SIM2S INHIBITS BREAST CANCER PROGRESSION BY MAINTAINING GENOMIC STABILITY AND PREVENTING AN EPITHELIAL MESENCHYMAL TRANSITION

A Dissertation

by

SCOTT JAMES PEARSON

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Weston W. Porter
Committee Members,	Scott V. Dindot
	Stephen H. Safe
	David W. Threadgill
Intercollegiate Faculty Chair,	David W. Threadgill

December 2019

Major Subject: Genetics

ABSTRACT

Breast cancer is the most common cancer in women, affecting approximately one in eight women during their lifetime. There is increasing evidence that genomic instability is a precursor to breast cancer progression from initial hyperplasia to invasive cancer. Previous studies have demonstrated that individuals who carry mutations in the critical DNA damage repair proteins, BRCA1 and BRCA2, are more likely to develop high-grade ductal carcinoma in situ (DCIS), which is characterized a neoplastic mammary lesion that is confined to the ductal-lobular system of the breast. In addition to the higher occurrences of DCIS in these individuals, they are more likely to progress from DCIS to the more malignant invasive ductal carcinoma (IDC). Here we identify SIM2s (Singleminded-2s; a member of the bHLH/PAS family of transcription factors) as a novel member of the homologous recombination (HR) pathway of DNA damage repair. In this study we show that SIM2s is stabilized through interaction with ATM, which leads to phosphorylation of SIM2s in response to dsDNA damage. Once stabilized, SIM2s interacts with members of the HR pathway and enhances RAD51 recruitment to the site of DNA damage. Inhibition of this process through the mutation of *SIM2s* at the ATM recognition site (S115) or loss of *SIM2s* leads to a significant decrease in dsDNA repair via HR and prolonged presence of dsDNA breaks. Moreover, we found that SIM2s is necessary to maintain replication fork stability during replication stress, with loss of *SIM2s* resulting in a significant increase in replication fork collapses, leading to further unresolved dsDNA breaks. Finally, we found that loss or mutation of *SIM2s* leads to an epithelial mesenchymal transition (EMT) that is characterized by a decrease in E-cadherin and induction of the basal marker, K14. In

addition, we observed increased invasion and metastasis using both our DCIS progression model cell line and in xenograft models. Together, these results identify SIM2s as a novel player in the DNA damage repair pathway and suggest that loss of *SIM2s* connects DCIS progression to IDC through increased genomic instability, EMT, and metastasis.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Porter, and my committee members, Dr. Dindott, Dr. Safe, and Dr. Threadgill for their guidance and support throughout the course of this research. In addition, I would like to thank the members of Dr. Porter's lab, Dr. Gile's lab, and Dr. Safe's lab for their support and assistance throughout this process. Dr. Barhoumi and Dr. Burghardt at the Texas A&M University College of Veterinary Medicine & Biomedical Sciences Image Analysis Laboratory, supported by NIH-NCRR (1S10RR22532-01) grant, were of tremendous assistance in the acquisition of the confocal images presented in this work.

I would also like to acknowledge the Histology Core Facility at Texas A&M University College of Veterinary Medicine & Biomedical Sciences for tissue preparation and H&E staining and Dr. Behbod at the University of Kansas Medical Center for providing the human samples, tissue microarray, and pathology information. We acknowledge support from the University of Kansas (KU) Cancer Center's Biospecimen Repository Core Facility staff for helping obtain human specimens. The authors also acknowledge support from the KU Cancer Center's Cancer Center Support Grant (P30 CA168524).

Finally, I would like to thank my family and friends through their continued support, encouragement and love throughout this process.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Weston Porter, Scott Dindot of Veterinary Pathobiology, Stephen Safe of Veterinary Physiology and Pharmacology, and David Threadgill of Molecular and Cellular Medicine. The human tissue microarrays were collected in collaboration with the University of Kansas and analyzed by Dr. Kelly Scribner. All other work conducted for the thesis (or) dissertation was completed by the student independently.

Funding Sources

This work and graduate study was made possible in part by the National Cancer Institute through R21CA190941, R01HD083952, and R21CA185460. In addition, further funding was provided by the Department of Defense (DOD-CDMRP) W81XWH-11-1-0158. In addition, fellowship funding was provided by the Texas A&M University College of Veterinary Medicine.

NOMENCLATURE

53BP1	p53 Binding Protein
a-NHEJ	Alternative non-homologous end-joining
APE1	Apurinic/apyrimidinic endonuclease
AHR	Aryl-hydrocarbon receptor
ARNT	Aryl hydrocarbon receptor nuclear translocator
ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia and Rad3-related
BC	RAD51B-RAD51C complex
BCDX2	RAD51B-RAD51C-RAD51D-XRCC2 complex
bHLH	Basic-helix-loop-helix
bp	Base pairs
BRCA1	Breast Cancer Associated 1
BRCA2	Breast Cancer Associated 2
CDK	Cyclin dependent kinase
CHK1	Checkpoint kinase 1
CHK2	Checkpoint kinase 2
CME	Central midline element
c-NHEJ	Classical non-homologous end-joining
CNS	Central nervous system
CRY1	Cryptochrome 1
CRY2	Cryptochrome 2

CSA	Cockayne syndrome type A
CSB	Cockayne syndrome type B
CtIP	CTBP interacting protein
CX3	RAD51C-XRCC3 complex
DCIS	Ductal carcinoma in situ
DDR	DNA-damage repair
DRE	Dioxin response element
DS	Down Syndrome
DS	Down's syndrome
DSB	Double-stranded break
DSCR	Down's syndrome critical region
DX2	RAD51D-XRCC2 complex
EDHB17	Estradiol-17
EGFR	Epithelial growth factor receptor
EMT	Epithelial mesenchymal transition
ER	Estrogen receptor
ERCC1	Excision repair cross-complementation group 1
EXO1	Exonuclease 1
EZH2	Enhancer of zeste homolog 2
FEN1	Flap endonuclease 1
FIH1	Factor inhibiting HIF1α
HIF	HIF1α and HIF2α
HIF1α	Hypoxia-inducible factor 1α

HIF2α	Hypoxia-inducible factor 2α
HJ	Holliday Junction
HOX2	Homeobox 2
HR	Homologous Recombination
HRE	Hypoxia response element
HU	Hydroxyurea
ICL	Inter-strand crosslink
IDC	Invasive ductal carcinoma
Ku	The Ku70/Ku80 heterodimer
LIG1	DNA ligase 1
LIG3	DNA ligase 3
MMEJ	Microhomology-mediated end-joining
MMP	Matrix metalloprotease
MRI	Magnetic resonance imaging
MRN	MRE11-RAD50-NBS1 complex
NBS1	Nijmegen breakage syndrome protein 1
NF-κB	Nuclear factor κ-light-chain-enhancer of activated B cells
NHEJ	Non-homologous end-joining
NLS	Nuclear localization signal
NPAS1-4	neuronal PAS domain protein 1-4
ODD	Oxygen-dependent degradation
pADPr	Poly (ADP-ribose)
PALB2	Partner and localizer of BRCA2

PARG	Poly (ADP-ribose) glycohydrolase
PARP	Poly (ADP-ribose) polymerase
PARPi	PARP inhibitor
PAS	"PER, ARNT, SIM"
PAS	PER-ARNT-SIM
PCNA	Proliferating cell nuclear antigen
PER	Period
PHD	Prolyl hydroxylase
PIN1	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1
Pol	Polymerase
PR	Progesterone receptor
PRKN	Parkin
RAR	Retinoic acid receptor
RNR	Ribonucleotide reductase
ROS	Reactive oxygen species
RP	Rad51 paralog
RPA	Replication protein A
RPC	Replication factor C
SIM	Single-minded
SIM2	Single-minded 2
SMARCAD1	SWI/SNF-Related, Matrix-Associated Actin-Dependent Regulator Of Chromatin, Subfamily A, Containing DEAD/H Box 1
SNP	Single nucleotide polymorphism

SOD	Superoxide dismutase
SSA	Single-strand annealing
SSB	Single-stranded break
SWS1	Short-wavelength sensitive opsin
SWSAP1	SWS1-Associated Protein
TFIIH	Transcription factor IIH
TIRF	Total internal reflection fluorescence
TNBC	Triple-negative breast cancer
TOPIII	Topoisomerase 3
VBC/Cul-2	VHL, elongin B, elongin C, Cullin-2, and Rbx1
VHL	von Hippel-Lindau
WHO	World Health Organization
wt	Wild type
XAB2	XPA binding protein 2
XPA	Xeroderma pigmentosum complementation group A
XPB	Xeroderma pigmentosum complementation group B
XPD	Xeroderma pigmentosum complementation group D
XPF	Xeroderma pigmentosum complementation group F
XPG	Xeroderma pigmentosum complementation group G
XRCC1	X-ray repair cross-complementing protein 1

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Breast Cancer

History of Breast Cancer

In 2019 there are expected to be 891,480 new cases of cancer in women in the United States of America. Of these, 268,600 are predicted to be breast cancer; making it the most common form of cancer amongst woman. In fact, breast cancer is expected to affect 1 in 8 women during their lifetime. However, due increased screenings and early detection, as well as a better understanding of the underlying mechanisms involved in disease progression the 5-year survival rate has increased from 75.31% in 1975 to 90.90% in 2010 (https://seer.cancer.gov/statfacts/html/breast.html).

In addition to increased survival, there has also been an increase in patient quality of life as technology has improved. The first recorded case of breast cancer can be traced back to the Edwin Smith Papyrus, a texting dating to 1600BC (although suspected to be a copy of an early text from 2500-3000BC) Egypt outlining 48 medical cases and their suggested treatments[2, 3]. Here, case 45 describes "bulging tumors on the breast", where the author writes "There is no treatment[3]."

Further, advancements in the classification of the disease were made by Hippocrates and Galen around 400BC[2]. However, with the onset of the Middle Ages and a prohibition on autopsies, Hippocrates' theory stating the imbalance of the 4 humors (blood, phlegm, yellow bile, and black bile) being the root cause of cancer persisted for over 1300 years[4]. It wasn't until the 10th century when physicians like Ibn Sina and Abu Al-Qasium

Al-Zaharawi (Albucasis) from the Islamic empire started reviewing ancient Greek text and continued progress into understanding breast cancer[5, 6].

Progress in our understanding of breast cancer continued to develop throughout the next few centuries as our comprehension of surgery advanced. However, the next great leap forward was not until 1838, when Johannes Müller showed that cancer was made up of cells. His student, Rudolph Virchow, would later go on to show that these cells were derived from other, normal cells[4]. During this same time, the advent of anesthesia allowed for more aggressive tumor resections. In 1882, William Banks reported on the necessity for the removal of axillary lymph nodes in order to eradicate invasive tumors[7]. A little over 10 years later, in 1894, William S. Halstead established the predominant standard of care for the next 80 years when he recommended radical mastectomy to remove the *pectoralis major* intact so as not to fragment the tumor upon removal[8].

The end of the 20th century, with a better understanding of the molecular mechanisms underlying breast cancer including the hormone mediated response of some tumors, saw a trend away from radical mastectomy, and towards lumpectomies with adjuvant treatments[6, 9, 10]. Since this time numerous factors have been discovered that aid in tumor diagnosis and classification, which supports better treatment and management of newly discovered tumors.

Tumor Classification

With such a long history on the treatment of breast cancer, it should come as no surprise that numerous classification systems have arisen for the disease. Initially these classification systems were based on subjective observations by the physician. However,

over time the classification systems used to diagnosis tumors have become more objective, allowing for strong correlations to be made between tumor grade and patient outcomes[11].

Initial clinical detection and identification of breast malignancies are done through a variety of imaging techniques. Although mammography has become the gold standard over the years, it can be an ineffective technique in women under the age of 40 or for woman with mammographically dense breast. The most robust method for detection in women of all ages is ultrasound. This technique also allows for the detection of disease within the axillary. Finally, magnetic resonance imaging (MRI) can be used to enhance the resolution of the extent of disease. In combination these techniques predominantly identify invasive carcinomas with stellate and circular structures lacking calcifications[11].

Further grading can be accomplished through biopsies or during tumor resection through histological analysis. Here severity of the disease is broken down into categories assessing tubule and gland formation, nuclear pleomorphisms, and mitotic counts. Each category is ranked on a scale from one to three, with three being the most severe. The first category (tubule and gland formation) relates to the relative morphology of the tissue and scores it based on the presence of normal ductal formations. The second category (nuclear pleomorphisms) looks at a small section of the tissue containing the most nuclear pleomorphisms and assesses the regularity of nuclear shape and size, with lower scores having uniform nuclei that are similar in size to neighboring epithelial cells. The last category (mitotic counts) enumerates the number of dividing cells on the leading edge of the tumor. Finally, the scores from each of these categories are combined and higher scores are associated with more severe malignancies with poor differentiation and meager prognosis[11].

Together the information gathered from tissue grading, as well as the origin of the tumor allows for the tumor to be classified into 1 of 88 different tumor classifications. Broadly, these are simplified into 8 categories, loosely based on the cell type of the initial lesion[11]. However diverse, initial tumor formation follows a similar pattern. The genesis of a new tumor starts with a hyperplasia, or an abnormal dividing of cells while still maintaining the general shape and order of the cell type. Eventually as these cells divide, a cell might gain additional mutations that progresses the tumor into a dysplasia. At this point, depending on the severity of the dysplasia, cells start to lose their defined shape and start to regress towards a more stem-like appearance. From here, epithelial tumors can form a carcinoma *in situ,* with one of the most defined being ductal carcinoma *in situ* (DCIS; Figure 1)[11].

DCIS represents a heterogeneous group of diseases characterized by a neoplastic mammary lesion that is confined to the ductal-lobular system of the breast and is a non-obligate pre-cursor to invasive ductal carcinoma (IDC)[12]. This heterogeneity within the disease poses an interesting clinical challenge in accurately grading the tumor, as tumor grade varies drastically even within a small area of the duct[13-15]. This has been shown to

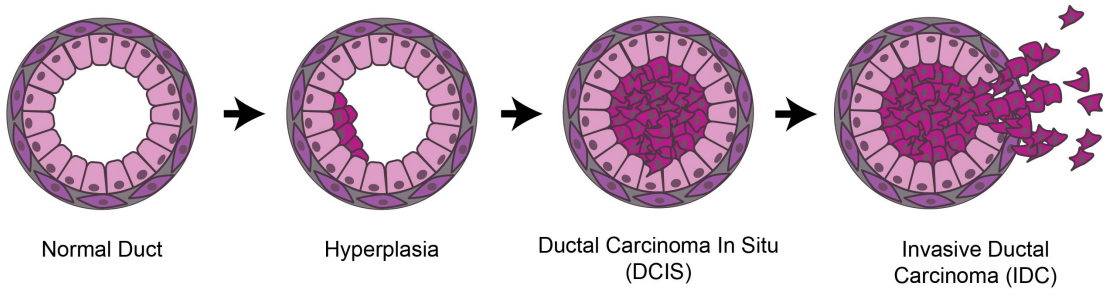


Figure 1 Tumor progression model.

result in biopsies not accurately representing the tumor as a whole[14]. Historically DCIS has been categorized by elevated levels of calcification and necrosis, with increasing levels of each corresponding to a more severe disease state[16]. However, there have also been contradictory studies arguing the significance of each of these methods of tumor evaluation. Fisher et al. found that tumor margin and the presence of comedo necrosis was sufficient to predict a DCIS to IDC transition[17]. However, previously, Solin et al. argued that tumor margin was dispensable, and risk of IDC could be determined through the presence of necrosis and nuclear grade[18]. In contrast, Sneige et al. suggested that risk of progression could be determined through the presence of necrosis and fibrosis while Silverstein et al. found nuclear grade, necrosis, tumor size, and margin width to be necessary to accurately categorize pre-invasive tumors[19, 20].

Clearly, there has been a historic issue with respect to the categorization of DCIS, and thus the potential of determining which DCIS will progress into an IDC. DCIS is afforded only one level of classification by the World Health Organization (WHO), even though the scope of DCIS severity spans from a small ductal hyperplasia to the complete obstruction of the ductal structure by a high-level dysplasia with extensive necroses and calcification. Interestingly, the problems in the classification of DCIS were reported nearly a decade ago[21]. In a study by Clark et al., the viability of molecular testing and classification of DCIS, a technique that had already been demonstrated to be useful for the categorization of IDC ten years prior, was assessed[21-23]. They were able to demonstrate that using 16 biomarkers, they could categorize DCIS samples into luminal A, luminal B, $Her2^{+}$, and triple-negative breast cancer (TNBC) subgroups; subgroups that were already

used to describe IDC. This allowed for the better overall analysis of tumorigenic potential and DCIS severity (Figure 2)[21].

With advancements in molecular technologies, the expression of just a few factors, including human epidermal growth factor receptor 2 (HER2), p53, and estrogen receptor (ER), were used to gauge early disease severity[24]. However, this classification does not necessarily predict which DCIS will progress into IDC. Two studies conducted in 2010 reached similar conclusions, showing that while it is possible to differentiate between early and late-stage DCIS, as well as categorize the tumor subtype, molecular testing of known tumorigenic factors is inadequate to differentiate between invasive and non-invasive forms of the disease[25, 26]. More recently, a microarray study of 160 samples (59 DCIS, 85 IDC, and 16 mixed) again concluded that molecular analysis was sufficient to categorize disease subtype, but was inadequate at differentiating between late DCIS and IDC[27]. However, their technique of first grouping samples by subtype and then looking at diagnosis to differentiate gene expression profiles between DCIS and IDC yielded some success for some subtypes, making this the closest study to accurately predict which DCIS will progress into IDC[27]. Yet a more recent single-cell analysis conducted in 2018, was unable to differentiate between DCIS and IDC[28]. This later study was able to demonstrate that disease heterogeneity persists through progression from DCIS to IDC, adding another layer of complexity for treating the disease.

Together, these studies demonstrated the complexity of the disease, and why it has proven to be a struggle to predict DCIS outcome. With no adequate solution to differentiate between the two disease states, clinicians must determine if monitoring the

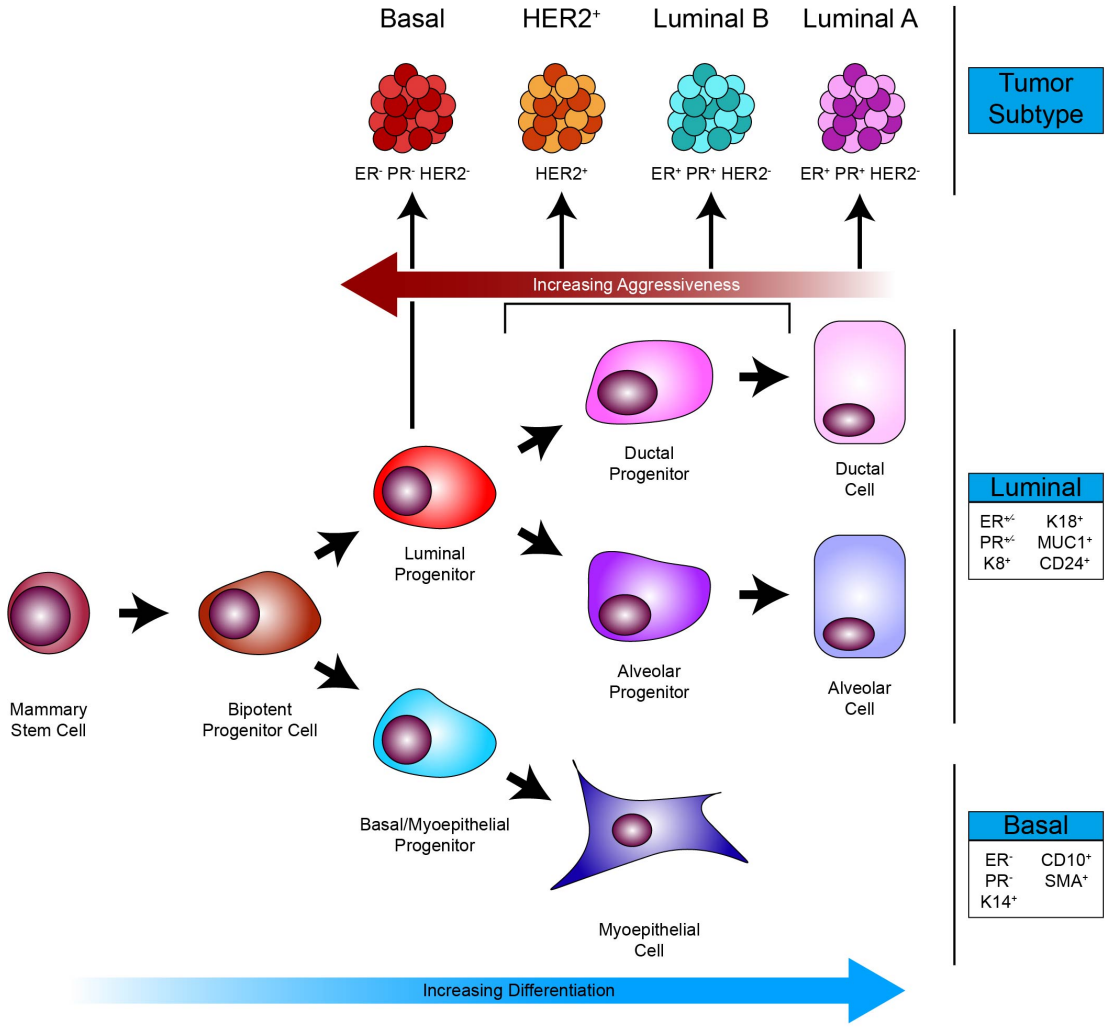


Figure 2 Mammary gland development and tumor subtype.

DCIS is sufficient or if radical therapy is warranted. In recent decades, this limitation has become more pronounced with increased use of screening mammography. The widespread adoption of this screening technology by 1975 had reduced patient morbidity by as much as 70%[29]. However, it also resulted in a sharp increase in the diagnosis of DCIS, which is easily identified by this technology. Specifically the diagnosis of DCIS rose from less than 1% of breast cancers to 15-25% of all diagnosed cases, accounting for over 54,944 new cases in 2013 in the United States of America[12, 30, 31]. However, it has been reported that as few as 20% of diagnosed DCIS will progress to IDC, leading to the potential for gross overtreatment and reduced quality of life for a disease that may never affect the individual[32, 33]. As screening technologies continue to improve, it is likely that a higher percentage of new breast cancer cases will be DCIS, as retrospective MRI analysis of mammographies have shown that as much as 48% of DCIS cases are missed with mammography alone[34].

Currently, treatment for DCIS does not drastically differ from that of IDC, with neoadjuvant chemotherapy followed by radical mastectomy still being used in many regions[35]. However, there has been a growing trend in more developed regions of the world preforming simple excision of the primary DCIS lesion followed by adjuvant radiotherapy, as an increasing numbers of studies have demonstrated similar patient outcomes with these less invasive techniques for the treatment of DCIS[36-39]. Not surprisingly, these breast-saving procedures are better tolerated by patients, both physically and emotionally[40]. Interestingly, there have been contradictory studies regarding the efficacy of the use of radiotherapy in low-risk individuals. Whereas the NSABP-17 trial showed a 40% improvement in patient outcome for those treated with radiotherapy after excision, other

studies have shown no significant change in clinical outcomes[41-45]. One crucial factor that may have been overlooked in many of these studies is the heredity of the patients as three separate studies have shown that excision alone is sufficient to treat low-stage DCIS in individuals with eastern lineages[46-48].

This idea of heredity strongly influencing cancer outcome was not new, with studies conducted in the mid-1990's and early 2000's demonstrating a higher incidence of TNBC in African-American women[49, 50]. Unfortunately, this disparity is also associated with higher-grade tumors at diagnosis and increased patient morbidity, even when adjusted for socioeconomic factors and mistrust of the medical field[50]. Part of this increased morbidity is due to the nature of TNBC. At its basis TNBC is defined as ER and progesterone receptor (PR) negative with aberrant expression of HER2[51]. Clinical, this presents as a highly invasive tumor with poor response to traditional endocrine based chemotherapeutics[52]. Past reports have demonstrated that fewer than 30% of woman have recurrent free survival at 5-years past initial treatment, with almost all ultimately succumbing to the disease[53].

Paradoxically, patients with TNBC tend to respond well to non-endocrine based neoadjuvant and adjuvant therapies[54]. However, the inherent heterogeneity of the disease, and formation of multiple metastatic niches ultimately results in the disease being nearly impossible to treat, with continuous treatment with sequential chemotherapeutic agents being the current standard of care in order to prevent metastatic outgrowth[55]. TNBC is particularly sensitive to platinum salts, a potent class of DNA cross-linking agents, with the efficacy of these treatments suggesting a strong level of genomic instability inherent in TNBC[56, 57]. Indeed, based on current knowledge of the disease, mutations in multiple genes

involved in DNA damage repair (DDR) are concentrated in TNBC[58]. These genes include ataxia telangiectasia mutated (*ATM*), breast cancer associated 1 and 2 (*BRCA1/BRCA2*), *TP53,* and aberrant expression of *RAD51*[51, 59]. This misregulation of DDR elements can also, in part, explain the high incidence of TNBC in African American women. A study in 2012 found that African-American woman have an increased incidence of a pathogenic variant of *EZH2* (enhancer of zeste homologue 2), resulting in increased levels of the protein[60]. Earlier studies showed that EZH2 expression is known to decrease the efficacy of DDR by reducing the expression of RAD51 paralogs (reviewed later), resulting in decreased levels of homologous recombination (HR) based repair[61].

DNA Damage Repair and Cancer

A recurring theme in breast cancer progression is increasing levels of genomic instability. This phenomenon is the result of a cells inability to properly contend with breaks in DNA and other forms of genotoxic stress. It is estimated that ~55,000 DNA single-stranded breaks (SSB) and 10-50 DNA double-stranded breaks (DSB) occur daily in every nucleated cell[62, 63]. These breaks occur due to both endogenous events as well as genotoxic exogenous agents. As mentioned above, TNBC has a higher incidence of genomic instability, which is highlighted by their sensitivity to platinum salts. This class of genotoxic agents is known to activate multiple repair pathways due to the formation of inter-strand crosslinked (ICL) DNA, with one of the earliest being the activation of SSB repair[64].

Single Stranded Break Repair

Even without interference from chemotherapeutic agents, a cell is still subject to a multitude of SSBs. These breaks are one of the driving forces in tumorigenesis, with each

DNA lesion possibly giving rise to a new mutation that could allow the cell to proliferate faster and establish itself in new environments. About $1/5^{th}$ of the nearly 55,000 SSBs that occur can be attributed to spontaneous hydrolysis of nucelotides[65]. This results in the formation of an abasic site, which in addition to lacking the necessary encoding information, is also a potent inhibitor of transcription and replication. The other 80% of breaks that occur are spread out among a variety of processes, such as the spontaneous deamination of cytosine to uracil/thymine can occur, which leads to non-Watson and Crick pairing. Although easily repairable, this reaction has a high likelihood of inserting a permanent mutation into the DNA sequence, if its opposing nucleotide is excised during repair instead of itself[66].

Another significant source of DNA damage within a cell can be attributed to the formation of reactive oxygen species (ROS). These arise from a variety of cellular sources, but are most concentrated around mitochondrial oxidative phosphorylation[67]. Here, reactive superoxide anions are created, which can react with a variety of macromolecules within a cell. Superoxide dismutase (SOD) readily converts superoxide anions into the less reactive H_2O_2[68]. However, upon contact with Fe^{2+}, H_2O_2 is catalyzed to form a highly reactive hydroxide (OH^-) group which readily reacts with DNA and can result in many variety of DNA lesions inlcuding: formation of an oxidized abasic site, 2-deoxyribonolactone, formation of thymine glycol, the addition of a hydroxyl group to the C8-position of guanine, and a SSB[69-79]. Each of these different DNA lesions, presents a different challenge for DDR machinery. As such, cells have developed multiple, highly conserved SSB repair pathways in order to repair the DNA and allow transcription and synthesis to continue.

Nucleotide Excision Repair

Many of the aforementioned forms of DNA damage can be repaired through nucleotide excision repair (NER). Most importantly in regards to cancer treatments, NER is the pathway of SSB repair that is activated in response to ICL, like those created by platinum salts[64]. NER removes bulky DNA lesions through the removal of short oligonucleotides, leaving a gap that must be filled in through repair synthesis[80-84]. Regardless of the mechanism by which the SSB occurred, initial steps in DNA break recognition are the same. Early detection of the break can be attributed to PARP1[85]. From there, PARP1 will activate itself before creating branched chains of poly (ADP-ribose) (pADPr)[86]. These in turn are quickly dissolved by poly (ADP-ribose) glycohydrolase (PARG), leading to continues regulation of PARP1, ensuring that DDR pathways are only activated when needed[87]. pADPr in turn functions as a signal for the recruitment of other factors necessary for SSB repair to bind to and be recruited to the site of DNA damage[88]. From here, depending on what type of DNA damage has occurred, NER diverges from the other two major pathways of SSB repair: base excision repair (BER), and mismatch repair (MMR).

NER can be divided into two sub-pathways depending on how the factors involved in initiating NER: global NER and transcription-coupled NER[89, 90]. As the name suggests transcription-coupled NER is linked to regions of the genome being actively transcribed and with the help of cockayne syndrome type A (CSA), cockayne syndrome type B (CSB), and XPA (xeroderma pigmentosum complementation group A) binding protein 2 (XAB2) is brought to site of DNA damage by RNA polymerase[91]. Conversely, global NER can be initiated anywhere along the genome, and is commonly seen at sites of UV damage. It is

initiated by the XPC-RAD23B complex which is able to recognize thermodynamically destabilized duplexed DNA[91, 92].

Once initiated transcription factor IIH (TFIIH) is recruited to the site of DNA damage and opens the DNA with xeroderma pigmentosum complementation group B (XPB). This allows the xeroderma pigmentosum complementation group D (XPD) subunit of TFIIH to recognize bulky DNA lesions and recruit XPA, replication protein A (RPA), and xeroderma pigmentosum complementation group G (XPG). Excision repair cross-complementation group 1 (ERCC1) and xeroderma pigmentosum complementation group F (XPF) are then recruited to initiate a 5' incision into the lesion allowing for Pol κ, Pol δ, or Pol ε to fill in the gap created by XPB. Once the polymerase comes to the end of the gap, XPG makes a 3' incision and LIG1 or LIG3 ligates the nicks[91].

Generally this DDR pathway is considered to by error-free[93]. However, over the last 40 years there have been many studies demonstrating that under certain situations NER introduces mutations during the repair process[94-100]. Important for the resolution of ICLs created by platinum salts, Sarkar et al. demonstrated that mutagenesis associated with NER is highest in ICL repair, thus greatly adding to the mutagenic burden put on a cell by treatment with platinum salts, potentially leading to cell death as factors that are necessary for cell viability become non-functional[100].

Base Excision Repair

In certain situations, BER is capable of preventing the formation of ICLs, thus negating the need for NER to resolve them[101]. BER handles these through the resolution of precursors to ICL lesions, meaning that once the ICL has formed, BER is no longer capable of repairing the lesion and other pathways are necessary. In normal cellular

functions, BER is one of the most critical pathways for repairing ssDNA lesions[68]. In addition to removing lesions from DNA, it is also the primary mechanism through which SSBs are repaired[102]. After localization of PARP1 to the site of DNA damage, activated PARP1 is able to recruit X-ray repair cross-complementing protein 1 (XRCC1) through its BRCT domain[103]. XRCC1 acts as a scaffolding protein and uses its BRCT domain to recruit other factors that are necessary for DNA repair. In its simplest form, or short-patch repair, a DNA glycosylase is recruited to the damage site and removes damaged bases, leaving behind an abasic site[104, 105]. The abasic site is then removed by the endonuclease apurinic/apyrimidinic endonuclease (APE1), allowing for polymerase (Pol) β to fill in the gap[106]. Finally, DNA ligase 1 (LIG1) or DNA ligase 3 (LIG3) seals the nick, finishing DDR[107, 108].

In some cases, the formation of a 5'-flap of up to 13 nucleotides is created[68]. This is termed long-patch BER. In these instances, replication factor C (RFC) and proliferating cell nuclear antigen (PCNA) are recruited, before the gap is filled by Pol δ or Pol ε. Finally, the flap is removed by flap endonuclease 1 (FEN1) and the nicks are sealed via LIG1. The fate determining mechanisms that dictate whether a DNA lesion will be repaired by short-patch or long-patch BER are not well understood[68]. Although mutations in this pathway do not greatly affect the mutagenic burden placed on a cell by treatment with platinum salts, other factors have been identified that directly target mutations in this pathway, leading to the potential induction of tumor cell death. Many of these chemotherapeutic agents are alkylating drugs, such as temozolomide, N-methyl-N'-nitro-N-nitrosoguanidine, dacarbazine, chlorambucil, nimustine, carmustine, lomustine,

fotemustine, and cyclophosphamide[109]. Failure of BER to repair the adducts formed by these agents can result in the formation of DSBs as replication forks stall at the lesion.

Mismatch Repair

Interestingly, the last major form of SSB repair, sensitizes cells to platinum salts[110]. In 1998 it was found that loss of this pathway provides protection for cells against many chemotherapeutic agents[111]. The next year, it was found that MMR does so by inhibiting RAD52/RAD51-dependent recombinational bypass, which would otherwise allow a cell to repair ICLs through recombination[110]. In addition, MMR has also been implicated in the resolution of cyclobutene pyrimidine dimers formed by UV radiation, and damage caused by chemotherapeutics[112-114]. This potentially provides a therapeutic challenge, in that treatment with platinum salts or other chemotherapeutic agents can select for tumor cells that have deficiencies in MMR, and thus can develop resistance to these treatments.

Mutations in the MMR pathway provide another issue since it is the main pathway for correcting the incorporation of incorrect nucleotides during replication. On its own, DNA polymerases have an error rate of ~10^{-7} per bp[115]. Due to MMR this error rate plummets to 10^{-9}-10^{-10} errors per bp[112]. Thus, mutations in this pathway can provide for a significant uptick in the development of mutations, but not at the same extreme rate that is seen with the treatment of DNA lesion causing chemotherapeutic agents, leading to the ability of a tumor to slowly select for mutations that increase tumorgenicity without having a mutational burden that is incompatible with cell viability.

In eukaryotes, the mechanisms underlying MMR are not completely understood; however, it is known that there is substantial overlap in machinery used in MMR and in other forms of DDR, such as single-strand annealing (SSA) and non-homologous end-

joining (NHEJ)[112]. Briefly, what is known about MMR is that recognition of DNA lesions is initiated through MutSα (for mismatches/lesions up to 3 nucleotides in length; MSH2/6 in humans) or MutSβ (for mismatches/lesions up to 13 nucleotides in length; MSH2/3 in humans)[112, 116-118]. MutSβ is the factor that is responsible for the recognition of ICLs[119]. From here, MutLα (MLH1/PMS2 in humans) is recruited, whose endonuclease activity is thought to be activated by the recruitment of PCNA for incision into the DNA[120-124]. Excision of the nucleotides on one of the DNA strands is then performed by exonuclease 1 (EXO1)[125]. Again, the exact mechanism that determines which strand EXO1 excises is not fully understood in eukaryotes, however, during active replication, lack of DNA methylation on the daughter strand plays a substantial role, and selectively removes nucleotides from the daughter strand, using the parental strand as the template[112]. In addition, the formation of Okazaki fragments during asymmetrical replication provides easy access to 5' entry sites for EXO1, making for preferential excision of the lagging strand[126, 127]. Once a gap is created by EXO1, Pol δ fills in the missing nucleotides and LIG1 seals the nicks that were created[128].

Double-Stranded Break Repair

Although small mutations may be inserted into the genome by SSB repair, these pathways are preferable to the formation of a DSB. DSB are considered to be more highly mutagenic and cytotoxic to the cell and can form when a DNA/RNA polymerase encounters a DNA lesion. This collision will lead to the stalling of the polymerase machinery, and other pathways of DDR to become activated[78, 88, 102]. In some cases, the utilization of specific polymerases allows for the replisome to bypass the lesion, resulting in the incorporation of random nucleotides, and thus the non-faithful inheritance of genetic

material to daughter cells[129]. In other instances for cells undergoing replication, the stalling of the DNA fork can lead to replication fork collapse, fragmentation of the chromosome, aneuploidy, translocation events, or a number of other chromosomal alterations that can greatly disrupt normal gene expression[78]. Although much work has been done to understand these pathways, there are still gaps in our knowledge regarding all the factors involved in these processes. Improved insights of these pathways will enhance our knowledge of factors that contribute to tumor progression, as well assist in finding new therapeutic targets[130].

Of the mechanisms that are utilized to resolve replication lesions, the most important, in regard to maintaining genomic fidelity, is homologous recombination (HR)[131-133]. However, HR is not the most prominent form of DSB repair. Initial experiments looking into HR were focused around the resolution of DNA DSBs. In some cases, programmed DSB breaks will arise in order to facilitate normal cellular processes, such as during mitosis as Holliday junctions (HJ) are resolved[134]. In instances like this, the cell uses HR in order to repair the DSB, while maintaining genomic fidelity[62]. However, many DSBs are not programmed, instead arising from stalling of polymerases at SSB or other ssDNA lesions[78, 88]. The increase in cytotoxicity/mutagenicity associated with DSB is due to the lack of the cell inherently knowing how to accurately piece the fragmented chromosome back together. Many of the pathways that have developed focus predominately on the reestablishment of a congruent chromosome, and thus protecting unprotected DNA ends from nuclease activity, and less on maintaining genomic fidelity.

In order to study these pathways, researchers commonly induce DNA DSB breaks with γ-irradiation (IR)[135]. IR forms DSB through two different methods: The first is

through the direct interaction of high-energy particles with the phosphodiester backbone, leading to break formation[62]. The second, is through the creation of free radicals after water molecules are split by IR in proximity to DNA[136, 137]. Each of these processes more frequently results in the generation of SSB; however, opposed SSB in close proximity to each other will spontaneously result in DSB, and activate DSB repair[62, 63].

Pathway determination and early DSB repair

Although HR is the most accurate form of DSB repair, it is not the predominant method. Within a cell, the majority of DSBs are repaired via NHEJ. This is due, in part, to the necessity of vast regions of homologous sequences to be present within the cell for HR to occur, and this is a condition that usually only exists between late S-phase and M-phase of the cell cycle. Interestingly, even in conditions that are considered favorable for HR, the ratio of NHEJ:HR remains as high as 4:1[138].

However, even with the vast majority of DSBs being repaired via NHEJ, mutation of factors in HR have a stronger association with breast cancer progression than factors in NHEJ[58, 139-144]. Of the factors mentioned above that are commonly mutated in TNBC, all but one, p53, plays a role in HR, and only one, ATM, is involved in NHEJ[145, 146]. This fact underlies the importance of HR in maintaining genomic stability, since it is not only involved in the repair of DSB, but also the stabilization and resolution of stalled replication forks[147, 148]. But with the requirement of a large region of homology to be present on a sister chromatid for HR to occur, it only makes sense that pathway determination in DSB repair is reliant on the cell cycle. Specifically, it has been determined that cyclin dependent kinases (CDKs) play a crucial role in pathway determination[149, 150].

Early in G1 of the cell cycle, CDK4 and CDK6 are the predominant kinases present. This gradually changes to CDK2 as G1 progresses. In cycling cells, CDK1 levels start to increase as the cell enters S-phase and remains elevated till the end of M-phase. The dynamic interaction between these kinases and their inhibitors allows for the tight control of the cell cycle with set checkpoints in place to inhibit cells from progressing into further stages of the cell cycle with DNA damage present (reviewed in[151]).

Important to DSB repair pathway determination is the ability of CDK1 to phosphorylate CTBP interacting protein (CtIP) on T847 and S327 and antagonize p53 binding protein (53BP1) through the actions of SMARCAD1 (SWI/SNF-Related, Matrix-Associated Actin-Dependent Regulator Of Chromatin, Subfamily A, Containing DEAD/H Box 1)[149, 150, 152-155]. The phosphorylation of S327 on CtIP supports BRCA1 recruitment, which stimulates DNA end-resection[154-158]. Resection is further promoted by the CDK1-dependent phosphorylation of exonucleases EXO1 and Nijmegen breakage syndrome protein 1 (NBS1)[159-161]. Conversely, CDK2 stimulates peptidyl-prolyl cis-trans isomerase NIMA-interacting 1 (PIN1) which inhibits DNA end-resection and thus inhibits HR[162]. These crucial steps in determining to what extent end-resection occurs is what ultimately dictates which form of DSB break repair will occur. Although the two major forms of DSB repair are NHEJ (herein referred to as classical-NHEJ; c-NHEJ) and HR, there exists a number of other pathways, sharing many similarities to these two pathways[163, 164]. These pathways, collectively termed alternative-NHEJ (a-NHEJ), include microhomology-mediated end-joining (MMEJ) and SSA pathways[163, 164]. A major defining factor of each of these pathways is the amount of end-resection that is necessary for each of them to occur,

with c-NHEJ needing ≤4bp of overlap, MMEJ needing 2-20bp, SSA needing >25bp, and HR needing >100bp[165].

These a-NHEJ pathways are generally regulated to a small portion of the cell cycle, due to the necessity of factors from both c-NHEJ and HR to be present[146]. This unique nature of the a-NHEJ pathways actually results in them being some of the most mutagenic pathways[146]. For example, although SSA allows for the largest amount of homology to be found amongst all the NHEJ pathways, it is also the most error-prone pathway, with Artemis cleaving vast regions of DNA in order to force ligation of a short region of homology[146]. Interestingly, inhibition of SSA is not facilitated till later in the HR pathway, meaning that SSA is also promoted by recruitment of BRCA1. Multiple studies have demonstrated that in conditions where end-resection is favored, inhibition of a-NHEJ pathways is not achieved until BRCA2/RAD51 bind at the site of the DNA lesion[156, 166-168]. This, in part, explains how the presence of a functional HR pathway is vital to genomic stability. Irregularities in this pathway, activate some of the most mutagenic pathways involved in DDR.

Regardless of the choice pathway, initial break recognition is facilitated through the same pathway. The MRE11-RAD50-NBS1 (MRN) complex is thought to be the first factors recruited to the site of DNA damage, allowing for ATM recruitment and subsequent autophosphorylation at Ser1981[169]. Once activated, ATM phosphorylates as many as 20 other factors involved in DDR, including histone 2aX (H2AX) into γH2AX[170]. However, this MRN-centric model of dsDNA break recognition is contradictory to other studies showing that NBS1 brings the MRN complex to sites of DNA damage through its

ability to bind γH2AX through its BRCT domain, which suggests that ATM exists at sites of dsDNA breaks prior to the MRN complex being recruited[171].

Irrespective of the initial recognition of the dsDNA break, the phosphorylation of H2AX allows for the loosening of the chromatin structure permitting the recruitment of repair machinery. Although γH2AX is most concentrated immediately flanking the sites of DNA damage, γH2AX has also been observed as far out as 2-30 megabases from the site of DSB[172]. From here 53BP1 is recruited, and phosphorylated by ATM[173]. At this point, HR and c-NHEJ diverge with p53BP1 and BRCA1 antagonizing one another to determine pathway selection, in part due to their differing requirements for end-resection as well as the necessity for an extra copy of genetic material to be present within the cell for HR to occur[174-176].

Homologous Recombination

The process by which RAD51 is recruited to sites of DSBs is quite extensive, and not fully elucidated, with any aberration in these pathways resulting in less efficient DDR and the potential for another pathway to facilitate the repair (Figure 3). The primary requirement for HR to be successful is the presence of a homologous region of genetic material that the cell can use as a template while repairing the DSB. As such, HR factors are seen to be more active in late S-phase through the start of M-phase, as this is the short time within the cell cycle that a duplicate copy of genetic material is present. As noted above, for HR to occur, extensive end-resection, that can extend for 10s of kilobases, is necessary[177]. Initial 5' and 3' end resection occurs through activities by CtIP and the MRE11 subunit of the MRN complex[178]. Both of these factors remain within close proximity of the DSB and are capable of facilitating resection for up to 1kb from the site of

the break[178]. This initial end-resection serves two major purposes. The first, and most obvious, being to initiate end-resection and remove factors that antagonize further resection. The other function of these factors is to remove covalent adducts that may have formed at the 5' ends of breaks, thus allowing resection to occur[179].

As stated above, CtIP activation and recruitment is facilitated by CDK1[150]. However, the activation and recruitment of MRE11 is much more regulated, with the two other subunits within the MRN complex, serving to regulate MRE11 activity. The first step that must occur for MRE11 recruitment is the translocation of MRE11 from the cytoplasm to the nucleus. Since there is no nuclear localization signal (NLS) on MRE11, it relies on the NLS in NBS1 for its translocation[180]. Once within the nucleus, DNA-binding by the MRN complex is achieved by the coiled-coil domain of RAD50 and the intermolecular binding of the Zn^{2+}-hook to linear DNA molecules[181-183]. RAD50, serves a second purpose, in that, when bound to ATP, it masks the active site of MRE11, thus preventing end-resection[184]. However, when MRE11 is bound to RAD50 it catalyzes ATP hydrolysis activity of RAD50, thus leading to conformational changes in both MRE11 and RAD50, exposing the nuclease site on MRE11 and allowing end-resection to proceed[184, 185]. As neither the MRN complex nor CtIP migrates past the site of the initial break, end-resection past ~1kb is facilitated through the recruitment of other factors.

With initial end-resection having been achieved, further resection is carried out by EXO1 and DNA2, with EXO1 having 5'-3' nuclease activity and DNA2 contributing predominately to 3'-5' resection[186-189]. Interestingly, DNA2 also exhibits 5'-3' exonuclease activity, however, in the presence of RPA, this activity is attenuated, while 3'-5' activity is enhanced[190, 191]. Furthermore, in vertebrates, RPA is necessary for DNA2 recruitment to

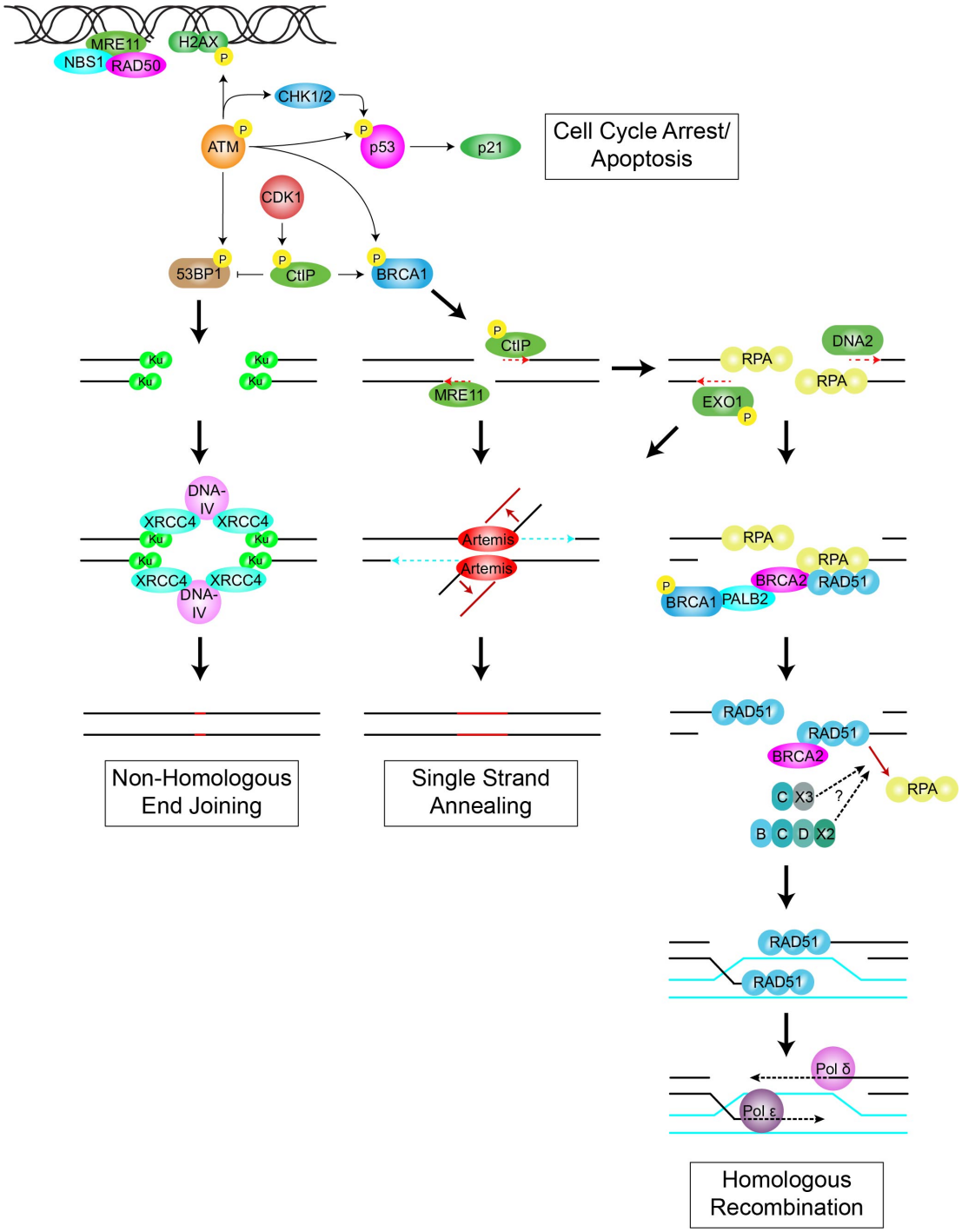


Figure 3 Homologous recombination and double-stranded break repair.

the site of DNA breaks, thus leading to DNA2's 3'-5' resection bias[192]. Co-fractionation experiments have revealed that the MRN complex also binds to DNA2, meaning that in addition to supplying the appropriate 5' overhangs necessary for DNA2 activity, the MRN complex also recruits DNA2 to the site of DNA damage[191, 193, 194].

With 3'-5' exonuclease activity being covered by DNA2, EXO1 facilitates 5'-3' resection. EXO1 is recruited to sites of DNA damage by CtIP[195]. Like DNA2, EXO1 activity is stimulated by the presence of RPA, however RPA is not necessary for EXO1 to be recruited to sites of DNA damage[189]. RPA does, however, protect ssDNA from forming non-beneficiary structures with EXO1, allowing resection to proceed[189]. Without this extensive end-resection, RAD51 may not be able to find sites of homology, forcing non-optimal pairing or unresolved dsDNA damage[196].

However, if sufficient end-resection to facilitate homology-search is achieved, the next steps in HR are facilitated by the recruitment of BRCA1. As mentioned above, BRCA1 is considered the first major HR factor in the fate determination between HR and c-NHEJ. In this capacity it serves to recruit other factors necessary for HR either through direct protein interactions or by making conditions favorable for the binding of other factors to the repairisome utilizing its E3 ubiquitin ligase activity[197, 198]. With its position at the top of the HR pathway, and thus responsible for the recruitment of all factors involved in HR, mutations in BRCA1 have been shown to lead to some of the most severe mammary tumors, being consistently more invasive and of higher-grade at diagnosis than other tumors associated with mutations in HR machinery[199]

Part of the reason for this increase in severity in tumors associated with BRCA1 is due to the E3 ubiquitin ligase activity associated with BRCA1, which is achieved through

the dimerization of BRCA1 with BARD1, forming a RING domain[197, 200, 201]. This pairing is then able to form unique lysine[6]-lysine ubiquitin chains, which are uncommon and may function as a signal for factor recruitment[202, 203]. Not surprisingly, loss of BRCA1 leads to rapid accumulation of 53BP1 at sites of dsDNA breaks, and thus greatly attenuates a cells ability to repair breaks via HR[176]. Interestingly, concurrent loss of both 53BP1 and BRCA1 rescues cells from this BRCA1$^{-/-}$-phenotype, and allows HR to continue[176]. As 53BP1 prevents end-resection at the site of dsDNA breaks, this phenotypic rescue of BRCA1$^{-/-}$ by loss of 53BP1 illustrates the importance of end-resection to occur in order for HR to proceed.

Initial understanding of the importance of BRCA1 was reported on in the 1970's. During this time an inheritable form of breast cancer was known to exist, with multiple studies showing that an individual's risk of the disease was substantially increased by being related to someone with the disease[204-206]. In 1990, Hall et al. found the first susceptibility locus at 17q21, with subsequent confirmation coming from Narod et al.[140, 207]. Interestingly, these reports failed to postulate on the possibility of a novel gene in this region. Instead, Hall et al. speculated that these familial tumors were associated with other known tumor-associated genes in the region, such has *ERBB2 (*gene from which HER2 is expressed*)*, estradiol-17β (*EDHB17*), homeobox 2 (*HOX2*), retinoic acid receptor α (*RARα),* and *wnt3*[140].

Shortly thereafter, this region would be identified as the BRCA1 locus, with subsequent studies soon identifying another breast cancer susceptibility locus on chromosome 13 that was termed BRCA2[208, 209]. That same year (1994) Miki et al. successfully cloned BRCA1 and identify many of its domains, including the "acidic blob"

DNA binding domain[210]. Further studies would go on to identify multiple familial mutations in BRCA1 that increased breast cancer susceptibility[210-213]. Prominent amongst these were the discovery of the 187delAG mutation and the subsequent association with the Ashkenazi Jewish population, allowing for a strong hereditary link to be formed between families from different regions[213-215].

Although the seriousness of the risks associated with mutations in the *BRCA1* and *BRCA2* had been determined to be 91% and 70% incidence by the age of 70 respectively, the cellular role of BRCA1 and BRCA2's had yet to be identified[216]. In 1997 multiple works by Scully et al. linked BRCA1 to DDR, showing the co-localization of RAD51 and BRCA1 at sub-nuclear foci in response to DNA damage[217-219]. Although other DNA damage response factors, including p53, had previously been associated with cancer, BRCA mutations were unique in that there was a strong association between heredity and aggressive forms of the disease[220].

The involvement of BRCA2 in DDR would be established in 1999 with Yuan et al. demonstrating its necessity to facilitate RPA-RAD51 exchange at sites of DSBs in vertebrates[221]. As mentioned above, RPA binds to ssDNA in order to protect it from nucleases and to prevent the formation of inhibitory secondary or tertiary structures in the ssDNA[192]. However, for HR to occur, RAD51 must first compete with RPA, which has a significantly higher ssDNA binding affinity. Part of this high affinity for binding ssDNA comes from the four ssDNA-binding domains in RPA (three on the RPA70 subunit and one on the RPA32 subunit)[222-224]. This structure is stabilized by the third subunit of RPA, RPA14[225].

Interestingly, RAD51 is capable of binding to one of the ssDNA-binding domains of RPA70, thus lowering the binding affinity RPA has for ssDNA, and allowing for RPA-RAD51 exchange[226]. The mechanism by which this binding-site on RPA becomes detached from ssDNA, and thus exposed for RAD51 to bind to, is highly debated. One possible mechanism focuses on the multitude of structural conformations RPA utilizes to bind ssDNA, which possibly exist in equilibrium[227]. This theory suggests that while transitioning from one conformation to another, the ssDNA-binding domain on RPA70 becomes exposed, allowing RAD51 access, and thus transiently facilitating strand exchange[228]. It is unlikely that this mechanism plays any significant role in RPA-RAD51 exchange, because in the absence of RAD52, BRCA2, or other factors known to facilitate RPA-RAD51 exchange, RAD51 does not have a significant presence on ssDNA[221, 229-231].

Much of the work showing that mediators are necessary for RPA-RAD51 exchange was carried out in bacteria (RecFOR complex) and fungus (with the Brh2 protein)[232-234]. In these models, the RecFOR and Brh2 proteins recognize the ssDNA-dsDNA junction and use this structure to facilitate RAD51 filament growth as it replaces RPA[234]. However, in yeast and vertebrates, this system becomes much more convoluted[234]. In yeast, RAD52, the primary molecule responsible for RPA removal, lacks a definitive localization domain, and instead is known to wrap ssDNA around itself, destabilizing the ssDNA-RPA interaction[235]. At the same time, RAD52 stimulates RAD51 binding to newly exposed ssDNA through a direct interaction of RAD52 and RAD51[236].

In vertebrates, even though the RAD52 structure remains highly conserved, its essential function is diminished, and BRCA2 was found to be necessary for RAD51 recruitment[221, 237]. Later it was found that this action is facilitated by the BRCA2-DSS1

complex, although the exact function of DSS1 has not been elucidated[221, 229-231]. Interestingly, loss of BRCA2 in this system allows for partial restoration of RAD52 function, showing some level of redundancy in RPA-RAD51 exchange in vertebrates[229]. This redundancy, could, in part, explain the controversy surrounding the necessity of BRCA2 for effective HR[238]. However, as seen by the significant increase in breast cancer incidence, and higher levels of genomic instability in individuals found to have pathogenic variants of BRCA2, BRCA2 is necessary for efficient HR to proceed[216]. These observations underpin the lack of understanding that still exists in the mechanisms surrounding RAD51 recruitment in HR. Much of our understanding of these pathways comes from organisms that achieve RAD51 recruitment via machinery that does not function the same way in vertebrate systems. In addition, due to gene duplication events and divergent evolution, many factors that are present at the site of DSBs in vertebrates do not exist in other organisms, so thus can only be researched in vertebrate systems[239, 240]. This reinforces the necessity of further research into the discovery of novel factors and their function in HR.

Later studies investigating the roles of BRCA2 in HR have demonstrated that it has multiple roles in regulating the nucleation of RAD51 onto ssDNA. These are facilitated by eight BRC binding domains on BRCA2, which can be assembled into two groups, BRC1-4 and BRC5-8, with each group serving a distinctly different role in RAD51 recruitment[241]. BRC1-4 are capable of binding RAD51 monomers, targeting them to ssDNA and away from dsDNA through reduction of RAD51s ATPase activity. BRC5-8 on the other hand bind to the RAD51-ssDNA filament and serve to stabilize its structure[242]. DSS1 has been shown to support this function of BRCA2[231].

With a better understanding of the pathways that BRCA1 and BRCA2 are involved in, researchers began looking into the categorization and classification of familial BRCA-related tumors. A distinct phenotype started to present itself, which differed wildly from that of sporadic breast cancers. BRCA-associated tumors were found to regularly be TNBC, with pushing margins, lymphatic infiltration, mutations in *TP53*, and upregulation of epithelial growth factor receptor (EGFR) and *cMYC*[58, 143, 243-247]. Interestingly, there was found to be a reverse correlation between age at onset and tumor-severity, with younger women having higher grade infiltrates at presentation, potentially suggesting that these tumors are driven by *BRCA* mutagenesis, while tumors developing later in life are initiated by other, more common factors[199].

BRCA-associated tumors also tend to be less differentiated with a more mesenchymal or basal-like phenotype[248]. Fundamental to this is the general lack of a specific tumor subtype associated with these tumors[130]. While initial reports showed a low incidence of DCIS being present in *BRCA*-associated tumors[249], more recent reports have demonstrated a vast majority of *BRCA*-associated tumors are positive for DCIS, possibly due to the increased resolution of newer screening mammography techniques[199]. Moreover, these tumors are also significantly more likely to be positive for IDC than their sporadic counterpart[250]. Collectively, the phenotype associated with *BRCA*-related tumors have been termed "BRCAness," and may aid in the determination of which DCIS will proceed into an IDC[250].

However, BRCA1 and BRCA2 are clearly not the only factors involved in HR. Another crucial factor that has been shown to interact in this pathway is the partner and localizer of BRCA2 (PALB2). PALB2 serves as a bridge, bringing the effector molecules

of HR to the growing repairisome through multiple domains. It does so through binding BRCA1 with its coiled-coil domain and BRCA2 with its WD40 β-propeller domain[251-255]. Interestingly, this step acts as a fate determining step, as prior to BRCA2 recruitment both SSA and HR are promoted, however both BRCA2 and RAD51 have been shown to antagonize SSA activity, thus HR is favored once these factors are recruited[156, 167]. Another less discussed role PALB2 plays in DDR is in support of Pol η, thus enhancing DNA synthesis at sites of repair[256].

Unsurprisingly, loss-of function germline mutations in PALB2 have also been associated with an increased risk of breast cancer, with a higher incidence of TNBC in these cases than in sporadic breast cancer[167]. Studies like this, and many others led to the notion that, while the BRCAness phenotype was initially relegated only to tumors with mutations in *BRCA*, other factors may contribute to this phenotype[58]. Indeed, with the wide implementation of sequencing technologies, tumors that strongly resembled the BRCAness phenotype were found to have wild-type (wt) *BRCA*, but instead had mutations in other HR-related factors[130, 257-259]. Recently, this led to the notion that the BRCAness phraseology should be opened up to include mutations and the misregulation of factors that function in the same pathways of BRCA1/2[130].

Some of the newest factors that have been shown to present with a BRCAness phenotype when mutated in breast cancer, are some of the oldest known HR-associated factors, the RAD51 paralogs (RP). These factors have been notoriously hard to study in vertebrate systems, as loss of any one of them results in embryonic lethality[260-264]. Oddly, concurrent abrogation of *TP53* can rescue this phenotype, although tissue culture of

primary cells from these mice has remained elusive[260]. The increase in survival observed with loss of *TP53* is thought to be due the inhibition of DNA damage-induced apoptosis.

Currently, five classical RPs have been identified, with a few other non-canonical RPs having recently been identified. The five canonical paralogs share 20-30% protein sequence identity with RAD51, and are thought to have arisen by gene duplication events, which then allowed for gene divergence[239, 240]. Amongst all of these proteins, all show high levels of conservation in their Walker A and Walker B motifs, illustrating the importance of their role in ATP binding and hydrolysis[239, 240, 265, 266].

The five classical RPs are known to form two distinct complexes consisting of either RAD51B-RAD51C-RAD51D-XRCC2 (BCDX2 complex) or RAD51C-XRCC3 (CX3 complex)[267-270]. In addition, a number of subcomplexes have been identified, including the RAD51B-RAD51C (BC complex) and the RAD51D-XRCC2 (DX2 complex)[268, 271-273]. Stochastically, each complex contains only one copy of each paralog, and these assemble through the binding of the C-terminus of one RP to the N-terminus of the next[274, 275]. In addition to these complexes, an additional complex is formed by the non-canonical RPs. They are involved in the formation of the Shu complex, which consists of SWSAP1/SWS1(SWS1-Associated Protein/Short-wavelength sensitive opsin)[276, 277]. As opposed to the Walker A/B motifs in the classical RP complexes, the Shu complex contains a highly conserved zinc-finger like binding motif allowing it to moderate RAD51 activities in a yet to be determined way[276-278].

In addition to its involvement in multiple RP complexes, RAD51C directly interacts with PALB2 and associates with the BRCA1-PALB2-BRCA2 complex[279]. Interestingly, mutation of PALB2 to abrogate this interaction leads to an accumulation of

BRCA2 foci but a decrease in RAD51 recruitment at sites of DNA damage, suggesting that RAD51C is required for the nucleation of RAD51[279]. This early involvement of RAD51C is also supported by the necessity of the BCDX2 complex for RAD51 recruitment[280]. Of note, this same study demonstrated that the CX3 complex is dispensable for RAD51 recruitment and nucleoprotein filament formation but is still necessary for HR to occur. This advocates for a late role of the CX3 complex in HR, through either the stabilization of the nucleoprotein filament complex, homology search, or filament dissociation after repair, making RAD51C also necessary in late-stage HR. In line with this, the CX3 complex has been shown to facilitate DNA strand exchange[275].

However, this notion that CX3 is only necessary in late-stage HR has been contradicted by other studies suggesting that CX3 is necessary for nucleoprotein filament formation[274, 275, 278]. These studies argue that the presence of the Walker A/B motif on both the BCDX2 and CX3 complexes demonstrate their role in catalyzing RAD51 filament formation through ATP hydrolysis. In addition to a possible role for both the BCDX2 and CX3 complex in filament formation, the Shu complex has also been shown to be necessary for this process[281]. Complicating this matter, the mechanistic role each of these proteins plays in RAD51 recruitment has not been established[278]. With many of these studies having been done in different model systems, it is hard to directly compare their conclusions. It has also been suggested that different types of DNA lesions lead to recruitment of different RP complexes[282]. Before a more definitive role for each of these factors can be established, many more studies will have to be conducted.

With recruitment of all the necessary factors, RAD51-RPA exchange proceeds. The RAD51 nucleoprotein filament takes on a right-handed helical structure around ssDNA,

with each RAD51 molecule binding three nucleotides[283, 284]. This presynaptic filament is then brought to dsDNA to begin homology search by standard Watson-Crick base pairing[283-285]. TIRF (total internal reflection fluorescence) microscopy in yeast has demonstrated that the presynaptic filament samples eight nucleotides at a time, with a match of at least seven nucleotides being needed to yield a positive match[283, 284].

Once a homologous sequence is identified, and a HJ is formed, DNA synthesis can occur (Figure 4). This process can take many forms, and is the time during which mutations are most likely to arise from HR. In the simplest model, asynchronous synthesis, a single strand invades to form a HJ. This strand is then elongated, ligated, and released before being used as a template to repair its complementary strand[129]. Similar to this process is synchronous synthesis, where both the 5' and 3' ends of a DNA break invade into and pair with their homologous strand, both extending at the same time, before being ligated and released. In both of these cases, no crossover events occur, meaning no factors are needed to facilitate dissolution of the HJ[129]. The last structure that is known to form during DNA synthesis after a dsDNA break is the double HJ[286]. This mimics the structure seen during meiosis and as such requires TopIII and the RecQ helicase to resolve the intertwined strands[286, 287].

Regardless of which approach is taken, DNA polymerases are recruited to the site of replication by RAD54 after RAD51-mediated strand invasion[288]. The most well studied polymerase in this process is Pol δ, which is capable of extending up to 80% of all invading strands[129, 289, 290]. Crucial to this role is PCNA, which loads onto DNA with the c-terminus positioned towards the 3' end and then binds Pol δ, orienting it in the proper direction for elongation[291].

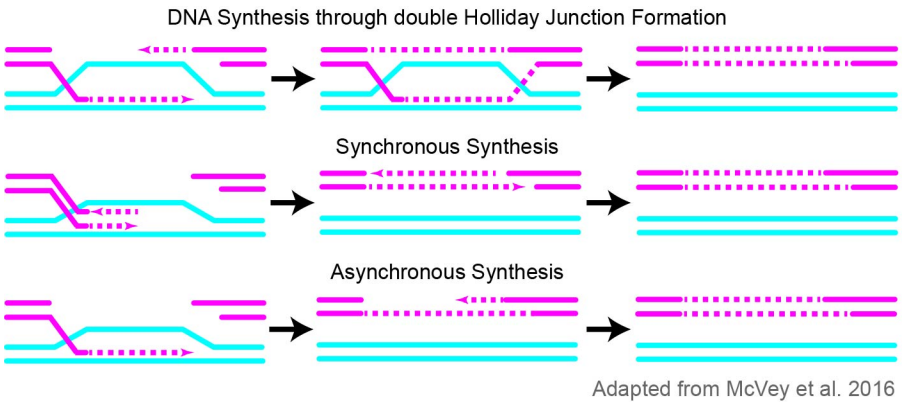


Figure 4 Nascent DNA synthesis in response to DNA damage.

The other major polymerase known to be involved in response to HR is Pol ε[129]. Unlike Pol δ, Pol ε is not capable of displacement synthesis, meaning that only Pol δ is capable of lagging strand DNA synthesis[292]. However, both of these polymerases contain 3'-5' proofreading activity, making them both exceptional at error-free DNA synthesis[129, 293]. In some processes, such as asynchronous synthesis, Pol δ and Pol ε are not recruited to sites of DNA damage, and instead Pol α, Pol ζ, or Pol η are recruited. Although these polymerases share varying degrees of similarity to Pol δ and Pol ε, they lack any innate proofreading activity, and are thus more error-prone than their counterparts, leading to higher rates of mutation[129].

Replication Stress and Resolution

Interestingly, DDR is not the only known role for HR. Recent studies have shown that replication stress is more sensitive to perturbations in BRCA1 levels than other established BRCA1 roles[294]. More specifically, a mutation in a single copy of *BRCA1* is sufficient to reduce replication fork stability, whereas loss of homology in BRCA1 is needed before significant deficiencies are observed in homology directed DNA damage

repair[294]. Moreover, a common factor in replication stress stabilization and repair after prolonged stalling of replication forks is the accumulation of non-DNA-damage-associated RAD51 and other members of the HR pathway. This suggests a recombination-based attempt to resolve the stalled fork[295-299]. Although the role HR repair-elements are crucial for maintaining genomic fidelity, the initial, and possibly more important increase in genomic instability seen during cancer progression in BRCA-associated tumors could be due to the role HR plays in enhancing replication fork stability and the resolution of stalled replication forks.

DNA replication during mitosis is a highly regulated process. With the introduction of larger and larger genomes, it became necessary for organisms to possess multiple origins of replication (Ori) in order to replicate their genomes quickly. However, this results in the possibility of Ori firing multiple times during a single S phase. In order to prevent this, eukaryotes evolved the minichromosome maintenance complex 2-7 (MCM2-7). These helicases bind at the Ori and once activated by CDK2 recruit the replisome[300]. However, once initiation has occurred, a number of other factors, such as DNA adducts, cross-links, DNA breaks, can prevent replication from continuing to completion. Most of these factors are removed before the replisome encounters the lesion through one of the DNA repair pathways. However, in some cases this does not occur, and the DNA lesion must be resolved for replication to proceed.

By far, the simplest way for the replisome to move past a lesion is by bypassing it altogether. Polymerases such as Pol η, Pol ζ, Pol ι, and Pol κ all have this ability, but in doing so introduce mutations[139, 301]. Not surprisingly, their activity is suppressed in favor of less mutagenic pathways. Interestingly, the HR pathway of DNA damage repair plays a

crucial role in the protection and resolution of stalled replication forks[139, 298]. Upon sensing replication stress, the ssDNA, which is quickly coated by RPA[302], that is generated at the stalled replication fork is sensed by ataxia telangiectasia and RAD3-related (ATR)[303, 304]. This in turn sets off a signaling cascade, with ATM also being activated, and recruiting many of the factors previously addressed in break repair[303, 304]. Like in dsDNA breaks, CHK1 and CHK2 are activated, leading to cell cycle arrest via p53 activation and p21 upregulation[139, 303, 304]. In addition, BRCA1 is recruited and directly suppresses bypass replication through the recruitment of RFC[305]. In cases where BRCA1 is mutated and unable to localize to sites of replication stress, 53BP1 and other NHEJ factors are recruited, resolving the replication stress through NHEJ[306]. Unsurprisingly, this is more mutagenic than if factors involved in HR had been recruited. However, it has been shown that concurrent mutation of 53BP1 and BRCA1 rescues the $BRCA1^{-/-}$ phenotype and allows HR factors to bind at sites of replication stress[306].

Before replication restart can occur, the replication fork must first be stabilized, as stalled forks are prone to collapse which can lead to dsDNA break formation, chromosomal fragmentation, and aneuploidy[139]. Previous reports have shown that recruitment of BRCA2 by BRCA1-PALB2 at replication forks stabilizes many of these structures and prevents replication fork collapse[307]. Although collapse of the replication fork does not inhibit replication progression, as new Ori will fire in response, it does introduce fragmented pieces of a chromosome into the cell[300, 307].

Non-homologous end-joining

In instances where HR does not occur, it is replaced with NHEJ, with the most common form of NHEJ being c-NHEJ. Although not always mutagenic, c-NHEJ is

considered to be significantly more mutagenic than HR, regularly leading to the loss of a few nucleotides every time it occurs. Within vertebrates, NHEJ has two primary purposes. The first is for V(D)J recombination within T- and B-cell development[308]. The other, more pertinent role is in DDR. With most cells within our bodies existing in G1 or G0 of the cell cycle, c-NHEJ is the primary form of DDR our cells use. This form of DDR is considered to be error-prone and thus has a high likelihood of introducing mutations as DDR is occurring.

Upon induction of a DSB, a heterodimer consisting of Ku70 and Ku80 (hereinto forth referred to as Ku) quickly binds to the break through a DNA-binding domain on Ku70[309, 310]. Ku can be detected at sites of DNA damage within seconds of the formation of a DSB[309]. It has been speculated that the rapidity of this binding is due to the high affinity (2×10^9 M^{-1}) of this complex for DNA as well as the abundance of Ku within the cells (~500,000 molecules/cell)[310-313]. Once bound, Ku works as a backbone for the recruitment of other factors involved in c-NHEJ as well as preventing extensive end-resection, and thus preventing other DDR pathways[146].

The simplest and most conservative form of c-NHEJ, generally resulting in no genetic material being added or lost, is blunt-end ligation[146]. This process consists of the Ku-XRCC4-DNA ligase IV complex. After initial binding of Ku to the DSB, Ku promotes binding of XRCC4-DNA ligase IV to the DNA ends[311]. XRCC4 binds to DNA ligase IV via a region adjacent to DNA ligase IV's two BRCT domains, forming a 2:1 XRCC4:DNA ligase IV ratio[314]. In this way, it is though that XRCC4 is able to bridge the two DNA blunt-ends, and facilitate DNA ligation by DNA ligase IV[146].

In instances where there are incompatible DNA ends, either due to mismatched overhangs or modifications of the phosphodiester backbone, end-processing must occur. One common method for this to occur is through the use of a nuclease[146]. Within c-NHEJ, EXO1, Mre11, and EXD2 endonucleases are involved in nucleotide removal, however in all cases Artemis endonuclease is necessary and is used to cleave unwanted ends or to resect DNA ends to facilitate microhomology (≤4 base pairs [bp])[315]. Artemis is recruited to the site of DSB by DNA-PKcs, after autophosphorylation[316, 317]. At this point, DNA-PKcs phosphorylates Artemis, which is thought activate Artemis by revealing its catalytic region[146, 318].

Once active, Artemis facilitates DNA end-resection in a multitude of different ways, the simplest of which being the removal of ssDNA-overhangs in order to form blunt ends[319, 320]. Interestingly, this is handled differently between 3' and 5' ends, possibly due to how the Artemis-DNA-PKcs complex binds to DNA-ends[321]. Whereas the 3' end is cut to leave a 4bp overhang, the 5' end is cleaved at the ss-dsDNA junction, forming a blunt end for the Ku-XRCC4-DNA ligase IV complex[319, 320, 322]. However, in some instances, where chemical alterations in the DNA structure make ligation of the DNA ends impossible Artemis is able to further resect DNA ends in order to remove the alteration and expose an area for microhomology. It is through these mechanisms through which genetic material is lost and is why c-NHEJ is considered to be mutagenic.

In addition to recruiting nucleases to remove nucleotides interfering with c-NHEJ, Ku is also able to recruit polymerases Pol μ and Pol λ[319]. Pol λ serves a more traditional role in the template-based addition of nucleotides to fill in gaps left after strand annealing and ligation[323]. Pol μ, due to a variation of loop 1 is able to add nucleotides in a template

independent manner[323, 324]. In this way Pol μ is able to add nucleotides to generate regions of microhomology that are favorable for Ku-XRCC4-DNA ligase IV mediated ligation[325]. Factors such as XLF are able to aid in this process through the stabilization of mismatched ends and their subsequent ligation[326].

BRCAness and PARP inhibition

The presence of mutations in factors involved in DDR-repair elements sensitizes a cell to genotoxic agents. Thus, the importance of fully understanding these pathways is crucial. More specifically, an increase of our understanding of the factors that contribute to BRCAness is important as it opens up treatment options to novel therapeutic targets. The defects in HR seen in these tumors forces them to use more error prone pathways. By taking advantage of this feature and using targeted therapeutics to specifically target this deficiencies, researchers have found that they can force these tumor cells to gain mutations at a rate that is incompatible with cell viability, while leaving most somatic cells relatively untouched[58, 130].

Many of the original treatments, and still strongly utilized therapeutics, revolve around the application of platinum salts in order to create ICLs that cannot be adequately resolved with a defective HR pathway. These are generally fairly well tolerated, although leucopenia, neutropenia and anemia have all been reported[327]. In recent years, researchers have also found that HR-deficient tumors are susceptible to synthetic lethality treatment with PARP (poly (ADP-ribose) polymerase) inhibition (PARPi)[86, 259, 328-335]. These studies have demonstrated that with a dependency of cells to have a deficiency in HR for this form of synthetic lethality treatment to be effective, these drugs are very well tolerated at even high doses. The idea of synthetic lethality arose from the observations that concurrent

mutations in associated pathways could lead to cell death, even though a single mutation in either pathway alone still resulted in a viable cell[336]. In the case of synthetic lethality through PARPi, multiple labs have discovered that while loss of the HR-pathway alone through mutations in *BRCA1/2* still produced a viable cell, forcing more dsDNA breaks through with inhibition of ssDNA break pathways resulted in loss of cell viability[329, 337]. Preliminary results from studies with PARP inhibitors in combinational therapies have shown a 51% pathological complete response (up from 26% in control individuals) in TNBC individuals[338]. With such a strong response the PARPi Olaparib (AZD2281) has only recently gained approval by the Unites States Food and Drug Administration for use in BRCA-associated tumors[339].

Although this is a significant step forward in the treatment of BRCA-associated breast cancer, there are mutations in other factors that have been shown to sensitize tumor cells to PARPi[259, 331-335]. It is likely, that as our understanding of these factors grows, many of them will be recognized as a suitable target for PARPi treatment in breast cancer. Although much is already known about the HR pathway, there obviously still exists large gaps in our understanding of crucial steps in this pathway. Even factors whose roles were thought to be well established have come into question in recent years. A prime example is the necessity of BRCA2. As described above, BRCA2 is necessary for RAD51 recruitment; however, multiple studies have shown that RAD51 is still recruited to DDR foci in the absence of BRCA2[238, 340]. Thus, a more complete understanding of these pathways is still needed, since a clearer picture of which factors are involved in DDR will lead to advancements in the treatment of breast cancer.

SIM2s

Single-minded 2 (SIM2) is a homolog of the *drosophila* regulator of neurogenesis and midline development: *single-minded* (*sim*)[341-344]. A series of studies conducted during embryogenesis of *drosophila* between 1988 and 1991 demonstrated that *sim* plays a crucial role in the development of the central nervous system (CNS) by regulating midline cells[342-345]. These studies showed that sim is a bHLH/PAS (basic-helix-loop-helix/PER-ARNT-SIM; period, aryl hydrocarbon receptor nuclear translocator) transcription factor that localizes to the nucleus of midline cells and were necessary for proper differentiation and gene expression during development[342-344].

With such an important role in the development of the CNS, it is unsurprising that loss of *sim* in *drosophila* is embryonic lethal[345]. This trend continues into vertebrates, where two *sim* homologs exist: SIM1 and SIM2[346]. In situ hybridization experiments in murine models revealed that *Sim* can be detected in the CNS as early as embryotic-day 8[346, 347]. However, *SIM2* expression is not relegated to these tissues, with kidney, skeletal muscle, lung, and breast also showing *SIM2* expression[346, 348, 349]. Moreover, loss of SIM2 does not result in prenatal lethality, but instead is lethal early after birth in ~99% of pups due to the presence of cleft pallet leading to aerophagia[350].

These findings of early lethality with loss of *SIM2* is not surprising as the bHLH-PAS family of transcription factors are an important family of transcriptional regulators that are involved in multiple cellular processes, including circadian rhythms, hypoxia, toxin metabolism, and more[351]. Members of this family are typified by the inclusion of a DNA binding bHLH domain followed by a series of two PAS domains: PASA and PASB (Figure 5a)[352]. The bHLH binding domain allows these transcription factors to bind to

protypical E-box motifs[353]. Canonically, members of this family are regulated through factors interacting with the PAS domains, with a few crucial exceptions (discussed below)[351, 354-356]. In the case of AHR (aryl hydrocarbon receptor), on of the best described bHLH/PAS family members, binding of the ligands, such as xenobiotics and metabolites, to the PASB region leads to its activation and subsequent translocation to the nucleus[352, 357]. This later induces the transcription of genes required for xenobiotic metabolism, such as the cytochrome P450 family of enzymes[357, 358].

In other cases, factors binding to the PASB domain result in the negative regulation of the bHLH-PAS transcription factor. For example, the interaction between the bHLH-PAS family member CLOCK and cryptochrome 1/2 (CRY1/CRY2) at the PASB domain represses CLOCK-BMAL1 induced transcription, which includes CRY1/CRY2, and thus allows for cell autonomous autoregulation of circadian rhythms[359, 360]. In regards to the negative regulation of SIM2, the PASB region is thought to be the target of ubiquitination by ubiquitin ligases such as PRKN (Parkin), leading to its subsequent proteasomal degradation[352, 361].

In contrast to the promiscuous binding nature of the PASB domain, the PASA domain plays a crucial role in the dimerization with other bHLH-PAS family members (Figure 5b)[352, 362]. These pairings are fairly specific for each bHLH-PAS family member and allows the bHLH-PAS family to be divided into two classes (class 1 and class 2) with a member from class 1 dimerizing with a member from class2. Class 1 members are the signal-regulated subunit, and include AHR, HIF1α and HIF2α (hypoxia inducible factors 1α and 2α, SIM1, SIM2, NPAS1-4 (neuronal PAS domain protein 1-4), and CLOCK.

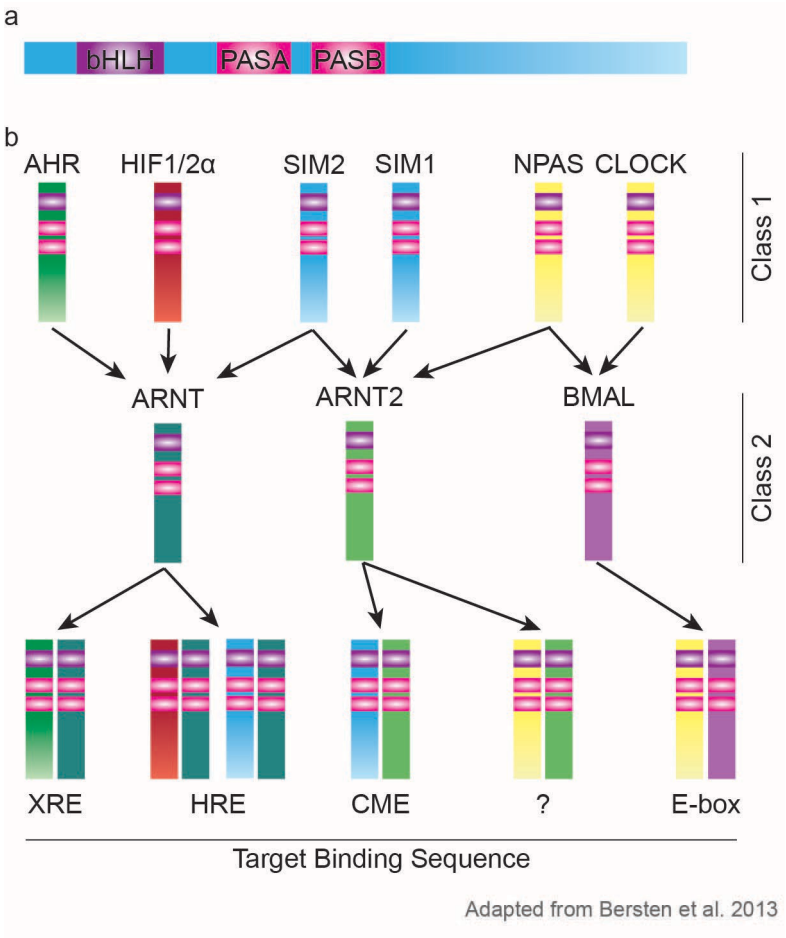


Figure 5 bHLH-PAS family members.
(a) Basic schematic of domain layout on bHLH-PAS family transcription factors. (b) Schematic showing bHLH-PAS family members and their binding partners as well as the response element they bind to (XRE – Xenobiotic response element, HRE – Hypoxia response element, CME – Central midline element).

Class 2 members are more ubiquitous, and less tissue specific. These include ARNT, ARNT2, and BMAL[352].

SIM2 specifically has been shown to dimerize with ARNT and ARNT2 allowing it to bind to the central midline element (CME)[346, 352]. SIM2's binding capacity for ARNT leads to it competing with other bHLH/PAS family members in order to heterodimerize with ARNT and regulate gene expression via interaction with hypoxia response elements (HRE), dioxin response element (DRE), or central midline elements (CME)[347, 363]. This is an interesting dynamic, as unlike other bHLH-PAS members, SIM proteins are transcriptional repressors, and thus their competitive binding for ARNT represses the corresponding pathways[346, 347, 364].

By binding to the HRE, SIM2 directly competes with HIF1α and HIF2α (hereinto forth referred to as HIF), thus interfering with hypoxic response. One study suggests that this competition results in SIM2 activation leading to pro-survival signals for tumor cells in specific tissues[365]. This is in direct contrast to other findings, suggesting that the pro-oncogenic/anti-oncogenic effect these molecules exert on a cell cannot be so rigidly defined. It is well established that the tumor microenvironment is highly hypoxic, and thus exhibits increased expression of HIF[366]. This increase in HIF expression is most likely the cause of reduced inhibition from FIH1 (factor inhibiting HIF1α), an enzyme which is activated by high concentrations of O_2[367]. The subsequent signaling cascades have both pro-oncogenic and anti-oncogenic effects, with HIF1α signaling activating p53 signaling and thus apoptosis, while HIF2α signaling activates MYC signaling (pro-proliferative) and angiogenesis, allowing for increased nutrient supply to the tumor[368-370].

In addition to the hypoxic response, HIF is also upregulated in response to other cellular stressors, such as glucose deprivation[371]. Once activated through this mechanism, HIF activates pro-glycolytic pathways and suppresses oxidative phosphorylation[371]. This has the added effect of reducing the presence of ROS in the cell, and thus has a pro-survival effect[370, 371]. This reduction in ROS levels is also enhanced by HIF2α-dependent promotion of SOD expression, thus lowering ROS levels and decreasing p53-induced death[370]. Together, these features partially outline a condition in cancer metabolism, referred to as the Warburg effect. This is characterized by the tendency of a tumor cell to favor glycolysis over oxidative phosphorylation, in order to prevent ROS formation and generate metabolites that are necessary for rapid cell growth, while sacrificing ATP synthesis[372, 373]. Meaningfully, it has also been suggested that SIM2 plays a role in metabolism, again aligning SIM2 and HIF signaling; however, it is yet to be determined the exact role SIM2 plays in metabolism[374].

Due to the contradictory signaling of HIF and SIM2, it would not be surprising if SIM2 supported oxidative phosphorylation. These disparities between HIF and SIM2 are again seen in the inflammatory response, where NF-κB (nuclear factor κ-light-chain-enhancer of activated B cells). In regard to HIF, a combination of studies has demonstrated that hypoxic conditions can lead to NF-κB activation, which is then able to activate *HIF1α*. In turn HIF promotes NF-κB, creating a positive feedback loop (Figure 6)[375-377]. In contrast our laboratory has demonstrated that *SIM2s* expression leads to the downregulation of NF-κB and in turn NF-κB activation leads to decreased levels of *SIM2* (Wyatt et al. 2019, unpublished data; *in review*)[378]. In this respect, it is not surprising that SIM2 and HIF compete for the same HREs; however, this repressive property of mammalian SIM

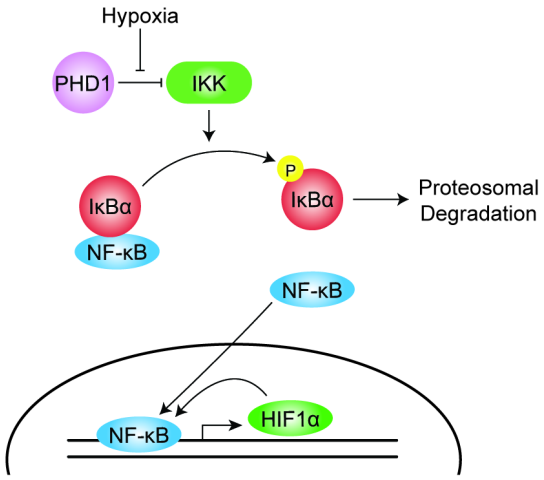


Figure 6 HIF-dependent regulation of NF-κB.

proteins drastically differs from that of *drosophila* sim, which is a transcriptional activator[379]. In the case of SIM2 these inhibitory properties are the result of two repressive domains at the C-terminus, a proline/alanine rich domain and a proline/serine rich domain (Figure 7a)[347].

Interestingly, *SIM2* is expressed in two isoforms, *SIM2*-long (*SIM2l)* and *SIM2*-short (*SIM2s*; Figure 7a)[380]. The variation in these two isoforms exists at the c-terminus with *SIM2l* incorporating exon 11, while *SIM2s* transcribes past exon 10 but terminates before exon 11 (Figure 7b). This results in the loss of the proline/alanine rich repression domain, with the proline/serine rich repression domain remaining intact. This change reverses the role of SIM2 at CME, with SIM2s enhancing expression of genes from this regulator[349]. Conversely, the presence of the proline/serine rich domains on SIM2s is still sufficient to suppresses transcription from DRE and HRE[349].

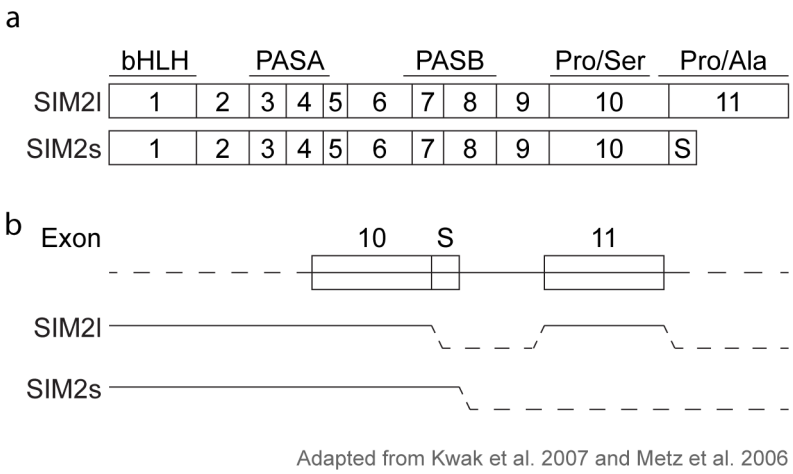


Figure 7 SIM2 isoforms.
(a) Graphical representation of exon inclusion of *SIM2l* vs *SIM2s*. (b) Schematic representing differential transcription between *SIM2l and SIM2s.*

In addition, these isoforms exhibit tissue specific expression, with *SIM2l* being the predominant form in skeletal muscle and brain, while *SIM2s* is the predominant form in breast tissue[349, 381]. Although tissue specificity is not abnormal in bHLH-PAS family members, the lack of known regulators of *SIM2* is unique[351, 352]. Of the aforementioned class 1 bHLH-PAS family members, AHR responds to xenobiotic stressors, HIF responds to hypoxia, and CLOCK/NPAS is regulated by circadian rhythms, leaving SIM as the only bHLH-PAS family member with no known mechanism of activation[349, 351, 352, 357, 363].

Of these factors, HIF has been shown to have the most in common with SIM2. Great strides have been made towards the understanding of its down-stream signaling, in part because researchers have shown the mechanisms governing the regulation of HIF itself. Many of the original studies looking into the regulation of HIF were conducted in the 1990's, when two studies conducted in 1996 came to the conclusion that hypoxic conditions did not effect *HIF* mRNA levels, disrupting the canonical thinking of HIF

regulation[382, 383]. Of these two reports, Huang et al. speculated that HIF was rapidly degraded at the protein level unless stabilized, potentially through a phosphorylation event in response to hypoxia[383]. The next year, Salceda and Caro found that treatment of cells with MG-132, a potent 26S proteasome complex inhibitor, stabilized HIF in normoxic conditions[384]. As the 26S proteasome complex targets ubiquitinated proteins, it is unsurprising that in 2002 Groulx and Lee discovered that the VHL (von Hippel-Lindau) subunit of the VBC/Cul-2 (VHL, elongin B, elongin C, Cullin-2, and Rbx1) E3-ubiquitin ligase complex targeted HIF for proteasomal degradation in normoxic conditions[385]. This oxygen dependent targeting of VHL to HIF is indirectly facilitated by PHD (prolyl hydroxylase), which hydroxylates the ODD (oxygen-dependent degradation) domain of HIF, increasing its affinity for VHL[386, 387].

As stated above, SIM2, at least in part, mimics this method of regulation in that it is constantly expressed, however is rapidly turned over through the actions of the E3-ubiquitin ligase PRNK[352, 361]. However, under which conditions PRKN targets SIM2 is still unknown. Meaningfully, the studies investigating the regulation of HIF were made possible by a greater understanding of the conditions in which HIF was activated, thus allowing scientist to study perturbations in the cellular state to determine how HIF was being activated. To date, it is still unknown what factors lead to the activation of SIM2. This gap in our understanding of *SIM* regulation, has made the study of SIM2s difficult as upregulation of *SIM2s* is only achievable through the artificial amplification of *SIM2s* using constitutive active or inducible promoters or the whole-cell inhibition of proteasomal degradation by treatment with MG-132, with prolonged treatments inducing apoptosis, all

of which ultimately leads to non-biologically relevant levels of *SIM2* within a cell and inherently disrupting cellular homeostasis.

SIM2 and disease

The location of *SIM2* within the Downs syndrome (DS) critical region (DSCR) on chromosome 21 (chromosome 16 in mice) has led to ample speculation on the role of SIM2 in DS[350, 388, 389]. DS is broadly defined as a trisomy of chromosome 21 in humans, although the region necessary for an individual to display all the hallmarks of DS has been distilled down to a 1.6-megabase region in chromosome 21q22.2[390, 391]. Research into the contribution trisomy of SIM2 has on DS-like phenotypes has resulted in conflicting observations, with one group reporting a significant decline in spatial learning with no effect on anxiety-related behavior, while another group reported an increase in anxiety-related behavior, reduced pain tolerance, and no change in spatial learning with overexpression of SIM2[388, 389]. It should be noted that these two groups achieved SIM2 overexpression in vastly different ways, with Chrast et al. having a more biologically relevant trisomy of SIM2 achieved with the introduction of a yeast artificial chromosome containing SIM2[389]. In contrast Ema et al. drove *SIM2* overexpression via a cytomegalovirus promoter, yielding unrealistically high expression of *SIM2*[388]. As such it is likely that the finding of increased anxiety-related behavior and decreased pain tolerance are more realistic findings.

Besides obvious learning deficiencies and craniofacial abnormalities, individuals with DS also possess a unique tumor profile. A high incidence of leukemia in individuals with DS was first recorded in 1957 and has since been well documented[392-395]. Interestingly, these individuals also present with more uncommon forms of leukemia, such

as myelodysplastic syndrome and acute megakaryocytic leukemia[393, 396-398]. A study of the Danish Cytogenetic Register found that leukemia accounts for 97% of malignancies in children under the age of 15 with DS, and accounts for 60% of malignancies in the total DS population[399].

While individuals with DS are predisposed to leukemia, trisomy of the DSCR appears to have a protective effect against many solid tumors[399]. Of those tumors observed within individuals with DS, breast cancer was the least common, suggesting that there may be a gene in this region that acts as a tumor suppressor in the breast[399]. Analysis of differential gene expression in individuals with DS has shown that SIM2 is one of the few factors to be upregulated in DS, showing high variability in numerous tissues[400, 401]. This has led to numerous studies investigating the role of SIM2 in tumor progression[378, 402-408]. Meaningfully, the majority of these reports show that increased expression of SIM2 is linked to highly aggressive tumors, with studies looking at both isoforms of SIM2 in prostate cancer showing that increased SIM2 expression is associated with a poorer outcome[403, 406, 408]. Similar studies in pancreatic cancer and colon cancer have also demonstrated that in both cell culture and xenograft models, inhibition of SIM2 stops cell growth and can even lead to the induction of apoptosis[404, 405].

Although individuals with DS have a global decrease in solid tumor incidence, this does not apply to all solid tumors, with some tissues exhibiting a higher incidence of carcinogenesis[399]. Key among these is pancreatic cancer, with pancreatic cancer being the 5th leading cause of cancer-related death in individuals with DS[409]. As normal pancreatic tissue does not express *SIM2*, it is unsurprising that its dysregulation as a result of trisomy of the DSCR leads to tumor progression[405]. This same phenomenon, however, cannot so

easily explain the slight decrease in colon cancer incidence in individuals with DS, another tissue which shows low *SIM2* levels in normal tissue[410].

Amongst tissues that exhibit a substantial decrease in cancer incidence in patients with DS, breast cancer is key among them[399]. Unlike pancreatic tissue and the colon, breast or mammary tissue typically expresses *SIM2*, with *SIM2s* being the major isoform present[381]. Previously, our laboratory has shown that the *SIM2s* plays a pivotal role in mammary gland development, pushing lactogenic development and pushing for mammary epithelial differentiation[378, 411]. In sharp contrast to the aforementioned SIM2-associated cancers, our laboratory has shown that *SIM2* is down-regulated in primary breast cancer, and that loss of *SIM2s* expression is associated with an epithelial mesenchymal transition (EMT) both in normal breast and breast cancer cell lines[349, 378, 381, 407, 411-413]. This EMT is punctuated by a decrease in E-cadherin and K18 expression as well as an increase in N-Cadherin, Vimentin and matrix metalloprotease (MMP) levels[407]. In addition, using a xenograft model, mice injected with breast cancer cells containing *shSIM2* showed a significant increase in the frequency of lung metastasis, showing that loss of *SIM2* in the mammary cells is associated with a significantly more invasive phenotype[413].

Although research has been conducted into the role of SIM2s in numerous cancers, the mechanism behind SIM2s activation has remained elusive[378, 402, 406]. One interesting avenue that has not been explored is the relationship of SIM2s and DDR. Fascinatingly there have been studies showing a role of HIF in DDR. Most notable amongst these was the finding that HIF1α expression leads to a decrease in BRCA1 while also directly enhancing Ku transcription, effectively promoting NHEJ[414-416]. With so many DDR factors being associated with breast cancer progression, and the parallels between HIF and SIM2

we hypothesized that the role SIM2s plays in prevention of breast cancer progression could be due to SIM2s aiding in the maintenance of genomic stability. Here, for the first time, we show a novel role for SIM2s activation in response to DNA damage and genomic stress, while regulating the recruitment of factors necessary for HR.

CHAPTER II

METHODS

Cell Culture

MCF7 and SUM159 cells were from American Type Culture Collection (ATCC) and were maintained in accordance with their guidelines. MCF10-DCIS.com cells were generously donated by Dr. Dan Medina (Baylor College of Medicine, Houston, TX, USA).

Primary Mammary Epithelial Cell (MEC) Isolation

Primary MECs were isolated from the #3, #4, and #5 mammary gland tissues and placed in wash buffer (1 × DMEM/F12 (Life Tech), 5% FBS (Atlanta Biological), 50 µg/mL (Life Technologies)) and mechanically homogenized with #10 scapels (Feather). Glands were then placed in 2 mg/mL Collegenase A (Roche) in wash buffer and incubated at 37°C with shaking for ~1.5 hours. Organoids were pelleted at 600 × g for 10 minutes, and supernatant was aspirated. Free nucleic acids were then digested with DNAseI treatment (100 µg/mL DNAse (Sigma), DMEM/F12). Organoids were washed in wash buffer 4-times and subsequently pelleted by pulse spinning at 450 × g. Organoids were then digested in 1 mg/mL trypsin (Life Technologies) at 37°C for ~20 minutes before being brought up to 10mL in growth media (DMEM/F12, 10% FBS, 100 units/mL penicillin/streptomycin (Life Technologies), 5 µg/mL insulin (Sigma), 50 µg/mL gentamicin (Life Technologies), 1 µg/mL hydrocortisone, 10 ng/mL mouse epidermal growth factor (EGF; Life Technologies) and single cells were pelleted at 450 × g for 3 minutes. MECs were washed twice more in growth media and pelleted again. MECs were finally plated on 10 cm tissue culture dishes and cultured at 32°C and 5% CO_2.

Generation of Cell Lines

Point mutations in the *SIM2* gene were generated via long cDNA synthesis (Invitrogen). Plasmids were amplified using Subcloning Efficiency™ DH5α™ competent cells (Life Technologies). Plasmid DNA was isolated using the HiPure Plasmid Maxiprep kit (Life Technologies) or the ZymoPURE Plasmid DNA Isolation Kit (Zymo Research). 10 μg of plasmid was mixed with GeneJuice (EMD Millipore) in 1 mL of Opti-MEM (Life Technologies) and incubated at room temperature for 15 minutes. This mixture was then added onto Phoenix-AMPHO lentiviral packaging cells (ATCC). Cells were incubated for 24 hours at 32°C and 5% CO_2. Media was collected and filtered through a 0.45 μm filter. The recommended amount of Sequabrene (Sigma) was added to the filtered media. The media was then added to the indicated target cells in 6-well plates. Plates were centrifuged at 200 × g for 60 minutes and allowed to incubate overnight at 32°C and 5% CO_2. Media was again collected from the packaging cells the next day, and target cells were transduced a second time, as described above. Puromycin selection (2 μg/mL) was started the following day and maintained for at least a week[407].

shSIM2 Sequences

The *shSIM2* was generated by inserting 5' - GAT CCG GTC GTT CTT TCT TCG AAT TTC AAG AGA ATT CGA AGA AAG AAC GAC CTC TTT TTT GGA AA-3' into pSilencer U6-retro 5.1 shRNA vector (Ambion) and control cells (*pSIL*) were generated by inserting a nonspecific scrambled sequence into the same vector. Plasmids were then packaged into lentivirus using Phenix HEK293-Ampho packaging cells as previously described[407]. The specificity of this *shSIM2* was confirmed by generating a second *shSIM2* in the same manner, with the new sequence being 5'-GAT CCG GTC ACC ACC AAA

TAC TAC TTC AAG AGA-3' and packaging into pSilencer U6-retro 5.1 shRNA vector (Ambion)[411].

drGFP Homologous Recombination Assay

Plasmids were amplified and harvested as above. 8 μg of drGFP (Addgene) and pCBA-SCEI (Addgene) were mixed with GeneJuice (EMD Millipore) in Opti-MEM (Life Technologies) and incubated at room temperature for 15 minutes. The mixture was added to HEK293 cells and incubated at 37°C for 2 hours. Cells were trypsinized and pelleted prior to resuspension in PBS. The fraction of GFP reversion was assessed by flow cytometry (BD Acurri C6). Ten thousand events were counted.

RNA Isolation and Real-Time qPCR (RT-qPCR)

RT-qPCR was conducted as previously described[413]. In brief, total RNA isolation was performed using the High Pure RNA Isolation Kit for tissue or for cells (Roche) and following manufacturer's instructions. Reverse transcription was performed using 1 μg of RNA template with the iScript cDNA Synthesis Kit (Bio-Rad). Quantitative PCR was performed on 50 ng cDNA using goTaq (Promega) in a CFX384 (BioRad). Data were normalized to β-actin. Relative expression analysis was conducted using the ΔΔCt method relative to lowest expression. Primers can be found in Table 1.

Clonogenic Survival Assay

Cells were plated at equal numbers on 60 mm tissue culture dishes and irradiated using a RS 2000 (Rad Source) at indicated doses. Cells were immediately washed with PBS, trypsinized (Life Technology), pelleted at 200 × g (Eppendorf 5710R), and resuspended in growth media. Viable cells were counted using a Cellometer Auto 1000 (Nexcelom Bioscience LLC) prior to plating 500 cells/plate on 10cm tissue culture

Table 1 RT-qPCR Primer List.

	Forward	Reverse
53BP1	AAGCCAGGCAAGAGAATGAGGC	GGCTGTTGACTCTGCCTGATTG
ACTB	GCAACGAGCGGTTCCG	CCCAAGAAGGAAGGCTGGA
ATM	CAGGGTAGTTTAGTTGAGGTTGACAG	CTATACTGGTGGTCAGTGCCAAAGT
β-2-Globulin	CGCTCCGTGGCCTTAGC	AATCTTTGGAGTACGCTGGATAGC
β-Actin	GTTTGAGACCTTCAACACCCC	GTGGCCATCTCTTGCTCGAAGTC
BRCA1	AGATGTGTGAGGCACCTGTGG	CACTCTAAGCTCCTGGCACTGGTAGAG
BRCA2	GTTCCCTCTGCGTGTTCTCA	CCATCCACCATCAGCCAACT
CDH1	CACAGACGCGGACGATGAT	GATCTTGGCTGAGGATGGTGTAA
MMP1	GATGGGAGGCAAGTTGAAAA	CTGGTTGAAAAGCATGAGCA
MMP3	TTCCTGATGTTGGTCACTTCAGA	TCCTGTATGTAAGGTGGGTTTTCC
MMP9	TGGGCAAGGGCGTCGTGGTTC	TGGTGCAGGCGGAGTAGGATT
RAD51	GGGAATTCTGAAAGCCGCTG	CCTGGCTTACGCTCCACTTC
SIM2s	AAGGTGGGCGGATCACCT	CAGCTTGTTGGCAGGCTTG
TP53	CCCCTCCTGGCCCCTGTCATCTTC	GCAGCGCCTCACAACCTCCGTCAT

dishes. Cells were incubated at 37°C and 5% CO_2 for 7 days and then washed with PBS before being fixed with 4% paraformaldehyde (Santa Cruz). Cells were stained with 0.01% crystal violet (Sigma-Aldrich) and washed twice with H_2O. Dishes were imaged using a ChemiDoc MP (Bio-Rad), and colonies were counted using ImageJ software. Percent survival was calculated as previously described[417]:

$$\frac{Number\ of\ colonies\ formed\ after\ treatment}{Number\ of\ cells\ seeded\ \times \left(\frac{Number\ of\ control\ colonies\ formed}{Number\ of\ cells\ seeded}\right) \times 100\%}$$

Survival Assay

10,000 cells were plated into 60 mm plates and dosed with the indicated concentrations of Olaparib (Cayman Chemical Company). Cells were incubated at 37°C

and 5% CO_2 for 4 days before being trypsinized and counted on a Cellometer Auto 1000 (Nexcelom Bioscience LLC).

Immunoblotting and Zymography

Cells were washed with cold PBS and lysed in high salt buffer (50 mM HEPES, 500 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 1% Triton X-100, pH 7.5) supplemented with 1 mM Na_3VO_4 (Sigma) and 1 mM complete ULTRA tablets mini EDTA-free Easy pack (Roche). Protein concentration was determined using the DC Protein Assay (Bio-Rad) with BSA as a standard. Equal amounts of protein were mixed with 6× sample buffer (375mM Tris-HCl, pH 6.8, 9% SDS, 50% glycerol, 0.03% bromophenol blue, 10% β-mercaptoethanol), warmed at 95°C for 5 min, and loaded on 6% or 10% SDS-PAGE gels. After electrophoretic separation, proteins were blotted onto PVDF membrane (Bio-Rad), blocked with 5% non-fat milk (Walmart) in TBST, and incubated at 4°C overnight with the indicated antibodies in TBST. Antibody sources are summarized in Table 2. Membranes were then washed and incubated with secondary anti-mouse or rabbit antibodies. Blots were washed again and imaged on a ChemiDoc MP (Bio-Rad) after incubating in ProSignal Pico ECL Spray (Genesee Scientific) for 3 minutes.

Table 2 Antibody List.

Target	Manufacturer	Product Number	Dilution	Application
α-Tubulin	Thermo Fisher Scientific	A11126	1:500	WB
ATM	Abcam	AB32420	1:1000	WB
β-Actin	Cell Signaling Technology	3700S	1:5000	WB
BRCA1	Abcam	AB131360	1:200 1:500	IF WB
BRCA1	Abcam	AB16780	-	CO-IP

Table 2 Continued

Target	Manufacturer	Product Number	Dilution	Application
BrdU	BS-bioscience	347580	1:25	IF
BrdU	Abcam	Ab6326	1:400	IF
E-Cadherin	Abcam	AB15148	1:100	IHC
FLAG	Cell Signaling Technology	8146S	1:500 1:200	WB IF
Keratin 14	BioLegend	905301	1:100	IHC
Lamin B1	Cell Signaling Technology	13435S	1:500	WB
p21/Waf1/Cip1	Cell Signaling Technology	2947S	1:1000	WB
p53	Abcam	AB26	1:1000	WB
p53BP1 (s1778)	Cell Signaling Technology	2675S	1:200	IF
pATM (s1981)	Abcam	AB81292	1:1000 1:1000	WB IF
p-p53 (s15)	Abcam	AB1431	1:1000	WB
pSerine	Abcam	AB9332	1:1000	WB
RAD51	Abcam	AB133534	1:200 1:500	IF WB
RAD51	Abcam	AB1837	-	CO-IP
RPA	Abcam	Ab2175	1:200	IF
SIM2	Aviva	ARP38551_P050	1:500	WB
SIM2	Millipore	AB4145	1:1000	WB
γH2AX (s139)	Abcam	AB2893	1:200	IF
γH2AX (s139)	Cell Signaling Technology	2577S	1:500	IF
Alexa-488 goat anti-rabbit IgG	Life Technologies	A11034	1:1000	IF
Alexa-568 goat anti-mouse IgG	Life Technologies	A11004	1:1000	IF
Anti-Mouse 2^{O}	Cell Signaling Technology	7073	1:2000	WB
Anti-Rabbit 2^{O}	Cell Signaling Technology	7074	1:2000	WB

ATM Inhibition

Indicated cells were pre-treated with 10 μM KU55933 (Cayman Chemical Company) for 2 hours before being dosed with 2GYs ionizing radiation (IR). Protein was harvested from plates as described above and analyzed via immunoblotting.

Co-Immunoprecipitation

For co-immunoprecipitation of FLAG, cells were lysed in high salt buffer 12 hours after treatment. 2 g of protein lysate were mixed with 200 μl of anti-FLAG M2 beads (Sigma, M8823) or IgG control beads (Cell Signaling, 5873S) after equilibrating the beads, IP was conducted according to the manufacturer's instructions β-mercaptoethanol was added prior to boiling for an additional 3 minutes and immunoblotting.

Co-immunoprecipitation of all other factors were conducted as follows: all steps were done on ice or at 4°C. All beads were washed 3X with five volumes TBS before use. Cells were lysed in RIPA buffer containing 1 mM Na_3VO_4 (Sigma) and 1 mM complete ULTRA tablets mini EDTA-free Easy pack (Roche) and agitated for 30 minutes prior to centrifugation at 10,000 X g for 10 minutes. Protein concentrations were determined via DC protein assay (Bio-Rad), and 100 μg of protein was added to IgG control beads (Cell Signaling, 5873S or 8726S) or 6 μg of the indicated antibody before incubating overnight. Magnetic beads (Active Motif, 53033) were then added to the antibody/protein mixture and allowed to incubate for an additional 4 hours. Tubes were then placed on a magnetic separator, and beads were washed 3X with TBS before being resuspended and boiled for 5 minutes in 2X Laemmli sample buffer lacking reducing agent. β-mercaptoethanol was then added, and samples were again boiled for 5 minutes before immunoblotting.

Immunofluorescent (IF) Staining of Cells

For IF assays, cells were fixed with 2% paraformaldehyde (Santa Cruz) for 15 min, permeated with 0.2% Triton X-100 (Sigma) for 15 min, and then blocked in 5% BSA (Fisher Bioreagents) in TBST at 4 °C overnight. The cells were incubated with indicated primary antibodies (Table 2) for 3 hours. After washing with TBST three times, samples were incubated the indicated secondary antibodies for 1 hour at room temperature and then stained with Hoechst 33342 (Life Technologies). Coverslips were mounted onto slides with ProLong Gold Antifade Mountant (Life Technologies). All images were taken using a using a Zeiss 780 confocal microscope[407].

Xenografts

Female Nu/Nu mice (4-6 months of age) were purchased from Jackson Laboratory and singly housed. A sample size of 5 mice per cohort has been calculated to achieve 80% power to detect differences of 1.4 at a p-value of 0.05. The indicated cells were mixed 1:1 with Matrigel (Corning) and kept on ice. Seventy-five thousand cells were then injected into each flank of nude mice. Mice were monitored daily for tumor growth and were euthanized once tumors reached a critical size, in accordance with IACUC procedures. During all subsequent experiments, scientists were blinded, and only made aware of the groups after data had been collected.

Immunostaining Tissue Sections

Tissues collected from mice were preserved in 4% paraformaldehyde and submitted to the Texas A&M University Veterinary Medicine & Biomedical Sciences Histology Laboratory for paraffin embedding, tissue sectioning, mounting, and H&E staining. Unstained sections were processed for immunostaining by first fixing sections to slides at

60°C for 30 minutes. The sections were then rehydrated by washing slides twice for 5 minutes in xylenes, twice for 3 minutes in 100% ethanol (EtOH), once for 3 minutes in 95% EtOH, once for 3 minutes in 70% EtOH, and finally once for 3 minutes in PBS. Antigen retrieval was performed in a 10mM sodium citrate solution (pH 6.0) inside a pressure cooker (Deni 9700) for 5 minutes on high. After a 5-minute wash in PBS, peroxidases were blocked in 3% hydrogen peroxide for 6 minutes. Slides were briefly rinsed in PBS before being blocked for 45 minutes in PBS-T containing 10% donor horse serum (Atlanta Biologicals). Primary antibodies (Table 2) were diluted in blocking solution and incubated on sections overnight at 4°C. Slides were then washed in PBS-T for 10 minutes before being incubated for 45 minutes in biotinylated secondary (anti-Mouse IgG, BMK2202, Vector or anti-Rabbit IgG, BA1000, Vector) antibodies diluted at 1:250 in PBS-T containing 1% donor horse serum. Sections were then incubated for 30 minutes in Vectastain ABC solution (PK-6200, Vector), according to the manufacturer's protocol. After rinsing for 5 minutes in PBS, sections were stained with DAB (SK-4100, Vector). Sections were washed for 5 minutes in water before being counter-stained in methyl salicylate for 3 minutes. Sections were then dehydrated by washing for 5 minutes in 95% EtOH, 100% EtOH, and finally in xylenes. Coverslips were attached with Permount (Electron Microscopy Sciences). Sections were imaged using a Zeiss Axio Imager.Z1 with 40× plan-apochromat objective or a Zeiss 780 confocal microscope. Quantification of nuclear intensity was done in ImageJ[378].

Tissue Micrographs

Tissue micrographs containing DCIS and DCIS with localized IDC were provided in collaboration with Dr. Fariba Behbod (University of Kansas Medical Center, Kansas

City, KS). ER, PR, HER2, p53, and Ki67 status of patients was known. Following placement in preservation media (LiforCell, Lifeblood Medical, Inc.), biopsy tissue was stored at 4°C until processing, as previously described[418, 419].

Cell Migration and Invasion

Invasion and migration assays were then conducted as previously described[407]. In brief, 25,000 cells suspended in serum-free media and were plated in FluoroBlok Cell Culture Inserts (Corning) that either contained Matrigel (invasion) or did not (migration). Cells were then incubated at 37°C and 5% CO_2 for either 20 hours (invasion) or 6 hours (migration). Cells that passed through the membrane were then stained with Hoechst 33342 (Life Technologies), imaged using a Zeiss Axio Imager.Z1 with 40× plan-apochromat objective, and then enumerated.

DNA Fiber Analysis

DNA combing assays were performed as previously described using IdU (Sigma) and CldU (Sigma) with the indicated modifications in time-points[420]. In brief, the cells were dosed with the indicated reagents (IdU, CldU, HU, DMSO), at the indicated dosages, for the indicated amounts of time depending on the experiment being conducted (Figure 19). Cells were then washed with PBS and trypsinized and collected in a 15 mL conical tube before being washed again with ice-cold PBS, brought to a concentration of 400 cells/ µL and placed on ice. 2 µL of cells were then pipetted onto a charged microscope slide and allowed to dry almost completely. 15 µL of lysis solution (0.2 M Tris pH 7.4, 50 mM EDTA, 0.5% SDS) was added and slides were incubated at room temperature for 10 minutes. Slides were then tilted to a 25° angle, allowing DNA fibers to run down slide, and

allowed to dry completely. DNA was then fixed in a 3:1 methanol:acidic acid solution for 2 minutes, and then removed and allowed to dry overnight.

The next day slides were placed at -20°C and incubated for a minimum of 24 hours before proceeding to the next step. Slides were then treated with 2.5 M HCl for 30 minutes, washed with 0.1% PBST (PBS-Tween) for 3 minutes, and then incubated in PBS two times for 3 minutes. Slides were blocked in 5% BSA (bovine serum albumin) for 30 minutes. DNA was then probed with the indicated primary antibodies for 1 hour, before washing 2 times with PBS for 3 minutes each (Table 2). Finally, secondary antibody was added and incubated for 1 hour (Table 2). Slides were washed 2 more times in PBS for 3 minutes and then images were captured using a Zeiss 780 confocal microscope, and fiber lengths were measured in ImageJ.

Anaphase Bridges

Cells were maintained at 37°C and 5% CO_2. First, cells were synchronized using a di-thymidine block. Briefly, cells were incubated in 2 mM thymidine (Cayman Chemical) for 19 hours, washed, and cultured again in normal media for 9 hours. Afterwards, 2 mM thymidine was reapplied for an additional 17 hours. Cells were washed again, and normal media was added for 9 additional hours. Finally, cells were fixed with 4% paraformaldehyde (Santa Cruz) and stained with Hoescht 33342 (Life Technologies). Images were captured using a Zeiss 780 confocal microscope.

Statistical Analysis

To address scientific rigor, all experiments in cell lines and xenografts were conducted in biological triplicates at a minimum, technical duplicates, and repeated three times. Normal distribution was confirmed before conducting two-tailed Student's t-test.

Likelihood ratio and Pearson statistical test were used for goodness of fit comparisons. Significance was considered at $p<0.05$.

Study Approval

Animal studies were approved by the Texas A&M University Laboratory Animal Care Committee. Patients gave written informed consent for participation in this University of Kansas Medical Center Institutional Review Board approved study, which allowed collection of de-identified surgical tissue for research.

CHAPTER III

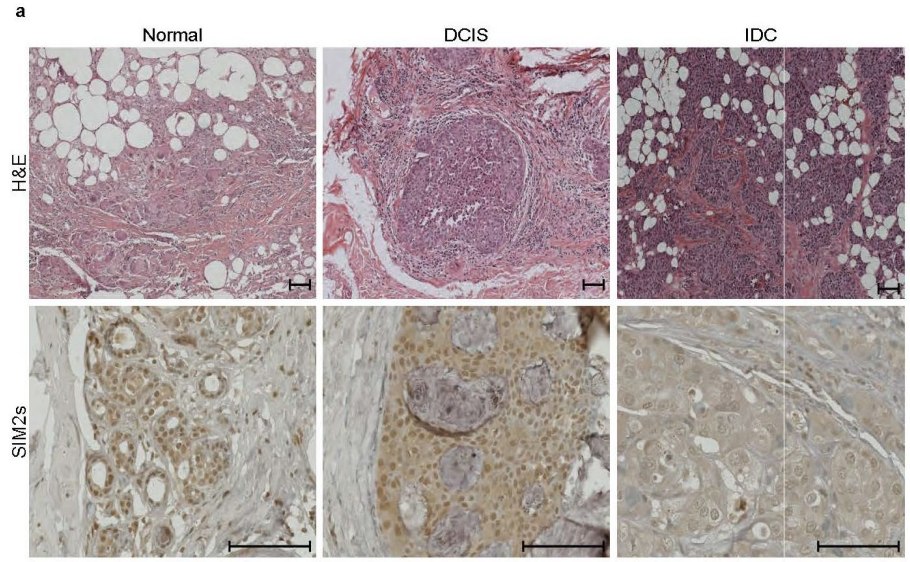
ATM-DEPENDENT ACTIVATION OF SIM2S REGULATES HOMOLOGOUS RECOMBINATION AND EPITHELIAL-MESENCHYMAL TRANSITION*

SIM2s Expression is Lost During Progression to IDC

To investigate the expression of SIM2s during progression from normal breast tissue through DCIS to IDC, we analyzed human tissue microarrays containing normal breast, DCIS, DCIS with local invasion, and IDC for differences in SIM2s by IHC in 27 patients diagnosed with DCIS or IDC. The results showed high, nuclear expression of SIM2s in normal tissue. Although SIM2s expression remained high in DCIS, a shift towards cytoplasmic localization was apparent. Interestingly, SIM2s loss was apparent in sections containing IDC (Figure 8a). Statistical analysis of SIM2s staining revealed a significant correlation between SIM2s expression and the state of breast disease (Normal, DCIS, or IDC) ($p<6.7E^{-07}$) (Figure 8b). SIM2s expression significantly correlated with ER ($p<0.0239$) and PR ($p<0.0156$) expression in DCIS samples, but not HER2 ($p<0.1377$), indicating a relationship between SIM2s and luminal stage breast cancer. SIM2s expression was also significantly correlated with p53 expression ($p<0.0095$) and inversely related to Ki67 expression ($p<0.0124$) in DCIS samples (Figure 8c). Importantly, loss of SIM2s correlated with increased micro-invasion and metastasis, supporting a role for SIM2s in inhibiting breast cancer progression (Figure 8d; $p<0.0079$).

---

*Reprinted with permission from "ATM-dependent activation of SIM2s regulates homologous recombination and epithelial–mesenchymal transition" by Pearson et al., 2019. *Oncogene*, Volume 38, 2611-2626, Copyright 2019 by Springer Nature.



b

Type	SIM2s Expression: Nuclear	Both	Cyto	None
Normal	17	3	0	5
DCIS	1	6	7	5
IDC	0	1	5	4
Pearson	2.02E-05*			
Likelihood ratio	6.17E-07*			

c

	ER +	ER -	PR +	PR -	Ki-67 +	Ki-67 -	p53 +	p53 -
SIM2 Positive	9	2	8	3	1	4	0	5
SIM2 Negative	2	5	1	6	10	2	7	5
Pearson	0.0239*		0.0156*		0.0128*		0.026*	
Likelihood ratio	0.0219*		0.0119*		0.0124*		0.0095*	

d

Variable	SIM2s: Pos	Neg
Invasion		
Positive	1	9
Negative	6	3
Pearson	0.0106*	
Likelihood ratio	0.0079*	
Metastasis		
Positive	0	4
Negative	7	8
Pearson	0.0856	
Likelihood ratio	0.0386*	

Figure 8 Loss of SIM2s correlates with increased metastasis and invasion.
(a) Histological staining of human normal, DCIS, and IDC samples for H&E and SIM2s. Samples were imaged at 25× magnification. Scale bars, 100µm. (b) Statistical analysis of breast type with SIM2s expression. Tissue microarrays were analyzed categorically to compare the location of SIM2s staining versus the breast type. (c) Statistical analysis of SIM2s status correlated with ER, PR, Ki-67, and p53 status in DCIS specific pathology reports (n=18). (d) Prognosis of micro invasion and/or metastasis was compared with binomial SIM2s staining. Likelihood ratio and Pearson's chi-squared tests were performed to test correlations. *p-value<0.05. Reprinted with permission from[434].

SIM2s Regulation of the DDR Pathway

To determine the impact of SIM2s on DDR, we utilized breast cancer cell lines that express differing levels of endogenous *SIM2s* with SUM159 (low), MCF7 (high), and

MCF10-DCIS.com (DCIS.com-medium)[381, 413]. We modified *SIM2s* levels by over-expression (*SIM2s-FLAG*) or with a previously validated *shSIM2*[407]. DNA damage was induced with IR and colony formation assays were performed[381, 407, 413, 421]. The results show that sh*SIM2* sensitized cells to IR (Figure 9a, b), whereas over-expression of *SIM2s* had a protective effect in SUM159 cells (Figure 9c). Interestingly, there was no significant change in DCIS.com-SIM2s-FLAG cells, possibly due to the presence of endogenous SIM2s in control cells (Figure 9d).

To investigate this protective effect of SIM2s, we examined how loss of *SIM2s* affected genomic stability. As DNA damage is detected, ATM phosphorylates histone H2AX (γH2AX), which then forms foci at the site of DNA damage[422, 423]. We found that MCF7-*shSIM2* cells treated with 2GYs IR had an increase in basal γH2AX foci compared to control cells ($p<0.00815$) and recovered slower from DNA damage compared to control cells ($p<0.00163$; Figure 9e). Although the DCIS.com-*shSIM2* lacked the initial basal elevation of γH2AX foci, they did have significantly more γH2AX foci 6 hours after treatment with 2GYs IR ($p<0.00143$; Figure 9f). The delayed resolution of γH2AX foci has been previously associated with the ATM-dependent DDR within heterochromatin structures, where repair is both temporally and spatially separated from other DDR in order to prevent inappropriate cross-over events[424-426].

SIM2s is Phosphorylated by ATM and Stabilized in Response to IR

The bHLH-PAS family of transcription factors plays an important role in development and in response to environmental cues; however, mechanisms that activate SIM2s have remained elusive[363, 427]. To better understand the role SIM2s may play in DDR, we assessed SIM2s levels in DCIS.com-*SIM2s-FLAG* and MCF7 cells treated with

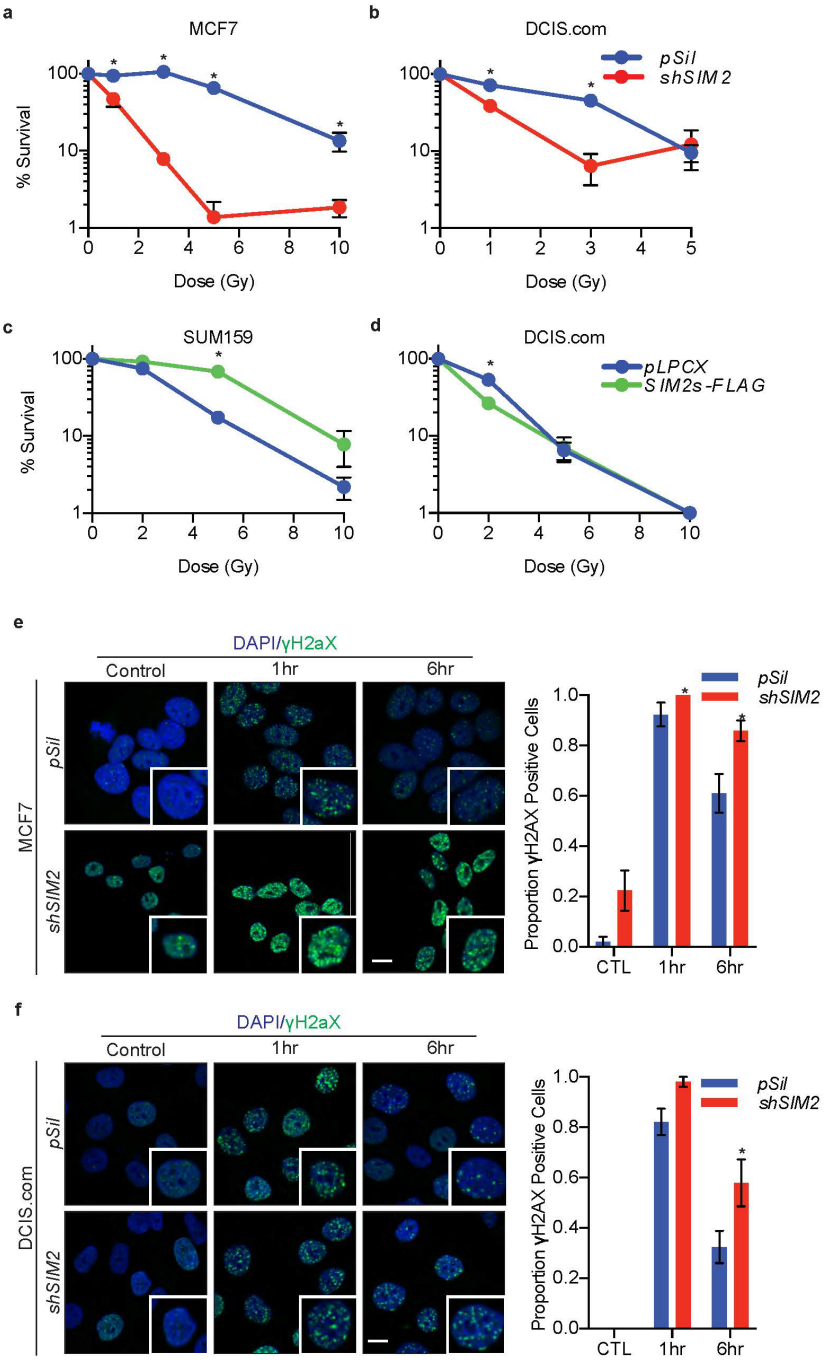


Figure 9 SIM2s increases cell survivability by reducing DNA damage.
(a-d) Clonogenic survival assays of (a) MCF7 and (b) DCIS.com cells containing *shSIM2* or control were treated with the indicated doses of irradiation. Cells were then plated and allowed to grow for 7 days before being counted. (c) SUM159 and (d) DCIS.com cells over-expressing *SIM2s-FLAG* or control were treated with the indicated doses of IR and sampled as before. (e) MCF7 and (f) DCIS.com cells containing *shSIM2* or control were treated with 2GYs IR or left untreated and fixed at the indicated time points. Cells were then immunostained for γH2AX, and cells containing more than 10 foci were counted as positive (n=50). Scale bars, 10 μm. Values indicate the mean ± SE with n=3 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

IR. IF analysis showed an elevated level of nuclear SIM2s-FLAG after treatment with 2GYs IR (Figure 10a). SIM2s activation was further confirmed by western blot analysis in MCF7 and DCIS.com treated with 2GYs IR (Figure 10b). The efficiency of our previously validated *shSIM2* construct was further confirmed in MCF7 and DCIS.com cells, with no SIM2s detected 12 hours after 2GYs IR treatment (Figure 10b)[407]. SUM159-*SIM2s-FLAG* and DCIS.com-*SIM2s-FLAG* cells also showed SIM2s stabilization after treatment with 2GYs IR (Figure 10c).

We next sought to determine if the increase in SIM2s levels was due to an up-regulation of *SIM2s* gene expression. Interestingly, in both normal MCF7 and DCIS.com cells, we observed no change in *SIM2s* levels after treatment with 2GYs IR (Figure 10d). SIM2s is ubiquitinated by the E3 ubiquitin ligase, PRKN, leading to its rapid proteasome-dependent degradation[361]. Proteasomal inhibition with 10 μM (R)-MG132 for 2 hours inhibits this pathway and resulted in increased SIM2s levels (Figure 10e). This led us to theorize that SIM2s could be stabilized post-translationally. To test this, we treated DCIS.com-*SIM2s-FLAG* cells with 50 μg/mL cycloheximide (CHX) to inhibit translation at the indicated time points after treatment with 2GYs of IR and harvested all samples 12 hours after IR treatment. No significant change in SIM2s levels was observed, supporting our hypothesis that SIM2s is stabilized post-translationally (Figure 10f).

To investigate the possibility that SIM2s is stabilized by a post-translational modification in response to IR, we analyzed SIM2s for phosphorylation sites of known DDR kinases and identified 12 total ATM consensus sites, three of which are highly conserved across *Mus Musculus, Homo Sapiens, and Xenopus laevis* (S35, S115, S363) (Figure 11a)[428]. To determine if ATM interacts with and phosphorylates SIM2s,

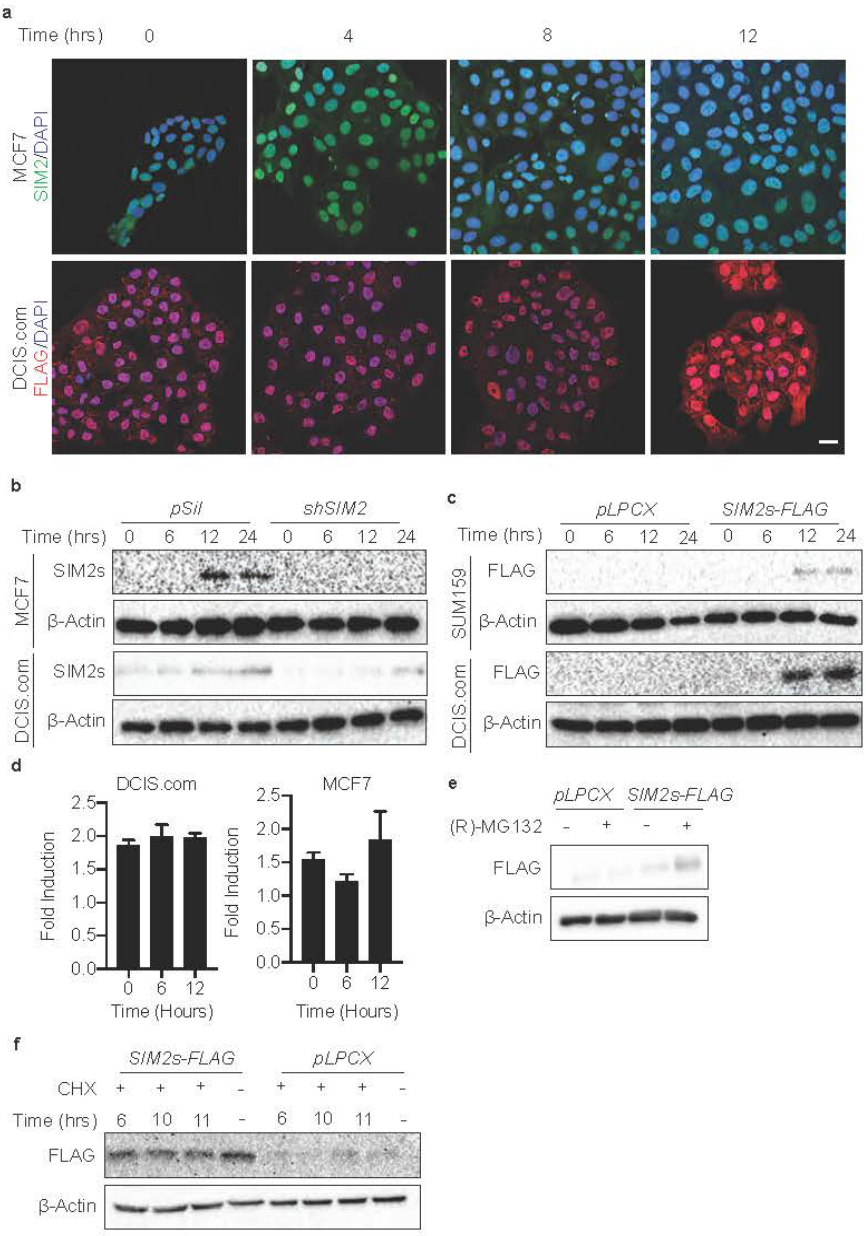


Figure 10 SIM2s is stabilized with IR treatment.
(a) WT MCF7 cells or DCIS.com cells over-expressing SIM2s-FLAG were treated with 2GYs of IR and fixed at the indicated time points before immunostaining for the presence of FLAG. Scale bar, 20 μm. (b) Efficiency of SIM2s knockdown by shSIM2 was confirmed in MCF7 and DCIS.com cells via western blot. Stabilization of endogenous SIM2s 12 hours after IR treatment was further confirmed via western blot. (c) The FLAG epitope was confirmed to not interfere with SIM2s stabilization in the DCIS.com and SUM159 cells over-expressing SIM2s-FLAG. (d) DCIS.com and MCF7 cells were treated with 2GYs of IR, and RNA was harvested at the indicated time points. RT-qPCR analysis of SIM2s indicates no change in SIM2s mRNA after treatment with 2GYs IR. (e) Western blot analysis of DCIS.com SIM2s-FLAG and pLPCX control cells treated for 2 hours with 10 μm (R)-MG132 or DMSO. (f) Confirmation of SIM2s stabilization at the protein level was confirmed in DCIS.com cells over-expressing SIM2s-FLAG. Cells were treated with 2GYs IR and then dosed with 50μg/mL of the translation inhibitor, cycloheximide (CHX), at the indicated time points before being harvested 12 hours post-IR. Values indicate the mean ± SE with n=3 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

DCIS.com-*SIM2s-FLAG* cells were treated with 2GYs of IR and assessed for nuclear co-localization of FLAG and pATM via IF (Figure 11b). This interaction was confirmed through immunoprecipitation of FLAG in DCIS.com-*SIM2s-FLAG* cells treated with 2GYs IR followed by immunoblotting for ATM (Figure 11c). Additionally, an increase in phospho-serine residues at the predicted molecular weight of SIM2s (~65 kilodaltons) was observed after treatment with 2GYs IR (Figure 11c). To determine if ATM is necessary for IR-induced SIM2s stabilization, we pretreated SUM159-*SIM2s-FLAG* cells with KU55933, a selective ATM inhibitor, for 2 hours prior to treatment with 2GYs of IR. Western blot analysis showed a significant reduction in SIM2s stabilization in KU55933 treated cells in response to IR (Figure 11d). These findings suggest that SIM2s is stabilized in response to IR through phosphorylation.

We next generated a DCIS.com-*SIM2s-FLAG* cell line with serine to alanine mutations at all the predicted ATM consensus sites (S35, S115, S203, S216, S309, S343, S352, S361, S363, S392, S393, S426; SIM2sΔ12). To test if the abrogation of these consensus sites effected SIM2s stabilization in response to IR, we treated SIM2sΔ12 cells with 2GY IR and isolated protein 12 hours later. Interestingly, there was no increase in SIM2s levels with these 12 sites mutated (Figure 11e).

To narrow down which serine residue is necessary for SIM2s stabilization we generated DCIS.com-*SIM2s-FLAG* cells lines containing point mutations at the three highly conserved ATM-consensus sites: serine 35 (S35A), 115 (S115A), or 363 (S363A). Utilizing these cells lines, we sought to determine if one of these mutations inhibited SIM2s stabilization. Immunoblot analysis found that S115A no longer showed SIM2s stabilization 12 hours after treatment with 2GYs IR (Figure 11e). As inefficient

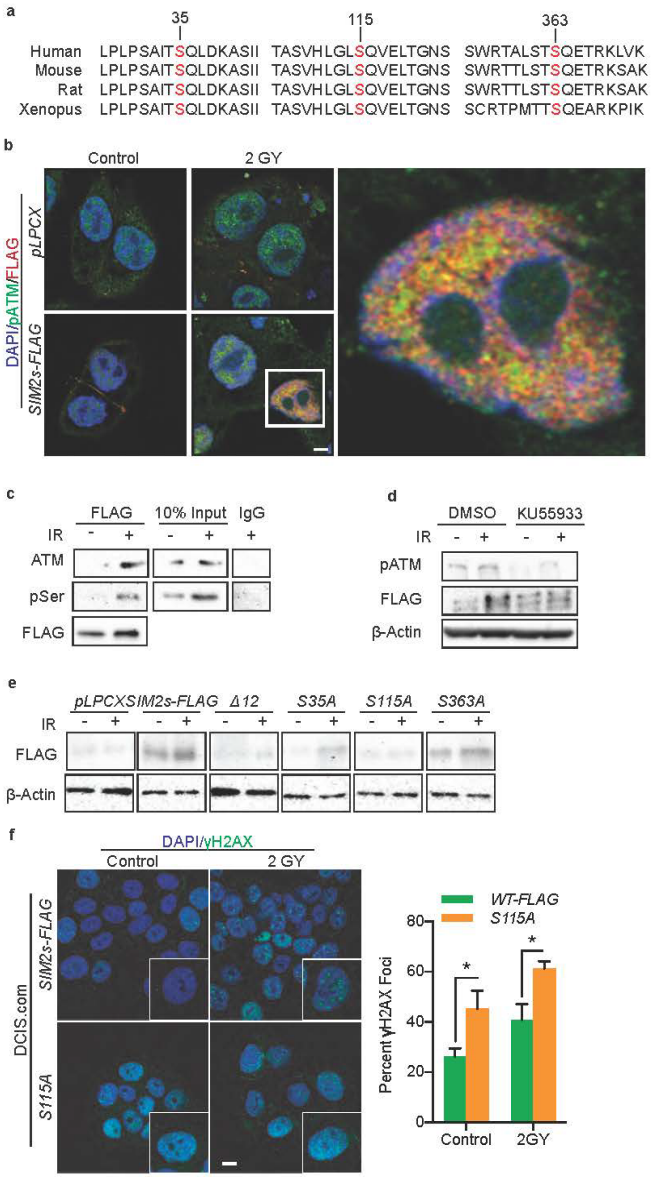


Figure 11 S115 is necessary for SIM2s stabilization.
(a) *SIM2s* was analyzed for ATM consensus sequences. Three highly conserved sites were found and confirmed with NetPhos 3.1[1]. (b) Co-localization of SIM2s and pATM 12 hours after IR treatment was confirmed in DCIS.com cells. Scale bar, 10 μm. (c) Immunoprecipitation of SIM2s-FLAG 12 hours after treatment with 2GYs IR in DCIS.com cells by a FLAG specific antibody shows SIM2s interacts with ATM in response to IR. In addition, an increase in pSer is seen in response to IR at the molecular weight of SIM2s-FLAG. (d) Treatment with 10 μm of KU55933 inhibited SIM2s stabilization in response to 2GYs IR. (e) Serine to alanine point mutations were generated in SIM2s-FLAG at S35, S115, S363, or at all predicted ATM consensus sites (Δ12) and then stably transfected into DCIS.com cells. Mutants were analyzed for SIM2s stabilization 12 hours post-treatment with 2GYs IR, with only the S115A mutant not responding to IR. (f) DCIS.com cells over-expressing *SIM2s-FLAG* or *SIM2s-S115A* were treated with 2GYs IR and fixed 6 hours later before being immunostained for the presence of γH2AX foci. Cells containing 10 or more γH2AX foci were counted as positive (n=50). Scale bar, 10 μm. Values indicate the mean ± SE with n=3 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

transduction of S115A could have led to these findings, we verified *SIM2s* RNA levels in these cells via RT-qPCR and found a significant increase in *SIM2s* RNA when compared to the *pLPCX* control cells (Figure 12). The interaction of SIM2s

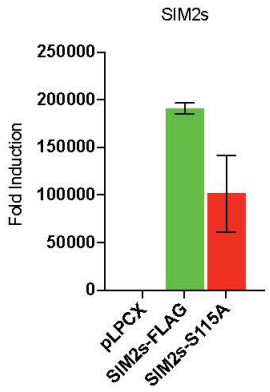


Figure 12 Verification of effective transduction SIM2s vectors.
DCIS.com cells were transduced with either *pLPCX-SIM2s-FLAG*, *pLPCX-SIM2s-S115A-FLAG*, or *pLPCX-empty.* RNA was isolated and then *SIM2s* levels were analyzed via RT-PCR. Values indicate the mean ± SE with n=3. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

and ATM, suggests that SIM2s may be acting downstream of ATM; however, western blot analysis revealed that loss of *SIM2s* does not affect p53 activation (data not shown).

The loss of stabilization observed with S115A led us to predict that cells containing this mutation would behave similarly to cells containing *shSIM2* and have impaired DDR. Comparing DCIS.com-*SIM2s-FLAG* to S115A revealed that cells containing S115A had significantly more γH2AX foci 6 hours after 2GYs IR than cells over-expressing *SIM2s-FLAG* (Figure 11f).

Loss of SIM2s Impairs Homologous Recombination

As SIM2s is a transcription factor we sought to determine if SIM2s regulates DDR gene expression. Interestingly, loss of *SIM2s* did not have a negative effect on *ATM, TP53, 53BP1, BRCA1, BRCA2,* or *RAD51* levels 6 hours after treatment with 2GYs IR (Figure 13). This finding led us to hypothesize that SIM2s directly interacts with proteins involved in DDR. Utilizing IF, we observed that SIM2s co-localizes with γH2AX in response to IR (Figure 14a). The localization of this interaction to the nuclear periphery is consistent with a role for SIM2s in aiding ATM repair of damaged heterochromatin via HR[429]. To further test this, we immunoprecipitated FLAG from IR treated DCIS.com-*SIM2s-FLAG* 12 hours after treatment and found that IR enhanced SIM2s interaction with BRCA1 (Figure 14b).

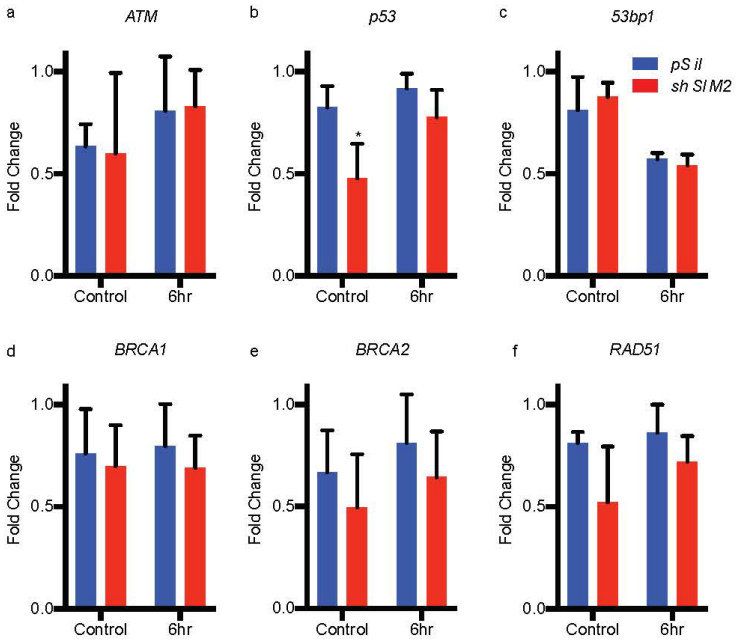


Figure 13 SIM2s does not affect common DDR gene expression.
(a-f) DCIS.com cells containing *shSIM2* or scrambled plasmid were dosed with 2GYs IR or mock, and RNA was harvested 6 hours later. RT-qPCR analysis was performed for the indicated mRNA (n=3). Values indicate the mean ± SE. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

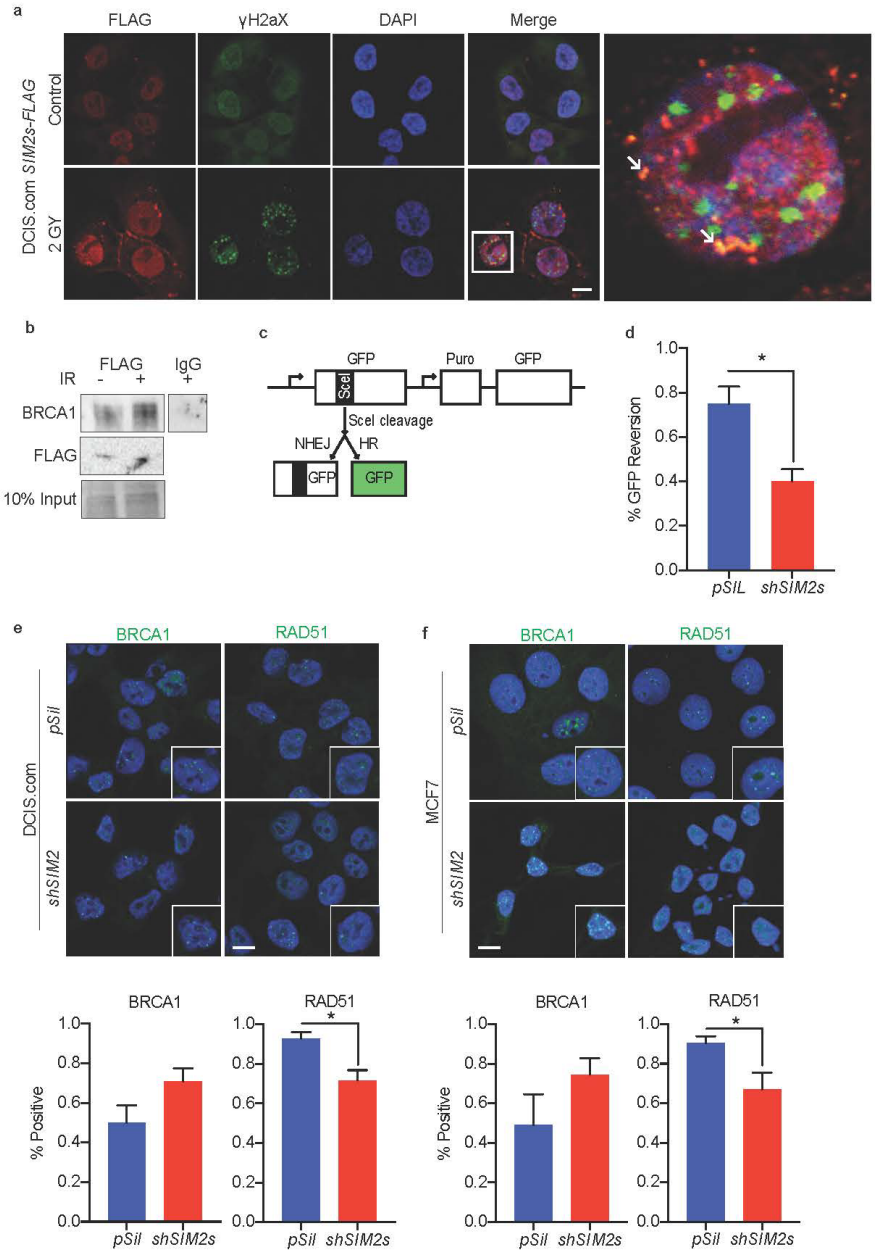


Figure 14 SIM2s is involved in HR and is necessary for RAD51 recruitment.
(a) DCIS.com cells over-expressing *SIM2s-FLAG* were treated with 2GYs IR and probed for γH2AX and FLAG 12 hours later. Arrows indicate sites of co-localization. Scale bar, 10 μm. (b) The interaction of SIM2s with BRCA1 was further assessed by immunoprecipitation of SIM2s with a FLAG specific antibody from DCIS.com *SIM2s-FLAG* cell lysate 12 hours after 2GYs IR or mock treatment. (c) Diagram of drGFP locus reporter plasmid structure. (d) The effect that loss of SIM2s has on HR was assessed using the drGFP reporter system in conjunction with a plasmid encoding SceI. The presence of GFP was analyzed via flow cytometry (10,000 events) and represented as the percentage of cells expressing GFP. (e) DCIS.com or (f) MCF7 cells containing a scrambled vector or *shSIM2* were treated with 2GYs of IR and probed for BRCA1 or RAD51 6 hours later (n=50). Cells containing 10 or more foci were counted as positive. Scale bars, 10 μm. Values indicate the mean ± SE with n=3 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

To investigate the effect of SIM2s on HR, we utilized a GFP reporter system to assay the efficiency of HR in cells containing *shSIM2*[430]. This system contains an expressed non-functional GFP with a SceI cleavage site and an unexpressed functional GFP. Upon treatment with SceI, cells capable of HR use the functional GFP template for repair. In this way, HR efficacy can be assessed by quantification of GFP activity in cells following transduction of cells containing drGFP with SceI (Figure 14c). Here, we employed this system to investigate whether depletion of *SIM2s* inhibits HR. Our results showed a significant decrease in GFP reversion in cells containing *shSIM2* compared to controls, suggesting that loss of *SIM2s* impairs HR ($p<0.022$; Figure 14d).

The dynamics that determine which DDR pathway is chosen have been well studied and depend upon the location of the break, the stage of the cell cycle, and the break type[431]. In instances where BRCA1 is lost, cells are no longer able to remove p53BP1 and a dramatic increase in p53BP1 foci is observed[176]. Interestingly, we observed no change in p53BP1 foci in cells containing *shSIM2* (Figure 15). This finding suggests that SIM2s may be acting downstream of the NHEJ/HR fate determining steps. In line with this hypothesis, we observed that loss of *SIM2s* leads to a significant decrease in RAD51 recruitment in DCIS.com ($p<0.012$) and MCF7 ($p<0.003$) cells after treatment with IR, whereas there was no change in BRCA1 recruitment (Figure 14e, f, Figure 16). Together, these data show that loss of *SIM2s* impairs HR by decreasing the efficiency of RAD51 loading in response to IR.

Loss of SIM2s Sensitizes Cells to Treatment with PARP Inhibitors

Recent studies have shown that loss of factors crucial for HR, including BRCA1, BRCA2, and RAD51, sensitizes cells to PARP inhibitors in pre-clinical and clinical

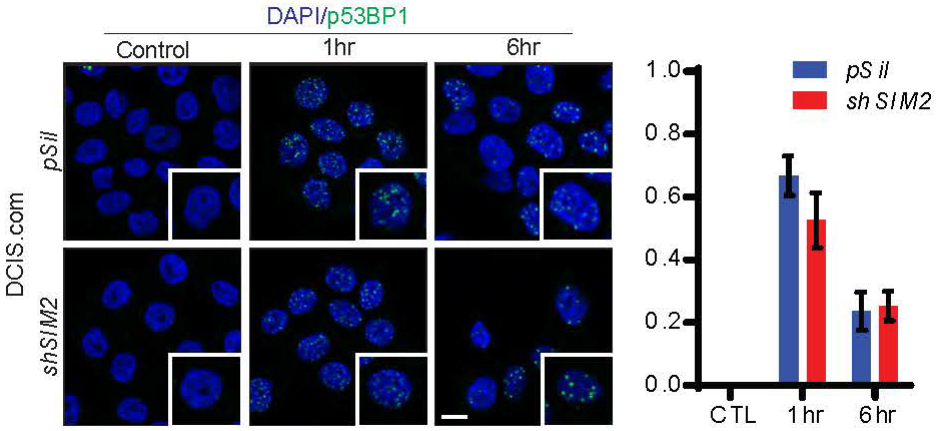


Figure 15 Loss of SIM2s does not affect p53BP1 foci.
DCIS.com cells containing *shSIM2* or control were treated with 2GYs IR or mock and fixed at the indicated time points. Cells were then immunostained for p53BP1, and cells containing more than 10 foci were counted as positive (n=50). Values indicate the mean ± SE. Scale bar, 10 µm. Student's t-test was performed to test significance. Reprinted with permission from[434].

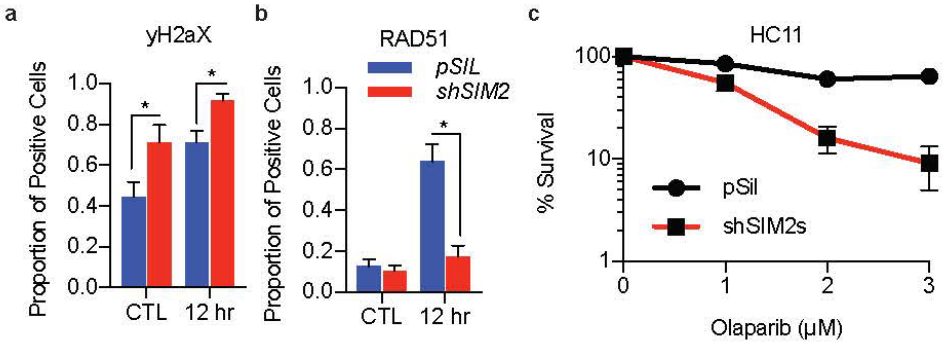


Figure 16 Verification of shSIM2 experiments with additional *shSIM2*.
HC11 cells containing *shSIM2* or control were treated with 2GYs IR or left untreated and fixed at the indicated time point. Cells were then immunostained for (a) γH2AX or (b), and cells containing more than 10 foci were counted as positive (n=50). (c) Percent cell survival in HC11 cells containing *shSIM2* or control with increasing doses of the PARP inhibitor Olaparib. Values indicate the mean ± SE with n=3 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

trials[328, 332, 335, 337, 432]. PARPs are necessary in early DDR to poly-ADP-ribosylate target histones, leading to the destabilization of chromatin structure and exposure of damaged DNA for invasion of DDR machinery[433]. This combination of loss of both single-stranded break repair systems as well as impaired DSB repair leads to the accrual of DNA damage that eventually overwhelms and kills the cell. As such, we hypothesized that loss of *SIM2s* would sensitize cells to treatment with PARP inhibitors. Consequently, we dosed *shSIM2* containing cells with the indicated doses of Olaparib, a potent PARP1/PARP2 inhibitor, and assayed for differences in proliferation[335]. Loss of *SIM2s* significantly reduced cell survival in a dose-dependent manner, further supporting a role for SIM2s in HR (Figure 16, Figure 17).

SIM2s Mutation at S115 Leads to EMT

We have previously shown that SIM2s is an inhibitor of EMT in normal breast and breast cancer cell lines and in the mouse mammary gland, which are further supported by our current findings (Figure 8)[381, 407, 412]. As we have identified a role for SIM2s in the recruitment of RAD51, we sought to determine if the S115A mutation impairs the EMT

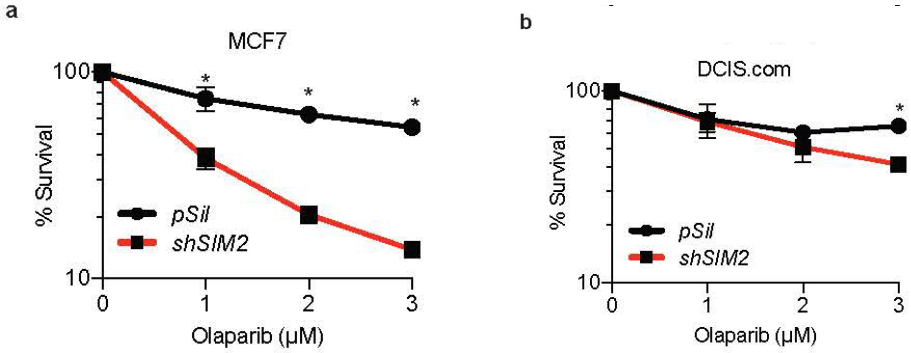


Figure 17 Loss of SIM2s sensitizes cells to synthetic lethal treatments.
Percent cell survival in (a) MCF7 and (b) DCIS.com cells containing *shSIM2* or control with increasing doses of the PARP inhibitor Olaparib. Values indicate the mean ± SE with n=3 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

inhibitory function of SIM2s. Using Boyden chamber assays, we found that the S115A mutant resulted in a significant increase in the ability of DCIS.com cells to not only migrate ($p<0.0009$), but also to invade ($p<0.0485$), through an extra-cellular matrix (Figure 18a, b). We have shown that the invasive phenotype observed with loss of *SIM2s* can be attributed to an increase in MMP activity[381, 407]. Here, we found that S115A cells have an increase in *MMP1* ($p<0.006$) and *MMP9* ($p<0.05$) expression (Figure 18c-e). This also correlated with an increase in MMP9 activity in these cells, as demonstrated by gelatin zymography (Figure 18f).

To determine if these characteristics translated into an increased invasive potential *in vivo*, control, DCIS.com-*SIM2s-FLAG,* or *S115A* were injected into the flanks of nude mice and monitored for tumor growth. Histological analysis of tumors revealed the *SIM2s-FLAG* tumors formed distinct lobular structures, whereas the S115A tumors exhibited a more invasive phenotype with increased areas of necrosis (Figure 18g). Staining of sections from *SIM2s-FLAG* tumors revealed a single layer of keratin 14 (K14) positive cells, whereas S115A tumors stained positive for K14 with no defined organization (Figure 18g). Further, we found that S115A tumors contained decreased E-cadherin compared to *SIM2s-FLAG*, suggesting induction of EMT (Figure 18g). RT-qPCR analysis of RNA isolated from tumors showed a significant increase of *CDH1* in *SIM2s-FLAG* ($p<0.02$) but not in S115A (Figure 18h) tumors. Analysis of tumor growth, showed decreased tumor growth in both *SIM2s-FLAG* ($p<0.0001$) and S115A ($p<0.002$) tumors, as compared to controls (Figure 18i). This finding is consistent with our previous studies, which showed that DCIS.com-shSIM2 xenografts leads to increased invasion and metastasis, independent of tumor growth[413].

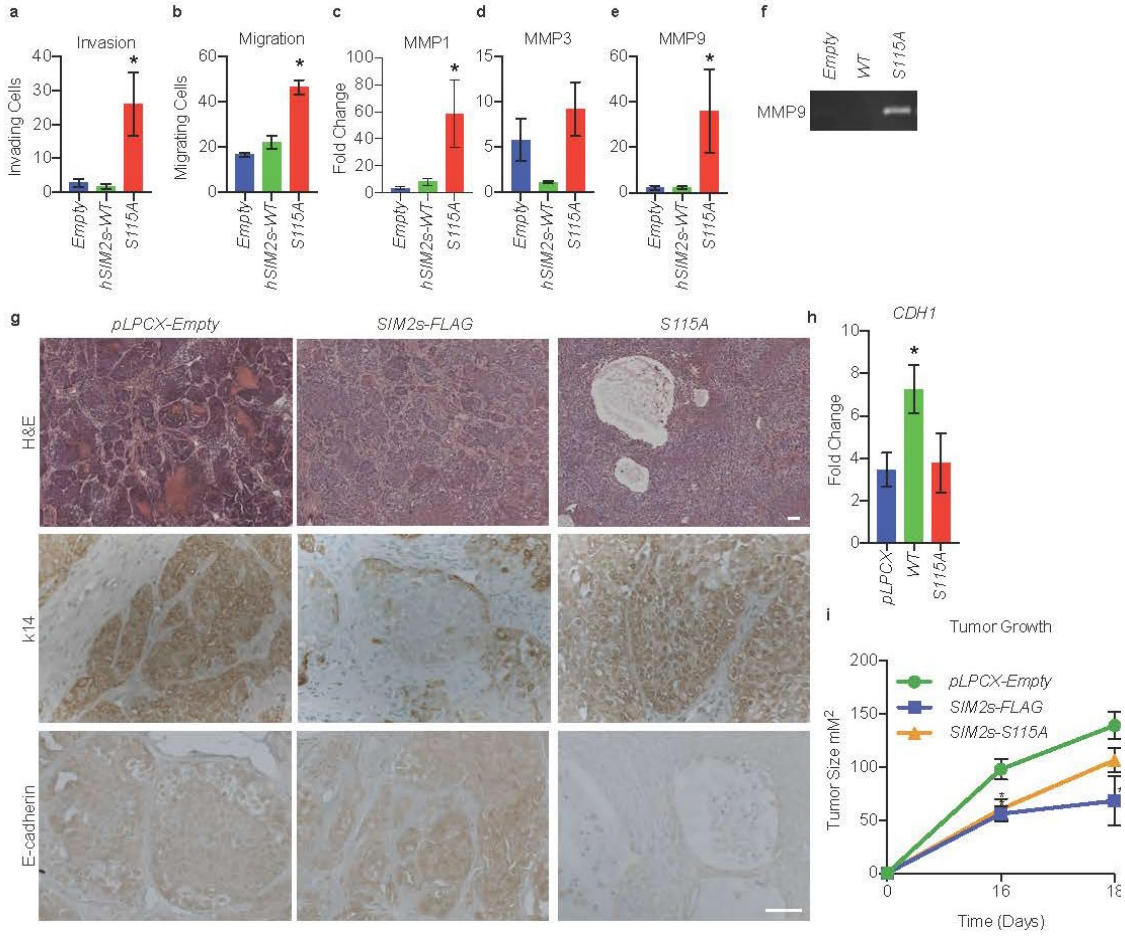


Figure 18 Mutation of SIM2s at S115 promotes basal marker expression and tumor metastasis. (a) DCIS.com cells containing the indicated constructs were plated in Boyden chamber inserts containing Matrigel and measured for invasive potential. (b) DCIS.com cells containing the indicated plasmids were plated in Boyden chamber inserts and measured for migratory potential. (c-e) DCIS.com cells containing the indicated constructs were harvested for RNA and analyzed for the expression of (c) *MMP1,* (d) *MMP3*, and (e) *MMP9* via RT-qPCR. (f) Matrix-metalloprotease activity was confirmed via gelatin zymograph in DCIS.com cells containing *SIM2s-FLAG*, *S115A*, or empty vector. (g) H&E staining of tumors isolated from the indicated mice. Scale bar, 100 μm. Tumor sections were stained for the indicated basal and luminal markers. Scale bar, 50 μm. (h) RNA was isolated from the largest tumor from each mouse and analyzed for *CDH1* expression via RT-qPCR. (i) Seventy-five thousand DCIS.com cells containing *SIM2s-FLAG*, *SIM2s-S115A,* or empty vector were injected into the flanks of nude mice, and tumor growth was measured. (j) Percent and (k) enumeration of mice positive for lung metastasis in the indicated tumors was measured by RT-qPCR analysis using a human specific β-2-globulin primer that does not cross-react with mouse. (l) Lungs were sectioned before being H&E stained. Top: slide scanned whole lung sections. Bottom: lung sections under 10X magnification. Scale bar, 200 μm. Values indicate the mean ± SE with n=3 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05. Reprinted with permission from[434].

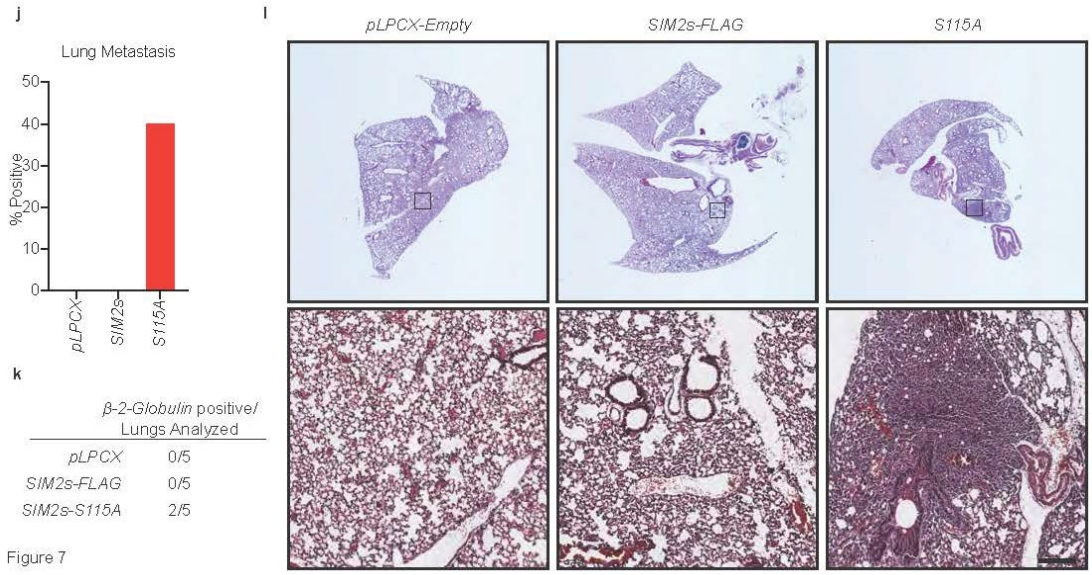


Figure 18 Continued.

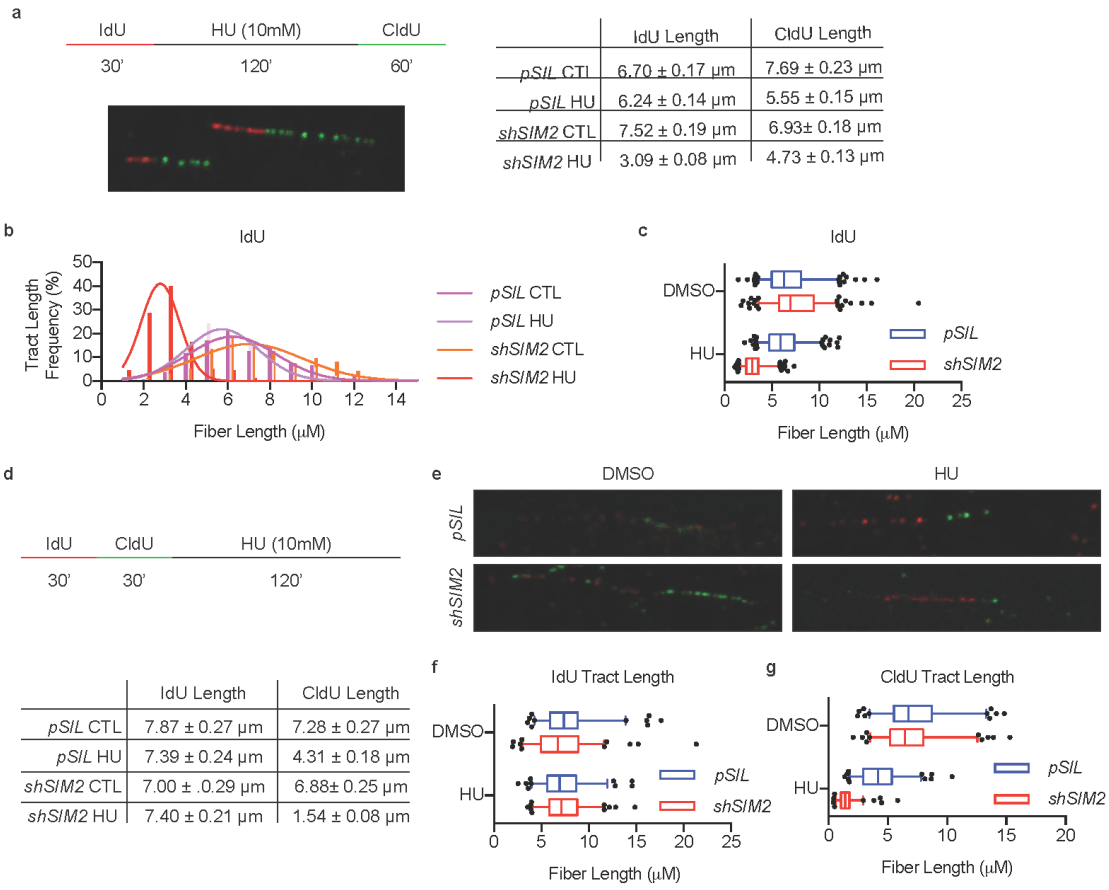
As previous studies have shown that loss of *SIM2s* leads to an increase in lung metastasis in xenografts, we hypothesized that S115A tumors would have increased lung metastasis[413]. As such, RNA was isolated from mouse lungs, and RT-qPCR analysis was performed with human-specific β-2-globulin primers[413]. We found an increase in human specific β-2-globulin positive lungs in mice that were injected with S115A cells compared to either the *SIM2s-FLAG* or control groups (Figure 18j, k). To further validate these findings, we stained lung sections gathered from these mice with H&E. In two of the five mice that were injected with *S115A*, dense masses of cells, with defined boundaries could be identified in the lung tissues (Figure 18l). Together, these findings link SIM2s-dependent DDR and HR to inhibition of EMT and DCIS progression.

CHAPTER IV

LOSS OF *SIM2s* INHIBITS RAD51 BINDING AND LEADS TO UNRESOLVED REPLICATION STRESS AND INCREASED GENOMIC INSTABILITY

Loss of SIM2s Leads to an Increase in Replication Fork Collapse but Does Not Affect Replication Restart Speed

It has been previously demonstrated that members of the HR DDR pathway are associated with maintaining genomic stability through the resolution of replicative stress[295-299, 305]. Having recently discovered SIM2s as a novel protein involved in HR, we hypothesized that loss of *SIM2s* (through inclusion of *shSIM2)* would result in a decrease in genomic stability[434]. To test this, we pulse labeled our previously established MCF7-*shSIM2* and MCF7-*pSIL-scambled* cell lines with the thymidine analog IdU (5-Iodo-2'-deoxyuridine) for 30 minutes. Then, control groups were immediately treated with CldU (5-Chloro-2'-deoxyuridine) for the indicated time, and treatment groups were treated with HU (a potent antineoplastic agent that inhibits DNA replication through the inhibition of ribonucleoside diphosphate reductase [RNR]) for 2 hours. Finally, treatment groups were pulse labeled for the indicated time with CldU (Figure 19a)[407, 434]. A minimum of 100 tracts were measured for each group and analyzed for overall length (Figure 19b). To correct for differences in replication speed between cells containing *shSIM2* and control cells, tract lengths of HU-treated groups were normalized to their basal counterparts for statistical analysis. Using this method, we were able to assess changes in both replication restart, by measuring CldU tract lengths, and replication fork stability under replication stress, by measuring IdU tract lengths. Interestingly, when we assessed DNA replication



	IdU Length	CldU Length
pSIL CTL	6.70 ± 0.17 μm	7.69 ± 0.23 μm
pSIL HU	6.24 ± 0.14 μm	5.55 ± 0.15 μm
shSIM2 CTL	7.52 ± 0.19 μm	6.93± 0.18 μm
shSIM2 HU	3.09 ± 0.08 μm	4.73 ± 0.13 μm

	IdU Length	CldU Length
pSIL CTL	7.87 ± 0.27 μm	7.28 ± 0.27 μm
pSIL HU	7.39 ± 0.24 μm	4.31 ± 0.18 μm
shSIM2 CTL	7.00 ± .0.29 μm	6.88± 0.25 μm
shSIM2 HU	7.40 ± 0.21 μm	1.54 ± 0.08 μm

Figure 19 Loss of *SIM2s* leads to an increase in replication fork collapse.
(a) Illustration of experimental design and raw measurements collected for BrdU analogue pulse labeling in MCF7-*shSIM2* and MCF7-*pSIL* for experiments in Figure 19b-c. (b and c) Visualization of IdU tract lengths from MCF7-shSIM2 and MCF7-pSIL cells treated with DMSO or 10 mM HU. (d and e) Illustration of experimental design and raw measurements for BrdU analogue pulse labeling experiments for Figure 19f-g. (f) Visualization of IdU tract lengths from MCF7-*shSIM2* and MCF7-*pSIL* cells treated with DMSO or 10mM HU. (g) Visualization of CldU tract lengths from MCF7-*shSIM2* and MCF7-*pSIL* cells treated with DMSO or 10 mM HU. Values indicate the median with the interquartile range. Whiskers span 5-95th percentile; n=100 unless otherwise stated. Student's t-test was performed to test significance. *p-value<0.05.

restart 30 minutes after HU release, no visible CldU tracts were observable, leading us to extend CldU pulse labeling to 60 minutes (data not shown).

We first analyzed the ability of *shSIM2* cells to restart replication after release from HU treatment. Interestingly, we found no significant change in replication restart between *shSIM2* and control groups (Figure 20a, b). However, when we assessed replication fork stability, as observed by the IdU tract length, we detected a dramatic decrease in the IdU tract length of MCF7-*shSIM2* cells that were treated with HU (Figure 19b, c). This shortening of the tract length of replicated DNA has been previously attributed to failure of a cell to maintain the stability of the stalled replication fork, leading to its collapse and subsequent re-replication[295]. To confirm this finding, we again pulse labeled MCF7-*shSIM2* and MCF7-*pSIL* control cells with IdU for 30 minutes, and then immediately pulse labeled with CldU for another 30 minutes before treating cells for 2 hours with DMSO or 10 mM HU (Figure 19d). The length of IdU tracts and CldU tracts that were immediately adjacent to IdU tracts, which eliminated any newly firing forks, were then measured (Figure 19d-g). A slight decrease in IdU replication length was observed in MCF7-*shSIM2* cells treated with DMSO, supporting differing replication speeds between the two cell lines

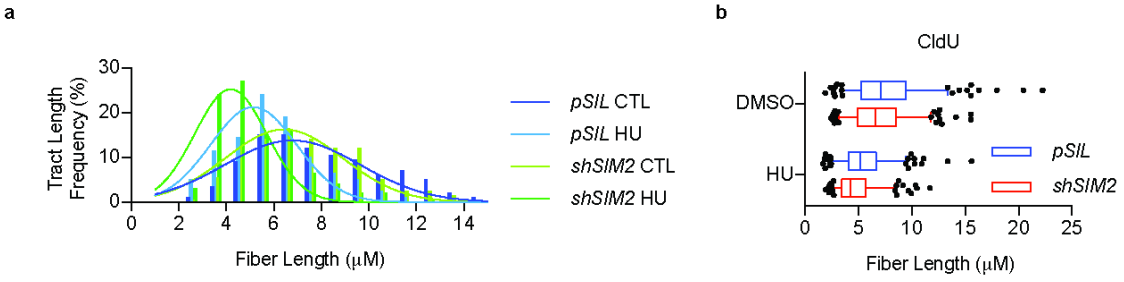


Figure 20 Loss of *SIM2s* does not affect replication-fork restart time.
(a-b) Visualization of CldU tract lengths from MCF7-*shSIM2* and MCF7-*pSIL* cells treated with DMSO or 10mM HU.

(Figure 19d, f). Of note, we found a significant decrease in the CldU tract length in MCF7-*shSIM2* cells treated with HU, supporting our previous finding (Figure 19d, g).

Loss of SIM2s Leads to an Increase in Stalled Forks and Newly Firing Origins

It has been reported that stalled or collapsed replication forks can lead to formation of a gap between IdU and CldU labeling, possibly due to the firing of a new origin downstream of the stalled fork[257]. Upon measuring the length of gaps between the two pulse labels (from Figure 19a), we found that, although there is no difference between treated and untreated cells, cells lacking SIM2s had significantly larger gaps than control cells (Figure 21a, b). Since larger gap sizes have been attributed to multiple causes, we next tested MCF7-*shSIM2* and control cells for the frequency of elongating replication forks, stalled replication forks, and newly firing origins (Figure 21a, c)[257]. Cells containing

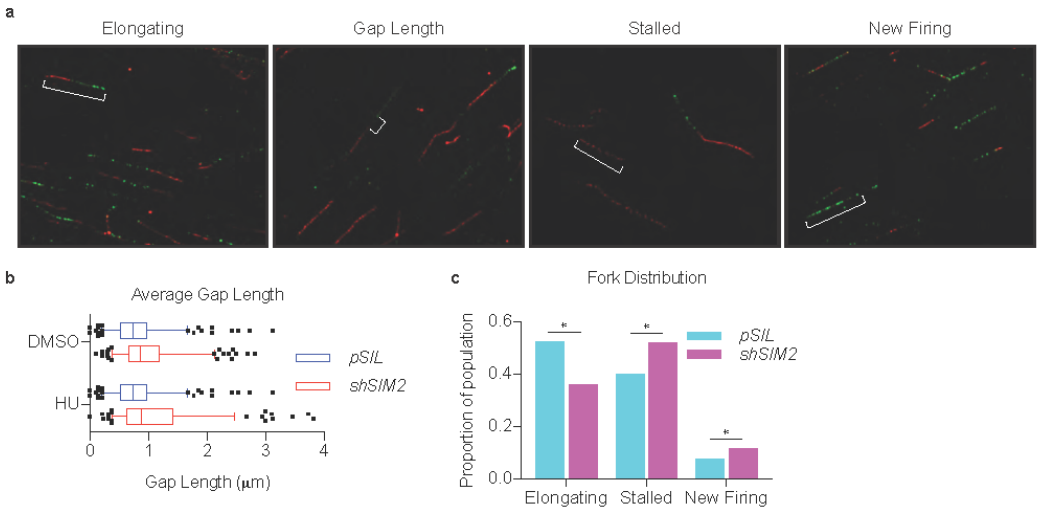


Figure 21 Loss of *SIM2s* increases the incidence of aberrations associated with stalled replication forks. (a) DNA fibers from MCF7 cells treated as in Figure 16a were analyzed for the presence of elongating strands, length of the gap between adjacent labeling, presence of stalled forks, and newly firing forks. (b) A significant increase in gap length was observed in cells containing *shSIM2*, with no difference found between DMSO and HU treatment groups. Values indicate the median with the interquartile range. Whiskers span 5-95th percentile; n=100. Student's t-test was performed to test significance. (c) Finally, *shSIM2* containing cells had lower incidence of actively elongating tracts, with a significant increase in stalled forks and newly firing origins. Likelihood ratio and Pearson's chi-squared tests were performed to test correlations. *p-value<0.05.

shSIM2 exhibited a higher frequency of both stalled replication forks, as well as an increase in the presence of newly firing replication forks. Both of these findings suggest that loss of *SIM2s* leads to genomic instability that culminates in the inability to resolve replication stress.

Loss of SIM2s Disrupts DNA Replication

To further characterize the effect loss of *SIM2s* has on replicating cells, we analyzed MCF7-*shSIM2* and control cells during anaphase. Previously, it has been shown that loss of factors involved in HR contributes to sister-chromosome non-disjunction, or the inability for sister-chromosomes to fully separate during mitosis. Traditional DNA staining is sufficient to reveal these abnormalities, which can present as DNA bridges, lagging strands, or acentric chromosomes (Figure 22a). After synchronization using a di-thymidine block, cells were stained with Hoechst 33342 and analyzed for the presence of anaphase abnormalities. Here, we observed a significant increase in the fraction of cells containing DNA bridges, lagging strands, and acentric chromosomes in cells containing *shSIM2* (Figure 22b).

Loss of SIM2s Decreases RAD51 Recruitment

Previous studies have demonstrated that there is significant overlap between DDR pathways and the stabilization and resolution of replication stress[435, 436]. With the increase in genomic instability associated with loss of *SIM2*, we sought to determine if loss of *SIM2s* correlates with a reduction in DDR factors. To start, we looked at γH2AX foci formation within the nucleus of MCF7-*shSIM2* and MCF7-*pSIL* cells. As a previous study

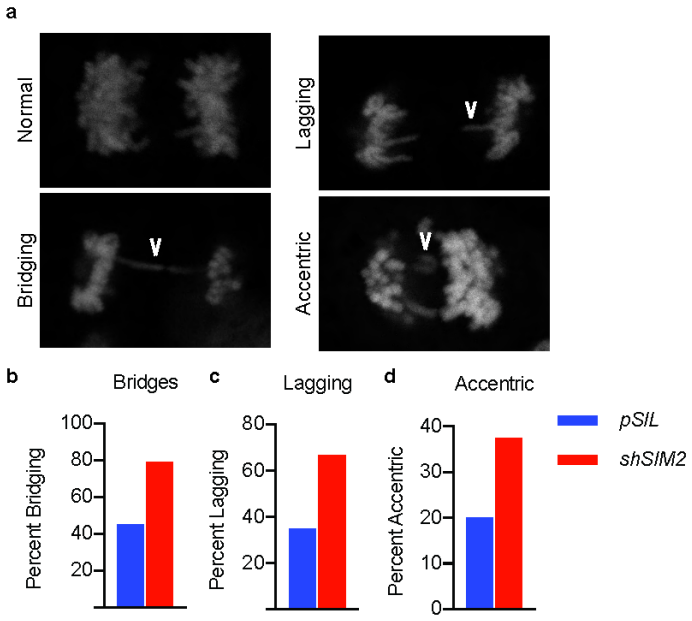


Figure 22 Loss of *SIM2s* disrupts DNA replication.
(a) Representative images showing normal, bridging, lagging, and acentric chromosomes during mitosis. MCF7-*shSIM2* and MCF7-*pSIL* cells were synchronized and fixed during anaphase before being analyzed for the presence of (b) bridging strands, (c) lagging strands, and (d) acentric chromosomes. Likelihood ratio and Pearson's chi-squared tests were performed to test correlations. n=20. *p-value<0.05.

has demonstrated that γH2AX levels rise between 16-48 hours after HU treatment, we dosed our cells with 0.5 mM HU for 24 hours before fixation and immunofluorescent staining[435]. Interestingly, we observed a significant increase in γH2AX foci in cells containing *shSIM2* that were treated with HU (Figure 23). This finding is likely due to an increase in unresolved stalled replication forks and an increase in dsDNA breaks[434]. Of further note, MCF7-*shSIM2* cells exhibited an increase in BRCA1 and RPA foci after HU treatment, whereas p53BP1 showed no change between control and *shSIM2* cells (Figure 23). Surprisingly, loss of *SIM2s* in MCF7 cells resulted in a significant reduction in RAD51 foci after HU treatment (Figure 23).

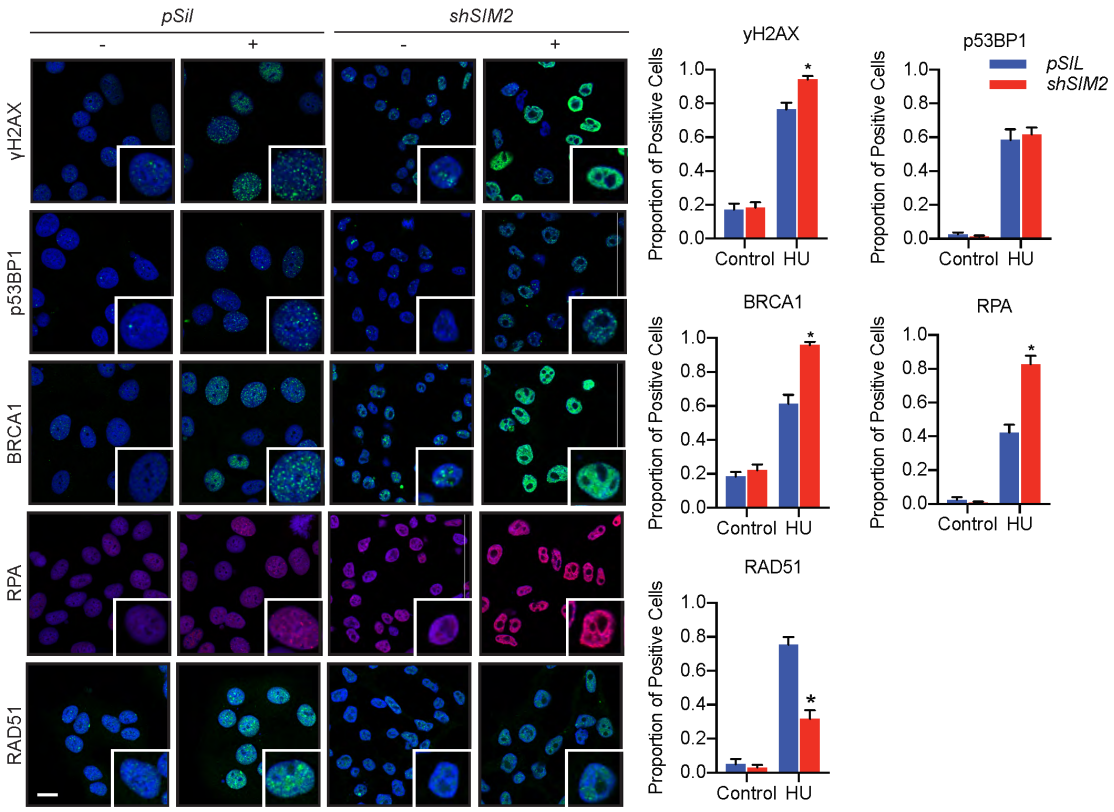


Figure 23 Loss of *SIM2s* leads to a decrease in RAD51 foci in MCF7 cells.
MCF7 cells containing *shSIM2* or *pSIL* were treated with 0.5 mM HU, fixed 24 hours later, and finally probed with the indicated antibody. Scale bars, 10 μm. Values indicate the mean ± SE with n=3. Student's t-test was performed to test significance. *p-value<0.05.

SIM2s is Necessary for RAD51 Recruitment in Response to Genotoxic Stress in Primary Mammary Epithelial Cells

To confirm our finding that RAD51 is reduced in cells lacking SIM2s, we utilized a mammary tissue specific conditional *Sim2* knockout mouse, which was generated via a "floxed" $Sim2^{fl/fl}$ allele. *Sim2* is conditionally deleted for the duration of lactation by crossing $Sim2^{fl/fl}$ mice with $Wap^{Cre/+}$ mice, which express Cre recombinase under the control of the whey acidic protein (*Wap*) promoter. *Wap* is specifically expressed in mammary alveolar epithelial cells from mid-pregnancy through lactation, and thus allows

for conditional knockout of *Sim2.* To visualize Cre recombinase activity, *Wap*$^{Cre/+}$;*Sim2*$^{fl/fl}$ (*Sim2*$^{fl/fl}$) and Wap$^{Cre/+}$; Sim2$^{+/+}$ (control) mice were genetically tagged with Gt(ROSA)26Sor$^{tm4(ACTB-tdTomato,-EGFP)luo}$/J (mTmG)[437]. Confirmation of efficient loss of the *Sim2s* locus after pregnancy was visualized in tissue sections using immunofluorescence and was confirmed via RT-qPCR (Figure 24a, b). Primary mammary epithelial cells (MEC) were isolated from mice during late pregnancy (day 18) and treated with 0.5 mM HU or DMSO for 24 hours prior to immunostaining for RAD51. As we observed in MCF7 cells, loss of *SIM2s* led to a significant reduction in RAD51 foci in cells treated with HU (Figure 24c).

Loss of SIM2s Increases γH2AX Levels in Mammary Tissue

With a significant decrease in RAD51 with the loss of *SIM2s*, we hypothesized that prolonged absence of SIM2s would lead to an increase in genomic instability, resulting in elevated levels of DNA damage. To test this, we isolated the fourth inguinal mammary glands of *Sim2*$^{fl/fl}$ and control mice at lactation day 18, allowing MECs to progress through pregnancy and peak lactation; two stages that metabolically stress the mammary tissue and result in the elevation of factors associated with HR [438]. Sections were then probed for γH2AX. Loss of *SIM2* resulted in significantly higher intensities of γH2AX, suggesting that decreased levels of SIM2s result in elevated levels of genomic instability (Figure 24d).

SIM2s Interacts with RAD51 and is Necessary for RAD51 Binding to the Chromosome

It has been established that RAD51 translocates to the nucleus before binding to dsDNA breaks in response to DNA damage[439]. To test where in this process loss of *SIM2s* interferes with RAD51 loading, we isolated cytoplasmic, soluble nucleus, and insoluble

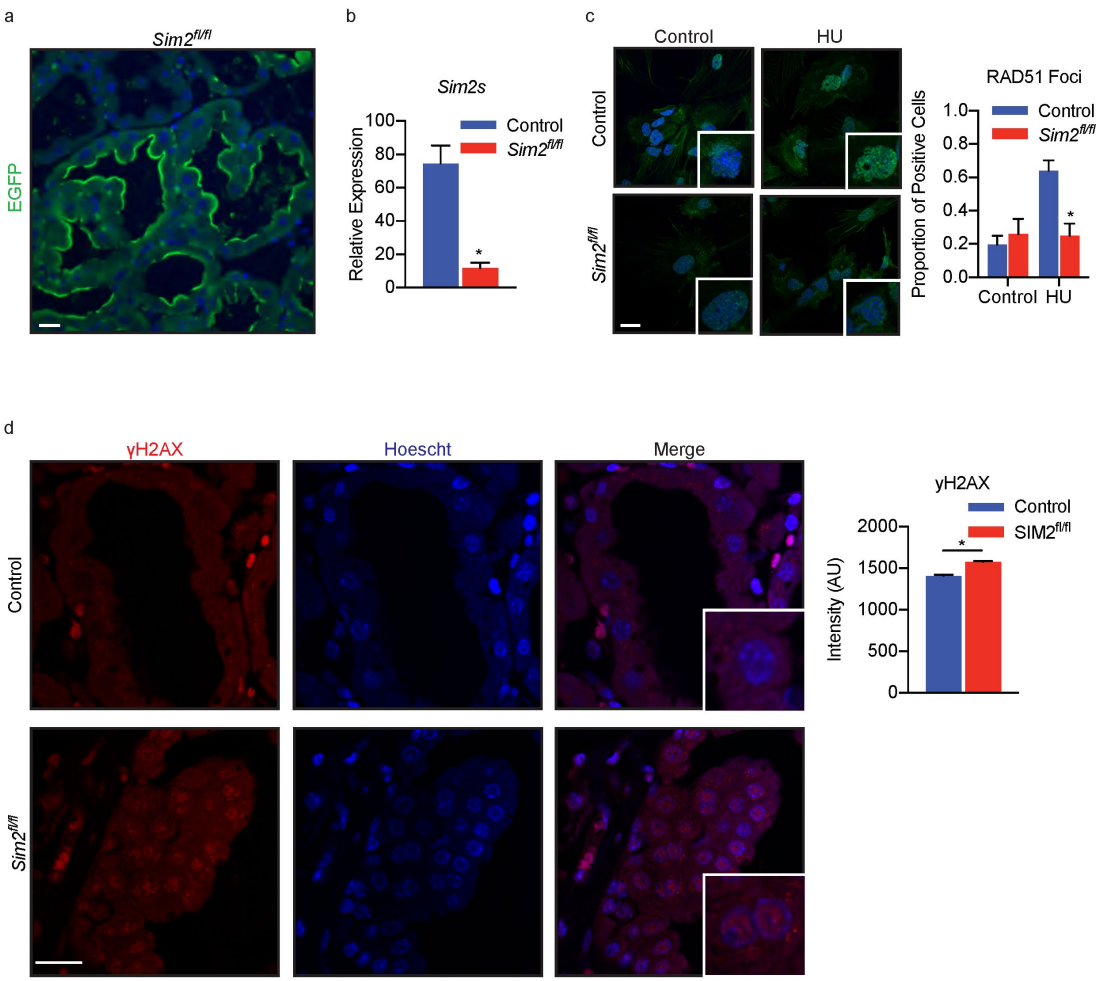


Figure 24 Loss of *SIM2s* in a mouse model decreases RAD51 recruitment and increases genomic instability.
(a) Recombination of the SIM2 locus was confirmed in late-stage pregnant mice by the presence of eGFP. Scale bars, 20 µm. (b) RNA isolated from mammary glands of control or $SIM2^{fl/fl}$ mice was analyzed via RT-qPCR for the presence of *Sim2s* mRNA. (c) MECs isolated from control and late-stage pregnancy $SIM2^{fl/fl}$ mice were treated with 0.5 mM HU for 24 hours before being assessed for the presence of RAD51 foci. Scale bars, 10 µm. (d) Mammary glands were collected from lactating $SIM2^{fl/fl}$ and control mice and assessed for the presence of γH2AX. Scale bars, 20 µm. Values indicate the mean ± SE with n=3. Student's t-test was performed to test significance. *p-value<0.05.

nuclear (chromatin) fractions from MCF7-*shSIM2* and MCF7-*pSIL* cells that had been treated with DMSO or 0.5 mM HU for 24 hours (Figure 25a). Counterintuitive to the decrease in RAD51 foci we observed in response to DNA damage in the *shSIM2 cells*, we detected an increase in RAD51 levels in the cytoplasm, both basally and with treatment of HU. We also observed no change in the ability of RAD51 to translocate to the nucleus (Figure 25a). However, loss of *SIM2s* led to a significant decrease in the levels of RAD51 found in the insoluble/chromatin fraction of the nucleus (Figure 25a).

Having previously shown that SIM2s interacts with BRCA1, we next hypothesized that SIM2s may be necessary for RAD51 to interact with other proteins within the repairosome. However, MCF7 cells containing *shSIM2* that were treated with 0.5 mM HU did not show a significant decrease in the ability of RAD51 to bind to BRCA1, as observed through immunoprecipitation of BRCA1 (Figure 25b).

Based on this finding, we next hypothesized that SIM2s may directly interact with RAD51. As no previous studies show the kinetics of SIM2s in response to HU treatment, we first analyzed SIM2s levels in response to 0.5 mM HU treatment over time. Immunofluorescent analysis of FLAG in SUM159 cells overexpressing SIM2s-FLAG showed an increase in nuclear FLAG after treatment with 0.5 mM HU (Figure 25c, d). These findings were also confirmed via western blot analysis in SUM159 and MCF7 cells, where we saw increased levels of SIM2s 2-4 hours after treatment with 0.5 mM HU (Figure 25e, f). Having shown that SIM2s levels peak 2 hours after treatment with HU, we next immunoprecipitated RAD51 from MCF7 cells treated with DMSO or HU for 2 hours and probed for SIM2s to test for their interaction. Interestingly, we observed that RAD51

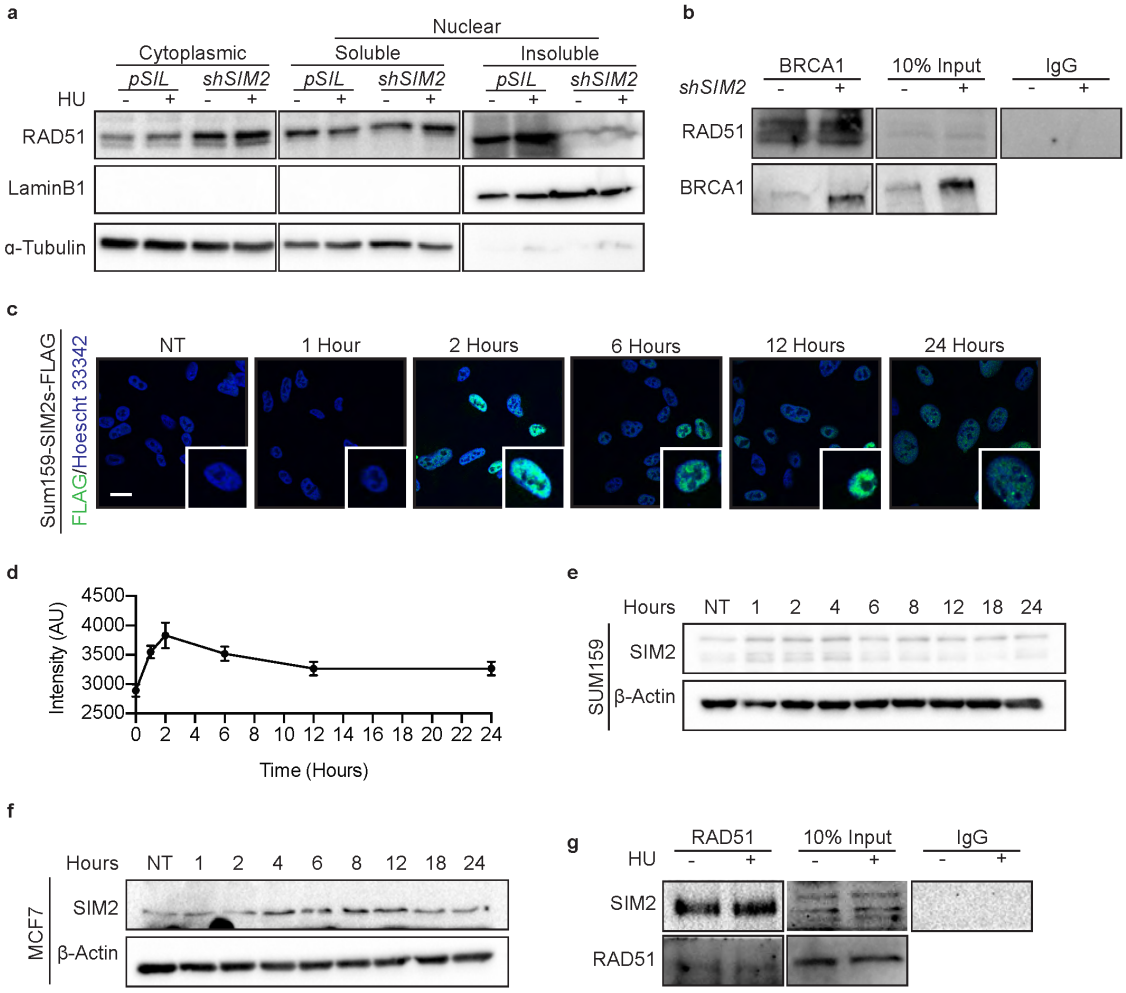


Figure 25 SIM2s interacts with RAD51 and is necessary for RAD51 nucleation.
(a) MCF7-*shSIM2* and MCF7-*pSIL* cells were treated with DMSO or 0.5 mM HU for 24 hours and then fractionated before being probed for RAD51. α-Tubulin and LaminB1 were used as loading controls to verify efficient separation of fractions. (b) MCF7-*shSIM2* and MCF7-*pSIL* cells were treated with 0.5 mM HU and harvested 2 hours later. BRCA1 was immunoprecipitated and lysates were probed for the indicated proteins. (c) Stabilization and localization of SIM2s was assessed in SUM159-*SIM2s-FLAG* treated with 0.5 mM HU and fixed at the indicated timepoints before being probed for FLAG. (d) Quantification of nuclear FLAG from Figure 6C. (e-f) Western blot analysis of SIM2s stabilization in (e) SUM159-*SIM2s-FLAG* and (f) MCF7 cells in response to 0.5mM HU treatment. (g) RAD51 was immunoprecipitated in MCF7 cells after 2 hours of treatment with HU or DMSO and lysates were probed for the indicated proteins.

interacts with SIM2s both basally and with HU treatment (Figure 25g), confirming that SIM2s interacts with the RAD51 complex (Figure 26).

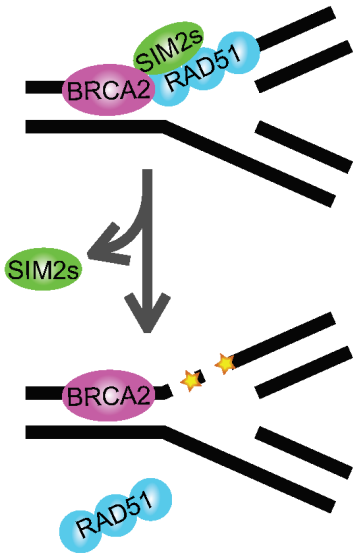


Figure 26 Theoretical model for SIM2s-RAD51 interaction during replication stress.

CHAPTER V

CONCLUSIONS

It is well recognized that genomic instability is a crucial factor in breast cancer progression[440]. A great deal of research has focused on key components of the DDR pathway, including BRCA1, BRCA2, PTEN, and RAD51, as mutations in these factors can increase an individual's lifetime risk of developing breast cancer 10-fold[441-445]. Tumors with mutations in these key components are largely TNBC and are significantly more aggressive than their HR-competent counterparts. However, loss or mutation of factors in the HR pathway also opens up new therapeutic strategies, such as the recently approved PARP inhibitors. These unique phenotypes have led to the "BRCAness" nomenclature, which is gaining traction in the description of HR-deficient tumors[58, 130]. Of note, these mutations are not only found in familial breast cancer, but are commonly found in more aggressive sporadic tumors[141, 446].

In the first part of these studies, we show that loss of *SIM2s* leads to increased genomic instability and correlates with progression from DCIS to IDC (Figure 8 and Figure 9). Of interest, loss of *SIM2s* hinders HR by impairing RAD51 nucleation in response to IR, abrogating DDR through HR (Figure 14). Failure of a cell to initiate this crucial step has been linked with an increased incidence of SNPs and chromosomal abnormalities[447-451]. These findings parallel previous studies demonstrating that the loss of BRCA1 results in increased γH2AX foci and sensitivity to mutagenic agents (Figure 9 and Figure 14)[452-455]. Other studies have also suggested that *BRCA* mutant carriers exhibit higher-grade DCIS cases than are observed in sporadic tumors[250]. This correlation between

defects in HR machinery and more severe tumorigenesis may explain the negative correlation between SIM2s levels and progression to IDC (Figure 8).

A mutation in a single BRCA1/2 allele is sufficient to result in carcinogenesis; however, a single functional copy of BRCA1/2 is also sufficient to maintain HR functionality[294]. Thus, BRCA-mutated tumor progression is thought to predominantly occur through loss of heterozygosity (LOH). Yet, the mechanistic pathways underlying LOH are vague and revolve around accruing DNA damage. More recent studies have shown that replication stress is more sensitive to perturbations in BRCA1 levels than other established BRCA1 roles. More specifically, a mutation in a single copy of BRCA1 is sufficient to reduce replication fork stability[294]. This finding supports the notion that although the role BRCA plays in HR is crucial for maintaining genomic fidelity, the initial, and possibly more important, increase in genomic instability seen during cancer progression in BRCA-associated tumors could be due to its role in maintaining replication fork stability. The rapid increase in genomic instability observed with *BRCA* mutations mimics those observed with loss of *SIM2s*, underpinning its importance in this pathway (Figure 24).

Accordingly, we show that loss of *SIM2s* sensitizes replication forks to genotoxic stress, leading to increased replication fork collapse, much like loss of *BRCA1/2* (Figure 19, Figure 20, and Figure 21). Replication fork collapse is also associated with abnormal separation of sister chromatids during mitosis, which is known to result in chromatin fragmentation and aneuploidy (Figure 22)[295]. These findings parallel those previously observed with *BRCA1* mutations, with familial BRCA1/2-associated tumors having a higher instance of DNA deletions and chromosomal translocations than sporadic tumors[294].

The rapid flux in genomic integrity that is observed in BRCA-mutated tumors predisposes them to mutations in *TP53,* ER, PR, and ERBB2, and thus biases them toward highly-invasive TNBC with a poor clinical prognosis[58].

The direct role SIM2s plays within this pathway remains unknown and requires further investigation. Here, we show that loss of *SIM2s* does not affect the ability of RAD51 to translocate to the nucleus in response to replication stress (Figure 25); however, loss of *SIM2s* does result is reduced RAD51 binding on DNA. Previous studies have reported that the dynamics of RAD51 binding and unbinding to DNA requires RPs, in part due to the drastic decrease in RAD51s innate ATP hydrolysis activity during its evolution from RecA (the bacterial RAD51 homolog)[456]. Thus, the loss of RAD51 foci in response to genotoxic stress with loss of *SIM2s* suggests that SIM2s may be interacting directly with RAD51 or indirectly acting on RAD51 through the interaction/regulation of one of the RP complexes. Interestingly, high levels of *SIM2s* and XRCC2 are each individually associated with poor clinical outcome in colorectal cancer patients[457]. This parallel may suggest that SIM2s is involved in the regulation of the BCDX2 complex.

The regulation of orphan bHLH-PAS family members, including SIM2s, remains elusive. Some reports demonstrate that SIM2s is constitutively active, alluding to protein level regulation of SIM2s[352, 458, 459]. Here we have demonstrated a novel mechanism for the stabilization of SIM2s in response to IR through phosphorylation of SIM2s at serine 115 by ATM (Figure 11). The discovery of a new factor involved in ATM signaling pathways could have clinical implications, since it has been suggested that improper ATM signaling results in a 2-5 fold increase in breast cancer risk, in individuals with Ataxia-telangiectasia[426, 460, 461].

ATM has previously been identified as a necessary component in the repair of heterochromatic DNA DSBs[424]. Interestingly, the repair of heterochromatic DNA via HR is both temporally and spatially separated from NHEJ, as HR repair sites are moved to the nuclear periphery prior to repair[431, 462]. This occurs to maintain genomic fidelity by preventing improper recombination events in highly repetitive sequences[462]. The delayed response of γH2AX resolution with loss of *SIM2s* that we observed is consistent with this temporal shift and suggests that SIM2s is involved in the resolution of breaks that occur in heterochromatin (Figure 9).

Although mutation of a single HR factor is sufficient to predispose an individual to breast cancer, it is generally accepted that a secondary mutation is necessary to initiate tumor development. In fact, this trend is observed with *BRCA1*, *BRCA2*, and *RAD51C* mutations and now in *SIM2s*[463-465]. Interestingly, mutation of *TP53* combined with mutation of any of these factors is sufficient to instigate tumor development and results in drastic shift from cellular quiescence toward TNBC[58, 463-465].

Here, we have demonstrated that loss of *SIM2s* in a xenograft model results in an EMT, characterized by decreased levels of E-Cadherin, increased activity of matrix-metalloproteinases (MMPs), and increased invasion and migration potential (Figure 18). EMT, typified by the activation of molecular pathways that promote loss of epithelial character, plays an important role in breast cancer invasion and metastatic progression[466]. During cancer progression, these processes are thought to promote breakdown of the basement membrane, leading to increased invasion into the surrounding stroma and metastasis[467, 468]. Previously, EMT and DDR were assumed to be mechanistically independent. More recently, evidence has suggested that EMT-inducing transcription

factors, including SLUG and TWIST, play a role in cancer progression by promoting EMT and overriding DDR through inhibition of *CDKN1A* (p21)[469]. Interestingly, loss of *SIM2s* similarly correlates with a more metastatic phenotype (Figure 8 and Figure 18). The S115A point mutation is sufficient to imbue mesenchymal characteristics on what would normally be a highly differentiated xenograft (Figure 18). This transition includes an increase in the basal marker K14 and a decrease in E-cadherin. S115A tumors also exhibit large necrotic centers, a condition associated with group 2 and 3 (intermediate and high-grade) lesions on the Van Duys classification scale (Figure 18)[19]. Further, the S115A mutation impairs DDR, showing a significant decrease in γH2AX resolution after treatment with IR (Figure 11). The dual role of this SNP in both impairing DDR and transitioning cells to a more mesenchymal phenotype, suggests that SIM2s may help to bridge the gap between DDR and EMT.

Due to the strong association between mutations in HR repair elements and early-onset breast carcinogenesis, genetic testing for *BRCA1* and *BRCA2* mutations has been suggested for individuals with breast cancer under the age of 60. However, an argument should be made to broaden the scope of genetic testing for individuals with early-onset breast cancer. In these individuals, multigene analysis of factors involved in DDR is warranted based on the correlation between elevated TNBC incidence and mutations in *BARD1, BRIP1, PALB2,* and the *RAD51* paralogs *RAD51C* and *RAD51D*[470]. The definitive role of RAD51 in alleviating replication stress, protecting damaged DNA from nucleases, and promoting genomic stability has long been established[471]. However, due to the embryonic lethality of *RAD51*$^{-/-}$ and RP knockout models, very little progress has been made toward understanding the regulation and nucleation of RAD51[278]. For example,

RAD51C and XRCC3 have been known to play a role in HR for decades; but, due to the difficulty in studying genes that are critical for development, their involvement in replication fork restart has only recently been discovered[278]. As a result, the development of treatments directly targeting these mutations has not been possible.

Here, we provide evidence that SIM2s impedes DCIS progression by inhibiting an EMT and simultaneously plays a crucial role in the maintenance of genomic stability through the regulation of RAD51. As DCIS is a non-obligate precursor to IDC, the identification of new diagnostic markers that correlate with an increased risk for progression to IDC is of critical importantance[472]. Given the strong correlation between loss of *SIM2s* and metastatic potential, SIM2s is an ideal candidate for genetic screenings.

These findings also suggest that tumors lacking wt *SIM2s* may be responsive to targeted therapeutic treatments. As previously mentioned, a unique therapeutic advantage of cancers with mutations in HR proteins is their sensitivity to synthetic lethality treatments[86]. Two leading classes of drugs that have shown promising results are the platinum salts and PARPi. In fact, the PARPi Olaparib (AZD2281) only recently gained approval by the Unites States Food and Drug Administration for use in BRCA-associated tumors, and its therapeutic potential has not been fully realized[339]. Numerous other studies have shown that synthetic lethality treatments targeting mutations in *XRCC2, RAD51,* and *RAD51C* are effective[333, 434, 457]. Here, we propose the addition of mutations in *SIM2s* to this list, as loss of *SIM2s* sensitizes cells to synthetic lethality treatments (Figure 17).

A new strategy for treatment of tumors found to lack crucial DDR repair elements, is combinational therapy with PARPi and platinum salts. Utilization of these treatments in individuals with *BRCA1/2* mutations has shown varying levels of success. For example,

preliminary results from combinational therapy studies with PARPi have shown 51% pathological complete response (up from 26% in control individuals) in TNBC individuals[328, 337, 338]. Interestingly, RAD51 levels can be used as a predictive indicator of the efficacy of these treatments in breast cancer[330]. Specifically, BRCA-mutated tumors that express low levels of RAD51, and thus have low recombinase activity, predict treatment efficacy[334, 473, 474]. This finding highlights the importance of fully understanding the factors involved in the regulation of HR and of continuing to identify novel elements within this pathway, such as SIM2s. These efforts will ultimately improve patient outcomes and our understanding of the intricacies involved in replication stress.

REFERENCES

1. Blom, N., Gammeltoft, S. & Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* **294**, 1351-1362 (1999).

2. van Middendorp, J.J., Sanchez, G.M. & Burridge, A.L. The Edwin Smith papyrus: a clinical reappraisal of the oldest known document on spinal injuries. *Eur Spine J* **19**, 1815-1823 (2010).

3. Breasted, J.H. *The Edwin Smith Sugical Papyrus*, Vol. 1. (The University of Chicago Press; 1930).

4. Sudhakar, A. History of Cancer, Ancient and Modern Treatment Methods. *J Cancer Sci Ther* **1**, 1-4 (2009).

5. Ades, F., Tryfonidis, K. & Zardavas, D. The past and future of breast cancer treatment-from the papyrus to individualised treatment approaches. *Ecancermedicalscience* **11**, 746 (2017).

6. Lakhtakia, R. A Brief History of Breast Cancer: Part I: Surgical domination reinvented. *Sultan Qaboos Univ Med J* **14**, e166-169 (2014).

7. Banks, W.M. On Free Removal of Mammary Cancer, with Extirpation of the Axillary Glands as a Necessary Accompaniment. *Br Med J* **2**, 1138-1141 (1882).

8. Halsted, W.S. I. The Results of Operations for the Cure of Cancer of the Breast Performed at the Johns Hopkins Hospital from June, 1889, to January, 1894. *Ann Surg* **20**, 497-555 (1894).

9. Fisher, B. *et al.* Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. *The New England journal of medicine* **320**, 822-828 (1989).

10. Veronesi, U. *et al.* Quadrantectomy versus lumpectomy for small size breast cancer. *Eur J Cancer* **26**, 671-673 (1990).

11. Lakhani, S.R., International Agency for Research on, C. & Organització Mundial de la, S. *WHO Classification of Tumours of the Breast*. (International Agency for Research on Cancer, Lyon; 2014).

12. Cowell, C.F. *et al.* Progression from ductal carcinoma in situ to invasive breast cancer: Revisited. *Molecular oncology* (2013).

13. Lennington, W.J., Jensen, R.A., Dalton, L.W. & Page, D.L. Ductal carcinoma in situ of the breast. Heterogeneity of individual lesions. *Cancer* **73**, 118-124 (1994).

14. Allred, D.C. *et al.* Ductal carcinoma in situ and the emergence of diversity during breast cancer evolution. *Clin Cancer Res* **14**, 370-378 (2008).

15. Miller, N.A. *et al.* In situ duct carcinoma of the breast: clinical and histopathologic factors and association with recurrent carcinoma. *Breast J* **7**, 292-302 (2001).

16. Van Cleef, A. *et al.* Current view on ductal carcinoma in situ and importance of the margin thresholds: A review. *Facts, views & vision in ObGyn* **6**, 210-218 (2014).

17. Fisher, E.R. *et al.* Pathologic findings from the National Surgical Adjuvant Breast Project (NSABP) Protocol B-17. Intraductal carcinoma (ductal carcinoma in situ). The National Surgical Adjuvant Breast and Bowel Project Collaborating Investigators. *Cancer* **75**, 1310-1319 (1995).

18. Solin, L.J. *et al.* Ductal carcinoma in situ (intraductal carcinoma) of the breast treated with breast-conserving surgery and definitive irradiation. Correlation of pathologic parameters with outcome of treatment. *Cancer* **71**, 2532-2542 (1993).

19. Silverstein, M.J. *et al.* Prognostic classification of breast ductal carcinoma-in-situ. *Lancet* **345**, 1154-1157 (1995).

20. Sneige, N. *et al.* Ductal carcinoma in situ treated with lumpectomy and irradiation: histopathological analysis of 49 specimens with emphasis on risk factors and long term results. *Hum Pathol* **26**, 642-649 (1995).

21. Clark, S.E. *et al.* Molecular subtyping of DCIS: heterogeneity of breast cancer reflected in pre-invasive disease. *Br J Cancer* **104**, 120-127 (2011).

22. Perou, C.M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752 (2000).

23. Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10869-10874 (2001).

24. Burstein, H.J., Polyak, K., Wong, J.S., Lester, S.C. & Kaelin, C.M. Ductal carcinoma in situ of the breast. *The New England journal of medicine* **350**, 1430-1441 (2004).

25. Muggerud, A.A. *et al.* Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer. *Mol Oncol* **4**, 357-368 (2010).

26. Moelans, C.B., de Weger, R.A., Monsuur, H.N., Maes, A.H. & van Diest, P.J. Molecular differences between ductal carcinoma in situ and adjacent invasive breast

carcinoma: a multiplex ligation-dependent probe amplification study. *Anal Cell Pathol (Amst)* **33**, 165-173 (2010).

27. Lesurf, R. *et al.* Molecular Features of Subtype-Specific Progression from Ductal Carcinoma In Situ to Invasive Breast Cancer. *Cell Rep* **16**, 1166-1179 (2016).

28. Casasent, A.K. *et al.* Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell* **172**, 205-217 e212 (2018).

29. Rosner, D., Bedwani, R.N., Vana, J., Baker, H.W. & Murphy, G.P. Noninvasive breast carcinoma: results of a national survey by the American College of Surgeons. *Ann Surg* **192**, 139-147 (1980).

30. Kumar, G., Redick, M. & Dixon, G.D. New techniques for mammography screening: advantages and limitations. *Missouri medicine* **102**, 138-141 (2005).

31. Rudloff, U. *et al.* Nomogram for predicting the risk of local recurrence after breast-conserving surgery for ductal carcinoma in situ. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 3762-3769 (2010).

32. Page, D.L., Dupont, W.D., Rogers, L.W., Jensen, R.A. & Schuyler, P.A. Continued local recurrence of carcinoma 15-25 years after a diagnosis of low grade ductal carcinoma in situ of the breast treated only by biopsy. *Cancer* **76**, 1197-1200 (1995).

33. Sanders, M.E., Schuyler, P.A., Dupont, W.D. & Page, D.L. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer* **103**, 2481-2484 (2005).

34. Kuhl, C.K. *et al.* MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. *Lancet* **370**, 485-492 (2007).

35. Ernster, V.L., Barclay, J., Kerlikowske, K., Wilkie, H. & Ballard-Barbash, R. Mortality among women with ductal carcinoma in situ of the breast in the population-based surveillance, epidemiology and end results program. *Arch Intern Med* **160**, 953-958 (2000).

36. Allred, D.C. *et al.* Adjuvant tamoxifen reduces subsequent breast cancer in women with estrogen receptor-positive ductal carcinoma in situ: a study based on NSABP protocol B-24. *J Clin Oncol* **30**, 1268-1273 (2012).

37. Punglia, R.S., Schnitt, S.J. & Weeks, J.C. Treatment of ductal carcinoma in situ after excision: would a prophylactic paradigm be more appropriate? *J Natl Cancer Inst* **105**, 1527-1533 (2013).

38. Warnberg, F. *et al.* Effect of radiotherapy after breast-conserving surgery for ductal carcinoma in situ: 20 years follow-up in the randomized SweDCIS Trial. *J Clin Oncol* **32**, 3613-3618 (2014).

39. Fisher, B. *et al.* Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. *The New England journal of medicine* **312**, 665-673 (1985).

40. Badruddoja, M. Ductal carcinoma in situ of the breast: a surgical perspective. *Int J Surg Oncol* **2012**, 761364 (2012).

41. Fisher, B. *et al.* Lumpectomy compared with lumpectomy and radiation therapy for the treatment of intraductal breast cancer. *The New England journal of medicine* **328**, 1581-1586 (1993).

42. Julien, J.P. *et al.* Radiotherapy in breast-conserving treatment for ductal carcinoma in situ: first results of the EORTC randomised phase III trial 10853. EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group. *Lancet* **355**, 528-533 (2000).

43. Houghton, J. *et al.* Radiotherapy and tamoxifen in women with completely excised ductal carcinoma in situ of the breast in the UK, Australia, and New Zealand: randomised controlled trial. *Lancet* **362**, 95-102 (2003).

44. Smith, B.D. *et al.* Effectiveness of radiation therapy in older women with ductal carcinoma in situ. *J Natl Cancer Inst* **98**, 1302-1310 (2006).

45. Vargas, C. *et al.* Factors associated with local recurrence and cause-specific survival in patients with ductal carcinoma in situ of the breast treated with breast-conserving therapy or mastectomy. *Int J Radiat Oncol Biol Phys* **63**, 1514-1521 (2005).

46. Jiveliouk, I., Corn, B., Inbar, M. & Merimsky, O. Ductal carcinoma in situ of the breast in Israeli women treated by breast-conserving surgery followed by radiation therapy. *Oncology* **76**, 30-35 (2009).

47. Kayani, N. & Bhurgri, Y. Ductal carcinoma in situ (DCIS) in Karachi. *J Pak Med Assoc* **55**, 199-202 (2005).

48. Ruggiero, R. *et al.* Ductal carcinoma in situ of the breast: our experience. *G Chir* **30**, 121-124 (2009).

49. Newman, B. *et al.* The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. *Breast Cancer Res Treat* **35**, 51-60 (1995).

50. Furberg, H., Millikan, R., Dressler, L., Newman, B. & Geradts, J. Tumor characteristics in African American and white women. *Breast Cancer Res Treat* **68**, 33-43 (2001).

51. Bianchini, G., Balko, J.M., Mayer, I.A., Sanders, M.E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol* **13**, 674-690 (2016).

52. Hammond, M.E. *et al.* American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* **28**, 2784-2795 (2010).

53. Bonotto, M. *et al.* Measures of outcome in metastatic breast cancer: insights from a real-world scenario. *Oncologist* **19**, 608-615 (2014).

54. Carey, L.A. *et al.* The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin Cancer Res* **13**, 2329-2334 (2007).

55. Sparano, J.A. *et al.* Randomized phase III trial of ixabepilone plus capecitabine versus capecitabine in patients with metastatic breast cancer previously treated with an anthracycline and a taxane. *J Clin Oncol* **28**, 3256-3263 (2010).

56. Sikov, W.M. *et al.* Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (Alliance). *J Clin Oncol* **33**, 13-21 (2015).

57. Isakoff, S.J. *et al.* TBCRC009: A Multicenter Phase II Clinical Trial of Platinum Monotherapy With Biomarker Assessment in Metastatic Triple-Negative Breast Cancer. *J Clin Oncol* **33**, 1902-1909 (2015).

58. Turner, N., Tutt, A. & Ashworth, A. Hallmarks of 'BRCAness' in sporadic cancers. *Nat Rev Cancer* **4**, 814-819 (2004).

59. Ding, S.L. *et al.* Abnormality of the DNA double-strand-break checkpoint/repair genes, ATM, BRCA1 and TP53, in breast cancer is related to tumour grade. *Br J Cancer* **90**, 1995-2001 (2004).

60. Pang, J. *et al.* Invasive breast carcinomas in Ghana: high frequency of high grade, basal-like histology and high EZH2 expression. *Breast Cancer Res Treat* **135**, 59-66 (2012).

61. Zeidler, M. *et al.* The Polycomb group protein EZH2 impairs DNA repair in breast epithelial cells. *Neoplasia* **7**, 1011-1019 (2005).

62. Cannan, W.J. & Pederson, D.S. Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin. *J Cell Physiol* **231**, 3-14 (2016).

63. Vilenchik, M.M. & Knudson, A.G. Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12871-12876 (2003).

64. Dronkert, M.L. & Kanaar, R. Repair of DNA interstrand cross-links. *Mutat Res* **486**, 217-247 (2001).

65. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715 (1993).

66. Olinski, R., Jurgowiak, M. & Zaremba, T. Uracil in DNA--its biological significance. *Mutat Res* **705**, 239-245 (2010).

67. Murphy, M.P. How mitochondria produce reactive oxygen species. *Biochem J* **417**, 1-13 (2009).

68. Abbotts, R. & Wilson, D.M., 3rd Coordination of DNA single strand break repair. *Free Radic Biol Med* **107**, 228-244 (2017).

69. Wallace, D.C. Mitochondria and cancer. *Nat Rev Cancer* **12**, 685-698 (2012).

70. Dizdaroglu, M., Jaruga, P., Birincioglu, M. & Rodriguez, H. Free radical-induced damage to DNA: mechanisms and measurement. *Free Radic Biol Med* **32**, 1102-1115 (2002).

71. Gates, K.S. An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem Res Toxicol* **22**, 1747-1760 (2009).

72. Pogozelski, W.K. & Tullius, T.D. Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety. *Chem Rev* **98**, 1089-1108 (1998).

73. Aguiar, P.H. *et al.* Oxidative stress and DNA lesions: the role of 8-oxoguanine lesions in Trypanosoma cruzi cell viability. *PLoS Negl Trop Dis* **7**, e2279 (2013).

74. Basu, A.K., Loechler, E.L., Leadon, S.A. & Essigmann, J.M. Genetic effects of thymine glycol: site-specific mutagenesis and molecular modeling studies. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 7677-7681 (1989).

75. Duncan, B.K. & Miller, J.H. Mutagenic deamination of cytosine residues in DNA. *Nature* **287**, 560-561 (1980).

76. Engelward, B.P. *et al.* Cloning and characterization of a mouse 3-methyladenine/7-methyl-guanine/3-methylguanine DNA glycosylase cDNA whose gene maps to chromosome 11. *Carcinogenesis* **14**, 175-181 (1993).

77. Hu, J. *et al.* Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes. *J Biol Chem* **280**, 40544-40551 (2005).

78. Kuzminov, A. Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8241-8246 (2001).

79. Warren, J.J., Forsberg, L.J. & Beese, L.S. The structural basis for the mutagenicity of O(6)-methyl-guanine lesions. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19701-19706 (2006).

80. Boyce, R.P. & Howard-Flanders, P. Release of Ultraviolet Light-Induced Thymine Dimers from DNA in E. Coli K-12. *Proceedings of the National Academy of Sciences of the United States of America* **51**, 293-300 (1964).

81. Cleaver, J.E. Defective repair replication of DNA in xeroderma pigmentosum. *Nature* **218**, 652-656 (1968).

82. Pettijohn, D. & Hanawalt, P. Evidence for Repair-Replication of Ultraviolet Damaged DNA in Bacteria. *J Mol Biol* **9**, 395-410 (1964).

83. Rasmussen, R.E. & Painter, R.B. Evidence for Repair of Ultra-Violet Damaged Deoxyribonucleic Acid in Cultured Mammalian Cells. *Nature* **203**, 1360-1362 (1964).

84. Setlow, R.B. & Carrier, W.L. The Disappearance of Thymine Dimers from DNA: An Error-Correcting Mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **51**, 226-231 (1964).

85. Caldecott, K.W. Protein ADP-ribosylation and the cellular response to DNA strand breaks. *DNA Repair (Amst)* **19**, 108-113 (2014).

86. Morales, J. *et al.* Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases. *Crit Rev Eukaryot Gene Expr* **24**, 15-28 (2014).

87. Davidovic, L., Vodenicharov, M., Affar, E.B. & Poirier, G.G. Importance of poly(ADP-ribose) glycohydrolase in the control of poly(ADP-ribose) metabolism. *Exp Cell Res* **268**, 7-13 (2001).

88. Caldecott, K.W. DNA single-strand break repair. *Exp Cell Res* **329**, 2-8 (2014).

89. Gillet, L.C. & Scharer, O.D. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev* **106**, 253-276 (2006).

90. Hanawalt, P.C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nature reviews. Molecular cell biology* **9**, 958-970 (2008).

91. Scharer, O.D. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol* **5**, a012609 (2013).

92. Sugasawa, K. *et al.* A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev* **15**, 507-521 (2001).

93. Kozmin, S.G. & Jinks-Robertson, S. The mechanism of nucleotide excision repair-mediated UV-induced mutagenesis in nonproliferating cells. *Genetics* **193**, 803-817 (2013).

94. Eckardt, F. & Haynes, R.H. Induction of pure and sectored mutant clones in excision-proficient and deficient strains of yeast. *Mutat Res* **43**, 327-338 (1977).

95. James, A.P. & Kilbey, B.J. The timing of UV mutagenesis in yeast: a pedigree analysis of induced recessive mutation. *Genetics* **87**, 237-248 (1977).

96. James, A.P., Kilbey, B.J. & Prefontaine, G.J. The timing of UV mutagenesis in yeast: continuing mutation in an excision-defective (rad1-1) strain. *Mol Gen Genet* **165**, 207-212 (1978).

97. Eckardt, F., Teh, S.J. & Haynes, R.H. Heteroduplex repair as an intermediate step of UV mutagenesis in yeast. *Genetics* **95**, 63-80 (1980).

98. Cohen-Fix, O. & Livneh, Z. Biochemical analysis of UV mutagenesis in Escherichia coli by using a cell-free reaction coupled to a bioassay: identification of a DNA repair-dependent, replication-independent pathway. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 3300-3304 (1992).

99. Heidenreich, E., Eisler, H., Lengheimer, T., Dorninger, P. & Steinboeck, F. A mutation-promotive role of nucleotide excision repair in cell cycle-arrested cell populations following UV irradiation. *DNA Repair (Amst)* **9**, 96-100 (2010).

100. Sarkar, S., Davies, A.A., Ulrich, H.D. & McHugh, P.J. DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta. *EMBO J* **25**, 1285-1294 (2006).

101. McHugh, P.J., Gill, R.D., Waters, R. & Hartley, J.A. Excision repair of nitrogen mustard-DNA adducts in Saccharomyces cerevisiae. *Nucleic Acids Res* **27**, 3259-3266 (1999).

102. Caldecott, K.W. Single-strand break repair and genetic disease. *Nat Rev Genet* **9**, 619-631 (2008).

103. Caldecott, K.W., Aoufouchi, S., Johnson, P. & Shall, S. XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' in vitro. *Nucleic Acids Res* **24**, 4387-4394 (1996).

104. Lindahl, T. An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 3649-3653 (1974).

105. Jacobs, A.L. & Schar, P. DNA glycosylases: in DNA repair and beyond. *Chromosoma* **121**, 1-20 (2012).

106. Matsumoto, Y. & Kim, K. Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science* **269**, 699-702 (1995).

107. Dianov, G., Price, A. & Lindahl, T. Generation of single-nucleotide repair patches following excision of uracil residues from DNA. *Molecular and cellular biology* **12**, 1605-1612 (1992).

108. Dianov, G. & Lindahl, T. Reconstitution of the DNA base excision-repair pathway. *Curr Biol* **4**, 1069-1076 (1994).

109. Kondo, N., Takahashi, A., Ono, K. & Ohnishi, T. DNA damage induced by alkylating agents and repair pathways. *J Nucleic Acids* **2010**, 543531 (2010).

110. Durant, S.T. *et al.* Dependence on RAD52 and RAD1 for anticancer drug resistance mediated by inactivation of mismatch repair genes. *Curr Biol* **9**, 51-54 (1999).

111. Fink, D., Aebi, S. & Howell, S.B. The role of DNA mismatch repair in drug resistance. *Clin Cancer Res* **4**, 1-6 (1998).

112. Liu, D., Keijzers, G. & Rasmussen, L.J. DNA mismatch repair and its many roles in eukaryotic cells. *Mutat Res* **773**, 174-187 (2017).

113. Wang, H., Lawrence, C.W., Li, G.M. & Hays, J.B. Specific binding of human MSH2.MSH6 mismatch-repair protein heterodimers to DNA incorporating thymine- or uracil-containing UV light photoproducts opposite mismatched bases. *J Biol Chem* **274**, 16894-16900 (1999).

114. Duckett, D.R. *et al.* Human MutSalpha recognizes damaged DNA base pairs containing O6-methylguanine, O4-methylthymine, or the cisplatin-d(GpG) adduct. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6443-6447 (1996).

115. Kunkel, T.A. DNA replication fidelity. *J Biol Chem* **279**, 16895-16898 (2004).

116. Harfe, B.D. & Jinks-Robertson, S. DNA mismatch repair and genetic instability. *Annual review of genetics* **34**, 359-399 (2000).

117. Li, G.M. DNA mismatch repair and cancer. *Front Biosci* **8**, d997-1017 (2003).

118. McCulloch, S.D., Gu, L. & Li, G.M. Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. *J Biol Chem* **278**, 3891-3896 (2003).

119. Surtees, J.A. & Alani, E. Mismatch repair factor MSH2-MSH3 binds and alters the conformation of branched DNA structures predicted to form during genetic recombination. *J Mol Biol* **360**, 523-536 (2006).

120. Jeon, Y. *et al.* Dynamic control of strand excision during human DNA mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 3281-3286 (2016).

121. Kadyrov, F.A., Dzantiev, L., Constantin, N. & Modrich, P. Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126**, 297-308 (2006).

122. Plotz, G., Raedle, J., Brieger, A., Trojan, J. & Zeuzem, S. N-terminus of hMLH1 confers interaction of hMutLalpha and hMutLbeta with hMutSalpha. *Nucleic Acids Res* **31**, 3217-3226 (2003).

123. Plotz, G., Raedle, J., Brieger, A., Trojan, J. & Zeuzem, S. hMutSalpha forms an ATP-dependent complex with hMutLalpha and hMutLbeta on DNA. *Nucleic Acids Res* **30**, 711-718 (2002).

124. Plotz, G., Piiper, A., Wormek, M., Zeuzem, S. & Raedle, J. Analysis of the human MutLalpha.MutSalpha complex. *Biochem Biophys Res Commun* **340**, 852-859 (2006).

125. Schmidt, T.T. & Hombauer, H. Visualization of mismatch repair complexes using fluorescence microscopy. *DNA Repair (Amst)* **38**, 58-67 (2016).

126. Liberti, S.E., Larrea, A.A. & Kunkel, T.A. Exonuclease 1 preferentially repairs mismatches generated by DNA polymerase alpha. *DNA Repair (Amst)* **12**, 92-96 (2013).

127. Pavlov, Y.I., Mian, I.M. & Kunkel, T.A. Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol* **13**, 744-748 (2003).

128. Longley, M.J., Pierce, A.J. & Modrich, P. DNA polymerase delta is required for human mismatch repair in vitro. *J Biol Chem* **272**, 10917-10921 (1997).

129. McVey, M., Khodaverdian, V.Y., Meyer, D., Cerqueira, P.G. & Heyer, W.D. Eukaryotic DNA Polymerases in Homologous Recombination. *Annual review of genetics* **50**, 393-421 (2016).

130. Lord, C.J. & Ashworth, A. BRCAness revisited. *Nat Rev Cancer* **16**, 110-120 (2016).

131. Mizuno, K., Lambert, S., Baldacci, G., Murray, J.M. & Carr, A.M. Nearby inverted repeats fuse to generate acentric and dicentric palindromic chromosomes by a replication template exchange mechanism. *Genes Dev* **23**, 2876-2886 (2009).

132. Payen, C., Koszul, R., Dujon, B. & Fischer, G. Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet* **4**, e1000175 (2008).

133. Slack, A., Thornton, P.C., Magner, D.B., Rosenberg, S.M. & Hastings, P.J. On the mechanism of gene amplification induced under stress in Escherichia coli. *PLoS Genet* **2**, e48 (2006).

134. de Massy, B. Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. *Annual review of genetics* **47**, 563-599 (2013).

135. Sax, K. Chromosome Aberrations Induced by X-Rays. *Genetics* **23**, 494-516 (1938).

136. Yamaguchi, H., Uchihori, Y., Yasuda, N., Takada, M. & Kitamura, H. Estimation of yields of OH radicals in water irradiated by ionizing radiation. *J Radiat Res* **46**, 333-341 (2005).

137. Ward, J.F. DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog Nucleic Acid Res Mol Biol* **35**, 95-125 (1988).

138. Beucher, A. *et al.* ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. *EMBO J* **28**, 3413-3427 (2009).

139. Gaillard, H., Garcia-Muse, T. & Aguilera, A. Replication stress and cancer. *Nat Rev Cancer* **15**, 276-289 (2015).

140. Hall, J.M. *et al.* Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* **250**, 1684-1689 (1990).

141. Hilton, J.L. *et al.* Inactivation of BRCA1 and BRCA2 in ovarian cancer. *J Natl Cancer Inst* **94**, 1396-1406 (2002).

142. Kuchenbaecker, K.B. *et al.* Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA* **317**, 2402-2416 (2017).

143. Lakhani, S.R. *et al.* The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol* **20**, 2310-2318 (2002).

144. Peshkin, B.N., Alabek, M.L. & Isaacs, C. BRCA1/2 mutations and triple negative breast cancers. *Breast Dis* **32**, 25-33 (2010).

145. Heyer, W.D., Ehmsen, K.T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annual review of genetics* **44**, 113-139 (2010).

146. Chang, H.H.Y., Pannunzio, N.R., Adachi, N. & Lieber, M.R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews. Molecular cell biology* **18**, 495-506 (2017).

147. Badie, S. *et al.* BRCA2 acts as a RAD51 loader to facilitate telomere replication and capping. *Nat Struct Mol Biol* **17**, 1461-1469 (2010).

148. Carr, A.M. & Lambert, S. Replication stress-induced genome instability: the dark side of replication maintenance by homologous recombination. *J Mol Biol* **425**, 4733-4744 (2013).

149. Cannavo, E. & Cejka, P. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature* **514**, 122-125 (2014).

150. Huertas, P., Cortes-Ledesma, F., Sartori, A.A., Aguilera, A. & Jackson, S.P. CDK targets Sae2 to control DNA-end resection and homologous recombination. *Nature* **455**, 689-692 (2008).

151. Malumbres, M. Cyclin-dependent kinases. *Genome Biol* **15**, 122 (2014).

152. Chen, X. *et al.* Enrichment of Cdk1-cyclins at DNA double-strand breaks stimulates Fun30 phosphorylation and DNA end resection. *Nucleic Acids Res* **44**, 2742-2753 (2016).

153. Ferrari, M. *et al.* Functional interplay between the 53BP1-ortholog Rad9 and the Mre11 complex regulates resection, end-tethering and repair of a double-strand break. *PLoS Genet* **11**, e1004928 (2015).

154. Yu, X., Wu, L.C., Bowcock, A.M., Aronheim, A. & Baer, R. The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. *J Biol Chem* **273**, 25388-25392 (1998).

155. Yu, X. & Chen, J. DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. *Molecular and cellular biology* **24**, 9478-9486 (2004).

156. Stark, J.M., Pierce, A.J., Oh, J., Pastink, A. & Jasin, M. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Molecular and cellular biology* **24**, 9305-9316 (2004).

157. Schlegel, B.P., Jodelka, F.M. & Nunez, R. BRCA1 promotes induction of ssDNA by ionizing radiation. *Cancer Res* **66**, 5181-5189 (2006).

158. Cruz-Garcia, A., Lopez-Saavedra, A. & Huertas, P. BRCA1 accelerates CtIP-mediated DNA-end resection. *Cell Rep* **9**, 451-459 (2014).

159. Tomimatsu, N. *et al.* Phosphorylation of EXO1 by CDKs 1 and 2 regulates DNA end resection and repair pathway choice. *Nat Commun* **5**, 3561 (2014).

160. Wohlbold, L. *et al.* Chemical genetics reveals a specific requirement for Cdk2 activity in the DNA damage response and identifies Nbs1 as a Cdk2 substrate in human cells. *PLoS Genet* **8**, e1002935 (2012).

161. Falck, J. *et al.* CDK targeting of NBS1 promotes DNA-end resection, replication restart and homologous recombination. *EMBO Rep* **13**, 561-568 (2012).

162. Steger, M. *et al.* Prolyl isomerase PIN1 regulates DNA double-strand break repair by counteracting DNA end resection. *Molecular cell* **50**, 333-343 (2013).

163. Ivanov, E.L., Sugawara, N., Fishman-Lobell, J. & Haber, J.E. Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae. *Genetics* **142**, 693-704 (1996).

164. Ma, J.L., Kim, E.M., Haber, J.E. & Lee, S.E. Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. *Molecular and cellular biology* **23**, 8820-8828 (2003).

165. Pannunzio, N.R., Watanabe, G. & Lieber, M.R. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *J Biol Chem* **293**, 10512-10523 (2018).

166. Lee, S.E. *et al.* Saccharomyces Ku70, mre11/rad50 and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* **94**, 399-409 (1998).

167. Tutt, A. *et al.* Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. *EMBO J* **20**, 4704-4716 (2001).

168. Dronkert, M.L. *et al.* Mouse RAD54 affects DNA double-strand break repair and sister chromatid exchange. *Molecular and cellular biology* **20**, 3147-3156 (2000).

169. Falck, J., Coates, J. & Jackson, S.P. Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* **434**, 605-611 (2005).

170. Kobayashi, J., Antoccia, A., Tauchi, H., Matsuura, S. & Komatsu, K. NBS1 and its functional role in the DNA damage response. *DNA Repair (Amst)* **3**, 855-861 (2004).

171. Kobayashi, J. *et al.* NBS1 localizes to gamma-H2AX foci through interaction with the FHA/BRCT domain. *Curr Biol* **12**, 1846-1851 (2002).

172. Rogakou, E.P., Boon, C., Redon, C. & Bonner, W.M. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *The Journal of cell biology* **146**, 905-916 (1999).

173. FitzGerald, J.E., Grenon, M. & Lowndes, N.F. 53BP1: function and mechanisms of focal recruitment. *Biochem Soc Trans* **37**, 897-904 (2009).

174. Bouwman, P. *et al.* 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. *Nat Struct Mol Biol* **17**, 688-695 (2010).

175. Bunting, S.F. *et al.* 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141**, 243-254 (2010).

176. Chapman, J.R., Sossick, A.J., Boulton, S.J. & Jackson, S.P. BRCA1-associated exclusion of 53BP1 from DNA damage sites underlies temporal control of DNA repair. *J Cell Sci* **125**, 3529-3534 (2012).

177. Chung, W.H., Zhu, Z., Papusha, A., Malkova, A. & Ira, G. Defective resection at DNA double-strand breaks leads to de novo telomere formation and enhances gene targeting. *PLoS Genet* **6**, e1000948 (2010).

178. Symington, L.S. End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb Perspect Biol* **6** (2014).

179. Mimitou, E.P. & Symington, L.S. DNA end resection: many nucleases make light work. *DNA Repair (Amst)* **8**, 983-995 (2009).

180. Tsukamoto, Y., Mitsuoka, C., Terasawa, M., Ogawa, H. & Ogawa, T. Xrs2p regulates Mre11p translocation to the nucleus and plays a role in telomere elongation and meiotic recombination. *Mol Biol Cell* **16**, 597-608 (2005).

181. Anderson, D.E., Trujillo, K.M., Sung, P. & Erickson, H.P. Structure of the Rad50 x Mre11 DNA repair complex from Saccharomyces cerevisiae by electron microscopy. *J Biol Chem* **276**, 37027-37033 (2001).

182. de Jager, M. *et al.* Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Molecular cell* **8**, 1129-1135 (2001).

183. Hopfner, K.P. *et al.* The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature* **418**, 562-566 (2002).

184. Lim, H.S., Kim, J.S., Park, Y.B., Gwon, G.H. & Cho, Y. Crystal structure of the Mre11-Rad50-ATPgammaS complex: understanding the interplay between Mre11 and Rad50. *Genes Dev* **25**, 1091-1104 (2011).

185. Mockel, C., Lammens, K., Schele, A. & Hopfner, K.P. ATP driven structural changes of the bacterial Mre11:Rad50 catalytic head complex. *Nucleic Acids Res* **40**, 914-927 (2012).

186. Bernstein, K.A., Gangloff, S. & Rothstein, R. The RecQ DNA helicases in DNA repair. *Annual review of genetics* **44**, 393-417 (2010).

187. Szankasi, P. & Smith, G.R. A DNA exonuclease induced during meiosis of Schizosaccharomyces pombe. *J Biol Chem* **267**, 3014-3023 (1992).

188. Tran, P.T., Erdeniz, N., Symington, L.S. & Liskay, R.M. EXO1-A multi-tasking eukaryotic nuclease. *DNA Repair (Amst)* **3**, 1549-1559 (2004).

189. Cannavo, E., Cejka, P. & Kowalczykowski, S.C. Relationship of DNA degradation by Saccharomyces cerevisiae exonuclease 1 and its stimulation by RPA and Mre11-Rad50-Xrs2 to DNA end resection. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E1661-1668 (2013).

190. Cejka, P. *et al.* DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature* **467**, 112-116 (2010).

191. Niu, H. *et al.* Mechanism of the ATP-dependent DNA end-resection machinery from Saccharomyces cerevisiae. *Nature* **467**, 108-111 (2010).

192. Chen, H., Lisby, M. & Symington, L.S. RPA coordinates DNA end resection and prevents formation of DNA hairpins. *Molecular cell* **50**, 589-600 (2013).

193. Chiolo, I. *et al.* Srs2 and Sgs1 DNA helicases associate with Mre11 in different subcomplexes following checkpoint activation and CDK1-mediated Srs2 phosphorylation. *Molecular and cellular biology* **25**, 5738-5751 (2005).

194. Shim, E.Y. *et al.* Saccharomyces cerevisiae Mre11/Rad50/Xrs2 and Ku proteins regulate association of Exo1 and Dna2 with DNA breaks. *EMBO J* **29**, 3370-3380 (2010).

195. Eid, W. *et al.* DNA end resection by CtIP and exonuclease 1 prevents genomic instability. *EMBO Rep* **11**, 962-968 (2010).

196. Bhargava, R., Onyango, D.O. & Stark, J.M. Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends Genet* **32**, 566-575 (2016).

197. Christensen, D.E., Brzovic, P.S. & Klevit, R.E. E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. *Nat Struct Mol Biol* **14**, 941-948 (2007).

198. Prakash, R., Zhang, Y., Feng, W. & Jasin, M. Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harb Perspect Biol* **7**, a016600 (2015).

199. Mavaddat, N. *et al.* Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). *Cancer Epidemiol Biomarkers Prev* **21**, 134-147 (2012).

200. Wu, L.C. *et al.* Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat Genet* **14**, 430-440 (1996).

201. Brzovic, P.S., Rajagopal, P., Hoyt, D.W., King, M.C. & Klevit, R.E. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nat Struct Biol* **8**, 833-837 (2001).

202. Wu-Baer, F., Lagrazon, K., Yuan, W. & Baer, R. The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *J Biol Chem* **278**, 34743-34746 (2003).

203. Nishikawa, H. *et al.* Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. *J Biol Chem* **279**, 3916-3924 (2004).

204. Kelsey, J.L. A review of the epidemiology of human breast cancer. *Epidemiol Rev* **1**, 74-109 (1979).

205. Claus, E.B., Risch, N.J. & Thompson, W.D. Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol* **131**, 961-972 (1990).

206. Adami, H.O., Hansen, J., Jung, B. & Rimsten, A. Familiality in breast cancer: a case-control study in a Sweden population. *Br J Cancer* **42**, 71-77 (1980).

207. Narod, S.A. *et al.* Familial breast-ovarian cancer locus on chromosome 17q12-q23. *Lancet* **338**, 82-83 (1991).

208. Wooster, R. *et al.* Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science* **265**, 2088-2090 (1994).

209. Easton, D.F., Bishop, D.T., Ford, D. & Crockford, G.P. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* **52**, 678-701 (1993).

210. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).

211. Castilla, L.H. *et al.* Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer. *Nat Genet* **8**, 387-391 (1994).

212. Friedman, L.S. *et al.* Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nat Genet* **8**, 399-404 (1994).

213. Struewing, J.P. *et al.* Detection of eight BRCA1 mutations in 10 breast/ovarian cancer families, including 1 family with male breast cancer. *Am J Hum Genet* **57**, 1-7 (1995).

214. Simard, J. *et al.* Common origins of BRCA1 mutations in Canadian breast and ovarian cancer families. *Nat Genet* **8**, 392-398 (1994).

215. Struewing, J.P. *et al.* The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nat Genet* **11**, 198-200 (1995).

216. Ford, D., Easton, D.F., Bishop, D.T., Narod, S.A. & Goldgar, D.E. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet* **343**, 692-695 (1994).

217. Scully, R. *et al.* Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* **90**, 425-435 (1997).

218. Scully, R. *et al.* BRCA1 is a component of the RNA polymerase II holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 5605-5610 (1997).

219. Scully, R. *et al.* Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**, 265-275 (1997).

220. Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C.C. p53 mutations in human cancers. *Science* **253**, 49-53 (1991).

221. Yuan, S.S. *et al.* BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res* **59**, 3547-3551 (1999).

222. Bochkareva, E., Korolev, S., Lees-Miller, S.P. & Bochkarev, A. Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J* **21**, 1855-1863 (2002).

223. Bochkareva, E., Belegu, V., Korolev, S. & Bochkarev, A. Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding. *EMBO J* **20**, 612-618 (2001).

224. Bastin-Shanower, S.A. & Brill, S.J. Functional analysis of the four DNA binding domains of replication protein A. The role of RPA2 in ssDNA binding. *J Biol Chem* **276**, 36446-36453 (2001).

225. Weisshart, K. *et al.* Partial proteolysis of simian virus 40 T antigen reveals intramolecular contacts between domains and conformation changes upon hexamer assembly. *J Biol Chem* **279**, 38943-38951 (2004).

226. Stauffer, M.E. & Chazin, W.J. Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA. *J Biol Chem* **279**, 25638-25645 (2004).

227. Blackwell, L.J., Borowiec, J.A. & Mastrangelo, I.A. Single-stranded-DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase. *Molecular and cellular biology* **16**, 4798-4807 (1996).

228. Fanning, E., Klimovich, V. & Nager, A.R. A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res* **34**, 4126-4137 (2006).

229. Jensen, R.B., Ozes, A., Kim, T., Estep, A. & Kowalczykowski, S.C. BRCA2 is epistatic to the RAD51 paralogs in response to DNA damage. *DNA Repair (Amst)* **12**, 306-311 (2013).

230. Thorslund, T. *et al.* The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat Struct Mol Biol* **17**, 1263-1265 (2010).

231. Liu, J., Doty, T., Gibson, B. & Heyer, W.D. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat Struct Mol Biol* **17**, 1260-1262 (2010).

232. Morimatsu, K. & Kowalczykowski, S.C. RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Molecular cell* **11**, 1337-1347 (2003).

233. Yang, H., Li, Q., Fan, J., Holloman, W.K. & Pavletich, N.P. The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature* **433**, 653-657 (2005).

234. Wright, W.D., Shah, S.S. & Heyer, W.D. Homologous recombination and the repair of DNA double-strand breaks. *J Biol Chem* **293**, 10524-10535 (2018).

235. New, J.H., Sugiyama, T., Zaitseva, E. & Kowalczykowski, S.C. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* **391**, 407-410 (1998).

236. Song, B. & Sung, P. Functional interactions among yeast Rad51 recombinase, Rad52 mediator, and replication protein A in DNA strand exchange. *J Biol Chem* **275**, 15895-15904 (2000).

237. Yuan, R. *et al.* Coordinate alterations in the expression of BRCA1, BRCA2, p300, and Rad51 in response to genotoxic and other stresses in human prostate cancer cells. *Prostate* **40**, 37-49 (1999).

238. Whelan, D.R. *et al.* Spatiotemporal dynamics of homologous recombination repair at single collapsed replication forks. *Nat Commun* **9**, 3882 (2018).

239. Thompson, L.H. & Schild, D. The contribution of homologous recombination in preserving genome integrity in mammalian cells. *Biochimie* **81**, 87-105 (1999).

240. Thacker, J. A surfeit of RAD51-like genes? *Trends Genet* **15**, 166-168 (1999).

241. Bignell, G., Micklem, G., Stratton, M.R., Ashworth, A. & Wooster, R. The BRC repeats are conserved in mammalian BRCA2 proteins. *Hum Mol Genet* **6**, 53-58 (1997).

242. Carreira, A. & Kowalczykowski, S.C. Two classes of BRC repeats in BRCA2 promote RAD51 nucleoprotein filament function by distinct mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 10448-10453 (2011).

243. Foulkes, W.D. *et al.* Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J Natl Cancer Inst* **95**, 1482-1485 (2003).

244. Crook, T. *et al.* p53 mutation with frequent novel condons but not a mutator phenotype in BRCA1- and BRCA2-associated breast tumours. *Oncogene* **17**, 1681-1689 (1998).

245. Lakhani, S.R. *et al.* Pathology of ovarian cancers in BRCA1 and BRCA2 carriers. *Clin Cancer Res* **10**, 2473-2481 (2004).

246. Palacios, J. *et al.* Immunohistochemical characteristics defined by tissue microarray of hereditary breast cancer not attributable to BRCA1 or BRCA2 mutations: differences from breast carcinomas arising in BRCA1 and BRCA2 mutation carriers. *Clin Cancer Res* **9**, 3606-3614 (2003).

247. Lakhani, S.R. *et al.* Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. *Clin Cancer Res* **11**, 5175-5180 (2005).

248. Sorlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8418-8423 (2003).

249. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. *Lancet* **349**, 1505-1510 (1997).

250. Yang, R.L. *et al.* DCIS in BRCA1 and BRCA2 mutation carriers: prevalence, phenotype, and expression of oncodrivers C-MET and HER3. *J Transl Med* **13**, 335 (2015).

251. Xia, B. *et al.* Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Molecular cell* **22**, 719-729 (2006).

252. Oliver, A.W., Swift, S., Lord, C.J., Ashworth, A. & Pearl, L.H. Structural basis for recruitment of BRCA2 by PALB2. *EMBO Rep* **10**, 990-996 (2009).

253. Sy, S.M., Huen, M.S. & Chen, J. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7155-7160 (2009).

254. Zhang, F., Fan, Q., Ren, K. & Andreassen, P.R. PALB2 functionally connects the breast cancer susceptibility proteins BRCA1 and BRCA2. *Mol Cancer Res* **7**, 1110-1118 (2009).

255. Zhang, F. *et al.* PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol* **19**, 524-529 (2009).

256. Buisson, R. *et al.* Cooperation of breast cancer proteins PALB2 and piccolo BRCA2 in stimulating homologous recombination. *Nat Struct Mol Biol* **17**, 1247-1254 (2010).

257. He, J., Kang, X., Yin, Y., Chao, K.S. & Shen, W.H. PTEN regulates DNA replication progression and stalled fork recovery. *Nat Commun* **6**, 7620 (2015).

258. Cancer Genome Atlas Research, N. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-615 (2011).

259. Mendes-Pereira, A.M. *et al.* Synthetic lethal targeting of PTEN mutant cells with PARP inhibitors. *EMBO Mol Med* **1**, 315-322 (2009).

260. Lim, D.S. & Hasty, P. A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Molecular and cellular biology* **16**, 7133-7143 (1996).

261. Tsuzuki, T. *et al.* Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6236-6240 (1996).

262. Deans, B., Griffin, C.S., Maconochie, M. & Thacker, J. Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. *EMBO J* **19**, 6675-6685 (2000).

263. Pittman, D.L. & Schimenti, J.C. Midgestation lethality in mice deficient for the RecA-related gene, Rad51d/Rad51l3. *Genesis* **26**, 167-173 (2000).

264. Shu, Z., Smith, S., Wang, L., Rice, M.C. & Kmiec, E.B. Disruption of muREC2/RAD51L1 in mice results in early embryonic lethality which can Be partially rescued in a p53(-/-) background. *Molecular and cellular biology* **19**, 8686-8693 (1999).

265. Cartwright, R., Tambini, C.E., Simpson, P.J. & Thacker, J. The XRCC2 DNA repair gene from human and mouse encodes a novel member of the recA/RAD51 family. *Nucleic Acids Res* **26**, 3084-3089 (1998).

266. Adam, J., Deans, B. & Thacker, J. A role for Xrcc2 in the early stages of mouse development. *DNA Repair (Amst)* **6**, 224-234 (2007).

267. Liu, N., Schild, D., Thelen, M.P. & Thompson, L.H. Involvement of Rad51C in two distinct protein complexes of Rad51 paralogs in human cells. *Nucleic Acids Res* **30**, 1009-1015 (2002).

268. Masson, J.Y. *et al.* Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev* **15**, 3296-3307 (2001).

269. Liu, Y., Tarsounas, M., O'Regan, P. & West, S.C. Role of RAD51C and XRCC3 in genetic recombination and DNA repair. *J Biol Chem* **282**, 1973-1979 (2007).

270. Kurumizaka, H. *et al.* Homologous pairing and ring and filament structure formation activities of the human Xrcc2*Rad51D complex. *J Biol Chem* **277**, 14315-14320 (2002).

271. Schild, D., Lio, Y.C., Collins, D.W., Tsomondo, T. & Chen, D.J. Evidence for simultaneous protein interactions between human Rad51 paralogs. *J Biol Chem* **275**, 16443-16449 (2000).

272. Sigurdsson, S. *et al.* Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. *Genes Dev* **15**, 3308-3318 (2001).

273. Miller, K.A., Sawicka, D., Barsky, D. & Albala, J.S. Domain mapping of the Rad51 paralog protein complexes. *Nucleic Acids Res* **32**, 169-178 (2004).

274. Masson, J.Y., Stasiak, A.Z., Stasiak, A., Benson, F.E. & West, S.C. Complex formation by the human RAD51C and XRCC3 recombination repair proteins. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8440-8446 (2001).

275. Kurumizaka, H. *et al.* Homologous-pairing activity of the human DNA-repair proteins Xrcc3.Rad51C. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 5538-5543 (2001).

276. Martino, J. & Bernstein, K.A. The Shu complex is a conserved regulator of homologous recombination. *FEMS Yeast Res* **16** (2016).

277. McClendon, T.B., Sullivan, M.R., Bernstein, K.A. & Yanowitz, J.L. Promotion of Homologous Recombination by SWS-1 in Complex with RAD-51 Paralogs in Caenorhabditis elegans. *Genetics* **203**, 133-145 (2016).

278. Sullivan, M.R. & Bernstein, K.A. RAD-ical New Insights into RAD51 Regulation. *Genes (Basel)* **9**, 629 (2018).

279. Park, J.Y. *et al.* Breast cancer-associated missense mutants of the PALB2 WD40 domain, which directly binds RAD51C, RAD51 and BRCA2, disrupt DNA repair. *Oncogene* **33**, 4803-4812 (2014).

280. Chun, J., Buechelmaier, E.S. & Powell, S.N. Rad51 paralog complexes BCDX2 and CX3 act at different stages in the BRCA1-BRCA2-dependent homologous recombination pathway. *Molecular and cellular biology* **33**, 387-395 (2013).

281. Gaines, W.A. *et al.* Promotion of presynaptic filament assembly by the ensemble of S. cerevisiae Rad51 paralogues with Rad52. *Nat Commun* **6**, 7834 (2015).

282. Godin, S.K. *et al.* The Shu complex promotes error-free tolerance of alkylation-induced base excision repair products. *Nucleic Acids Res* **44**, 8199-8215 (2016).

283. Qi, Z. *et al.* DNA sequence alignment by microhomology sampling during homologous recombination. *Cell* **160**, 856-869 (2015).

284. Lee, J.Y. *et al.* DNA RECOMBINATION. Base triplet stepping by the Rad51/RecA family of recombinases. *Science* **349**, 977-981 (2015).

285. Smith, M.J. & Rothstein, R. Poetry in motion: Increased chromosomal mobility after DNA damage. *DNA Repair (Amst)* **56**, 102-108 (2017).

286. Liberi, G. & Foiani, M. The double life of Holliday junctions. *Cell Res* **20**, 611-613 (2010).

287. Chu, W.K. & Hickson, I.D. RecQ helicases: multifunctional genome caretakers. *Nat Rev Cancer* **9**, 644-654 (2009).

288. Li, X. & Heyer, W.D. RAD54 controls access to the invading 3'-OH end after RAD51-mediated DNA strand invasion in homologous recombination in Saccharomyces cerevisiae. *Nucleic Acids Res* **37**, 638-646 (2009).

289. Sebesta, M., Burkovics, P., Haracska, L. & Krejci, L. Reconstitution of DNA repair synthesis in vitro and the role of polymerase and helicase activities. *DNA Repair (Amst)* **10**, 567-576 (2011).

290. Fabre, F., Boulet, A. & Faye, G. Possible involvement of the yeast POLIII DNA polymerase in induced gene conversion. *Mol Gen Genet* **229**, 353-356 (1991).

291. Moldovan, G.L., Pfander, B. & Jentsch, S. PCNA, the maestro of the replication fork. *Cell* **129**, 665-679 (2007).

292. Goodman, M.F. & Woodgate, R. Translesion DNA polymerases. *Cold Spring Harb Perspect Biol* **5**, a010363 (2013).

293. Kesti, T., Flick, K., Keranen, S., Syvaoja, J.E. & Wittenberg, C. DNA polymerase epsilon catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. *Molecular cell* **3**, 679-685 (1999).

294. Pathania, S. *et al.* BRCA1 haploinsufficiency for replication stress suppression in primary cells. *Nat Commun* **5**, 5496 (2014).

295. Petermann, E., Orta, M.L., Issaeva, N., Schultz, N. & Helleday, T. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Molecular cell* **37**, 492-502 (2010).

296. Lundin, C. *et al.* Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells. *Molecular and cellular biology* **22**, 5869-5878 (2002).

297. Helleday, T. Pathways for mitotic homologous recombination in mammalian cells. *Mutat Res* **532**, 103-115 (2003).

298. Zhang, J. The role of BRCA1 in homologous recombination repair in response to replication stress: significance in tumorigenesis and cancer therapy. *Cell Biosci* **3**, 11 (2013).

299. Laulier, C., Cheng, A. & Stark, J.M. The relative efficiency of homology-directed repair has distinct effects on proper anaphase chromosome separation. *Nucleic Acids Res* **39**, 5935-5944 (2011).

300. Boos, D., Frigola, J. & Diffley, J.F. Activation of the replicative DNA helicase: breaking up is hard to do. *Curr Opin Cell Biol* **24**, 423-430 (2012).

301. Pan, Q., Fang, Y., Xu, Y., Zhang, K. & Hu, X. Down-regulation of DNA polymerases kappa, eta, iota, and zeta in human lung, stomach, and colorectal cancers. *Cancer Lett* **217**, 139-147 (2005).

302. Toledo, L.I. *et al.* ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell* **155**, 1088-1103 (2013).

303. Zou, L. & Elledge, S.J. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300**, 1542-1548 (2003).

304. Jazayeri, A. *et al.* ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nature cell biology* **8**, 37-45 (2006).

305. Pathania, S. *et al.* BRCA1 is required for postreplication repair after UV-induced DNA damage. *Molecular cell* **44**, 235-251 (2011).

306. Adamo, A. *et al.* Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Molecular cell* **39**, 25-35 (2010).

307. Lomonosov, M., Anand, S., Sangrithi, M., Davies, R. & Venkitaraman, A.R. Stabilization of stalled DNA replication forks by the BRCA2 breast cancer susceptibility protein. *Genes Dev* **17**, 3017-3022 (2003).

308. Malu, S., Malshetty, V., Francis, D. & Cortes, P. Role of non-homologous end joining in V(D)J recombination. *Immunol Res* **54**, 233-246 (2012).

309. Mari, P.O. *et al.* Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 18597-18602 (2006).

310. Downs, J.A. & Jackson, S.P. A means to a DNA end: the many roles of Ku. *Nature reviews. Molecular cell biology* **5**, 367-378 (2004).

311. Mimori, T., Hardin, J.A. & Steitz, J.A. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J Biol Chem* **261**, 2274-2278 (1986).

312. Walker, J.R., Corpina, R.A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607-614 (2001).

313. Blier, P.R., Griffith, A.J., Craft, J. & Hardin, J.A. Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *J Biol Chem* **268**, 7594-7601 (1993).

314. Grawunder, U., Zimmer, D. & Lieber, M.R. DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains. *Curr Biol* **8**, 873-876 (1998).

315. Biehs, R. *et al.* DNA Double-Strand Break Resection Occurs during Non-homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination. *Molecular cell* **65**, 671-684 e675 (2017).

316. Goodarzi, A.A. *et al.* DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *EMBO J* **25**, 3880-3889 (2006).

317. Gu, J. *et al.* DNA-PKcs regulates a single-stranded DNA endonuclease activity of Artemis. *DNA Repair (Amst)* **9**, 429-437 (2010).

318. Ma, Y. *et al.* The DNA-dependent protein kinase catalytic subunit phosphorylation sites in human Artemis. *J Biol Chem* **280**, 33839-33846 (2005).

319. Ma, Y. *et al.* A biochemically defined system for mammalian nonhomologous DNA end joining. *Molecular cell* **16**, 701-713 (2004).

320. Lu, H. *et al.* A biochemically defined system for coding joint formation in V(D)J recombination. *Molecular cell* **31**, 485-497 (2008).

321. Chang, H.H., Watanabe, G. & Lieber, M.R. Unifying the DNA end-processing roles of the artemis nuclease: Ku-dependent artemis resection at blunt DNA ends. *J Biol Chem* **290**, 24036-24050 (2015).

322. Blommers, M.J. *et al.* Effects of base sequence on the loop folding in DNA hairpins. *Biochemistry* **28**, 7491-7498 (1989).

323. Nick McElhinny, S.A. *et al.* A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Molecular cell* **19**, 357-366 (2005).

324. Moon, A.F. *et al.* Structural insight into the substrate specificity of DNA Polymerase mu. *Nat Struct Mol Biol* **14**, 45-53 (2007).

325. Gu, J. *et al.* XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *EMBO J* **26**, 1010-1023 (2007).

326. Tsai, C.J., Kim, S.A. & Chu, G. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7851-7856 (2007).

327. Petrelli, F., Barni, S., Bregni, G., de Braud, F. & Di Cosimo, S. Platinum salts in advanced breast cancer: a systematic review and meta-analysis of randomized clinical trials. *Breast Cancer Res Treat* **160**, 425-437 (2016).

328. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-921 (2005).

329. Helleday, T., Bryant, H.E. & Schultz, N. Poly(ADP-ribose) polymerase (PARP-1) in homologous recombination and as a target for cancer therapy. *Cell Cycle* **4**, 1176-1178 (2005).

330. Cruz, C. *et al.* RAD51 foci as a functional biomarker of homologous recombination repair and PARP inhibitor resistance in germline BRCA-mutated breast cancer. *Ann Oncol* **29**, 1203-1210 (2018).

331. De Vos, M., Schreiber, V. & Dantzer, F. The diverse roles and clinical relevance of PARPs in DNA damage repair: current state of the art. *Biochem Pharmacol* **84**, 137-146 (2012).

332. Du, Y. *et al.* Blocking c-Met-mediated PARP1 phosphorylation enhances anti-tumor effects of PARP inhibitors. *Nat Med* **22**, 194-201 (2016).

333. Min, A. *et al.* RAD51C-deficient cancer cells are highly sensitive to the PARP inhibitor olaparib. *Mol Cancer Ther* **12**, 865-877 (2013).

334. Murata, S. *et al.* Predictors and Modulators of Synthetic Lethality: An Update on PARP Inhibitors and Personalized Medicine. *Biomed Res Int* **2016**, 2346585 (2016).

335. Wang, J. *et al.* Loss of CtIP disturbs homologous recombination repair and sensitizes breast cancer cells to PARP inhibitors. *Oncotarget* **7**, 7701-7714 (2016).

336. McCabe, N. *et al.* Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. *Cancer Res* **66**, 8109-8115 (2006).

337. Bryant, H.E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913-917 (2005).

338. Rugo, H.S. *et al.* Adaptive Randomization of Veliparib-Carboplatin Treatment in Breast Cancer. *The New England journal of medicine* **375**, 23-34 (2016).

339. Ledermann, J. *et al.* Olaparib maintenance therapy in platinum-sensitive relapsed ovarian cancer. *The New England journal of medicine* **366**, 1382-1392 (2012).

340. Feng, Z. *et al.* Rad52 inactivation is synthetically lethal with BRCA2 deficiency. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 686-691 (2011).

341. Crews, S., Franks, R., Hu, S., Matthews, B. & Nambu, J. Drosophila single-minded gene and the molecular genetics of CNS midline development. *J Exp Zool* **261**, 234-244 (1992).

342. Crews, S.T., Thomas, J.B. & Goodman, C.S. The Drosophila single-minded gene encodes a nuclear protein with sequence similarity to the per gene product. *Cell* **52**, 143-151 (1988).

343. Nambu, J.R., Franks, R.G., Hu, S. & Crews, S.T. The single-minded gene of Drosophila is required for the expression of genes important for the development of CNS midline cells. *Cell* **63**, 63-75 (1990).

344. Nambu, J.R., Lewis, J.O., Wharton, K.A., Jr. & Crews, S.T. The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. *Cell* **67**, 1157-1167 (1991).

345. Thomas, J.B., Crews, S.T. & Goodman, C.S. Molecular genetics of the single-minded locus: a gene involved in the development of the Drosophila nervous system. *Cell* **52**, 133-141 (1988).

346. Ema, M. *et al.* Two new members of the murine Sim gene family are transcriptional repressors and show different expression patterns during mouse embryogenesis. *Molecular and cellular biology* **16**, 5865-5875 (1996).

347. Moffett, P., Reece, M. & Pelletier, J. The murine Sim-2 gene product inhibits transcription by active repression and functional interference. *Molecular and cellular biology* **17**, 4933-4947 (1997).

348. Fan, C.M. *et al.* Expression patterns of two murine homologs of Drosophila single-minded suggest possible roles in embryonic patterning and in the pathogenesis of Down syndrome. *Mol Cell Neurosci* **7**, 1-16 (1996).

349. Metz, R.P., Kwak, H.I., Gustafson, T., Laffin, B. & Porter, W.W. Differential transcriptional regulation by mouse single-minded 2s. *J Biol Chem* **281**, 10839-10848 (2006).

350. Shamblott, M.J., Bugg, E.M., Lawler, A.M. & Gearhart, J.D. Craniofacial abnormalities resulting from targeted disruption of the murine Sim2 gene. *Developmental dynamics : an official publication of the American Association of Anatomists* **224**, 373-380 (2002).

351. Kewley, R.J., Whitelaw, M.L. & Chapman-Smith, A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *Int J Biochem Cell Biol* **36**, 189-204 (2004).

352. Bersten, D.C., Sullivan, A.E., Peet, D.J. & Whitelaw, M.L. bHLH-PAS proteins in cancer. *Nat Rev Cancer* **13**, 827-841 (2013).

353. Pongratz, I., Antonsson, C., Whitelaw, M.L. & Poellinger, L. Role of the PAS domain in regulation of dimerization and DNA binding specificity of the dioxin receptor. *Molecular and cellular biology* **18**, 4079-4088 (1998).

354. Crews, S.T. Control of cell lineage-specific development and transcription by bHLH-PAS proteins. *Genes Dev* **12**, 607-620 (1998).

355. Crews, S.T. & Fan, C.M. Remembrance of things PAS: regulation of development by bHLH-PAS proteins. *Curr Opin Genet Dev* **9**, 580-587 (1999).

356. Taylor, B.L. & Zhulin, I.B. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* **63**, 479-506 (1999).

357. Hao, N. *et al.* Xenobiotics and loss of cell adhesion drive distinct transcriptional outcomes by aryl hydrocarbon receptor signaling. *Mol Pharmacol* **82**, 1082-1093 (2012).

358. Nguyen, L.P. & Bradfield, C.A. The search for endogenous activators of the aryl hydrocarbon receptor. *Chem Res Toxicol* **21**, 102-116 (2008).

359. Huang, N. *et al.* Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. *Science* **337**, 189-194 (2012).

360. Sato, T.K. *et al.* Feedback repression is required for mammalian circadian clock function. *Nat Genet* **38**, 312-319 (2006).

361. Okui, M. *et al.* Transcription factor single-minded 2 (SIM2) is ubiquitinated by the RING-IBR-RING-type E3 ubiquitin ligases. *Exp Cell Res* **309**, 220-228 (2005).

362. McIntosh, B.E., Hogenesch, J.B. & Bradfield, C.A. Mammalian Per-Arnt-Sim proteins in environmental adaptation. *Annu Rev Physiol* **72**, 625-645 (2010).

363. Woods, S.L. & Whitelaw, M.L. Differential activities of murine single minded 1 (SIM1) and SIM2 on a hypoxic response element. Cross-talk between basic helix-loop-helix/per-Arnt-Sim homology transcription factors. *J Biol Chem* **277**, 10236-10243 (2002).

364. Jain, S., Dolwick, K.M., Schmidt, J.V. & Bradfield, C.A. Potent transactivation domains of the Ah receptor and the Ah receptor nuclear translocator map to their carboxyl termini. *J Biol Chem* **269**, 31518-31524 (1994).

365. Farrall, A.L. & Whitelaw, M.L. The HIF1alpha-inducible pro-cell death gene BNIP3 is a novel target of SIM2s repression through cross-talk on the hypoxia response element. *Oncogene* **28**, 3671-3680 (2009).

366. Bertout, J.A., Patel, S.A. & Simon, M.C. The impact of O2 availability on human cancer. *Nat Rev Cancer* **8**, 967-975 (2008).

367. Mahon, P.C., Hirota, K. & Semenza, G.L. FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity. *Genes Dev* **15**, 2675-2686 (2001).

368. An, W.G. *et al.* Stabilization of wild-type p53 by hypoxia-inducible factor 1alpha. *Nature* **392**, 405-408 (1998).

369. Gordan, J.D. *et al.* HIF-alpha effects on c-Myc distinguish two subtypes of sporadic VHL-deficient clear cell renal carcinoma. *Cancer Cell* **14**, 435-446 (2008).

370. Majmundar, A.J., Wong, W.J. & Simon, M.C. Hypoxia-inducible factors and the response to hypoxic stress. *Molecular cell* **40**, 294-309 (2010).

371. Gordan, J.D., Thompson, C.B. & Simon, M.C. HIF and c-Myc: sibling rivals for control of cancer cell metabolism and proliferation. *Cancer Cell* **12**, 108-113 (2007).

372. Liberti, M.V. *et al.* A Predictive Model for Selective Targeting of the Warburg Effect through GAPDH Inhibition with a Natural Product. *Cell Metab* **26**, 648-659 e648 (2017).

373. Warburg, O., Wind, F. & Negelein, E. The Metabolism of Tumors in the Body. *J Gen Physiol* **8**, 519-530 (1927).

374. Lu, B., Asara, J.M., Sanda, M.G. & Arredouani, M.S. The role of the transcription factor SIM2 in prostate cancer. *PLoS One* **6**, e28837 (2011).

375. Nizet, V. & Johnson, R.S. Interdependence of hypoxic and innate immune responses. *Nat Rev Immunol* **9**, 609-617 (2009).

376. Cummins, E.P. *et al.* Prolyl hydroxylase-1 negatively regulates IkappaB kinase-beta, giving insight into hypoxia-induced NFkappaB activity. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 18154-18159 (2006).

377. Rius, J. *et al.* NF-kappaB links innate immunity to the hypoxic response through transcriptional regulation of HIF-1alpha. *Nature* **453**, 807-811 (2008).

378. Scribner, K.C., Wellberg, E.A., Metz, R.P. & Porter, W.W. Singleminded-2s (Sim2s) promotes delayed involution of the mouse mammary gland through suppression of Stat3 and NFkappaB. *Mol Endocrinol* **25**, 635-644 (2011).

379. Franks, R.G. & Crews, S.T. Transcriptional activation domains of the single-minded bHLH protein are required for CNS midline cell development. *Mech Dev* **45**, 269-277 (1994).

380. Chrast, R. *et al.* Cloning of two human homologs of the Drosophila single-minded gene SIM1 on chromosome 6q and SIM2 on 21q within the Down syndrome chromosomal region. *Genome Res* **7**, 615-624 (1997).

381. Kwak, H.I. *et al.* Inhibition of breast cancer growth and invasion by single-minded 2s. *Carcinogenesis* **28**, 259-266 (2007).

382. Salceda, S., Beck, I. & Caro, J. Absolute requirement of aryl hydrocarbon receptor nuclear translocator protein for gene activation by hypoxia. *Arch Biochem Biophys* **334**, 389-394 (1996).

383. Huang, L.E., Arany, Z., Livingston, D.M. & Bunn, H.F. Activation of hypoxia-inducible transcription factor depends primarily upon redox-sensitive stabilization of its alpha subunit. *J Biol Chem* **271**, 32253-32259 (1996).

384. Salceda, S. & Caro, J. Hypoxia-inducible factor 1alpha (HIF-1alpha) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes. *J Biol Chem* **272**, 22642-22647 (1997).

385. Groulx, I. & Lee, S. Oxygen-dependent ubiquitination and degradation of hypoxia-inducible factor requires nuclear-cytoplasmic trafficking of the von Hippel-Lindau tumor suppressor protein. *Molecular and cellular biology* **22**, 5319-5336 (2002).

386. Bruick, R.K. & McKnight, S.L. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* **294**, 1337-1340 (2001).

387. Epstein, A.C. *et al.* C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).

388. Ema, M. *et al.* Mild impairment of learning and memory in mice overexpressing the mSim2 gene located on chromosome 16: an animal model of Down's syndrome. *Hum Mol Genet* **8**, 1409-1415 (1999).

389. Chrast, R. *et al.* Mice trisomic for a bacterial artificial chromosome with the single-minded 2 gene (Sim2) show phenotypes similar to some of those present in the partial trisomy 16 mouse models of Down syndrome. *Hum Mol Genet* **9**, 1853-1864 (2000).

390. Epstein, R., Davisson, M., Lehmann, K., Akeson, E.C. & Cohn, M. Position of Igl-1, md, and Bst loci on chromosome 16 of the mouse. *Immunogenetics* **23**, 78-83 (1986).

391. Korenberg, J.R. *et al.* Molecular definition of a region of chromosome 21 that causes features of the Down syndrome phenotype. *Am J Hum Genet* **47**, 236-246 (1990).

392. Krivit, W. & Good, R.A. Simultaneous Occurrence of Mongolism and Leukemia - Report of a Nationwide Survey. *Ama J Dis Child* **94**, 289-293 (1957).

393. Fong, C.T. & Brodeur, G.M. Down's syndrome and leukemia: epidemiology, genetics, cytogenetics and mechanisms of leukemogenesis. *Cancer genetics and cytogenetics* **28**, 55-76 (1987).

394. Iselius, L., Jacobs, P. & Morton, N. Leukaemia and transient leukaemia in Down syndrome. *Human genetics* **85**, 477-485 (1990).

395. Bain, B. Down's Syndrome-Transient Abnormal Myelopoiesis and Acute Leukaemia. *Leuk Lymphoma* **3**, 309-317 (1991).

396. Robison, L.L. *et al.* Down syndrome and acute leukemia in children: a 10-year retrospective survey from Childrens Cancer Study Group. *The Journal of pediatrics* **105**, 235-242 (1984).

397. Hasle, H., Kerndrup, G. & Jacobsen, B.B. Childhood myelodysplastic syndrome in Denmark: incidence and predisposing conditions. *Leukemia* **9**, 1569-1572 (1995).

398. Zipursky, A., Thorner, P., De Harven, E., Christensen, H. & Doyle, J. Myelodysplasia and acute megakaryoblastic leukemia in Down's syndrome. *Leukemia research* **18**, 163-171 (1994).

399. Hasle, H., Clemmensen, I.H. & Mikkelsen, M. Risks of leukaemia and solid tumours in individuals with Down's syndrome. *Lancet* **355**, 165-169 (2000).

400. Chen, R. *et al.* FitSNPs: highly differentially expressed genes are more likely to have variants associated with disease. *Genome Biol* **9**, R170 (2008).

401. Chatterjee, A. *et al.* Potential contribution of SIM2 and ETS2 functional polymorphisms in Down syndrome associated malignancies. *BMC Med Genet* **14**, 12 (2013).

402. Aleman, M.J. *et al.* Inhibition of Single Minded 2 gene expression mediates tumor-selective apoptosis and differentiation in human colon cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 12765-12770 (2005).

403. Arredouani, M.S. *et al.* Identification of the transcription factor single-minded homologue 2 as a potential biomarker and immunotherapy target in prostate cancer. *Clin Cancer Res* **15**, 5794-5802 (2009).

404. DeYoung, M.P., Tress, M. & Narayanan, R. Identification of Down's syndrome critical locus gene SIM2-s as a drug therapy target for solid tumors. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 4760-4765 (2003).

405. DeYoung, M.P., Tress, M. & Narayanan, R. Down's syndrome-associated Single Minded 2 gene as a pancreatic cancer drug therapy target. *Cancer Lett* **200**, 25-31 (2003).

406. Halvorsen, O.J. *et al.* Increased expression of SIM2-s protein is a novel marker of aggressive prostate cancer. *Clin Cancer Res* **13**, 892-897 (2007).

407. Laffin, B. *et al.* Loss of singleminded-2s in the mouse mammary gland induces an epithelial-mesenchymal transition associated with up-regulation of slug and matrix metalloprotease 2. *Molecular and cellular biology* **28**, 1936-1946 (2008).

408. Long, Q. *et al.* Protein-coding and microRNA biomarkers of recurrence of prostate cancer following radical prostatectomy. *Am J Pathol* **179**, 46-54 (2011).

409. Yang, Q., Rasmussen, S.A. & Friedman, J.M. Mortality associated with Down's syndrome in the USA from 1983 to 1997: a population-based study. *Lancet* **359**, 1019-1025 (2002).

410. Satge, D., Sasco, A.J., Vekemans, M.J., Portal, M.L. & Flejou, J.F. Aspects of digestive tract tumors in Down syndrome: a literature review. *Dig Dis Sci* **51**, 2053-2061 (2006).

411. Wellberg, E., Metz, R.P., Parker, C. & Porter, W.W. The bHLH/PAS transcription factor singleminded 2s promotes mammary gland lactogenic differentiation. *Development* **137**, 945-952 (2010).

412. Gustafson, T.L. *et al.* Ha-Ras transformation of MCF10A cells leads to repression of Singleminded-2s through NOTCH and C/EBPbeta. *Oncogene* (2009).

413. Scribner, K.C., Behbod, F. & Porter, W.W. Regulation of DCIS to invasive breast cancer progression by Singleminded-2s (SIM2s). *Oncogene* **32**, 2631-2639 (2013).

414. Wirthner, R., Wrann, S., Balamurugan, K., Wenger, R.H. & Stiehl, D.P. Impaired DNA double-strand break repair contributes to chemoresistance in HIF-1 alpha-deficient mouse embryonic fibroblasts. *Carcinogenesis* **29**, 2306-2316 (2008).

415. Rohwer, N., Zasada, C., Kempa, S. & Cramer, T. The growing complexity of HIF-1alpha's role in tumorigenesis: DNA repair and beyond. *Oncogene* **32**, 3569-3576 (2013).

416. Koshiji, M. *et al.* HIF-1alpha induces cell cycle arrest by functionally counteracting Myc. *EMBO J* **23**, 1949-1956 (2004).

417. Franken, N.A., Rodermond, H.M., Stap, J., Haveman, J. & van Bree, C. Clonogenic assay of cells in vitro. *Nat Protoc* **1**, 2315-2319 (2006).

418. Elsarraj, H.S. *et al.* NEMO, a Transcriptional Target of Estrogen and Progesterone, Is Linked to Tumor Suppressor PML in Breast Cancer. *Cancer Res* **77**, 3802-3813 (2017).

419. Valdez, K.E. *et al.* Human primary ductal carcinoma in situ (DCIS) subtype-specific pathology is preserved in a mouse intraductal (MIND) xenograft model. *J Pathol* **225**, 565-573 (2011).

420. Parra, I. & Windle, B. High resolution visual mapping of stretched DNA by fluorescent hybridization. *Nat Genet* **5**, 17-21 (1993).

421. Gustafson, T.L. *et al.* Ha-Ras transformation of MCF10A cells leads to repression of Singleminded-2s through NOTCH and C/EBPbeta. *Oncogene* **28**, 1561-1568 (2009).

422. Burma, S., Chen, B.P., Murphy, M., Kurimasa, A. & Chen, D.J. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J Biol Chem* **276**, 42462-42467 (2001).

423. Wang, H. *et al.* A Perspective on Chromosomal Double Strand Break Markers in Mammalian Cells. *Jacobs J Radiat Oncol* **1** (2014).

424. Goodarzi, A.A. *et al.* ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin. *Molecular cell* **31**, 167-177 (2008).

425. Morrison, C. *et al.* The controlling role of ATM in homologous recombinational repair of DNA damage. *EMBO J* **19**, 463-471 (2000).

426. Renwick, A. *et al.* ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet* **38**, 873-875 (2006).

427. Epstein, D.J. *et al.* Members of the bHLH-PAS family regulate Shh transcription in forebrain regions of the mouse CNS. *Development* **127**, 4701-4709 (2000).

428. Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S. & Brunak, S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **4**, 1633-1649 (2004).

429. Ryu, T. *et al.* Heterochromatic breaks move to the nuclear periphery to continue recombinational repair. *Nature cell biology* **17**, 1401-1411 (2015).

430. Pierce, A.J., Johnson, R.D., Thompson, L.H. & Jasin, M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev* **13**, 2633-2638 (1999).

431. Lorkovic, Z.J. & Berger, F. Heterochromatin and DNA damage repair: Use different histone variants and relax. *Nucleus* **8**, 583-588 (2017).

432. Tutt, A. *et al.* Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. *Lancet* **376**, 235-244 (2010).

433. Schreiber, V., Dantzer, F., Ame, J.C. & de Murcia, G. Poly(ADP-ribose): novel functions for an old molecule. *Nature reviews. Molecular cell biology* **7**, 517-528 (2006).

434. Pearson, S.J. *et al.* ATM-dependent activation of SIM2s regulates homologous recombination and epithelial-mesenchymal transition. *Oncogene* (2018).

435. Gagou, M.E., Zuazua-Villar, P. & Meuth, M. Enhanced H2AX phosphorylation, DNA replication fork arrest, and cell death in the absence of Chk1. *Mol Biol Cell* **21**, 739-752 (2010).

436. Her, J., Ray, C., Altshuler, J., Zheng, H. & Bunting, S.F. 53BP1 Mediates ATR-Chk1 Signaling and Protects Replication Forks under Conditions of Replication Stress. *Molecular and cellular biology* **38** (2018).

437. Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593-605 (2007).

438. Eilon, T. & Barash, I. Forced activation of Stat5 subjects mammary epithelial cells to DNA damage and preferential induction of the cellular response mechanism during proliferation. *J Cell Physiol* **226**, 616-626 (2011).

439. Gildemeister, O.S., Sage, J.M. & Knight, K.L. Cellular redistribution of Rad51 in response to DNA damage: novel role for Rad51C. *J Biol Chem* **284**, 31945-31952 (2009).

440. Loeb, L.A. A mutator phenotype in cancer. *Cancer Res* **61**, 3230-3239 (2001).

441. Gatz, S.A. & Wiesmuller, L. p53 in recombination and repair. *Cell Death Differ* **13**, 1003-1016 (2006).

442. Golding, S.E. *et al.* Double strand break repair by homologous recombination is regulated by cell cycle-independent signaling via ATM in human glioma cells. *J Biol Chem* **279**, 15402-15410 (2004).

443. Yoshida, K. & Miki, Y. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Sci* **95**, 866-871 (2004).

444. Newman, B., Austin, M.A., Lee, M. & King, M.C. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 3044-3048 (1988).

445. Society, A.C. in American Cancer Society (Atlanta; 2018).

446. Burgess, M. & Puhalla, S. BRCA 1/2-Mutation Related and Sporadic Breast and Ovarian Cancers: More Alike than Different. *Front Oncol* **4**, 19 (2014).

447. Bergamaschi, A. *et al.* Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes Chromosomes Cancer* **45**, 1033-1040 (2006).

448. Fridlyand, J. *et al.* Breast tumor copy number aberration phenotypes and genomic instability. *BMC Cancer* **6**, 96 (2006).

449. Jonsson, G. *et al.* Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. *Cancer Res* **65**, 7612-7621 (2005).

450. Tirkkonen, M. *et al.* Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. *Cancer Res* **57**, 1222-1227 (1997).

451. Kim, T.M. *et al.* RAD51 mutants cause replication defects and chromosomal instability. *Molecular and cellular biology* **32**, 3663-3680 (2012).

452. Zamborszky, J. *et al.* Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions. *Oncogene* (2016).

453. Zhou, C., Smith, J.L. & Liu, J. Role of BRCA1 in cellular resistance to paclitaxel and ionizing radiation in an ovarian cancer cell line carrying a defective BRCA1. *Oncogene* **22**, 2396-2404 (2003).

454. Trenz, K., Schutz, P. & Speit, G. Radiosensitivity of lymphoblastoid cell lines with a heterozygous BRCA1 mutation is not detected by the comet assay and pulsed field gel electrophoresis. *Mutagenesis* **20**, 131-137 (2005).

455. Savage, K.I. *et al.* Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability. *Molecular cell* **54**, 445-459 (2014).

456. Tombline, G. & Fishel, R. Biochemical characterization of the human RAD51 protein. I. ATP hydrolysis. *J Biol Chem* **277**, 14417-14425 (2002).

457. Xu, K. *et al.* XRCC2 promotes colorectal cancer cell growth, regulates cell cycle progression, and apoptosis. *Medicine (Baltimore)* **93**, e294 (2014).

458. Denison, M.S., Soshilov, A.A., He, G., DeGroot, D.E. & Zhao, B. Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicol Sci* **124**, 1-22 (2011).

459. Bardos, J.I. & Ashcroft, M. Negative and positive regulation of HIF-1: a complex network. *Biochim Biophys Acta* **1755**, 107-120 (2005).

460. Swift, M., Reitnauer, P.J., Morrell, D. & Chase, C.L. Breast and other cancers in families with ataxia-telangiectasia. *The New England journal of medicine* **316**, 1289-1294 (1987).

461. Thompson, D. *et al.* Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst* **97**, 813-822 (2005).

462. Tsouroula, K. *et al.* Temporal and Spatial Uncoupling of DNA Double Strand Break Repair Pathways within Mammalian Heterochromatin. *Molecular cell* **63**, 293-305 (2016).

463. Liu, X. *et al.* Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 12111-12116 (2007).

464. Cheung, A.M. *et al.* Brca2 deficiency does not impair mammary epithelium development but promotes mammary adenocarcinoma formation in p53(+/-) mutant mice. *Cancer Res* **64**, 1959-1965 (2004).

465. Kuznetsov, S.G., Haines, D.C., Martin, B.K. & Sharan, S.K. Loss of Rad51c leads to embryonic lethality and modulation of Trp53-dependent tumorigenesis in mice. *Cancer Res* **69**, 863-872 (2009).

466. Shook, D. & Keller, R. Mechanisms, mechanics and function of epithelial-mesenchymal transitions in early development. *Mech Dev* **120**, 1351-1383 (2003).

467. Huber, M.A., Kraut, N. & Beug, H. Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr Opin Cell Biol* **17**, 548-558 (2005).

468. Lee, J.M., Dedhar, S., Kalluri, R. & Thompson, E.W. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *The Journal of cell biology* **172**, 973-981 (2006).

469. Ansieau, S. *et al.* Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer cell* **14**, 79-89 (2008).

470. Shimelis, H. *et al.* Triple-Negative Breast Cancer Risk Genes Identified by Multigene Hereditary Cancer Panel Testing. *J Natl Cancer Inst* **110**, 855-862 (2018).

471. Hashimoto, Y., Ray Chaudhuri, A., Lopes, M. & Costanzo, V. Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis. *Nat Struct Mol Biol* **17**, 1305-1311 (2010).

472. Cowell, C.F. *et al.* Progression from ductal carcinoma in situ to invasive breast cancer: revisited. *Mol Oncol* **7**, 859-869 (2013).

473. Graeser, M. *et al.* A marker of homologous recombination predicts pathologic complete response to neoadjuvant chemotherapy in primary breast cancer. *Clin Cancer Res* **16**, 6159-6168 (2010).

474. Maacke, H. *et al.* Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer. *Int J Cancer* **88**, 907-913 (2000).