated DNA repair pathway members and **b-d** cumulative quantification of protein levels normalized to Gapdh loading control and relative to WT BM (fold change). Total number of DNA damage immunofluorescent foci in γH2ax^{+} cells for **e,** 53Bp1 and **f,** Rad51. **g.** Expression of *Polq* mRNA in Brca1 HET and KO BM relative to WT controls (fold change) determined by qRT-PCR (normalized using 18s rRNA). Data are represented as the mean and statistical testing is depicted as two-sided, unpaired t-tests; *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.

from total BM and showed there was an increase in the transcript levels of this atypical alt-EJ polymerase in both HET and KO mice (Figure 5.9g). This is consistent with *POLQ* levels seen after siRNA mediated knockdown of *BRCA1* or *BRCA2* in 293T cells (Ceccaldi et al., 2015). Taken together, this is strong correlative evidence for increased use of both NHEJ and alt-EJ in Brca1-deficient cells, which would increase error rates and reduce genome integrity. Importantly, HET BM, which does not experience HSPC

loss and has only mild damage phenotypes, already shows upregulation or more error-prone pathways.

Discussion

In this study, I have begun to elucidate the molecular mechanisms that contribute to the widespread chromosomal aberrations and translocations that lead to bone marrow failure (BMF) or HMs in *Brca1*-deficient mice. I first establish that KO HSPCs are lost due to apoptosis at both the progenitor and stem cell level, with a potential delay in LT-HSC attrition due to their less proliferative status. I provide evidence that unperturbed *Brca1* KO BM has high levels of DSBs on par with gamma irradiation, which leads to a G1 cell cycle block due to a DNA damage response (DDR) including stabilized p53 and elevated p21 protein. I established this phenotype using multiple methods and showed similar deficiencies across most HSPC subpopulations. Furthermore, I implicate replication stress response failures due to the dual role of BRCA1 in HR and replication fork protection as a potential explanation for the source of DSBs that are aberrantly repaired by more error-prone pathways. Finally, I show that HET BM cells have mild deficiencies in DNA replication and repair that are likely exacerbated by replication stress.

Genomic instability in *BRCA1*-deficient cells is attributed to the use of more error prone DSB repair pathways in the absence of HR activity. Alt-EJ is a highly error-prone DNA repair pathway that has been increasingly implicated as a back-up when HR is lost, and it is upregulated and essential in *BRCA1*-deficient cells (Ceccaldi et al., 2015). Polq acts on resected DSB ends and relies on small microhomologies, leading to

deletions and insertions at repair sites and potential for translocations (Wood and Doublié, 2016). Depending on the context, alt-EJ promotes translocations in an *IgH-Myc* translocation model (Yousefzadeh et al., 2014) but can also suppress translocations at deprotected telomere ends or at CRISPR/Cas9 induced DSBs (Mateos-gomez et al., 2014). Moreover, Polq also plays a role in replication where it has been shown to associate with replication origins and MCM proteins(Fernandez-Vidal et al., 2014), reduce genomic instability at G-quadruplex sites (Koole et al., 2014), and is recruited to stalled forks by PARP1 (Mateos-gomez et al., 2014). Polq also has higher affinity for ssDNA and fork structures than dsDNA (Ceccaldi et al., 2015). My findings that *Polq*, Lig3, and Parp1 are upregulated in Brca1-deficient HSPCs supports a role for alt-EJ activity in these cells. The differential use of alt-EJ as a back-up pathway relative to NHEJ or HR could explain the different outcomes of BMF and HMs. Interestingly, HET cells also showed elevated damage and a surprisingly high level of Parp1 protein. This could result from the presence of Brca1 in HETs initiating end resection in S-phase DSBs to prevent NHEJ, but under replicative stress, there is insufficient HR activity, shifting the balance toward alt-EJ-mediated repair. It will be important to determine the role for alt-EJ in the high levels of genomic instability in Brca1-deficient HSPCs, and I have begun crossing the *Brca1* mouse model to *Polq*$^{-/-}$ mice to assess the effects on genome stability, hematopoiesis, and rates of BMF.

Germline mutations in *BRCA1* are well known for increased risk to breast and ovarian cancer due to loss of HR and increased genomic instability, but earlier studies suggested that the second copy of *BRCA1* was almost always lost in malignant tissues by a variety of mechanisms. This led to the prevailing idea that LOH for *BRCA1* is

required before genomic instability and malignancies arise, whereas having a single functional copy of *BRCA1* is not considered haplo-insufficient. Nevertheless, there is increasing evidence that mutations acquired during *BRCA1*-heterozygosity are required for cell survival past LOH events and likely contribute to carcinogenesis in these patients with germline mutations (Martins et al., 2012; Norquist et al., 2010). Therefore, my findings of even mild deficiencies in HET BM cells are important, because heterozygosity for *BRCA1* is the more clinically relevant paradigm. Importantly, I show that although HET BM cells do not exhibit increased cell death or HSPC losses, they still have more DSBs compared to WT BM, with some evidence for an active p53/p21-mediated DDR. Additionally, DNA damage foci and Western blotting show elevated expression of alternative DSB repair pathways NHEJ and alt-EJ in both HET and KO BM. These results suggest that although HET BM cells encounter mildly increased levels of DSBs, they cope with this level of DNA damage and are spared the more serious consequences of widespread genomic instability preceding transformation or BMF in KO animals. Nevertheless, the upregulation of NHEJ and alt-EJ factors suggests the use of more error-prone pathways, which may contribute to genomic instability and somatic mutations in *Brca1* heterozygous cells. This could explain cooperating mutations that may be acquired before an LOH event and promote a transition to a malignancy.

It is particularly important to understand whether genomic instability arises in heterozygous *BRCA1*-deficient cells, as this would impact risk for both primary and secondary cancers. For example, if patients with germline mutations in *BRCA1* are treated for a primary cancer with chemotherapeutics that induce DNA damage or

replication stress, this may not be well tolerated in non-cancer cells with only one intact copy of *BRCA1*. Indeed, work from our laboratory and others has shown that patients developing a HM after prior treatment for breast cancer are more likely to carry germline mutations in *BRCA1/2* (Churpek et al., 2016). Therefore, my results suggest that *BRCA1*$^{+/-}$ cells may be normal under steady state conditions but experience an induced haploinsufficiency under damage or replicative stress, as has been shown at replication forks in mammary epithelial cells (Sedic et al., 2015). Therefore, elucidating the molecular mechanisms underlying genomic instability in heterozygous *BRCA1*-deficient cells can help tailor frontline treatments to circumvent risks of secondary cancers, or at least increase surveillance after primary treatments for early detection of adverse events. It is also possible that risk for hematologic malignancies is elevated independently of prior treatment as chronic inflammation and other factors can also produce replicative stress, which is important for guiding optimal therapeutic strategies.

Synthetic lethality is an exciting new treatment modality that is exemplified by the success of PARPis that were first approved for use in patients with *BRCA1/2* mutations that developed breast or ovarian cancer. This approach targets a cancer cell's over-dependence on back-up DNA repair pathways that rely on PARP1, whereas normal cells with intact HR remain unaffected. My findings that HET BM cells experience mild dysfunction that is exacerbated by replication stress warrants caution with this synthetic lethal strategy. It is possible that the "normal", non-cancer cells in patients with germline mutations in *BRCA1* are still impacted by PARPi use. This is also consistent with the newer mechanistic paradigm of PARPi efficacy being a result of PARP1 trapping on

gDNA leading to replication stress in BRCA1-deficient cells (Cong et al., 2021; Helleday, 2011).

This is especially important as PARPis are being used increasingly in maintenance therapy, which could increase risk for t-MNs due to off-target effects on these non-cancer cells. In line with this, some of the most frequent adverse events following PARPi treatment include neutropenia, thrombocytopenia, and anemia (Wang and Li, 2021) and there is now evidence for risk of developing MDS and AML (Morice et al., 2021). This includes an FDA warning to monitor for hematopoietic toxicities, including MDS/AML, when treating patients with olaparib. Therefore, this mechanistic work provides a potential explanation to the off-target effects on hematopoietic cells observed in the clinical use of PARPis.

A more detailed understanding of how replication forks fail and which DDR pathways are activated may uncover new targets for synthetic lethal strategies. For example, my findings implicate alt-EJ activity in the absence of HR, suggesting that targeting Polq may provide similar success to, or even augment, PARPi use in these patients. Indeed, while working on this project, two independent groups recently developed Polq inhibitors and showed that they can target cancer cells successfully with HR-deficiencies and increase the efficacy of PARPi in combination treatments (Zatreanu et al., 2021; Zhou et al., 2021). Furthermore, my work suggests that BRCA1-loss also has profound effects at the replication fork independent of HR-activity and DDR responses, suggesting that replication stressors may also enhance sensitivity to PARPi treatments. Nevertheless, as mentioned for PARPi, caution is warranted as inhibition of Polq could have deleterious effect on “normal” cell in patients who are

heterozygous for HR-repair factors. Future studies could also evaluate the impact of losing other HR factors that have additional functions in replication forks both for causing replication-mediated genomic instability or as potential for new synthetic lethal targeting in drug development.

Additional work is still required to confirm replication fork failures and discern which alternative DSB repair pathway is actually contributing to chromosomal translocations and somatic mutations in *Brca1*-deficient cells. Although the experiments I conducted up to this point provide strong correlative evidence for the use of NHEJ or alt-EJ at DSBs, a true functional and quantitative readout of repair pathway use would enhance these findings. For example, alternative immunofluorescent approaches with higher resolution to provide a true co-localization of repair foci to γH2ax may help resolve which proteins are primarily recruited to DSBs and not just upregulated at the expression level. Another approach would be to use plasmid reporter constructs transfected into murine BM or to assess bait DSBs such as crossing to the I-SceI mouse model or using CRISPR/Cas9 targeting of a particular genomic locus. These types of experiments would allow for direct examination of repair product outcomes and provide important mechanistic detail about error rates depending on repair pathway usage in Brca1-deficient BM.

Future work could also continue to evaluate replication fork failures as the source for DSBs in this model. As mentioned previously, at least one additional technical replicate and two biological replicates will be required to confirm the nascent fiber assay results that I include in this chapter. It may also be necessary to sort a more specific subpopulation of HSPCs for these experiments, as small variations in the proliferative

dynamics can contribute confounding variability to fiber measurements. Finally, future work could address specifically the type of fork failures occurring in stressed HET BM cells. For example, the use of other replication stressors or chemotherapeutic agents in place of HU in nascent DNA assays could provide important mechanistic insights into the risk for genomic instability in patients with germline *BRCA1* mutations.

CHAPTER VI

Discussion and Future Directions

In the previous chapters, I investigated the effects of mutations or dysregulation of homologous recombination (HR) pathway members and the impact on genomic instability in lymphocytes and other hematopoietic cells. My findings from these studies highlight the intricate relationship between the DNA damage response and DNA replication stress responses, and the overlapping activities at double strand breaks (DSBs) and stalled forks. Historically, much of our understanding of HR pathway members derive from studies on recombinational repair of DSBs to maintain genomic integrity. Over the last few decades, it has become increasingly clear that HR proteins have distinct and essential functions in maintaining replication forks under stress (Kolinjivadi et al., 2017; Prado, 2018). My studies add to this paradigm and provide evidence that replication is involved in genomic instability that results from failure of HR pathway members and signaling proteins, such as BRCA1 and CHK2. Below I discuss the implications of this mechanistic work on the etiology of genomic instability in hematopoietic cells and the role of altered DNA repair and replication stress programs in the progression to hematopoietic malignancies (HMs).

Implications for defining risk alleles that predispose to hematopoietic malignancies

Although many germline pathogenic variants have been well described for predisposition to solid malignancies, recent years have accelerated the identification of

hereditary factors in the development of hematopoietic cancers. This has included the identification of germline mutations in genes involved in transcriptional machinery, ribosomopathies, cellular proliferation, telomere biology, and DNA repair pathways (Feurstein et al., 2016; Roloff et al., 2021). However, there are frequently variants that are identified in genes within families that are of uncertain significance in their contribution to particular cancers, which must be tested for segregation or characterized for functional contributions to understand patient risks. Here I provide substantial evidence for two DNA repair factors having deleterious functional effects on hematopoietic cells that can explain the identification of mutations or variants in patients with HMs. In Chapter IV, I show that the *CHEK2* p.I200T and p.S428P alleles are associated with an increased risk for HM, and I provide functional evidence in mouse modeling that the p.I200T allele contributes to clonal hematopoiesis (CH) and HMs. In Chapter V, I provide additional mechanistic evidence showing that loss of *Brca1* can lead to high levels genomic instability through increased replication fork failure and alternative DNA repair pathway usage, even in heterozygote carriers. This further supports *BRCA1* mutations as being important in HM development. In both the *BRCA1* and *CHEK2* studies, altered replication stress responses and DNA repair pathway use is implicated in the etiology of HMs from mutations in these two genes. The work in the above chapters strengthens the mechanistic association of these genes with failures in DNA repair leading to hematopoietic malignancies, but also highlights their functional role in controlling replication stress responses. This suggests that replication fork dysregulation contributes to genomic instability in hematopoietic cells and is part of the etiology of HMs arising in germline carriers of pathogenic variants in these genes.

Therefore, my work viewing new variants in DNA repair genes or genes with unknown biology in this replication context may be a useful paradigm for functionally characterizing risk. Importantly, the entirety of the *ATM-CHEK2-BRC*A DNA repair axis has been increasingly implicated for predisposition to HMs, and my work supports future variant characterization taking into consideration the replication functions of affected genes as well.

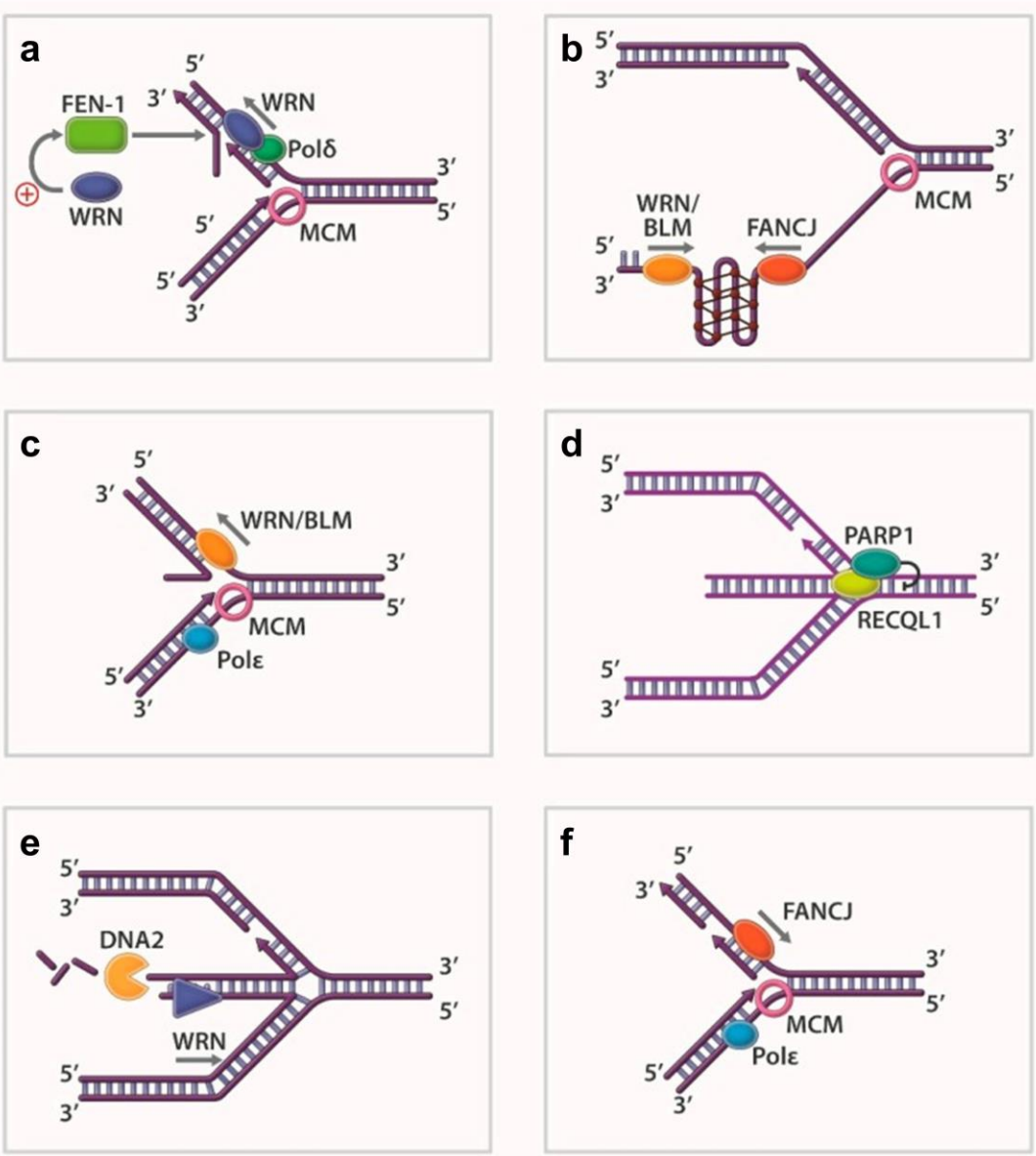
The value of considering replication related functions of these germline mutations that predispose to HMs is also exemplified by recent studies of *DDX41*. The function of this protein was not well characterized when risk alleles were first identified for predisposition to leukemias and lymphomas, although it was implicated in the cGAS/STING immune pathway for sensing of cytosolic DNA (Omura et al., 2016). Nevertheless, this protein contains a DEAD-box helicase domain that suggests it participates in RNA unwinding. Recent reports have now shown that deficiencies in DDX41 can lead to global increases in R-loop structures, which serve as replication impediments during S-phase and lead to STING-mediated activation of inflammatory programs in HSPCs (Mosler et al., 2021; Weinreb et al., 2021). Therefore, replication fork impediments are a significant component of the genomic instability and risk to HM associated with germline variants in this gene.

Interestingly, several other hereditary syndromes associated with DNA helicases also support an overlapping role in DNA repair and replication fork regulation in the etiology of HMs. For example, *BLM*, *RECQL4*, and *FANCJ/BRIP1* are all helicases where pathogenic variants have been identified that predispose to HMs as part of their tumor spectrums(Brosh and Matson, 2020). For example, BLM is best known for its role

in HR repair where it has affinity for double holiday junctions, but it can also unwind complex DNA structures that serve as replication impediments like G-quadruplexes, D-loops, and telomere DNA (Figure 6.1) (Vindigni and Hickson, 2009). BLM can interact with RAD51 to participate in strand invasion and also promotes DNA end resection by EXO1 and DNA2 (Kaur et al., 2021). Furthermore, multiple studies have identified BLM at stalled forks where it interacts with FA proteins to resolve replication blockages (Kaur et al., 2021). Similarly, FANCJ/BRIP1 is best known for its role in the FA pathway for ICL repair and is linked to DSB repair through its association with BRCA1 (Brosh and Cantor, 2014). However, increasing evidence also places FANCJ at stalled replication forks where it participates in resolution of G-quadruplexes and in fork bypass of replication blockages (Figure 6.1b,f) (Brosh and Cantor, 2014; Peng et al., 2018; Schwab et al., 2013). As many of these factors have important roles at stalled replication forks, my findings that Brca1 deficiency has deleterious effects at replication forks further supports a role for some of these other replication stress factors in the development of HMs.

Taken together, there is significant overlap in the functional roles of genes initially identified in HR-repair of DSBs with roles in managing stalled or collapsed replication forks. Therefore, my findings support the idea that germline mutations in HR genes are deleterious in function due to replication failures, which may act synergistically with impairments in DNA repair to degrade genomic stability. Future work should continue to view the function of these genome maintenance proteins in the context of replication stress and consider other replication fork modifying factors for roles in disease etiologies when either germline or somatic mutations are found.

Figure 6.1 Helicase roles at replication forks



Schematic of replication fork structures and the role of DNA helicases. **a**. WRN or BLM (blue) unwinds 5' flaps to promote strand displacement. WRN/BLM also interacts with FEN-1 (green) to stimulate 5' endonuclease activity in Okazaki fragment processing. **b.** WRN/BLM (orange) collaborates with FANCJ (red) to resolve G-quadruplex (G4) structure on lagging stand template. **c.** WRN/BLM (orange) unwind lagging stand duplex to initiate fork regression at stalled forks. **d.** RECQL1 (yellow) reverses stalled replication for into chicken-foot structure, which is antagonized by PARP1 (green). **e.** WRN (blue) promotes DNA2 (orange) processing to facilitate fork restart from revered fork. **f.** FANCJ (red) promotes fork elongation by unknown mechanism at stalled fork. (MCM replication complex represented as pink circle.) Adapted from Brosh et al.

The role of DNA repair factors in predisposition to clonal hematopoiesis

Clonal hematopoiesis is characterized by the outgrowth of bone marrow (BM) clones with some of the well characterized somatic ‘driver’ mutations in specific genes (i.e. DNMT3A, TET2, ASXL1, or PPM1D) and is a risk factor for progression to myelodysplastic syndrome (MDS) and HMs (Bick et al., 2020). Several recent large studies have attempted to identify predisposition alleles, with germline mutations in *CHEK2* emerging as a potential risk factor both for CH and myeloproliferative neoplasms (MPNs) (Bao et al., 2020; Bick et al., 2020). My findings in Chapter IV that *Chek2* p.I161T mice develop CH strongly supports the role of mutations in this DNA repair and cell cycle regulator in contributing to somatic mutations and clonal outgrowth in BM. It will be interesting to see if other germline mutations in DNA repair factors or replication regulators also contribute risk to CH, including the other *CHEK2* variants identifies our patient cohorts. As such, I am currently testing the *Brca1* heterozygote mice for signs of CH. Excitingly, while preparing this report, a preprint of another genome wide association study evaluating more than 200,000 individuals for loci that are associated with CH was published on preprint severs. This study also implicated variants in *CHEK2*, as well as several other DNA repair genes including *ATM* and *PARP1* (Kar et al., 2022). Therefore, my findings fit with these other recent studies identifying DNA repair factors, and specifically *CHEK2* mutations, as being functionally related to the acquisition of CH mutations. Furthermore, the effects on cell cycle signaling may contribute to the clonal outgrowths that are characteristic of CH. Furthermore, as not all CH goes on to develop into MDS or HMs, the mouse model I report here will be a valuable tool to test additional factors that contribute to CH

development and progression, like inflammation or DNA damage, as was discussed in more detail in Chapter IV.

Replication as sources for endogenous DSBs in translocations in hematopoietic cells

Chromosomal translocations are overrepresented in HMs and are common drivers of tumorigenesis in these cancers. These translocations frequently involve the T-cell receptor (*TCR*) or the Immunoglobulin (*IG*) loci due to the activity of RAG recombinases and AID that are engaged in programmed DSBs and somatic mutations as a normal part receptor diversification during development. Nevertheless, translocations require the fusion of two separate DSBs from disparate loci and the mechanisms leading to genomic instability at these partner loci remains incomplete. Here I provide substantial evidence that replication-mediated DSBs are involved in the etiology of translocation in hematopoietic cells. In Chapter III, I found that DSBs in the *Tcr* locus are Rag-mediated whereas *Myc-Pvt1* breaks are Rag-independent. I further linked *Myc-Pvt1* breaks to Tcf1-mediated downregulation of HR factors that increase levels of replication-associated DSBs. I found similar dysfunction at replication forks in *Brca1*-deficient murine BM and implicate replication fork collapse in the widespread chromosomal aberrations that arise in these cells. Taken together, these findings support a significant role for replication mediated DSBs in the genomic instability in hematopoietic cells that experience stalled replication forks. These findings may help provide a mechanistic explanation for other germline or somatic mutations in DNA repair genes that are associated with HMs, particularly if they play additional roles in resolving

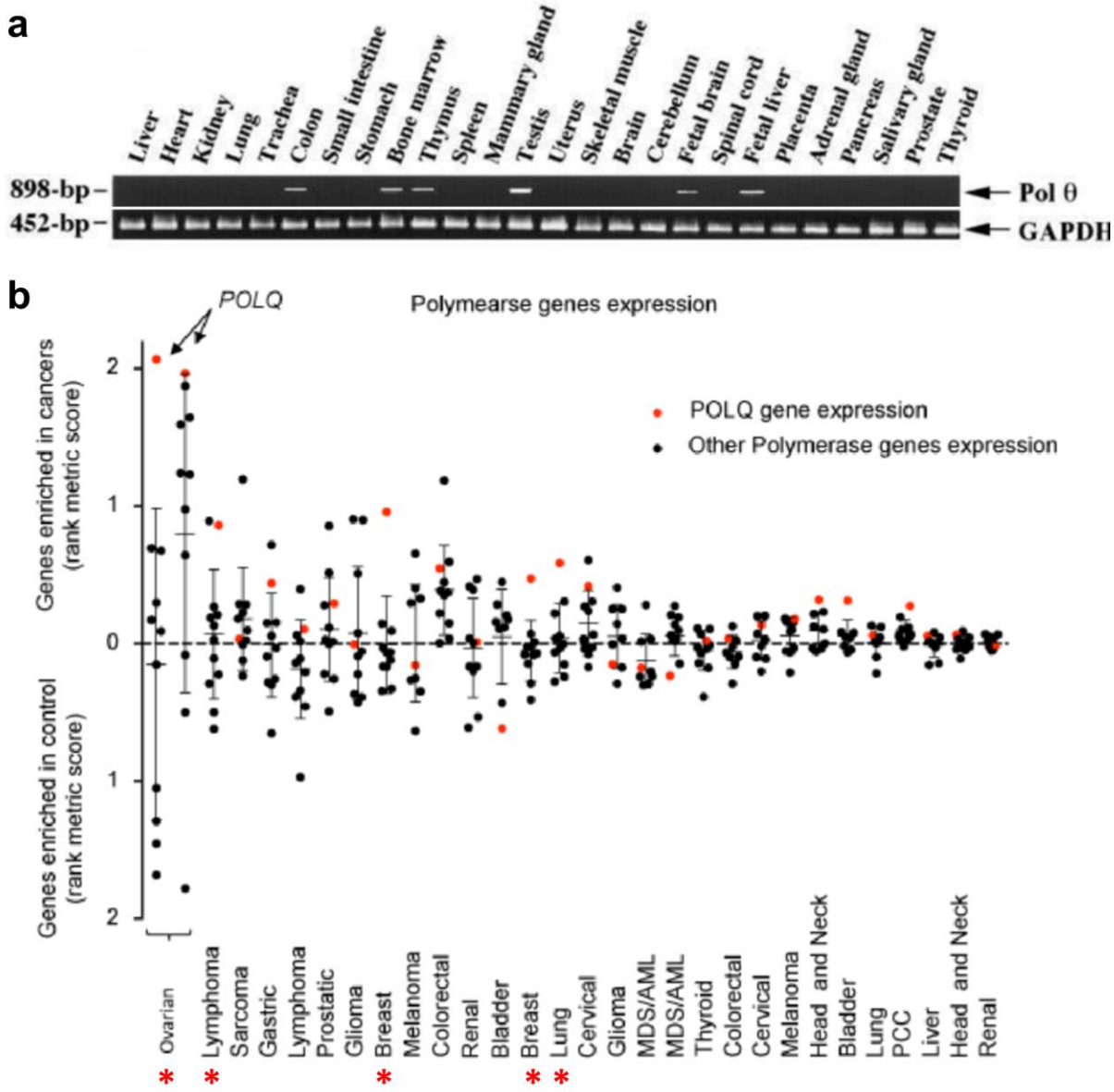
replication stress. Future work should take into consideration the types of damage that can occur at replication forks if HR factors are mutated, and include replication in functional testing to characterize new variants or suspected cellular deficits.

The contribution of alternative DSB repair pathway usage to genomic instability in hematopoietic cells

Another factor involved in genomic instability in HSPCs is which DNA damage response (DDR) and replication stress responses are expressed and activated in cells that lose HR activity. Differential expression rates and use of back up pathways like non-homologous end joining (NHEJ) or alternative end joining (alt-EJ) could account for differences in mutation rates or the types of mutations that develop. For example, my findings in Chapter V show that alt-EJ factors are expressed in *Brca1*-deficient BM and likely contribute to chromosomal aberrations in these cells. Work from other laboratories have shown that *Polq* is expressed more highly at baseline in hematopoietic tissues and could be a cellular response that is poised to act at replication-mediated DSBs should HR-activity fail (Figure 6.2a). Furthermore, *Polq* expression is also more highly elevated in ovarian cancers and lymphoma (Figure 6.2b). Nevertheless, comprehensive analysis of DDR pathway members by tissue type remains to be determined, and it will be especially important for future work to compare relative DDR pathway expression in the absence of particular HR-factors or when stress responses are otherwise impaired.

Differential expression of DDR pathways may already be fundamental in hematopoietic systems, as prior work has shown than quiescent stem cells and proliferating progenitors utilize different DDR pathways for DSBs. As quiescent cells are

Figure 6.2 Expression of *POLQ* in normal human tissues and human cancers



a. RT-PCR for expression of *POLQ* (*POLθ*) with each lane representing one of 26 normal human tissues, with *GAPDH* on the bottom as a loading control. Adapted from Kawamura et al. **b.** Gene set enrichment analysis for expression of polymerases across cancers and control samples in 28 independent data sets from 19 different cancer types (x-axis). Enrichment values (represented as a single dot for each polymerase gene in data set) were determined using the rank metric score to compare expression values between cancers and control samples. Dots above the dashed line reflect enrichment in cancer samples, whereas dots below the dashed line show gene expression enriched in control samples. Dots in red are Polq and red stars indicate cancers where *POLQ* is the most enriched polymerase. Adapted from Ceccaldi et al.

maintained in G0/G1 phase of the cell cycle, they predominately express NHEJ factors which are thought to promote error-prone repair (Mohrin et al., 2010). Nevertheless, hematopoietic stem cells can access error-free repair when they exit dormancy and cycle. Progenitors, which cycle more regularly, are more likely to be expressing HR related factors in S-phase. Therefore, depending on the cell type that is affected by HR failures, the outcomes of error-prone repair could depend on the previously existing expression programs in the cell. For example, hematopoietic stem cells harboring DNA repair pathway mutations may have to balance entering S-phase for error-free repair with vulnerability due to replication associated damage. Given that developing lymphocytes engage in receptor rearrangements later in development, which includes the induction of DSBs, their transcriptional and epigenetic programs could be poised to express different cohorts of DDR factors than other cell types. To what extent lymphocytes or particular hematopoietic stem and progenitor cell (HSPC) subpopulations inherently depend on other DDR pathways in the absence of HR remains to be determined. My findings in Chapter V suggest that alt-EJ may play an important role in DDR responses in HSPCs, and future studies should evaluate stem and progenitor cells for their relative usage of alt-EJ when HR pathways are inhibited.

Furthermore, my findings in Chapter III support the role of epigenetic reprogramming in developing thymocytes altering access to DNA repair and replication stress programs contributing to genomic instability and the biogenesis of translocations. As thymic development requires bursts of proliferation followed by G1 arrest for RAG expression and receptor rearrangements, the transcriptional and epigenetic programs between these distinct cellular states must be tightly controlled. Therefore, I show that

when β-catenin stabilization redistributes Tcf1 from its normal roles, thymocytes lose access to HR factors, leading to failed replication stress responses and aberrant repair programs. This supports the idea that when DNA repair pathways fail or lineage-defining transcription factors are dysregulated, the alternative and back up pathways available to a cell may in part depend on the previously existing epigenetic and transcriptional programs in these developing cells. In line with this idea, it would be interesting for future studies to compare the epigenetic states of and hematopoietic stem and progenitor cells (HSPCs) in *Brca1*-deficient cells for the regulation of alt-EJ factors like *Polq* and *Parp1*.

Implications for tissue specific risk for cancer development

One of the major outstanding questions in germline predisposition for cancer development is why perturbations in general genome maintenance genes only affect a limited number of tissues. For example, mutations in *BRCA1/2* are well known for outsized risk for breast and ovarian cancers, in addition to smaller risk for stomach, pancreas, prostate and colon cancers (Hu et al., 2021; King et al., 2003; Thompson et al., 2002).There are many potential explanations for this behavior, including but not limited to, expression levels by tissue, exposure to hormones, differential metabolism, expression of alternative repair pathways, and the relative proliferation rates of the tissues.

My findings in the previous chapters support replication as an important mediator of genomic instability in HR-deficient hematopoietic cells that is likely playing a role in cancer risk. This may in part explain the additional risk to hematopoietic tissues in

BRCA1/2 mutation carriers and the involvement of other HR repair factors in the etiologies of HMs. Breast and ovarian tissues experience high levels of proliferation in monthly cycling behavior and hematopoietic systems also require massive daily outputs to maintain blood homeostasis, suggesting these tissues may be more vulnerable to replication-mediated instability. In line with this, *BRCA1*$^{+/-}$mammary epithelial cells (MECs) have been shown to be haploinsufficient for replication stress responses whereas keratinocytes are not (Sedic et al., 2015). This study suggested that tissue specific risk to mammary cells in people with a germline *BRCA1* mutation was derived from this dysfunctional replication. In Chapter V, my findings suggest that HSPCs that are heterozygous for *Brca1* are also haploinsufficent at stressed replication forks. Furthermore, replication stress exacerbated the altered hematopoiesis phenotypes of *Brca1* knockout mice. Together, my findings support the idea that heterozygosity for *BRCA1* may have outsized effects on highly proliferative tissues. This may also account for the acquisition of additional somatic mutations under *BRCA1* heterozygous conditions that have been implicated in transformation after an LOH event (Martins et al., 2012). Although proliferation alone cannot explain tissue risk as other proliferative tissues are still less affected, deeper mechanistic understanding of the overlapping roles of HR protein in replication have provided some clues for understanding tissue specificity.

Understanding tissue specific risk will likely have to take multiple cellular pathways into account, and lymphocytes may be more vulnerable to replication stress. My findings support the idea that hematopoietic cells are especially sensitive to

replication-mediated DNA damage due to combined reliance on the replication roles of HR proteins and the relative expression of backup DDR pathways in these cells.

Implications for synthetic lethal approaches

The discovery and successful translation of the first synthetic lethal cancer therapy approach with PARPi has led to increased interest in identifying and exploiting other synthetic lethal interactions in human cancers. This includes both identifying new uses for PARPi in patients without *BRCA1/2* mutations and the search for new targets, often within DDR pathways.

Pathway-specific function of HR proteins have also been critical in understanding the mechanism of action of certain therapies and the potential for developing resistance. For example, the synthetic lethal strategy behind PARPis in *BRCA1*-deficient cells was derived from PARP1's role in other "backup" DDR pathways, like single strand annealing or alt-EJ. Nevertheless, while this may in part be the case, extensive functional work in recent years have revealed that the mechanism of PARPis may be more closely linked to replication blockages. As such, the development of new PARPis have frequently been designed to increase PARP1-trapping on DNA, which can sometimes be linked to PARPi efficacy. This implies that the deleterious effects of *BRCA1* loss are related to its role in replication, and gaps or blockages that impact DNA replication are toxic in these cells. My findings in Chapter V support the notion replication fork failures are a significant component of cellular disfunction in *Brca1*-deficient HSPCs. This contributes to the functional evidence supporting replication blockages as particularly deleterious in *BRCA1*-deficient cells that are likely contributing

to PARP1 efficacy. Although I have not yet directly tested the effect of PARPi on replication forks in *Brca1*-deficient HSPCs, future work could test fork fidelity and collapse under PARPi conditions.

Importantly, my findings in Chapter V may partially explain the worrisome hematopoietic toxicities and risk to MDS/AML in patients being treated with olaparib. As this PARPi is typically used in patients with a germline mutation *in BRCA1/2*, although the cancer cells may experience loss of heterozygosity (LOH), the rest of a patient's cells still have one functional copy. Nevertheless, my work on heterozygous *Brca1* HSPCs showed that some mild dysfunction was present in replication fork protection assays. Therefore, these hematopoietic cells may be more sensitive to replication stress than expected and may experience off-target deleterious effects from PARPi due to replication fork blockages. Additionally, I found that *Brca1*$^{+/-}$ HSPC also show increased expression of some NHEJ and alt-EJ factors, which could contribute to genomic instability in the patient's "normal" hematopoietic cells. Interestingly, PARP1 is also a part of the alt-EJ repair pathway that appears to be expressed more highly in hematopoietic tissues (Figure 6.1a). It is possible that PARPi has deleterious effects in *Brca1*$^{+/-}$ hematopoietic cells due to inhibition of alt-EJ. Future work should evaluate the effects of PARPi on alt-EJ mediated repair in HSPCs to test if this is responsible for leukemogenesis or altered hematopoiesis seen in patients. These findings also suggest that maintenance therapy on PARPi should be performed with extreme caution in heterozygous carriers of *BRCA1/2* mutations or potentially other HR proteins, where minimal curative doses may be required to prevent off target hematopoietic toxicities.

The success of PARPi in patients with *BRCA1/2* mutations has led clinicians to search for other contexts where these drugs will be effective in patients without these mutations. The results I present in Chapter III support the idea that WNT/β-catenin signaling can lead to downregulation of HR-factors in thymocytes. I further show that the lymphomas arising in these mice are indeed PARPi sensitive. Future work should evaluate whether WNT/ β -catenin signaling can lead to downregulation of HR factors in other hematopoietic cells, or in other tissues. Although I provide correlative evidence from drug inhibition databases that leukemia and lymphoma cell lines with WNT signaling are more sensitive to PARPi, the expression of HR factors and sensitivity in β-catenin stabilized leukemia cell lines should be tested directly. I have already designed experiments to treat human PTEN mutant T-ALL xenografts, which have stabilized β-catenin, with PARPi to determine efficacy in human T-ALL samples. It will be interesting to test whether WNT/β-catenin leads to HR-deficiencies in other hematopoietic cells or in other tissues. Furthermore, WNT/β-catenin signaling may serve as a biomarker of PARPi efficacy, or WNT signaling could be targeted to induce sensitivity to PARPi.

For expanding PARPi use, clinicians have also begun to define tumors by their “BRCAness” and to develop HR deficiency (HRD) scores to identify tumors with similar functional characteristics to BRCA-deficient cells and are therefore likely to respond to PARPi (Lord and Ashworth, 2016; Telli et al., 2016). Clinically, HRD is defined by three factors that are associated with HR loss: sum of loss-of-heterozygosity, telomeric allelic imbalance, and large-scale state transition (Telli et al., 2016). HRD scores have successfully been used in the clinic to predict patient responses to therapeutic treatments, however they are still imperfect measures (Coleman et al., 2017, 2019;

González-Martín et al., 2019; Mirza et al., 2016; Ray-Coquard et al., 2019). My findings supporting the functional overlap of HR proteins in replication stress responses encourages the consideration of replication measures and mutations in replication-associate genes for refining HRD scores or predicting PARPi efficacy. For example, PARPi has already been expanded for use in metastatic castration-resistant prostate cancer patients based on mutations in an additional set of DDR genes including *ATM, BARD1, BRIP1, CDK12, CHEK1, CHEK2, FANCL, PALB2*, and *RAD51B/C/D*(de Bono et al., 2020). Many of these DDR proteins interact with BRCA1/2 and may also have roles in replication stress responses, which could contribute to PARPi efficacy. Future work should continue to evaluate the role of replication failure in the efficacy of PARPi to expand their treatment value to additional patients with mutations that affect replication stress responses.

Recent studies have also attempted to replicate the success of PARPi by identifying new synthetic lethal interactions in patients with a DDR gene mutation, usually by targeting other components of the DDR(Ashworth and Lord, 2018). One promising example is the identification of alt-EJ repair pathways as a key mediator of survival in *BRCA*-deficient cells. I had also found in my studies presented in Chapter V that *Polq* and Parp1 were overexpressed in *Brca1*-deficient HSCPs, which suggested these genes may serve as ideal synthetic lethal targets in *Brca1*-deficint tissue. To test this postulate, I have already begun crossing *Brca1*-deficient mice with *Polq* knockout animals. Excitingly, while working on these studies, two independent groups have developed Polq inhibitors that successfully target *BRCA1*-deficient cancer cells and act synergistically with PARPis (Zatreanu et al., 2021; Zhou et al., 2021). Additionally, my

findings that proliferation and replication stress are significant components of genomic instability in cancers with DDR mutations suggest that Polq may play a role in resolving replication fork blockages in the absence of *Brca1*. Indeed, Polq itself contains a helicase domain and may play a role in replication fork stress responses, which could partially explain the success of these recent studies (Wang et al., 2019). My results and accumulating evidence from the field suggests that screening replication fork remodelers and other components of the replication stress response for synthetic lethal interactions would be a promising approach to identify new targets. Furthermore, validation of hits from current DDR synthetic lethal screens may require looking at previously unknown replication functions for DDR pathway members.

Implications for patient care and cancer risks

Understanding the different functional roles of HR proteins has important implications for human patient health, especially in carriers of germline mutations in these genome maintenance genes. First, mechanistically linking particular variants in specific genes to risk for cancer development is critical for early surveillance. In families with strong histories of cancer, the identification of a germline variant allows for genetic testing for the familial deleterious alleles to inform patient risk. More intense cancer surveillance strategies lead to earlier detection of cancers before they become more aggressive and difficult to treat. Similarly, as germline DDR gene mutations may make patients more susceptible to secondary cancers due to off-target effects of DNA damaging therapeutics on non-cancer cells, surveillance post-treatment is critical for the detection of secondary cancers. With sufficient information of cellular response to

therapies, the selection of alternative front-line treatments may even be able to reduce the risk for secondary malignancies.

Understanding cellular behavior in the context of DDR deficiencies is also critical for designing appropriate cancer treatment strategies. The presence of particular mutations or fusion genes in cancer cells is sometimes used for the classification of disease and the selection of therapeutic intervention, as in AML (Grimwade and Hills, 2009; Mrózek et al., 2004). Furthermore, genomic instability in cancer cells is one of the mechanisms providing clonal selective advantages leading to resistance and treatment escape (Lipinski et al., 2016). The ability to predict drug efficacy based on patient-specific DDR responses could improve outcomes and allow for more selective targeting of cancer cells. Therefore, my findings supporting replication-mediated genomic instability provides additional context both for understanding current treatment outcomes and for designing new therapeutic approaches.

References

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S. a J.R., Behjati, S., Biankin, A. V, Bignell, G.R., Bolli, N.N., Borg, A., Borresen-Dale, A.-L., et al. (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Alexandrov, L.B., Kim, J., Haradhvala, N.J., Huang, M.N., Tian Ng, A.W., Wu, Y., Boot, A., Covington, K.R., Gordenin, D.A., Bergstrom, E.N., et al. (2020). The repertoire of mutational signatures in human cancer. Nature *578*, 94–101.

Anglana, M., Apiou, F., Bensimon, A., and Debatisse, M. (2003). Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. Cell *114*, 385–394.

Aoki, K., Aoki, M., Sugai, M., Harada, N., Miyoshi, H., Tsukamoto, T., Mizoshita, T., Tatematsu, M., Seno, H., Chiba, T., et al. (2007). Chromosomal instability by β-catenin/TCF transcription in APC or β-catenin mutant cells. Oncogene *26*, 3511–3520.

Ashihara, E., Takada, T., and Maekawa, T. (2015). Targeting the canonical Wnt/β-catenin pathway in hematological malignancies. Cancer Sci. *106*, 665–671.

Ashworth, A., and Lord, C.J. (2018). Synthetic lethal therapies for cancer: what's next after PARP inhibitors? Nat. Rev. Clin. Oncol. *15*, 564–576.

Atkinson, J., and McGlynn, P. (2009). Replication fork reversal and the maintenance of genome stability. Nucleic Acids Res. *37*, 3475–3492.

Bakker, S.T., De Winter, J.P., and Te Riele, H. (2013). Learning from a paradox: Recent insights into Fanconi anaemia through studying mouse models. DMM Dis. Model. Mech. *6*, 40–47.

Bannon, S.A., and DiNardo, C.D. (2016). Hereditary Predispositions to Myelodysplastic Syndrome. Int. J. Mol. Sci. *17*.

Bao, E.L., Nandakumar, S.K., Liao, X., Bick, A.G., Karjalainen, J., Tabaka, M., Gan, O.I., Havulinna, A.S., Kiiskinen, T.T.J., Lareau, C.A., et al. (2020). Inherited myeloproliferative neoplasm risk affects haematopoietic stem cells. Nature *586*, 769–775.

Bartek, J., Bartkova, J., and Lukas, J. (2007). DNA damage signalling guards against activated oncogenes and tumour progression. Oncogene *26*, 7773–7779.

Bell, D.W., Varley, J.M., Szydlo, T.E., Kang, D.H., Wahrer, D.C., Shannon, K.E., Lubratovich, M., Verselis, S.J., Isselbacher, K.J., Fraumeni, J.F., et al. (1999). Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. Science (80-.). *286*, 2528–2531.

Bellei, B., Cota, C., Amantea, A., Muscardin, L., and Picardo, M. (2006). Association of p53 Arg72Pro polymorphism and β-catenin accumulation in mycosis fungoides. Br. J. Dermatol. *155*, 1223–1229.

Belver, L., and Ferrando, A. (2016). The genetics and mechanisms of T cell acute lymphoblastic leukaemia. Nat. Rev. Cancer *16*, 494–507.

Bick, A.G., Weinstock, J.S., Nandakumar, S.K., Fulco, C.P., Bao, E.L., Zekavat, S.M., Szeto, M.D., Liao, X., Leventhal, M.J., Nasser, J., et al. (2020). Inherited causes of clonal haematopoiesis in 97,691 whole genomes. Nature *586*, 763–768.

de Bono, J., Mateo, J., Fizazi, K., Saad, F., Shore, N., Sandhu, S., Chi, K.N., Sartor, O., Agarwal, N., Olmos, D., et al. (2020). Olaparib for Metastatic Castration-Resistant Prostate Cancer. N. Engl. J. Med. *382*, 2091–2102.

Bredemeyer, A.L., Sharma, G.G., Huang, C.Y., Helmink, B.A., Walker, L.M., Khor, K.C., Nuskey, B., Sullivan, K.E., Pandita, T.K., Bassing, C.H., et al. (2006). ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. Nature *442*, 466–470.

Brosh, R.M., and Cantor, S.B. (2014). Molecular and cellular functions of the FANCJ DNA helicase defective in cancer and in Fanconi anemia. Front. Genet. *5*, 1–14.
Brosh, R.M., and Matson, S.W. (2020). History of DNA helicases. Genes (Basel). *11*, 1–45.

Brosh, R.M., Bellani, M., Liu, Y., and Seidman, M.M. (2017). Fanconi Anemia: A DNA repair disorder characterized by accelerated decline of the hematopoietic stem cell compartment and other features of aging. Ageing Res. Rev. *33*, 67–75.

Burrell, R. a, McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. Nature *501*, 338–345.
Cahill, D.P., Kinzler, K.W., Vogelstein, B., and Lengauer, C. (1999). Genetic instability and darwinian selection in tumours. Trends Biochem. Sci. *24*, 57–60.

Cancer Genome Atlas Research Network, Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A.G., Hoadley, K., Triche, T.J., Laird, P.W., et al. (2013). Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. N. Engl. J. Med. *368*, 2059–2074.

Ceccaldi, R., Liu, J.C., Amunugama, R., Hajdu, I., Primack, B., Petalcorin, M.I.R., O'Connor, K.W., Konstantinopoulos, P.A., Elledge, S.J., Boulton, S.J., et al. (2015). Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. Nature *518*, 258–262.

Ceccaldi, R., Rondinelli, B., and D'Andrea, A.D. (2016a). Repair Pathway Choices and

Consequences at the Double-Strand Break. Trends Cell Biol. *26*, 52–64.

Ceccaldi, R., Sarangi, P., and D'Andrea, A.D. (2016b). The Fanconi anaemia pathway: new players and new functions. Nat. Rev. Mol. Cell Biol. *17*, 337–349.

Chan, D.A., and Giaccia, A.J. (2011). Harnessing synthetic lethal interactions in anticancer drug discovery. Nat. Rev. Drug Discov. *10*, 351–364.

Chapman, J.R., Taylor, M.R.G., and Boulton, S.J. (2012). Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. Mol. Cell *47*, 497–510.

Chen, C.-C., Feng, W., Lim, P.X., Kass, E.M., and Jasin, M. (2018). Homology-Directed Repair and the Role of BRCA1, BRCA2, and Related Proteins in Genome Integrity and Cancer. Annu. Rev. Cancer Biol. *2*, 313–336.

Chen, K., Wallis, J.W., McLellan, M.D., Larson, D.E., Kalicki, J.M., Pohl, C.S., McGrath, S.D., Wendl, M.C., Zhang, Q., Locke, D.P., et al. (2009). BreakDancer: An algorithm for high-resolution mapping of genomic structural variation. Nat. Methods *6*, 677–681.

Cheng, N.C., van de Vrugt, H.J., van der Valk, M.A., Oostra, A.B., Krimpenfort, P., de Vries, Y., Joenje, H., Berns, A., and Arwert, F. (2000). Mice with a targeted disruption of the Fanconi anemia homolog Fanca. Hum. Mol. Genet. *9*, 1805–1811.

Churpek, J.E., Marquez, R., Neistadt, B., Claussen, K., Lee, M.K., Churpek, M.M., Huo, D., Weiner, H., Bannerjee, M., Godley, L.A., et al. (2016). Inherited mutations in cancer susceptibility genes are common among survivors of breast cancer who develop therapy-related leukemia. Cancer *122*, 304–311.

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

Coleman, R.L., Oza, A.M., Lorusso, D., Aghajanian, C., Oaknin, A., Dean, A., Colombo, N., Weberpals, J.I., Clamp, A., Scambia, G., et al. (2017). Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet (London, England) *390*, 1949–1961.

Coleman, R.L., Fleming, G.F., Brady, M.F., Swisher, E.M., Steffensen, K.D., Friedlander, M., Okamoto, A., Moore, K.N., Efrat Ben-Baruch, N., Werner, T.L., et al. (2019). Veliparib with First-Line Chemotherapy and as Maintenance Therapy in Ovarian Cancer. N. Engl. J. Med. *381*, 2403–2415.

Cong, K., Peng, M., Kousholt, A.N., Turchi, J.J., Rothenberg, E., Cantor, S.B., Cong, K., Peng, M., Kousholt, A.N., Lee, W.T.C., et al. (2021). Article Replication gaps are a key determinant of PARP inhibitor synthetic lethality with BRCA deficiency Article Replication gaps are a key determinant of PARP inhibitor synthetic lethality with BRCA

deficiency. Mol. Cell 1–17.

Coombs, C.C., Zehir, A., Devlin, S.M., Kishtagari, A., Syed, A., Jonsson, P., Hyman, D.M., Solit, D.B., Robson, M.E., Baselga, J., et al. (2017). Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. Cell Stem Cell *21*, 374-382.e4.

Couch, F.J., Shimelis, H., Hu, C., Hart, S.N., Polley, E.C., Na, J., Hallberg, E., Moore, R., Thomas, A., Lilyquist, J., et al. (2017). Associations Between Cancer Predisposition Testing Panel Genes and Breast Cancer. JAMA Oncol *3*, 1190–1196.

Cybulski, C., Górski, B., Huzarski, T., Masojć, B., Mierzejewski, M., Debniak, T., Teodorczyk, U., Byrski, T., Gronwald, J., Matyjasik, J., et al. (2004). CHEK2 is a multiorgan cancer susceptibility gene. Am J Hum Genet *75*, 1131–1135.

Daley, J.M., Gaines, W.A., Kwon, Y., and Sung, P. (2014). Regulation of DNA Pairing in Homologous Recombination. Cold Spring Harb. Perspect. Biol. *6*, 1–15.

Dalgaard, J.Z. (2012). Causes and consequences of ribonucleotide incorporation into nuclear DNA. Trends Genet. *28*, 592–597.

Deans, A.J., and West, S.C. (2011). DNA interstrand crosslink repair and cancer. Nat. Rev. Cancer *11*, 467–480.

Desrichard, A., Bidet, Y., Uhrhammer, N., and Bignon, Y.J. (2011). CHEK2 contribution to hereditary breast cancer in non-BRCA families. Breast Cancer Res *13*, R119.

Dev, H., Chiang, T.W.W., Lescale, C., de Krijger, I., Martin, A.G., Pilger, D., Coates, J., Sczaniecka-Clift, M., Wei, W., Ostermaier, M., et al. (2018). Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. Nat. Cell Biol. *20*, 954–965.

Dharan, N.J., Yeh, P., Bloch, M., Yeung, M.M., Baker, D., Guinto, J., Roth, N., Ftouni, S., Ognenovska, K., Smith, D., et al. (2021). HIV is associated with an increased risk of age-related clonal hematopoiesis among older adults. Nat Med *27*, 1006–1011.

Dose, M., Emmanuel, A.O., Chaumeil, J., Zhang, J., Sun, T., Germar, K., Aghajani, K., Davis, E.M., Keerthivasan, S., Bredemeyer, A.L., et al. (2014). β-Catenin induces T-cell transformation by promoting genomic instability. Proc. Natl. Acad. Sci. U. S. A. *111*, 391–396.

Dynan, W.S., and Yoo, S. (1998). Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. Nucleic Acids Res. *26*, 1551–1559.

Emmanuel, A.O., Arnovitz, S., Haghi, L., Mathur, P.S., Mondal, S., Quandt, J., Okoreeh, M.K., Maienschein-Cline, M., Khazaie, K., Dose, M., et al. (2018). TCF-1 and HEB

cooperate to establish the epigenetic and transcription profiles of CD4 + CD8 + thymocytes. Nat. Immunol. *19*, 1366–1378.

Escobar, G., Mangani, D., and Anderson, A.C. (2020). T cell factor 1: A master regulator of the T cell response in disease. Sci. Immunol. *5*.

Escribano-Díaz, C., Orthwein, A., Fradet-Turcotte, A., Xing, M., Young, J.T.F., Tkáč, J., Cook, M.A., Rosebrock, A.P., Munro, M., Canny, M.D., et al. (2013). A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. Mol. Cell *49*, 872–883.

Fackenthal, J.D., and Olopade, O.I. (2007). Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. Nat Rev Cancer *7*, 937–948.

Falck, J., Lukas, C., Protopopova, M., Lukas, J., Selivanova, G., and Bartek, J. (2001). Functional impact of concomitant versus alternative defects in the Chk2-p53 tumour suppressor pathway. Oncogene *20*, 5503–5510.

Farmer, H., McCabe, N., Lord, C.J., Tutt, A.N.J., Johnson, D. a, Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C., et al. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917–921.

Fernandez-Vidal, A., Guitton-Sert, L., Cadoret, J., Drac, M., Schwob, E., Baldacci, G., Cazaux, C., and Hoffmann, J.-S. (2014). A role for DNA polymerase θ in the timing of DNA replication. Nat. Commun. *5*, 4285.

Feurstein, S., Drazer, M.W., and Godley, L.A. (2016). Genetic predisposition to leukemia and other hematologic malignancies. Semin. Oncol. *43*, 598–608.

Filippini, S.E., and Vega, A. (2013). Breast cancer genes: beyond BRCA1 and BRCA2. Front. Biosci. (Landmark Ed. *18*, 1358–1372.

Friedenson, B. (2007). The BRCA1/2 pathway prevents hematologic cancers in addition to breast and ovarian cancers. BMC Cancer *7*, 152.

Friedenson, B. (2016). Comment on 'The incidence of leukaemia in women with BRCA1 and BRCA2 mutations: an International Prospective Cohort Study.' Br. J. Cancer *115*, e2–e2.

Fugmann, S.D., Lee, A.I., Shockett, P.E., Villey, I.J., and Schatz, D.G. (2000). The RAG proteins and V(D)J recombination: Complexes, ends, and transposition. Annu. Rev. Immunol. *18*, 495–527.

Gekas, C., D'Altri, T., Aligué, R., González, J., Espinosa, L., and Bigas, A. (2016). β-Catenin is required for T cell leukemia initiation and MYC transcription downstream of Notch1. Leukemia 1–9.

Germar, K., Dose, M., Konstantinou, T., Zhang, J., Wang, H., Lobry, C., Arnett, K.L., Blacklow, S.C., Aifantis, I., Aster, J.C., et al. (2011). T-cell factor 1 is a gatekeeper for T-cell specification in response to Notch signaling. Proc. Natl. Acad. Sci. U. S. A. *108*, 20060–20065.

Giambra, V., Jenkins, C.E., Lam, S.H., Hoofd, C., Belmonte, M., Wang, X., Gusscott, S., Gracias, D., and Weng, A.P. (2015). Leukemia stem cells in T-ALL require active Hif1α and Wnt signaling. Blood *125*, 3917–3927.

Girard, E., Eon-Marchais, S., Olaso, R., Renault, A.L., Damiola, F., Dondon, M.G., Barjhoux, L., Goidin, D., Meyer, V., Le Gal, D., et al. (2019). Familial breast cancer and DNA repair genes: Insights into known and novel susceptibility genes from the GENESIS study, and implications for multigene panel testing. Int J Cancer *144*, 1962–1974.

Girardi, T., Vicente, C., Cools, J., and De Keersmaecker, K. (2017). The genetics and molecular biology of T-ALL. Blood *129*, 1113–1123.

González-Martín, A., Pothuri, B., Vergote, I., DePont Christensen, R., Graybill, W., Mirza, M.R., McCormick, C., Lorusso, D., Hoskins, P., Freyer, G., et al. (2019). Niraparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. N. Engl. J. Med. *381*, 2391–2402.

Gostissa, M., Alt, F.W., and Chiarle, R. (2011). Mechanisms that Promote and Suppress Chromosomal Translocations in Lymphocytes.

Graux, C., Cools, J., Michaux, L., Vandenberghe, P., and Hagemeijer, A. (2006). Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast. Leukemia *20*, 1496–1510.

Grimwade, D., and Hills, R.K. (2009). Independent prognostic factors for AML outcome. Hematol. Am. Soc. Hematol. Educ. Progr. 385–395.

Groen, R.W.J.J., Oud, M.E.C.M.C.M., Schilder-Tol, E.J.M.M., Overdijk, M.B., Ten Berge, D., Nusse, R., Spaargaren, M., and Pals, S.T. (2008). Illegitimate WNT pathway activation by β-catenin mutation or autocrine stimulation in T-cell malignancies. Cancer Res. *68*, 6969–6977.

Guo, W., Lasky, J.L., Chang, C.J., Mosessian, S., Lewis, X., Xiao, Y., Yeh, J.E., Chen, J.Y., Iruela-Arispe, M.L., Varella-Garcia, M., et al. (2008). Multi-genetic events collaboratively contribute to Pten-null leukaemia stem-cell formation. Nature *453*, 529–533.

Guo, Z., Dose, M., Kovalovsky, D., Chang, R., O'Neil, J., Look, A.T., Von Boehmer, H., Khazaie, K., Gounari, F., Neil, J.O., et al. (2007). B-Catenin Stabilization Stalls the

Transition From Double-Positive To Single-Positive Stage and Predisposes Thymocytes To Malignant Transformation. Blood *109*, 5463–5472.

Gutierrez, A., Sanda, T., Grebliunaite, R., Carracedo, A., Salmena, L., Ahn, Y., Dahlberg, S., Neuberg, D., Moreau, L.A., Winter, S.S., et al. (2009). High frequency of PTEN, PI3K, and AKT abnormalities in T-cell acute lymphoblastic leukemia. Blood *114*, 647–650.

Hall, M.J., Bernhisel, R., Hughes, E., Larson, K., Rosenthal, E.T., Singh, N.A., Lancaster, J.M., and Kurian, A.W. (2021). Germline Pathogenic Variants in the Ataxia Telangiectasia Mutated (ATM) Gene are Associated with High and Moderate Risks for Multiple Cancers. Cancer Prev. Res. (Phila). *14*, 433–440.

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

Harada, N., Tamai, Y., Ishikawa, T.O., Sauer, B., Takaku, K., Oshima, M., and Taketo, M.M. (1999). Intestinal polyposis in mice with a dominant stable mutation of the β-catenin gene. EMBO J. *18*, 5931–5942.

Havranek, O., Kleiblova, P., Hojny, J., Lhota, F., Soucek, P., Trneny, M., and Kleibl, Z. (2015). Association of germline CHEK2 gene variants with risk and prognosis of Non-Hodgkin Lymphoma. PLoS One *10*, 1–15.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol. Cell *38*, 576–589.

Helleday, T. (2011). The underlying mechanism for the PARP and BRCA synthetic lethality: Clearing up the misunderstandings. Mol. Oncol. *5*, 387–393.

Hinds, D.A., Barnholt, K.E., Mesa, R.A., Kiefer, A.K., Do, C.B., Eriksson, N., Mountain, J.L., Francke, U., Tung, J.Y., Nguyen, H.M., et al. (2016). Germ line variants predispose to both JAK2 V617F clonal hematopoiesis and myeloproliferative neoplasms. Blood *128*, 1121–1128.

Hofstatter, E.W., Domchek, S.M., Miron, A., Garber, J., Wang, M., Componeschi, K., Boghossian, L., Miron, P.L., Nathanson, K.L., and Tung, N. (2011). PALB2 mutations in familial breast and pancreatic cancer. Fam. Cancer *10*, 225–231.

Hosokawa, H., and Rothenberg, E. V (2018). Cytokines, Transcription Factors, and the Initiation of T-Cell Development. Cold Spring Harb. Perspect. Biol. *10*, 1–20.

Hu, C., Hart, S.N., Gnanaolivu, R., Huang, H., Lee, K.Y., Na, J., Gao, C., Lilyquist, J., Yadav, S., Boddicker, N.J., et al. (2021). A Population-Based Study of Genes

Previously Implicated in Breast Cancer. N. Engl. J. Med. *384*, 440–451.

Ikawa, T., Hirose, S., Masuda, K., Kakugawa, K., Satoh, R., Shibano-Satoh, A., Kominami, R., Katsura, Y., and Kawamoto, H. (2010). An essential developmental checkpoint for production of the T cell lineage. Science *329*, 93–96.

Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). A Landscape of Pharmacogenomic Interactions in Cancer. Cell *166*, 740–754.

Iqbal, J., Nussenzweig, A., Lubinski, J., Byrski, T., Eisen, A., Bordeleau, L., Tung, N.M., Manoukian, S., Phelan, C.M., Sun, P., et al. (2016). The incidence of leukaemia in women with BRCA1 and BRCA2 mutations: an International Prospective Cohort Study. Br. J. Cancer *114*, 1160–1164.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071–1078.

Jagtap, P., and Szabó, C. (2005). Poly(ADP-ribose) polymerase and the therapeutic effects of its inhibitors. Nat. Rev. Drug Discov. *4*, 421–440.

Jaiswal, S., and Ebert, B.L. (2019). Clonal hematopoiesis in human aging and disease. Science *366*, 139–148.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V, Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., et al. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med *371*, 2488–2498.

Jaiswal, S., Natarajan, P., Silver, A.J., Gibson, C.J., Bick, A.G., Shvartz, E., McConkey, M., Gupta, N., Gabriel, S., Ardissino, D., et al. (2017). Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. N Engl J Med *377*, 111–121.

Janiszewska, H., Bak, A., Pilarska, M., Heise, M., Junkiert-Czarnecka, A., Kuliszkiewicz-Janus, M., Całbecka, M., Jaźwiec, B., Wołowiec, D., Kuliczkowski, K., et al. (2012). A risk of essential thrombocythemia in carriers of constitutional CHEK2 gene mutations. Haematologica *97*, 366–370.

Janiszewska, H., Bąk, A., Skonieczka, K., Jaśkowiec, A., Kiełbiński, M., Jachalska, A., Czyżewska, M., Jaźwiec, B., Kuliszkiewicz-Janus, M., Czyż, J., et al. (2018). Constitutional mutations of the CHEK2 gene are a risk factor for MDS, but not for de novo AML. Leuk Res *70*, 74–78.

Jasin, M., and Rothstein, R. (2013). Repair of strand breaks by homologous recombination. Cold Spring Harb. Perspect. Biol. *5*, a012740.

Johnson, J.L., Georgakilas, G., Petrovic, J., Kurachi, M., Cai, S., Harly, C., Pear, W.S.,

Bhandoola, A., Wherry, E.J., and Vahedi, G. (2018). Lineage-Determining Transcription Factor TCF-1 Initiates the Epigenetic Identity of T Cells. Immunity *48*, 243-257.e10.

Jones, R.M., and Petermann, E. (2012). Replication fork dynamics and the DNA damage response. Biochem. J. *443*, 13–26.

Kaczmarek-Ryś, M., Ziemnicka, K., Hryhorowicz, S.T., Górczak, K., Hoppe-Gołębiewska, J., Skrzypczak-Zielińska, M., Tomys, M., Gołąb, M., Szkudlarek, M., Budny, B., et al. (2015). The c.470 T > C CHEK2 missense variant increases the risk of differentiated thyroid carcinoma in the Great Poland population. Hered. Cancer Clin. Pract. *13*, 8.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. Nature *502*, 333–339.

Kar, S.P., Quiros, P.M., Gu, M., Jiang, T., Langdon, R., Iyer, V., Barcena, C., Vijayabaskar, M.S., Fabre, M.A., Carter, P., et al. (2022). Genome-wide analyses of 200,453 individuals yields new insights into the causes and consequences of clonal hematopoiesis. MedRxiv 2022.01.06.22268846.

Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443.

Karczewski, K.J., Solomonson, M., Chao, K.R., Goodrich, J.K., Tiao, G., Lu, W., Riley-Gillis, B.M., Tsai, E.A., Kim, H.I., Zheng, X., et al. (2021). Systematic single-variant and gene-based association testing of 3,700 phenotypes in 281,850 UK Biobank exomes. MedRxiv 2021.06.19.21259117.

Kastan, M.B., and Bartek, J. (2004). Cell-cycle checkpoints and cancer. Nature *432*, 316–323.

Kaur, E., Agrawal, R., and Sengupta, S. (2021). Functions of BLM Helicase in Cells: Is It Acting Like a Double-Edged Sword? Front. Genet. *12*, 1–14.

Kaveri, D., Kastner, P., Dembélé, D., Nerlov, C., Chan, S., and Kirstetter, P. (2013). β-Catenin activation synergizes with Pten loss and Myc overexpression in Notch-independent T-ALL. Blood *122*, 694–704.

Kiiski, J.I., Pelttari, L.M., Khan, S., Freysteinsdottir, E.S., Reynisdottir, I., Hart, S.N., Shimelis, H., Vilske, S., Kallioniemi, A., Schleutker, J., et al. (2014). Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. Proc. Natl. Acad. Sci. U. S. A. *111*, 15172–15177.

King, M.-C., Marks, J., Mandell, J., and Group, N.Y.B.C.S. (2003). Breast and ovarian

cancer risks due to inherited mutations in BRCA1 and BRCA2. Science *302*, 643–646.

Kinnersley, B., Kamatani, Y., Labussière, M., Wang, Y., Galan, P., Mokhtari, K., Delattre, J.Y., Gousias, K., Schramm, J., Schoemaker, M.J., et al. (2016). Search for new loci and low-frequency variants influencing glioma risk by exome-array analysis. Eur J Hum Genet *24*, 717–724.

Kolinjivadi, A.M., Sannino, V., de Antoni, A., Técher, H., Baldi, G., and Costanzo, V. (2017). Moonlighting at replication forks – a new life for homologous recombination proteins BRCA1, BRCA2 and RAD51. FEBS Lett. *591*, 1083–1100.

Koole, W., van Schendel, R., Karambelas, A.E., van Heteren, J.T., Okihara, K.L., and Tijsterman, M. (2014). A polymerase theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat. Commun. *5*, 1–10.

Kote-Jarai, Z., Salmon, a, Mengitsu, T., Copeland, M., Ardern-Jones, a, Locke, I., Shanley, S., Summersgill, B., Lu, Y.-J., Shipley, J., et al. (2006). Increased level of chromosomal damage after irradiation of lymphocytes from BRCA1 mutation carriers. Br. J. Cancer *94*, 308–310.

Kreslavsky, T., Gleimer, M., Miyazaki, M., Choi, Y., Gagnon, E., Murre, C., Sicinski, P., and von Boehmer, H. (2012). β-Selection-Induced Proliferation Is Required for αβ T Cell Differentiation. Immunity *37*, 840–853.

Küppers, R. (2005). Mechanisms of B-cell lymphoma pathogenesis. Nat. Rev. Cancer *5*, 251–262.

Larmonie, N.S.D., Dik, W.A., Meijerink, J.P.P., Homminga, I., van Dongen, J.J.M., and Langerak, A.W. (2013). Breakpoint sites disclose the role of the V(D)J recombination machinery in the formation of T-cell receptor (TCR) and non-TCR associated aberrations in T-cell acute lymphoblastic leukemia. Haematologica *98*, 1173–1184.

Ledermann, J., Harter, P., Gourley, C., Friedlander, M., Vergote, I., Rustin, G., Scott, C.L., Meier, W., Shapira-Frommer, R., Safra, T., et al. (2014). Olaparib maintenance therapy in patients with platinum-sensitive relapsed serous ovarian cancer: A preplanned retrospective analysis of outcomes by BRCA status in a randomised phase 2 trial. Lancet Oncol. *15*, 852–861.

Lee, J.-H., and Paull, T.T. (2007). Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. Oncogene *26*, 7741–7748.

Lee, J.S., Collins, K.M., Brown, A.L., Lee, C.H., and Chung, J.H. (2000). hCds1-mediated phosphorylation of BRCA1 regulates the DNA damage response. Nature *404*, 201–204.

Leggett, R.M., Clavijo, B.J., Clissold, L., Clark, M.D., and Caccamo, M. (2014). Next

clip: An analysis and read preparation tool for nextera long mate pair libraries. Bioinformatics *30*, 566–568.

Lemaçon, D., Jackson, J., Quinet, A., Brickner, J.R., Li, S., Yazinski, S., You, Z., Ira, G., Zou, L., Mosammaparast, N., et al. (2017). MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. Nat. Commun. *8*, 860.

Lento, W., Congdon, K., Voermans, C., Kritzik, M., and Reya, T. (2013). Wnt signaling in normal and malignant hematopoiesis. Cold Spring Harb. Perspect. Biol. *5*, a008011–a008011.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, J., Williams, B.L., Haire, L.F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M.B., Jackson, S.P., Smerdon, S.J., Ridgeway, T., et al. (2002). Structural and Functional Versatility of the FHA Domain in DNA-Damage Signaling by the Tumor Suppressor Kinase Chk2. *9*, 1045–1054.

Liang, M., Zhang, Y., Sun, C., Rizeq, F.K., Min, M., Shi, T., and Sun, Y. (2018). Association Between CHEK2*1100delC and Breast Cancer: A Systematic Review and Meta-Analysis. Mol. Diagn. Ther. *22*, 397–407.

Liao, H., Ji, F., Helleday, T., and Ying, S. (2018). Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments. EMBO Rep. *19*, 1–18.

Lieber, M.R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu. Rev. Biochem. *79*, 181–211.

Lieber, M.R. (2016). Mechanisms of human lymphoid chromosomal translocations. Nat. Rev. Cancer *16*, 387–398.

Lindahl, T., and Barnes, D.E. (2000). Repair of endogenous DNA damage. Cold Spring Harb. Symp. Quant. Biol. *65*, 127–133.

Lipinski, K.A., Barber, L.J., Davies, M.N., Ashenden, M., Sottoriva, A., and Gerlinger, M. (2016). Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine. Trends in Cancer *2*, 49–63.

Liu, C., Wang, Q.-S., and Wang, Y.-J. (2012). The CHEK2 I157T variant and colorectal cancer susceptibility: a systematic review and meta-analysis. Asian Pac. J. Cancer Prev. *13*, 2051–2055.

Liu, W., Krishnamoorthy, A., Zhao, R., and Cortez, D. (2020). Two replication fork remodeling pathways generate nuclease substrates for distinct fork protection factors. Sci. Adv. *6*, 1–11.

Loberg, M.A., Bell, R.K., Goodwin, L.O., Eudy, E., Miles, L.A., SanMiguel, J.M., Young, K., Bergstrom, D.E., Levine, R.L., Schneider, R.K., et al. (2019). Sequentially inducible mouse models reveal that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. Leukemia *33*, 1635–1649.

Lord, C.J., and Ashworth, A. (2016). BRCAness revisited. Nat. Rev. Cancer *16*, 110–120.

Lord, C.J., and Ashworth, A. (2017). PARP inhibitors: Synthetic lethality in the clinic. Science (80-.). *355*, 1152–1158.

Lossaint, G., Larroque, M., Ribeyre, C., Bec, N., Larroque, C., Décaillet, C., Gari, K., and Constantinou, A. (2013). FANCD2 binds MCM proteins and controls replisome function upon activation of s phase checkpoint signaling. Mol. Cell *51*, 678–690.

Love, P.E., and Bhandoola, A. (2011). Signal integration and crosstalk during thymocyte migration and emigration. Nat. Rev. Immunol. *11*, 469–477.

Ma, A., Pena, J.C., Chang, B., Margosian, E., Davidson, L., Alt, F.W., and Thompson, C.B. (1995). Bclx regulates the survival of double-positive thymocytes. Proc. Natl. Acad. Sci. U. S. A. *92*, 4763–4767.

Mamrak, N.E., Shimamura, A., and Howlett, N.G. (2017). Recent discoveries in the molecular pathogenesis of the inherited bone marrow failure syndrome Fanconi anemia. Blood Rev. *31*, 93–99.

Martin, R.W., Orelli, B.J., Yamazoe, M., Minn, A.J., Takeda, S., and Bishop, D.K. (2007). RAD51 up-regulation bypasses BRCA1 function and is a common feature of BRCA1-deficient breast tumors. Cancer Res. *67*, 9658–9665.

Martins, F.C., De, S., Almendro, V., Gönen, M., Park, S.Y., Blum, J.L., Herlihy, W., Ethington, G., Schnitt, S.J., Tung, N., et al. (2012). Evolutionary pathways in BRCA1-associated breast tumors. Cancer Discov. *2*, 503–511.

Mateos-gomez, P.A., Gong, F., Nair, N., Miller, K.M., Lazzerini-denchi, E., and Sfeir, A. (2014). Mammalian polymerase θ promotes alternative NHEJ and suppresses recombination. Nature *518*, 254–257.

Mateos-Gomez, P.A., Kent, T., Deng, S.K., Mcdevitt, S., Kashkina, E., Hoang, T.M., Pomerantz, R.T., and Sfeir, A. (2017). The helicase domain of Polθ counteracts RPA to promote alt-NHEJ. Nat. Struct. Mol. Biol. *24*, 1116–1123.

Matsuoka, S., Huang, M., and Elledge, S.J. (1998). Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. Science *282*, 1893–1897.

McNerney, M.E., Godley, L.A., and Le Beau, M.M. (2017). Therapy-related myeloid neoplasms: when genetics and environment collide. Nat. Rev. Cancer *17*, 513–527.

Meijers-Heijboer, H., van den Ouweland, A., Klijn, J., Wasielewski, M., de Snoo, A., Oldenburg, R., Hollestelle, A., Houben, M., Crepin, E., van Veghel-Plandsoen, M., et al. (2002). Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet *31*, 55–59.

Meisel, M., Hinterleitner, R., Pacis, A., Chen, L., Earley, Z.M., Mayassi, T., Pierre, J.F., Ernest, J.D., Galipeau, H.J., Thuille, N., et al. (2018). Microbial signals drive pre-leukaemic myeloproliferation in a Tet2-deficient host. Nature *557*, 580–584.

Merelli, I., Guffanti, A., Fabbri, M., Cocito, A., Furia, L., Grazini, U., Bonnal, R.J., Milanesi, L., and McBlane, F. (2010). RSSsite: A reference database and prediction tool for the identification of cryptic Recombination Signal Sequences in human and murine genomes. Nucleic Acids Res. *38*, W262-7.

Mertens, F., Johansson, B., Fioretos, T., and Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. Nat. Rev. Cancer *15*, 371–381.

Mgbemena, V.E., Signer, R.A.J., Wijayatunge, R., Laxson, T., Morrison, S.J., and Ross, T.S. (2017). Distinct Brca1 Mutations Differentially Reduce Hematopoietic Stem Cell Function. Cell Rep. *18*, 947–960.

Michl, J., Zimmer, J., and Tarsounas, M. (2016). Interplay between Fanconi anemia and homologous recombination pathways in genome integrity. EMBO J. *35*, 909–923.

Mielke, L.A., Liao, Y., Clemens, E.B., Firth, M.A., Duckworth, B., Huang, Q., Almeida, F.F., Chopin, M., Koay, H.F., Bell, C.A., et al. (2019). TCF-1 limits the formation of Tc17 cells via repression of the MAF–RORγt axis. J. Exp. Med. *216*, 1682–1699.

Mijic, S., Zellweger, R., Chappidi, N., Berti, M., Jacobs, K., Mutreja, K., Ursich, S., Ray Chaudhuri, A., Nussenzweig, A., Janscak, P., et al. (2017). Replication fork reversal triggers fork degradation in BRCA2-defective cells. Nat. Commun. *8*, 859.

Mirza, M.R., Monk, B.J., Herrstedt, J., Oza, A.M., Mahner, S., Redondo, A., Fabbro, M., Ledermann, J.A., Lorusso, D., Vergote, I., et al. (2016). Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. N. Engl. J. Med. *375*, 2154–2164.

Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer *7*, 233–245.

Mohiuddin, I.S., and Kang, M.H. (2019). DNA-PK as an Emerging Therapeutic Target in

Cancer. Front. Oncol. *9*, 635.

Mohrin, M., Bourke, E., Alexander, D., Warr, M.R., Barry-Holson, K., Le Beau, M.M., Morrison, C.G., and Passegué, E. (2010). Hematopoietic Stem Cell Quiescence Promotes Error-Prone DNA Repair and Mutagenesis. Cell Stem Cell *7*, 174–185.

Morice, P.M., Leary, A., Dolladille, C., Chrétien, B., Poulain, L., González-Martín, A., Moore, K., O'Reilly, E.M., Ray-Coquard, I., and Alexandre, J. (2021). Myelodysplastic syndrome and acute myeloid leukaemia in patients treated with PARP inhibitors: a safety meta-analysis of randomised controlled trials and a retrospective study of the WHO pharmacovigilance database. Lancet Haematol. *8*, e122–e134.

Mosimann, C., Hausmann, G., and Basler, K. (2009). β-Catenin hits chromatin: Regulation of Wnt target gene activation. Nat. Rev. Mol. Cell Biol. *10*, 276–286.

Mosler, T., Conte, F., Longo, G.M.C., Mikicic, I., Kreim, N., Möckel, M.M., Petrosino, G., Flach, J., Barau, J., Luke, B., et al. (2021). R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. Nat. Commun. *12*, 7314.

Moynahan, M.E., and Jasin, M. (2010). Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. Nat. Rev. Mol. Cell Biol. *11*, 196–207.
Mrózek, K., Heerema, N.A., and Bloomfield, C.D. (2004). Cytogenetics in acute leukemia. Blood Rev. *18*, 115–136.

Nalepa, G., and Clapp, D.W. (2018). Fanconi anaemia and cancer: An intricate relationship. Nat. Rev. Cancer *18*, 168–185.

Neelsen, K.J., Zanini, I.M.Y., Mijic, S., Herrador, R., Zellweger, R., Ray Chaudhuri, A., Creavin, K.D., Blow, J.J., and Lopes, M. (2013). Deregulated origin licensing leads to chromosomal breaks by rereplication of a gapped DNA template. Genes Dev. *27*, 2537–2542.

Ng, O.H., Erbilgin, Y., Firtina, S., Celkan, T., Karakas, Z., Aydogan, G., Turkkan, E., Yildirmak, Y., Timur, C., Zengin, E., et al. (2014). Deregulated WNT signaling in childhood T-cell acute lymphoblastic leukemia. Blood Cancer J. *4*, e192.

Niraj, J., Färkkilä, A., and D'Andrea, A.D. (2019). The Fanconi Anemia Pathway in Cancer. Annu. Rev. Cancer Biol. *3*, 457–478.

Norquist, B.M., Garcia, R.L., Allison, K.H., Jokinen, C.H., Kernochan, L.E., Pizzi, C.C., Barrow, B.J., Goff, B.A., and Swisher, E.M. (2010). The molecular pathogenesis of hereditary ovarian carcinoma. Cancer *116*, 5261–5271.

Nussenzweig, A., and Nussenzweig, M.C. (2010). Origin of Chromosomal Translocations in Lymphoid Cancer. Cell *141*, 27–38.

Obazee, O., Archibugi, L., Andriulli, A., Soucek, P., Małecka-Panas, E., Ivanauskas, A., Johnson, T., Gazouli, M., Pausch, T., Lawlor, R.T., et al. (2019). Germline BRCA2 K3326X and CHEK2 I157T mutations increase risk for sporadic pancreatic ductal adenocarcinoma. Int J Cancer *145*, 686–693.

Oguro, H., Ding, L., and Morrison, S.J. (2013). SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. Cell Stem Cell *13*, 102–116.

Oliver, A.W., Knapp, S., and Pearl, L.H. (2007). Activation segment exchange: a common mechanism of kinase autophosphorylation? Trends Biochem. Sci. *32*, 351–356.

Olsen, J.H., Hahnemann, J.M., Børresen-Dale, A.L., Brøndum-Nielsen, K., Hammarström, L., Kleinerman, R., Kääriäinen, H., Lönnqvist, T., Sankila, R., Seersholm, N., et al. (2001). Cancer in patients with ataxia-telangiectasia and in their relatives in the nordic countries. J. Natl. Cancer Inst. *93*, 121–127.

Omura, H., Oikawa, D., Nakane, T., Kato, M., Ishii, R., Ishitani, R., Tokunaga, F., and Nureki, O. (2016). Structural and Functional Analysis of DDX41: a bispecific immune receptor for DNA and cyclic dinucleotide. Sci. Rep. *6*, 34756.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631–644.

Palomero, T., Sulis, M.L., Cortina, M., Real, P.J., Barnes, K., Ciofani, M., Caparros, E., Buteau, J., Brown, K., Perkins, S.L., et al. (2007). Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. Nat. Med. *13*, 1203–1210.

Panday, A., Willis, N.A., Elango, R., Menghi, F., Duffey, E.E., Liu, E.T., and Scully, R. (2021). FANCM regulates repair pathway choice at stalled replication forks. Mol. Cell *81*, 2428-2444.e6.

Páral, P., Faltusová, K., Molík, M., Renešová, N., Šefc, L., and Nečas, E. (2018). Cell cycle and differentiation of Sca-1+ and Sca-1- hematopoietic stem and progenitor cells. Cell Cycle *17*, 1979–1991.

Parmar, K., D'Andrea, A., and Niedernhofer, L.J. (2009). Mouse models of Fanconi anemia. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *668*, 133–140.

Pathania, S., Bade, S., Le Guillou, M., Burke, K., Reed, R., Bowman-Colin, C., Su, Y., Ting, D.T., Polyak, K., Richardson, A.L., et al. (2014). BRCA1 haploinsufficiency for replication stress suppression in primary cells. Nat. Commun. *5*, 5496.

Peng, M., Cong, K., Panzarino, N.J., Nayak, S., Calvo, J., Deng, B., Zhu, L.J., Morocz, M., Hegedus, L., Haracska, L., et al. (2018). Opposing Roles of FANCJ and HLTF

Protect Forks and Restrain Replication during Stress. Cell Rep. *24*, 3251–3261.

Pietras, E.M., Reynaud, D., Kang, Y.A., Carlin, D., Calero-Nieto, F.J., Leavitt, A.D., Stuart, J.A., Göttgens, B., and Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. Cell Stem Cell *17*, 35–46.

Poli, J., Tsaponina, O., Crabbé, L., Keszthelyi, A., Pantesco, V., Chabes, A., Lengronne, A., and Pasero, P. (2012). dNTP pools determine fork progression and origin usage under replication stress. EMBO J. *31*, 883–894.

Prado, F. (2018). Homologous Recombination: To Fork and Beyond. Genes (Basel). *9*.
Qian, J., Wang, Q., Dose, M., Pruett, N., Kieffer-Kwon, K.-R., Resch, W., Liang, G., Tang, Z., Mathé, E., Benner, C., et al. (2014). B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. Cell *159*, 1524–1537.

Qiu, J., Papatsenko, D., Niu, X., Schaniel, C., and Moore, K. (2014). Divisional history and hematopoietic stem cell function during homeostasis. Stem Cell Reports *2*, 473–490.

Quandt, J., Arnovitz, S., Haghi, L., Woehlk, J., Mohsin, A., Okoreeh, M., Mathur, P.S., Emmanuel, A.O., Osman, A., Krishnan, M., et al. (2021). Wnt–β-catenin activation epigenetically reprograms Treg cells in inflammatory bowel disease and dysplastic progression. Nat. Immunol. *22*, 471–484.

Quinet, A., Lemaçon, D., and Vindigni, A. (2017). Replication Fork Reversal: Players and Guardians. Mol. Cell *68*, 830–833.

Rafnar, T., Gudbjartsson, D.F., Sulem, P., Jonasdottir, A., Sigurdsson, A., Jonasdottir, A., Besenbacher, S., Lundin, P., Stacey, S.N., Gudmundsson, J., et al. (2011). Mutations in BRIP1 confer high risk of ovarian cancer. Nat. Genet. *43*, 1104–1107.

Ram-Wolff, C., Martin-Garcia, N., Bensussan, A., Bagot, M., and Ortonne, N. (2010). Histopathologic diagnosis of lymphomatous versus inflammatory erythroderma: A morphologic and phenotypic study on 47 skin biopsies. Am. J. Dermatopathol. *32*, 755–763.

Ray-Coquard, I., Pautier, P., Pignata, S., Pérol, D., González-Martín, A., Berger, R., Fujiwara, K., Vergote, I., Colombo, N., Mäenpää, J., et al. (2019). Olaparib plus Bevacizumab as First-Line Maintenance in Ovarian Cancer. N. Engl. J. Med. *381*, 2416–2428.

Ray Chaudhuri, A., Callen, E., Ding, X., Gogola, E., Duarte, A.A., Lee, J.-E., Wong, N., Lafarga, V., Calvo, J.A., Panzarino, N.J., et al. (2016). Replication fork stability confers chemoresistance in BRCA-deficient cells. Nature *535*, 382–387.

Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J.P. (2006). GenePattern 2.0 [2]. Nat. Genet. *38*, 500–501.

Río, P., Segovia, J.C., Hanenberg, H., Casado, J.A., Martínez, J., Göttsche, K., Cheng, N.C., Van de Vrugt, H.J., Arwert, F., Joenje, H., et al. (2002). In vitro phenotypic correction of hematopoietic progenitors from Fanconi anemia group A knockout mice. Blood *100*, 2032–2039.

Risso, D., Ngai, J., Speed, T.P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. Nat. Biotechnol. *32*, 896–902.

Robbiani, D.F., and Nussenzweig, M.C. (2013). Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. Annu. Rev. Pathol. *8*, 79–103.

Robbiani, D.F., Bothmer, A., Callen, E., Reina-San-Martin, B., Dorsett, Y., Difilippantonio, S., Bolland, D.J., Chen, H.T., Corcoran, A.E., Nussenzweig, A., et al. (2008). AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell *135*, 1028–1038.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Roerink, S.F., Schendel, R. Van, and Tijsterman, M. (2014). Polymerase theta-mediated end joining of replication-associated DNA breaks in C . elegans. 954–962.

Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S., and Bonner, W.M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J. Biol. Chem. *273*, 5858–5868.

Roloff, G.W., Drazer, M.W., and Godley, L.A. (2021). Inherited Susceptibility to Hematopoietic Malignancies in the Era of Precision Oncology. JCO Precis. Oncol. 107–122.

Rose, M., Burgess, J.T., O'Byrne, K., Richard, D.J., and Bolderson, E. (2020). PARP Inhibitors: Clinical Relevance, Mechanisms of Action and Tumor Resistance. Front. Cell Dev. Biol. *8*.

Rothenberg, E. V. (2000). Stepwise specification of lymphocyte developmental lineages. Curr. Opin. Genet. Dev. *10*, 370–379.

Rothenberg, E. V. (2019). Programming for T-lymphocyte fates: Modularity and mechanisms. Genes Dev. *33*, 1117–1135.

Rowley, J.D. (1973). Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining.

Nature *243*, 290–293.

Rudd, M.F., Sellick, G.S., Webb, E.L., Catovsky, D., and Houlston, R.S. (2006). Variants in the ATM-BRCA2-CHEK2 axis predispose to chronic lymphocytic leukemia. Blood *108*, 638–644.

Saiki, R., Momozawa, Y., Nannya, Y., Nakagawa, M.M., Ochi, Y., Yoshizato, T., Terao, C., Kuroda, Y., Shiraishi, Y., Chiba, K., et al. (2021). Combined landscape of single-nucleotide variants and copy number alterations in clonal hematopoiesis. Nat Med *27*, 1239–1249.

Salesse, S., and Verfaillie, C.M. (2002). BCR/ABL: from molecular mechanisms of leukemia induction to treatment of chronic myelogenous leukemia. Oncogene *21*, 8547–8559.

Samra, B., Jabbour, E., Ravandi, F., Kantarjian, H., and Short, N.J. (2020). Evolving therapy of adult acute lymphoblastic leukemia: State-of-the-art treatment and future directions. J. Hematol. Oncol. *13*, 1–17.

Savola, P., Lundgren, S., Keränen, M.A.I., Almusa, H., Ellonen, P., Leirisalo-Repo, M., Kelkka, T., and Mustjoki, S. (2018). Clonal hematopoiesis in patients with rheumatoid arthritis. Blood Cancer J *8*, 69.

Sawyer, S.L., Tian, L., Kähkönen, M., Schwartzentruber, J., Kircher, M., University of Washington Centre for Mendelian Genomics, FORGE Canada Consortium, Majewski, J., Dyment, D.A., Innes, A.M., et al. (2015). Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. Cancer Discov. *5*, 135–142.

van Schendel, R., van Heteren, J., Welten, R., and Tijsterman, M. (2016). Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining. PLoS Genet. *12*, e1006368.

Schlacher, K., Christ, N., Siaud, N., Egashira, A., Wu, H., and Jasin, M. (2011). Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. Cell *145*, 529–542.

Schwab, R.A., Nieminuszczy, J., Shin-ya, K., and Niedzwiedz, W. (2013). FANCJ couples replication past natural fork barriers with maintenance of chromatin structure. J. Cell Biol. *201*, 33–48.

Scott, S.P., Bendix, R., Chen, P., Clark, R., Dork, T., and Lavin, M.F. (2002). Missense mutations but not allelic variants alter the function of ATM by dominant interference in patients with breast cancer. Proc. Natl. Acad. Sci. U. S. A. *99*, 925–930.

Scully, R., and Livingston, D.M. (2000). In search of the tumour-suppressor functions of BRCA1 and BRCA2. Nature *408*, 429–432.

Scully, R., Panday, A., Elango, R., and Willis, N.A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. Nat. Rev. Mol. Cell Biol. *20*, 698–714.

Sedic, M., and Kuperwasser, C. (2016). BRCA1-hapoinsufficiency: Unraveling the molecular and cellular basis for tissue-specific cancer. Cell Cycle *15*, 621–627.

Sedic, M., Skibinski, A., Brown, N., Gallardo, M., Mulligan, P., Martinez, P., Keller, P.J., Glover, E., Richardson, A.L., Cowan, J., et al. (2015). Haploinsufficiency for BRCA1 leads to cell-type-specific genomic instability and premature senescence. Nat. Commun. *6*, 7505.

Seppälä, E.H., Ikonen, T., Mononen, N., Autio, V., Rökman, A., Matikainen, M.P., Tammela, T.L., and Schleutker, J. (2003). CHEK2 variants associate with hereditary prostate cancer. Br J Cancer *89*, 1966–1970.

Shah, D.K., and Zuniga-Pflucker, J.C. (2014). An Overview of the Intrathymic Intricacies of T Cell Development. J. Immunol. *192*, 4017–4023.

Shao, Z., Davis, A.J., Fattah, K.R., So, S., Sun, J., Lee, K.-J., Harrison, L., Yang, J., and Chen, D.J. (2012). Persistently bound Ku at DNA ends attenuates DNA end resection and homologous recombination. DNA Repair (Amst). *11*, 310–316.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). Ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics *15*, 284.

Silver, A.J., Bick, A.G., and Savona, M.R. (2021). Germline risk of clonal haematopoiesis. Nat Rev Genet *22*, 603–617.

Singh, S., Jakubison, B., and Keller, J.R. (2020). Protection of hematopoietic stem cells from stress-induced exhaustion and aging. Curr. Opin. Hematol. *27*, 225–231.

Singhal, D., Hahn, C.N., Feurstein, S., Wee, L.Y.A., Moma, L., Kutyna, M.M., Chhetri, R., Eshraghi, L., Schreiber, A.W., Feng, J., et al. (2021). Targeted gene panels identify a high frequency of pathogenic germline variants in patients diagnosed with a hematological malignancy and at least one other independent cancer. Leukemia.

Smith, J., Mun Tho, L., Xu, N., and A. Gillespie, D. (2010). The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer (Elsevier Inc.).
So, A., Le Guen, T., Lopez, B.S., and Guirouilh-Barbat, J. (2017). Genomic rearrangements induced by unscheduled DNA double strand breaks in somatic mammalian cells. FEBS J. *284*, 2324–2344.

Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P.,

and Ebert, B.L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood *126*, 9–16.

Steinke, F.C., Yu, S., Zhou, X., He, B., Yang, W., Zhou, B., Kawamoto, H., Zhu, J., Tan, K., and Xue, H.H. (2014). TCF-1 and LEF-1 act upstream of Th-POK to promote the CD4 + T cell fate and interact with Runx3 to silence Cd4 in CD8 + T cells. Nat. Immunol. *15*, 646–656.

Stok, C., Kok, Y.P., van den Tempel, N., and van Vugt, M.A.T.M. (2021). Shaping the BRCAness mutational landscape by alternative double-strand break repair, replication stress and mitotic aberrancies. Nucleic Acids Res. *49*, 4239–4257.

Stolarova, L., Kleiblova, P., Janatova, M., Soukupova, J., Zemankova, P., Macurek, L., and Kleibl, Z. (2020). CHEK2 Germline Variants in Cancer Predisposition: Stalemate Rather than Checkmate. Cells *9*.

Stracker, T.H., Usui, T., and Petrini, J.H.J. (2009). Taking the time to make important decisions: the checkpoint effector kinases Chk1 and Chk2 and the DNA damage response. DNA Repair (Amst). *8*, 1047–1054.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. *102*, 15545–15550.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med *12*, e1001779.

Sulli, G., Di Micco, R., and di Fagagna, F. d'Adda (2012). Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer. Nat Rev Cancer *12*, 709–720.

Sutcliffe, E.G., Stettner, A.R., Miller, S.A., Solomon, S.R., Marshall, M.L., Roberts, M.E., Susswein, L.R., Arvai, K.J., Klein, R.T., Murphy, P.D., et al. (2020). Differences in cancer prevalence among CHEK2 carriers identified via multi-gene panel testing. Cancer Genet *246–247*, 12–17.

Swift, M., Chase, C.L., and Morrell, D. (1990). Cancer predisposition of ataxia-telangiectasia heterozygotes. Cancer Genet. Cytogenet. *46*, 21–27.

Syljuåsen, R.G., Sørensen, C.S., Hansen, L.T., Fugger, K., Lundin, C., Johansson, F., Helleday, T., Sehested, M., Lukas, J., and Bartek, J. (2005). Inhibition of human Chk1 causes increased initiation of DNA replication, phosphorylation of ATR targets, and DNA breakage. Mol. Cell. Biol. *25*, 3553–3562.

Teachey, D.T., and O'Connor, D. (2020). How i treat newly diagnosed T-cell acute lymphoblastic leukemia and T-cell lymphoblastic lymphoma in children. Blood *135*, 159–166.

Telli, M.L., Timms, K.M., Reid, J., Hennessy, B., Mills, G.B., Jensen, K.C., Szallasi, Z., Barry, W.T., Winer, E.P., Tung, N.M., et al. (2016). Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. Clin. Cancer Res. *22*, 3764–3773.

Teng, G., Maman, Y., Resch, W., Kim, M., Yamane, A., Qian, J., Kieffer-Kwon, K.R., Mandal, M., Ji, Y., Meffre, E., et al. (2015). RAG Represents a Widespread Threat to the Lymphocyte Genome. Cell *162*, 751–765.

Thangavel, S., Berti, M., Levikova, M., Pinto, C., Gomathinayagam, S., Vujanovic, M., Zellweger, R., Moore, H., Lee, E.H., Hendrickson, E.A., et al. (2015). DNA2 drives processing and restart of reversed replication forks in human cells. J. Cell Biol. *208*, 545–562.

Thompson, D., Easton, D.F., and Breast Cancer Linkage Consortium (2002). Cancer Incidence in BRCA1 mutation carriers. J. Natl. Cancer Inst. *94*, 1358–1365.

Timakhov, R.A., Tan, Y., Rao, M., Liu, Z., Altomare, D.A., Pei, J., Wiest, D.L., Favorova, O.O., Knepper, J.E., and Testa, J.R. (2009). Recurrent chromosomal rearrangements implicate oncogenes contributing to T-cell lymphomagenesis in Lck-MyrAkt2 transgenic mice. Genes Chromosom. Cancer *48*, 786–794.

Toledo, L.I., Altmeyer, M., Rask, M.-B., Lukas, C., Larsen, D.H., Povlsen, L.K., Bekker-Jensen, S., Mailand, N., Bartek, J., and Lukas, J. (2013). ATR prohibits replication catastrophe by preventing global exhaustion of RPA. Cell *155*, 1088–1103.

Topatana, W., Juengpanich, S., Li, S., Cao, J., Hu, J., Lee, J., Suliyanto, K., Ma, D., Zhang, B., Chen, M., et al. (2020). Advances in synthetic lethality for cancer therapy: Cellular mechanism and clinical translation. J. Hematol. Oncol. *13*, 1–22.

Tseng, Y.Y., Moriarity, B.S., Gong, W., Akiyama, R., Tiwari, A., Kawakami, H., Ronning, P., Reuland, B., Guenther, K., Beadnell, T.C., et al. (2014). PVT1 dependence in cancer with MYC copy-number increase. Nature *512*, 82–86.

Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell *168*, 644–656.

Tubbs, A., Sridharan, S., van Wietmarschen, N., Maman, Y., Callen, E., Stanlie, A., Wu, W., Wu, X., Day, A., Wong, N., et al. (2018). Dual Roles of Poly(dA:dT) Tracts in Replication Initiation and Fork Collapse. Cell *174*, 1127-1142.e19.

Vaandrager, J.W., Schuuring, E., Philippo, K., and Kluin, P.M. (2000). V(D)J recombinase-mediated transposition of the BCL2 gene to the IGH locus in follicular lymphoma. Blood *96*, 1947–1952.

Varley, J.M. (2003). Germline TP53 mutations and Li-Fraumeni syndrome. Hum. Mutat. *21*, 313–320.

Vasanthakumar, A., Arnovitz, S., Marquez, R., Lepore, J., Rafidi, G., Asom, A., Weatherly, M., Davis, E.M., Neistadt, B., Duszynski, R., et al. (2015). Brca1 deficiency causes bone marrow failure and spontaneous hematologic malignancies in mice. Blood *127*, blood-2015-03-635599.

van der Veeken, J., Glasner, A., Zhong, Y., Hu, W., Wang, Z.-M.M., Bou-Puerto, R., Charbonnier, L.-M.M., Chatila, T.A., Leslie, C.S., and Rudensky, A.Y. (2020). The Transcription Factor Foxp3 Shapes Regulatory T Cell Identity by Tuning the Activity of trans-Acting Intermediaries. Immunity *53*, 971-984.e5.

Vermeer, M.H., Van Doorn, R., Dijkman, R., Mao, X., Whittaker, S., Van Voorst Vader, P.C., Gerritsen, M.J.P., Geerts, M.L., Gellrich, S., Söderberg, O., et al. (2008). Novel and highly recurrent chromosomal alterations in Sézary syndrome. Cancer Res. *68*, 2689–2698.

Vindigni, A., and Hickson, I.D. (2009). RecQ helicases: multiple structures for multiple functions? HFSP J. *3*, 153–164.

Van Vlierberghe, P., Pieters, R., Beverloo, H.B., and Meijerink, J.P.P.P. (2008). Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. Br. J. Haematol. *143*, 153–168.

Walter, D., Lier, A., Geiselhart, A., Thalheimer, F.B., Huntscha, S., Sobotta, M.C., Moehrle, B., Brocks, D., Bayindir, I., Kaschutnig, P., et al. (2015). Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. Nature *520*, 549–552.

Walton, K.D., Wagner, K.U., Rucker, E.B., Shillingford, J.M., Miyoshi, K., and Hennighausen, L. (2001). Conditional deletion of the bcl-x gene from mouse mammary epithelium results in accelerated apoptosis during involution but does not compromise cell function during lactation. Mech. Dev. *109*, 281–293.

Wang, C., and Li, J. (2021). Haematologic toxicities with PARP inhibitors in cancer patients: an up-to-date meta-analysis of 29 randomized controlled trials. J. Clin. Pharm. Ther. *46*, 571–584.

Wang, Y., Dai, B., and Ye, D. (2015). CHEK2 mutation and risk of prostate cancer: a systematic review and meta-analysis. Int. J. Clin. Exp. Med. *8*, 15708–15715.

Wang, Z., Song, Y., Li, S., Kurian, S., Xiang, R., Chiba, T., and Wu, X. (2019). DNA polymerase (POLQ) is important for repair of DNA double-strand breaks caused by fork collapse. J. Biol. Chem. *294*, 3909–3919.

Weber, B.N., Chi, A.W.S., Chavez, A., Yashiro-Ohtani, Y., Yang, Q., Shestova, O., and Bhandoola, A. (2011). A critical role for TCF-1 in T-lineage specification and differentiation. Nature *476*, 63–69.

Weinreb, J.T., Ghazale, N., Pradhan, K., Gupta, V., Potts, K.S., Tricomi, B., Daniels, N.J., Padgett, R.A., De Oliveira, S., Verma, A., et al. (2021). Excessive R-loops trigger an inflammatory cascade leading to increased HSPC production. Dev. Cell *56*, 627-640.e5.

Wong, T.N., Miller, C.A., Jotte, M.R.M.M., Bagegni, N., Baty, J.D., Schmidt, A.P., Cashen, A.F., Duncavage, E.J., Helton, N.M., Fiala, M., et al. (2018). Cellular stressors contribute to the expansion of hematopoietic clones of varying leukemic potential. Nat. Commun. *9*, 1–10.

Wood, R.D., and Doublié, S. (2016). DNA polymerase θ (POLQ), double-strand break repair, and cancer. DNA Repair (Amst). *44*, 22–32.

Wu, X., Webster, S.R., and Chen, J. (2001). Characterization of tumor-associated Chk2 mutations. J Biol Chem *276*, 2971–2974.

Xing, S., Li, F., Zeng, Z., Zhao, Y., Yu, S., Shan, Q., Li, Y., Phillips, F.C., Maina, P.K., Qi, H.H., et al. (2016). Tcf1 and Lef1 transcription factors establish CD8+ T cell identity through intrinsic HDAC activity. Nat. Immunol. *17*, 695–703.

Yang, F., Long, N., Anekpuritanang, T., Bottomly, D., Savage, J.C., Lee, T., Solis-Ruiz, J.M., Borate, U., Wilmot, B., Tognon, C.E., et al. (2021). Identification and prioritization of myeloid malignancy germline variants in a large cohort of adult AML patients. Blood.

Yousefzadeh, M.J., and Wood, R.D. (2013). Mini review DNA polymerase POLQ and cellular defense against DNA damage. DNA Repair (Amst). *12*, 1–9.

Yousefzadeh, M.J., Wyatt, D.W., Takata, K. ichi, Mu, Y., Hensley, S.C., Bylund, O., Doublie, S., Tomida, J., Mcbride, K.M., Wood, R.D., et al. (2014). Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ. PLoS Genet. *10*.

Yu, T.W., and Anderson, D. (1997). Reactive oxygen species-induced DNA damage and its modification: a chemical investigation. Mutat. Res. *379*, 201–210.

Zatreanu, D., Robinson, H.M.R., Alkhatib, O., Boursier, M., Finch, H., Geo, L., Grande, D., Grinkevich, V., Heald, R.A., Langdon, S., et al. (2021). Polθ inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance. Nat. Commun. *12*, 3636.

Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. Nat. Cell Biol. *16*, 2–9.

Zhang, J., Willers, H., Feng, Z., Ghosh, J.C., Kim, S., Weaver, D.T., Chung, J.H., Powell, S.N., and Xia, F. (2004). Chk2 phosphorylation of BRCA1 regulates DNA double-strand break repair. Mol. Cell. Biol. *24*, 708–718.

Zhao, X., Shan, Q., and Xue, H.-H.H. (2021). TCF1 in T cell immunity: a broadened frontier. Nat. Rev. Immunol. *13*.

Zhou, J., Gelot, C., Pantelidou, C., Li, A., Yücel, H., Davis, R.E., Farkkila, A., Kochupurakkal, B., Syed, A., Shapiro, G.I., et al. (2021). A first-in-class Polymerase Theta Inhibitor selectively targets Homologous-Recombination-Deficient Tumors. Nat. Cancer *2*, 598–610.

Zhunussova, G., Afonin, G., Abdikerim, S., Jumanov, A., Perfilyeva, A., Kaidarova, D., and Djansugurova, L. (2019). Mutation Spectrum of Cancer-Associated Genes in Patients With Early Onset of Colorectal Cancer. Front Oncol *9*, 673.