A Genome-Wide Study of Homologous Recombination in Mammalian Cells Identifies RBMX, a Novel Component of the DNA Damage Response

A genome-wide study of homologous recombination in mammalian cells identifies RBMX, a novel component of the DNA damage response

Abstract

Repair of DNA double-strand breaks is critical to the maintenance of genomic stability, and failure to repair these DNA lesions can cause loss of chromosome telomeric regions, complex translocations, or cell death. In humans this can lead to severe developmental abnormalities and cancer. A central pathway for double-strand break repair is homologous recombination (HR), a mechanism that operates during the S and G2 phases of the cell cycle and primarily utilizes the replicated sister chromatid as a template for repair. Most knowledge of HR is derived from work carried out in prokaryotic and eukaryotic model organisms. To probe the HR pathway in human cells, we performed a genome-wide siRNA-based screen; and through this screen, we uncovered cellular functions required for HR and identified proteins that localize to sites of DNA damage. Among positive regulators of HR, we identified networks of pre-mRNA-processing factors and canonical DNA damage response effectors. Within the former, we found RBMX, a heterogeneous nuclear ribonucleoprotein (hnRNP) that associates with the spliceosome, binds RNA, and influences alternative splicing. We found that RBMX is required for cellular resistance to genotoxic stress, accumulates at sites of DNA damage in a poly(ADP-ribose) polymerase 1-dependent manner and through multiple domains, and promotes HR by facilitating proper BRCA2 expression. Screen data also revealed that the mammalian recombinase RAD51 is commonly off-targeted by siRNAs, presenting a cautionary note to those studying HR with RNAi and highlighting the vulnerability of RNAi screens to off-target effects in general. Candidate validation through secondary screening with independent reagents successfully circumvented the effects of off-targeting and set a new standard for reagent redundancy in RNAi screens.

TABLE OF CONTENTS

FIGURES

TABLES

SUPPLEMENTARY TABLES

COMMON ABBREVIATIONS

nucleotide	nt
untranslated region	UTR
DNA damage response	DDR
homologous recombination	HR
non-homologous end-joining	NHEJ
canonical non-homologous end-joining	c-NHEJ
alternative non-homologous end-joining	a-NHEJ
microhomology-mediated end-joining	MMEJ
single-strand annealing	SSA
mismatch repair	MMR
base excision repair	BER
nucleotide excision repair	NER
single-strand break repair	SSBR
gene conversion	GC
single Holliday junction	HJ
double Holliday junction	dHJ
double-strand break	DSB
double-strand break repair	DSBR
synthesis-dependent strand annealing	SDSA
break-induced replication	BIR
structure-specific endonuclease	SSE
RNA-binding protein	RBP
RNA recognition motif	RRM
alternative splicing	AS
single-stranded DNA	ssDNA
double-stranded DNA / RNA	dsDNA / dsRNA
messenger RNA	mRNA
RNA interference	RNAi
small interfering RNA	siRNA
microRNA	miRNA
RNA polymerase II	RNAPII
RNAPII carboxy-terminal domain	CTD
Escherichia coli	*E. coli*
Saccharomyces cerevisiae	*S. cerevisiae*

ACKNOWLEDGEMENTS

First, I acknowledge my advisor, Dr. Stephen J. Elledge. My training in the Elledge lab has been an experience of exponential personal and scientific growth, the summation of which can only be described as invaluable. From Steve, I have learned and observed first-hand that the careful and prepared consideration of any problem can and ultimately will yield a logical solution. It is to Steve's guidance and encouragement that I attribute the scientific knowledge, skill, and confidence I have gained these six years. For his efforts, advice and mentorship I am very grateful.

Thank you to Dr. Agata Smogorzewska, who patiently taught me how to use both my brain and my hands as more efficient tools for experimental science. I greatly appreciate the investment she made in my scientific education. I am also grateful to members of the Elledge lab past and present, a collection of individuals who are both uniquely brilliant and kind. A special thank you to the postdoctoral researchers who helped me along the way, including Drs. Cecilia Cotta-Ramusino (my Italian kindred spirit), Alberto Ciccia, Andrew Elia, Mike Emanuele, Nicole Solimini, Andrea Bredemeyer, Claudio Thoma, Kristen Hurov, and Rob McDonald, as well as my fellow Elledge lab trainees, Dr. Michael Schlabach, Dr. Danny Chou, Dr. Anna Mazzucco, Yifan Wang, and Laura Sack. In particular, I thank my long-time bay mates Drs. Natasha Pavlova and Ben Larman for sharing their boundless creativity and enthusiasm, from which I daily drew inspiration.

I thank my collaborators Drs. Randall W. King and Frederic D. Sigoillot, whose computational contributions greatly influenced the work presented herein. I thank my dissertation committee members, Drs. Johannes Walter, Marc Vidal, and David M. Livingston for providing both their creative insights and the yearly opportunity to observe three great minds in action. I appreciate the time and energy of my defense examiners, Drs. Johannes Walter, Angelika Amon, Ralph Scully, and Galit Lahav. I thank them for critically reading this text. A special thank you to Dr. Johannes Walter for serving as the chair of both my dissertation and defense committees and to Dr. Ralph Scully for participating in my preliminary qualifying exam as well as my defense.

PREFACE

I am grateful to all the professors and scientists at MIT who contributed to my early training, especially to Dr. Angelika Amon. It was in her lab and under her mentorship that I first discovered an enthusiasm for experimental science. Her willingness to train young and novice undergraduates has, and will not doubt continue, to make ever-expanding contributions to the field of biology.

Finally, I am grateful to my family for their support throughout all my studies, especially to my mother, Susan Adamson, who first showed me the rewarding nature of thoughtful persistence and to my father, Keith Adamson, from whose vast intelligence I have benefited in unquantifiable ways. I thank my brother, Ben Adamson, for setting the bar so high, and my sisters, Kate and Christine Adamson, for their wholehearted and irreplaceable friendships. In life as in HR, sisters are the best possible source of repair. Finally, I thank my husband, Aaron Malnarick, for his enduring and unconditional support. I dedicate this work to Eleanor Nelson, whose curiosity was unbounded by generation.

"Dans les champs de l'observation
le hasard ne favorise que les esprits prepares."

-Louis Pasteur

Chapter One

Introduction

I. DNA damage

1-1. The DNA molecule

To understand DNA damage and repair, one must begin by understanding the structure of DNA itself. Deoxyribonucleic acid (DNA) is a macromolecule that encodes the genetic instructions necessary for coordinating cellular life in all organisms, prokaryote to eukaryote. Fundamentally, DNA is composed of five organic elements –hydrogen, oxygen, nitrogen, carbon, and phosphorous– which are organized into highly ordered substructures called nucleotides (Figure 1). Single DNA molecules are polymers of repetitive nucleotides each containing a five-carbon sugar residue (deoxyribose), an aromatic nitrogenous base called a nucleobase (or base), and a phosphate group. Alternating phosphate and sugar residues form the backbone of the DNA polymer, with a phosphodiester bond spanning the fifth and third carbons of deoxyribose moieties in adjacent nucleotides. The deoxyribose is covalently linked to one of four bases, adenine, thymine, guanine or cytosine, by a β-N-glycosidic bond from its first carbon. Thymine (T) and cytosine (C) are aromatic ring compounds known as pyrimidines that have nitrogen at ring positions 1 and 3. Adenine (A) and guanine (G) are larger bases called purines that are composed of a pyrimidine fused to an imidazole ring.

In vivo cellular DNA rarely exists as a single molecule but forms a paired helical structure in which two DNA polymers (or strands) are linked by matched nucleobases in an inverted, anti-parallel alignment (Figure 1). Hydrogen bonds and base-stacking interactions between base pairs mediate linkage of the stands to form an ordered structure known as the double-helix. This structure was first describe by James D. Watson and Francis H. C. Crick in a 1953 Nature publication as, “two helical chains each coiled round the same axis. . . .related by a dyad [of bases] perpendicular to the fibre axis . . . .[These] follow right handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions” (Watson and Crick, 1953). In properly aligned molecules, pyrimidine bases are paired

exclusively with purines (T to A and G to C), so that "if the sequence of bases on one chain is given, then the sequence of the other chain is automatically determined" (Watson and Crick, 1953).

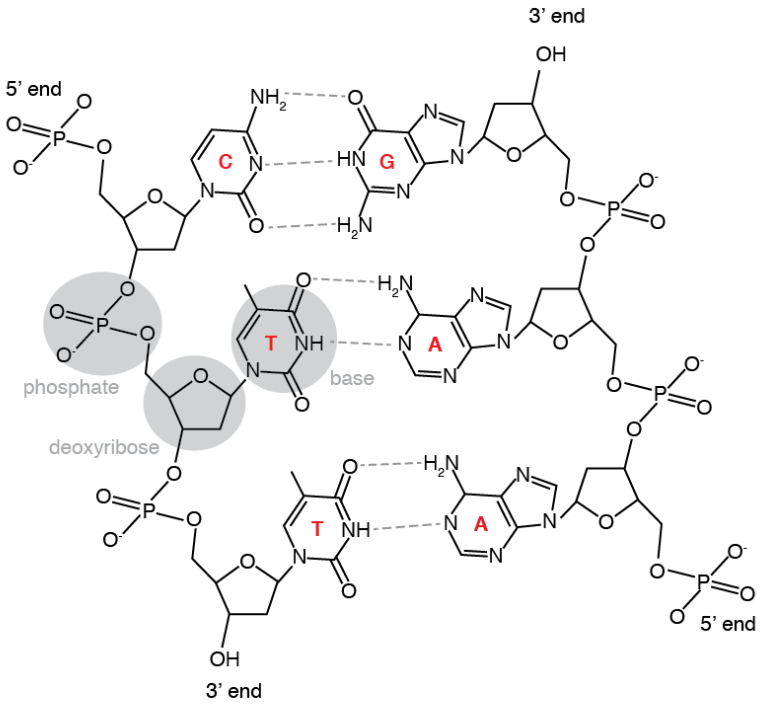


Figure 1. Schematic of anti-parallel pairing between two DNA strands. Each DNA strand contains three nucleotides. C, G, T and A indicate cytosine, guanine, thymine and adenine, respectively. The phosphate group, deoxyribose, and nucleobase (base) of one nucleotide moiety are indicated. 5' phosphate and 3' hydroxyl groups are also indicated. DNA structure is discussed in greater detail in Chapter 1, 1-1.

1-2. DNA lesions and repair pathways

As with any ordered structure, DNA is subject to error, modification and deterioration. The dynamic chemical nature of the cellular environment and frequent *in vivo* processing provide an endogenous baseline of DNA damage that is compounded by exposure to sources of environmental stress. The maintenance of DNA fidelity, however, is a task essential to both cellular function and survival because as a "blueprint" molecule DNA sits at the foundation of a dynamic central dogma that turns nucleic acid-encoded information into functional cellular processes (Crick, 1970). Even one mutated nucleobase (out of ~4.6×10^6 for the bacteria *Escherichia coli*, ~1.2×10^7 for the budding yeast *Saccharomyces cerevisiae*, and ~3.3×10^9 for humans) placed in the right genomic position can have functional consequences that are deleterious even fatal to an organism, and the accumulation of many

mutations, especially in the gamete cells of multicellular organisms, amplifies the probability of such problematic events.

To maintain cellular function throughout the growth of a cell population or development of an organism, DNA must also be faithfully replicated and distributed as cells divide; however, errors in DNA replication are not uncommon. In *Escherichia coli* (*E. coli*), DNA polymerases have been estimated to incorporate the wrong nucleotide into newly replicated DNA –causing a mismatched nucleobase dyad– at a rate of approximately 10^{-7} - 10^{-8} (Kunkel, 2004), which is an estimated 1 mutation per replication event. Left unrepaired such mismatches will result in full the mutation of nucleotide pairs during the next round of replication (mutation fixation), after which they cannot be distinguished from properly preserved nucleotides in the surrounding DNA sequence. While a majority of these errors are fixed by the proofreading activities of DNA polymerases, mutations that escape proofreading must be repaired by other means.

Happily for all cellular life, comprehensive mechanisms of DNA repair that fix a variety of DNA lesions, including nucleobase mismatches, have evolved and been conserved. In general, DNA lesions are diverse, but they can be classified in two ways, by their structural manifestation and by the mechanisms that engage them for repair (Figure 2). Structurally, DNA lesions are divided into two major categories: those that involve only one strand of the DNA double-helix and those that affect both, and both single- and double-strand DNA lesions can be further subcategorized based on their chemical makeup. Among single-strand DNA (ssDNA) lesions are base-base mismatches from DNA replication errors, apurinic / apyrimidinic (AP) sites formed by base hydrolysis, structural base alterations caused by chemical modification, and regions of missing nucleotides known as single-strand breaks. Chemical modifications to DNA bases include small changes generated by deamination, alkylation (commonly methylation) or oxidation, larger modifications that distort helical structure (such as addition of bulky adducts, including aromatic amines and hydrocarbons), and covalent intrastrand linkages between nucleobases (such as pyrimidine dimers). Base-base mismatches from replication errors are repaired by a mechanism aptly named mismatch repair (MMR), AP sites and small base modifications caused by

endogenous reactive oxygen species, X-rays, or chemical agents are repaired through base excision repair (BER), and bulky adducts and intrastrand crosslinks, which arise from exposure to UV-light and environmental chemicals (often found in soot, tar and cigarette smoke), rely on nucleotide excision repair (NER) for repair (Figure 2). The mechanism for repair of single-strand breaks (SSBs) is called single-strand break repair (SSBR). This mechanism is often engaged downstream of BER and NER once damaged bases have been removed.

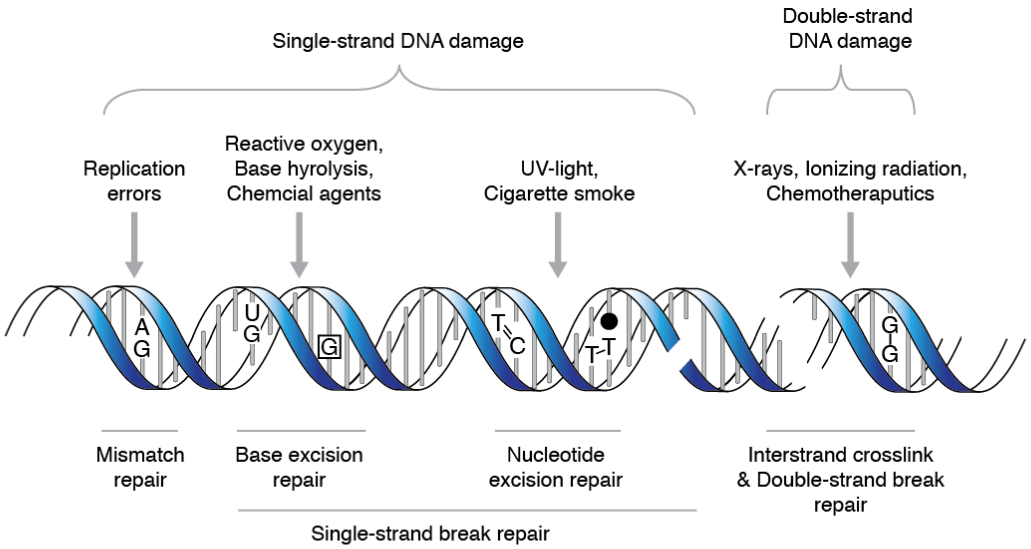


Figure 2. DNA lesions and repair pathways. Specific DNA lesions are categorized by possible exogenous and endogenous sources (above) and corresponding repair pathways (below). Lesion examples (from left to right): mismatched bases, an apurinic / apyrimidinic (AP) site, cytosine deamination forming uracil, methylation of guanine forming 7-methylguanine, 2 pyrimidine dimers, covalent attachment of a bulky adduct, a single-strand break, a double-strand break (DSB), an interstrand crosslink. DNA lesions and repair pathways are discussed in greater detail in Chapter 1, 1-2.

Although the particulars of MMR, BER, NER and SSBR vary, these mechanisms all share one important feature: genomic fidelity is maintained by engaging the complementary and *undamaged* DNA strand as a template for repair. In these mechanisms nucleotides are removed from the damaged strand and DNA polymerases facilitate DNA re-synthesis guided by the readily available compliment strand. In contrast, repair of double-strand DNA (dsDNA) lesions requires the repair of missing or damaged bases without an obvious template. Double-strand DNA lesions include double-strand breaks (DSBs) and covelant interstrand crosslinks (ICLs) caused by abortive topoisomerase activity, collapsed replication forks, exposure to radiation (X- and γ–rays), and various chemical agents, such as nitrogen mustard and

chemotherapeutics (chlorambucil, mitomycin C, cisplatin and bleomycin). DSBs are widely considered to be the most cytotoxic form of DNA lesion (Jackson, 2002; Khanna and Jackson, 2001), and cleavage of both strands present a unique challenge for repair. To repair this form of DNA damage cells have evolved at least three types of mechanism: non-homologous end-joining (NHEJ), homologous recombination (HR), and single-strand annealing (SSA). The two primary and most well studied pathways are NHEJ and HR. NHEJ is a direct and "error-prone" religation of DSB ends, and HR is a cell cycle-regulated mechanism that uses a newly replicated DNA strand (sister chromatid) as a template for "error-free" repair by DNA synthesis. The details of NHEJ, HR, SSA, and their sub-pathways are described in greater detail below (Chapter 1, III).

1-3. Medical rationale for the study of DNA repair

Although the various mechanisms of DNA repair have been defined in copious detail (reviewed herein), there remains much to be learned and potential medical applications inspire continued study. Loss-of-function and hypomorphic mutations in many repair genes have been causally associated with human pathologies. Among these are neurodegenerative syndromes and defects in neurodevelopment thought to arise from a strong dependency of non-proliferative neuronal cells on DNA repair, immunodeficiency disorders caused by defects in antigen receptor molecule generation by NHEJ mechanisms, and premature ageing disorders (Ciccia and Elledge, 2010).

Perhaps the most common pathology associated with DNA repair dysfunction, however, is cancer. In particular, germline mutations in MMR genes (*MSH2*, *MSH6*, and *MLH1*) carry risk associated with hereditary nonpolyposis colorectal cancer (HNPCC). NER gene mutation (*XPA-G*) causes xeroderma pigmentosum (XP), a disorder associated with an increased risk of skin cancer and melanoma (Ciccia and Elledge, 2010). Mutations in double strand break repair genes (*ATM, NBS1, LIG4,* and *ARTEMIS*) predispose to lymphomas and leukemias, and in the HR pathway specifically, *BRCA1* and *BRCA2* mutations have been found to be strongly predictive of hereditary breast and ovarian cancer (Ciccia and Elledge, 2010). Germline mutations in additional HR-associated genes, including *BACH1,*

PALB2, *RAD51C, CHK2, ATM, NBS1*, and *RAD50,* have also been correlated with breast cancer (Ciccia and Elledge, 2010). The normal cellular functions of each of these genes are discussed in detail throughout Chapter 1. Interestingly, because somatic mutations in some DNA repair-associated genes have been found in tumor samples at frequencies suggestive of cancer-driving events, for example *ATM* mutations in primary lung adenocarcinoma, the development of sporadic cancers, although not predisposed by any known germline mutations, is also thought to be linked to dysregulation of DNA repair at some level (Ciccia and Elledge, 2010; Ding et al., 2008).

In this manner, the study of DNA repair can be thought of both as an academic exercise aimed at achieving a better understanding of our own cellular nature and an endeavor of current medical relevancy. To this point, much consideration has recently been given to the potential for developing cancer treatments that target specific DNA repair deficiencies with a synthetic lethality-based approach. The defining logic here being that cancer cells deficient in one repair pathway will have an increased dependency on alternative repair mechanisms, more so than normal cells, and by targeting these mechanisms a cytotoxic differential between normal and malignant cells may be reached (Luo et al., 2009).

II. The vertebrate DNA damage response (DDR)

2-1. Signaling and checkpoint activation

The first step towards efficiently managing DNA damage, regardless of the physical lesion sustained, is recognition of the lesion site. To this end, a broad cellular response has evolved that coordinates DNA lesion identification and facilitates subsequent repair; in aggregate this system is referred to as the DNA damage response (DDR) (Ciccia and Elledge, 2010; Zhou and Elledge, 2000). Herein, discussion is focused primarily on mechanisms of DNA repair as the primary objective of the DDR; however, it should be understood that the DDR is a diverse mechanism. Cell fate in response to DNA damage is variable and depends on cell type and condition, as well as the variety and number of

DNA lesions sustained. Conditions of high genotoxic stress or unrepairable DNA lesions can lead to outcomes of regulated cell death through apoptosis or an irreversible condition of cell quiescence called senescence. In multicellular organisms these mechanisms serve to remove damaged cells, which may have acquired deleterious mutations, from contributing to future cell lineages. Apoptosis and senescence are achieved in part through DDR-regulated transcriptional control; and recent evidence suggests that the DDR also regulates splicing and metabolic function after damage (Ciccia and Elledge, 2010).

Protein mediators of the DDR can be broken down into three subcategories: lesion sensors, signal transducers and outcome effectors (Zhou and Elledge, 2000). Sensors, the "first responders" of the DDR, recognize and bind distinct DNA structures at the sites of DNA damage in order to recruit and activate transducer enzymes; transducers then initiate signaling cascades that serve to functionally coordinate effector proteins, many of which are directly responsible for lesion repair. Distinct sensor proteins have evolved with varied affinities for different DNA lesions and repair intermediates. The MRN complex (MRE11-RAD50-NBS1) and the Ku70-Ku80 heterodimer, for example, bind DNA ends present at DSBs (Blier et al., 1993; de Jager et al., 2001; Mimori and Hardin, 1986), while the heterotrimeric complex RPA recognizes and binds single-strand DNA formed at various lesion types and stalled / broken replication forks (Wold, 1997). The poly(ADP-ribose) polymerase PARP1 senses both single- and double-strand breaks (D'Amours et al., 1999), while the MMR heterodimers MSH2/MSH6 and MSH2/MSH3 are specific for identification of mismatched and improperly inserted or deleted nucleotides (Germann et al., 2010).

Once activated, DDR transducer enzymes go on to modify downstream effector proteins through post-translational modifications, including phosphorylation, ubiquitylation, sumoylation, PARylation, and possibly others, to disseminate the message that DNA damage has occurred and to functionally initiate response mechanisms (Branzei et al., 2008; D'Amours et al., 1999; Dianov et al., 2011; Galanty et al., 2009; Morris et al., 2009). Central transducer enzymes in vertebrates are the phosphatidylinositol-3-kinase (PI3K) family proteins ATM (ataxia telangiectasia mutated), ATR (ATM and Rad3 related), and DNA-PK_{cs} (DNA-dependent protein kinase, catalytic subunit), which are activated in response to DNA

damage through the MRN complex, RPA, and the Ku70-Ku80 heterodimer, respectively (Lieber, 2010; Smith et al., 2010).

ATM activation occurs in response to the formation of DSBs, which are DNA substrates particular to the MRN sensor complex (Figure 3b) (Lamarche et al., 2010). MRN is composed of MRE11, RAD50 and NBS1 (Falck et al., 2005). MRE11 has two C-terminal DNA-binding domains and an affinity for DNA ends that is stimulated by RAD50 and NBS1 (de Jager et al., 2001; Paull and Gellert, 1999). In response to DSB formation, nuclear MRN binds DNA ends and through a direct interaction with the C-terminal region of NBS1 facilitates ATM recruitment and subsequent ATM activation (Falck et al., 2005). In undamaged cells, ATM is maintained as an inactive homodimer, a structure that conformationally inhibits the kinase domain by holding it in an inhibitory state near serine residue 1981. However, once at sites of DNA damage, autophosphorylation of S1981 facilitates dimer disruption and consequent kinase activation (Bakkenist and Kastan, 2003). ATM autophosphorylation primarily requires binding to MRN; however, recent work has shown that the sensor protein PARP1 also partially controls ATM activation and mediates the earliest recruitment of MRE11 and NBS1 to DNA damage, which occurs within the first minute after damage (Haince et al., 2007; Haince et al., 2008).

Like ATM, ATR is activated through a mechanism of recruitment to sites of DNA damage; ATR however, is recruited to ssDNA structures coated with RPA, a single-strand binding heterotrimer (Figure 3e). *In vivo,* ATR forms an obligate heterodimer with the ATR-interacting protein ATRIP, and consistent with this, the stabilities of ATR and ATRIP are interdependent and phenotypes caused by individual ATR or ATRIP loss are observed to be strikingly similar (Cimprich and Cortez, 2008). ATRIP binds the large subunit of RPA (RPA1) to facilitate ATR localization to ssDNA structures (Ball et al., 2007; Cortez et al., 2001; Zou and Elledge, 2003), and colocalization of ATR-ATRIP with the independently recruited TopBP1 promotes activation of ATR kinase activity through stimulatory binding between ATR-ATRIP and TopBP1 (Kumagai et al., 2006; Mordes et al., 2008). TopBP1 accumulates on RPA-coated ssDNA through the coordinated actions of the RAD17-RFC2/4 damage-specific clamp loader and the 9-1-1 complex (RAD9-RAD1-HUS1). First, RAD17-RFC2/4 loads on to RPA-coated ssDNA structures,

preferentially those with 5'-recessed ssDNA / dsDNA junctions (Ellison and Stillman, 2003), and facilitates docking of the 9-1-1 complex (Bermudez et al., 2003). 9-1-1 then recruits TopBP1, at least partially through a constitutively phosphorylated serine residue at position 387 that binds a BRCT phospho-peptide binding domain on TopBP1 (St Onge et al., 2003). Work in *Xenopus laevis*, indicates that TopBP1 may also be recruited through a RAD9-binding independent mechanism (Lee and Dunphy, 2010), which could be mediated by the recently discovered 9-1-1- and TopBP1-interacting protein, RHINO (Cotta-Ramusino et al., 2011).

Once activated, ATM, ATR and DNA-PK_{cs} phosphorylate hundreds of proteins, including key repair effectors, primarily at S/T-Q amino acid consensus sites (Matsuoka et al., 2007). Importantly, ATM and ATR also activate the Chk2 and Chk1 transducer kinases through stimulatory phosphorylation on the following S/T residues: T68 (Chk2), S317 (Chk1), and S345 (Chk1) (Figure 3b,e) (Ahn et al., 2000; Guo et al., 2000; Matsuoka et al., 1998). Chk2 and Chk1 are serine-threonine kinases, so-called checkpoint kinases, that amplify DNA damage signaling initiated by ATR / ATM by phosphorylating an overlapping set of effector substrates (Smith et al., 2010; Zhou and Elledge, 2000). The transcriptional regulator and apoptotic mediator p53, for example, is a direct target of Chk2 (Chehab et al., 2000). As their names suggest, Chk2 and Chk1 also have well-established roles in initiating damage-specific G1, S, and G2 cell cycle arrest checkpoints (Boye et al., 2009; Jones and Petermann, 2012). After lesion recognition, temporary cell cycle arrests can be initiated as part of the DDR to allow time for DNA repair and prevent mutation fixation. After completion of repair, arrested cells can re-enter the cell cycle and continue proliferation with unaltered genetic information. To accomplish this, Chk1 and Chk2 inactivate members of the CDC25 dual-specificity phosphatase family (CDC25A/B/C) which normally promote cell cycle progression (Donzelli and Draetta, 2003). Specifically, Chk-mediated phosphorylation of CDC25A (serine-213) initiates phosphatase ubiquitylation and $SCF^{\beta\text{-}TRCP}$-dependent degradation (Jin et al., 2003; Sorensen et al., 2003), and phosphorylation of CDC25C (serine-216) by Chk1 / Chk2 causes 14-3-3-mediated inhibition possibly through protein relocalization (Matsuoka et al., 1998; Peng et al., 1997; Sanchez et al., 1997).

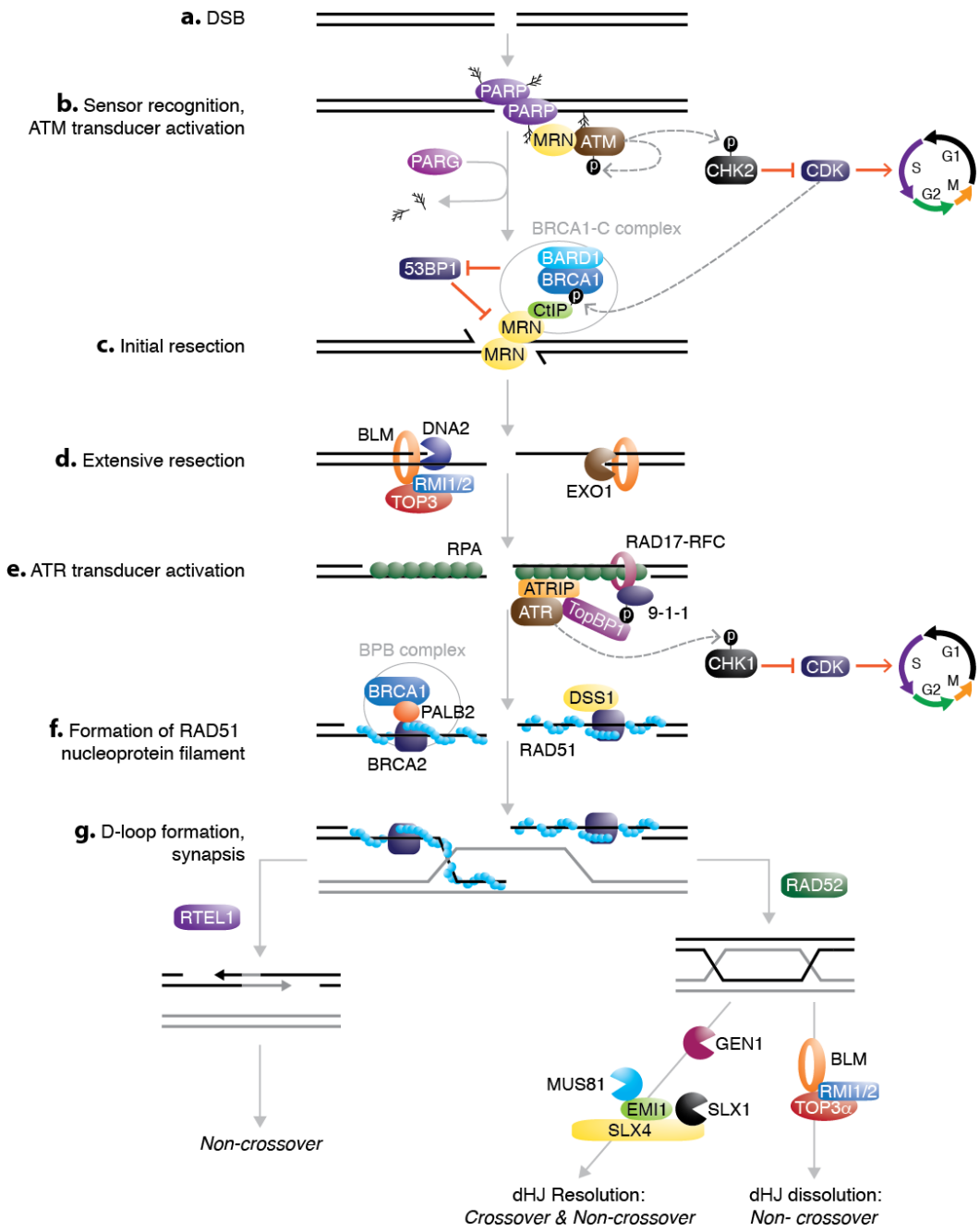


Figure 3. Schematic of double-strand break repair by homologous recombination. (a) DSB formation. **(b)** PARP facilitates early MRN recruitment and ATM activation. ATM phosphorylates CHK2 to initiate checkpoint arrest. **(c)** MRN and CtIP facilitate initial DSB end resection. Phosphorylated CtIP recruits the BRCA1-C complex. **(d)** EXO1 and DNA2 with BLM-TOP3-RMI1-RMI2 mediate extensive DSB end resection. **(e)** RPA binding to ssDNA activates ATR through RAD17-RFC, RAD9-RAD1-HUS1, TopBP1, and ATRIP. ATR phosphorylates CHK1 to initiate checkpoint arrest. **(f)** BRCA2 loads the RAD51 recombinase onto single-strand DSB ends forming the presynaptic nucleoprotein filament. **(g)** Invasion of the presynaptic filament and pairing with a region of complimentary homologous dsDNA constitutes synapsis. Postsynaptic repair can then proceed through either synthesis-dependent strand annealing (meditated by RTEL1) or second-end capture (mediated by RAD52) followed by double-strand break repair. These mechanisms are discussed in greater detail throughout Chapter 1.

During a normal cell cycle, CDC25A facilitates the transition from G1 to S phase, as well as progression through S phase, by activating the cell cycle kinase CDK2, which controls replication origin firing by promoting CDC45 loading into pre-replication complexes (Donzelli and Draetta, 2003; Jones and Petermann, 2012). CDC45 is a critical component of the helicase complex (also containing MCM2–7) that unwinds DNA during replication. After checkpoint activation, reduced CDC25A levels cause CDK2 inactivation and diminished replication origin initiation, thereby restricting progression from G1 into S phase (the G1 / S checkpoint). Similarly, inactivation of CDK2 during S phase slows DNA replication by restricting late origin firing (the intra-S checkpoint). A second, distinct pathway for promoting the intra-S checkpoint has also been described; this pathway is dependent on ATM-mediated phosphorylation of NBS1 and SMC1 (Yazdi et al., 2002). The G2 / M checkpoint, on the other hand, stalls cell cycle progression after completion of DNA replication in G2. Normally, CDC25A/B/C activate the late cell cycle kinase CDK1 at the G2 / M boundary, but in a similar fashion, Chk-mediated inactivation of CDC25C restrains CDK1 and prevents the transition from G2 to M (Donzelli and Draetta, 2003).

2-2. Visualization of the DNA damage response by immunofluorescence techniques

Protein accumulation at genomic regions of DNA damage, as described above for MRN, ATM, ATR-ATRIP, TopBP1, 9-1-1 and Rad17-RFC2/4 (Chapter 1, 2-1), is a common mechanism for engaging the activities of DDR proteins specifically within the chromatin at and surrounding DNA lesions. The finding that many proteins recruited to sites of DNA damage can be visualized as discrete nuclear foci through immunofluorescence techniques (Chen et al., 1996; Scully et al., 1997a; Scully et al., 1996) has greatly enhanced our ability to investigate DDR pathways. In some cases, these foci are evident in undamaged cells, especially in S phase when repair proteins accumulate at regions of replication stress; importantly however, focus formation is induced upon exposure to various genotoxic agents, including

ionizing radiation (IR), ultraviolet light (UV) and chemical crosslinking agents such as mitomycin C (MMC).

Immediately following DNA damage, ATM phosphorylates the histone variant H2AX on a C-terminal serine residue (S139) in chromatin specifically adjacent to DNA lesions (Rogakou et al., 1999). This occurs to the extent that nucleosomes with phosphorylated H2AX (γH2AX) can be found at lengths of millions of basepairs away, allowing γH2AX foci to be easily observed by immunofluorescence (Rogakou et al., 1999). A single damage focus marked by γH2AX colocalizes with many other repair proteins and yet is thought to represent the site of only one DNA lesion. Consistent with this, studies have demonstrated that DNA damage exposure directly correlates with the induction of foci (Asaithamby and Chen, 2009; Sedelnikova et al., 2002). Titrated exposure to IR (between 5 mGy and 1 Gy), for example, generates a linear correlation between dosage and the number of 53BP1 foci observed, with exposure to 1 Gy estimated to cause 20 IR-induced foci (so called, IRIF) marking 20 double-stand breaks (Asaithamby and Chen, 2009). Additionally, the induction of a single double-strand break by endonucleolytic cleavage has proved sufficient to recruit major DDR proteins to an observable focus (Soutoglou et al., 2007). Interestingly, the recruitment of specific DDR proteins (including NBS1, MRE11 and ATM) to chromatin in the absence of DNA damage has been observed to be sufficient for DDR activation, indicating the importance of localization as a DDR regulatory mechanism (Soutoglou and Misteli, 2008).

The adaptation of UV-A lasers to generate large tracks of DNA lesions within defined nuclear regions has advanced the immunofluorescence technique of observing of protein recruitment to DNA damage (Lukas et al., 2005) and, in doing so, has both simplified the identification of novel proteins recruited sites of DNA damage and made possible studies of the spatiotemporal regulation of protein recruitment (Bekker-Jensen et al., 2006; Lukas et al., 2004). In general, exposure to UV-A light (λ = 337-390) causes both single- and double-stand DNA breaks (Peak and Peak, 1990), and pretreatment of cells with halogenated thymidine analogues (bromodeoxyuridine, BrdU or iododeoxyuridine, IdU) and

Hoechst 33258 will further hypersensitizes cells to this source of DNA damage (Limoli and Ward, 1993).

III. Mechanisms of DNA repair

3-1. Single-strand break repair (SSBR) and the role of PARP

Poly(ADP-ribosyl)ation or PARylation is a posttranslational modification achieved through construction of branched poly(ADP-ribose) (PAR) structures on target proteins which is catalyzed by the successive transfer of ADP moieties from hydrolyzed NAD^+ coenzymes to glutamic acid residues. In humans, and largely conserved throughout vertebrates, there are at least 17 poly(ADP-ribose) polymerase (PARP) family member proteins defined by a paralogous catalytic region predominately located at their C-termini (Ame et al., 2004; Schreiber et al., 2006). The PARP1 and PARP2 members of this group are damage sensor proteins that bind single-strand breaks and initiate specialized DDR signaling through auto-PARylation and localized trans-PARylation of nuclear protein substrates (Schreiber et al., 2006). Both PARP1 and PARP2 also contain defined DNA-binding domains (Schreiber et al., 2006). The DNA-binding domain of PARP1 is composed of two zinc fingers that selectively bind DNA breaks; this well-characterized (so called "nick-sensing") motif is located on the protein's N-terminus (Gradwohl et al., 1990). Within seconds after DNA damage, PARP1 and PARP2 are recruited to DNA lesions (Aguilar-Quesada et al., 2007; Mortusewicz et al., 2007) and PARP catalytic activity increases 10-500 fold (Ame et al., 1999; D'Amours et al., 1999). Interestingly, PAR formation at sites of DNA damage is transient, and disassembly of PAR structures is mediated by the mammalian-conserved poly(ADP-ribose) glycohydrolase (PARG), which hydrolyzes branched and linear PAR chains into free ADP-ribose molecules (Miwa and Sugimura, 1971). An emerging role for PARP3 in the repair of double-strand breaks is discussed below (Chapter 1, 3-2-1).

While highly efficient PARP automodification comprises a majority of damage-induced PARylation activity (Ame et al., 1999; Ogata et al., 1981), other substrates are known. In particular, PARylation of histone H1 and H2B is thought to relax chromatin structure in response to DNA damage

and increase DNA lesion accessibility to proteins that mediate repair (Poirier et al., 1982). Functionally, PAR modification after DNA damage also facilitates PAR-docked protein-protein interactions that recruit proteins to sites of DNA damage to contribute to repair. Three PAR-binding domains have been defined. These are the macrodomain (Karras et al., 2005), the zinc-finger PBZ domain (Ahel et al., 2008), and the basic amino acid rich consensus sequence (Gagne et al., 2008; Pleschke et al., 2000). PAR-binding domains are found in many DDR-associated proteins, including ATM, DNA-PK, MRE11, MSH6, Ku70 and LIG3, of which all contain the amino acid motif (Gagne et al., 2008).

Downstream effectors of SSBR are recruited to single-strand DNA breaks in a PARP-dependent manner. Specifically, the SSBR scaffold protein XRCC1 associates with PAR structures near SSBs through a central PAR-binding amino acid motif (Masson et al., 1998; Okano et al., 2003) and through coordinated interactions recruits PNKP, APTX, Pol β, LIG3, and TDP1 (Caldecott, 2008; Hirano et al., 2007; Lan et al., 2004; Loizou et al., 2004; Mortusewicz et al., 2006). Once at breaks, PNKP, APTX, Pol β and TDP1 process single-strand DNA ends in preparation for gap filling by DNA synthesis, which is mediated by Pol β (and possibly other polymerases). Single-strand DNA nicks left after DNA synthesis are then ligated by LIG3 to complete repair. PNKP is a polynucleotide kinase 3'-phosphatase that both phosphorylates 5’ DNA ends and dephosphorylates 3’ ends, as is necessary for DNA synthesis and ligation; APTX is a diadenosine polyphosphate hydrolase that removes AMP moieties formed by abortive ligase activity at 5’ ends, and TDP1 is tyrosyl-DNA phosphodiesterase 1 that processes 3’ ends covalently linked to the enzyme topoisomerase I (Caldecott, 2008). Interestingly, Pol β is both a DNA polymerase and 5’ deoxyribose phosphate (dRP) lyase that removes dRP groups from 5’ DNA ends (Sobol et al., 2000). dRPs are structurally incomplete nucleotides lacking base attachment formed at SSB ends when abasic intermediates of base excision repair (BER) are cleaved. Two mechanisms of SSBR, “short-patch” and “long-patch,” have been defined. Short-patch SSBR requires gap filling of only one missing nucleotide, while “long-patch” repair involves a longer tract of DNA synthesis that causes displacement

of nucleotides 5' to the break site; these are subsequently removed by the endonuclease FEN1. Long-patch repair may also require distinct polymerase activity and LIG1 (Caldecott, 2008)

3-2. Double-strand break repair: Non-homologous end-joining (NHEJ)

Non-homologous end-joining is a method of double-strand break repair that proceeds by the conceptually straightforward process of dsDNA end re-ligation. This pathway is often thought of as the main alternative to homologous recombination, which is discussed in greater detail below (Chapter 1, 3-3). Unlike homologous recombination, however, NHEJ is a template-free mechanism of DSB repair and, as such, remains active throughout the cell cycle. There are two known pathways of NHEJ, canonical and alternative.

3-2-1. Canonical non-homologous end-joining (c-NHEJ)

Like single-strand break repair, and in fact similar to all mechanisms of DNA repair, NHEJ can be logically organized into four functional steps associated with distinct enzymatic activities: (1) break recognition and signaling, (2) nucleolytic end processing, (3) DNA polymerase activity, and (4) nucleotide ligation. In canonical NHEJ (c-NHEJ), break recognition is mediated by the Ku70-Ku80 sensor complex, which binds DSB ends with high affinity in a sequence-independent manner (dsDNA binding constant = 2.4×10^{9} M^{-1}) (Blier et al., 1993; Mimori and Hardin, 1986). Once at DSBs, Ku activates the independently recruited DNA-PK_{cs} through a stimulatory interaction (Yan et al., 2007b); after which, DNA-PK_{cs} goes on to phosphorylate several c-NHEJ-associated substrates, including Ku70, Ku80, XRCC4, XLF, ATREMIS and LIG4 (Hartlerode and Scully, 2009). ARTEMIS is a single-strand 5' to 3' exonuclease that acquires endonuclease activity with a preference for cleaving DNA overhang structures after DNA-PK_{cs}-mediated phosphorylation (Ma et al., 2002). This nucleolytic activity, as part of an ARTEMIS-DNA-PK_{cs} complex, mediates a limited amount of DSB end processing to regenerate DSB ends that frequently carry structural abnormalities in preparation for repair by re-ligation. This processing is regulated by both ARTEMIS phosphorylation and extensive DNA-PK_{cs}

autophosphorylation (Meek et al., 2008).

Although DSB end cleavage is an expected part of NHEJ, it is perhaps surprising that during c-NHEJ DSB ends are iteratively processed by both nuclease and polymerase activities (Lieber, 2010). In particular, members of the Pol X family of DNA polymerases, including Pol μ and λ , are recruited to DSBs by N-terminal BRCT domains that bind Ku-DNA complexes to participate in gap-filling activities (Ma et al., 2004). Pol μ has been shown to add nucleotides to DNA ends in a template-independent and template-discontinuous manner, suggesting that nucleotides are added as well as removed from DSB ends prior to ligation (Gu et al., 2007). An XLF-XRCC4-LIG4 complex efficiently ligates both blunt and overhanging dsDNA ends with or without end homology as the last step in c-NHEJ (Grawunder et al., 1997; Gu et al., 2007). This flexibility for ligation substrates, in addition to the combinatorial nuclease and template-independent polymerase activities at DSBs, makes NHEJ particularly prone to mutation at the ligation site, a fact that demands the necessity of additional and more error-free methods of DSB repair.

Recently, the poly(ADP-ribose) polymerase PARP3 has also been implicated in promoting double-strand break repair by canonical non-homologous end-joining mechanisms. In particular, a model has been suggested wherein PARP3 supports c-NHEJ through facilitating the recruitment of APLF and XRCC4-LIG4 to double-strand breaks (Rulten et al., 2011). The PARP3 catalytic domain is similar to domains in both PARP1 and PARP2, with 39% amino acid identity and 61% similarity to PARP1 (Augustin et al., 2003); however PARP3 has been reported to function primarily as a mono-ADP-ribosylase, with a limited capacity for PARylation (Rulten et al., 2011). Indicative of a role in DNA repair, PARP3 co-immunoprecipitates with several SSBR and NHEJ factors, including DNA-PK_{cs}, PARP1, LIG3, LIG4, Ku70, and Ku80, as well as and polycomb group proteins (Rouleau et al., 2007), which have also been implicated in DNA repair (Chou et al., 2010). Additionally, PARP3 accumulates at regions of DNA damaged by microirradiation, and PARP3 depletion has been shown to cause defective and delayed repair of DSBs (Boehler et al., 2011; Rulten et al., 2011).

3-2-2. Alternative non-homologous end-joining (a-NHEJ) and microhomology-mediated end-joining (MMEJ)

In the absence of proteins that facilitate canonical NHEJ in mammalian cells (including DNA-PK_{cs}, LIG4, XRCC4 and Ku80) DNA end-joining still occurs, albeit predominantly with slower kinetics (DiBiase et al., 2000; Wang et al., 2006). This indicates the existence of one or more alternative NHEJ pathways (a-NHEJ) that serve as back-up mechanisms to canonical NHEJ. Interestingly, three studies that evaluated alternative NHEJ observed that in XRCC4-deficient murine B cells and LIG4-deficient humans DSB repair during class switch recombination proceeds in a manner that increases both the frequency and length of homology evident at repair junctions when compared to controls (Pan-Hammarstrom et al., 2005; Soulas-Sprauel et al., 2007; Yan et al., 2007a). These observations are consistent with a pathway of a-NHEJ that relies on small regions of homology (so called "microhomologies") for repair. This pathway, referred to as microhomology-mediated end-joining (MMEJ), is to date poorly defined but is generally characterized by evidence of break point deletions and ~5-25-nt regions of break point homology (McVey and Lee, 2008). Because homology at DSB ends is not required for NHEJ in mammalian cells and end-joining events that carry no homology have been observed even in cells deficient for c-NHEJ factors (Soulas-Sprauel et al., 2007), the contribution of MMEJ to end-joining under wild-type conditions is not well understood; however, evidence suggests that this pathway may be of particular relevance to V(D)J recombination and class switch recombination (Corneo et al., 2007; McVey and Lee, 2008).

There is still much left to learn regarding the pathways of alternative NHEJ, and a particular challenge will be to define the proteins that regulate them. Interestingly, however, early evidence suggests that a-NHEJ is mediated (at least in part) by the activities of the SSBR effectors PARP and XRCC1-LIG3 (Audebert et al., 2004; Wang et al., 2006), perhaps most convincingly because DSB end-joining in $Lig4^{-/-}$ and $Ku80^{-/-}$ mammalian cells is decreased by PARP inhibition (Wang et al., 2006).

3-3. Double-strand break repair: homologous recombination (HR)

In contrast to NHEJ, homologous recombination is an "error-free" method of DSB repair initiated during the S and G2 phases of the cell cycle that engages the newly replicated sister chromatid as a template for repair by DNA synthesis (Heyer et al., 2010; San Filippo et al., 2008; Symington and Gautier, 2011). HR mechanisms are known to repair DBSs caused by ionizing radiation, various chemical agents (including many chemotherapeutics), collapsed replication forks and damaged replication intermediates, abortive topoisomerase activity, and endogenous or exogenous endonuclease activity.

3-3-1. A Historical Perspective

Genetic recombination has arguably been studied for over 100 years. It was first observed at the turn of the 20th century by William Bateson and Reginald C. Punnett as an exception to Gregor Mendel's rule of independent assortment, which defines the nature of genetic inheritance in its most simplistic form; it states that specific genes (known to Mendel only as amorphous particles) responsible for the manifestation of associated phenotypes are transferred from parent to offspring as independent units. Mendel derived this law from phenotypic analysis of his pioneering genetic crosses performed with the garden pea plant *Pisum sativum*; his results were published in the 1866 paper *Versuche über Pflanzenhybriden* / *Experiments on Plant Hybridization*. Almost 40 years later Bateson and Punnett conducted similar mating experiments using the pea species *Lathyrus odoratus*. They observed, however, that some inherited traits demonstrated "coupling or repulsion" in the offspring of particular crosses because they separated more or less frequently than the law of independent assortment would predict (Bateson et al., 1905). This phenomenon was soon described in terms of genetic linkage by Thomas Hunt Morgan, and the synthesis of this observation with the concept that genes exist on chromosomes led to the inference of meiotic recombination. Meiosis is a specialized cell cycle that produces haploid gamete cells (or spores) necessary for the sexual reproduction of diploid organisms, and genetic recombination in meiosis can result in bidirectional exchanges (crossovers) of genetic information between parental

homologous chromosomes. This was described by Morgan and colleagues as such, "If two factors lie in the same member of a chromosome pair we should expect them always to be found together in successive generations of a cross unless an interchange can take place between such a chromosome and the homologous chromosome derived from the other parent" (Morgan, 1911a, b; Morgan et al., 1915). The physical existence of these "interchanges" was proven in 1931 by Harriet B. Creighton and Barbara McClintock through a perceptive combination of inferred genetic and direct cytological evidence (Creighton and McClintock, 1931; McClintock, 1931).

In meiosis, one round of DNA replication is followed by two rounds of cell division that sequentially segregate homologous chromosomes (meiosis I) and then sister chromatids (meiosis II) into four haploid gamete or spore cells. This differs from mitosis. In mitosis, DNA replication is followed by a single segregation event that separates sisters into a pair of diploid daughter cells. In meiosis I, homologous chromosomes pair and align along the equatorial plane of the cell (the metaphase plate) before they are pulled into separate daughter cells by microtubules. Meiotic crossovers, which occur during the pachytene stage of meiosis I, establish stable connections between homologous chromosomes (chiasmata) that are necessary for proper alignment and subsequent segregation of the chromosomes (Marston and Amon, 2004). In this way meiotic crossovers serve as physical linkages relevant to the specialized mechanics of meiosis. However, meiotic crossovers have also been hypothesized to serve a role in evolution because they increase genetic diversity within diploid species over what can be achieved through the selection (in meiosis) and mixing (during fertilization) of parental chromosomes.

It is now understood that meiotic crossover events occur specifically at sequences flanking programmed double strand breaks initiated by the conserved type II topoisomerase SPO11 (Bergerat et al., 1997; Keeney et al., 1997) and that crossover formation is completed through a particular mechanism of homology-directed DSB repair, known as the double strand break repair (DSBR) pathway (Figure 4e) (Szostak et al., 1983). The key to this mechanism is that broken chromosomes –carrying SPO11-induced DSBs– engage their homolog pairs as templates for repair by DNA synthesis. In the process, complex DNA junctions and intermediate structures form between chromosomes. It is the resolution of these

structures necessary to once again separate chromosomes that causes the exchange of genetic information and formation of crossover products (Figure 4). Of note a second but unidirectional form of sequence exchange, gene conversion (GC), can also occur in which newly replicated DNA sequences are incorporated into broken chromosomes while the flanking sequences are left unchanged (non-crossover products) (Szostak et al., 1983).

Interestingly, somatic cells have co-opted a similar mechanism of recombination to deal with unscheduled DSBs. In these cells, however, recombination occurs primarily between homologous regions of sister chromatids and, as such, is restricted to the S and G2 phases of the mitotic cell cycle during which newly replicated sister chromatids are available for use as repair templates. Sisters, unlike homologs, have identical DNA sequences, and template-use restriction allows DSB repair to proceed through homologous recombination without generating major sequence alterations to broken chromatids, which can be detrimental to somatic cells. The idea that DNA repair in mitosis can occur through recombination was first proposed by Robin Holliday (Holliday, 1964) and was supported by genetic work in *E. coli* and the smut fungus *Ustilago maydis* that indicated an association between radiation sensitivity and defective genetic recombination (Clark and Margulies, 1965; Holliday, 1962; Holliday, 1967). Similar studies in *E. coli* and the budding yeast *Saccharomyces cerevisiae* (*S. cerevisiae*), subsequently identified the first genetic regulators of homologous recombination, including genes encoding the central HR recombinases RecA and RAD51 (Clark and Margulies, 1965; Game and Mortimer, 1974). Many of these regulators have since been found to be well conserved among all eukaryotes, with a degree of conservation in prokaryotes as well.

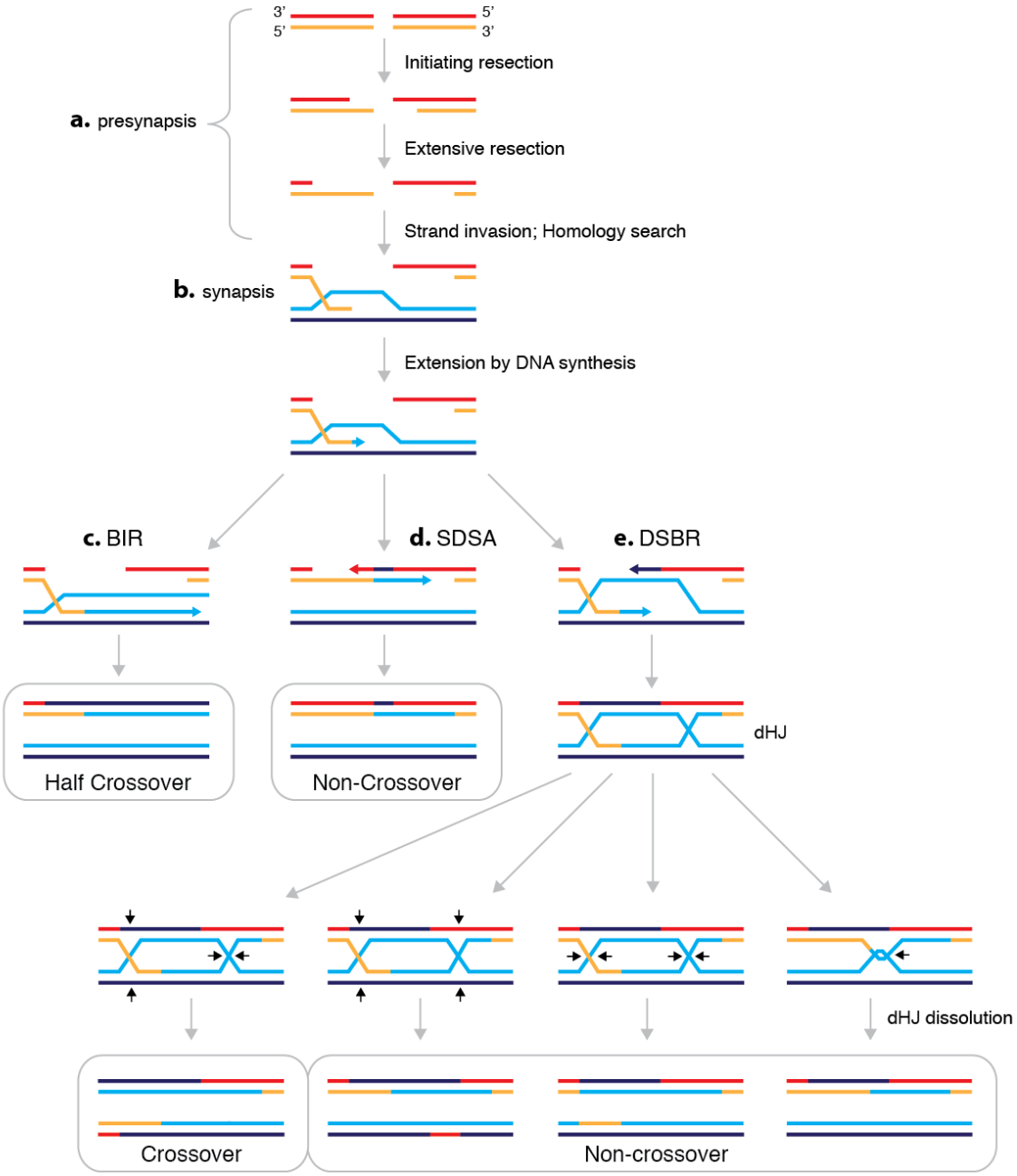


Figure 4. Schematic of DNA intermediate structures that arise during homologous recombination. **(a)** After DSB formation, presynaptic end processing occurs, first as limited initiating resection and then by extensive or "long-range" resection. A homology search coordinated by the recombinase RAD51 ensues. **(b)** Invasion of the presynaptic filament and pairing with a region of complimentary homologous dsDNA constitutes synapsis. **(c-e)** Postsynaptic repair can proceed through multiple pathways: **(c)** Break-induced replication (BIR), **(d)** Synthesis-dependent strand annealing (SDSA), and **(e)** Double-strand break repair (DSBR), which leads to the formation of double Holliday junctions (dHJs). Possible repair products, crossover and non-crossover, from each pathway are indicated. Black arrows indicate endonucleolytic cleavage events responsible for the resolution and dissolution of dHJs. These pathways are discussed in greater detail throughout Chapter 1.

3-3-2. Models of homologous recombination

There are three models of DSB repair by homologous recombination: the aforementioned double strand break repair model (DSBR), the synthesis-dependent strand annealing model (SDSA) and break-induced replication (BIR) (Figure 4). The classic DSBR model, first proposed by Szostak and colleagues (Szostak et al., 1983), is thought to be most active in meiosis and predominately result in the formation of crossover products, while the SDSA model is considered the primary repair pathway for gene conversion events without crossover formation in both mitosis and meiosis (McMahill et al., 2007). The BIR model is induced for repair when only one end of a double stand break has template homology, such as in the case of broken replication forks or eroded telomeres. Repair by BIR can result in non-reciprocal translocation products or "half-crossovers" and therefore is not expected to be substantially active in somatic cells. This pathway has been well characterized in yeast, but its contribution to repair in mammals is not yet fully understood (Bosco and Haber, 1998; McEachern and Haber, 2006). These three HR pathways, although distinct, can all be divided into two prominent stages: before and after annealing of the broken DNA to a region of homology, or pre and postsynapsis, and all three models of HR share a common mechanism of presynapsis.

3-3-3. Presynapsis: Resection

The initiating event that commits DSBs to repair by HR is 5' to 3' DNA end resection (Figure 3c,d, 4a). This process creates 3' ssDNA overhangs at break ends that are responsible for pairing with homologous DNA templates (synapsis) in order to generate heteroduplex DNA structures (D-loops) primed for DNA synthesis and subsequent repair (Figure 3g, 4b). In eukaryotes, resection is mediated by a set of conserved and redundant nucleases, MRE11, CtIP (Sae2 in the budding yeast *S. cerevisiae*; Ctp1 in the fision yeast *Schizosaccharomyces pombe*), EXO1 and DNA2, as well as the RecQ DNA helicase BLM (Sgs1 in *S. cerevisiae*; Rqh1 in *S. pombe*) (Symington and Gautier, 2011). The coordinated mechanism by which these redundant proteins promote resection has been largely defined by work in

yeast; however, evidence suggests that the prevailing "2-step" model is conserved in mammals. Briefly, MRE11 (as part of the MRN complex) and CtIP facilitate initial processing of DSB DNA ends (Figure 3c) especially in the context of complex or "dirty" ends that carry modified / damaged nucleotides or covalent DNA-protein crosslinks. This initial processing is then followed by extensive or "long-range" DNA resection meditated by two redundant pathways, one facilitated by EXO1 and the other by DNA2 in coordination with the BLM-TOP3-RMI1-RMI2 helicase complex (Sgs1-Top3-Rmi1 in *S. cerevisiae*) and the SMARCAD1 nucleosome remodeler (Figure 3d) (Costelloe et al., 2012). Extensive resection is thought to ensure that sufficient ssDNA is generated for proper homologous pairing as well as checkpoint activation through ATR (Figure 3e) (Symington and Gautier, 2011). ATR activation is discussed in Chapter 1, 2-1.

The redundant and interdependent nature of the resection machinery complicated early genetic and biochemical studies aimed at parsing the pathway, and the "2-step" model was not reached until the contributions of Exo1, Dna2, and Sgs1 were discovered in *S. cerevisiae* (Gravel et al., 2008; Mimitou and Symington, 2008; Zhu et al., 2008). Before this, studies focused primarily on the nuclease activities of the yeast MRN ortholog MRX (Mre11-Rad50-Xrs2). The conserved MRE11 subunit of MRX contains an N-terminal $Mn2^{+}$/$Mg2^{+}$-dependent phosphoesterase domain, and both yeast and human MRE11 proteins demonstrate 3' to 5' dsDNA exonuclease activity as well as ssDNA endonuclease activity *in vitro* (Moreau et al., 1999; Paull and Gellert, 1998; Trujillo et al., 1998). Yeast *mre11Δ*, *rad50Δ* and *xrs1Δ* mutants demonstrate strong sensitivity to IR and delayed but not deficient *in vivo* resection of DSBs induced by the site-specific endonuclease HO (Ivanov et al., 1994; Moreau et al., 1999). While these mutant phenotypes indicate a role for Mre11 and MRX in DNA end resection, the observed 3' to 5' exonuclease function is opposite the predicted activity, and *mre11* nuclease defective point mutants have markedly less sensitivity to IR and no defect in HO-induced DSB resection (Bressan et al., 1999; Llorente and Symington, 2004; Moreau et al., 1999).

The observation that Mre11 nuclease activity is required for resection of Spo11-induced DSBs, however, lead to the hypothesis that the nuclease activity of Mre11 is only required for resection when it

necessitates the processing of complex DSB ends, such as those formed by Spo11, while other functions of the Mre11 protein are required for resection in general (Krogh and Symington, 2004; Moreau et al., 1999). Unlike HO, which generates a "clean" DSB with free 3' hydroxyl groups and 5' phosphates, Spo11 remains covalently linked to DSB ends with a 5' orientation. In yeast, processing of these complex ends requires both Mre11 and Mre11 nuclease activity, as well as a second ssDNA endonuclease Sae2, which is stimulated *in vitro* by MRX (Lengsfeld et al., 2007). *sae2Δ* cells, similar to *mre11* nuclease mutants, demonstrate moderate sensitivity to IR and defective resection of Spo11-induced DSBs; however, they show only slightly delayed resection of HO-induced DSBs indicating that the role of Sae2 in resection is also primarily required at complex ends (Clerici et al., 2005; Symington and Gautier, 2011).

Recently, a model of bidirectional resection at Spo11-linked DSBs has been put forward (Garcia et al., 2011). This model proposes precise functions for the endonuclease and 3' to 5' exonuclease activities of Mre11; specifically, it suggests that Spo11-linked DNA strands are nicked by the endonuclease activity of Mre11 approximately 300 nucleotides from DSB ends and that these strands are then simultaneously resected by the respective 3' to 5' and 5' to 3' exonuclease activities of Mre11 and Exo1. This mechanism would eventually release a Spo11-linked oligonucleotide, which is experimentally observed (Garcia et al., 2011). To support this model, Garcia et al. generated and evaluated a useful endonuclease proficient but partially exonuclease deficient *mre11-H59S* mutant.

The discovery that Exo1 and Dna2-Sgs1 mediate redundant pathways of long-range resection in *S. cerevisiae* clarified the much-debated roles of Mre11 nuclease and Sae2 function in "clean" DSB end processing and inspired the "2-step" model of DNA resection (Gravel et al., 2008; Mimitou and Symington, 2008; Zhu et al., 2008). In this model, MRX and Sae2 coordinate the initial processing of all DSBs breaks ("clean" or "dirty") but leave bulk resection to the activities of Exo1 and Dna2-Sgs1 (Mimitou and Symington, 2008; Zhu et al., 2008). Together the work of Mimitou, Symington, Zhu and colleagues, showed that *sgs1Δ*, *dna2Δ* and *exo1Δ* mutants all have reduced rates of DNA resection approximately 3-28 kb away from HO-induced DSBs but have unaltered levels of resection immediately

adjacent to the break sites (as compared to controls); *rad50Δ* and *mre11Δ* cells, on the other hand, were shown to have defective resection initiation but normal long-range rates, which is indicative of the "2-step" model (Mimitou and Symington, 2008; Zhu et al., 2008). Importantly, the combined deletion of *SGS1* and *EXO1* was observed to synergistically reduce resection, and *sgs1Δ exo1Δ* double mutants were shown to process only a few hundred nucleotides at DSB ends in an Sae2- and Rad50-dependent manner (Gravel et al., 2008; Mimitou and Symington, 2008; Zhu et al., 2008). These genetic observations placed Sgs1 and Exo1 resection activities into separate pathways downstream of MRX; however, because *SGS1* and *DNA2* demonstrate no genetic interaction a single Dna2-Sgs1 pathway is thought to exist. This "2-step" model retrospectively explains the intermediate phenotypes observed in *mre11* nuclease deficient and *sae2Δ* mutants by supposing that Exo1 and Dna2-Sgs1 require MRX- and Sae2-mediated end processing to facilitate loading onto complex but not "clean" DSB ends (Symington and Gautier, 2011). Following these genetic studies, biochemical work showed that DNA resection can be reconstituted *in vitro* with purified Sgs1 and Dna2 in a manner dependent on RPA and stimulated by both MRX and Top3-Rmi1 (Cejka et al., 2010; Niu et al., 2010). Recently, the nucleosome remodeler Fun30 (SMARCAD1 in humans) has also been shown to contribute to long-range resection in collaboration with both Exo1 and Dna2-Sgs1 (Chen et al., 2012; Costelloe et al., 2012; Eapen et al., 2012).

In mammals, the MRN proteins NBS1 and RAD50, the Sae2 ortholog CtIP, and the nuclease activities of MRE11 are all required for embryonic viability and cell proliferation (Buis et al., 2008; Chen et al., 2005; Luo et al., 1999; Zhu et al., 2001) which has presented challenges to studying DNA end resection in mammalian cells. The use of conditional alleles and RNA interference (RNAi) has, nevertheless, allowed analysis of these orthologs. MRE11-dependent activation of ATM in mammalian cells (as described in Chapter 1, 2-1) does not require MRE11 nuclease activities; however unlike in yeast, murine cells that express mutant MRE11 but lack the associated nuclease activities phenocopy strong DSB repair defects observed in MRE11 null cells, including increased sensitivity to IR, deficiency in DNA end resection (as determined by IR-induced RPA foci), and reduced gene conversion at DSBs induced by the rare cutting I-SceI endonuclease (Buis et al., 2008). Unlike its yeast ortholog, CtIP also

appears to be required for DSB end resection and efficient HR in mammalian cells, although surprisingly no CtIP nuclease activity has yet been demonstrated (Sartori et al., 2007). Taken together, these data may indicate that CtIP and MRE11 both have larger roles in facilitating DSB end resection than their yeast orthologs or may suggest that DSB ends in mammalian cells require more complex processing than in yeast. However, because I-SceI (like Spo11) remains bound to DNA after cleavage and IR creates chemically modified DNA ends, it remains undetermined if CtIP and MRE11 are required for the resection of "clean" DSB ends.

Nevertheless, evidence suggests that the "2-step" model is conserved. In mammalian cells, as in yeast, EXO1 and BLM (the ortholog of Sgs1) facilitate *in vivo* DNA end processing as components of parallel resection machineries downstream of MRE11 and CtIP (Bolderson et al., 2009; Eid et al., 2010; Gravel et al., 2008). These pathways have been reconstituted *in vitro* using purified human proteins (Nimonkar et al., 2011; Nimonkar et al., 2008). From this work, we have learned that although DNA2 is a bidirectional exonuclease (Masuda-Sasa et al., 2006), DNA2 coupled with BLM *in vitro* and in the presence of RPA resects DNA specifically 5' to 3' in a manner dependent on the nuclease activity of DNA2, the helicase activity of BLM and stimulated by MRN. EXO1, on the other hand, is specifically a 5' to 3' exonuclease that is alone sufficient to degrade DNA *in vitro* (Lee and Wilson, 1999); however like DNA2, EXO1-mediated resection is stimulated both by MRN and BLM (Nimonkar et al., 2011; Nimonkar et al., 2008). The stimulation of EXO1 by BLM stands in contrast to the *in vivo* data from both yeast and human cells indicating that EXO1 and BLM operate in distinct resection pathways (Gravel et al., 2008; Mimitou and Symington, 2008; Zhu et al., 2008); however, this stimulation has been observed to be independent of BLM helicase activity and unable to increase the processivity of EXO1 (Nimonkar et al., 2011). EXO1 stimulation by BLM, therefore, does not represent enzymatic synergy and may simply result from a favorable physical interaction between proteins (Nimonkar et al., 2008). The stimulatory effects of MRN in these *in vitro* experiments was largely attributed to the ability of MRN to increase EXO1- and BLM-binding to DNA, which is consistent with *in vivo* data demonstrating that MRE11 and CtIP are required for EXO1 recruitment to DNA breaks (Eid et al., 2010). Importantly, this suggests that

EXO1 functions downstream of MRN and CtIP in agreement with yeast data and "2-step" model of nuclease activity.

In yeast Sgs1 is the only RecQ family helicase, but in humans there are five; these are BLM, WRN, RECQ4, RECQ1 and RECQ5 (Bernstein et al., 2010). While other human RecQ proteins have known roles in DNA repair, BLM is considered to be the most orthologous to Sgs1 because it maintains conserved associations with TOP3, RMI1 and RMI2 (orthologs of the yeast Sgs1-interaction proteins Top3 and Rmi1) (Bernstein et al., 2010). Consistent with this assumption, BLM-mediated stimulation of *in vitro* resection by EXO1 and DNA2 was found to be specific, and the WRN, RECQ5 and RECQ4 paralogs showed no similar effect (Nimonkar et al., 2011; Nimonkar et al., 2008). TOP3 is a type I topoisomerase that directly binds BLM (Johnson et al., 2000; Wu et al., 2000). RMI1/2 are OB-fold containing proteins essential for the stability of the BLM-TOP3-RMI1-RMI2 complex (Singh et al., 2008; Xu et al., 2008; Yin et al., 2005). RMI1 and RMI2 are also required for BLM localization to DNA breaks (Singh et al., 2008; Xu et al., 2008; Yin et al., 2005). OB-folds similar to those in RMI1/2 mediate ssDNA binding in both the RPA complex and BRCA2 protein (Chapter 1, 3-3-5); however, the OB-folds in RMI2 lack key amino acid residues critical for DNA binding, and neither RMI1 nor RMI2 have been shown to bind ssDNA. The BLM-TOP3-RMI1-RMI2 complex has multifaceted roles in mammalian homologous recombination, making it difficult to determine the specific functions (if any) of the TOP3 and RMI1/2 proteins in mammalian resection. In yeast however, *top3Δ* and *rmiΔ* mutants have resection defects similar to those in *sgs1Δ* cells (Zhu et al., 2008), and *in vitro* Top3 and Rmi1 stimulate resection by Dna2-Sgs1-RPA (Cejka et al., 2010; Niu et al., 2010). The function of Top3 in resection is not dependent on its topoisomerase activity (Niu et al., 2010), making it likely that these Sgs1/BLM accessory factors function in resection only to stabilize BLM at DSBs. A description of the role that the BLM-TOP3-RMI1-RMI2 complex has in postsynaptic homologous recombination can be found in Chapter 1, 3-3-7.

3-3-4. Presynapsis: Formation of the RAD51 nucleoprotein filament

The next step in homologous recombination is the assembly of RAD51 recombinase molecules onto newly resected ssDNA at DSB ends to form right-handed helical nucleoprotein filaments (Figure 3f). These presynaptic RAD51 filaments catalyze homology sampling and DSB end pairing with homologous regions of dsDNA (synapsis). They are essential for repair by HR. Initially, resected DSB ends are coated with RPA complexes, heterotrimers of RPA1, RPA2 and RPA3 that bind indiscriminately to ssDNA formed during repair or at replication forks (Figure 3e). RPA-binding stabilizes ssDNA and minimizes the formation of ssDNA secondary structures. In the context of HR, RAD51 molecules rapidly displace RPA.

RAD51 is a 37 kDa ATPase conserved from bacteria to humans. The first member of the conserved RAD51 family to be discovered was identified 50 years ago in a genetic screen for mutant alleles that cause recombination defects in *E. coli* almost (Clark and Margulies, 1965). This was RecA. Functionally, RecA was characterized as a recombinase because of its unusual *in vitro* activity to facilitate the invasion of ssDNA into dsDNA duplexes and promote heteroduplex pairing between homologous regions (Radding, 1989). RAD51 proteins were subsequently identified in eukaryotes and characterized with similar activities (Baumann et al., 1996; Benson et al., 1994; Sung, 1994). In general, RAD51 proteins are divergent on the N- and C-termini but have a conserved center region containing two domains that mediate ATP binding and hydrolysis (Walker A and B). Human RAD51 shares 56% amino acid homology (30% identity) with the bacterial RecA (Shinohara et al., 1993; Yoshimura et al., 1993).

RAD51 filament formation proceeds in two steps: nucleation followed by elongation. Our understanding of these processes has been greatly enhanced by studies of nucleoprotein formation using individual single- and double-strand DNA molecules (Holthausen et al., 2010). For example, it has been observed that both RAD51 and RecA nucleate stochastically along DNA, and that the cooperative binding of 4-5 RAD51 monomers constitutes a successful nucleation event (Galletto et al., 2006; Holthausen et al., 2010; Modesti et al., 2007). Subsequent elongation of RAD51 and RecA filaments is estimated to

proceed by 1-7 monomers at a time (Holthausen et al., 2010), and for RecA, elongation occurs bidirectionally with distinct end kinetics that cause cumulative growth in the 3' direction (Joo et al., 2006). Importantly, RAD51 and RecA demonstrate divergent rates of nucleation and elongation. RecA nucleation occurs rarely and with rate-limiting kinetics, but once nucleated, efficient RecA elongation generates long continuous filaments. RAD51, on the other hand, demonstrates more efficient nucleation followed by elongation rates that are limited by the density of adjacent nucleation clusters (Holthausen et al., 2010; Modesti et al., 2007; van der Heijden et al., 2007). In this manner, filament formation is more cooperative for RecA than RAD51, but once formed, RAD51 filaments are thought to have greater flexibility (Holthausen et al., 2010).

Active RecA and RAD51 nucleoprotein filaments can be generated *in vitro* with a binding stoichiometry of ~3 DNA nucleotides per recombinase monomer (Arata et al., 2009; Baumann and West, 1997; Zlotnick et al., 1993). RecA filaments generated under these conditions can then bind naked ssDNA molecules thereby increasing stoichiometry to ~7 nucleotides per monomer. This second strand capture forms joint filaments thought to mimic the structure of *in vivo* RecA / RAD51 bound to heteroduplex DNA at the sites of recombination (Zlotnick et al., 1993). DNA incorporated into recombinase filaments is both extended and underwound, an effect that likely facilitates heteroduplex pairing (Benson et al., 1994; Chen et al., 2008b).

As mentioned, RAD51 family proteins have conserved ATP-binding and DNA-dependent ATPase activity (Benson et al., 1994; Radding, 1989). ATP-binding occurs at the interface of RecA / RAD51 monomers to promote both filament formation and stabilization (Chen et al., 2008b); however, neither nucleation nor elongation requires ATP hydrolysis (Holthausen et al., 2010; Kowalczykowski and Eggleston, 1994). Perhaps surprisingly, ATPase activity is also dispensable for recombinase-mediated homology sampling and heteroduplex paring but is active in the process of recombinase disassociation from DNA. Data from a single molecule study of human RAD51 filaments suggest that RAD51 chains composed of both ATP- and ADP-bound RAD51 are stabilized by end monomers bound exclusively to ATP and support a model wherein ATP hydrolysis by end monomers triggers segmented bursts of

processive ADP-bound RAD51 dissociation, which occurs monomer-by-monomer until the next ATP-bound RAD51 is reached (van Mameren et al., 2009). In this light, RecA and RAD51 nucleoprotein formation, stabilization, and dissociation must be considered dynamic processes regulated by intermittent ATP hydrolysis events.

3-3-5. Presynapsis: Mediators of the RAD51 nucleoprotein filament

RAD51, unlike RecA, assembles onto both single- and double-strand DNA with equal efficiency (Benson et al., 1994); however, RAD51 binding to dsDNA is inhibitory for strand exchange (Sigurdsson et al., 2001) and therefore must be regulated in eukaryotic cells. The role of preferentially loading RAD51 onto ssDNA is attributed to mediator proteins, and in mammals, the most prominent of these is BRCA2 (Figure 3f). *In vivo* colocalization of BRCA2 and RAD51 in discrete subnuclear foci and a direct physical interaction between the proteins first suggested that BRCA2 might function as a RAD51 regulator (Scully et al., 1997b). Subsequently, murine and human cells deficient for BRCA2 were found to have increased sensitivity to DNA damaging agents and defects in DSB repair that were specific to gene conversion pathways (Moynahan et al., 2001; Sharan et al., 1997). Importantly, when present in the human germline, heterozygous loss-of-function mutations in *BRCA2* are known to convey an increased risk of breast and ovarian cancer, with a predicated probability of 40-80% for breast cancer (Fackenthal and Olopade, 2007).

BRCA2 is a 3418 amino acid protein, with no obvious ortholog in yeast or bacteria. Efforts to determine the mediator function of BRCA2 by *in vivo* exogenous expression or *in vitro* biochemical assays were initially limited by the large size of the protein, difficulties with expression, and poor solubility. Many BRCA2-mediated functions in homologous recombination were, however, correctly inferred from studies that investigated the behavior of BRCA2 fragments and domain fusions (Carreira et al., 2009; Pellegrini et al., 2002; Wong et al., 1997; Yang et al., 2002) as well as smaller orthologs of the protein, including Brh2 from *U. maydis* (Yang et al., 2005). Between residues 990 and 2100 human BRCA2 has 8 repeat regions (~30 amino acids each) called BRC domains that facilitate binding to

RAD51 (Wong et al., 1997), and at the C-terminal end of BRCA2 there is a DNA-binding domain (DBD) containing a tower domain and 3 ssDNA-binding OB-folds similar to those in RPA and RMI1/2 (Yang et al., 2002). The structure of the tower domain is indicative of dsDNA-binding (Yang et al., 2002). *In vitro*, BRCA2 BRC repeats stimulate RAD51 filament formation on ssDNA while limiting filament formation on dsDNA (Carreira et al., 2009), and studies of purified Brh2, which carries a conserved DBD and one BRC repeat, showed that this BRCA2 ortholog preferentially nucleates RAD51 onto ssDNA at dsDNA / ssDNA junctions (Yang et al., 2005). Contrary to these results, *in vivo* expression of BRC4 has been shown to disrupt RAD51 filament formation, as evident by loss of RAD51 nuclear foci after DNA damage (Chen et al., 1999); however, this result most likely represents inhibitory sequestration of RAD51 away from DNA by an isolated BRC repeat and illustrates the critical nature of coordinated BRCA2-RAD51 and BRCA2-DNA interactions (Pellegrini et al., 2002). A crystal structure of BRC4 bound to RAD51 revealed that BRC repeats structurally resemble the oligomerization domain of RAD51 (Pellegrini et al., 2002). This insight led to a functional model wherein BRCA2 accumulates at DSBs through direct DNA binding to provide a platform for initial RAD51 nucleation, overall reducing nucleation energy constraints. Evidence also suggests that BRCA2 stabilizes RAD51-ssDNA filaments. Consistent with this, RAD51 Walker domains in BRC-bound RAD51 adopt a more closed confirmation that may inhibit ATP hydrolysis (Carreira et al., 2009; Pellegrini et al., 2002).

In 2010, three groups successfully purified human BRCA2 as a full-length protein (Jensen et al., 2010; Liu et al., 2010; Thorslund et al., 2010) and i*n vitro* evaluations proved that the functionalities ascribed to BRCA2 by previous models were largely accurate. These studies confirmed that BRCA2 stimulates homology-mediated strand exchange by facilitating RAD51 nucleation onto ssDNA and supported the "platform" hypothesis by demonstrating that loading occurs *in vitro* at substoichiometric levels. Approximately 5-6 RAD51 monomers were estimated to bind one BRCA2 molecule (Jensen et al., 2010; Liu et al., 2010), and stimulation of RAD51 loading onto RPA-coated ssDNA was estimated to occur at a stoichiometry of one BRCA2 molecule to 33-100 RAD51 monomers (Jensen et al., 2010; Liu et al., 2010). These works also demonstrated that full-length BRCA2 preferentially binds ssDNA over

dsDNA favoring RAD51 filament formation on ssDNA substrates (both naked and RPA-coated) (Jensen et al., 2010; Thorslund et al., 2010) and confirmed that BRCA2 inhibits RAD51-mediated ATP hydrolysis (favoring filament stabilization) (Jensen et al., 2010). Surprisingly however, full-length BRCA2, unlike Brh2, did not show a strong preference for binding ssDNA / dsDNA junctions over ssDNA (Jensen et al., 2010; Liu et al., 2010), although in *in vitro* strand exchange experiments a slight preference was observed for substrates containing 3' or 5' DNA overhangs over ssDNA (Jensen et al., 2010). In combination with the finding (described in Chapter 1, 3-3-4) that human RAD51 demonstrates a higher nucleation rate than RecA and forms many disrupted RAD51 chains rather than longer continuous filaments (van der Heijden et al., 2007), the observation that BRCA2 does not preferentially act at DNA junctions becomes less surprising and, in fact, suggests that human BRCA2 functions by nucleating RAD51 filaments along the length of resected DNA.

DSS1 and PALB2 are two BRCA2-assocaiated proteins also considered mediators of presynaptic filament formation in mammalian cells (Figure 3f). DSS1 is a small (70 amino acid) acidic protein that was first identified as the product of one of three genes occupying a deletion region of chromosome 7 known to convey the autosomal dominant developmental disorder split hand / split foot disease (Crackower et al., 1996). At the cellular level, depletion of DSS1 causes increased chromosomal abnormalities and defective RAD51 focus formation after exposure to DNA damage (Gudmundsdottir et al., 2004). DSS1 binds the C-terminal DNA-binding region of BRCA2 (Marston et al., 1999). A functional role for this DSS1-BRCA2 interaction in the formation of RAD51-ssDNA filaments remains somewhat elusive; however, it may well be attributed to *in vivo* stabilization of the large BRCA2 protein and its insoluble DBD. In two separate studies DSS1 was copurified with Brh2 or BRCA2 (specifically a fragment containing the DBD) in order to increase solubility of the target peptide (Yang et al., 2002; Yang et al., 2005), and while depletion of DSS1 from mammalian cells was shown to cause reduction of wild-type BRCA2, it did not affect the stability of a BRCA2 fragment lacking the DBD (Li et al., 2006). Interestingly, in one biochemical study, DSS1 was also shown to have an *in vitro* stimulatory effect on

BRCA2-facilitated RAD51 nucleation suggesting a possible mechanistic role for DSS1 in presynaptic filament formation (Liu et al., 2010).

PALB2 is a second BRCA2-interacting protein required for BRCA2 protein stability. PALB2, however, specifically stabilizes the nuclear faction of BRCA2 in mammalian cells (Xia et al., 2006). PALB2 and BRCA2 colocalize at nuclear foci marking sites of DNA damage, and depletion of PALB2 causes defective BRCA2 focus formation as well as reduced gene conversion efficiency (Xia et al., 2006). While PALB2, like DSS1, is predominantly considered to be a "caretaker" of BRCA2 stability and nuclear function, PALB2 has recently also been shown to have binding affinities for RAD51 and DNA as well as *in vitro* activity for stimulating RAD51-mediated D-loop formation and synaptic capture in the absence of BRCA2 (Dray et al., 2010). Indicative of a critical role overall, heterozygous mutation of germline *PALB2* has been associated with breast cancer development in humans (Erkko et al., 2007; Rahman et al., 2007).

Additional proteins involved in the formation of the presynaptic filament are encoded by genes within the *RAD52* epistasis group. In *S. cerevisiae*, this group contains *RAD50, RAD51, RAD52, RAD54, RAD55, RAD57, RAD59, MRE11,* and *XRS2*. Genetic grouping of these genes was initially defined by phenotypes common to mutant alleles and indicative of similar functions in DNA repair (Game and Mortimer, 1974). We now know that several of these genes also have conserved human orthologs, *RAD52, RAD54, RAD51B, RAD51C, RAD51D, XRCC2* and *XRCC3,* that are likewise involved in presynaptic filament formation.

Although unrelated by sequence, the *S. cerevisiae* Rad52 protein is functionally similar to human BRCA2. It binds Rad51 (Shinohara et al., 1992), catalyzes an exchange of RPA for RAD51 on ssDNA, and promotes homology-directed strand pairing (New et al., 1998; Shinohara and Ogawa, 1998). In a similar manner, human RAD52 binds RAD51 and has been shown to stimulate RAD51-mediated pairing between homologous single- and double-strand DNA *in vitro* (in the presence and absence of RPA) (Benson et al., 1998; Shen et al., 1996). It is widely accepted, however, that human RAD52 does not predominantly function *in vivo* as a mediator of the presynaptic filament as it does in yeast. The most

striking evidences to this effect are disparate phenotypes associated with RAD52 loss in yeast and vertebrate cells. In yeast, *rad52* mutants demonstrate strongly defective HR, while vertebrate cells without *RAD52* have much milder defects (Rijkers et al., 1998; Yamaguchi-Iwai et al., 1998). Indicative of an *in vivo* requirement for Rad51 assembly, *rad52* yeast mutants do not form Rad51 foci during meiosis (Gasior et al., 1998). *RAD52*$^{-/-}$ DT40 chicken cells, on the other hand, demonstrate no defect in RAD51 foci formation after ionizing radiation compared to controls (Yamaguchi-Iwai et al., 1998). Taken together these data indicate that during vertebrate evolution the mediator functions of Rad52 found in *S. cerevisiae* were largely reassigned to BRCA2. Interestingly however, a genetic interaction between *RAD52* and *BRCA2* in the mammalian recombination pathway has recently suggested that mammalian RAD52 has retained a "backup" role in presynaptic filament formation. Specifically, *in vivo* depletion of RAD52 in BRCA2 deficient cells was shown to exacerbate RAD51 foci and homologous recombination defects (Feng et al., 2011). Although the specific contribution of RAD52 to this early step of homologous recombination in vertebrate cells remains unclear, it is worth noting here that RAD52 has well defined roles in other processes relevant to homology-directed repair. In particular, RAD52 mediates ssDNA annealing that is critical to postsynaptic HR mechanisms (Chapter 1, 3-3-7) and RAD51-independent single strand annealing (SSA), the details of which are discussed below (Chapter 1, 3-4).

3-3-6. Synapsis

Once formed, a single RAD51-coated presynaptic filament can invade dsDNA and initiate a poorly understood mechanism of homology sampling with the strand of opposite polarity (Figure 3g, 4b). When a region of sufficient homology is identified heteroduplex pairing (synaptic capture) occurs to prime the broken DNA end for repair by DNA synthesis. In meiosis, synapsis occurs between homologous chromosomes, representing an intermediate step in the generation of meiotic crossovers, but in mitosis this process occurs preferentially between sister chromatids in order to maintain fidelity of the somatic genome (Bzymek et al., 2010). The intermediate structure formed by disruption of dsDNA and displacement of the non-complementary strand is called a D-loop, and D-loop formation is the point after

which the three models of homologous recombination, double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA) and break-induced replication (BIR), diverge (Figure 4c-e).

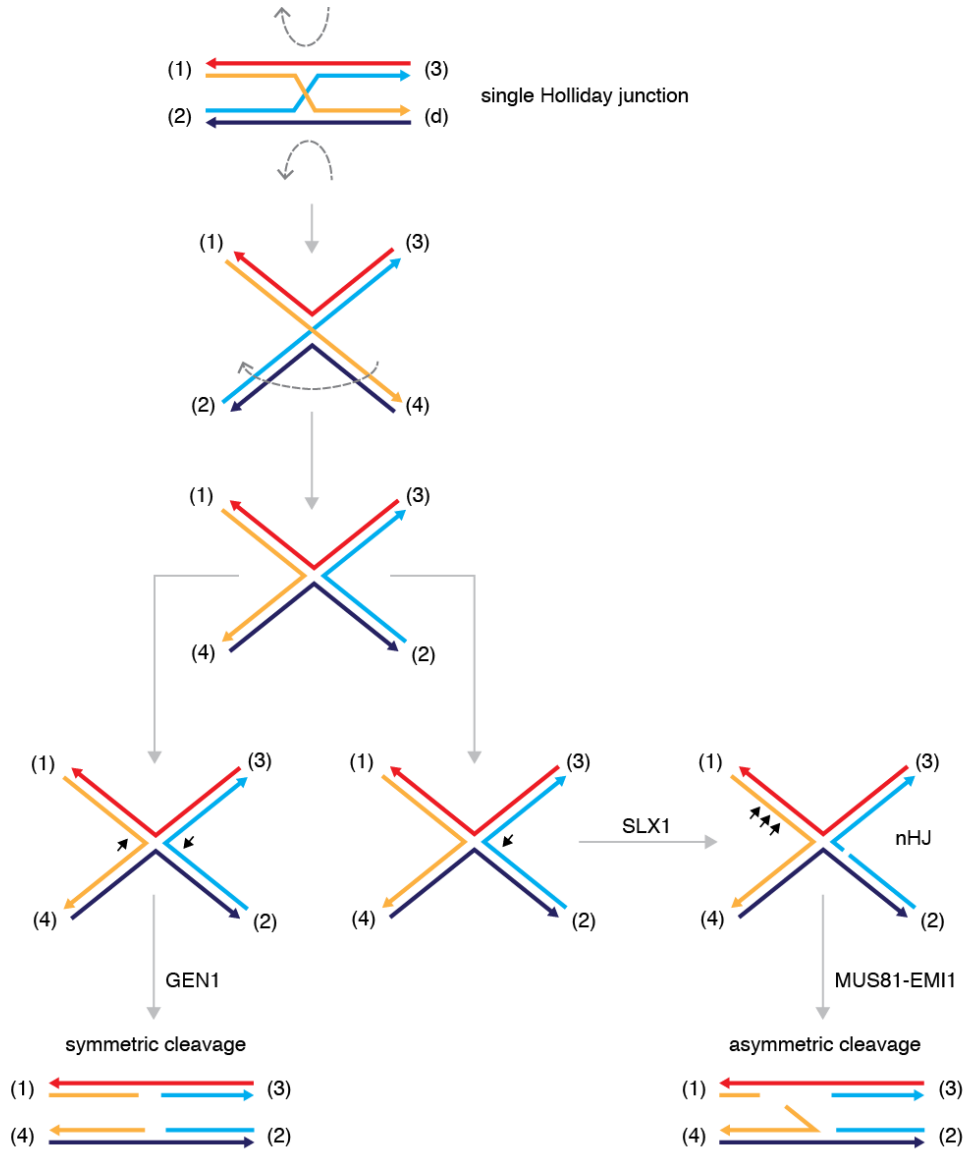


Figure 5. The Holliday junction. Dotted arrows indicate the directions in which illustrated strands can be logically manipulated to produce the next Holliday junction representation. (1)-(4) indicate specific dsDNA ends. Black arrows indicate endonucleolytic cleavage events mediated by GEN1, SLX1, and MUS81-EMI1. nHJ indicates a nicked Holliday junction.

3-3-7. Postsynapsis: Formation, resolution and dissolution of DNA intermediate structures

In the classic double-strand break repair (DSBR) model, DNA strands displaced by D-loop formation pair with the second resected ends of broken dsDNA molecules (Figure 4e). This process, called second end capture, allows for simultaneous repair of both DSB ends through parallel DNA synthesis and subsequent ligation. Biochemical analyses have shown that second end capture can be mediated *in vitro* by the ssDNA annealing activity of RAD52, which is discussed in greater detail below (Chapter 1, 3-4) (McIlwraith and West, 2008; Nimonkar et al., 2009; Reddy et al., 1997). A key intermediate structure in DSBR is the double Holliday junction (dHJ) (Szostak et al., 1983) (Figure 4), a complex physical linkage between dsDNA molecules that is composed of two adjacent DNA intersections, or single Holliday junctions (HJs) (Figure 5). By definition, a single Holliday junction is the point of strand exchange between two DNA helices that occurs when the single strands of one helix pair with complementary sequences in the other; or more simply, a single Holliday junction can be thought of as a four-stranded DNA cross structure (Figure 5). Single Holliday junctions were first proposed by Robin Holliday in the 1960's to conceptually model the DNA transactions that occur during gene conversion in fungi (Holliday, 1964), and since then both single and double Holliday junctions have been widely accepted as key HR intermediates (Szostak et al., 1983; West, 2009). Importantly, evidence of *in vivo* dHJ formation has been obtained from cells actively cycling in meiosis, and to a lesser extent in mitosis, suggesting the *in vivo* presence of a regulated DSBR-like repair pathway (Bzymek et al., 2010).

Logically, separation of two DNA helices interlinked by a single Holliday junction can happen in only two ways: by disrupting DNA base pairing along the lengths of two arms or by nicking two DNA strands with the same sequence polarity (5' to 3' or 3' to 5') near the junction site. This latter nick-based scenario is a process called Holliday junction resolution. Because of HJ symmetry, resolution of a single HJ can occur in two possible orientations determined by the non-complementary pair of strands on which the separating nicks occur. In the context of a single Holliday junction, the choice of nick orientation dictates only whether the newly freed heteroduplex arms must pass, or "cross over," each other during

duplex separation (Figure 5); however in the context of a double Holliday junction, dual orientation choice controls the extent of sequence exchange between the DNA helices and determines whether or not crossover products are formed (Figure 4). Resolving HJ nicks, made in either orientation, can also occur symmetrically or asymmetrically with regard to position along the chosen strand (Figure 5). If cut symmetrically, the HJ structure will be resolved into products ready for direct ligation; however, if made asymmetrically, the separated duplexes will carry a gap or DNA flap and require additional processing prior to ligation. Importantly, the nick orientation and position constraints of single Holliday junction resolution are similar to those of dHJ resolution with the added consideration of an adjacent junction.

Enzymatic dHJ disassembly is a critical and much considered step in the DSBR model of HR. To this end, two distinct pathways of dHJ disassembly, resolution and dissolution, have been found. These are controlled by distinct enzymatic activities, the first of which was discovered in 1982 and belonged to the endonuclease VII protein from T4 bacteriophage, an HJ resolvase (Mizuuchi et al., 1982). Almost a decade later, the first cellular HJ resolvase RuvC was discovered in *E. coli* (Connolly et al., 1991). Because these first-identified resolvases were shown to cleave HJ structures with symmetrically positioned nicks, they became known as "classic" HJ resolvases. Within the last 20-30 years, however, asymmetric resolvases with important roles in HR have also been identified. The first indication of classic resolvase activity in mammalian cells was found in protein purifications from homogenized calf thymus in 1990 (Elborough and West, 1990). However, due to major redundancies in eukaryotic resolution activity, and among postsynapsis HR mechanisms in general, the particular enzyme responsible for this classic activity, the so-called ResA, was not identified until almost two decades later when GEN1 (an XPG-type endonuclease) was finally characterized (Ip et al., 2008). *In vitro,* GEN1 and the *S. cerevisiae* ortholog Yen1 cleave static HJs symmetrically with nicks positioned 1 nucleotide away from the junction point in the 3' direction (Figure 5) (Ip et al., 2008).

Before GEN1 and Yen1 were identified, MUS81 was the best-studied mediator of HJ resolution in eukaryotes. MUS81 is a well-conserved ERCC4-type endonuclease that acts as part of a heterodimeric complex with the non-catalytic EME1. In 2001, the Mus81 ortholog in fission yeast was shown to have

structure-specific resolvase activity (Boddy et al., 2001) and human MUS81 was put forward as a promising candidate for mammalian ResA (Chen et al., 2001). Soon afterward however, MUS81-mediated HJ resolution was shown to be asymmetric (Constantinou et al., 2002), and purification of the MUS81-EME1 heterodimer revealed that, unlike ResA, this complex does not efficiently cleave static HJ structures *in vitro* but preferentially cleaves 3'-flap and replication fork structures (Ciccia et al., 2003). Interestingly, these preferred MUS81 substrates both contain a DNA branch point adjacent to a free 5' phosphate and resemble HJ structures with a single strand nick near the junction (nicked HJ or nHJ) (Figure 5). Current evidence suggests that MUS81 contributes to the resolution of nicked Holliday junctions by asymmetric cleavage of the DNA strand opposite the nick (~3-6-nt 5' of the branch point) (Hollingsworth and Brill, 2004).

Recent work has implicated the UvrC-type endonuclease SLX1 in the *in vivo* generation of nHJ substrates for asymmetric cleavage by MUS81-EME1 (Figure 5) (Svendsen and Harper, 2010). *SLX1* was first identified in yeast as a gene that incurs synthetic lethality with *SGS1* and *TOP3* (SLX, synthetic lethal of unknown function) (Mullen et al., 2001). A related gene, also found in this work, was shown to be synthetic-lethal with *SLX1*, as well as *SGS1* and *TOP3*. This gene was *SLX4*. These genetic interactions, in combination with evidence that Slx1 and Slx4 physically interact *in vivo*, led the authors of this early study to predict that Slx1 and Slx4 act together in a heterodimeric complex to process HR intermediate DNA structures (Mullen et al., 2001). Later, this same group demonstrated that an Slx1-Slx4 heterodimer can function *in vitro* as a structure-specific endonuclease (SSE) with a preference for cleaving 5'-flap and replication fork structures (Fricke and Brill, 2003). Notably, these structures are inverted MUS81-EME1 substrates. When the Slx4 ortholog BTBD12 (renamed SLX4) was isolated from human cells, it was shown that, in addition to SLX1, human SLX4 binds the heterodimeric endonucleases MUS81-EME1 and ERCC4-ERCC1 (Fekairi et al., 2009; Munoz et al., 2009a; Svendsen et al., 2009). Interestingly, SLX4 complexes were shown to cleave static HJs symmetrically (2-nt 3' to the branch point) and nHJ asymmetrically (3-5-nt 5' to the branch point). Through a set of clever experiments that physically separated SLX4 interactions into SSE subcomplexes, the former symmetric activity was

ascribed to SLX1-SLX4 and the later to MUS81-EME1-SLX4 (Svendsen et al., 2009). From this, a model of ordered HJ cleavage was proposed in which HJs are first cleaved by SLX1-SLX4 to generate nHJs and then resolved by asymmetric MUS81-EME1-SLX4 cleavage (Figure 5) (Svendsen et al., 2009). Interestingly, no HJ resolvase activity was ascribed to the ERCC4-ERCC1-SLX4 subcomplex.

As mentioned previously, resolution of dHJs can result in the formation of distinct crossover and non-crossover products depending on the orientation of DNA strand cutting (Figure 4). Non-crossover products retain a majority of their original presynaptic sequence, while crossover products, generated through the swapping of entire sequence arms with their synaptic partners, can have substantially altered sequence information. Because crossover formation is inherently mutagenic, it may be evolutionarily advantageous in meiosis; however, for the same reason it must be tightly repressed in mitosis. For this reason, any dHJs that arise during mitosis must be carefully deconstructed into non-crossover products. An appealing model for crossover control is one of SSE regulation that dictates the orientation of dHJ resolution within the different cell cycle programs. However, a second method of dHJ disassembly, which is exclusive to non-crossover product formation, may simply out compete resolution for dHJ substrates in times when crossover formation is inappropriate. This method is dHJ dissolution.

In contrast to dHJ resolution, in which adjacent DNA junctions are individually cleaved and religated, dHJ dissolution proceeds by the active migration of dHJ junctions into a single catenated point that is then decatenated to separate DNA helices (Figure 4). dHJ branch migration and subsequent decatenation are coordinated by the helicase and the topoisomerase activities of BLM-TOP3-RMI1-RMI2 complex, also called the dissolvasome (Figure 3) (Karow et al., 2000; Wu and Hickson, 2003). This multifaceted complex is described above in its alternative capacity for generating long-range resection in eukaryotic cells (Chapter 1, 3-3-3). Interestingly, human cells deficient in BLM demonstrate hyperrecombination between sister chromatids (Ellis et al., 1999; German et al., 1977), a phenotype observed by the increased occurrence of sister chromatid crossover structures (SCEs). This phenotype is contrary to what would be expected for a singular BLM requirement in the resection step of HR, but can be explained by the expectation that crossover-prone resolution is expected to predominate in the absence

of dissolution. In fact, increased SCEs are strongly indicative of dissolution defects. TOP3-dependent and BLM-, RMI1/2-stimulated dHJ dissolution has been elegantly demonstrated *in vitro* using a small dHJ substrate (Raynard et al., 2006; Singh et al., 2008; Wu and Hickson, 2003; Xu et al., 2008).

The simplest model for repression of crossover formation during postsynaptic repair is suppression of dHJ formation altogether. To this end, the synthesis-dependent strand annealing (SDSA) repair model has been proposed (Figure 4d), wherein D-loop formation at repair sites is not succeeded by second end capture, and repair by DNA synthesis occurs with serial rather than parallel coordination to prevent dHJ intermediates. As in DSBR, synapsis occurs when an invading RAD51-coated DNA strand anneals to a region of dsDNA homology. In SDSA, however, this invading strand is displaced from the D-loop structure after extension by DNA synthesis so that it may reanneal to the complementary strand of the broken DNA molecule and serve as a template for the completion of repair.

In yeast, the Srs2 helicase promotes SDSA. In higher eukaryotes, SDSA is regulated by a functional analog of Srs2, the conserved RAD3-like helicase RTEL1, presumably though its capacity to disrupt D-loop structures (Figure 3). Similar to the discovery of *SLX1*, *RTEL1* was found in a screen for genes that incur synthetic lethality with a conserved *BLM / SGS1* ortholog, in this case the nematode *Caenorhabditis elegans him-6* (Barber et al., 2008). *In vitro,* RTEL1 disrupts D-loop structures without disrupting RAD51 nucleoprotein filaments and with a preference for 3' invasion structures notably reminiscent of *in vivo* repair intermediates (Barber et al., 2008; Youds et al., 2010). Interestingly, *C. elegans rtel-1* mutants were shown to produce an increased number of double- and triple-crossover events per chromatid in meiosis when compared to controls (Youds et al., 2010) indicating that SDSA also functions in meiosis, as was proposed in 2001 by Neil Hunter and Nancy Kleckner to explain their observation that strand invasion of single DSB ends occurs in yeast meiosis (Hunter and Kleckner, 2001). This suggests that SDSA may drive repression of crossover formation in either cell cycle program by direct competition with DSBR. How the choice between SDSA and DSBR is regulated will be an interesting question for future work.

3-3-8. The role of BRCA1

BRCA1 is a critical HR mediator and tumor suppressor in mammals (Huen et al., 2010). In humans, *BRCA1* is similar to *BRCA2* in that heterozygous germline mutations in either gene convey strong predispositions to cancer (Venkitaraman, 2002), a fact that is incorporated in both gene names, Breast Cancer 1/2, early onset. However unlike BRCA2, the mechanism of BRCA1 function in HR is not straightforward, and because of this, BRCA1 has been the subject of intense study over the last decade.

BRCA1 is an 1863 amino acid protein that binds several important DDR proteins through multiple identifiable domains and, in this way, can be thought of as a structural scaffold for coordinating a diverse set of DDR processes. BRCA1 binds the ubiquitin E3 ligase BARD1 through an N-terminal RING domain (Wu et al., 1996), and *in vivo* these proteins form an obligate heterodimer (Yu and Baer, 2000). Suggestive of highly cooperative *in vivo* functions, mouse models have shown that the tumor suppressor effects of BRCA1 and BARD1 are relatively indistinguishable (Shakya et al., 2008). BRCA1-BARD1 also has been identified as the core subunit of at least three distinct BRCA1 supercomplexes, BRCA1-A, BRCA1-B and BRCA1-C. These multiprotein complexes are formed by the mutually exclusive binding of phosphorylated Abraxas, BACH1 and CtIP proteins (Greenberg et al., 2006; Huen et al., 2010) to a phospho-peptide-binding module of tandem BRCT repeats present on the BRCA1 C-terminus (Manke et al., 2003; Yu et al., 2003). The functions of these BRCA1 supercomplexes in DNA repair, as well as the interrelationships between them, are only beginning to be understood; however, their mechanisms of action in the DDR are likely ones of local regulation because, like many mediators of the DDR, BRCA1 and its associated proteins, accumulate at sites of DNA damage (Scully et al., 1997b).

The BRCA1-A complex, consisting of BRCA1-BARD1, Abraxas, RAP80, BRCC36, BRE and NBA1, is perhaps the best characterized of the BRCA1 supercomplexes, and it is responsible for a majority of BRCA1 accumulation at DNA breaks (Dong et al., 2003; Feng et al., 2009; Kim et al., 2007; Shao et al., 2009; Sobhian et al., 2007; Wang et al., 2009; Wang et al., 2007). The detailed mechanism by which BRCA1-A recruitment occurs is as follows: ATM, activated by the MRN complex, phosphorylates serine-139 of histone H2AX in chromatin immediately adjacent to DNA lesions (Rogakou

et al., 1999). This phospho-mark then recruits the MDC1 protein through direct binding to an MDC1 BRCT domain, and MDC1, in turn, binds ATM and is phosphorylated (Lou et al., 2006; Stewart et al., 2003; Stucki et al., 2005). Next, the phosphorylated region of MDC1 recruits RNF8 through FHA-mediated binding, and once there, RNF8 together with UBC13 initiate polymerization of K63-linked poly-ubiquitin in the local chromatin (Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007; Wang and Elledge, 2007). These ubiquitin chains then recruit RNF168 through binding of RNF168 ubiquitin interaction motifs (MIUs); after which, RNF168 and UBC13 amplify the initial ubiquitylation signal (Doil et al., 2009; Pinato et al., 2009; Stewart et al., 2009) by ubiquitylating the K13-15 residues of histones H2A and H2AX in the chromatin surrounding DNA lesions (Mattiroli et al., 2012). This facilitates BRCA1-A complex recruitment in a manner dependent on the ubiquitin interaction motifs (UIMs) of its RAP80 subunit (Figure 6) (Kim et al., 2007; Sobhian et al., 2007; Wang et al., 2007). Of note, the critical interaction between RNF8 and UBC13 is mediated by phosphorylation-induced binding of HERC2 to an RNF8 FHA domain (Bekker-Jensen et al., 2010).

Due in part to the functional nature of their protein subunits, the BRCA1-B (BRCA1-BARD1, BACH1 and TOPBP1) and BRCA1-C (BRCA1-BARD1, CtIP, and MRN) complexes have been described as regulators of the S phase replication checkpoint and DNA end resection, respectively (Huen et al., 2010). Consistent with this, depletion of BRCA1 from mammalian cells compromises the formation of ssDNA and loading of RAD51 at DSBs (Schlegel et al., 2006), and evidence suggests that the interaction between BRCA1 and CtIP is necessary for both processes (Chen et al., 2008a; Yun and Hiom, 2009). An additional interaction between BRCA1 and PALB2, facilitated by a coiled-coiled region in BRCA1, physically links BRCA1 to BRCA2 (Figure 3f). To date, it is unclear whether the BRCA1-PALB2-BRCA2 (BPB) complex is distinct from BRCA1-A, -B and -C; however, the interaction between BRCA1 and PALB2 has convincingly been shown to contribute RAD51 presynaptic filament assembly and HR efficiency (Sy et al., 2009; Zhang et al., 2009), and BRCA1 mutations that disrupt BRCA1-PALB2 binding have been identified in cancer patients (Sy et al., 2009). Different groups have reported conflicting evidence relevant to whether BRCA1 controls PALB2 and BRCA2 localization to

sites of DNA damage (Sy et al., 2009; Zhang et al., 2009), and the mechanism by which BRCA1-PALB2 binding controls HR warrants additional investigation.

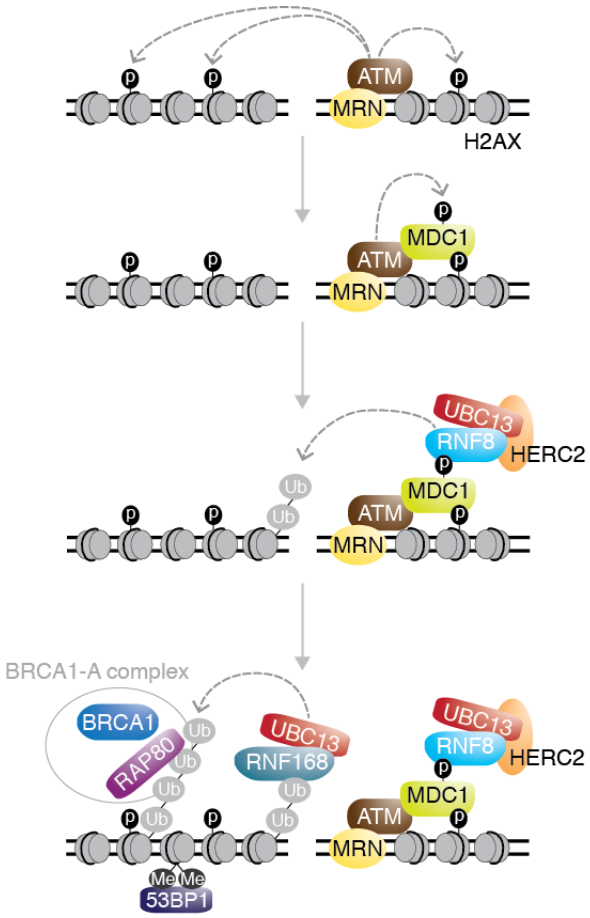


Figure 6. Schematic of BRCA1-A complex and 53BP1 recruitment to DNA damage. ATM is activated by the MRN complex and phosphorylates H2AX to recruit MDC1. ATM phosphorylates MDC1 to recruit RNF8. RNF8-HERC2-UBC13 builds K63-Ub chains to recruit RNF168. RNF168-UBC13 amplifies the ubiquitylation signal to recruit 53BP1 and the BRCA1-A complex. These mechanisms are discussed in greater detail in Chapter 1, 3-3-8.

Although the BRCA1-C and BPB complexes seem to have clear roles in promoting HR, BRCA1 recruitment to sites of DNA damage has been shown to depend predominantly on the BRCA1-A complex, and interestingly, (as highlighted by work presented herein) key components of BRCA1-A are not required for HR. This suggests one of two possibilities: (1) efficient HR requires relatively minimal BRCA1-C and / or BPB at DNA breaks, the recruitment of which occurs independently of BRCA1-A or (2) these complexes contribute to HR without local action. A recent study suggests that the former possibility is most likely (Hu et al., 2011). The authors of this study found that disruption of BRCA1 localization to IR-induced foci through depletion of the RAP80 BRCA1-A subunit actually increases

colocalization of BRCA1 with CtIP and BACH1 at remaining foci. They also observed that while CtIP (BRCA1-C) and BACH1 (BRCA1-B) depletion decreases HR efficiency, depletion of RAP80, Abraxas or BRCC36 (BRCA1-A) have the opposite effect and increase HR. These data suggest that BRCA1 is recruited to sites of DNA damage in a BRCA1-A independent manner as part of BRCA1-C, -B or possible other complexes and present the possibility that the various BRCA1 complexes fine-tune DSB pathway choice at DSBs, with BRCA1-C and BRCA1-B promoting and BRCA1-A suppressing HR (Hu et al., 2011).

The BRCA1 and BARD1 proteins both contain RING domains, which are zinc finger structural regions associated with ubiquitin E3 ligase activity; as mentioned above, the BRCA1 RING binds BARD1. Both proteins are also capable of promoting ubiquitylation individually, and BRCA1-BARD1 heterodimerization increases E3 ligase activity over what either protein is capable of alone (Hashizume et al., 2001). Because both proteins appear to have essential roles in HR, it is tempting to speculate that BRCA1-BARD1 mediated ubiquitylation, perhaps of DDR target substrates, contributes to DSB repair; however, the role of this activity remains unclear, and recent evidence suggests that BRCA1 ligase activity is not required for HR. Specifically, mutation of the BRCA1 RING domain in a manner that abrogates BRCA1 E3 ligase activity but does not affect BARD1 binding (BRCA1-I26A) confers no substantial change to MMC sensitivity, RAD51 foci formation, or gene conversion efficiency (Reid et al., 2008). Additionally, this mutation does not appear to promote cancer formation in mice as would be expected for a genetic HR defect (Shakya et al., 2011). A second BRCA1 RING mutation (C61G) that abrogates both BRCA1 E3 ligase activity and BARD1 binding, however, is pathogenic in both mice and humans (Drost et al., 2011). Taken together these data suggest that although the BRCA1-BARD1 interaction may contribute to HR, BRCA1 ligase activity does not (Elia and Elledge, 2012).

3-4. Double-strand break repair: Single-strand annealing (SSA)

Single-strand annealing (SSA) is an alterative method of homology-mediated DSB repair that is conceptually much less complicated than HR. Briefly, when resection of DSB ends generates ssDNA overhangs containing complementary regions of DNA, direct annealing between these regions can promote RAD51-independent repair. This repair pathway is facilitated primarily by RAD52, which has been shown *in vitro* to catalyze RPA-stimulated ssDNA annealing (Baumann and West, 1999; Benson et al., 1998). RAD52, from both yeast and human cells, oligomerizes into ring-like structures that have been shown through electron microscopy and crystal structure analysis to contain ~7-11 unit monomers (Singleton et al., 2002; Stasiak et al., 2000). These oligomers can then combine into higher order structures through further RAD52 self-association (Van Dyck et al., 2001). Visualization of RAD52 in the presence of DNA has revealed that RAD52 oligomerization occurs selectively on DNA ends in a protective manner that facilitates DNA end-to-end fusions (Van Dyck et al., 1999, 2001). *In vivo*, the SSA pathway has been observed primarily from evaluating repair products formed when DSBs are induced between direct repeats (Fishman-Lobell et al., 1992).

3-5. Double-strand break repair: Regulation of pathway choice

Although much is known about the pathways that facilitate DSB repair (NHEJ, HR, and SSA), somewhat less is understood of the mechanisms that regulate repair pathway choice. With regard to HR, the predominant method of regulation seems to be tight coupling to the S and G2 phases of the cell cycle, and consistent with this, it has been observed in yeast that G2-arrested cells are gene conversion proficient, while G1-arrested cells are GC deficient (Aylon et al., 2004). Additionally, it has been found that cyclin-dependent kinase (CDK) activity, which is high in G2 and low in G1, is required for resection in both yeast and mammalian cells, indicating that conserved mechanisms of HR regulation focused on the resection machinery have evolved (Ira et al., 2004; Jazayeri et al., 2006). CDK-dependent phosphorylation of CtIP on serine-327 during S and G2 regulates formation of the resection-promoting

BRCA1-C complex (Figure 3c) (Chen et al., 2008a; Yu and Chen, 2004; Yun and Hiom, 2009), and cell cycle regulated CDK1 phosphorylation of threonine-847 on CtIP (S267 in Sae2) promotes resection through an unknown mechanism that operates at the level of CtIP recruitment to DNA damage (Huertas et al., 2008; Huertas and Jackson, 2009). Regulation of HR through resection in particular, suggests that the formation of ssDNA at DSBs represents a physical commitment to homology-directed repair (either HR or SSA over NHEJ). This idea is supported by the observation that cells depleted of HR effectors upstream of resection can bypass the G2 / M checkpoint in an NHEJ-dependent manner, but those deficient for HR downstream of resection cannot (Cotta-Ramusino et al., 2011).

CDK1 activity has also been shown to promote HR through suppression of a Ku70-Ku80-mediated inhibitory effect that acts on MRX-controlled resection (Clerici et al., 2005; Tomita et al., 2003). Genetic evidence in yeast suggests that MRX and Ku70-Ku80 compete to regulate resection in the positive and negative directions, respectively (Tomita et al., 2003), and this antagonistic relationship likely represents a direct physical competition for DSB end binding. Consistent with this, Ku binds ssDNA much less efficiently than dsDNA (Mimori and Hardin, 1986), and Ku deletion in mammalian cells increases HR mediated gene conversion (Pierce et al., 2001).

Recent evidence has uncovered a second competitive mechanism of HR regulation in mammalian cells involving BRCA1 and 53BP1 (Figure 3c). 53BP1 is a key transducer of the DDR required for induction of the damage-specific G2 / M checkpoint (Wang et al., 2002), and as discussed above in Chapter 1, 3-3-8 the BRCA1 protein is required for proper end resection due to roles enacted as part of the BRCA1-C and / or BPB complexes. Interestingly, deletion of 53BP1 in BRCA1 deficient cells has been found to rescue homologous recombination defects, as indicated by the reestablishment of both RAD51 foci formation and resection-indicative RPA phosphorylation (Bouwman et al., 2010; Bunting et al., 2010), suggesting that 53BP1 negatively modulates the HR. Because 53BP1 is recruited to sites of DNA damage by the same ubiquitylation cascade that promotes BRCA1-A complex recruitment (Figure 6) (Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007; Wang and Elledge, 2007), this observation is consistent with the hypothesis that BRCA1-A recruitment to DNA lesions (as opposed to BRCA1-C)

also causes repression of HR (discussed above in Chapter 1, 3-3-8). Interestingly, a recent study that used super-resolution microscopy to observe IR-induced foci determined that S phase damage foci have distinct central compartments that stain for BRCA1 but not 53BP1, indicating that a direct physical competition between the BRCA1-C / BPB complexes and 53BP1 may exist at DNA breaks (Chapman et al., 2012).

Control through cell cycle-regulated CDK phosphorylation and antagonistic relationships between repair pathway effectors have emerged as the primary models for regulation of homologous recombination and DSB repair pathway choice. With regard to the cell cycle regulated control of HR however, an intriguing question remains unanswered, which is how the cell distinguishes between a replicated portion of DNA during S phase (which would presumably be eligible for repair by HR due to the presence of an intact sister chromatid) and an as yet unreplicated section of DNA that cannot engage a sister chromatid for repair (and would therefore presumably be restricted from engaging the HR machinery). Additionally, although we now have some understanding of how HR is controlled through resection; little is know about the regulation of post-resection choices, such as those between HR and SSA and between DSBR, SDSA, and BIR. No doubt post-translational events that modulate these choices will also exist. Additional work to address these issues will be of considerable interest.

Chapter Two

A genetic screen for regulators of mammalian homologous recombination

Britt Adamson[1], Agata Smogorzewska[1,2], Stephen J. Elledge[1]

[1]Division of Genetics, Brigham and Women's Hospital, Department of Genetics, Harvard University Medical School, Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA

[2]Laboratory of Genome Maintenance, The Rockefeller University, New York, New York 10065, USA

Work herein can be found in the following publication:

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol *14*, 318-328.

Tables S1-4 can be found on the accompanying data CD included with this dissertation

The screen presented herein was designed by Britt Adamson, Agata Smogorzewska and Stephen J. Elledge and was conducted by Britt Adamson in collaboration with Agata Smogorzewska. Data analysis was performed by Britt Adamson. We thank Dr. Maria Jasin for DR-U2OS cells and Dr. Philip Ng and Dr. Frank Graham for the AdNGUS24i. We thank the Institute of Chemistry and Cell Biology-Longwood screening facility, including Caroline Shamu, Stewart Rudnicki, Sean M. Johnston and Tiao Xie for screening support.

CHAPTER TWO

I. Introduction

1-1. Screen Rationale

As discussed previously (Chapter 1, 3-3-1), the study of homologous recombination is historically linked with the advent of genetics as a scientific discipline, and as such, its conserved mechanisms have, in one way or another, long been a central focus of biological inquiry. Highly efficient homologous recombination in yeast has for decades provided an almost ideal and genetically tractable model system for the direct investigation of HR pathways. The practical advantages of studying HR in yeast are numerous. Easy interchange between diploid and haploid states as well as efficient gene targeting allow for both forward and reverse genetic approaches; straightforward and inexpensive techniques for laboratory growth and selection provide a realistic platform for scalable and systematic inquiry; and biological conservation within the meiotic and mitotic homologous recombination pathways gives weight to relevant findings. Moreover, regulated gene conversion events of natural origin, which occur within the mating (*MAT*) locus of haploid yeast cells, provide an almost ideal experimental readout for the evaluation of HR effector function.

Although HR is highly conserved mechanistically and with regard to the sequences, structures, and functions of key effectors, evolutionary pressures have diverged HR in mammals from lower eukaryotes to some extent. So while genetic work in yeast, primarily *S. cerevisiae*, has generated many fundamental insights into the mechanistic underpinnings of HR and DNA repair, such work has limitations with regard to its applicability to human health, and experimentally delineating mammalian-specific regulators of HR has, therefore, been of substantial interest. Particular motivation in this regard has been driven by knowledge that several human HR effectors are also critical tumorsuppressors, the most famous examples of which are BRCA2 and BRCA1 (Chapter 1, 3-3-5 and 3-3-8). These proteins lack obvious orthologs in yeast but are required for proper HR in vertebrates, and in humans, loss-of-

function heterozygous germline mutations in either gene carries significantly increased risk of cancer development (Venkitaraman, 2002).

In general, the most useful tool for identifying novel pathway effectors in model organisms is genome-scale genetic screening, and because human cell culture recapitulates many aspects of *in vivo* cell growth, it is a ready platform for functional human genetics. The adaptation of genetic methods to cell culture, however, has presented some unique challenges. In particular, barriers to bi-allelic mutation of minimally diploid mammalian cells has limited forward genetic approaches; and although the isolation of haploid cancer cell lines now presents an opportunity for mutagenic screening (Carette et al., 2009; Reiling et al., 2011), this application has, to date, not been proven for the use of identifying non-selectable phenotypes, such as sensitivities to DNA damaging agents caused by loss-of-function HR mutations. The low efficiency of gene targeting in mammalian cell culture combined with the unfeasibility of manipulating successfully targeted genes to homozygosity has also limited traditional reverse genetic approaches; and while the generation of mammalian cells carrying diploid deletion mutations is possible on a gene-by-gene scale through whole-organism mouse genetics, the scale of production that would be necessary to conduct reverse genetic screening with genome-wide coverage using this approach is prohibitive. The discovery of gene silencing by RNA interference (RNAi) in the late 1990s / early 2000s and the recent availability of genome-wide RNAi libraries, however, has provided the first opportunity to systematically evaluate mammalian genetics in cell culture and, as yet, is the only practical method for genome-wide interrogation of mammalian pathways using reverse genetics. Here we present a genome-wide small interfering RNA (siRNA)-based screen for regulators of homologous recombination in human cells that has identified a highly validated list of mammalian HR genes and generated novel insights into RNAi screening technology.

1-2. RNAi-mediated gene silencing and genome-wide RNAi screening technologies

RNAi is an endogenous method of posttranscriptional gene silencing caused by short double-strand RNA molecules (dsRNA). It was first observed in plants (Napoli et al., 1990; van der Krol et al., 1990), later characterized in *C. elegans* (Fire et al., 1998), and has since been identified in many eukaryotes, including plants, fungi, animals, and of particular interest, humans. Perhaps the most important aspect of RNAi has been its aforementioned application to experimental biology, wherein the engineered reutilization of its conserved, endogenous pathways has allowed genetic inquiry in otherwise genetically intractable biological systems.

In simple terms, RNAi can be described as the complementary pairing of short RNA molecules (usually between 19-23-nt) to targeted protein-coding messenger RNAs (mRNAs) that elicits gene silencing through message degradation or translational repression; mechanisms of RNAi are discussed in Chapter 3, Section I. In the experimental setting, introduction of exogenous dsRNA engineered to repress particular mRNAs through sequence targeting achieves directed gene silencing, commonly referred to as gene "knockdown" or protein depletion. This technique has been widely used for posttranscriptional genetic studies in model organisms, including *Drosophila* and *C. elegans*, as well as in mammalian cell culture, and the generation of genome-wide RNAi libraries has prompted large-scale genetic screening in these systems. Such RNAi-based screening has substantially impacted biological discovery within the last decade and has contributed to our collective understanding of many cellular processes, including cell viability and proliferation, cancer biology, the cell cycle, stress responses, cell death, signal transduction, and RNAi biology itself (Mohr et al., 2010).

Two varieties of exogenous RNAi reagents and two corresponding methods of cell delivery predominate in cell culture based RNAi screens: viral transduction of small hairpin RNAs (shRNAs) (Silva et al., 2008) and lipid transfection of small interfering RNAs (siRNAs). While both shRNAs and siRNAs elicit gene knockdown through the same endogenous pathway, shRNAs and siRNAs have distinct entry points into the RNAi machinery. shRNAs are designed as hairpin structures within the

context of an endogenous microRNA gene (microRNAs are discussed in Chapter 3, 1-2). Predominantly, these are transduced into the genome of a host cell by viral infection where they are stably encoded as DNA. This shRNA-encoding DNA can then be transcribed into RNA by host polymerases and processed into short dsRNA molecules by the host RNAi machinery. siRNAs, on the other hand, are short dsRNA molecules synthesized *in vitro* for direct introduction into mammalian cells by lipid transfection. In both cases, one strand of the resulting short dsRNA is incorporated into a protein complex called RISC (RNA-induced silencing complex) to coordinating gene silencing. The technical differences between siRNA and shRNA technologies provide two distinct but complementary strategies for genetic screening in mammalian cells: arrayed and pooled (Mohr et al., 2010). Batch infection with targeted shRNA libraries has proven successful for pooled selection- or dropout-based screens, while arrayed transfection of siRNA libraries using microtiter plates has been most applicable to high-throughput screening of phenotypes that are not resolvable by changes in cell proliferation. In order to probe the mammalian HR pathway with specificity, we chose an siRNA array-based strategy using the Dharmacon human siGENOME siRNA library.

II. Results

2-1. A genome-wide siRNA screen to identify regulators of homologous recombination

For our screen we used a well-characterized GFP-based reporter (DR-GFP) (Figure 7a) (Pierce et al., 1999; Xia et al., 2006). DR-GFP carries two mutant versions of *GFP*; one with two premature stop codons and an internal I-SceI endonuclease restriction site (*SceGFP*), the other with 3' and 5' end truncations (*iGFP*) (Pierce et al., 1999). Neither *SceGFP* nor *iGFP* express a functional protein; however, a gene conversion event between the mutants –generated by recombinational repair of an I-SceI-induced DSB– can reconstitute wild-type *GFP*. In this manner nuclear GFP expression is an accurate and relatively simple readout for HR. For our purposes, we employed the osteosarcoma cell line DR-U2OS that has a single, stably integrated copy of DR-GFP (Xia et al., 2006), and we drove expression of I-SceI with an adenovirus (AdNGUS24i).

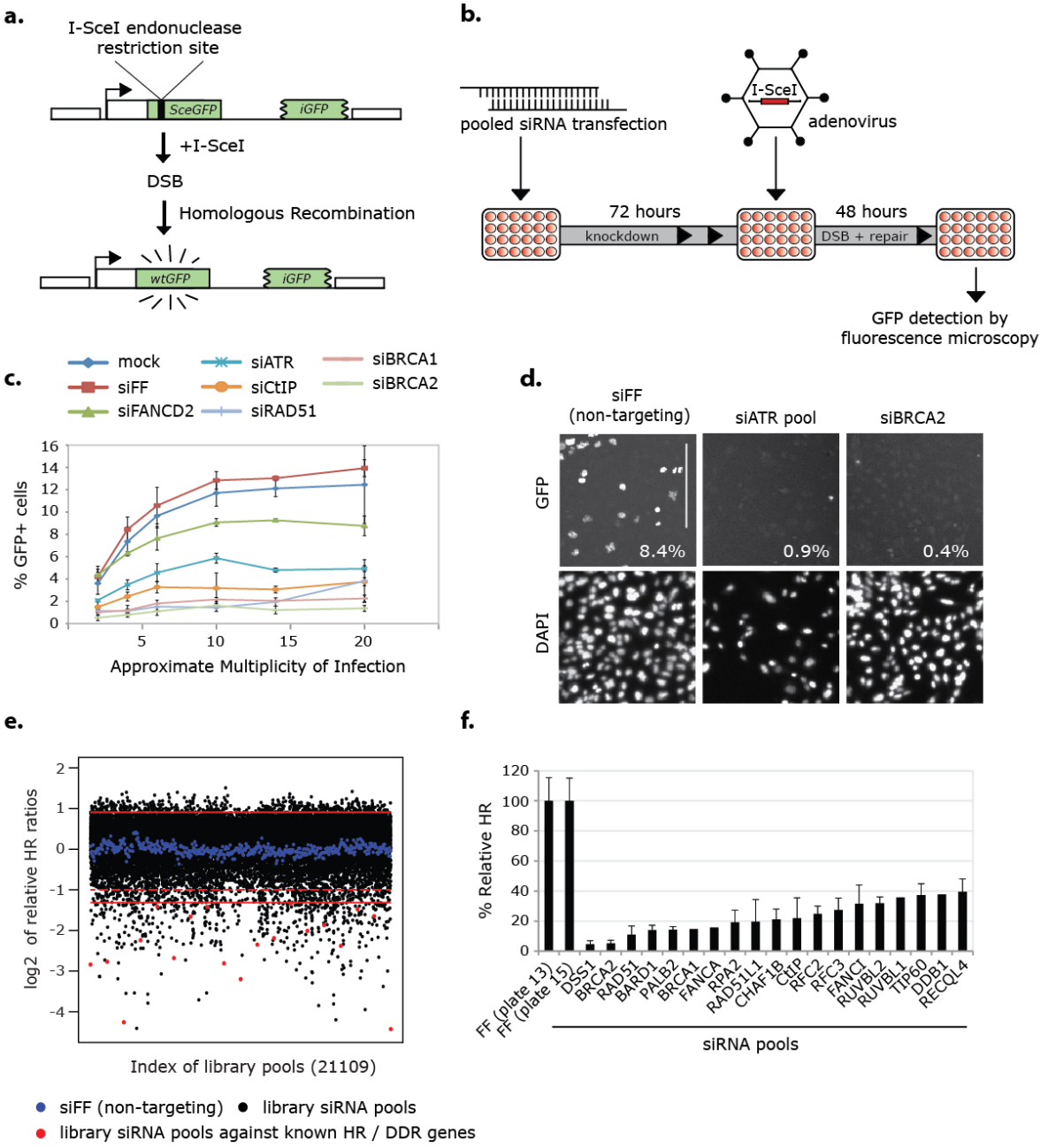
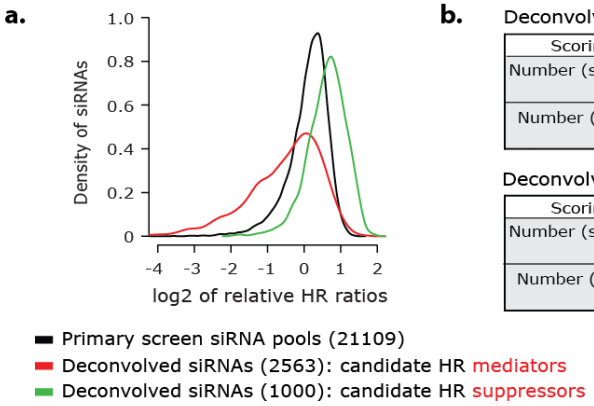


Figure 7. A genome-wide siRNA screen for mammalian homologous recombination (HR) genes. (a) Schematic of DR-GFP construct. **(b)** Schematic of the high-throughput (HTP) HR screen. Arrayed pools of siRNAs were reverse transfected into DR-U2OS cells in 384 well plates. Cells were infected with the I-SceI expressing adenovirus AdNGUS24i at an MOI of ~10 after 72 hours and 48 hours later were fixed and imaged. **(c)** DR-GFP cells transfected with the indicated siRNA pools in HTP were infected with AdNGUS24i at the estimated MOIs, and % GFP+ cells were determined by high-throughput imaging. Assay results plateaued at an MOI of ~10. Error bars represent ± s.d. across three replicates. **(d)** DR-U2OS cells were transfected with positive (siATR pool and siBRCA2) and negative (siFF) control siRNAs and assayed for HR in high throughput. Images were taken on the automated screening platform and are presented in representative portions. % GFP+ cells –as calculated from full images– are included. Scale bar represents 300 μm. **(e)** Relative HR ratios (presented as log2 values) from 22109 library siRNA pools and 476 negative controls (siFF). Solid red lines indicate 2 s.d. from the screen-wide mean of relative HR ratios (presented as log2 values). These were used as cutoff values to determine pools scoring with increased or decreased HR. The region between the dotted red line and the lower solid red line indicates a scoring range from which most additional candidate HR mediators were selected. **(f)** Known regulators of HR or the DDR that scored under the 2 s.d.-based cutoff.

We screened the Dharmacon human siGENOME siRNA library in triplicate, which is arrayed as 21,121 single-target pools of 4 distinct siRNAs. Briefly, DR-U2OS cells were plated in 384 well plates, reverse transfected with siRNAs, and infected with AdNGUS24i at a multiplicity of infection (MOI) of ~10 (Figure 7b), an MOI that at least by visual inspection was not observed to substantially affect cell survival. At this high titer changes in cell number caused by siRNA transfection should have little effect on assay results (Figure 7c). Cells were fixed, stained with Hoechst, and the percentage of GFP+ cells per well was determined by fluorescence microscopy on an automated platform (Figure 7d). The average percentage of GFP+ cells from each experimental triplicate was normalized to that from on-plate, non-targeting control wells transfected with an siRNA against firefly luciferase (siFF) to obtain a relative HR ratio for each library pool (Figure 7e, Table S1).

2-2. Identification and validation of candidate HR regulators

Hits from the screen were defined as siRNA pools that decreased or increased relative HR >2 standard deviations (s.d.) from the screen-wide mean (cutoff values ~ 40% or 188% relative HR). From these, 510 candidate HR mediators and 484 candidate HR suppressors were identified (Table S1). Indicative of a successful screen, we recovered 19 genes known to be involved in HR and the DDR, including *RAD51*, *BRCA1* and *BRCA2* (Figure 7e,f). We extended the list of candidate mediators by 131 genes corresponding to siRNA pools that trended in the screen (primarily with 40-50% relative HR) (Tables S1-2). These additional genes had been identified in previous DDR screens. Next we deconvolved the 641 siRNA pools against candidate mediators and the strongest 250 pools against candidate suppressors (including 1 duplicate pool) and rescreened each siRNA individually (Tables S2-3). As expected, siRNAs from both candidate sets enriched for the appropriate phenotype (Figure 8a).

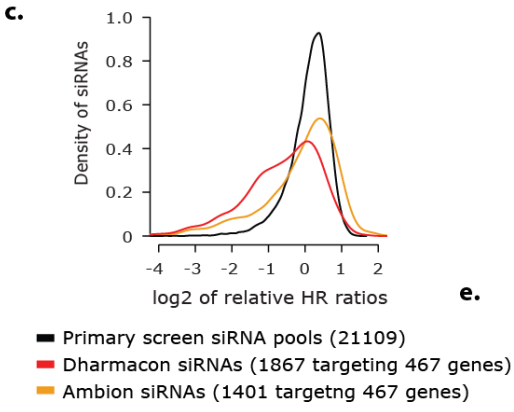


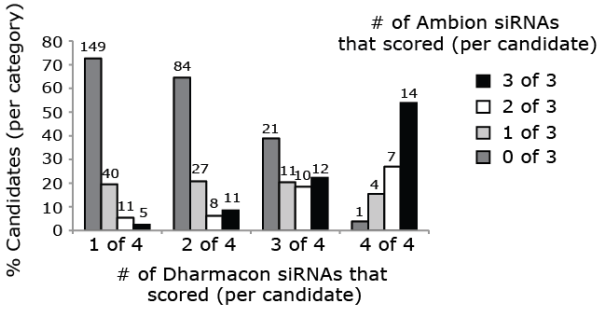
b. Deconvolved siRNAs against candidate HR mediators (641 pools)

Scoring siRNAs	4 of 4	3 of 4	2 of 4	1 of 4
Number (strong cutoff)	15	24	60	258
% of Total	2%	4%	9%	40%
Number (weak cutoff)	29	56	153	284
% of Total	5%	9%	24%	44%

Deconvolved siRNAs against candidate HR suppressors (250 pools)

Scoring siRNAs	4 of 4	3 of 4	2 of 4	1 of 4
Number (strong cutoff)	4	17	67	105
% of Total	2%	7%	27%	42%
Number (weak cutoff)	6	46	82	88
% of Total	2%	18%	33%	35%





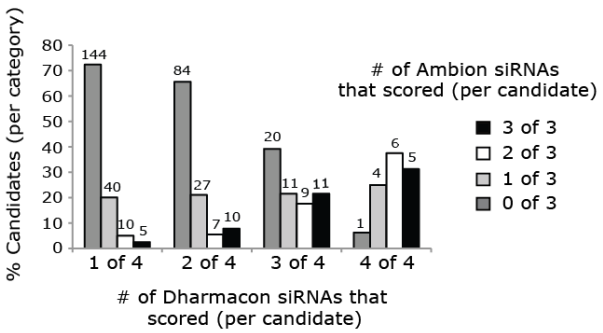


Figure 8. Rescreen and validation of candidate HR genes. **(a)** Distributions of log2 relative HR ratios for three sets of screened siRNAs: primary screen pools (black), deconvolved Dharmacon siRNAs against candidate HR mediators (red), and deconvolved Dharmacon siRNAs against candidate HR suppressors (green). **(b)** The number and percentage of Dharmacon siRNA pools that rescored with 1, 2, 3 or 4 siRNAs after deconvolution. Results analyzed using strong (2 s.d. from the screen-wide mean-based: 40% and 188% relative HR) or weak (1.5 s.d.-based: 59% and 169%) cutoff values. **(c)** Distributions of log2 relative HR ratios for primary screen pools (black: as presented in a), individual Dharmacon siRNAs against 467 candidate mediators (red), and individual Ambion siRNAs against the same candidate mediators (orange). **(d)** Percentage of candidate HR mediator genes that scored with 0, 1, 2, or 3 Ambion siRNAs in categories of those that scored with 1, 2, 3, or 4 Dharmacon siRNAs. Weak cutoff was used for scoring. Number of candidates indicated. **(e)** Same as (d) after 22 known HR and DDR genes were excluded from analysis.

We evaluated the rescreened siRNAs using both strong and weak phenotype cutoffs (Figure 8b). Strong siRNAs were those that rescored below (for mediator siRNAs) or above (for suppressor siRNAs) the 2 s.d.-based thresholds from the primary screen (40% and 188% relative HR, respectively). Weak cutoffs were based on 1.5 s.d. from the primary screen mean (<59% or >169% relative HR for mediators and suppressors, respectively). We considered candidate siRNA pools validated if ≥3 individual siRNAs (out of 4) rescored with at least a weak HR value (14% of pools for HR mediators and 20% of pools for HR suppressors) (Figure 8b). The higher validation rate for siRNA pools targeting HR suppressors is likely a result of rescreening only the strongest 250 pools. Additionally, the 510 candidate mediators chosen according to the screen cutoff yielded a higher validation rate than the 131 select candidates added to the mediator list, with validation rates of 15% and 8% respectively.

To further evaluate the candidate mediator list and better understand the 68% of pools that rescored with 1-2 siRNAs, we conducted a second round of rescreening using siRNAs from the Ambion Silencer Select library targeting 467 candidate mediators (3 siRNAs / gene) (Table S4). These siRNAs also enriched for reduced HR but at a level substantially less than observed among Dharmacon siRNAs targeting the same 467 genes (Figure 8c). We reason that the independently selected Ambion siRNAs had a reduced incidence of off-target effects (discussed in Chapter 3) and were, therefore, more likely to score true positives. Importantly, candidates that validated with 3-4 (of 4) Dharmacon siRNAs had greater likelihood of scoring with 2-3 (of 3) Ambion siRNAs (over candidates that scored with fewer Dharmacon siRNAs), even when known HR and DDR mediators were not considered (Figure 8d,e). After eliminating reagents predicted to be false-positives through analysis discussed in Chapter 3, we refined our list of validated candidate HR mediators to 121 that scored with at least 3 of 7 combined Ambion and Dharmacon siRNAs (Table 1).

Table 1: Candidate HR mediators validated with at least 3 of 7 siRNAs

Entrez Gene Symbol	Entrez Gene ID	Entrez Gene Symbol	Entrez Gene ID	Entrez Gene Symbol	Entrez Gene ID
BRCA1	672	PPARBP	5469	FZD9	8326
BRCA2	675	RAD51	5888	GNG5	2787
PALB2	79728	RFC1	5981	GPR35	2859
PHF5A	84844	RFC3	5983	GSTM5	2949
RBBP8	5932	RPA2	6118	HLA-DQA2	3118
RFC2	5982	RUVBL2	10856	HTATIP	10524
SF3A2	8175	SF3A1	10291	IKZF4	64375
SF3B1	23451	SF3B14	51639	KIAA0947	23379
ATR	545	SF3B2	10992	KRT81	3887
BARD1	580	SF3B3	23450	LSM2	57819
CDC73	79577	SNRPC	6631	MAB21L2	10586
HNRPC	3183	SNRPD2	6633	MDS032	55850
MFAP1	4236	SUPT6H	6830	MED19	219541
POLR2E	5434	USPL1	10208	MED31	51003
POLR2G	5436	ZNF207	7756	MED9	55090
PRPF6	24148	CDC40	51362	MFGE8	4240
RBMX	27316	CSNK1A1	1452	NUDT11	55190
RFC4	5984	DHX8	1659	PABPN1	8106
SF3A3	10946	FANCA	2175	PCDHGA1	56114
SF3B4	10262	FUBP1	8880	PFAS	5198
SHFM1	7979	GPR112	139378	PITPNM3	83394
SNRPA1	6627	LSM4	25804	PLCL2	23228
SNRPB	6628	LSM6	11157	POLR2H	5437
SON	6651	MAEA	10296	POLR2I	5438
UBL5	59286	PGD	5226	PSMA5	5686
ASCC3L1	23020	PIK3R2	5296	PSMB4	5692
CRNKL1	51340	POLR2F	5435	PSMD14	10213
CRSP2	9282	PRPF31	26121	RANBP5	3843
CRSP6	9440	PRPF8	10594	RAP1B	5908
CRSP8	9442	PSMB7	5695	RHOA	387
CUL1	8454	RBM25	58517	RPTN	126638
INTS2	57508	SURB7	9412	RUVBL1	8607
KIAA1604	57703	WBP11	51729	SALL1	6299
MED10	84246	XAB2	56949	SART1	9092
MED4	29079	C12orf32	83695	SCPEP1	59342
MED6	10001	C4orf21	55345	SIRPB1	10326
MED8	112950	CCL3	6348	SMR3B	10879
PABPC1	26986	CD274	29126	SNW1	22938
PAF1	54623	CELSR3	1951	SOS1	6654
POLR2B	5431	FLJ32549	144577	SRRM2	23524
				ST6GALNAC2	10610

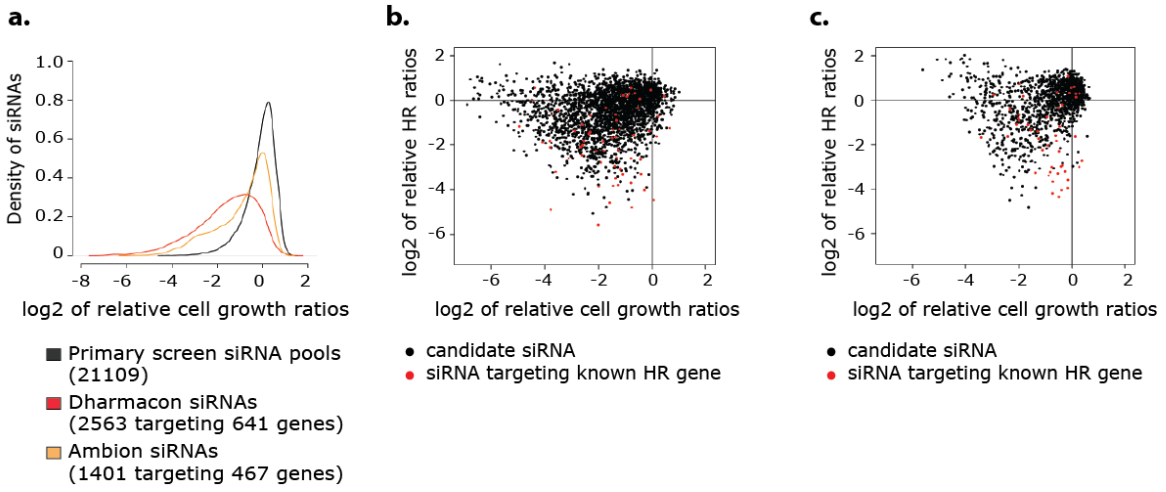


Figure 9. Cell growth effects of siRNAs during HTP screening. (a) Distributions of log2 relative cell growth ratios for primary screen pools (black), deconvolved Dharmacon siRNAs against candidate HR mediators (red), individual Ambion siRNAs against candidate HR mediators (orange). **(b)** Scatter plot comparing log2 relative HR and log2 relative cell growth ratios for 2563 Dharmacon siRNAs against 641 candidate HR mediators. No direct correlation between the phenotypes is observed. siRNAs indicated in red target BRCA2, CtIP, DSS1, PALB2, BARD1, RPA2, ATR, BRCA1, RAD51, RUVBL2, TIP60, FANCA, RUVBL1, FANCI, RAD51L1, and ATM. **(c)** Same as (d) but for 1401 Ambion siRNAs against 467 candidate HR mediators.

We also evaluated siRNA toxicity throughout and observed no correlation between cell growth and relative HR levels (Figure 9a-c, Tables S1-4). However, depletion of many HR proteins is toxic to cells, and we observed that siRNAs against candidate mediators from our screen were enriched for those that adversely affected cell growth. Of 519 siRNA pools that scored for decreased HR in the primary screen, 14% resulted in final cell numbers that were less than 25% of control cells, compared to 2% of pools from the rest of the screen, which is a 7-fold enrichment for decreased cell growth (Fisher's Exact Test p-value = 1.30×10^{-41}). Toxicity was, likewise, prevalent in both rescreening analyses of candidate HR mediators (Figure 9a-c). Importantly though, we observed no correlation between relative cell growth and relative HR even among siRNAs targeting known HR proteins.

2-3. Network analysis

Next we evaluated candidate genes for enrichment of functional categories and interaction networks using Ingenuity Pathway Analysis (IPA, Ingenuity Systems, www.ingenuity.com) and Gene Ontology terms (GO terms) assigned by DAVID Bioinformatics Resources 6.7 (david.abcc.ncifcrf.gov) (see Chapter 2, 4-3 for a description of the gene lists submitted to these analyses and references). Both mediators and suppressors were enriched for genes functionally categorized as DNA replication, recombination, and repair by IPA which we expected for mediators of HR, but not necessarily for suppressors as little is known about what activities limit recombination (Figure 10a,b). Among candidate mediators, two gene networks with known HR genes were identified (Figure 10c,d). Interestingly, these suggest a role for factors associated with DDB1 and the CUL4A ubiquitin ligase in HR and highlight roles for the RFC DNA clamp loader and the TIP60 acetylase complex. Overall, nine components of the TIP60 complex scored or trended in the primary screen: TIP60, RUVBL1, RUVBL2, DMAP1, Brd8, p400, ING3, MRGBP and MRG15 (Table S1), which is perhaps not surprising as the acetyltransferase activity of TIP60 is known to play a role in both ATM activation (Sun et al., 2005) and histone remodeling at DSBs through γH2AX acetylation (van Attikum and Gasser, 2009).

As mentioned, candidate mediators were also evaluated using an analysis of GO terms. Specifically, we evaluated candidates for enrichment of genes annotated with GO terms indicative of functional roles in homologous recombination (HR), double-strand break repair (DSBR), checkpoint regulation, and the DNA-damage response (DDR). HR, DSBR and DDR genes, but not checkpoint genes, were significantly enriched among 433 candidate mediators selected by the primary screen (and returned by DAVID with unique identifiers) (Fisher's Exact Test *p*-values = 0.001, 0.006, 5.67×10^{-5}, and 0.058 respectively). As indicated by IPA analysis, several components of RFC (RFC1-3) scored in the primary screen; however, we note here that many other proteins involved in checkpoint activation, including the 9-1-1 complex, RAD17 and TopBP1 did not. Although concrete conclusions cannot be made regarding what does not score in RNAi screens because siRNA reagents have inherently incomplete sensitivities (discussed in Chapter 3, 1-1), we note that defective HR in the absence of RFC proteins may

represent dysfunctional DNA replication and not an effect on HR. Nevertheless, the recently characterized 9-1-1- and TopBP1-interacting checkpoint protein RHINO (C12ORF32) was identified by our screen (Table 1) (Cotta-Ramusino et al., 2011), perhaps suggesting that some checkpoint regulators have more substantial roles in HR than others.

Of 96 library genes annotated as functionally relevant to HR processes by DAVID (using GO:0000724~double-strand break repair via homologous recombination, GO:0000725~recombinational repair, and GO:0006310~DNA recombination), 9.4% scored among candidate mediators identified in the primary screen (a 15-fold enrichment). A discussion of known HR genes that did and did not score in the primary screen can be found in Chapter 2, Section III.

By IPA the most significantly enriched category among candidate mediators was RNA post-transcriptional modification, which also produced a strong interaction network (Figure 10a,e). Although the involvement of RNA-processing proteins in the DDR is poorly understood, several large-scale genetic and proteomic analyses of the DDR have shown similar enrichments (Hurov et al., 2010; Matsuoka et al., 2007; Paulsen et al., 2009). A role in promoting HR could explain the enrichment observed in all three screens. Among HR suppressors a small network containing phosphatases emerged, and it is possible that these act to limit the activity of HR-promoting kinases (Figure 10f).

IPA also identified categories of candidate genes that may relate to the design of the screen but not HR. The DR-GFP based HR assay depends on infection of an adenovirus, expression of I-SceI and GFP, and normal cell cycle progression; and it is, therefore less likely that candidates functionally categorized under infection mechanism, gene expression, or cell cycle regulation represent biological true positives (Figure 10a,b). Additionally, a strong network of transcriptional proteins was identified among candidate HR mediators (Figure 10g).

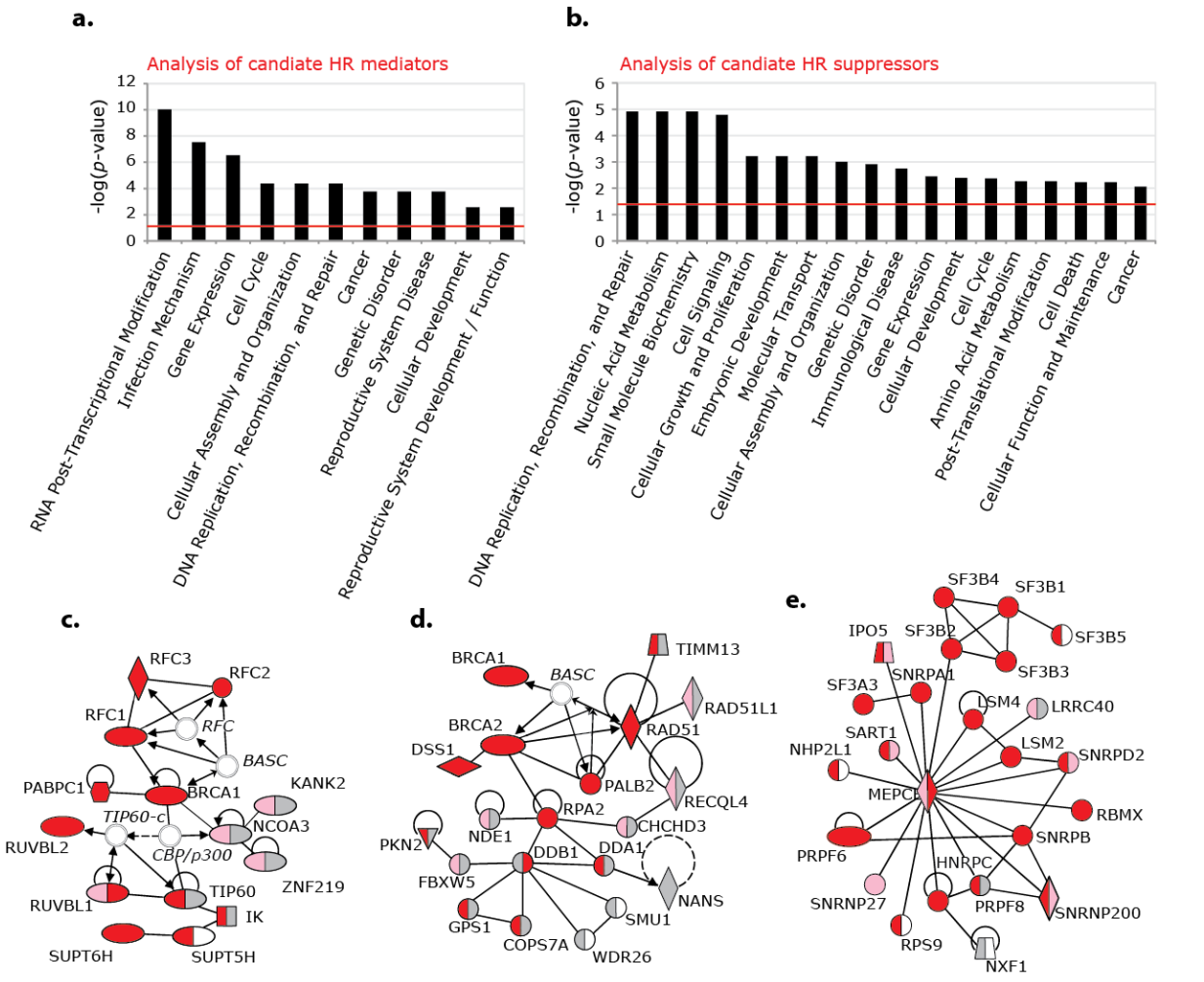
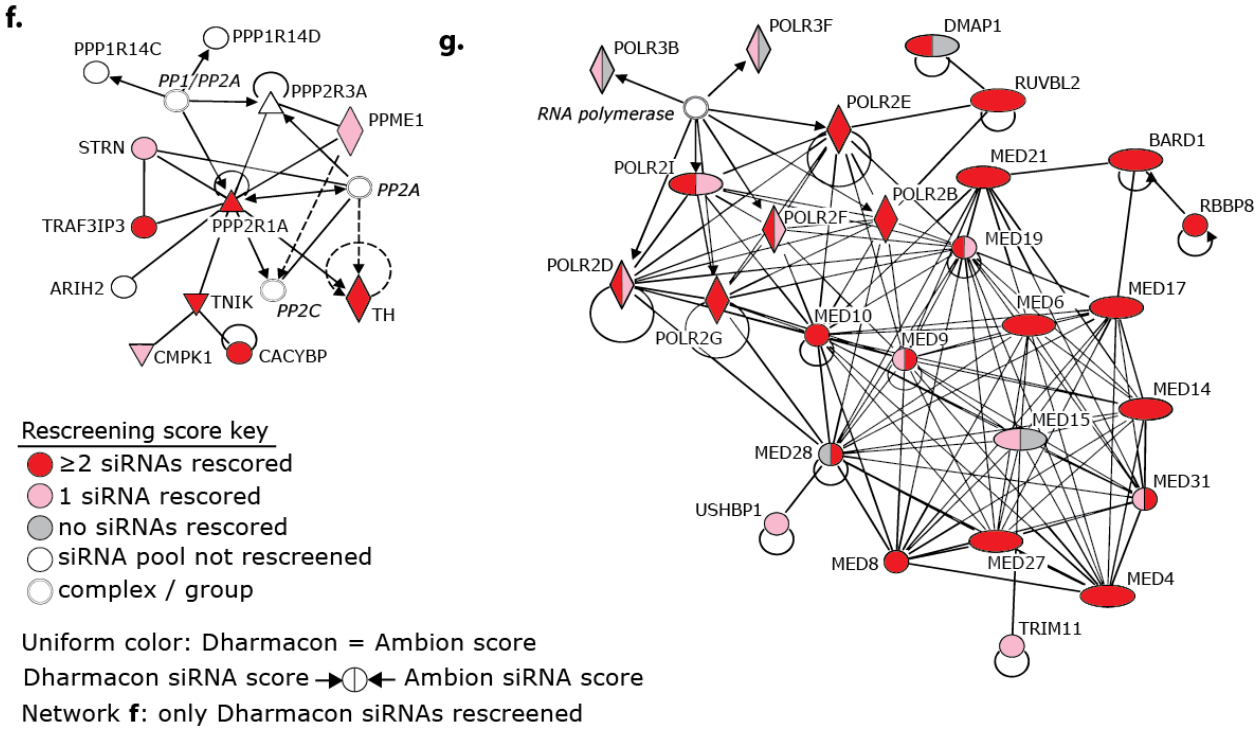


Figure 10. Functional categories and interaction networks among HR candidates identified by the primary screen and generated using Ingenuity Pathway Analysis (IPA). (a) Select functional gene categories enriched among candidate HR mediators. Red line represents a threshold of 1.3 = -log(p-value = 0.05). **(b)** Same as (a) for candidate HR suppressors. **(c-g)** Interaction networks: **(c)** DDR network including components of the TIP60 complex and the RFC DNA clamp loader from candidate mediators, **(d)** DDR and HR network including canonical HR proteins and CUL4A ubiquitin ligase-associated proteins from candidate mediators, **(e)** pre-mRNA processing network from candidate mediators, **(f)** network containing phosphatases from candidate suppressors, and **(g)** transcription factor network from candidate mediators. Dharmacon and Ambion rescreen data are indicated by color, except in (f) where no Ambion data are represented. Color key: red indicates a candidate that rescored with ≥2 siRNAs (out of 4 Dharmacon and 3 Ambion), pink indicates that 1 siRNA rescored, gray indicates that 0 siRNAs rescored, and white indicates that a candidate was not rescreened. Line key: solid lines indicate direct interactions, dashed lines indicate indirect actions, arrows indicate the direction of interactions, and lines without arrowheads indicate binding. Shape key: ovals indicate transcription regulators, hexagons indicate translational regulators, diamonds (vertical) indicate enzymes, diamonds (horizontal) indicate peptidases, trapezoids indicate transporters, triangles indicate phosphatases, inverted triangles indicate kinases, squares indicate cytokines, circles indicate other, double circles indicate groups of proteins or complexes.

Figure 10 (continued)



III. Discussion

Here we describe an unbiased genetic screen to identify regulators of the mammalian HR machinery that produced two candidate sets (510 positive and 484 negative regulators), and we present these lists as a well-curated resource. Importantly, of 510 candidate mediators, we validated 121 with ≥3 individual siRNAs, compiling a highly validated list of known and as yet uncharacterized HR proteins.

Among the candidate mediators here identified are several known HR genes, including CtIP, BRCA2, PALB2, DSS1, BRCA1, BARD1, ATR, RPA2, RAD51 and TIP60, as well as additional components of the TIP60 complex. Interestingly, these are all known to promote HR upstream of or coincident with formation of the RAD51 presynaptic filament, suggesting a bias for presynaptic regulators over downstream effectors among candidates identified by our screen. Perhaps surprisingly, then, we also observe that none of the resection enzymes (MRE11, BLM, DNA2, EXO1) known to

promote presynaptic resection in mammals came through the screen. Although (as mentioned above) no substantial conclusions can be made with regard to what does not score in RNAi-based screens, we note that both resection and postsynaptic HR pathways are highly redundant (discussed in Chapter 1). From this we conclude that the DR-GFP assay may lack the sensitivity required to identify novel effectors of redundant HR mechanisms using strategies based on single-target, genome-wide RNAi knockdown. In support of this, we observed that several resection, resolution, and dissolution effectors, including DNA2, NBS1, TOP3, EME1, GEN1 and SLX4, trended in the primary screen with minimal HR defects (relative HR 50-70%) but do not score.

Because several *bona fide* mediators of HR have previously been characterized with very weak HR defects using the DR-GFP assay (Smogorzewska et al., 2007; Svendsen et al., 2009), we also conclude that siRNA pools yielding intermediary phenotypes in our primary screen may represent mild but biologically significant effects. We have not as yet undertaken the task of systematically evaluating these candidates because the phenotypic range between 50-70% encompasses 1,678 siRNA pools, making their individual prioritization both costly and work-intensive. Nevertheless, novel HR components may be identified within this large candidate set in the future, and improvements to our ability to perform gene interaction studies in mammalian cells may, in particular, present a useful approach for such work. A format applicable to shRNA-based HR screening that could be used towards screening this set with higher efficiency is presented in Chapter 5, Section III.

During the course of our work, a similar genetic screen (also using the DR-GFP system in mammalian cells to evaluated HR) was published (Slabicki et al., 2010). Different from our approach, the authors of this study used an arrayed library of esiRNA reagents. These are single-target siRNA pools generated from the nucleolytic cleavage of long dsRNA molecules. This library contained 16,707 esiRNAs, all targeting genes annotated by unique Ensembl identifiers. To compare the results of this study with our work, genes targeted by the Dharamcon siRNA library (and annotated by unique Entrez gene IDs) were re-annotated with Ensembl identifiers (Flicek et al., 2013) using the Bioconductor AnnotationDbi package (Gentleman et al., 2004; Herve Pages, 2012) in R (R Core Team, 2012). After

manually correcting for duplicate Ensembl gene IDs, 17,754 genes remained; however, of these only 14,452 were among those also evaluated by Buchholz and colleagues. This gene set includes 41 of the 44 validated candidates reported by Slabicki et al. to decrease HR and 106 of those on our validated mediator list. A comparison of these yields seven common genes, a small but significant overlap (Fisher's Exact Test p-value = 1.71×10^{-8}). These are 4 well-characterized HR proteins (BRCA1, RAD51, CtIP, DSS1), 2 mRNA-processing related proteins (CWC22, SNRNP200) and the proteasomal subunit PSMD14, which has previously been implicated in HR (Jacquemont and Taniguchi, 2007). The presence of *CWC22* and *SNRNP200* on this list supports our finding that RNA-processing proteins have some functional role in promoting HR, which we address more thoroughly in Chapter 4.

Several factors could explain the limited candidate overlap observed between these published screens. First, annotation of the human genome is by no means complete and different candidate sets are expected from RNAi libraries that do not target identical gene sets, as is the case for those discussed here. Moreover, computational comparison of human gene identifiers from different annotation sources remains a challenge and may affect our ability to faithfully compare candidate sets, as illustrated by the 18 genes identified as HR mediators that could not be cross examined. Methodological differences between screens could also have influenced results. The most obvious of these are the use of different cell lines (HeLa / U2OS) and the method / timing of I-SceI delivery during the DR-GFP assay. In our work, cells were transfected with siRNAs 72 hours before I-SceI transduction; however, Slabicki et al., transfected I-SceI into cells coincident with esiRNAs and fixed cells for analysis 72 hours later. This difference may have affected the degree of protein depletion achieved before DSB formation in the DR-GFP assay and, in this manner, may also have affected the lists of candidates identified.

Lastly, it is important to consider that false positives and false negatives caused by the limitations of RNAi technologies may have impacted the reproducibility of candidates between these screens. In particular, a distinct set of false positives caused by different sets of off-target effects in each library may have influenced results. RNAi off-target effects will be discussed further in Chapters 3 and 5. Use of high reagent redundancy in our work likely limited the inclusion false positives on our validated

candidate list; however, the library used by Slabicki et al. for primary screening contained only one esiRNA reagent per target gene and validation of their candidate set was likewise conducted with only a single, independent (and non-overlapping) reagent per gene. Because only two reagents were evaluated per target gene in the this study, both off-target effects and reagent failure may have strongly influenced its results. Of note, the limited overlap discussed here is not unique and other similar but independent RNAi-based screens have also yielded few common candidates (Mohr et al., 2010; Sigoillot and King, 2011).

Interestingly, of the 34 validated candidates from the esiRNA screen that were absent from our validated set (which could be evaluate by our data), 8 (24%) scored in our primary screen using a weak cutoff value (59% relative HR). Overall, this cutoff scores approximately 7.3% of evaluated targets with unique identifiers from our screen. Of the 99 validated candidates from our screen that were absent from the candidate list presented by Slabicki et al. (and that could be evaluated by their work), 23 (23%) were among the 7.3% of targets with the lowest HR score in their primary screen. The trending of these validate candidates suggests that many may represent true positives.

Overall, we expect that the candidate HR mediators and suppressors presented herein will aid future characterization of HR. Among the uncharacterized genes on our mediators list is *C4OFR21*, which has been shown to influence crosslink repair in a genome-wide screen for sensitivity to mitomycin C; intriguingly, the encoded protein carries a predicted DNA-binding domain and helicase-like region (Smogorzewska et al., 2010). Additionally, as it is known that NHEJ proteins suppress HR (Pierce et al., 2001), we also expect that our suppressor list will yield positive regulators of NHEJ. OTUB1, a deubiquitinating enzyme that inhibits HR (Nakada et al., 2010), and RAP80, whose HR suppressive effects are discussed in Chapter 1, 3-3-8, are tellingly present on this list. Lastly, we note that the discovery of novel HR suppressor genes will be of particular interest for genetic engineering studies that aim to improve gene targeting technology in mammalian cells.

IV. Materials and Methods

4-1. Cell culture

DR-U2OS cells (Xia et al., 2006) provided by Maria Jasin (Memorial Sloan-Kettering Cancer Center) were grown in McCoy's 5A media supplemented with 10% fetal bovine serum (FBS) for HTP screening.

4-2. High-throughput (HTP) screening

The primary screen was performed using 21,121 siRNA pools from the Dharmacon human siGENOME siRNA library (G-005000-05) at 50 nM. Dharmacon and Ambion rescreens were conducted at 20 nM. DR-U2OS cells were plated on 384 well plates at 700 cells / well and reverse transfected with siRNAs using Oligofectamine^{TM} Transfection Reagent. Positive (siATR, siBRCA2) and negative (siFF) controls were added to each plate. After 72 hours, cells were infected with the I-SceI carrying adenovirus AdNGUS24i (provided by Frank Graham, McMaster University) at an estimated MOI of 10; 48 hours after infection, cells were fixed in 3.7% formaldehyde and stained with Hoechst 33342 at a dilution of 1:5000 (Invitrogen). Changes were made to this protocol for screening Ambion siRNAs: (1) 500 DR-U2OS cells / well were plated (to adjust for reduced toxicity of Ambion siRNAs observed in controls), (2) the Lipofectamine^{TM} RNAiMAX Transfection Reagent was used, and (3) AdNGUS24i was used at an MOI of ~15 (to adjust for differences between viral preparations).

For data collection automated imaging of screen plates (2-4 images per well in 2 channels) was conducted on an Image Express Micro microscope (Molecular Devices) at 4X magnification (488 nm and 350 nm wavelengths were used to detect GFP expression and Hoechst 33342 stained DNA, respectively). Automated counting of GFP+ and Hoechst stained nuclei was performed for each image with Metamorph Cell Scoring software (Molecular Devices Inc.) and a ratio of GFP+ to Hoechst stained (total) nuclei was calculated for each well using all corresponding images. Primary screen pools, Ambion siRNAs and deconvolved Dharmacon siRNAs against candidate suppressors were evaluated in triplicate, while individual Dharmacon siRNAs against candidate mediators were evaluated in duplicate. To normalize day-to-day variability, each triplicate (or duplicate) average of % GFP+ was normalized to the average %

GFP+ from on-plate negative control wells (siFF); standard deviation was calculated and propagated for each. These normalized values are presented as relative HR ratios. Images from the primary screen and candidate mediator Dharmacon rescreen that produced relative HR ratios with high standard deviations were selectively visually inspected, and data from images that were found to contain an irregularity (for example, were out of focus) were deleted from analysis.

A cell number / well value was also calculated for each HTP well as the sum of Hoechst 33342 stained nuclei from all corresponding images, and a relative cell growth ratio for each siRNA pool was calculated by normalizing the average cell number / well of corresponding experimental wells to that of on-plate negative control wells (siFF). The number of images taken of experimental and corresponding control wells was kept the same.

4-3. Candidate selection

Primary screen data was collected and processed as described above (Chapter 2, 4-2). All relative HR ratios from the primary screen were compiled, and from this, 519 pools that decreased relative HR >2 s.d. from the screen-wide mean and 486 pools that increased relative HR >2 s.d. (including 2 duplicate pools against SMAD1 and TIAM2) were identified (screen mean = 1.14, s.d. = 0.37) (Table S1). siRNA pools that were unavailable for validation, corresponded to discontinued gene entries in the NCBI gene database (www.ncbi.nlm.nih.gov), or (as discussed above: Chapter 2, 4-2) were determined by visual analysis to be based on poor quality imaging were eliminated. The 510 and a subset of the strongest 486 siRNA pools remaining (against 510 and 484 HR mediator and suppressor candidates, respectively) were deconvolved into individual duplexes and rescreened (Tables S2-3). For this 131 pools (against 131 genes) were added to the 510 candidate pools against HR mediators (for 641 candidates, 2564 siRNAs total). Selection of the 131 additional candidates is described in Chapter 2, 2-2). Only the 250 pools (including 1 duplicate pool against SMAD1) that most strongly increased HR were rescreened. siRNAs from the Ambion Silencer Select library targeting 467 (of 641) candidate HR mediators were also screened (3 siRNAs / gene, 1401 siRNAs total). Selection of these 467 candidates was based on data

from our primary screen and Dharmacon rescreen analysis, as well as data from related DNA damage screens, GESS analysis and published information about each gene.

Ingenuity Pathway Analysis (IPA) and Gene Ontology term (Ashburner et al., 2000) analysis was conducted on the candidate HR mediator and / or suppressor sets generated by applying a 2 s.d. cutoff to the primary screen data and prior to the expansion / editing of the candidate lists outlined above (519 mediators and 484 suppressors) (Table S1). Both candidate sets were uploaded into IPA software and scanned for functional enrichment and interaction networks based on information in the Ingenuity® Knowledge Base. Select functional enrichment categories from IPA are displayed in Figure 10a,b. Enrichment *p*-values were determined using the Fisher's Exact Test. Select protein networks from IPA are displayed in Figure 10c-g. Network nodes are colored according to the number of siRNAs that scored using a weak cutoff value based on 1.5 s.d. from the primary screen mean. These networks and functional analyses were generated through the use of IPA (Ingenuity Systems, www.ingenuity.com).

Gene Ontology terms descriptive of biological processes (GOTERM_BP) were obtained for genes targeted by the Dharamcon library by uploading all Entrez Gene IDs listed in Table S1 into the DAVID Bioinformatics Resources 6.7 Analysis Wizard (Huang da et al., 2009a, b). The 15,747 genes subsequently analyzed were those returned with unique identifiers. Functional gene sublists were generated from these 15,747 genes using GO terms indicative of roles in homologous recombination (HR), double-strand break repair (DSBR), checkpoint regulation, and the DNA-damage response (DDR): HR (GO:0000724~double-strand break repair via homologous recombination, GO:0000725~recombinational repair, GO:0006310~DNA recombination), DSBR (GO:0006302~double-strand break repair, GO:0000729~DNA double-strand break processing), checkpoint regulation (GO:0000077~DNA damage checkpoint, GO:0031570~DNA integrity checkpoint, GO:0031572~G2/M transition DNA damage checkpoint, GO:0031576~G2/M transition checkpoint, GO:0031573~intra-S DNA damage checkpoint), DDR (GO:0006281~DNA repair, GO:0006282~regulation of DNA repair, GO:0006974~response to DNA damage stimulus, GO:0030330~DNA damage response signal transduction by p53 class mediator, GO:0042770~DNA damage response signal transduction,

GO:0042772~DNA damage response signal transduction resulting in transcription, GO:0045739~positive regulation of DNA repair, GO:0008630~DNA damage response). These sublists were then compare to GO term annotation of candidate mediators identified by the primary screen (described above) and present among the genes returned by DAVID with unique identifiers.

Chapter Three

RAD51 is a predominantly off-targeted transcript in mammalian cells

Britt Adamson[1], Agata Smogorzewska[1,2], Frederic D. Sigoillot[3], Randall W. King[3], Stephen J. Elledge[1]

[1]Division of Genetics, Brigham and Women's Hospital, Department of Genetics, Harvard University Medical School, Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA

[2]Laboratory of Genome Maintenance, The Rockefeller University, New York, New York 10065, USA

[3]Department of Cell Biology, Harvard University Medical School, Boston, Massachusetts 02115, USA

Work discussed herein can be found in the following publications:

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol *14*, 318-328.

Sigoillot, F.D., Lyman, S., Huckins, J.F., Adamson, B., Chung, E., Quattrochi, B., and King, R.W. (2011). A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. Nat Methods *9*, 363-366.

GESS analysis discussed herein was designed and conducted by Frederic D. Sigoillot and Randall W. King. All experimental work was conducted by Britt Adamson. Agata Smogorzewska contributed to the initial discovery of RBMX accumulation at sites of DNA damage. We thank Dr. Peter Adams for HIRA and UBN1 antibodies.

CHAPTER THREE

I. Introduction

In general, co-opting the endogenous RNAi machinery to modulate gene expression and evaluate genetic phenotypes presents unique challenges, and in the context of RNAi-based screening, the presence of reagent-specific caveats particular to 20,000-80,000 reagents (or more) may substantially hinder the enrichment of biological true positives among screen candidates (Sigoillot and King, 2011). Following a brief discussion of the potential pitfalls relevant to large-scale RNAi technology and the manner in which these relate to endogenous mechanisms, I will discuss how *RAD51* depletion was identified as a prominent off-target effect in our screen and used towards the systematic elimination of false positives.

1-1. Common caveats to genetic studies using RNAi

Perhaps the most important caveats to posttranscriptional genetics using RNAi is that protein depletion using this technology, as opposed to genetic deletion, generates hypomorphic phenotypes, and variable sensitivities among RNAi reagents with the same intended target often cause inherently distinct phenotypic results. Although this can be beneficial for identifying functionalities within the context of toxic or otherwise lethal depletion events or for evaluating phenotypic ranges associated with titrated protein expression (akin to allelic series), it is entirely problematic with regard to evaluating large RNAi data sets because without direct measurement no *a priori* predications can be made as to which (or how many) RNAi reagents effectively deplete their intended target. Moreover, methods currently available for determining the depletion efficacies of specific reagent-target pairs are not scalable for high-throughput applications. So the common inclusion of ineffectual RNAi reagents in genome-wide screens ensures high false-negative rates. As with most caveats to RNAi (discussed herein), a standard solution to this problem is the use of high reagent redundancy, which was applied to our screen during candidate validation through the use of two rounds of rescreening (Chapter 2, 2-2). At a minimum, this approach

biases large-scale RNAi experiments towards the inclusion of multiple functional reagents per individual gene.

An additional concern pertaining to RNAi-based experiments is the possibility of unintentionally perturbing endogenous cellular mechanisms through *in vivo* exposure to exogenous RNAs (Sigoillot and King, 2011). In particular, short dsRNAs may, in select cases, non-specifically activate an interferon response present in mammalian cells that, as part of an ancient antiviral mechanism, responds to long dsRNA molecules of foreign origin. Alternatively, RNAi reagents may simply flood the endogenous RNAi machinery causing dysregulation of endogenous miRNAs (discussed below) and unrelated changes to gene expression. The inclusion of appropriate non-targeting and mock controls in RNAi-based experiments, however, should effectively manage these sequence-independent effects in most cases.

Lastly, the suboptimal specificity of shRNA and siRNA reagents must be considered. Although RNAi reagents are generally designed with perfect sequence complementarities to their intended mRNA targets, RNAi mechanisms in general allow for promiscuous and degenerate sequence-based targeting, and clear evidence of broad off-target gene silencing has been shown. Using mircoarray analysis, individual siRNAs at both high (100 nM) and low (4 nM) reagent concentrations have been observed to affect hundreds of unintended transcripts (Jackson et al., 2003), and wide-spread off-targeting among RNAi libraries has been estimated to substantially increase the false-positives rates of many RNAi-based screens (Sigoillot and King, 2011). To fully evaluate these off-target effects, however, we must first examine the targeting mechanisms used by the endogenous miRNA machinery.

1-2. Endogenous mircoRNAs

Common genome-encoded effectors of endogenous RNAi are ~21-23-nt microRNAs (miRNAs). These are differentially expressed within various cell and tissue types to mediate broad posttranscriptional gene regulation through the modulation of mRNA stability and translation (Carthew and Sontheimer, 2009). Mature miRNAs are processed from longer pri- and pre-miRNA molecules. In mammals, this occurs as follows: First, miRNA-encoding genes are transcribed into pri-miRNAs by RNA polymerase II

(RNAPII). pri-miRNAs contain secondary structures with one or more RNA hairpins that can then be cut into stem-loop structures (pre-miRNAs) by the endoribonuclease activity of DROSHA within the nucleus. pre-miRNAs are exported into the cytoplasm by exportin 5 where they are processed into open-ended dsRNA molecules (mature miRNAs) by the endoribonuclease DICER. One strand of a given miRNA, referred to as the "guide" strand, can then be incorporated into RISC; after which, the RISC subunit AGO2 mediates pairing to a complementary target mRNA and facilitates subsequent gene silencing through direct translational repression or AGO2-induced mRNA cleavage followed by exonucleolytic degradation. Determination of the guide strand is controlled by the thermodynamic properties of a given miRNA duplex and follows the general, but not absolute rule, that the guide strand has lower affinity for its complementary RNA strand at the 5' end. Rules that promote translational repression over mRNA cleavage, or *vice versa*, remain unclear; although, they appear to be guided in part by the degree of miRNA to mRNA pairing, with near perfect complementarity promoting cleavage.

1-3. Sequence-based targeting of the endogenous RNAi machinery

In animals, the target sites of endogenous miRNAs are predominantly found within the 3' untranslated regions (UTRs) of mRNAs, and to date, six modes of sequence-directed miRNA-mRNA pairing –all with target predictive capacities– have been identified (Bartel, 2009). These are defined as cleavage (or near perfect) pairing, centered pairing (Shin et al., 2010), and four modes of seed pairing: seed (Lim et al., 2005), seed with G-bulges (Chi et al., 2012), seed with 3'-supplementary, and seed with 3'compensatory pairing (Bartel, 2009). The miRNA seed region spans nucleotides 2-8 on the 5' end of the miRNA guide strand and facilitates the smallest and most predominant of known pairing modes. This minimal 7-nt requirement for gene silencing is thought to allow broad modulation of gene expression through the repression of hundreds of mRNA targets by single miRNAs (Lim et al., 2005), a concept that has recently been supported by the finding that miRNA-induced silencing can also occur at target sites with mRNA G-bulges at seed positions 5 and 6 (Chi et al., 2012). Pairing of the 3' miRNA end (centered on nucleotides 13-17) in combination with mismatch-containing seeds (compensatory pairing) further

expands the degenerate target repertoire (Bartel, 2009). On the other hand, perfect 6-8-nt seed matches with uninterrupted 3-4-nt supplemental pairing at miRNA 3' ends (nt 13-16) show increased efficacy of miRNA targeting in rare cases (Bartel, 2009), and atypical centered pairing of 11 contiguous nucleotides (4-14, 5-15) has been shown to target miRNAs without minimal 6-mer seeds (Shin et al., 2010). The increased informational complexity of centered pairing is thought necessary to overcome a conformational preference for 5' end pairing dictated by RISC (Shin et al., 2010); however, this no doubt also limits the number of messages targeted by this pairing mode.

1-4. Predictable sequence-based off-targeting

Unlike miRNA-target pairing, only two sequence-based modes of off-target pairing by exogenous RNAi reagents have been identified with predictive value: one through near perfect pairing to unintended transcripts and the other by 6-7-nt seed sequence matches to unintended 3'UTRs, so-called "miRNA-like" off-targeting (Birmingham et al., 2006). Initially, seed-based off-targeting was identified only for individual siRNAs (Birmingham et al., 2006; Jackson et al., 2003); however, these effects have increasingly been identified within genome-wide data sets (Adamson et al., 2012; Schultz et al., 2011; Sigoillot et al., 2011; Sudbery et al., 2010). Here we present the identification of a single prominently off-targeted transcript (*RAD51*) from our genome-wide screen using a recently developed computational algorithm (Sigoillot et al., 2011), and by eliminating screen reagents predicted to deplete *RAD51* through "mRNA-like" off-target effects, we enrich our candidate list for true positives. This work helps to set a new standard for data analysis from RNAi-based screens.

II. Results

2-1. Three candidate HR regulators localize to sites of DNA damage: HIRIP3, RBMX, DDX17

To begin characterizing candidates identified by our screen, we made GFP fusions of 22 candidate proteins and evaluated each for relocalization after DNA damage. Two candidate mediators, HIRIP3 and RBMX, and one suppressor, DDX17, accumulated at regions of DNA damaged by

microirradiation (Figure 11a). RBMX is an RNA-binding protein that associates with the spliceosome and plays a role in alternative splicing (Heinrich et al., 2009), and DDX17 is a DEAD-box RNA helicase that is phosphorylated in response to IR (Janknecht, 2010; Matsuoka et al., 2007). HIRIP3 is a poorly characterized histone chaperone that binds histones H2B and H3 and interacts with the histone chaperone HIRA (Lorain et al., 1998). Interestingly, HIRA and two additional HIRA-associated proteins (UBN1 and CAIN) also localized to DNA damage after microirradiation (Figure 11b).

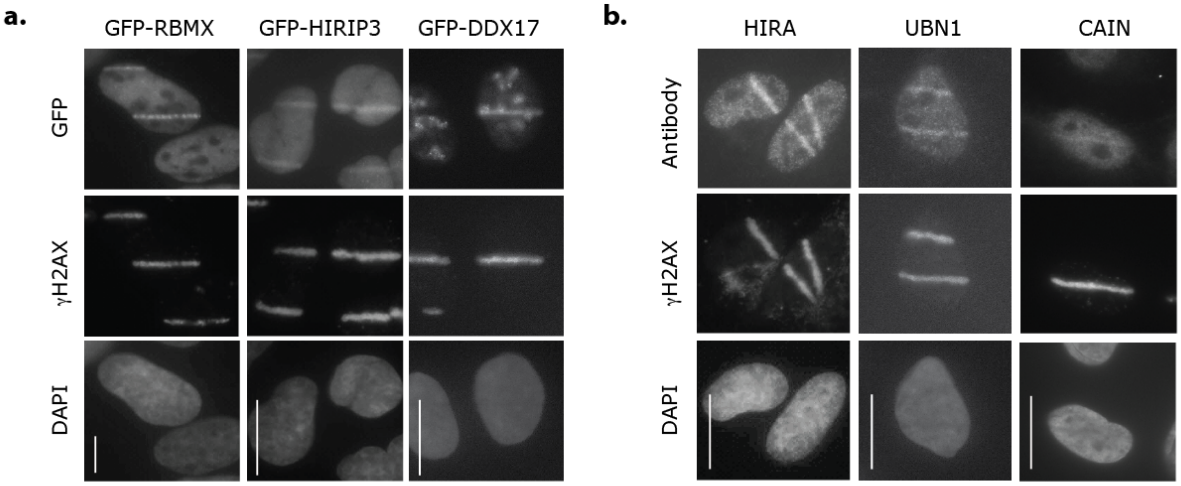


Figure 11. Screen candidates localize to sites of DNA damage. (a) Cells expressing GFP fusions of RBMX, HIRIP3 or DDX17 were microirradiated and prepared for immunofluorescence after 0-5 (RBMX and DDX17) and 15-20 minutes (HIRIP3) with an antibody against γH2AX. GFP was observed directly. Nuclei were stained with DAPI. Scale bars indicate 10 μm. Images were prepared from three separate experiments and are not intended for comparison. **(b)** Cells were microirradiated and prepared for immunofluorescence with antibodies against HIRA, UBN1 or CAIN and γH2AX within 30 minutes. Nuclei were stained with DAPI. Scale bars indicate 10 μ m. Images were prepared from three separate experiments and are not intended for comparison.

The HIRIP3-targeting siRNA pool in our primary screen gave a strong HR defect, but while seven siRNAs and three shRNAs were individually shown to deplete HIRIP3, only five of these caused substantial HR defects (Figure 12a-d), and none of three screened HIRIP3 Ambion siRNAs scored (Table S4). Indicative of an off-target effect, expression of siRNA-resistant HIRIP3 did not rescue the HR defect caused by siHIRIP3-2 (data not shown). The HIRA siRNA pool from our primary screen trended for decreased HR, but also indicative of a false positive, these siRNAs caused a range of HR defects that did not correlate with mRNA depletion once deconvolved (Figure 12g,h). Similarly, several UBN1- and

CAIN-targeting siRNAs caused HR defects likely to be associated with off-targeting (Figure 12f,g,i,j). Because of these data we suspected that HR might be particularly sensitive to off-target effects.

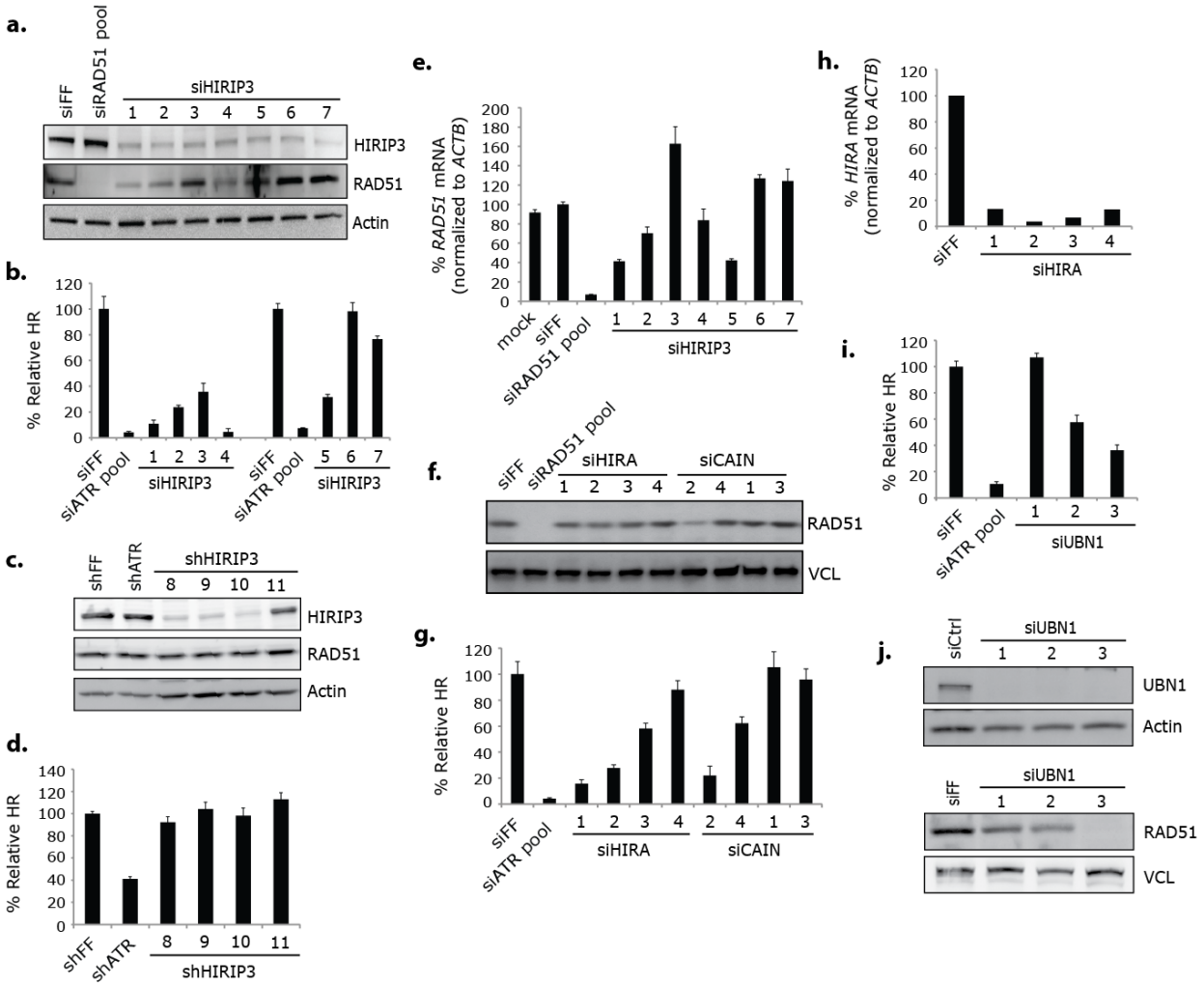


Figure 12. The DR-GFP HR assay is sensitive to off-target effects. **(a)** HIRIP3 / RAD51 western blot analysis prepared from U2OS cells transfected with the indicated siRNAs; an antibody against Actin was used as a loading control. **(b)** HR assay results from DR-U2OS cells transfected with the indicated siRNAs. Two separate experiments are represented. Experimental data is grouped with corresponding controls. Error bars represent ± s.d. across three replicates. **(c)** HIRIP3 / RAD51 western blot analysis prepared from DR-U2OS cells transduced with the indicated shRNAs; an antibody against Actin was used as a loading control. **(d)** HR assay results from cells in (c). Error bars represent ± s.d. across three replicates. **(e)** RT-qPCR of RAD51 (normalized to ACTB) from DR-U2OS cells transfected with indicated siRNAs. Primers used against RAD51 recognize four transcript variants. Error bars represent ± s.e.m. across three replicates. **(f)** Whole-cell extracts from U2OS cells transfected with the indicated siRNAs were immunoblotted with the indicated antibodies. Analysis shows siCAIN-2 off-target depletion of Rad51. **(g)** HR assay results from DR-U2OS cells transfected with the indicated siRNAs. A range of HR phenotypes is observed. Error bars represent ± s.d. across three replicates. **(h)** RT-qPCR of HIRA (normalized to ACTB) from DR-U2OS cells transfected with the indicated siRNAs. Data represent the mean of two replicates. **(i)** HR assay results from DR-U2OS cells transfected with the indicated siRNAs. Error bars represent ± s.d. across three replicates. **(j)** Whole-cell extracts from U2OS cells transfected with the indicated siRNAs were immunoblotted with the indicated antibodies. siCtrl does not target UBN1 and was used as a negative control. Two western blots of the same extracts are presented (panels are grouped accordingly).

2-2. Off-target *RAD51* depletion contributes to screen false positives

To search for off-target effects in our screen data, we used Genome-wide Enrichment of Seed Sequences (GESS) analysis, an algorithm that identifies 3'UTRs with enriched sequence complementary to the seed regions of siRNAs that score in genome-wide screens (over non-scoring siRNAs) (Sigoillot et al., 2011). As previously discussed (Chapter 3, 1-3), siRNA seed regions are nucleotides 2-8 of single siRNA strands incorporated into RISC, which according to siRNA design should be the antisense (guide) strands in most cases but could be either the sense or antisense (Sigoillot and King, 2011). GESS analysis capitalizes on the fact that siRNA seed complementarity to particular 3'UTRs can elicit repression of the associated transcripts through a pathway endogenously engaged by microRNA-containing RISC complexes (Bartel, 2009; Sigoillot et al., 2011).

GESS analysis of our Dharmacon rescreen data revealed that siRNAs against candidate mediators with strong HR defects (40% relative HR cutoff) were 3-fold enriched for 7-nt antisense seed sequences matches to the 3'UTR of *RAD51*, compared to those that did not rescore with a strong phenotype (an increase from 8% to 25%, Fisher's Exact Test *p*-value = 4.65×10^{-23}) (Figure 13a). The sense strands of strongly scoring Dharmacon siRNAs, however, gave no enrichment for seed matches to the *RAD51* 3'UTR (Fisher's Exact Test *p*-value = 0.5986), and no enrichment for seed region complementarities (from both strands) to the *RAD51* coding region (CDS) was observed (Fisher's Exact Test *p*-value = 0.8886) (Figure 13b). From this analysis, we predict a 17% false positive rate due to off-target *RAD51* depletion among our candidate Dharmacon siRNAs. Importantly, strong scoring Ambion siRNAs targeting candidate mediators were not enriched for antisense (or sense) seed complimentarily to the *RAD51* 3'UTR (Fisher's Exact Test *p*-value = 0.3526 antisense and 0.7485 sense), suggesting that this off-target effect did not confound the data from those reagents (Figure 13c). Of note, GESS analysis of Dharmacon siRNAs that scored for increased HR identified small but significant enrichments of seed matches to three transcript 3'UTRs: *ITGB1BP3*, *FAM153C* and *EDC3* (Figure 13d). These have not previously been implicated in HR.

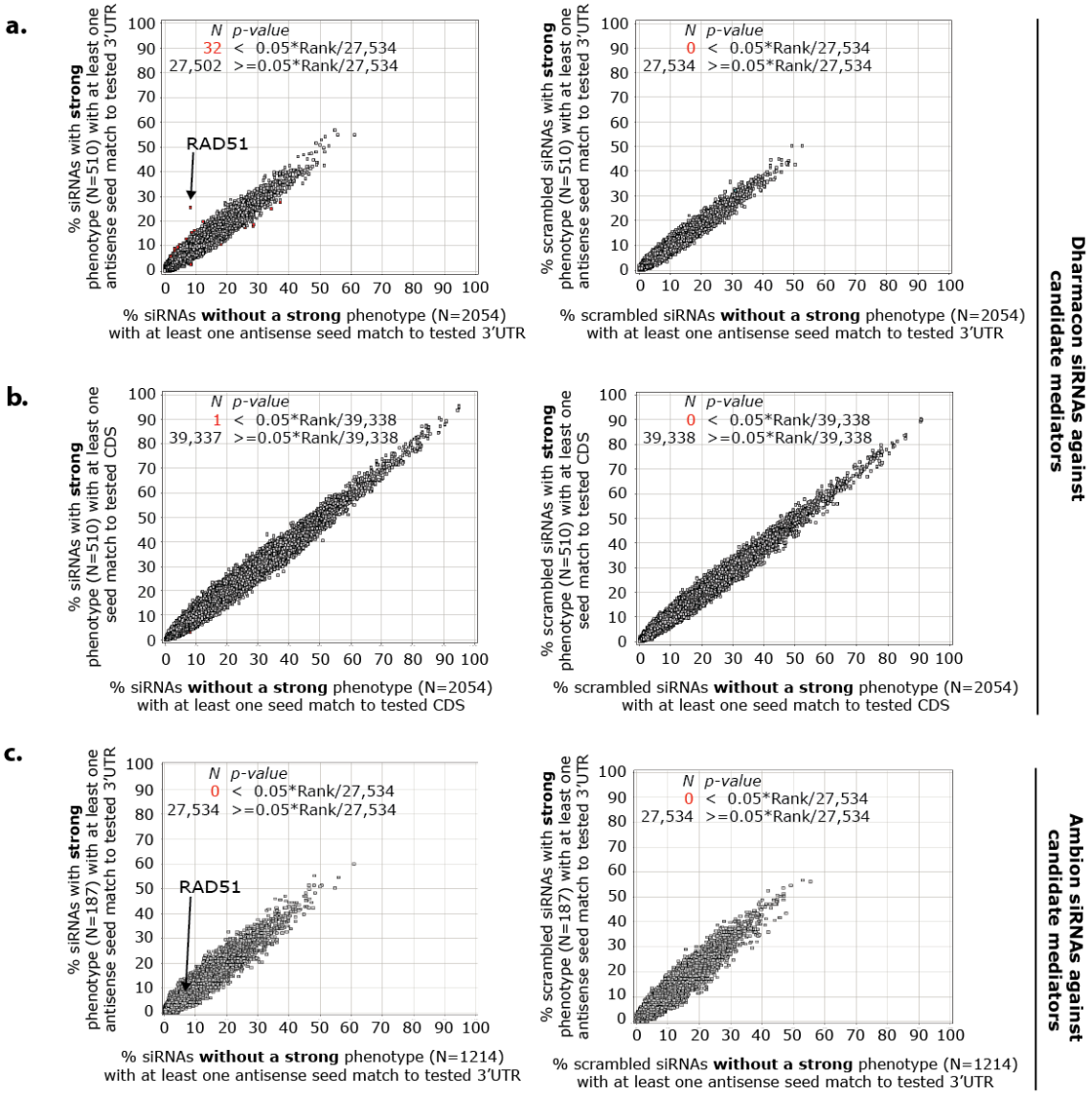
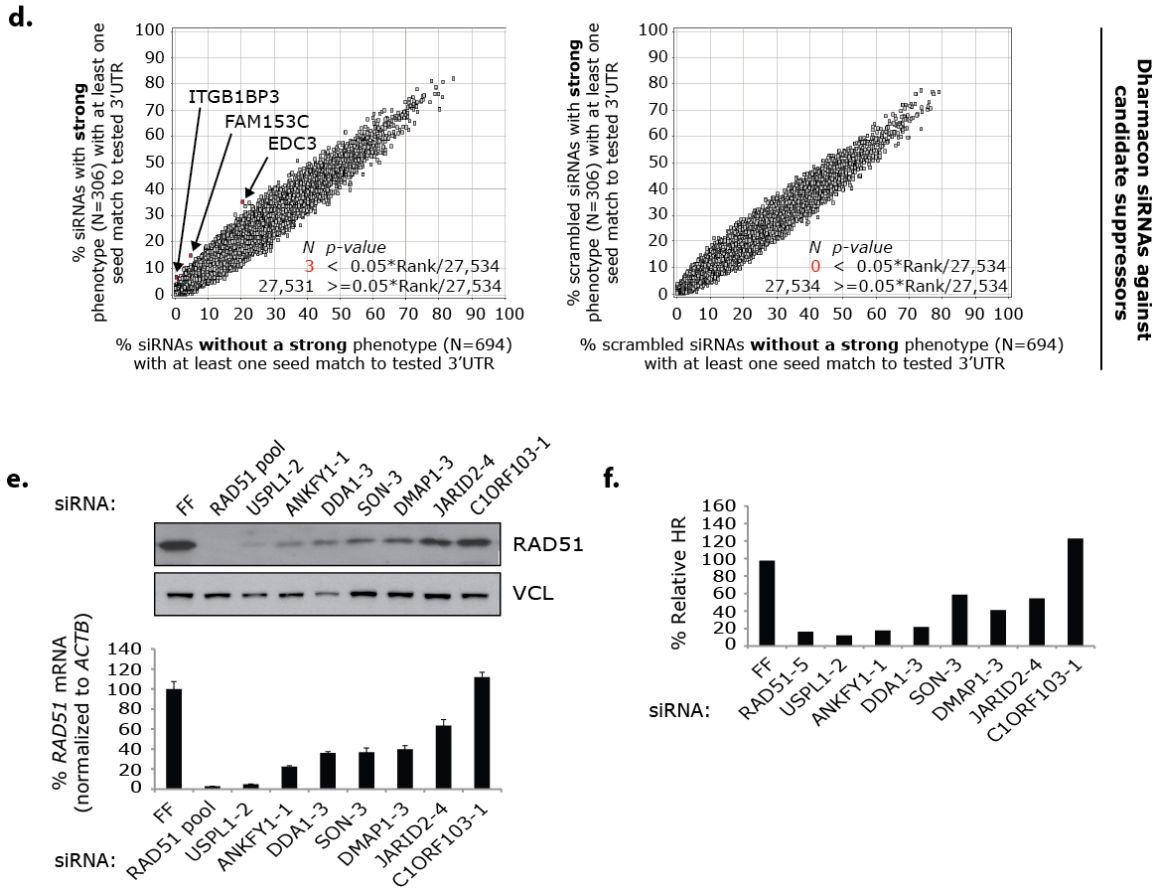


Figure 13. Off-target *RAD51* depletion was a major source of false positives among Dharmacon siRNAs. (a) GESS analysis of Dharmacon siRNAs rescreened against candidate mediator genes. Scatter plots represent the percentage of siRNAs in 2 groups that have at least one 7-nt antisense seed sequence match to 27,534 human 3'UTRs. Left plot: compares siRNAs that individually rescored with a strong phenotype (y-axis) to those that did not (x-axis). Right plot: compares the same 2 groups of siRNAs after scrambling the seed sequences and serves as a control. 3'UTRs that significantly enriched for seed matches are red. Arrow illustrates the 3'UTR of *RAD51*. **(b)** GESS analysis of Dharmacon siRNAs rescreened against candidate mediator genes as described in (a) using criteria of at least one 7-nt seed sequence match (either guide or passenger strand) to 39,338 human CDSs (coding sequences). One CDS significantly enriched for seed matches and is indicated in red. **(c)** GESS analysis of Ambion siRNAs rescreened against candidate mediator genes as described in (a) using criteria of at least one 7-nt antisense seed sequence match to 27,534 human 3'UTRs. No 3'UTR significantly enriched. **(d)** GESS analysis of Dharmacn siRNAs rescreened against candidate suppressor genes as described in (a) using criteria of at least one 7-nt seed sequence match (either strand) to 27,534 human 3'UTRs. Three 3'UTRs (labeled) significantly enriched for seed matches and are indicated in red. **(e)** RAD51 western blot analysis and RT-qPCR of *RAD51* (normalized to *ACTB*) prepared from U2OS cells transfected with seven Dharmacon siRNAs predicted to off-target RAD51 by a 7-nt antisense seed region match to the *RAD51* 3'UTR; an antibody against VCL was used as a loading control. Primers used against *RAD51* recognize four transcript variants. Error bars in RT-qPCR analysis represent ± s.e.m. across three replicates. **(f)** HR data from HTP Dharmacon rescreening analysis in DR-U2OS cells. Data represent the average of two replicates. Negative control (siFF) taken form the same rescreen plate as siDMAP1-3 and siC1ORF103-1.

Figure 13 (continued)



Predicted off-target RAD51 depletion was confirmed for 6 screened Dharmacon siRNAs with 7-nt antisense seed region matches to the 3'UTR of *RAD51* (of 7 tested), and both *RAD51* mRNA and protein depletion by these siRNAs correlated with the relative HR measurements determined for each during rescreening analysis (Figure 13e-f). Perhaps not surprisingly, we found that the HR defects caused by four HIRIP3-targeting siRNAs (including three from the primary screen pool), siUBN1-2, siUBN1-3, and siCAIN-2 also correlated with off-target RAD51 depletion (Figure 12a,b,e-g,i,j). To remove this strong off-target effect from confounding our screen results, we refined our list of validated candidate HR mediators to 121 that scored with at least 3 of 7 combined Ambion and Dharmacon siRNAs after eliminating Dharmacon siRNAs predicted to deplete RAD51 by a 7-nt antisense seed region match to the *RAD51* 3'UTR (Table 1).

III. Discussion

3-1. Off-target *RAD51* depletion

Using GESS analysis, we determined that *RAD51* is a prominently off-targeted transcript among Dharmacon siRNAs selected by our primary screen, and we found that this phenomenon could primarily be attributed to sequence complementarities between the antisense seed regions of these siRNA and the *RAD51* 3'UTR. Consistent with this, we experimentally confirmed off-target depletion of *RAD51* mRNA and protein for 6 Dharmacon siRNAs with antisense seed complementarily (of 7 tested) (Fig. 3e). We found no enrichment for seed matches among the sense strands of siRNAs from the screen however, suggesting that the Dharmacon siGENOME reagents we used successfully controlled siRNA strand loading into RISC using thermal asymmetry rules and limited additional off-target effects to some extent. However, of the seven HIRIP3-, UBN1-, and CAIN-targeting siRNAs found to cause off-target depletion of RAD51 in subsequent analyses, only siHIRIP3-5 had a 7-nt seed match to the *RAD51* 3'UTR, which indicates that *RAD51* off-targeting can occur without complete seed complementarity and suggests that the incidence of *RAD51* off-target effects in our screen is underrepresented by the GESS-derived estimation. Therefore additional mechanisms of *RAD51* off-targeting may also confound experimental results, perhaps mediated through pairing modes defined by incomplete seeds with G-bulges or aided by 3'compensatory sites.

We note here that GESS analysis could not be applied to data from the esiRNA-based HR screen published by Buchholz and colleagues and discussed in the discussion section of Chapter 2 because candidates from that work were not validated with RNAi reagents carrying individual target sequences (Slabicki et al., 2010); however, we estimate that on average approximately 1 of 16 siRNAs will have a 7-nt seed match to a 1 kilobase region of DNA (approximately the size of the *RAD51* 3'UTR), so multiple seed matches to *RAD51* are expected to be present in each of their esiRNA pools. In work presented elsewhere, our screen and subsequent validation data was used to demonstrate that GESS-derived information could aid the removal of off-target RNAi reagents from screen results (Sigoillot et al., 2011). Specifically, it was shown that by discarding GESS-predicted off-target effects the percentage of

candidates scoring with ≥2 Dharmacon siRNAs that validated with 2-3 Ambion siRNAs increased from 36% to 51% (using a strong cutoff threshold) (Sigoillot and King, 2011). Filtering our candidate set in this manner, however, did not remove any candidates that strongly validated with 2-3 Ambion siRNAs indicating that the initial selection of these candidates was not substantially influenced by off-target effects. This highlights the general importance of validating RNAi-based screens with independent reagents, even when predominantly off-targeted transcripts are not (or cannot be) identified. Overall these data suggest a new standard for validating candidates from RNAi-based screens.

Hypotheses regarding the cause of *RAD51* hypersensitivity to RNAi off-target effects and a discussion of other prominently off-targeted transcripts that have been identified, including the spindle assembly checkpoint mediator *MAD2* mRNA (Sigoillot et al., 2011), are included in Chapter 5.

3-2. HIRA-associated proteins and DNA repair

Although we showed that individual depletion of HIRIP3, HIRA and UBN1 does not result in defective HR, this investigation lead to the observation that these proteins accumulate at sites of DNA damage and therefore indicates that they may have functional roles in DNA repair. HIRA is a multifunctional histone chaperone that deposits the histone variant H3.3 into chromatin during DNA synthesis-independent nucleosome assembly (Ray-Gallet et al., 2002; Tagami et al., 2004), and in cooperation with UBN1 and ASF1A, HIRA coordinates the formation of senescence-associated heterochromatin foci (Banumathy et al., 2009; Zhang et al., 2005) thought to repress proliferation-promoting genes. HIRA has two orthologs in *S. cerevisiae*, Hir1 and Hir2, which are part of the transcriptionally repressive HIR complex that also controls histone deposition into chromatin (Eitoku et al., 2008; Prochasson et al., 2005). Similarly, UBN1 and CAIN are orthologs of HIR complex subunits (Hpc2 and Hir3) (Balaji et al., 2009; Banumathy et al., 2009), and because these can be biochemically copurified with HIRA, a conserved complex is thought to exist in mammalian cells (Tagami et al., 2004). HIRIP3 was first identified in humans as a HIRA-interacting protein (Lorain et al., 1998), and like HIRA, it is a histone chaperone. HIRIP3 binds core histones H2B and H3 (Lorain et al., 1998) and is the

predicted human ortholog of the *S .cerevisiae* protein Chz1, which preferentially deposits the histone variant H2AZ into chromatin (Luk et al., 2007).

Because DNA repair is conducted within the context of nuclear chromatin, much thought has been given to the manner in which repair effectors gain access to regions of broken DNA and how chromatin structure is restored after the completion of repair. Homologous recombination-associated chromatin modulation, which is no doubt required to shift or remove DSB-flanking nucleosomes during DNA end resection, is of particular interest (Ransom et al., 2010; Smerdon, 1991). Several chromatin remodelers, including the mammalian TIP60 complex (SWR1 and INO80 in yeast) as well as the recently characterized SWI/SNF-related SMARCAD1 protein (Costelloe et al., 2012), and several histone chaperones have already been implicated in processes associated with DNA repair. The FACT complex, for example, has been shown to facilitate exchange of γH2AX/H2B dimers for H2A/H2B in chromatin surrounding DNA breaks (Heo et al., 2008), and the ASF1 and CAF1 histone H3/H4 chaperones, although not required for repair *per se*, are thought to mediate DNA-synthesis dependent nucleosome assembly in regions of newly repaired DNA (Polo et al., 2006; Ransom et al., 2010).

Exactly what HIRA, a histone chaperone involved in DNA synthesis-independent nucleosome assembly, and its associated proteins are doing at DNA lesions associated with active DNA synthesis is unclear. One possibility is that they function as histone acceptors, binding and sequestering histones that have been newly freed from repair-associated chromatin; another is that, like ASF1 / CAF1, they mediate chromatin reconstruction after the completion of DNA repair. Neither of these scenarios would *a priori* require protein function for repair, but would instead require only the coordination of chromatin dynamics that occur downstream of DNA repair. Notably, defects in such processes would not necessarily be expected to generate phenotypes also indicative of defective repair. Another possibility is that HIRA and its associated proteins function only within specific DNA contexts that may not be amenable to repair by HR, at particular DNA lesion types for example or within regions of specific chromatin structure such as heterochromatin. Nevertheless, the possibility remains that these proteins will yet be found to positively regulate the repair of DSBs through pathways other than HR, and intriguingly, the HIRIP3 ortholog Chz1,

together with Swr1, has been observed to modulate levels of the histone variant H2AZ in chromatin surrounding DSBs in yeast (Kalocsay et al., 2009; Ransom et al., 2010; van Attikum and Gasser, 2009).

IV. Materials and Methods

4-1. Cell culture

Human U2OS and DR-U2OS osteosarcoma cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 units / ml of penicillin, and 0.1 mg / ml streptomycin (Invitrogen). Cell selection after viral transduction was conducted with puromycin at 1 ug / ml.

4-2. Plasmids, shRNAs, and siRNAs

RBMX and HIRIP3 cDNAs were from hORFeome V5.1. DDX17 cDNA was obtained in pOTB7 from Open Biosystems (Item MSH1011-59342). DDX17 was isolated by PCR and cloned into the pENTR™ / D-TOPO vector using the pENTR™ Directional TOPO® Cloning Kit (Invitrogen). Full-length RBMX, HIRIP3 and the 5' end of DDX17 were verified by sequencing. All three cDNAs were cloned into pMSCV-N-EGFP-GAW-PGK-PURO for expression using the Gateway recombination system. shRNAs were used in the pSMP-MSCV-PURO vector (Open Biosystems). Throughout all data chapters, siRNAs were transfected into cells at 20-50 nM using either Oligofectamine™ or Lipofectamine™ RNAiMAX Transfection Reagents (Invitrogen) according to manufacturer instructions and cells were processed for indicated experiments 2-3 days later. shRNA and siRNA sequences used in this chapter that are not listed in Tables S2-4 are listed in Table 2. Multiple negative (siFF) and positive control siRNAs were used throughout all experimental chapters (from Dharmacon, Invitrogen and Ambion) and are also listed in Table 2; the particular controls used in each experiment were for the most part matched to the vendor of the experimental reagents.

Table 2. RNAi Reagents

Target	Reagent Type	Source	Clone ID / Catalog Number	Sequence
FF	siRNA	Dharmacon	custom	CGUACGCGGAAUACUUCGAUU
FF	siRNA	Invitrogen	custom	GGAUUUCGAGUCGUCUUAAUGUAUA GAGGACCUAUGAUUAUGUCCGGUUA
FF	siRNA	Ambion	custom	CGUACGCGGAAUACUUCGAtt
RAD51	siRNA pool	Dharmacon	D-003530-02 D-003530-05 D-003530-07 D-003530-08	GAAGCUAUGUUCGCCAUUA GCAGUGAUGUCCUGGAUAA CCAACGAUGUGAAGAAAUU AAGCUAUGUUCGCCAUUAA
ATR	siRNA pool	Dharmacon	D-003202-05 D-003202-17 D-003202-31 D-003202-32	GAACAACACUGCUGGUUUG GCAACUCGCCUAACAGAUA UCUCAGAAGUCAACCGAUU GAAUUGUGUUGCAGAGCUU
ATR	siRNA pool	Invitrogen	HSS100876 HSS100877 HSS100878	UUUAGAUGAGGUUCUAGUAUUUCCC UAAAUUGGCUUCUUUACUCCAGACC UUAACAUGUUCUUACCCUCAGGUGG
ATR	siRNA pool	Ambion	s535 s536 s534	UAAAUUUUGCAUACUCAUCaa UCAGUAUCCAUUUCUACAAgg UUGACUUAAAAAUCGGCUCat
BRCA2	siRNA	Ambion	s224694	GGCUCAUACCCUCCAAUGAtt
HIRIP3-1	siRNA	Dharmacon	D-011481-01	GCAGUGAUGGCGAGAGUAA
HIRIP3-2	siRNA	Dharmacon	D-011481-02	UCAGCACGCUUACGCAUUC
HIRIP3-3	siRNA	Dharmacon	D-011481-03	ACAAGGAGCGCCUGAGUAU
HIRIP3-4	siRNA	Dharmacon	D-011481-04	GUAGCGACCCGGAGAGAAA
HIRIP3-5	siRNA	Invitrogen	HSS189204	GGUGGAGGGAAAUAAAGGAACUAAA
HIRIP3-6	siRNA	Invitrogen	HSS189205	GAAAGUGACUUGGAGAGGGAGGUAA
HIRIP3-7	siRNA	Invitrogen	HSS112359	GCCUCCUUGGAUGUUGCGAACAUCA
UBN1-1	siRNA	Invitrogen	HSS120994	GGGUGUAUGCCUAUCUUGCGUCAUU
UBN1-2	siRNA	Invitrogen	HSS120995	GGAUGCAGGCCAGAACUCUGUUUAA
UBN1-3	siRNA	Invitrogen	HSS120996	GCAGUUAGUGAAGACAGCGGCCAAA
HIRA-1	siRNA	Dharmacon	D-013610-01	GAAGGACUCUCGUCUCAUG
HIRA-2	siRNA	Dharmacon	D-013610-02	GGAGAUGACAAACUGAUUA
HIRA-3	siRNA	Dharmacon	D-013610-04	GAAAUUCUAGCUACUCUGA
HIRA-4	siRNA	Dharmacon	D-013610-05	GCGAUUCUGUCAAUAAAGA
CAIN-1	siRNA	Dharmacon	D-012454-01	GAACACAGCCCACGAGUAU
CAIN-2	siRNA	Dharmacon	D-012454-02	GGAGAGAGCUUGCUGGCCA
CAIN-3	siRNA	Dharmacon	D-012454-03	GGAUUGAUUUGUCGGACUA
CAIN-4	siRNA	Dharmacon	D-012454-04	GAUGUCAACCUCUGGUAUA
FF	shRNA	custom	custom	CCCGCCTGAAGTCTCTGATTAA
ATR	shRNA	Open Biosystems	V2HS_94661; RHS1764-9191741	ATAATGAATGATCTGGTCTGGT
HIRIP3-8	shRNA	Open Biosystems	V3LHS_334056; RHS4430-101065004	TATGTGACCAGTCTGGGGGTGC

Table 2 (continued). RNAi Reagents

Target	Reagent Type	Source	Clone ID / Catalog Number	Sequence
HIRIP3-9	shRNA	Open Biosystems	V3LHS_334058; RHS4430-101068486	TCTGAATTGAAGCGGAACCTTT
HIRIP3-10	shRNA	Open Biosystems	V3LHS_334057; RHS4430-101074418	TAGCGCTTCCAGTTCTGCCCGG
HIRIP3-11	shRNA	Open Biosystems	V3LHS_334055; RHS4430-101074132	TCCGACTCTGAATTGAAGCGGA

4-3. Antibodies

Primary antibodies used in this chapter are listed in Table 3. Secondary antibodies used for immunofluorescence were Alexa Fluor® 488 and 594 conjugated (Invitrogen) and for western blot analysis were HRP conjugated (Jackson Laboratory).

Table 3. Primary Antibodies

Antibody	Host Name	Source	Name / Catalog #	Information
HIRA	mouse	Peter D. Adams	WC cocktail	
UBN1	rabbit	Peter D. Adams	#1358	polyclonal raised to N-terminus
CAIN	rabbit	Calbiochem	1881-2173	
HIRIP3	rabbit	Santa Cruz	sc-98401	
γH2AX	rabbit	Bethyl	A300-081A	
γH2AX	mouse	Millipore	05-636	
RAD51	rabbit	Santa Cruz	sc-8349	
Actin	mouse	Santa Cruz	sc-8432	
VINCULIN	mouse	Sigma	V9131	

4-4. UV laser-induced DNA damage and immunofluorescence

UV laser-induced DNA damage was generated as previously described (Bekker-Jensen et al., 2006). Cells were sensitized to UV-A laser (λ = 355 nM) by 24 hour pre-treatment with 10 μM BrdU and microirradiated using a PALM MicroBeam with fluorescence illumination (Zeiss) at 40-45% laser power.

After damage, cells were allowed to recover for the indicated times at either room temperature (RT) or 37°C. Cells were extracted with 0.5% Triton X-100 only where indicated. Cells were then fixed in 3.7% formaldehyde for 10 minutes at RT. Fixed cells were washed with PBS, permeabilized in 0.5% NP-40, washed with PBS, and blocked with PBG (0.2% [w/v] cold fish gelatin, 0.5% [w/v] BSA in PBS) for 30 minutes prior to immunostaining with the indicated antibodies diluted in PBG. DNA was stained with DAPI by addition of Vectashield Mounting Medium (Vector Laboratories). GFP was observed directly. Images were collected on an Axioplan2 Zeiss microscope with an AxioCam MRM Zeiss digital camera and Axiovision 4.5-4.8 software. Images in this chapter were not intended for direct comparison.

4-5. HR Assay

The procedure by which low-throughput HR assays were conducted was similar to the HTP protocol (Chapter 2, 4-2), except: (1) DR-U2OS cells were either forward or reverse transfected in 6 well plates, (2) AdNGUS24i was used at an exact MOI of 10, (3) GFP+ ratios were determined ~36-48 hour post infection by FACS analysis on a BD LSRII Flow Cytometer (BD Biosciences).

4-6. RT-qPCR

RNA was isolated from cells using the RNAeasy Plus kit (Qiagen) and reverse transcribed into cDNA using SuperScript III Reverse Transcriptase (Invitrogen #18080-044) according to the manufacturer instructions. Quantitative RT-PCR (RT-qPCR) was performed using Platinum Cybergreen Super Mix with Rox dye (Invitrogen #11733-046) on an Applied Biosystems 7500 Fast PCR machine. RT-qPCR primers used were: RAD51 left primer 5'-CGTTCAACACAGACCACCAG; RAD51 right primer 5'-CGGTGGCACTGTCTACAATAAG; HIRA left primer 5’-AAGCCCCCAGAGAGCATT; HIRA right primer 5’-GTCACTTCATTCTCCACCTCAA; β-actin left primer 5’-GCTACGAGCTGCCTGACG; β-actin right primer 5’-GGCTGGAAGAGTGCCTCA.

4-7. Genome-wide Enrichment of Seed Sequences (GESS) analysis

GESS off-target analysis was conducted as described (Sigoillot et al., 2011). 2,564 Dharmacon siRNAs (from 641 pools) and 1401 Ambion siRNAs against candidate HR mediators, as well as 1000 Dharmacon siRNAs (from 250 pools, including 1 duplicate pool) against candidate HR suppressors, were each submitted to GESS analysis. Significance *p*-values were determined as follows: The Yates' Chi Square statistic and associated one-tailed *p*-value were calculated for each database sequence evaluated (3'UTR or CDS) if all siRNA categories (active siRNAs with or without matching and inactive siRNAs with or without matching) had more than 20 seed match events. Otherwise, a two-sided *p*-value was calculated from the Fisher's Exact Test. Transcript sequences were ranked from lowest to highest *p*-value and the statistical significance was determined by comparing the *p*-value to a *p*-value threshold (0.05) corrected for multiple hypothesis testing using the Benjamini and Hochberg (Simes') method.

Chapter Four

Characterization of the novel DNA damage response protein RBMX

Britt Adamson[1], Agata Smogorzewska[1,2], Stephen J. Elledge[1]

[1]Division of Genetics, Brigham and Women's Hospital, Department of Genetics, Harvard University Medical School, Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA

[2]Laboratory of Genome Maintenance, The Rockefeller University, New York, New York 10065, USA

Work discussed herein can be found in the following publication:

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol *14*, 318-328.

Experiments presented herein were conducted by Britt Adamson, except for Figure 18a which was performed by Agata Smogorzewska. We thank Dr. Simon Boulton for the PARP inhibitor.

I. Introduction

1-1. The RNA-processing machinery and the DNA damage response

Functional gene analysis of our screen candidates highlighted an increasingly common observation: that RNA-processing and RNA-binding (RBP) proteins play some (as yet mostly undefined) role in maintaining genomic stability. Our findings that the RBP RBMX and the RNA helicase DDX17 accumulate at regions of damaged DNA and regulate HR have indicated that specific components of the RNA-processing machinery are both regulated by and required for DDR function. Similar conclusions have been drawn from additional large-scale evaluations of the DDR. In particular, genome-wide RNAi screens have found that mammalian RNA processing proteins are required for cellular resistance to IR, maintenance of genomic stability and chromosome end protection (Hurov et al., 2010; Lackner et al., 2011; Paulsen et al., 2009), and proteomic studies have found that such proteins are targeted by ATM / ATR phosphorylation (Beli et al., 2012; Matsuoka et al., 2007). RNA-binding / editing proteins also comprise a major functional category of genes found to be transcriptionally altered in response to IR and UV, alongside the more anticipated categories of cell proliferation, DNA repair, stress response and signal transduction (Rieger and Chu, 2004). In yeast, a screen to identify reduction-of-function alleles that cause chromosome instability (CIN) and aberrant Rad52 foci formation also identified several mRNA-processing genes (Stirling et al., 2011a; Stirling et al., 2011b). This chapter describes current hypotheses for RBP function within the DDR and details our specific characterization of RBMX within the context of DNA repair.

1-1-1. Transcriptional control and splicing regulation as part of the DNA damage response

The DNA damage response has traditionally been viewed as a program of rapid posttranslational modification initiated to quickly coordinate the function of repair-relevant proteins followed by delayed transcriptional regulation –controlled by transcription factors like p53– to promote cell cycle checkpoint

activation and ultimately determine cell fate through proliferation control. Under this view, cellular RNA has only an intermediary role in the DDR, and changes to the RNA landscape induced by genotoxic stress occur merely as byproducts of transcriptionally regulated gene expression. Increasingly, however, evidence suggests that RNA itself has a greater functional complexity within the DDR, and recent work indicates that (1) RNA molecules themselves (specifically miRNAs) enact functional roles in the DDR, (2) posttranscriptional regulation of mRNA contributes to gene expression changes after DNA damage, and (3) transcriptional regulation in response to DNA damage has a functional significance beyond modulating protein expression.

Direct regulation of transcription is one way to alter protein expression in response to stimuli; however, as mentioned, alternative methods of posttranscriptional regulation are being found to operate within the DDR, including miRNA- and RBP-mediated translation repression, regulation of mRNA stability, and programs of alternative splicing. With regard to the first of these, a recent study of miRNA biogenesis found that ~25% of miRNAs in murine cells are upregulated in an ATM-dependent manner after exposure to DNA damage, and a subset of these were found to be regulated by ATM-mediated phosphorylation of the RNA-binding protein and splicing regulator KSRP (Zhang et al., 2011), which is also a key subunit in DROSHA- and DICER-containing miRNA-processing complexes (Trabucchi et al., 2009). Phosphorylation of KSRP was found to enhance protein association with specific pri-miRNAs in a manner thought to improve pri-miRNA maturation and thus promote specific transcript targeting. Overall, this suggests that a regulated program of miRNA-mediated gene repression may contribute to the DDR.

Direct RBP-mediated mRNA regulation has also been proposed as a model for gene expression control within the DDR (Reinhardt et al., 2011). A study by Yaffe and colleagues found that altered association of *GADD45α* mRNA with two RNA-binding proteins (hnRNPA0 and TIAR) as well as phosphorylation of the poly-A ribonuclease protein (PARN) after doxorubicin exposure promotes overall *GADD45α* mRNA stabilization (Reinhardt et al., 2010). The authors of this work suggest that PARN and TIAR (a translational inhibitor) suppress *GADD45α* mRNA stability and translation in undamaged cells,

but after DNA damage, decreased association with TIAR, negative regulation of PARN, and an increased association with hnRNPA0 stabilize the message. Importantly, GADD45α (growth arrest and DNA damage inducible protein, alpha) is a cell cycle checkpoint regulator induced by DNA damage. This example integrates two possible mechanisms of direct mRNA regulation by RBPs, control through (1) stability and (2) translation. However, RBPs may also regulate other aspects of RNA metabolism as part of the DDR, including nuclear export and spliceosome targeting (Reinhardt et al., 2011).

Emerging evidence suggests that splicing is controlled by the DDR. One study found that of 482 genes related to cancer, cell cycle, cell proliferation and cell death (all predicted to be alternatively spliced), 102 (22%) demonstrate some alternative splicing (AS) event after exposure to UV (Munoz et al., 2009b). This study also showed that damage-induced AS occurs in *trans* with regard to DNA lesions, and that it is correlated with, and possibly regulated by, changes to the phosphorylation state of the RNA polymerase II (RNAPII) C-terminal domain (CTD) and lower transcription elongation rates. A second study, which characterized the Ewing sarcoma protein (EWS) as a cotranscriptional splicing regulator, found that after UV exposure, AS of DDR-related genes, including *CHK2*, correlates with diminished EWS-mRNA binding (Paronetto et al., 2011). Taken together, these results suggest that UV-induced AS is a regulated component of the DDR that, in part, may be directed to specific transcripts by RBPs.

In one way or another, DNA damage no doubt prompts functionally relevant changes to the expression of specific proteins; however, evidence also indicates that bulk transcription is regulated by the DDR. In particular, UV irradiation has been observed to suppress the transcriptional activity of nuclear extracts in a manner controlled by the phosphorylation state of the RNAPII CTD (Rockx et al., 2000). The RNAPII CTD is composed of multiple tandem amino acid heptapeptide repeats (YSPTSPS). CTD phosphorylation of the second and fifth serine residues within these repeats occurs on the elongating form of the polymerase (RNAPIIo), but only hypophosphorylated RNAPII (RNAPIIa) can be recruited to transcription initiation complexes. Following UV, transcription is first inhibited through bulk RNAPIIa phosphorylation (initiation suppression) and then through RNAPIIo ubiquitylation and subsequent proteasomal degradation (Luo et al., 2001; Munoz et al., 2009b). Because RNAPIIo can be ubiquitylated

by BRCA1 / BARD1 *in vitro*, this bulk transcriptional suppression after DNA damage is thought to be BRCA1-dependent (Kleiman et al., 2005). Regulation of RNAPII in this manner may affect net elongation rates and, as mentioned above, contribute to damage-initiated programs of alternative splicing. Overall though, suppression of bulk transcription after DNA damage makes logical sense for two more obvious reasons: (1) The continued production of potentially damaged mRNAs could lead to aberrant and possibly detrimental protein function. (2) Transcription is an energetically costly process, the maintenance of which may not be prioritized under conditions of cell stress.

The presence of bulky transcriptional proteins near sites of DNA damage might also hinder repair processes, and their removal through local transcriptional repression could also be advantageous. Perhaps not surprisingly then, data from our work (not reported here) and others have shown that *cis*-acting programs of transcriptional silencing are enacted near DNA breaks (Chou et al., 2010; Shanbhag et al., 2010). Greenberg and colleagues demonstrated this phenomenon using a clever reporter construct (Shanbhag et al., 2010). They showed that active transcription, observed through the association of nascent transcripts with a protein fluorophore, is suppressed in chromatin flanking a single and separately marked DSB. In support of this, both our work and theirs showed that RNAPII CTD serine-2 phosphorylation is lost near DSBs to a greater extent than in the surrounding chromatin (Chou et al., 2010; Shanbhag et al., 2010). In our work, antibodies against marks of active transcription, including phospho-RNAPII-S2 and the 7-methylguanosine cap structure of nascent RNA transcripts, revealed that transcription is specifically suppressed within regions of DNA damaged by UV microirradiation. We observed that this occurs in a manner correlated with the recruitment of transcriptional repressor proteins, including polycomb group proteins, to damaged regions (Chou et al., 2010).

1-1-2. Transcription-coupled repair

Perhaps counterintuitively, transcriptional machineries can also function as sensors of DNA damage. This occurs predominantly in the context of bulky chemical alterations, such as those formed by UV and repaired by nucleotide excision repair (NER) (Figure 2) (de Laat et al., 1999; Lagerwerf et al.,

2011). NER recognizes bulky lesions in at least two distinct ways: (1) by a global genome (GG) NER sensor complex XPC–HR23B, which recognizes distortions of the DNA helical structure, and (2) by the elongating form of RNAPII, which stalls on DNA lesions within transcribed regions of the genome (de Laat et al., 1999; Lagerwerf et al., 2011). This second method of bulky lesion recognition is referred to as transcription-coupled (TC) NER. After stalling, RNAPIIo binds the protein CSB and facilitates assembly of general NER factors, including TFIIH, RPA, XPA and the endonucleases XPG and XPF / ERCC1. As part of TC-NER, CSB-dependent recruitment of CSA also facilitates the recruitment of TC-NER specific factors: HMGN1, XAB2 (splicing factor), and TFIIS. Once assembled, these proteins excise a short region of DNA (~30-nt) containing the lesion site, and coordinate repair by DNA gap filling (Lagerwerf et al., 2011).

Whether or not RNAPII dissociates from DNA during TC-NER is an open question. As discussed above (Chapter 4, 1-1-1), bulk RNAPIIo is known to be degraded after exposure to UV, and site specific transcription repression has been observed at DSBs; however, it has also been reported that RNAPII remains bound to chromatin near DNA lesions during NER in a manner that may facilitate resumption of transcription post repair (Lagerwerf et al., 2011). The stability of RNAPIIo near DNA lesions may depend on many factors, including the amount and type of DNA damage sustained, the speed of repair, or the density of blocked polymerases within a genomic region. Nevertheless, the role of RNAPII as a DNA damage sensor leads to the possibility that cotranscriptional RNA-processing / RNA-binding proteins may act in a parallel fashion by aiding the recruitment of DNA repair factors to sites of polymerase stalling.

1-1-3. R-loops and genomic instability

Much attention has recently been given to the idea that excessive RNA:DNA heteroduplex formation at both normally elongating and stalled RNA polymerases increases genomic instability (Aguilera and Garcia-Muse, 2012). RNA:DNA hybrids are not uncommon structures. They occur during normal cellular processes, including DNA replication and transcription (Aguilera and Garcia-Muse,

2012). However, rare RNA:DNA transcriptional byproducts, so-called R-loops, also form when nascent RNA strands "fold-back" and pair with complementary and negatively supercoiled regions of their DNA templates (Aguilera and Garcia-Muse, 2012). Structurally, these can be thought of as RNA corollaries to D-loops. R-loops have been described to have biologically positive functions in a few special cases. In mitochondria and *E. coli*, they serve as priming intermediates for DNA replication, and they are functional components of the immunoglobulin class switch machinery in vertebrate B cells (Aguilera and Garcia-Muse, 2012). Aberrant R-loop formation, however, is thought to cause genomic instability through displacement of ssDNA or through collisions with the replication machinery. Because of this R-loop formation is generally thought to be proactively limited.

An obvious method for repressing R-loop formation during transcription is one in which RBPs and the splicing machinery bind nascent mRNA transcripts to inhibit RNA:DNA association. This model is supported by genetic evidence in both yeast and mammalian cells. In yeast, mutation of mRNA cleavage and polyadenylation genes has been shown to cause both increased RNA:DNA hybrid formation and genomic instability-associated phenotypes, including chromosome instability (CIN) and aberrant Rad52 foci (Stirling et al., 2011a; Stirling et al., 2011b). Importantly, the expression of RNase H in these same mutants was shown to suppress CIN. RNase H is an enzyme with RNA-directed endonucleolytic activity that specifically degrades RNA from RNA:DNA hybrid molecules; and its activity in this case strongly argues that in the absence of RNA-processing, DNA damage is incurred by the formation of RNA:DNA intermediates. Consistent with this interpretation, depletion of mRNA-processing proteins from mammalian cells has been shown to cause aberrant accumulation of γH2AX foci in a manner that can be suppressed by RNase H (Paulsen et al., 2009), and RNase H-repressible genomic instability has been correlated with increased R-loop formation after depletion of the splicing regulator ASF / SF2 from vertebrate cells (Li and Manley, 2005). This model of RNA induced DNA-damage could explain the predominance of RNA-processing proteins among mediators of genomic stability identified in genetic screens (Hurov et al., 2010; Lackner et al., 2011; Paulsen et al., 2009; Stirling et al., 2011a; Stirling et al., 2011b).

A second possible method for R-loop suppression by the RNA-processing machinery is active disruption of aberrant DNA:RNA pairing by RNA helicases. In support of this, DDX1, a nuclear DEAD-box helicase with ATP-dependent RNA:RNA and DNA:RNA unwinding activity, has been shown to localize to sites of DNA damage in an manner that can be disrupted by RNase H treatment (Li et al., 2008); and recently, depletion of the DEAD-box RNA:DNA helicase SETX has been shown to directly increase R-loop formation (Skourti-Stathaki et al., 2011). Indicative of a role in maintaining genomic stability, human cells carrying *SETX* mutations demonstrate increased sensitivity to DNA-damaging agents, including H_2O_2, camptothecin and mitomycin C (Suraweera et al., 2007). Overall, these observations indicated that R-loop suppression is an important function of the RNA-processing machinery.

1-2. The RNA-binding protein RBMX

Because our screen identified a network of RNA-processing proteins and one RBP that localizes to sites of DNA damage, we next focused our studies on RBMX. RBMX is a nuclear hnRNP protein that regulates alternative splicing in at least two possible ways: one through RNA binding by an RNA recognition motif (RRM) and one independent of the RRM (Heinrich et al., 2009; Hofmann and Wirth, 2002; Nasim et al., 2003; Wang et al., 2004). X-linked RBMX and a Y chromosome-encoded paralog (RBMY) are conserved among mammals, and in humans, there are several intron-less retrogenes of RBMX present on various autosomes (Lingenfelter et al., 2001; Mazeyrat et al., 1999). Expression of RBMX and at least one of its retrogenes (RBMXL1) is ubiquitous throughout tissue types, while expression of RBMY is restricted to male germ cells indicating a role in spermatogenesis and possibly meiosis (Elliott, 2004; Lingenfelter et al., 2001; Mazeyrat et al., 1999). RBMX has also been proposed to be a tumor suppressor (Shin et al., 2006; Shin et al., 2007).

II. Results

2-1. RBMX accumulates at regions of DNA damage in a PARP-dependent manner

We observed GFP-RBMX localization to regions of DNA damaged by mircoirradiation in ~20-40% of U2OS cells (Figure 14a,d). RBMX microirradiation tracks (or "stripes") were also observed by fluorescent immunostaining of endogenous and Flag/Ha (FHA)-tagged RBMX after Triton X pre-extraction (Figure 14b,c). Background chromatin binding likely obscured detection of endogenous and FHA-RBMX stripes in the absence of pre-extraction, which is supported by the observation that the percentage of cells with GFP-RBMX tracks increased after endogenous RBMX was depleted by siRNA (Figure 14a). GFP-RBMX localization to DNA damage was transient, occurring 0-10 minutes after microirradiation (longer at room temperature); and following this initial recruitment GFP-RBMX was removed from damaged DNA causing a localization pattern we refer to as the "anti-stripe" (Figure 14d,e). HA-tagged and endogenous RBMX formed anti-stripes as well (Figure 14b,c). Unlike stripe formation, which is likely easily obscured by fluorescent immunostaining of pan-nuclear RBMX molecules, Triton X pre-extraction was not required to observed anti-stripes in these cells. We failed to observe RBMX accumulation at ionizing radiation-induced foci (data not shown). However, this could be explained by the transient nature of RBMX recruitment or a low signal-to-noise ratio. We previously showed that RNAPII forms anti-stripes after microirradiation in a manner correlated with transcriptional repression at sites of active DNA repair (Chou et al., 2010). Transcriptional repression at sites of DNA damage is discussed in Chapter 4, 1-1-1. Consistent with this interpretation, we also observed anti-stripe localization of other hnRNP proteins after microirradiation (Figure 15).

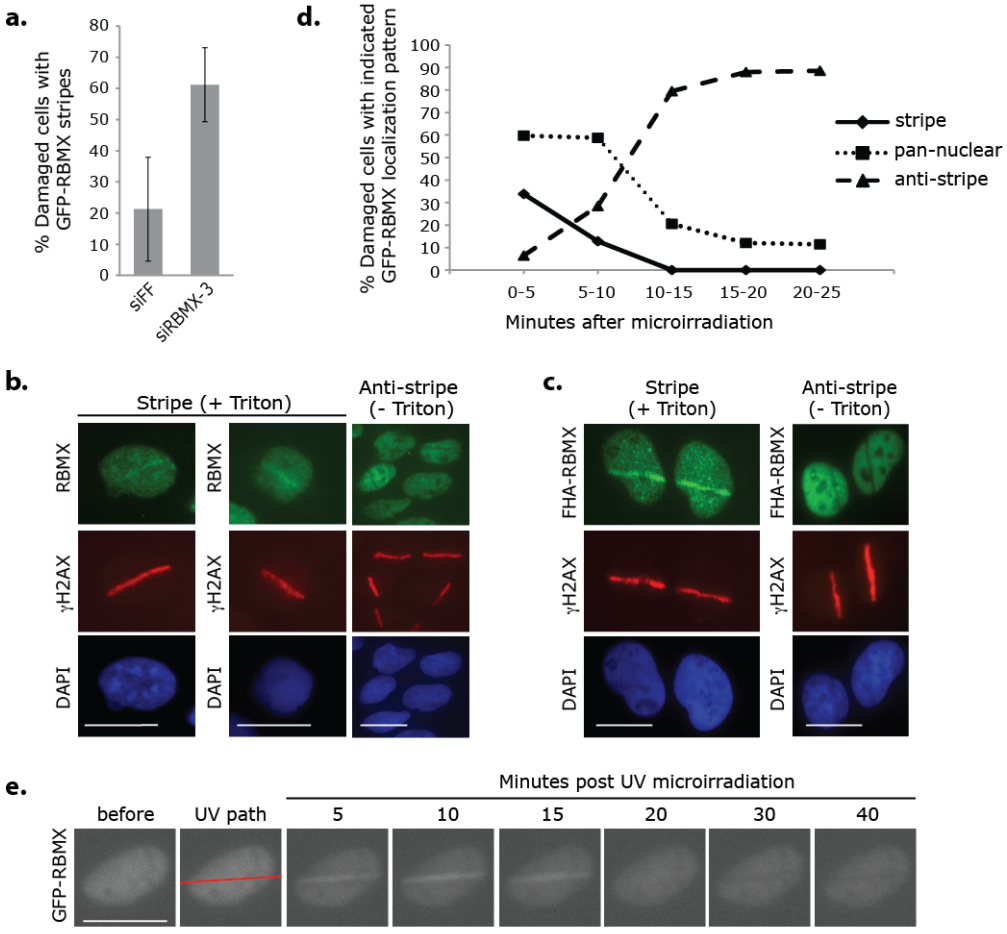


Figure 14. RBMX transiently accumulates at tracks of microirradiation-induced DNA damage. **(a)** U2OS cells expressing siRBMX-3 resistant GFP-RBMX and transfected with the indicated siRNAs were microirradiated (one at a time for 5 minutes) and processed for immunofluorescence with an antibody against γH2AX. Cells were evaluated for GFP-RBMX accumulation at γH2AX stained laser tracks (~60-190 cells / condition). Error bars represent ± s.d. across three replicates. **(b)** U2OS cells were microirradiated and processed for immunofluorescence with the indicated antibodies within 35 minutes. Cells were fixed directly or pre-extracted with 0.5% TritonX-100. Nuclei were stained with DAPI. Images were prepared from two separate experiments (panels are grouped accordingly). Exposure and processing were adjusted to best demonstrate stripes and anti-stripes. Scale bars indicate 10 μm. **(c)** U2OS cells expressing Flag/HA (FHA)-tagged RBMX were microirradiated and processed for immunofluorescence with antibodies against HA and γH2AX within 25 minutes. Cells were fixed directly or pre-extracted with 0.5% TritonX-100. Nuclei were stained with DAPI. Images were prepared from two separate experiments (panels are grouped accordingly) and are not intended for direct comparison. Scale bars indicate 10 μm. **(d)** U2OS cells expressing GFP-RBMX were microirradiated (one at a time for 5 minutes) and processed for immunofluorescence with an antibody against γH2AX at the indicated times. Cells were evaluated for GFP-RBMX accumulation at γ H2AX stained laser tracks (~130-190 cells / condition). Data represent the mean of two replicates. **(e)** Live cell imaging of GFP-RBMX recruitment to DNA damage at room temperature. U2OS cells expressing GFP-RBMX were laser microirradiated and imaged at indicated times after damage. Scale bar indicates 25 μm.

RBMX recruitment to DNA damage was independent of ATM signaling and H2AX but dependent on poly(ADP-ribose) polymerase 1 (PARP1) (Figure 16a-c). PARP1 is known to mediate the recruitment of several repair proteins to DNA damage *via* the transient polymerization of branched poly(ADP-ribose) (PAR) structures (discussed in Chapter 1, 3-1), and it is likely that RBMX localization is similarly facilitated. GFP-RBMX tracks occurred coincident with PAR formation at DNA breaks (Figure 16d), and inhibition of PARP1 –through use of the chemical inhibitor KU-0058948 or siRNA-mediated depletion– prevented GFP-RBMX track formation and caused early formation of anti-stripes (Figure 16a,c,e). KU-0058948 was used at 1 μM, a concentration that inhibits the activity of PARP1 and PARP3 *in vitro* (Loseva et al., 2010). Additionally, siRNAs against PARG, the PAR disassembly enzyme, increased the percentage of cells with GFP-RBMX tracks after microirradiation and prolonged localization at damage (Figure 16c,e).

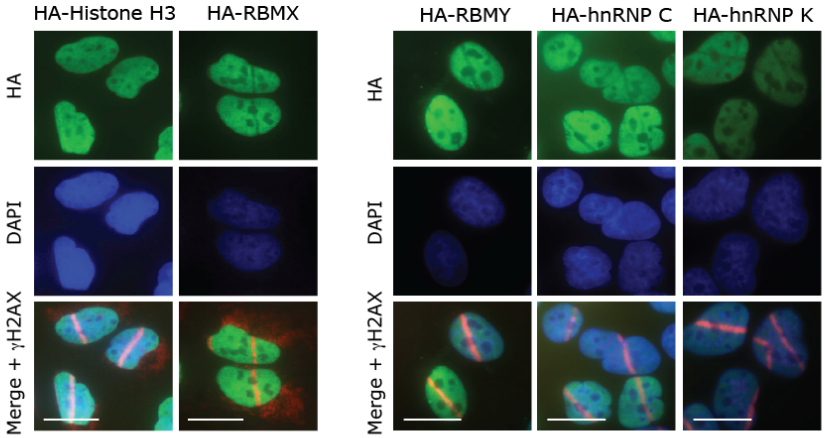


Figure 15. Anti-stripe formation with additional RNA-binding proteins. The Y-chromosome homolog of RBMX (RBMY) and two additional hnRNP proteins (hnRNP C and hnRNP K) form anti-stripes after microirradiation. U2OS cells expressing the indicated HA fusion proteins were microirradiated and processed for immunofluorescence with antibodies against HA and γH2AX within 25 minutes. HA-tagged histone H3 does not form anti-stripes and serves as a negative control. Images were prepared from two separate experiments (panels are grouped accordingly). Exposure and processing were adjusted to best demonstrate localization. Scale bars indicate 10 μm.

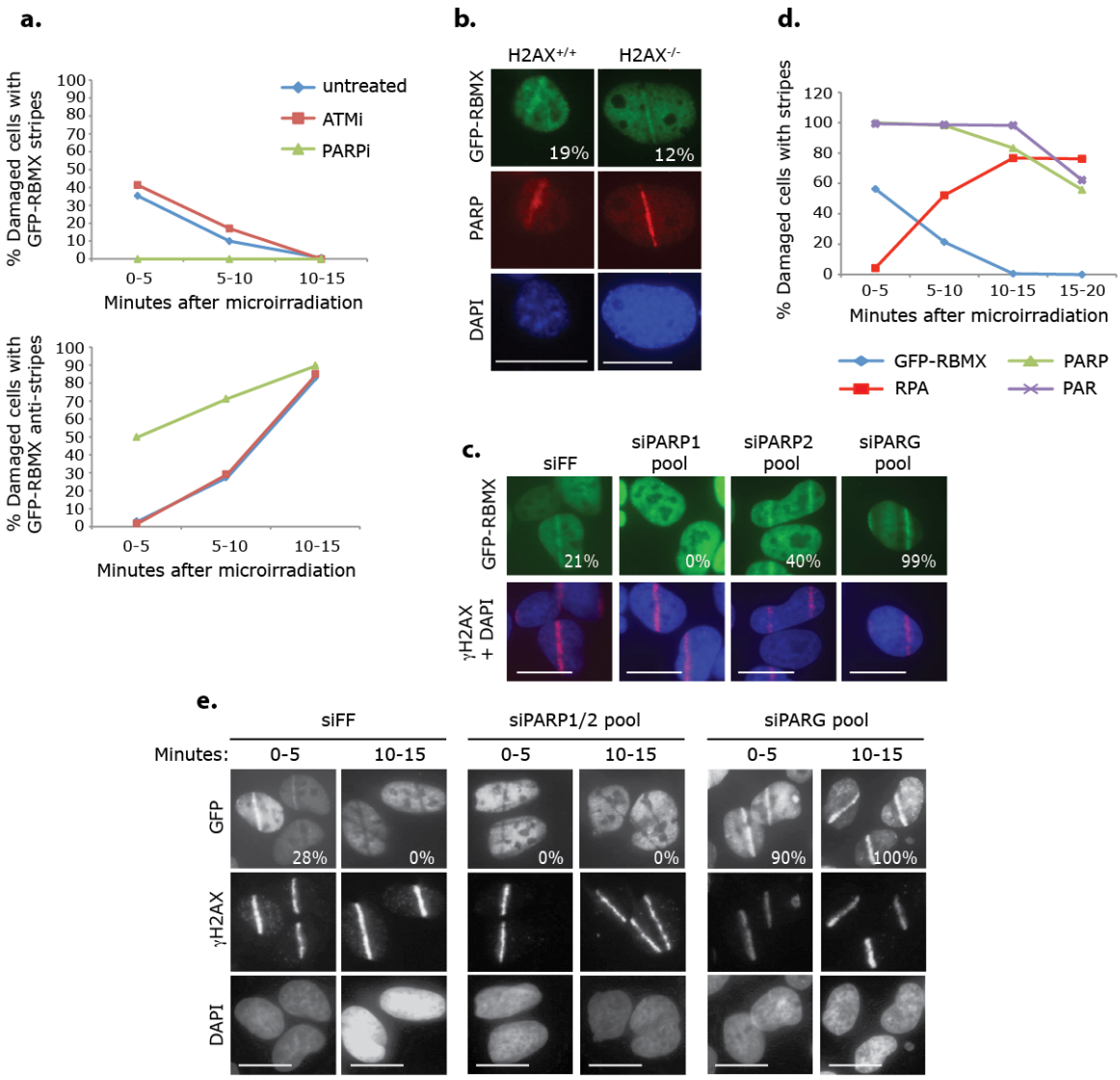


Figure 16. RBMX recruitment to DNA damage requires PARP1 activity but not H2AX or ATM signaling. **(a)** U2OS cells expressing GFP-RBMX were treated with a PARP inhibitor (PARPi), ATM inhibitor (ATMi), or DMSO (untreated, negative control) for 1 hour and then microirradiated (one at a time for 5 minutes) and processed for immunofluorescence with an antibody against γH2AX at the indicated times. Cells were evaluated for GFP-RBMX accumulation (top) or exclusion (bottom) at γH2AX stained laser tracks (~80-160 cells / condition, n=1). **(b)** H2AX-/- mouse embryonic fibroblasts and an isogenic cell line reconstituted with H2AX (H2AX+/+) expressing GFP-RBMX were microirradiated (one at a time for 5 minutes) and processed for immunofluorescence with an antibody against PARP. GFP was observed directly. Nuclei were stained with DAPI. The percentage of cells with GFP-RBMX accumulation at γ H2AX stained laser tracks is indicated for each condition (~80-140 cells / condition, n=1). Scale bars indicate 10 μm. **(c)** RBMX stripe formation was inhibited by PARP1-targeting but not PARP2-targeting siRNAs and stabilized by siRNAs against PARG. U2OS cells expressing GFP-RBMX and transfected with the indicated siRNAs were microirradiated (one at a time for 5 minutes) and processed for imaging with an antibody against γH2AX. GFP was observed directly. Nuclei were stained with DAPI. The percentage of cells with GFP-RBMX accumulation at γH2AX stained laser tracks is indicated for each condition (~90-300 cells / condition). Data represent the mean of two replicates except for siFF, which was performed in triplicate. Scale bars indicate 20 μm. **(d)** GFP-RBMX stripe formation was coincident with accumulation of PARP and PAR at laser tracks and prior to RPA accumulation. U2OS cells were microirradiated (one at a time for 5 minutes) and processed for immunofluorescence with antibodies against γH2AX and PARP, PAR or RPA at the indicated times. U2OS cells expressing GFP-RBMX were similarly microirradiated and processed for immunofluorescence with an antibody against γH2AX. GFP was observed directly. Damaged cells were evaluated for the indicated stripes at γH2AX stained laser tracks (~90-210 cells / condition, n=1). **(e)** U2OS cells expressing GFP-RBMX and transfected with the indicated siRNAs were microirradiated (one at a time for 5 minutes) and processed for imaging with an antibody against γH2AX at the indicated times. GFP was observed directly. Nuclei were stained with DAPI. The percentage of cells with GFP-RBMX accumulation at γH2AX stained laser tracks is indicated for each condition (~110-160 cells / condition, n=1). siPARP1/2 pool indicates a combined pool of siRNAs targeting PARP1 and PARP2 (4 duplexes each). Scale bars indicate 10 μm.

2-2. RBMX promotes homologous recombination

The RBMX siRNA pool in our primary screen decreased HR to 7% of controls, comparable to the effect of depleting BRCA2 and RAD51 (5% and 11%, respectively) (Table S1). All four siRNAs from this pool (siRBMX-1 through -4), as well as two Ambion siRNAs (siRBMX-5 and -6) and three independently selected shRNAs (shRBMX-7, -9 and -10), caused defective HR in a manner correlating with RBMX depletion (Figure 17a-d, Table S4). Importantly, we ruled out obvious off-target effects for these RNAi reagents by determining that RBMX depletion does not alter cell cycle distribution, correlate with *RAD51* depletion, or disrupt I-SceI expression (Figure 17e-h). We also found that expression of siRNA-resistant FHA-tagged RBMX rescued the siRBMX-3 associated HR defect (Figure 17i). There were no confounding effects on cell cycle distributions in the rescue assay (Figure 17j).

RBMX-targeting siRNAs also sensitized cells to DNA damaging agents that engage the HR machinery for repair, including DSB-inducing irradiation (IR), replication stress-inducing camptothecin, and several crosslinking agents (mitomycin C, chlorambucil, oxaliplatin, and carboplatin) (Figure 18a-c). The sensitivity to mitomycin C caused by siRBMX-3 was significantly attenuated by expression of siRNA-resistant FHA-RBMX (Figure 18b). Interestingly, RBMX depletion also caused sensitivity to ultraviolent light (UV) and tert-butyl hydroperoxide (tBHP), both of which cause DNA lesions not primarily repaired by HR (Figure 18a,c).

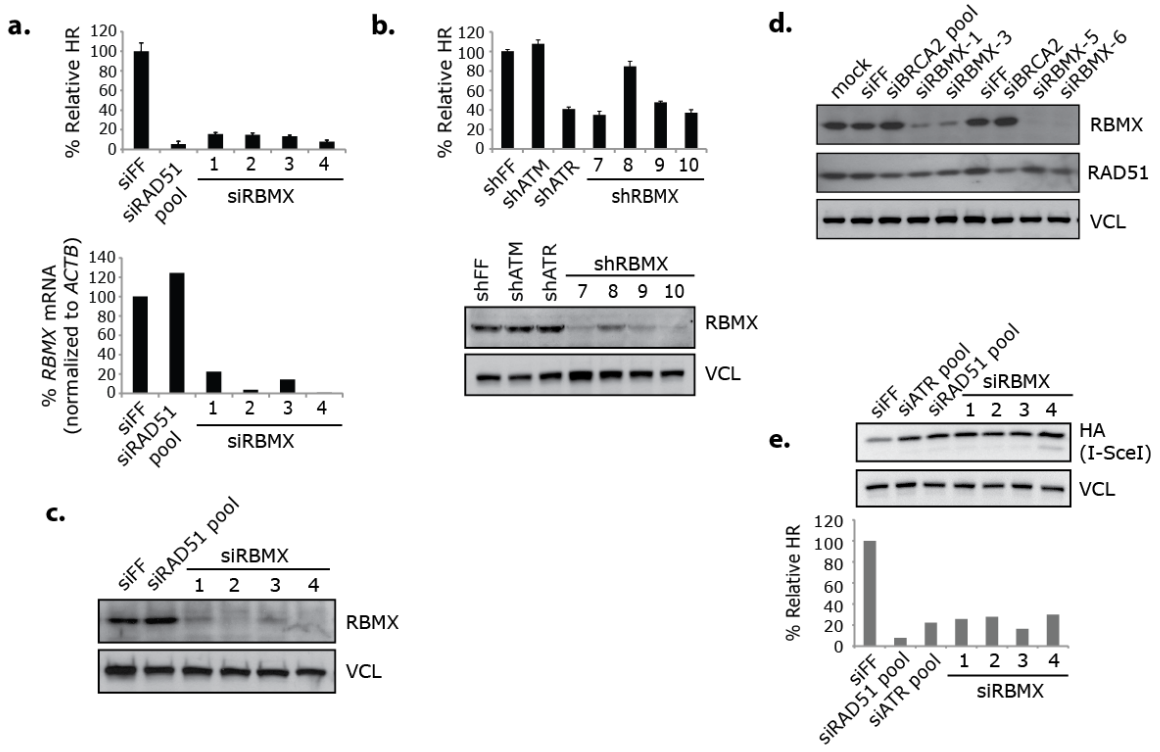
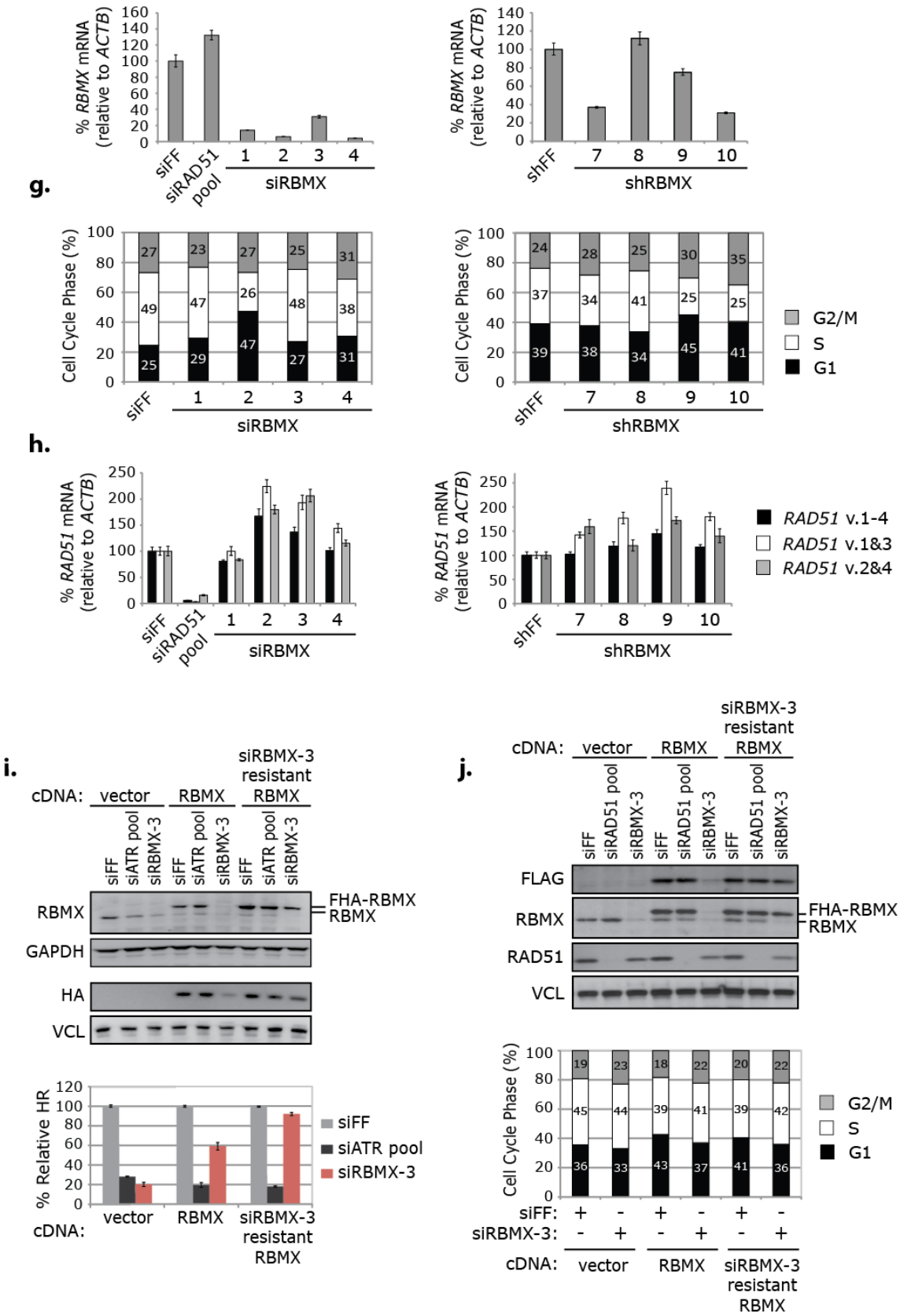


Figure 17. RBMX promotes homologous recombination. **(a)** HR assay results from DR-U2OS cells transfected with the indicated siRNAs and corresponding levels of *RBMX* mRNA as measured by RT-qPCR (normalized to *ACTB* mRNA). HR analysis error bars represent ± s.d. across three replicates and data in RT-qPCR analysis represent the mean of two replicates. **(b)** HR assay results from DR-U2OS cells transduced with the indicated shRNAs and corresponding levels of RBMX as measured by western blot analysis; an antibody against VCL was used as a loading control. Error bars represent ± s.d. across three replicates. **(c-d)** Whole-cell extracts from cells transfected with the indicated siRNAs were immunoblotted with the indicated antibodies. Data in (c) and (d) are from independent experiments. Extracts from (d) were separately evaluated in Figure 24b. **(e)** Whole-cell extracts from DR-U2OS cells transfected with the indicated siRNAs and infected with AdNGUS24i were immunoblotted with antibodies against HA and VCL (loading control) to determine levels of HA-tagged I-SceI expression and assayed for HR efficiency (n=1). **(f)** RT-qPCR of *RBMX* (normalized to *ACTB*) from DR-U2OS cells transfected with the indicated siRNAs or transduced with the indicated shRNAs. Error bars represent ± s.e.m. across three replicates. **(g)** Cell cycle analysis of DR-U2OS cells transfected with the indicated siRNAs or transduced with the indicated shRNAs (n=1). **(h)** RT-qPCR of *RAD51* (normalized to *ACTB*) from DR-U2OS cells transfected with the indicated siRNAs or transduced with the indicated shRNAs. The three primer sets used against RAD51 (v.1-4, v.1&3, v.2&4) have different specificities for four transcript variants of RAD51. Error bars represent ± s.e.m. across three replicates. The same siRNA transfected or shRNA transduced cells were used for (e-g) and Figure 20a-b. **(i)** DR-U2OS cells transduced with the indicated FHA-tagged cDNAs and then transfected with the indicated siRNAs were prepared for western blot analysis with the indicated antibodies and assayed for HR efficiency. Two western blots of the same extracts are presented (panels are grouped accordingly). HR assay results are presented normalized to corresponding siFF controls from cells with the same cDNA. Error bars represent ± s.d. across three replicates. **(j)** DR-U2OS cells transduced with the indicated FHA-tagged cDNAs and then transfected with the indicated siRNAs were prepared for western blot analysis with the indicated antibodies and submitted to cell cycle analysis. Cell cycle data represent the mean of two replicates.

Figure 17 (continued)



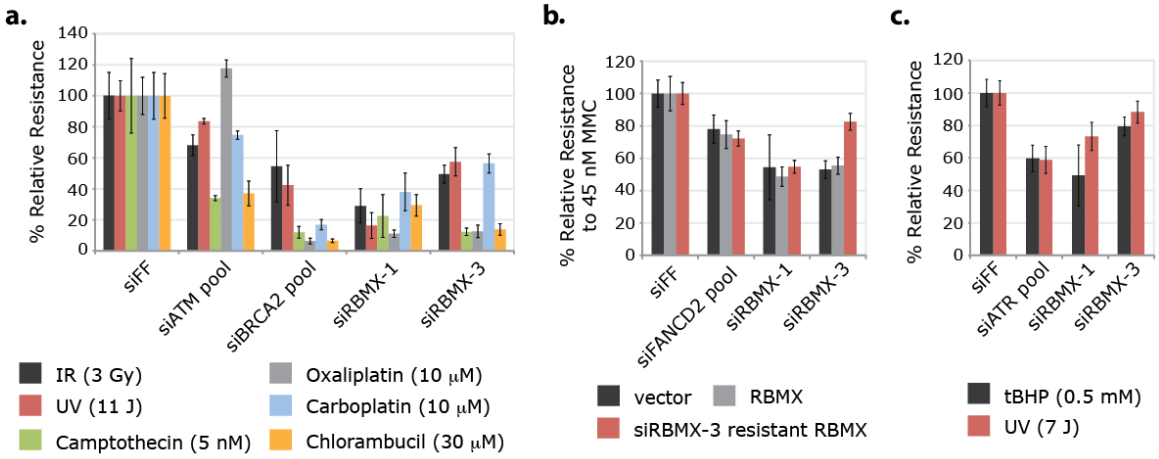


Figure 18. RBMX promotes resistance to various DNA damaging agents. (a) Multicolor competition assay for resistance to DNA damaging agents{Smogorzewska, 2007 #2}. Briefly, U2OS cells expressing GFP and transfected with the indicated siRNAs were mixed in a 1:1 ratio with dsRed U2OS cells transfected with siFF. Cell mixtures were treated with the indicated DNA damaging agents, and after 8 days, the relative survival of GFP to dsRed cells was determined by FACS analysis. GFP / dsRed ratios were normalized to those from undamaged cells to control for relative growth effects of the siRNAs. Error bars represent ± s.d. across three replicates. **(b)** DR-U2OS cells transduced with the indicated FHA-tagged cDNAs and then transfected with the indicated siRNAs were treated with 45 nM MMC, and after 7 days, the relative resistance as compared to untreated controls was determined by the CellTiter-Glo Luminescent Cell Viability Assay. Error bars represent ± s.d. across three replicates. **(c)** U2OS cells transfected with the indicated siRNAs were treated with tert-butyl hydroperoxide (tBHP) for 1 hour or UV, and after 5 days, the relative resistance as compared to untreated controls was determined by the CellTiter-Glo Luminescent Cell Viability Assay. Error bars represent ± s.d. across three replicates. Data from (a-c) are presented normalized to that from control transfected cells (siFF).

2-3. Structure-function analysis of RBMX

To determine the region(s) of RBMX responsible for promoting HR and facilitating localization to DNA damage, we tested a series of GFP- and FHA-tagged RBMX mutants (Figure 19). RBMX is composed of four identifiable regions: an N-terminal RNA recognition motif (RRM), a centrally located RBM1CTR region identified as common among RBMY-like hnRNPs, a C-terminus rich in serine, arginine, gylcine, and tyrosine residues, and a putative second RNA-binding domain at the C-terminal end (C-RBD) (Kanhoush et al., 2009). The canonical RRM of RBMX preferentially binds CC(A/C)-rich single-stranded RNA, and although RBMX can influence alternative splicing in an RRM-independent manner, evidence suggests that some splicing is directly facilitated through the RRM (Heinrich et al., 2009).

Deletion and point mutation of the conserved nucleotide-interacting residues in the RRM rendered the siRNA-resistant FHA-RBMX unable to rescue HR comparable to wild-type siRNA-resistant FHA-RBMX at the indicated levels of expression (Figure 19a-d); however, when evaluated at higher expression levels the deletion mutant showed better rescue of HR, indicating that this and other mutants may not be completely defective (data not shown). The RRM was dispensable for DNA damage localization, and two non-overlapping fragments of the amino and carboxy termini localized to tracks of DNA damage (Figure 19a). These data suggests that two distinct regions outside the RRM mediate RBMX recruitment to damage.

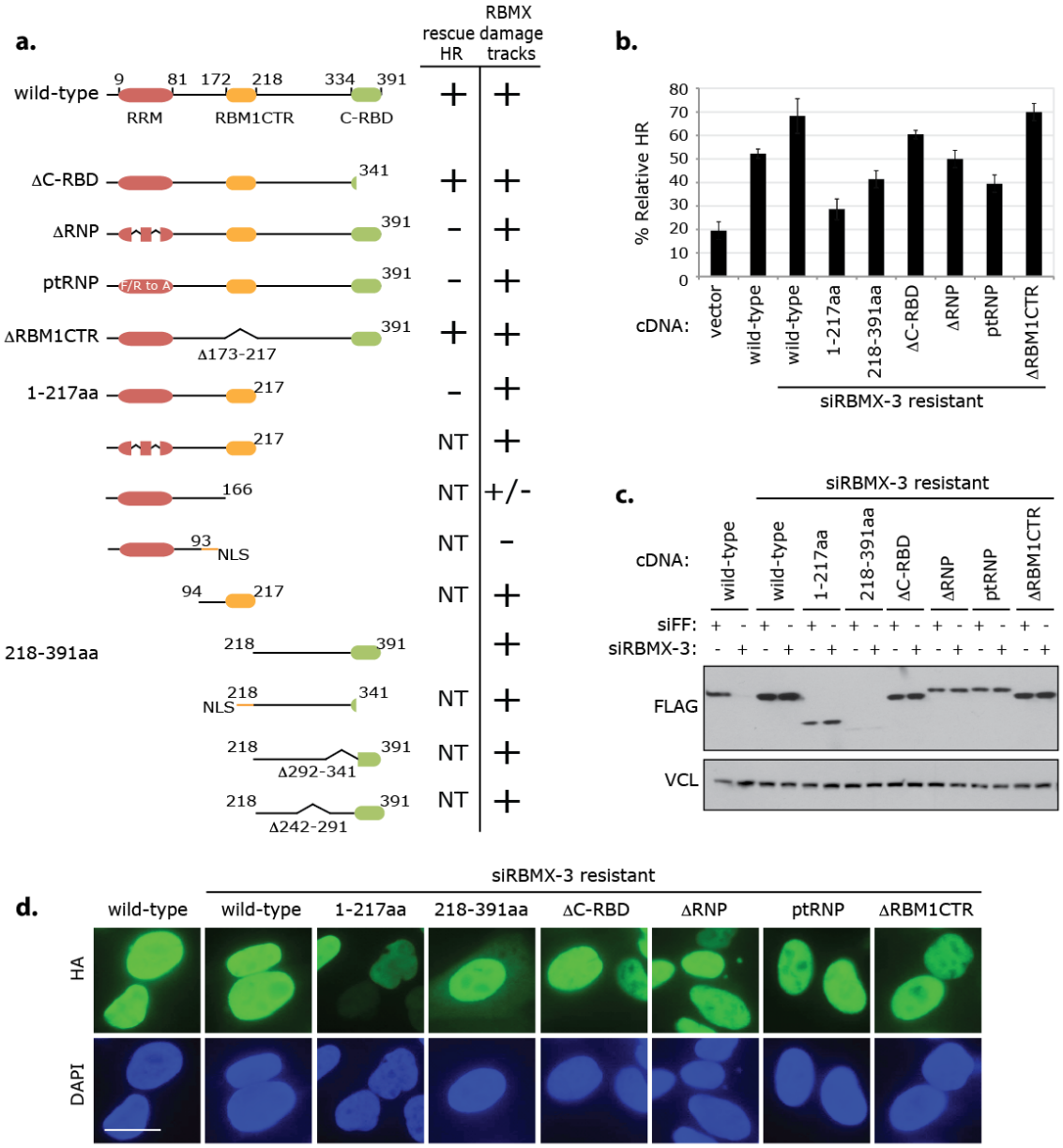


Figure 19. Structure-function analysis of RBMX. (a) A graphical representation of RBMX mutants and their corresponding phenotypes. Flag/HA-tagged mutants were evaluated for their capacity to attenuate defective HR caused by depletion of endogenous RBMX (b-d). GFP-tagged mutants were evaluated for localization to DNA damage after microirradiation (data is from 12 separate experiments preformed with various subgrouping of the mutants). NT indicates not tested. The ptRNP mutant carries point mutations of four aromatic residues important for the RNA binding of conserved RNA recognition motifs (RRMs): F11A, R49A, F51A and F53A. The ΔRNP mutant was deleted for two ribonucleoprotein (RNP) domains (amino acids 10-15 and 49-56) containing these conserved residues. Nuclear localization of all GFP-RBMX mutants except for the 1-93 amino acid truncation was observed. An NLS was added to the 1-93 and 218-341 amino acid fragments to aid nuclear localization; 218-341 GFP tracks were rarely observed without the NLS. +/- indicates that GFP damage tracks were rarely observed without depletion of endogenous RBMX. **(b)** DR-U2OS cells transduced with the indicated FHA-tagged cDNAs and then transfected with either siRBMX-3 or siFF were evaluated for HR efficiency. Data are presented normalized to corresponding siFF controls from cells with the same cDNA. Error bars represent ± s.d. across three replicates. **(c)** Whole-cell extracts from cells in (b) were immunoblotted with antibodies against FLAG and VCL (loading control) to confirm expression of RBMX (wild-type and mutant) transgenes. **(d)** Cells in (b) were processed for immunofluorescence with an antibody against HA to confirm nuclear localization of RBMX (wild-type and mutant) transgenes. Scale bar indicates 20 μm.

2-4. RBMX influences HR by facilitating proper expression of BRCA2

Next we evaluated the effect of RBMX depletion on known HR events, specifically RAD51 nucleation onto resected ssDNA and the coordinated upstream signaling. RBMX depletion caused defective formation of IR-induced RAD51 foci, which was attenuated by expression of siRNA-resistant FHA-RBMX (Figure 20a-e). Although the siRNAs used in these experiments (siRBMX-1 and -3) caused slight reductions to RAD51 protein levels (but not *RAD51* mRNA), the reduction caused by siRBMX-3 (unlike HR) was not rescued by siRNA-resistant FHA-RBMX (Figure 17d,h,j), indicating that there is a negligible, RBMX-independent effect of these reagents on RAD51 levels. RBMX was not required for RPA2 or CHK1 phosphorylation after IR or camptothecin treatment (Figure 21a-c), suggesting that in the absence of RBMX resection at breaks proceeds properly.

Next, we asked if RBMX localization to DNA damage is required for HR; and surprisingly, we found that HR efficiency was not decreased under PARP inhibition or depletion conditions that prevented RBMX accumulation at DNA damage (Figure 22a,b). PARP1 depletion also did not substantially alter HR in an RBMX-independent manner (Figure 22c).

As rapid recruitment of RBMX to DNA lesions was not important for HR, we reasoned that RBMX might promote HR by influencing protein expression through pre-requisite splicing events. While RBMX depletion had no effect on many repair proteins we tested, including PALB2, BRCA1 and RPA2 (and no apparent effect on BRCA1 foci formation) (Figure 23a-e), we found that levels of BRCA2 and to some extent ATR were decreased by RBMX siRNAs in a manner that could be rescued by expression of siRNA-resistant FHA-RBMX (Figure 24a-d). Because of these results, we evaluated siRNAs against seven additional pre-mRNA processing genes identified as candidate HR mediators by our screen and found that some of these also had an effect on BRCA2 expression (Figure 24e).

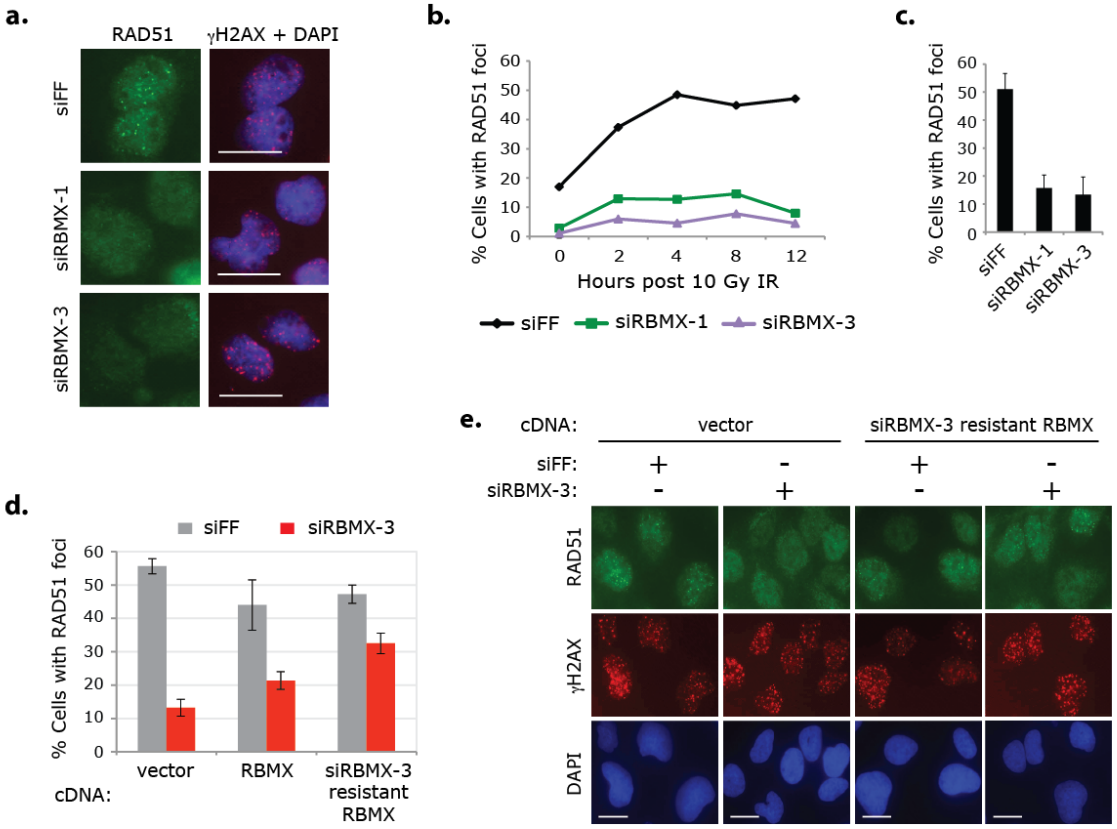


Figure 20. RBMX promotes formation of IR-induced RAD51 foci. (a) DR-U2OS cells transfected with the indicated siRNAs were treated with 10 Gy IR and processed for immunofluorescence with the indicated antibodies after 4 hours. Nuclei were stained with DAPI. Individual adjustment of color channels in γH2AX + DAPI was required to illustrate foci; identical adjustment parameters were used. Scale bars indicate 20 μm. Data from the same siRNA transfected cells are presented in (b) and Figure 17e-h. **(b)** DR-U2OS cells in (a) were processed for immunofluorescence with antibodies against RAD51 and γH2AX at the indicated times. The 0 hour time point was taken immediately after damage induction. The percentage of cells with RAD51 foci was determined by eye for each condition; because transfection with siRNAs against RBMX causes some changes to nuclear morphology, only normal shaped nuclei were included (~70-130 cells / condition). Data represent the mean of two replicates. Representative images of this experiment are presented (a). **(c)** DR-U2OS cells transfected with the indicated siRNAs were damaged and processed for immunofluorescence as described in (a). The percentage of cells with RAD51 foci was determined by eye for each condition, only normal shaped nuclei were included (~100-130 cells / condition). Error bars represent ± s.d. across four replicates. Data from the same siRNA transfected cells are presented in Figure 21b-c. **(d)** DR-U2OS cells transduced with the indicated FHA-tagged cDNAs were transfected with the indicated siRNAs and treated with 10 Gy IR; after 6 hours, cells were processed for immunofluorescence with antibodies against RAD51 and γH2AX and counted (200 cells / condition). Error bars represent ± s.d. across four replicates. **(e)** Representative images of cells in (d). Scale bars indicate 20 μm. Data from the same cells in (d-e) are presented in Figure 24c.

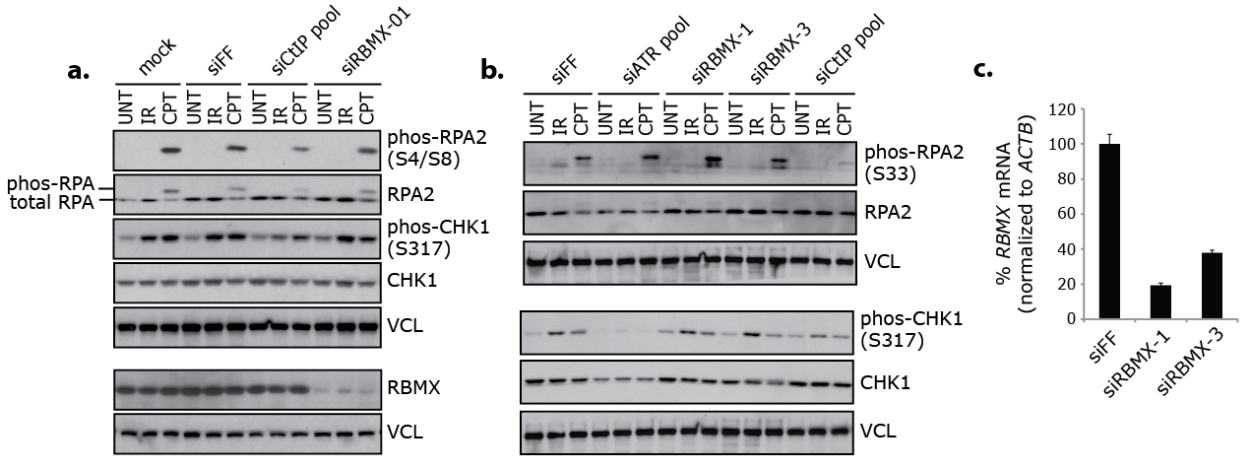
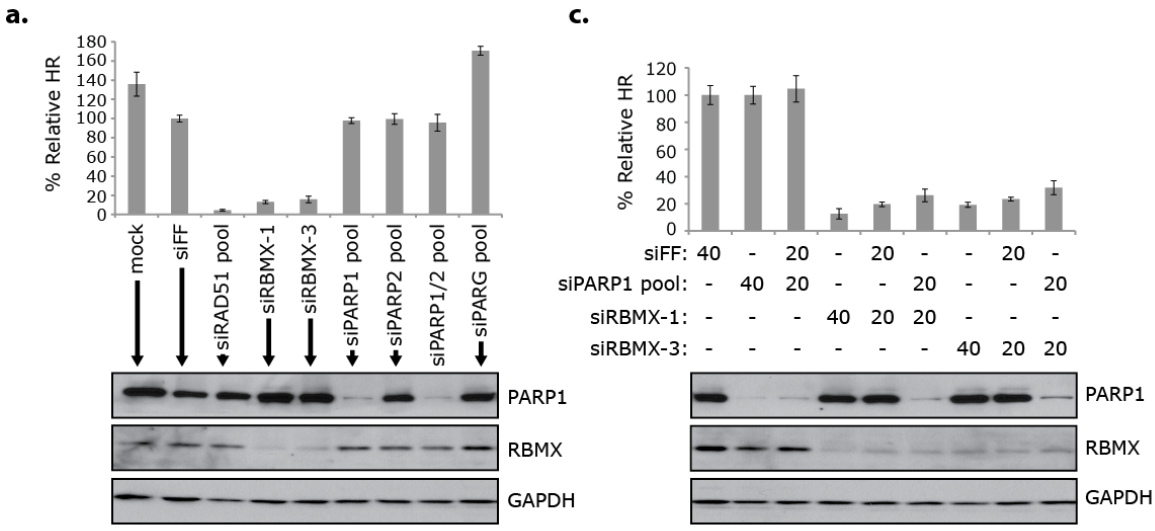


Figure 21. RBMX is not required for damage-induced RPA2 or CHK1 phosphorylation. (a-b) DR-U2OS cells transfected with the indicated siRNAs were treated with 10 Gy IR or 1 μM camptothecin (CPT) and after ~6.5 hours whole-cell extracts were collected and immunoblotted with the indicated antibodies. UNT indicates untreated. Data from (a) and (b) are from two independent experiments and two western blots of corresponding extracts are presented in both (panels are grouped accordingly). **(c)** RT-qPCR of RBMX (normalized to ACTB) from undamaged DR-U2OS cells evaluated in (b). Error bars represent ± s.e.m. across three replicates. Data from the same siRNA transfected cells in (b-c) is presented in Figure 20c.



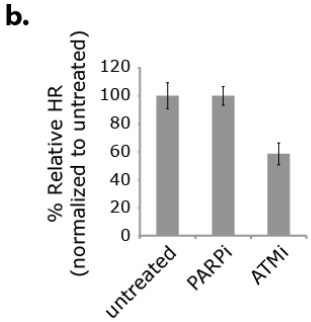


Figure 22. Neither inhibition nor depletion of PARP1 affects HR. (a) HR assay results and corresponding PARP1 / RBMX western blot analysis from DR-U2OS cells transfected with the indicated siRNAs at 20 nM. siPARP1/2 pool indicates a combined pool of siRNAs targeting PARP1 and PARP2 (4 duplexes each). Error bars represent ± s.d. across three replicates. **(b)** HR efficiency in DR-U2OS cells treated with a PARP inhibitor (PARPi), ATM inhibitor (ATMi), or DMSO (untreated, negative control). Inhibitors were added to cells at the same time as AdNGUS24i and again 12 hours later; cells were submitted to FACS analysis 36 hours post infection. Error bars represent ± s.d. across three replicates. **(c)** HR assay results and corresponding PARP1 / RBMX western blot analysis from DR-U2OS cells transfected with the indicated siRNAs at the indicated concentrations (20 or 40 nM). Error bars represent ± s.d. across three replicates.

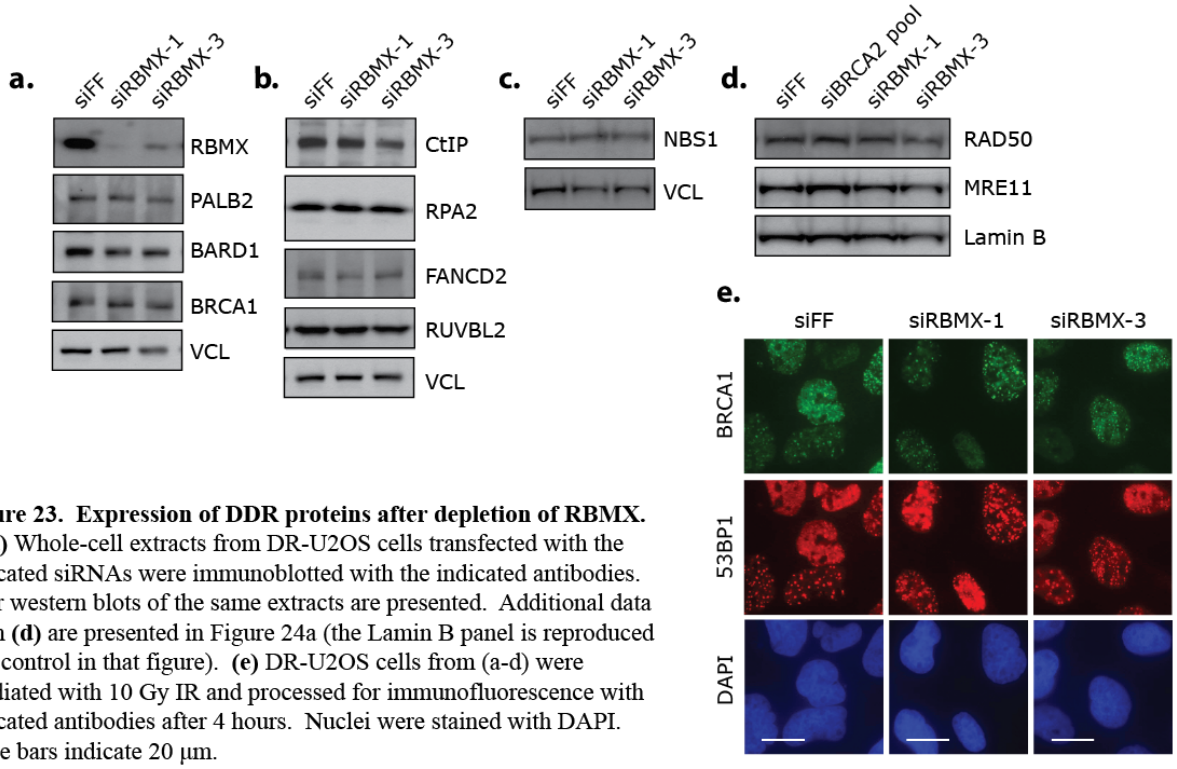


Figure 23. Expression of DDR proteins after depletion of RBMX. **(a-d)** Whole-cell extracts from DR-U2OS cells transfected with the indicated siRNAs were immunoblotted with the indicated antibodies. Four western blots of the same extracts are presented. Additional data from **(d)** are presented in Figure 24a (the Lamin B panel is reproduced as a control in that figure). **(e)** DR-U2OS cells from (a-d) were irradiated with 10 Gy IR and processed for immunofluorescence with indicated antibodies after 4 hours. Nuclei were stained with DAPI. Scale bars indicate 20 μm.

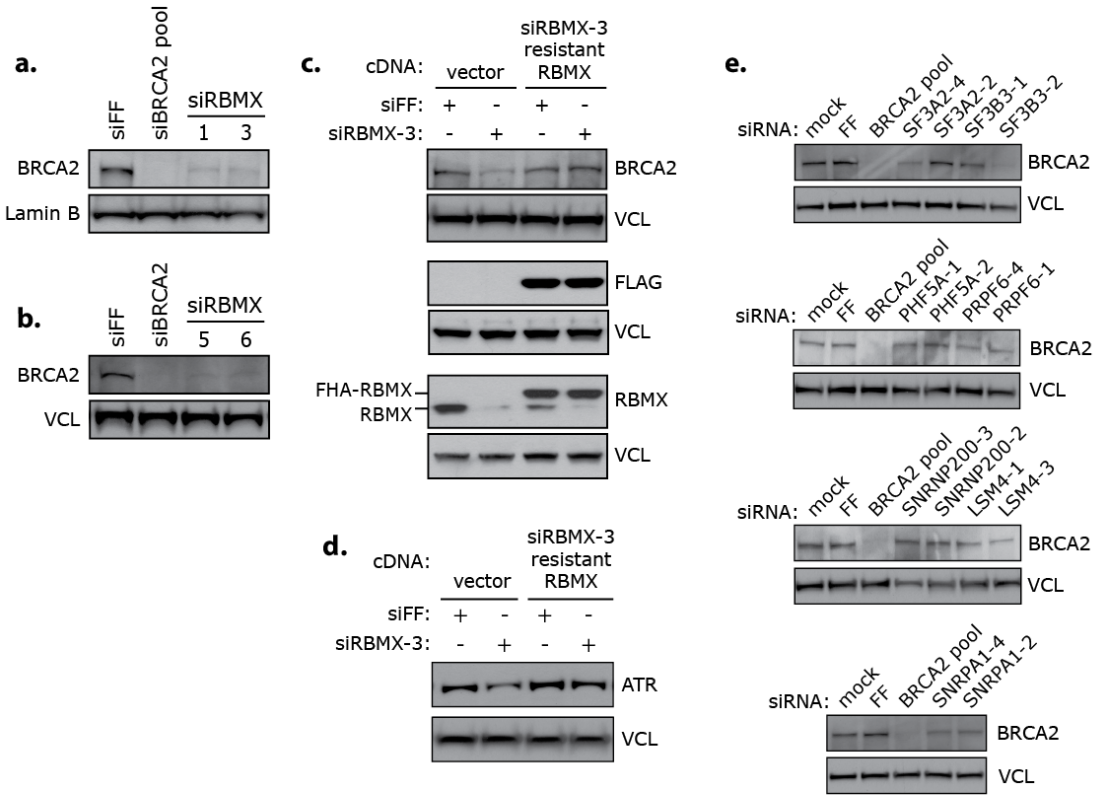


Figure 24. RBMX facilitates proper expression of BRCA2. (a-b) Whole-cell extracts from DR-U2OS cells transfected with the indicated siRNAs were immunoblotted with the indicated antibodies. **(a)** Extracts were also evaluated in Figure 23a-d. **(b)** Extracts were also evaluated in Figure 17d. **(c-d)** Whole-cell extracts from DR-U2OS cells transduced with the indicated FHA-tagged cDNAs and then transfected with the indicated siRNAs were immunoblotted with the indicated antibodies. Four western blots of the same extracts are presented (panels are grouped accordingly). Data from the same siRNA transfected cells are presented in Figure 20d-e. **(e)** Whole-cell extracts from cells transfected with the indicated siRNAs were immunoblotted with the indicated antibodies.

III. Discussion

The enrichment of RNA-processing factors among our candidate HR mediators and the identification of similar proteins in other large-scale studies of the DDR have made it difficult to ignore the notion that RNA processing and DNA repair functionally intersect (Beli et al., 2012; Hurov et al., 2010; Lackner et al., 2011; Matsuoka et al., 2007; Paulsen et al., 2009). Illustrative of the challenges inherent to evaluating the relationships between these two processes, we have found that RBMX indirectly regulates HR through an indirect (yet perhaps equally important mechanism), the expression of BRCA2. Our characterization of RBMX accumulation at sites of DNA damage, however, has successfully established a direct role for this RNA-binding protein (RBP) and splicing regulator in the DDR, adding to a rapidly expanding list of RBPs that have been similarly characterized.

In addition to RBMX, several RNA-processing and transcription-associated proteins have been shown to accumulate at sites of DNA damage within the last year, including two hnRNPU-like proteins (UL1 and UL2) (Polo et al., 2012), PPM1G (Beli et al., 2012), and Sp1 (Beishline et al., 2012). Similar to RBMX, UL1 and UL2 show two distinct localization patterns after DNA damage: (1) transient recruitment to sites of damage and (2) sustained exclusion from damaged regions (Polo et al., 2012). These proteins contain SAP (SAF-A/B, Acinus and PIAS) motifs, SPRY (SPIa / Ryanodine receptor) domains, an NK (putative nucleosides / nucleotide kinase) domain or BBS (BRD7-binding) site, and RGG (arginine and glycine-rich, RNA and ssDNA-binding) domains. They have protein sequence homology to hnRNPU, a heterogenous nuclear ribonucleoprotein that binds RNA through its RGG domain (Kiledjian and Dreyfuss, 1992). In general, hnRNPs are a loosely defined family of proteins categorized by common associations with nascent RNA polymerase II (RNAPII) transcripts and the spliceosome (Han et al., 2010; Rappsilber et al., 2002). Of note, RBMX is also an hnRNP family member (hnRNP G).

hnRNPUL exclusion from DNA damage was also correlated with local transcriptional suppression (similar to RBMX anti-stripes) and, interestingly, appeared to be RNA-binding dependent to

some extent (Polo et al., 2012). UL1 and UL2 accumulation at sites of DNA damage was enhanced by treatment with a Cdk-9 inhibitor (5,6-dichloro-1-b-D-ribofuranosylbenzimidazole or DRB) that blocks RNAPII transcriptional elongation and, like RBMX recruitment, was both RNA-binding and cell cycle independent. The dual recruitment patterns shown by both RBMX and the hnRNPULs therefore likely represent transcriptional silencing (anti-stripes) on the one hand and a transcriptionally-independent process functionally associated with DNA repair (stripes) on the other.

Interestingly, DRB acts by blocking RNAPII CTD hyperphosphorylation and has been shown, in at least one instance, to block UV-induced alternative splicing without affecting alternative splicing *per se* (Munoz et al., 2009b). From this an interesting hypothesis was proposed, that UV exposure may induce specific patterns of CTD heptapeptide phosphorylation (Phatnani and Greenleaf, 2006) that partially direct DNA damage-induced alternative splicing events (Munoz et al., 2009b). A tempting hypothesis relevant to our work then is that a subset of partially phosphorylated RNAPII stalled near UV laser-induced DNA lesions may coordinate the temporary recruitment or retention of specific cotranscriptional RNA-processing factors like RBMX to coordinate the initial steps of DNA repair. This would occur prior to anti-stripe formation and bulk RNAPII hyperphosphorylation and degradation that occurs after UV damage. Interestingly, coprecipitation and colocalization experiments have shown that splicing factors in general can associate with transcriptionally inactive and phosphorylated RNAPII through the CTD, even in the absence of pre-mRNA (Kim et al., 1997; Misteli and Spector, 1999; Phatnani and Greenleaf, 2006). Methods of distinguishing RNAPII CTD phosphorylation patterns may help address this hypothesis. Other possible roles for RBMX at sites of DNA damage are discussed in Chapter 5, 1-1.

Unlike RBMX, UL1 and UL2 recruitment to DNA damage was not reported to be PARP-dependent. These proteins were, however, shown to interact with components of the MRN complex and recruitment was shown to be largely MRN-dependent (Polo et al., 2012). Using rescue assays, hnRNPUL depletion was convincingly shown to cause phenotypes indicative of defective ATR signaling, HR, and resection. A model for hnRNPUL function in these processes was proposed wherein UL1 and UL2 promote resection through recruitment of the BLM helicase to DNA damage (Polo et al., 2012). This

conclusion hinged on three pieces of additional data: (1) UL1 coimmunoprecipitates with BLM, (2) depletion of UL1 / UL2 causes defects in BLM recruitment to sites of damage but does not reduce BLM expression and (3) co-depletion of UL1 and EXO1 but not UL1 and BLM increases cell sensitivity to CPT. Although this is an intriguing and well-supported model, the possibility that UL1 and UL2 mediate the essential splicing of a separate factor necessary for BLM recruitment has not yet been ruled out. This is particularly relevant in light of our finding that RBMX plays a role in promoting HR through BRCA2 expression. Of note, depletion of the BLM-TOP3-RMI1-RMI2 complex components RMI1 and RMI2 are known to disrupt BLM recruitment to damage and can be depleted in manner that has little to no effect on BLM expression levels (Singh et al., 2008; Xu et al., 2008; Yin et al., 2005).

Because RBMX appears to influence normal expression of BRCA2 and to some extent ATR, and because we found that siRNAs targeting other RNA-processing proteins also affect BRCA2 levels (albeit less significantly), we reason that depletion of splicing factors in general might cause phenotypes associated with DNA repair through aberrant essential splicing or alternative splicing of important repair effectors. A key future experiment will be to determine whether or not the observed HR defect caused by RBMX depletion is due wholly to altered *BRCA2* expression and / or splicing. This could be done by evaluating HR after RBMX depletion from cells expressing *BRCA2* from a cDNA construct.

Of note, *BRCA2* and *ATR*, like other large DNA damage response genes, contain many exons. According to published sequences, *BRCA2* (GenBank accession #U43746) has 27 exons (Tavtigian et al., 1996) and *ATR* (NCBI reference sequence: NM_001184.3) has 47; in contrast, human protein-coding genes have been reported to contain an average of 8.8 and a median of 7 exons (Lander et al., 2001). This high exon content may render expression of *BRCA2* and *ATR* particularly susceptible to misregulated splicing. Overall, we suggest that the enrichment of RNA processing proteins observed in our candidate list, as well as those from similar screens, may represent indirect but biologically significant components of DNA repair. The splicing function of RBMX, with particular attention to the possible effects of misregulated *BRCA2* and *ATR* splicing, is discussed in greater detail in Chapter 5, 1-2.

IV. Materials and Methods

4-1. Cell culture

Human U2OS and DR-U2OS osteosarcoma cells and mouse embryonic fibroblasts were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 units / ml of penicillin, and 0.1 mg / ml streptomycin (Invitrogen). Cell selection after viral transduction was conducted with puromycin at 1 ug / ml.

4-2. Plasmids, shRNAs, and siRNAs

Additional cDNAs (RBMY, hnRNP-K, hnRNP-C, and histone H3) were from hORFeome V5.1. Full-length histone H3 and the 5' ends of RBMY, hnRNP-K, and hnRNP-C were verified by sequencing. RBMX mutants with point mutations and internal deletions were generated using QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). Truncation fragments were generated by PCR and cloned into the pENTR™ / D-TOPO vector using pENTR™ Directional TOPO® Cloning Kit (Invitrogen). Mutations and truncations were verified by sequencing. The siRBMX-3 resistant RBMX cDNA was made by mutating the siRBMX-3 complementary site 5'-CAAGTTCTCGTGATACTAG to 5'-CTTCCAGCAGAGACACCCG. cDNAs were cloned into pMSCV-N-HA-Flag-GAW-IRES-PURO or pMSCV-N-EGFP-GAW-PGK-PURO for expression using the Gateway recombination system. shRNAs were used in the pSMP-MSCV-PURO vector (Open Biosystems). siRNA transfection was done as described in Chapter 3, 4-2. shRNA and siRNAs sequences used in this chapter that are not listed in Table 2 or Tables S2-4 are listed in Table 4.

Table 4. Additional RNAi Reagents

Target	Reagent Type	Source	Clone ID / Catalog Number	Sequence
PARP1	siRNA pool	Dharmacon	D-006656-02	GAAAGUGUGUUCAACUAAU
			D-006656-03	GCAACAAACUGGAACAGAU
			D-006656-04	GAAGUCAUCGAUAUCUUUA
			D-006656-17	GAUAGAGCGUGAAGGCGAA
PARP2	siRNA pool	Dharmacon	D-010127-13	AAGGAUUGCUCAAGGUAA
			D-010127-14	ACAGCUAGAUCUUCGGGUA
			D-010127-15	GCCAGAGACAGGAGUCGAA
			D-010127-16	ACAAUUGGGAAGAUCGAGA
PARG	siRNA pool	Dharmacon	D-011488-21	UGAGCUGUCAGGUGUUAAUA
			D-011488-22	CUGAGGAGCCGAGCGAAUA
			D-011488-23	GCAGUUGUCAGUUGGUACA
			D-011488-24	AUAAGCUGUUGCUACGAUA
FANCD2	siRNA pool	Dharmacon	D-016376-01	GGUCAGAGCUGUAUUAUUCUU
			D-016376-02	GAUAAGUUGUCGUCUAUUAUU
			D-016376-03	GCAGAACUUUGCCUACUUAUU
			D-016376-04	GAUCAACUCUCCUAAAGAUUU
RBMX-1	siRNA	Dharmacon	D-011691-01	UAUGGUAACUCACGUAGUG
RBMX-2	siRNA	Dharmacon	D-011691-02	CGAUAGAGAUGGAUAUGGU
RBMX-3	siRNA	Dharmacon	D-011691-03	CAAGUUCUCGUGAUACUAG
RBMX-4	siRNA	Dharmacon	D-011691-04	GUGGAAGUCGAGACAGUUA
RBMX-5	siRNA	Ambion	s26143	UCCCAUUCCACUGCUACUGcg
RBMX-6	siRNA	Ambion	s26144	UUGAUGGAUAGUCAUCACGtg
ATM	shRNA	Open Biosystems	V2HS_89366; RHS1764-9217463	TTAAATGACTGTATAGTCACCA
RBMX-7	shRNA	Open Biosystems	V3LHS_339669; RHS4430-101070238	TTTCTTGTCTGCCAACCCGATC
RBMX-8	shRNA	Open Biosystems	V3LHS_339674; RHS4430-101073396	TCTCTATCGCTATATCCTCTTG
RBMX-9	shRNA	Open Biosystems	V3LHS_339670; RHS4430-101075016	TATCCGTCACGTGAGCTGCTGT
RBMX-10	shRNA	Open Biosystems	V3LHS_645229; RHS4430-99890017	TTTTGTTTCTTTGAACTGGGAT

4-3. Antibodies and Inhibitors

Additional primary antibodies used in this chapter are listed in Table 5. Secondary antibodies used were previously described in Chapter 3, 4-3. The PARP inhibitor (KU-0058948, KuDOS Pharmaceuticals Ltd.) was used at 1 μM; this reagent was provided by Simon Boulton (Cancer Research UK, London Research Institute). The ATM inhibitor (KU-55933, Sigma) was used at 10 μM.

Table 5. Additional Primary Antibodies.

Antibody	Host Name	Source	Name / Catalog #
RBMX	goat	Santa Cruz	sc-14581
RAD51	rabbit	Santa Cruz	sc-8349
BRCA2	rabbit	Bethyl	A300-005A
BRCA2	mouse	Calbiochem	OP95
PALB2	rabbit	Bethyl	A301-246A
BARD1	rabbit	Bethyl	A300-263A
BRCA1	mouse	Calbiochem	OP92
BRCA1	mouse	Santa Cruz	sc-6954
CtIP	rabbit	Bethyl	A300-488A
FANCD2	rabbit	Novus	NB100-182
RUVBL2	rabbit	Bethyl	A302-536A
NBS1	rabbit	Novus	NB100-143
RAD50	rabbit	Novus	NB100-154
MRE11	rabbit	Novus	NB100-142
ATR	goat	Santa Cruz	sc-1887
53BP1	rabbit	Bethyl	A300-272A
RPA32 / RPA2	rabbit	Bethyl	A300-244A
RPA32 / RPA2	mouse	Abcam	ab16855
Phospho RPA32-S4/8	rabbit	Bethyl	A300-245A
Phospho RPA32-S33	rabbit	Bethyl	A300-246A
Phospho CHK1-Ser317	rabbit	Cell Signaling	2344
CHK1	mouse	Santa Cruz	sc-8408
PARP	rabbit	Cell Signaling	9542

Table 5 (continued). Additional Primary Antibodies.

Antibody	Host Name	Source	Name / Catalog #
PAR	mouse	Trevigen	4335-AMC-050
γH2AX	rabbit	Bethyl	A300-081A
γH2AX	mouse	Millipore	05-636
FLAG M2-Peroxidase HRP	mouse	Sigma	A8592
HA	mouse	Covance	MMS-101R
GAPDH	rabbit	Santa Cruz	sc-25778
Lamin B	goat	Santa Cruz	sc-6216
VINCULIN	mouse	Sigma	V9131

4-4. UV laser- and IR-induced DNA damage and immunofluorescence

UV laser-induced damage was generated as described in Chapter 3, 4-4. Ionizing irradiation (IR)-induced damage was generated by timed exposure to a Cesium-137 source. After damage, cells were allowed to recover for the indicated times at either room temperature (RT) or 37°C and then fixed, permeabilized, and immunostained as described in Chapter 3, 4-4. DNA was stained with DAPI by addition of Vectashield Mounting Medium (Vector Laboratories). GFP was observed directly. Images were collected on an Axioplan2 Zeiss microscope with an AxioCam MRM Zeiss digital camera and Axiovision 4.5-4.8 software. Images intended for comparison were prepared from the same experiment with the same exposure times, and were processed for brightness and contrast in an identical manner. Those not intended for comparison and not prepared in this way are indicated.

4-5. HR Assay

This assay was performed as described in Chapter 3, 4-5.

4-6. RT-qPCR

RT-qPCR was performed as described in Chapter 3, 4-6. Additional RT-qPCR primers used were: RBMX left primer 5'-CAGTTCGCAGTAGCAGTGGA; RBMX right primer 5'-

TCGAGGTGGACCTCCATAA; RAD51 left primers 5'-GGGAATTAGTGAAGCCAAAGC and 5'-GCCAAAGCTGATAAAATTCTGAC; RAD51 right primers 5'-TGGTGAAACCCATTGGAACT and 5'-GGAGGGTGCAGTAGACCAAG.

4-7. Cell cycle analysis

Cells were prepared for cell cycle analysis using the BD Pharmingen™ APC BrdU Flow Kit according to manufacturer instructions. Cell cycle profiles were obtained by FACS analysis on a BD LSRII Flow Cytometer (BD Biosciences).

4-8. Sensitivity assays

Multicolor competition assays were performed as previously described (Smogorzewska et al., 2007). CellTiter-Glo Luminescent Cell Viability Assay (Promega) was performed as follows: Cells were transfected with the indicated siRNAs and treated as indicated after 2-3 days. After recovery from treatment, the media was changed and CellTiter-Glo reagent was added (1:17.5 dilution). The resulting luminescent signal (proportional to the amount of ATP) was read on a VICTOR X5 Multilabel Plate Reader (PerkinElmer). The signal from each treatment condition was normalized to an untreated control to adjust for relative growth effects of the siRNAs. Data from both sensitivity assays are presented normalized to that from control transfected cells (siFF).

Chapter Five

Conclusions and Perspectives

I. The significance of RBMX

1-1. The role of RBMX at sites of DNA damage

The biochemical consequences of transiently accumulating RBMX at sites of DNA damage remain to be determined. Although this recruitment is phenotypically similar to that observed for the hnRNPUL proteins, RBMX does not appear to be required for resection and RBMX recruitment is not required for HR, making it unlikely that RBMX participates in mediating BLM recruitment to DSBs, as has been shown for the hnRNPULs. However, because RBMX and the hnRNPULs are functionally related through pre-mRNA processing and associations with the spliceosome (Rappsilber et al., 2002), and because they have DNA-damage localization patterns in common, it is reasonable to predict that they also share some DDR-associated functionality at DNA breaks. To further investigate this, an analysis of the potential functional and physical interactions between RBMX and the hnRNPULs within the context of the DDR should be conducted.

As discussed above (Chapter 4, Section I), several models have emerged to explain the possible functions of pre-mRNA-processing proteins in the maintenance of genomic stability, and the role of RBMX at DNA breaks may well align with one of these. One possibility is that RBMX functions as an assembly factor to promote the accumulation of general repair proteins at DNA lesions, possibly as part of a novel sensor mechanism initiated by RNAPII. RBMX recruitment to or retention on RNAPII stalled at DNA lesions could repress aberrant R-loop formation through non-specific RNA binding of nascent transcripts mediated by the RBMX RNA recognition motif (RRM). In this role, RBMX would presumably act as part of a local and regulated R-loop suppression mechanism that is controlled by the DDR. The existence of this (or a similar) mechanism is partially warranted by the observation that the presence of nicks within non-template DNA strands during transcription favors RNA:DNA binding over DNA:DNA reannealing (Roy et al., 2010), suggesting that DNA damage can promote R-loop formation in *cis*.

Because RBMX localization to DNA damage is PARP-dependent, RBMX may also serve as an effector of PARP activity at DNA breaks. In this capacity, RBMX may promote aspects of single-strand break repair (SSBR), alternative non-homologous end-joining (a-NHEJ), or even the early recruitment of MRN and activation of ATM. It is also possible that the ability of RBMX to accumulate at sites of DNA damage through at least two independent domains could facilitate DSB repair by bundling PAR structures together and restricting the movement of DSB ends. Alternatively, RBMX may promote the localization of specific noncoding RNAs to DSBs to facilitate repair in some manner. Since NHEJ factors like Ku are also known to bind RNA, these could be involved in tertiary interactions (Yoo and Dynan, 1998). Bulk comparison of RNAs bound to RBMX in the presence and absence of DNA damage might be an informative approach to evaluating the role of RBMX RNA binding in the DDR.

While additional investigations into the role of RBMX at DNA breaks may prove informative, such future works will no doubt require that one particular challenge first be circumnavigated: the separation of RBMX depletion phenotypes, such as defective HR, from the consequences of failure to localize RBMX at regions of DNA damage. To this end, we have begun the work of isolating an RBMX mutant capable of restoring HR and BRCA2 expression but unable to accumulate at laser-induced DNA damage through the initial structure / function studies presented herein (Figure 19). Although our progress has been limited by the ability of RBMX to localize to sites of DNA damage through multiple independent domains, our results suggests that one such domain likely resides between amino acids 94 and 166.

1-2. The role of RBMX in homologous recombination

Although RBMX accumulates at DNA lesions in a regulated way, our results indicate that the function of RBMX in HR is most likely mediated through the indirect regulation of BRCA2, possibly due to one or more pre-requisite splicing events. RBMX depletion also slightly reduces ATR expression. As discussed above (Chapter 4, Section III), *BRCA2* and *ATR* are large genes that may depend heavily on proper essential splicing for normal expression. With this in mind, we caution future studies that aim to

evaluate mRNA-processing proteins using depletion phenotypes within the context of the DDR and suggest that such work be tightly coupled to in-depth analyses of gene expression to avoid confounding direct DDR-phenotypes with splicing-specific effects.

A second possibility is that RBMX coordinates *BRCA2* and *ATR* alternative splicing under certain conditions, especially as a role for RBMX in alternative splicing (exon inclusion and exclusion) has been well established but a role in essential splicing is less clear (Heinrich et al., 2009; Hofmann and Wirth, 2002; Nasim et al., 2003; Wang et al., 2004). Consistent with this, we have observed that RBMX depletion does not affect the expression of several other proteins (Figure 23a-d). Additionally, the effect of RBMX on splice-site selection appears to be concentration dependent (Heinrich et al., 2009; Hofmann and Wirth, 2002) and RBMX expression has been shown to vary between tissue types in both humans (Nasim et al., 2003) and rats (Heinrich et al., 2009). This suggests that differential RBMX expression may control tissue-specific alternative splicing programs, a hypothesis entirely consistent with the fact that alternative splicing is known to vary between cell and tissue types in general (Stamm et al., 2005).

Alternatively spliced isoforms of *BRCA2* have been reported. A splice variant lacking exon 12 was identified in normal human tissue (Bieche and Lidereau, 1999), and within the Ensembl project database (www.ensembl.org) *BRCA2* is annotated with 6 isoforms, 3 of which are described as protein coding (*ATR* is annotated with 12 isoforms) (Flicek et al., 2013). In general, the alternative splicing of protein-coding transcripts can modulate the expression, structure, function, and modification of transcribed proteins, and as such, the successful translation of any alternatively spliced *BRCA2* and *ATR* transcripts may yield functional consequences. This becomes particularly relevant in light of the fact that germline *BRCA2* and *ATR* loss-of-function mutations are associated with familial breast cancer and Seckel syndrome, respectively (Ciccia and Elledge, 2010). Reports by several groups suggest that among these mutant genes are germline *BRCA2* variants that cause aberrant splicing (Bieche and Lidereau, 1999; Hofmann et al., 2003; Miki et al., 1996; Pensabene et al., 2009; Sanz et al., 2010).

Further investigation into the possibility that *BRCA2* is alternative spliced (both normally and pathologically) and into the splicing factors that may regulate such events is warranted. Once identified,

several approaches may be taken to evaluate and characterize the possible role of RBMX in *BRCA2* and / or *ATR* alternative splicing. These including large-scale evaluation of RBMX bound RNAs or comparison of the exon inclusion patterns of *BRCA2* and *ATR* transcripts in the presence and absence of RBMX. According to publically available data sets RBMX is not frequently mutated in cancer (Cerami et al., 2012); however, more extensive mining of cancer genome databases for RBMX alterations may be of interest.

II. Perspectives on RNAi screening and off-target effects

One unique and unforeseen contribution of our screening work has been the identification of *RAD51* as a prominently off-targeted transcript. Over the last several years, it has become increasingly apparent that poor reagent specificity is a major concern for RNAi-based screens (Chapter 3, 1-1), and our finding that pervasive off-targeting of a single transcript can confound ≥17% of candidate siRNAs serves as motivating example in favor of better reagent design. To this end, we have made available the list of siRNA reagents here identified to contain *RAD51* off-target effects (Table S2), and we note that computational mining of the associated sequences may be a productive means for identifying new sequence-based modes of mRNA recognition that contribute to RNAi off-targeting (akin to seed-3'UTR pairing). Thus far, we have eliminated the possibility that seed pairing to transcript coding regions (CDSs) contributes to predominant off-target effects, but the possibility remains that additional mechanisms of unintended *RAD51* pairing, mediated perhaps by incomplete seeds containing G-bulges or aided by 3'compensatory sites, may exist within our data set. Consistent with this idea, we have observed that some siRNAs lacking complete antisense seed matches to the *RAD51* 3'UTR affect RAD51 expression in an off-target manner (Chapter 3, 2-1).

In a separate work by Sigoillot et al., development of the GESS algorithm was published alongside an independent set of siRNA reagents that contain a pervasive off-target effect against the *MAD2* transcript in human cells (Sigoillot et al., 2011). MAD2 is a key mediator of the spindle assembly

checkpoint (SAC) and Sigoillot et al. identified *MAD2* off-target effects in two independent siRNA-based screens for mediators of SAC (Sigoillot et al., 2011; Tsui et al., 2009). This work also used GESS to identify a set of siRNAs that off-target the TGFβ receptor 2 transcript through seed complementarities from yet another siRNA-based screen (Schultz et al., 2011). Along with our data, results from these studies may prove useful in identifying novel sequence modes of off-target transcript pairing.

Interestingly, evaluation of RNAi off-target effects and their impact on high-throughput screening (in work by Sigoillot et al. and in general) has primarily focused on siRNA- and dsRNA-based work in mammalian and *Drosophila* cells (Perrimon and Mathey-Prevot, 2007; Sigoillot and King, 2011). (In *Drosophila* long dsRNAs typically ~500-nt in length can be used for RNAi-mediated silencing. These long dsRNAs are processed endogenously to generate shorter 20-22-nt dsRNAs that ultimately mediate mRNA knockdown.) Nevertheless, miRNA-like off-target effects can impact shRNA technologies as well (Jackson et al., 2006), and shRNA-based screens should be interpreted with similar caution. Of note however, off-targets may be controlled to some extent by viral transduction of shRNA reagents because this delivery technique allows shRNA expression to be achieved at relatively low copy numbers.

In general, the identification of several transcripts that are prominently off-targeted by RNAi argues that the development of off-target identification algorithms like GESS should be prioritized and that, once developed, they should be widely implemented into all RNAi screening procedures, regardless reagent-type. These efforts would be greatly aided by more in-depth and consistent reporting of RNAi screen results, including a move towards the full publication of reagent sequences. Increased availability to this type of information would substantially contribute to meta-analyses aimed at evaluating RNAi reagent behavior, both on- and off-target, and information gained from these analyses could be incorporated into RNAi reagent design to help limit confounding effects (Anderson et al., 2008; Sigoillot and King, 2011).

Recently, much thought has also been given to the design and implementation of new screening approaches that allow more accurate interpretation of RNAi results without directly improving RNAi reagent specificity, including strategies that incorporate increased reagent redundancy or improved library

sensitivity. Our work demonstrates that the incorporation of high reagent redundancy into screen design can successfully enrich for true positives without eliminating off-target effects *per se*. Specifically, we found that large-scale candidate validation using independent Ambion reagents (Chapter 3, 3-1) successfully eliminated false-positives that had initially been selected because of *RAD51* off-target effects. From this, we extrapolate that unidentified off-target effects were also circumvented using this approach. A second strategy for improving candidate selection from RNAi screens is the improvement of RNAi library sensitivity. To this end, the construction of genome-wide RNAi libraries that contain reagents with experimentally-indicated on-target depletion using approaches such as the recently developed sensor assay (Fellmann et al., 2011) will be incredibly useful. However, until such libraries become widely available the best strategies for successful screening will incorporate both high reagent redundancy and sequence-based off-target analyses.

The set of siRNAs among which pervasive *RAD51* off-targeting was initially identified had been selected by the *a priori* identification of a common phenotype; however through our studies of the HIRA-associated proteins, we have also observed that the interpretation of small-scale studies evaluating phenotypes caused by RNAi reagents chosen without selective pressure may also be confounded by *RAD51* off-targeting. In particular, of the 6 HIRIP3- and UBN1-targeting siRNAs chosen and evaluated separately from our screening work (not Ambion), 3 demonstrated off-target RAD51 depletion (Figure 12a-b,e,i,j). This argues that the individual validation of small-scale RNAi experiments is also paramount and highlights the necessity of using high reagent redundancy and rescue assays, or complementary approaches such as gene knockout or deletion systems, to rule out off-target effects during individual gene characterization.

Overall, we suggest that siRNA-based evaluations of the mammalian HR and SAC pathways be conducted and interpreted with a new level of rigor, as was done for our characterization of RBMX. Additionally, previous studies that may have been confounded by *RAD51* or *MAD2* off-targeting should be revisited. A spindle-checkpoint defect reported by our lab to be caused by siRNAs targeting the TAO1

kinase has in this way been shown to be the misleading result of *MAD2* off-targeting (Draviam et al., 2007; Hubner et al., 2010; Westhorpe et al., 2010).

Intriguingly, the identification of prominently off-targeted transcripts also presents the idea that some mRNA transcripts, like *RAD51*, are particularly sensitive to RNAi. It is tempting to speculate that some as yet unidentified structural element within the *RAD51* mRNA, perhaps in common with *MAD2* and *TGFβR2*, hypersensitizes *RAD51* to siRNA-mediated depletion, a miRNA enhancer of sorts. However, it is also possible that a shallow phenotypic threshold reached with minimal protein depletion more simply determines the heightened sensitivity of *RAD51* (and possibly *MAD2* and *TGFβR2*) to off-targeting. Consistent with this, the cooperativity of RAD51 filament assembly on ssDNA (as discussed in Chapter 1, 3-3-4) is expected to cause exquisite sensitivity to protein depletion, and the cell cycle regulated expression of RAD51 and MAD2 may render these proteins more susceptible overall to protein depletion, as compared to more stable proteins. Undoubtedly, however, a key aspect of *RAD51* and *MAD2* hypersensitivity to RNAi must be that both transcripts encode protein effectors that are critical to their pathways of inquiry.

III. Future genetic inquiry into the mammalian homologous recombination pathway

Although conserved mechanisms of HR have been comprehensively studied in model organisms for decades, we have successfully identified two lists of novel HR candidate genes (positive and negative) in mammalian cells (Table S1), demonstrating that regulators of HR (both direct and indirect) remain uncharacterized in this system. We note that during the course of our work many core facilitators of HR, including EXO1, DNA2, RMI2, GEN1, SLX1, SLX4 and RTEL1, have been newly identified and characterized using genetic approaches, indicating the breadth of HR-related discovery that remains ongoing. In this work, we have illustrated how redundancies in the HR pathway and a major *RAD51* off-target effect present unique challenges to genetic studies of mammalian HR. Nevertheless, we present our candidate lists as a foundation upon which future studies can build and suggest that the keys to future

RNAi-based inquiry into HR mechanisms will be the combined use of high reagent redundancy to limit off-targeting and analysis of synthetic interactions to parse redundancies.

Previously (Chapter 2, Section III) we proposed that the set of genes targeted by siRNA pools yielding intermediary phenotypes in our primary screen (1,678 siRNA pools = 50-70% relative HR) may also contain *bona fide* mediators of HR. However, to confidently identify novel HR effectors among this set using large-scale approaches, we estimate that a minimum of 10-20 reagents per gene would be necessary. While the labor and cost of this would prohibit use of the well-by-well screening format presented herein, we suggest that the DR-GFP reporter could be adapted for pooled shRNA screening, which is better suited to high reagent redundancy (discussed in Chapter 2, 1-2). Critical to this idea, the double-strand break that is induced in DR-GFP can be repaired in one of only four ways: (1) error-free NHEJ that reconstitutes the I-SceI recognition site, (2) error-prone NHEJ that introduces insertions or deletions at the I-SceI site, (3) HR or (4) SSA, which both replace I-SceI with sequence from wild-type GFP (Nakanishi et al., 2001). After PCR amplification, these repair events can be distinguished by sequence information at the break point; and because of this, phenotypic readouts from DR-GFP could theoretically be analyzed using a next generation sequencing (NGS)-based approach. Use of NGS data collection and an shRNA expression cassette carrying a tandem DR-GFP reporter could thus allow a simple and straightforward shRNA screen. This screen would proceed as follows: an shRNA library incorporated into the proposed DR-GFP cassette would be introduced into cells. These would be selected for stable integration of the cassette, infected with the I-SceI carrying adenovirus for 24-48 hours, and then collected. PCR preparation of the shRNA-DR-GFP cassette from isolated genomic DNA followed by massively parallel paired-end sequencing of the shRNA hairpin (on one end) and the I-SceI break point (on the other) could then generate measures of relative HR for 6,000-20,000 shRNAs in one experiment.

Our work provides a well-curated candidate list to guide future genetic studies of homologous recombination in mammalian cells and represents an important step forward in fully understanding the challenges that are specific to these endeavors. Altogether the biological insights presented herein will be

useful for exploring specific questions regarding HR regulation as well as developing new approaches towards that goal. Whatever the future of this work, however, the last century of study makes two things abundantly clear: our understanding of homologous recombination is not yet complete and our enthusiasm for inquiry into this complex mechanism is not yet exhausted.

References

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol *14*, 318-328.

Aguilar-Quesada, R., Munoz-Gamez, J.A., Martin-Oliva, D., Peralta, A., Valenzuela, M.T., Matinez-Romero, R., Quiles-Perez, R., Menissier-de Murcia, J., de Murcia, G., Ruiz de Almodovar, M.*, et al.* (2007). Interaction between ATM and PARP-1 in response to DNA damage and sensitization of ATM deficient cells through PARP inhibition. BMC Mol Biol *8*, 29.

Aguilera, A., and Garcia-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. Mol Cell *46*, 115-124.

Ahel, I., Ahel, D., Matsusaka, T., Clark, A.J., Pines, J., Boulton, S.J., and West, S.C. (2008). Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. Nature *451*, 81-85.

Ahn, J.Y., Schwarz, J.K., Piwnica-Worms, H., and Canman, C.E. (2000). Threonine 68 phosphorylation by ataxia telangiectasia mutated is required for efficient activation of Chk2 in response to ionizing radiation. Cancer Res *60*, 5934-5936.

Ame, J.C., Rolli, V., Schreiber, V., Niedergang, C., Apiou, F., Decker, P., Muller, S., Hoger, T., Menissier-de Murcia, J., and de Murcia, G. (1999). PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase. J Biol Chem *274*, 17860-17868.

Ame, J.C., Spenlehauer, C., and de Murcia, G. (2004). The PARP superfamily. Bioessays *26*, 882-893.

Anderson, E.M., Birmingham, A., Baskerville, S., Reynolds, A., Maksimova, E., Leake, D., Fedorov, Y., Karpilow, J., and Khvorova, A. (2008). Experimental validation of the importance of seed complement frequency to siRNA specificity. RNA *14*, 853-861.

Arata, H., Dupont, A., Mine-Hattab, J., Disseau, L., Renodon-Corniere, A., Takahashi, M., Viovy, J.L., and Cappello, G. (2009). Direct observation of twisting steps during Rad51 polymerization on DNA. Proc Natl Acad Sci U S A *106*, 19239-19244.

Asaithamby, A., and Chen, D.J. (2009). Cellular responses to DNA double-strand breaks after low-dose gamma-irradiation. Nucleic Acids Res *37*, 3912-3923.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T.*, et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25-29.

Audebert, M., Salles, B., and Calsou, P. (2004). Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. J Biol Chem *279*, 55117-55126.

Augustin, A., Spenlehauer, C., Dumond, H., Menissier-De Murcia, J., Piel, M., Schmit, A.C., Apiou, F., Vonesch, J.L., Kock, M., Bornens, M.*, et al.* (2003). PARP-3 localizes preferentially to the daughter centriole and interferes with the G1/S cell cycle progression. J Cell Sci *116*, 1551-1562.

Aylon, Y., Liefshitz, B., and Kupiec, M. (2004). The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. EMBO J *23*, 4868-4875.

Bakkenist, C.J., and Kastan, M.B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature *421*, 499-506.

Balaji, S., Iyer, L.M., and Aravind, L. (2009). HPC2 and ubinuclein define a novel family of histone chaperones conserved throughout eukaryotes. Mol Biosyst *5*, 269-275.

Ball, H.L., Ehrhardt, M.R., Mordes, D.A., Glick, G.G., Chazin, W.J., and Cortez, D. (2007). Function of a conserved checkpoint recruitment domain in ATRIP proteins. Mol Cell Biol *27*, 3367-3377.

Banumathy, G., Somaiah, N., Zhang, R., Tang, Y., Hoffmann, J., Andrake, M., Ceulemans, H., Schultz, D., Marmorstein, R., and Adams, P.D. (2009). Human UBN1 is an ortholog of yeast Hpc2p and has an essential role in the HIRA/ASF1a chromatin-remodeling pathway in senescent cells. Mol Cell Biol *29*, 758-770.

Barber, L.J., Youds, J.L., Ward, J.D., McIlwraith, M.J., O'Neil, N.J., Petalcorin, M.I., Martin, J.S., Collis, S.J., Cantor, S.B., Auclair, M.*, et al.* (2008). RTEL1 maintains genomic stability by suppressing homologous recombination. Cell *135*, 261-271.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bateson, W., Saunders, E.R., and Punnett , R.C. (1905). Experimental studies in the physiology of heredity. Reports to the Evolution Committee of the Royal Society, 1–55, 80–99.

Baumann, P., Benson, F.E., and West, S.C. (1996). Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. Cell *87*, 757-766.

Baumann, P., and West, S.C. (1997). The human Rad51 protein: polarity of strand transfer and stimulation by hRP-A. EMBO J *16*, 5198-5206.

Baumann, P., and West, S.C. (1999). Heteroduplex formation by human Rad51 protein: effects of DNA end-structure, hRP-A and hRad52. J Mol Biol *291*, 363-374.

Beishline, K., Kelly, C.M., Olofsson, B.A., Koduri, S., Emrich, J., Greenberg, R.A., and Azizkhan-Clifford, J. (2012). Sp1 Facilitates DNA Double-Strand Break Repair through a Nontranscriptional Mechanism. Mol Cell Biol *32*, 3790-3799.

Bekker-Jensen, S., Lukas, C., Kitagawa, R., Melander, F., Kastan, M.B., Bartek, J., and Lukas, J. (2006). Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. J Cell Biol *173*, 195-206.

Bekker-Jensen, S., Rendtlew Danielsen, J., Fugger, K., Gromova, I., Nerstedt, A., Lukas, C., Bartek, J., Lukas, J., and Mailand, N. (2010). HERC2 coordinates ubiquitin-dependent assembly of DNA repair factors on damaged chromosomes. Nat Cell Biol *12*, 80-86; sup pp 81-12.

Beli, P., Lukashchuk, N., Wagner, S.A., Weinert, B.T., Olsen, J.V., Baskcomb, L., Mann, M., Jackson, S.P., and Choudhary, C. (2012). Proteomic investigations reveal a role for RNA processing factor THRAP3 in the DNA damage response. Mol Cell *46*, 212-225.

Benson, F.E., Baumann, P., and West, S.C. (1998). Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. Nature *391*, 401-404.

Benson, F.E., Stasiak, A., and West, S.C. (1994). Purification and characterization of the human Rad51 protein, an analogue of E. coli RecA. EMBO J *13*, 5764-5771.

Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P.C., Nicolas, A., and Forterre, P. (1997). An atypical topoisomerase II from Archaea with implications for meiotic recombination. Nature *386*, 414-417.

Bermudez, V.P., Lindsey-Boltz, L.A., Cesare, A.J., Maniwa, Y., Griffith, J.D., Hurwitz, J., and Sancar, A. (2003). Loading of the human 9-1-1 checkpoint complex onto DNA by the checkpoint clamp loader hRad17-replication factor C complex in vitro. Proc Natl Acad Sci U S A *100*, 1633-1638.

Bernstein, K.A., Gangloff, S., and Rothstein, R. (2010). The RecQ DNA helicases in DNA repair. Annu Rev Genet *44*, 393-417.

Bieche, I., and Lidereau, R. (1999). Increased level of exon 12 alternatively spliced BRCA2 transcripts in tumor breast tissue compared with normal tissue. Cancer Res *59*, 2546-2550.

Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J.*, et al.* (2006). 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. Nat Methods *3*, 199-204.

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

Boddy, M.N., Gaillard, P.H., McDonald, W.H., Shanahan, P., Yates, J.R., 3rd, and Russell, P. (2001). Mus81-Eme1 are essential components of a Holliday junction resolvase. Cell *107*, 537-548.

Boehler, C., Gauthier, L.R., Mortusewicz, O., Biard, D.S., Saliou, J.M., Bresson, A., Sanglier-Cianferani, S., Smith, S., Schreiber, V., Boussin, F.*, et al.* (2011). Poly(ADP-ribose) polymerase 3 (PARP3), a newcomer in cellular response to DNA damage and mitotic progression. Proc Natl Acad Sci U S A *108*, 2783-2788.

Bolderson, E., Tomimatsu, N., Richard, D.J., Boucher, D., Kumar, R., Pandita, T.K., Burma, S., and Khanna, K.K. (2009). Phosphorylation of Exo1 modulates homologous recombination repair of DNA double-strand breaks. Nucleic Acids Res *38*, 1821-1831.

Bosco, G., and Haber, J.E. (1998). Chromosome break-induced DNA replication leads to nonreciprocal translocations and telomere capture. Genetics *150*, 1037-1047.

Bouwman, P., Aly, A., Escandell, J.M., Pieterse, M., Bartkova, J., van der Gulden, H., Hiddingh, S., Thanasoula, M., Kulkarni, A., Yang, Q.*, et al.* (2010). 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. Nat Struct Mol Biol *17*, 688-695.

Boye, E., Skjolberg, H.C., and Grallert, B. (2009). Checkpoint regulation of DNA replication. Methods Mol Biol *521*, 55-70.

Branzei, D., Vanoli, F., and Foiani, M. (2008). SUMOylation regulates Rad18-mediated template switch. Nature *456*, 915-920.

Bressan, D.A., Baxter, B.K., and Petrini, J.H. (1999). The Mre11-Rad50-Xrs2 protein complex facilitates homologous recombination-based double-strand break repair in Saccharomyces cerevisiae. Mol Cell Biol *19*, 7681-7687.

Buis, J., Wu, Y., Deng, Y., Leddon, J., Westfield, G., Eckersdorff, M., Sekiguchi, J.M., Chang, S., and Ferguson, D.O. (2008). Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation. Cell *135*, 85-96.

Bunting, S.F., Callen, E., Wong, N., Chen, H.T., Polato, F., Gunn, A., Bothmer, A., Feldhahn, N., Fernandez-Capetillo, O., Cao, L.*, et al.* (2010). 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell *141*, 243-254.

Bzymek, M., Thayer, N.H., Oh, S.D., Kleckner, N., and Hunter, N. (2010). Double Holliday junctions are intermediates of DNA break repair. Nature *464*, 937-941.

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

Carette, J.E., Guimaraes, C.P., Varadarajan, M., Park, A.S., Wuethrich, I., Godarova, A., Kotecki, M., Cochran, B.H., Spooner, E., Ploegh, H.L.*, et al.* (2009). Haploid genetic screens in human cells identify host factors used by pathogens. Science *326*, 1231-1235.

Carreira, A., Hilario, J., Amitani, I., Baskin, R.J., Shivji, M.K., Venkitaraman, A.R., and Kowalczykowski, S.C. (2009). The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. Cell *136*, 1032-1043.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Cejka, P., Cannavo, E., Polaczek, P., Masuda-Sasa, T., Pokharel, S., Campbell, J.L., and Kowalczykowski, S.C. (2010). DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. Nature *467*, 112-116.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E.*, et al.* (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov *2*, 401-404.

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

Chehab, N.H., Malikzay, A., Appel, M., and Halazonetis, T.D. (2000). Chk2/hCds1 functions as a DNA damage checkpoint in G(1) by stabilizing p53. Genes Dev *14*, 278-288.

Chen, C.F., Chen, P.L., Zhong, Q., Sharp, Z.D., and Lee, W.H. (1999). Expression of BRC repeats in breast cancer cells disrupts the BRCA2-Rad51 complex and leads to radiation hypersensitivity and loss of G(2)/M checkpoint control. J Biol Chem *274*, 32931-32935.

Chen, L., Nievera, C.J., Lee, A.Y., and Wu, X. (2008a). Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. J Biol Chem *283*, 7713-7720.

Chen, P.L., Liu, F., Cai, S., Lin, X., Li, A., Chen, Y., Gu, B., Lee, E.Y., and Lee, W.H. (2005). Inactivation of CtIP leads to early embryonic lethality mediated by G1 restraint and to tumorigenesis by haploid insufficiency. Mol Cell Biol *25*, 3535-3542.

Chen, X., Cui, D., Papusha, A., Zhang, X., Chu, C.D., Tang, J., Chen, K., Pan, X., and Ira, G. (2012). The Fun30 nucleosome remodeller promotes resection of DNA double-strand break ends. Nature *489*, 576-580.

Chen, X.B., Melchionna, R., Denis, C.M., Gaillard, P.H., Blasina, A., Van de Weyer, I., Boddy, M.N., Russell, P., Vialard, J., and McGowan, C.H. (2001). Human Mus81-associated endonuclease cleaves Holliday junctions in vitro. Mol Cell *8*, 1117-1127.

Chen, Y., Farmer, A.A., Chen, C.F., Jones, D.C., Chen, P.L., and Lee, W.H. (1996). BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner. Cancer Res *56*, 3168-3172.

Chen, Z., Yang, H., and Pavletich, N.P. (2008b). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature *453*, 489-484.

Chi, S.W., Hannon, G.J., and Darnell, R.B. (2012). An alternative mode of microRNA target recognition. Nat Struct Mol Biol *19*, 321-327.

Chou, D.M., Adamson, B., Dephoure, N.E., Tan, X., Nottke, A.C., Hurov, K.E., Gygi, S.P., Colaiacovo, M.P., and Elledge, S.J. (2010). A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage. Proc Natl Acad Sci U S A *107*, 18475-18480.

Ciccia, A., Constantinou, A., and West, S.C. (2003). Identification and characterization of the human mus81-eme1 endonuclease. J Biol Chem *278*, 25172-25178.

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

Cimprich, K.A., and Cortez, D. (2008). ATR: an essential regulator of genome integrity. Nat Rev Mol Cell Biol *9*, 616-627.

Clark, A.J., and Margulies, A.D. (1965). Isolation and Characterization of Recombination-Deficient Mutants of Escherichia Coli K12. Proc Natl Acad Sci U S A *53*, 451-459.

Clerici, M., Mantiero, D., Lucchini, G., and Longhese, M.P. (2005). The Saccharomyces cerevisiae Sae2 protein promotes resection and bridging of double strand break ends. J Biol Chem *280*, 38631-38638.

Connolly, B., Parsons, C.A., Benson, F.E., Dunderdale, H.J., Sharples, G.J., Lloyd, R.G., and West, S.C. (1991). Resolution of Holliday junctions in vitro requires the Escherichia coli ruvC gene product. Proc Natl Acad Sci U S A *88*, 6063-6067.

Constantinou, A., Chen, X.B., McGowan, C.H., and West, S.C. (2002). Holliday junction resolution in human cells: two junction endonucleases with distinct substrate specificities. EMBO J *21*, 5577-5585.

Corneo, B., Wendland, R.L., Deriano, L., Cui, X., Klein, I.A., Wong, S.Y., Arnal, S., Holub, A.J., Weller, G.R., Pancake, B.A., *et al.* (2007). Rag mutations reveal robust alternative end joining. Nature *449*, 483-486.

Cortez, D., Guntuku, S., Qin, J., and Elledge, S.J. (2001). ATR and ATRIP: partners in checkpoint signaling. Science *294*, 1713-1716.

Costelloe, T., Louge, R., Tomimatsu, N., Mukherjee, B., Martini, E., Khadaroo, B., Dubois, K., Wiegant, W.W., Thierry, A., Burma, S., *et al.* (2012). The yeast Fun30 and human SMARCAD1 chromatin remodellers promote DNA end resection. Nature *489*, 581-584.

Cotta-Ramusino, C., McDonald, E.R., 3rd, Hurov, K., Sowa, M.E., Harper, J.W., and Elledge, S.J. (2011). A DNA damage response screen identifies RHINO, a 9-1-1 and TopBP1 interacting protein required for ATR signaling. Science *332*, 1313-1317.

Crackower, M.A., Scherer, S.W., Rommens, J.M., Hui, C.C., Poorkaj, P., Soder, S., Cobben, J.M., Hudgins, L., Evans, J.P., and Tsui, L.C. (1996). Characterization of the split hand/split foot malformation locus SHFM1 at 7q21.3-q22.1 and analysis of a candidate gene for its expression during limb development. Hum Mol Genet *5*, 571-579.

Creighton, H.B., and McClintock, B. (1931). A Correlation of Cytological and Genetical Crossing-Over in Zea Mays. Proc Natl Acad Sci U S A *17*, 492-497.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561-563.

D'Amours, D., Desnoyers, S., D'Silva, I., and Poirier, G.G. (1999). Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. Biochem J *342 (Pt 2)*, 249-268.

de Jager, M., van Noort, J., van Gent, D.C., Dekker, C., Kanaar, R., and Wyman, C. (2001). Human Rad50/Mre11 is a flexible complex that can tether DNA ends. Mol Cell *8*, 1129-1135.

de Laat, W.L., Jaspers, N.G., and Hoeijmakers, J.H. (1999). Molecular mechanism of nucleotide excision repair. Genes Dev *13*, 768-785.

Dianov, G.L., Meisenberg, C., and Parsons, J.L. (2011). Regulation of DNA repair by ubiquitylation. Biochemistry (Mosc) *76*, 69-79.

DiBiase, S.J., Zeng, Z.C., Chen, R., Hyslop, T., Curran, W.J., Jr., and Iliakis, G. (2000). DNA-dependent protein kinase stimulates an independently active, nonhomologous, end-joining apparatus. Cancer Res *60*, 1245-1253.

Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B., *et al.* (2008). Somatic mutations affect key pathways in lung adenocarcinoma. Nature *455*, 1069-1075.

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D.H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J., *et al.* (2009). RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. Cell *136*, 435-446.

Dong, Y., Hakimi, M.A., Chen, X., Kumaraswamy, E., Cooch, N.S., Godwin, A.K., and Shiekhattar, R. (2003). Regulation of BRCC, a holoenzyme complex containing BRCA1 and BRCA2, by a signalosome-like subunit and its role in DNA repair. Mol Cell *12*, 1087-1099.

Donzelli, M., and Draetta, G.F. (2003). Regulating mammalian checkpoints through Cdc25 inactivation. EMBO Rep *4*, 671-677.

Draviam, V.M., Stegmeier, F., Nalepa, G., Sowa, M.E., Chen, J., Liang, A., Hannon, G.J., Sorger, P.K., Harper, J.W., and Elledge, S.J. (2007). A functional genomic screen identifies a role for TAO1 kinase in spindle-checkpoint signalling. Nat Cell Biol *9*, 556-564.

Dray, E., Etchin, J., Wiese, C., Saro, D., Williams, G.J., Hammel, M., Yu, X., Galkin, V.E., Liu, D., Tsai, M.S.*, et al.* (2010). Enhancement of RAD51 recombinase activity by the tumor suppressor PALB2. Nat Struct Mol Biol *17*, 1255-1259.

Drost, R., Bouwman, P., Rottenberg, S., Boon, U., Schut, E., Klarenbeek, S., Klijn, C., van der Heijden, I., van der Gulden, H., Wientjens, E.*, et al.* (2011). BRCA1 RING function is essential for tumor suppression but dispensable for therapy resistance. Cancer Cell *20*, 797-809.

Eapen, V.V., Sugawara, N., Tsabar, M., Wu, W.H., and Haber, J.E. (2012). The Saccharomyces cerevisiae Chromatin Remodeler Fun30 Regulates DNA End Resection and Checkpoint Deactivation. Mol Cell Biol *32*, 4727-4740.

Eid, W., Steger, M., El-Shemerly, M., Ferretti, L.P., Pena-Diaz, J., Konig, C., Valtorta, E., Sartori, A.A., and Ferrari, S. (2010). DNA end resection by CtIP and exonuclease 1 prevents genomic instability. EMBO Rep *11*, 962-968.

Eitoku, M., Sato, L., Senda, T., and Horikoshi, M. (2008). Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. Cell Mol Life Sci *65*, 414-444.

Elborough, K.M., and West, S.C. (1990). Resolution of synthetic Holliday junctions in DNA by an endonuclease activity from calf thymus. EMBO J *9*, 2931-2936.

Elia, A.E., and Elledge, S.J. (2012). BRCA1 as tumor suppressor: lord without its RING? Breast Cancer Res *14*, 306.

Elliott, D.J. (2004). The role of potential splicing factors including RBMY, RBMX, hnRNPG-T and STAR proteins in spermatogenesis. Int J Androl *27*, 328-334.

Ellis, N.A., Proytcheva, M., Sanz, M.M., Ye, T.Z., and German, J. (1999). Transfection of BLM into cultured bloom syndrome cells reduces the sister-chromatid exchange rate toward normal. Am J Hum Genet *65*, 1368-1374.

Ellison, V., and Stillman, B. (2003). Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5' recessed DNA. PLoS Biol *1*, E33.

Erkko, H., Xia, B., Nikkila, J., Schleutker, J., Syrjakoski, K., Mannermaa, A., Kallioniemi, A., Pylkas, K., Karppinen, S.M., Rapakko, K.*, et al.* (2007). A recurrent mutation in PALB2 in Finnish cancer families. Nature *446*, 316-319.

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

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

Fekairi, S., Scaglione, S., Chahwan, C., Taylor, E.R., Tissier, A., Coulon, S., Dong, M.Q., Ruse, C., Yates, J.R., 3rd, Russell, P*., et al.* (2009). Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. Cell *138*, 78-89.

Fellmann, C., Zuber, J., McJunkin, K., Chang, K., Malone, C.D., Dickins, R.A., Xu, Q., Hengartner, M.O., Elledge, S.J., Hannon, G.J*., et al.* (2011). Functional identification of optimized RNAi triggers using a massively parallel sensor assay. Mol Cell *41*, 733-746.

Feng, L., Huang, J., and Chen, J. (2009). MERIT40 facilitates BRCA1 localization and DNA damage repair. Genes Dev *23*, 719-728.

Feng, Z., Scott, S.P., Bussen, W., Sharma, G.G., Guo, G., Pandita, T.K., and Powell, S.N. (2011). Rad52 inactivation is synthetically lethal with BRCA2 deficiency. Proc Natl Acad Sci U S A *108*, 686-691.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fishman-Lobell, J., Rudin, N., and Haber, J.E. (1992). Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. Mol Cell Biol *12*, 1292-1303.

Flicek, P., Ahmed, I., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S*., et al.* (2013). Ensembl 2013. Nucleic Acids Res *41*, D48-55.

Fricke, W.M., and Brill, S.J. (2003). Slx1-Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1-Top3. Genes Dev *17*, 1768-1778.

Gagne, J.P., Isabelle, M., Lo, K.S., Bourassa, S., Hendzel, M.J., Dawson, V.L., Dawson, T.M., and Poirier, G.G. (2008). Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. Nucleic Acids Res *36*, 6959-6976.

Galanty, Y., Belotserkovskaya, R., Coates, J., Polo, S., Miller, K.M., and Jackson, S.P. (2009). Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. Nature *462*, 935-939.

Galletto, R., Amitani, I., Baskin, R.J., and Kowalczykowski, S.C. (2006). Direct observation of individual RecA filaments assembling on single DNA molecules. Nature *443*, 875-878.

Game, J.C., and Mortimer, R.K. (1974). A genetic study of x-ray sensitive mutants in yeast. Mutat Res *24*, 281-292.

Garcia, V., Phelps, S.E., Gray, S., and Neale, M.J. (2011). Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. Nature *479*, 241-244.

Gasior, S.L., Wong, A.K., Kora, Y., Shinohara, A., and Bishop, D.K. (1998). Rad52 associates with RPA and functions with rad55 and rad57 to assemble meiotic recombination complexes. Genes Dev *12*, 2208-2221.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J.*, et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol *5*, R80.

German, J., Schonberg, S., Louie, E., and Chaganti, R.S. (1977). Bloom's syndrome. IV. Sister-chromatid exchanges in lymphocytes. Am J Hum Genet *29*, 248-255.

Germann, M.W., Johnson, C.N., and Spring, A.M. (2010). Recognition of damaged DNA: structure and dynamic markers. Med Res Rev *32*, 659-683.

Gradwohl, G., Menissier de Murcia, J.M., Molinete, M., Simonin, F., Koken, M., Hoeijmakers, J.H., and de Murcia, G. (1990). The second zinc-finger domain of poly(ADP-ribose) polymerase determines specificity for single-stranded breaks in DNA. Proc Natl Acad Sci U S A *87*, 2990-2994.

Gravel, S., Chapman, J.R., Magill, C., and Jackson, S.P. (2008). DNA helicases Sgs1 and BLM promote DNA double-strand break resection. Genes Dev *22*, 2767-2772.

Grawunder, U., Wilm, M., Wu, X., Kulesza, P., Wilson, T.E., Mann, M., and Lieber, M.R. (1997). Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature *388*, 492-495.

Greenberg, R.A., Sobhian, B., Pathania, S., Cantor, S.B., Nakatani, Y., and Livingston, D.M. (2006). Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes. Genes Dev *20*, 34-46.

Gu, J., Lu, H., Tippin, B., Shimazaki, N., Goodman, M.F., and Lieber, M.R. (2007). XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. EMBO J *26*, 1010-1023.

Gudmundsdottir, K., Lord, C.J., Witt, E., Tutt, A.N., and Ashworth, A. (2004). DSS1 is required for RAD51 focus formation and genomic stability in mammalian cells. EMBO Rep *5*, 989-993.

Guo, Z., Kumagai, A., Wang, S.X., and Dunphy, W.G. (2000). Requirement for Atr in phosphorylation of Chk1 and cell cycle regulation in response to DNA replication blocks and UV-damaged DNA in Xenopus egg extracts. Genes Dev *14*, 2745-2756.

Haince, J.F., Kozlov, S., Dawson, V.L., Dawson, T.M., Hendzel, M.J., Lavin, M.F., and Poirier, G.G. (2007). Ataxia telangiectasia mutated (ATM) signaling network is modulated by a novel poly(ADP-ribose)-dependent pathway in the early response to DNA-damaging agents. J Biol Chem *282*, 16441-16453.

Haince, J.F., McDonald, D., Rodrigue, A., Dery, U., Masson, J.Y., Hendzel, M.J., and Poirier, G.G. (2008). PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. J Biol Chem *283*, 1197-1208.

Han, S.P., Tang, Y.H., and Smith, R. (2010). Functional diversity of the hnRNPs: past, present and perspectives. Biochem J *430*, 379-392.

Hartlerode, A.J., and Scully, R. (2009). Mechanisms of double-strand break repair in somatic mammalian cells. Biochem J *423*, 157-168.

Hashizume, R., Fukuda, M., Maeda, I., Nishikawa, H., Oyake, D., Yabuki, Y., Ogata, H., and Ohta, T. (2001). The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. J Biol Chem *276*, 14537-14540.

Heinrich, B., Zhang, Z., Raitskin, O., Hiller, M., Benderska, N., Hartmann, A.M., Bracco, L., Elliott, D., Ben-Ari, S., Soreq, H.*, et al.* (2009). Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. J Biol Chem *284*, 14303-14315.

Heo, K., Kim, H., Choi, S.H., Choi, J., Kim, K., Gu, J., Lieber, M.R., Yang, A.S., and An, W. (2008). FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. Mol Cell *30*, 86-97.

Herve Pages, M.C., Seth Falcon and Nianhua Li (2012). AnnotationDbi: Annotation Database Interface. R package version 1.18.4.

Heyer, W.D., Ehmsen, K.T., and Liu, J. (2010). Regulation of homologous recombination in eukaryotes. Annu Rev Genet *44*, 113-139.

Hirano, M., Yamamoto, A., Mori, T., Lan, L., Iwamoto, T.A., Aoki, M., Shimada, K., Furiya, Y., Kariya, S., Asai, H.*, et al.* (2007). DNA single-strand break repair is impaired in aprataxin-related ataxia. Ann Neurol *61*, 162-174.

Hofmann, W., Horn, D., Huttner, C., Classen, E., and Scherneck, S. (2003). The BRCA2 variant 8204G>A is a splicing mutation and results in an in frame deletion of the gene. J Med Genet *40*, e23.

Hofmann, Y., and Wirth, B. (2002). hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1. Hum Mol Genet *11*, 2037-2049.

Holliday, R. (1962). Mutation and replication in Ustilago maydis. Genetical Research, 472-486.

Holliday, R. (1964). A mechanism for gene conversion in fungi. Genet Res *89*, 285-307.

Holliday, R. (1967). Altered recombination frequencies in radiation sensitivie strains of Ustilago. Mutat Res *4*, 275-288.

Hollingsworth, N.M., and Brill, S.J. (2004). The Mus81 solution to resolution: generating meiotic crossovers without Holliday junctions. Genes Dev *18*, 117-125.

Holthausen, J.T., Wyman, C., and Kanaar, R. (2010). Regulation of DNA strand exchange in homologous recombination. DNA Repair (Amst) *9*, 1264-1272.

Hu, Y., Scully, R., Sobhian, B., Xie, A., Shestakova, E., and Livingston, D.M. (2011). RAP80-directed tuning of BRCA1 homologous recombination function at ionizing radiation-induced nuclear foci. Genes Dev *25*, 685-700.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Hubner, N.C., Wang, L.H., Kaulich, M., Descombes, P., Poser, I., and Nigg, E.A. (2010). Re-examination of siRNA specificity questions role of PICH and Tao1 in the spindle checkpoint and identifies Mad2 as a sensitive target for small RNAs. Chromosoma *119*, 149-165.

Huen, M.S., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M.B., and Chen, J. (2007). RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. Cell *131*, 901-914.

Huen, M.S., Sy, S.M., and Chen, J. (2010). BRCA1 and its toolbox for the maintenance of genome integrity. Nat Rev Mol Cell Biol *11*, 138-148.

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

Huertas, P., and Jackson, S.P. (2009). Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. J Biol Chem *284*, 9558-9565.

Hunter, N., and Kleckner, N. (2001). The single-end invasion: an asymmetric intermediate at the double-strand break to double-holliday junction transition of meiotic recombination. Cell *106*, 59-70.

Hurov, K.E., Cotta-Ramusino, C., and Elledge, S.J. (2010). A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. Genes Dev *24*, 1939-1950.

Ip, S.C., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M., and West, S.C. (2008). Identification of Holliday junction resolvases from humans and yeast. Nature *456*, 357-361.

Ira, G., Pellicioli, A., Balijja, A., Wang, X., Fiorani, S., Carotenuto, W., Liberi, G., Bressan, D., Wan, L., Hollingsworth, N.M.*, et al.* (2004). DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. Nature *431*, 1011-1017.

Ivanov, E.L., Sugawara, N., White, C.I., Fabre, F., and Haber, J.E. (1994). Mutations in XRS2 and RAD50 delay but do not prevent mating-type switching in Saccharomyces cerevisiae. Mol Cell Biol *14*, 3414-3425.

Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P.S. (2003). Expression profiling reveals off-target gene regulation by RNAi. Nat Biotechnol *21*, 635-637.

Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L., and Linsley, P.S. (2006). Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. RNA *12*, 1179-1187.

Jackson, S.P. (2002). Sensing and repairing DNA double-strand breaks. Carcinogenesis *23*, 687-696.

Jacquemont, C., and Taniguchi, T. (2007). Proteasome function is required for DNA damage response and fanconi anemia pathway activation. Cancer Res *67*, 7395-7405.

Janknecht, R. (2010). Multi-talented DEAD-box proteins and potential tumor promoters: p68 RNA helicase (DDX5) and its paralog, p72 RNA helicase (DDX17). Am J Transl Res *2*, 223-234.

Jazayeri, A., Falck, J., Lukas, C., Bartek, J., Smith, G.C., Lukas, J., and Jackson, S.P. (2006). ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. Nat Cell Biol *8*, 37-45.

Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. Nature *467*, 678-683.

Jin, J., Shirogane, T., Xu, L., Nalepa, G., Qin, J., Elledge, S.J., and Harper, J.W. (2003). SCFbeta-TRCP links Chk1 signaling to degradation of the Cdc25A protein phosphatase. Genes Dev *17*, 3062-3074.

Johnson, F.B., Lombard, D.B., Neff, N.F., Mastrangelo, M.A., Dewolf, W., Ellis, N.A., Marciniak, R.A., Yin, Y., Jaenisch, R., and Guarente, L. (2000). Association of the Bloom syndrome protein with topoisomerase IIIalpha in somatic and meiotic cells. Cancer Res *60*, 1162-1167.

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

Joo, C., McKinney, S.A., Nakamura, M., Rasnik, I., Myong, S., and Ha, T. (2006). Real-time observation of RecA filament dynamics with single monomer resolution. Cell *126*, 515-527.

Kalocsay, M., Hiller, N.J., and Jentsch, S. (2009). Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. Mol Cell *33*, 335-343.

Kanhoush, R., Beenders, B., Perrin, C., Moreau, J., Bellini, M., and Penrad-Mobayed, M. (2009). Novel domains in the hnRNP G/RBMX protein with distinct roles in RNA binding and targeting nascent transcripts. Nucleus *1*, 109-122.

Karow, J.K., Constantinou, A., Li, J.L., West, S.C., and Hickson, I.D. (2000). The Bloom's syndrome gene product promotes branch migration of holliday junctions. Proc Natl Acad Sci U S A *97*, 6504-6508.

Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. EMBO J *24*, 1911-1920.

Keeney, S., Giroux, C.N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell *88*, 375-384.

Khanna, K.K., and Jackson, S.P. (2001). DNA double-strand breaks: signaling, repair and the cancer connection. Nat Genet *27*, 247-254.

Kiledjian, M., and Dreyfuss, G. (1992). Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. EMBO J *11*, 2655-2664.

Kim, E., Du, L., Bregman, D.B., and Warren, S.L. (1997). Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. J Cell Biol *136*, 19-28.

Kim, H., Chen, J., and Yu, X. (2007). Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. Science *316*, 1202-1205.

Kleiman, F.E., Wu-Baer, F., Fonseca, D., Kaneko, S., Baer, R., and Manley, J.L. (2005). BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II. Genes Dev *19*, 1227-1237.

Kolas, N.K., Chapman, J.R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F.D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T.M.*, et al.* (2007). Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. Science *318*, 1637-1640.

Kowalczykowski, S.C., and Eggleston, A.K. (1994). Homologous pairing and DNA strand-exchange proteins. Annu Rev Biochem *63*, 991-1043.

Krogh, B.O., and Symington, L.S. (2004). Recombination proteins in yeast. Annu Rev Genet *38*, 233-271.

Kumagai, A., Lee, J., Yoo, H.Y., and Dunphy, W.G. (2006). TopBP1 activates the ATR-ATRIP complex. Cell *124*, 943-955.

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

Lackner, D.H., Durocher, D., and Karlseder, J. (2011). A siRNA-based screen for genes involved in chromosome end protection. PLoS One *6*, e21407.

Lagerwerf, S., Vrouwe, M.G., Overmeer, R.M., Fousteri, M.I., and Mullenders, L.H. (2011). DNA damage response and transcription. DNA Repair (Amst) *10*, 743-750.

Lamarche, B.J., Orazio, N.I., and Weitzman, M.D. (2010). The MRN complex in double-strand break repair and telomere maintenance. FEBS Lett *584*, 3682-3695.

Lan, L., Nakajima, S., Oohata, Y., Takao, M., Okano, S., Masutani, M., Wilson, S.H., and Yasui, A. (2004). In situ analysis of repair processes for oxidative DNA damage in mammalian cells. Proc Natl Acad Sci U S A *101*, 13738-13743.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lee, B.I., and Wilson, D.M., 3rd (1999). The RAD2 domain of human exonuclease 1 exhibits 5' to 3' exonuclease and flap structure-specific endonuclease activities. J Biol Chem *274*, 37763-37769.

Lee, J., and Dunphy, W.G. (2010). Rad17 plays a central role in establishment of the interaction between TopBP1 and the Rad9-Hus1-Rad1 complex at stalled replication forks. Mol Biol Cell *21*, 926-935.

Lengsfeld, B.M., Rattray, A.J., Bhaskara, V., Ghirlando, R., and Paull, T.T. (2007). Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. Mol Cell *28*, 638-651.

Li, J., Zou, C., Bai, Y., Wazer, D.E., Band, V., and Gao, Q. (2006). DSS1 is required for the stability of BRCA2. Oncogene *25*, 1186-1194.

Li, L., Monckton, E.A., and Godbout, R. (2008). A role for DEAD box 1 at DNA double-strand breaks. Mol Cell Biol *28*, 6413-6425.

Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell *122*, 365-378.

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

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature *433*, 769-773.

Limoli, C.L., and Ward, J.F. (1993). A new method for introducing double-strand breaks into cellular DNA. Radiat Res *134*, 160-169.

Lingenfelter, P.A., Delbridge, M.L., Thomas, S., Hoekstra, H.E., Mitchell, M.J., Graves, J.A., and Disteche, C.M. (2001). Expression and conservation of processed copies of the RBMX gene. Mamm Genome *12*, 538-545.

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

Llorente, B., and Symington, L.S. (2004). The Mre11 nuclease is not required for 5' to 3' resection at multiple HO-induced double-strand breaks. Mol Cell Biol *24*, 9682-9694.

Loizou, J.I., El-Khamisy, S.F., Zlatanou, A., Moore, D.J., Chan, D.W., Qin, J., Sarno, S., Meggio, F., Pinna, L.A., and Caldecott, K.W. (2004). The protein kinase CK2 facilitates repair of chromosomal DNA single-strand breaks. Cell *117*, 17-28.

Lorain, S., Quivy, J.P., Monier-Gavelle, F., Scamps, C., Lecluse, Y., Almouzni, G., and Lipinski, M. (1998). Core histones and HIRIP3, a novel histone-binding protein, directly interact with WD repeat protein HIRA. Mol Cell Biol *18*, 5546-5556.

Loseva, O., Jemth, A.S., Bryant, H.E., Schuler, H., Lehtio, L., Karlberg, T., and Helleday, T. (2010). PARP-3 is a mono-ADP-ribosylase that activates PARP-1 in the absence of DNA. J Biol Chem *285*, 8054-8060.

Lou, Z., Minter-Dykhouse, K., Franco, S., Gostissa, M., Rivera, M.A., Celeste, A., Manis, J.P., van Deursen, J., Nussenzweig, A., Paull, T.T.*, et al.* (2006). MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. Mol Cell *21*, 187-200.

Luk, E., Vu, N.D., Patteson, K., Mizuguchi, G., Wu, W.H., Ranjan, A., Backus, J., Sen, S., Lewis, M., Bai, Y.*, et al.* (2007). Chz1, a nuclear chaperone for histone H2AZ. Mol Cell *25*, 357-368.

Lukas, C., Bartek, J., and Lukas, J. (2005). Imaging of protein movement induced by chromosomal breakage: tiny 'local' lesions pose great 'global' challenges. Chromosoma *114*, 146-154.

Lukas, J., Lukas, C., and Bartek, J. (2004). Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time. DNA Repair (Amst) *3*, 997-1007.

Luo, G., Yao, M.S., Bender, C.F., Mills, M., Bladl, A.R., Bradley, A., and Petrini, J.H. (1999). Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. Proc Natl Acad Sci U S A *96*, 7376-7381.

Luo, J., Solimini, N.L., and Elledge, S.J. (2009). Principles of cancer therapy: oncogene and non-oncogene addiction. Cell *136*, 823-837.

Luo, Z., Zheng, J., Lu, Y., and Bregman, D.B. (2001). Ultraviolet radiation alters the phosphorylation of RNA polymerase II large subunit and accelerates its proteasome-dependent degradation. Mutat Res *486*, 259-274.

Ma, Y., Lu, H., Tippin, B., Goodman, M.F., Shimazaki, N., Koiwai, O., Hsieh, C.L., Schwarz, K., and Lieber, M.R. (2004). A biochemically defined system for mammalian nonhomologous DNA end joining. Mol Cell *16*, 701-713.

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M.R. (2002). Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell *108*, 781-794.

Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., and Lukas, J. (2007). RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. Cell *131*, 887-900.

Manke, I.A., Lowery, D.M., Nguyen, A., and Yaffe, M.B. (2003). BRCT repeats as phosphopeptide-binding modules involved in protein targeting. Science *302*, 636-639.

Marston, A.L., and Amon, A. (2004). Meiosis: cell-cycle controls shuffle and deal. Nat Rev Mol Cell Biol *5*, 983-997.

Marston, N.J., Richards, W.J., Hughes, D., Bertwistle, D., Marshall, C.J., and Ashworth, A. (1999). Interaction between the product of the breast cancer susceptibility gene BRCA2 and DSS1, a protein functionally conserved from yeast to mammals. Mol Cell Biol *19*, 4633-4642.

Masson, M., Niedergang, C., Schreiber, V., Muller, S., Menissier-de Murcia, J., and de Murcia, G. (1998). XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. Mol Cell Biol *18*, 3563-3571.

Masuda-Sasa, T., Imamura, O., and Campbell, J.L. (2006). Biochemical analysis of human Dna2. Nucleic Acids Res *34*, 1865-1875.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y*., et al.* (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science *316*, 1160-1166.

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

Mattiroli, F., Vissers, J.H., van Dijk, W.J., Ikpa, P., Citterio, E., Vermeulen, W., Marteijn, J.A., and Sixma, T.K. (2012). RNF168 Ubiquitinates K13-15 on H2A/H2AX to Drive DNA Damage Signaling. Cell *150*, 1182-1195.

Mazeyrat, S., Saut, N., Mattei, M.G., and Mitchell, M.J. (1999). RBMY evolved on the Y chromosome from a ubiquitously transcribed X-Y identical gene. Nat Genet *22*, 224-226.

McClintock, B. (1931). The Order of the Genes C, Sh and Wx in Zea Mays with Reference to a Cytologically Known Point in the Chromosome. Proc Natl Acad Sci U S A *17*, 485-491.

McEachern, M.J., and Haber, J.E. (2006). Break-induced replication and recombinational telomere elongation in yeast. Annu Rev Biochem *75*, 111-135.

McIlwraith, M.J., and West, S.C. (2008). DNA repair synthesis facilitates RAD52-mediated second-end capture during DSB repair. Mol Cell *29*, 510-516.

McMahill, M.S., Sham, C.W., and Bishop, D.K. (2007). Synthesis-dependent strand annealing in meiosis. PLoS Biol *5*, e299.

McVey, M., and Lee, S.E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet *24*, 529-538.

Meek, K., Dang, V., and Lees-Miller, S.P. (2008). DNA-PK: the means to justify the ends? Adv Immunol *99*, 33-58.

Miki, Y., Katagiri, T., Kasumi, F., Yoshimoto, T., and Nakamura, Y. (1996). Mutation analysis in the BRCA2 gene in primary breast cancers. Nat Genet *13*, 245-247.

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature *455*, 770-774.

Mimori, T., and Hardin, J.A. (1986). Mechanism of interaction between Ku protein and DNA. J Biol Chem *261*, 10375-10379.

Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. Mol Cell *3*, 697-705.

Miwa, M., and Sugimura, T. (1971). Splitting of the ribose-ribose linkage of poly(adenosine diphosphate-robose) by a calf thymus extract. J Biol Chem *246*, 6362-6364.

Mizuuchi, K., Kemper, B., Hays, J., and Weisberg, R.A. (1982). T4 endonuclease VII cleaves holliday structures. Cell *29*, 357-365.

Modesti, M., Ristic, D., van der Heijden, T., Dekker, C., van Mameren, J., Peterman, E.J., Wuite, G.J., Kanaar, R., and Wyman, C. (2007). Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. Structure *15*, 599-609.

Mohr, S., Bakal, C., and Perrimon, N. (2010). Genomic screening with RNAi: results and challenges. Annu Rev Biochem *79*, 37-64.

Mordes, D.A., Glick, G.G., Zhao, R., and Cortez, D. (2008). TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. Genes Dev *22*, 1478-1489.

Moreau, S., Ferguson, J.R., and Symington, L.S. (1999). The nuclease activity of Mre11 is required for meiosis but not for mating type switching, end joining, or telomere maintenance. Mol Cell Biol *19*, 556-566.

Morgan, T.H. (1911a). Random Segregation Versus Coupling in Mendelian Inheritance. Science *34*, 384.

Morgan, T.H. (1911b). The Origin of Nine Wing Mutations in Drosophila. Science *33*, 496-499.

Morgan, T.H., Sturtevant, A.H., Muller, H.J., and Bridges, C.B. (1915). The Mechanism of Mendelian Heredity Henry Holt and Company

Morris, J.R., Boutell, C., Keppler, M., Densham, R., Weekes, D., Alamshah, A., Butler, L., Galanty, Y., Pangon, L., Kiuchi, T.*, et al.* (2009). The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. Nature *462*, 886-890.

Mortusewicz, O., Ame, J.C., Schreiber, V., and Leonhardt, H. (2007). Feedback-regulated poly(ADP-ribosyl)ation by PARP-1 is required for rapid response to DNA damage in living cells. Nucleic Acids Res *35*, 7665-7675.

Mortusewicz, O., Rothbauer, U., Cardoso, M.C., and Leonhardt, H. (2006). Differential recruitment of DNA Ligase I and III to DNA repair sites. Nucleic Acids Res *34*, 3523-3532.

Moynahan, M.E., Pierce, A.J., and Jasin, M. (2001). BRCA2 is required for homology-directed repair of chromosomal breaks. Mol Cell *7*, 263-272.

Mullen, J.R., Kaliraman, V., Ibrahim, S.S., and Brill, S.J. (2001). Requirement for three novel protein complexes in the absence of the Sgs1 DNA helicase in Saccharomyces cerevisiae. Genetics *157*, 103-118.

Munoz, I.M., Hain, K., Declais, A.C., Gardiner, M., Toh, G.W., Sanchez-Pulido, L., Heuckmann, J.M., Toth, R., Macartney, T., Eppink, B.*, et al.* (2009a). Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. Mol Cell *35*, 116-127.

Munoz, M.J., Perez Santangelo, M.S., Paronetto, M.P., de la Mata, M., Pelisch, F., Boireau, S., Glover-Cutter, K., Ben-Dov, C., Blaustein, M., Lozano, J.J.*, et al.* (2009b). DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. Cell *137*, 708-720.

Nakada, S., Tai, I., Panier, S., Al-Hakim, A., Iemura, S., Juang, Y.C., O'Donnell, L., Kumakubo, A., Munro, M., Sicheri, F.*, et al.* (2010). Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1. Nature *466*, 941-946.

Nakanishi, K., Moran, A., Hays, T., Kuang, Y., Fox, E., Garneau, D., Montes de Oca, R., Grompe, M., and D'Andrea, A.D. (2001). Functional analysis of patient-derived mutations in the Fanconi anemia gene, FANCG/XRCC9. Exp Hematol *29*, 842-849.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell *2*, 279-289.

Nasim, M.T., Chernova, T.K., Chowdhury, H.M., Yue, B.G., and Eperon, I.C. (2003). HnRNP G and Tra2beta: opposite effects on splicing matched by antagonism in RNA binding. Hum Mol Genet *12*, 1337-1348.

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

Nimonkar, A.V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P., and Kowalczykowski, S.C. (2011). BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. Genes Dev *25*, 350-362.

Nimonkar, A.V., Ozsoy, A.Z., Genschel, J., Modrich, P., and Kowalczykowski, S.C. (2008). Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. Proc Natl Acad Sci U S A *105*, 16906-16911.

Nimonkar, A.V., Sica, R.A., and Kowalczykowski, S.C. (2009). Rad52 promotes second-end DNA capture in double-stranded break repair to form complement-stabilized joint molecules. Proc Natl Acad Sci U S A *106*, 3077-3082.

Niu, H., Chung, W.H., Zhu, Z., Kwon, Y., Zhao, W., Chi, P., Prakash, R., Seong, C., Liu, D., Lu, L.*, et al.* (2010). Mechanism of the ATP-dependent DNA end-resection machinery from Saccharomyces cerevisiae. Nature *467*, 108-111.

Ogata, N., Ueda, K., Kawaichi, M., and Hayaishi, O. (1981). Poly(ADP-ribose) synthetase, a main acceptor of poly(ADP-ribose) in isolated nuclei. J Biol Chem *256*, 4135-4137.

Okano, S., Lan, L., Caldecott, K.W., Mori, T., and Yasui, A. (2003). Spatial and temporal cellular responses to single-strand breaks in human cells. Mol Cell Biol *23*, 3974-3981.

Pan-Hammarstrom, Q., Jones, A.M., Lahdesmaki, A., Zhou, W., Gatti, R.A., Hammarstrom, L., Gennery, A.R., and Ehrenstein, M.R. (2005). Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. J Exp Med *201*, 189-194.

Paronetto, M.P., Minana, B., and Valcarcel, J. (2011). The Ewing sarcoma protein regulates DNA damage-induced alternative splicing. Mol Cell *43*, 353-368.

Paull, T.T., and Gellert, M. (1998). The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. Mol Cell *1*, 969-979.

Paull, T.T., and Gellert, M. (1999). Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. Genes Dev *13*, 1276-1288.

Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E.*, et al.* (2009). A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol Cell *35*, 228-239.

Peak, J.G., and Peak, M.J. (1990). Ultraviolet light induces double-strand breaks in DNA of cultured human P3 cells as measured by neutral filter elution. Photochem Photobiol *52*, 387-393.

Pellegrini, L., Yu, D.S., Lo, T., Anand, S., Lee, M., Blundell, T.L., and Venkitaraman, A.R. (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. Nature *420*, 287-293.

Peng, C.Y., Graves, P.R., Thoma, R.S., Wu, Z., Shaw, A.S., and Piwnica-Worms, H. (1997). Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. Science *277*, 1501-1505.

Pensabene, M., Spagnoletti, I., Capuano, I., Condello, C., Pepe, S., Contegiacomo, A., Lombardi, G., Bevilacqua, G., and Caligo, M.A. (2009). Two mutations of BRCA2 gene at exon and splicing site in a woman who underwent oncogenetic counseling. Ann Oncol *20*, 874-878.

Perrimon, N., and Mathey-Prevot, B. (2007). Matter arising: off-targets and genome-scale RNAi screens in Drosophila. Fly (Austin) *1*, 1-5.

Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev *20*, 2922-2936.

Pierce, A.J., Hu, P., Han, M., Ellis, N., and Jasin, M. (2001). Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. Genes Dev *15*, 3237-3242.

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

Pinato, S., Scandiuzzi, C., Arnaudo, N., Citterio, E., Gaudino, G., and Penengo, L. (2009). RNF168, a new RING finger, MIU-containing protein that modifies chromatin by ubiquitination of histones H2A and H2AX. BMC Mol Biol *10*, 55.

Pleschke, J.M., Kleczkowska, H.E., Strohm, M., and Althaus, F.R. (2000). Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. J Biol Chem *275*, 40974-40980.

Poirier, G.G., de Murcia, G., Jongstra-Bilen, J., Niedergang, C., and Mandel, P. (1982). Poly(ADP-ribosyl)ation of polynucleosomes causes relaxation of chromatin structure. Proc Natl Acad Sci U S A *79*, 3423-3427.

Polo, S.E., Blackford, A.N., Chapman, J.R., Baskcomb, L., Gravel, S., Rusch, A., Thomas, A., Blundred, R., Smith, P., Kzhyshkowska, J.*, et al.* (2012). Regulation of DNA-end resection by hnRNPU-like proteins promotes DNA double-strand break signaling and repair. Mol Cell *45*, 505-516.

Polo, S.E., Roche, D., and Almouzni, G. (2006). New histone incorporation marks sites of UV repair in human cells. Cell *127*, 481-493.

Prochasson, P., Florens, L., Swanson, S.K., Washburn, M.P., and Workman, J.L. (2005). The HIR corepressor complex binds to nucleosomes generating a distinct protein/DNA complex resistant to remodeling by SWI/SNF. Genes Dev *19*, 2534-2539.

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, URL http://wwwR-projectorg/.

Radding, C.M. (1989). Helical RecA nucleoprotein filaments mediate homologous pairing and strand exchange. Biochim Biophys Acta *1008*, 131-145.

Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., Reid, S., Spanova, K., Barfoot, R., Chagtai, T.*, et al.* (2007). PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet *39*, 165-167.

Ransom, M., Dennehey, B.K., and Tyler, J.K. (2010). Chaperoning histones during DNA replication and repair. Cell *140*, 183-195.

Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res *12*, 1231-1245.

Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Mol Cell *9*, 1091-1100.

Raynard, S., Bussen, W., and Sung, P. (2006). A double Holliday junction dissolvasome comprising BLM, topoisomerase IIIalpha, and BLAP75. J Biol Chem *281*, 13861-13864.

Reddy, G., Golub, E.I., and Radding, C.M. (1997). Human Rad52 protein promotes single-strand DNA annealing followed by branch migration. Mutat Res *377*, 53-59.

Reid, L.J., Shakya, R., Modi, A.P., Lokshin, M., Cheng, J.T., Jasin, M., Baer, R., and Ludwig, T. (2008). E3 ligase activity of BRCA1 is not essential for mammalian cell viability or homology-directed repair of double-strand DNA breaks. Proc Natl Acad Sci U S A *105*, 20876-20881.

Reiling, J.H., Clish, C.B., Carette, J.E., Varadarajan, M., Brummelkamp, T.R., and Sabatini, D.M. (2011). A haploid genetic screen identifies the major facilitator domain containing 2A (MFSD2A) transporter as a key mediator in the response to tunicamycin. Proc Natl Acad Sci U S A *108*, 11756-11765.

Reinhardt, H.C., Cannell, I.G., Morandell, S., and Yaffe, M.B. (2011). Is post-transcriptional stabilization, splicing and translation of selective mRNAs a key to the DNA damage response? Cell Cycle *10*, 23-27.

Reinhardt, H.C., Hasskamp, P., Schmedding, I., Morandell, S., van Vugt, M.A., Wang, X., Linding, R., Ong, S.E., Weaver, D., Carr, S.A.*, et al.* (2010). DNA damage activates a spatially distinct late cytoplasmic cell-cycle checkpoint network controlled by MK2-mediated RNA stabilization. Mol Cell *40*, 34-49.

Rieger, K.E., and Chu, G. (2004). Portrait of transcriptional responses to ultraviolet and ionizing radiation in human cells. Nucleic Acids Res *32*, 4786-4803.

Rijkers, T., Van Den Ouweland, J., Morolli, B., Rolink, A.G., Baarends, W.M., Van Sloun, P.P., Lohman, P.H., and Pastink, A. (1998). Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation. Mol Cell Biol *18*, 6423-6429.

Rockx, D.A., Mason, R., van Hoffen, A., Barton, M.C., Citterio, E., Bregman, D.B., van Zeeland, A.A., Vrieling, H., and Mullenders, L.H. (2000). UV-induced inhibition of transcription involves repression of transcription initiation and phosphorylation of RNA polymerase II. Proc Natl Acad Sci U S A *97*, 10503-10508.

Rogakou, E.P., Boon, C., Redon, C., and Bonner, W.M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol *146*, 905-916.

Rouleau, M., McDonald, D., Gagne, P., Ouellet, M.E., Droit, A., Hunter, J.M., Dutertre, S., Prigent, C., Hendzel, M.J., and Poirier, G.G. (2007). PARP-3 associates with polycomb group bodies and with components of the DNA damage repair machinery. J Cell Biochem *100*, 385-401.

Roy, D., Zhang, Z., Lu, Z., Hsieh, C.L., and Lieber, M.R. (2010). Competition between the RNA transcript and the nontemplate DNA strand during R-loop formation in vitro: a nick can serve as a strong R-loop initiation site. Mol Cell Biol *30*, 146-159.

Rulten, S.L., Fisher, A.E., Robert, I., Zuma, M.C., Rouleau, M., Ju, L., Poirier, G., Reina-San-Martin, B., and Caldecott, K.W. (2011). PARP-3 and APLF function together to accelerate nonhomologous end-joining. Mol Cell *41*, 33-45.

San Filippo, J., Sung, P., and Klein, H. (2008). Mechanism of eukaryotic homologous recombination. Annu Rev Biochem *77*, 229-257.

Sanchez, Y., Wong, C., Thoma, R.S., Richman, R., Wu, Z., Piwnica-Worms, H., and Elledge, S.J. (1997). Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. Science *277*, 1497-1501.

Sanz, D.J., Acedo, A., Infante, M., Duran, M., Perez-Cabornero, L., Esteban-Cardenosa, E., Lastra, E., Pagani, F., Miner, C., and Velasco, E.A. (2010). A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. Clin Cancer Res *16*, 1957-1967.

Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., and Jackson, S.P. (2007). Human CtIP promotes DNA end resection. Nature *450*, 509-514.

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

Schreiber, V., Dantzer, F., Ame, J.C., and de Murcia, G. (2006). Poly(ADP-ribose): novel functions for an old molecule. Nat Rev Mol Cell Biol *7*, 517-528.

Schultz, N., Marenstein, D.R., De Angelis, D.A., Wang, W.Q., Nelander, S., Jacobsen, A., Marks, D.S., Massague, J., and Sander, C. (2011). Off-target effects dominate a large-scale RNAi screen for modulators of the TGF-beta pathway and reveal microRNA regulation of TGFBR2. Silence *2*, 3.

Scully, R., Chen, J., Ochs, R.L., Keegan, K., Hoekstra, M., Feunteun, J., and Livingston, D.M. (1997a). Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. Cell *90*, 425-435.

Scully, R., Chen, J., Plug, A., Xiao, Y., Weaver, D., Feunteun, J., Ashley, T., and Livingston, D.M. (1997b). Association of BRCA1 with Rad51 in mitotic and meiotic cells. Cell *88*, 265-275.

Scully, R., Ganesan, S., Brown, M., De Caprio, J.A., Cannistra, S.A., Feunteun, J., Schnitt, S., and Livingston, D.M. (1996). Location of BRCA1 in human breast and ovarian cancer cells. Science *272*, 123-126.

Sedelnikova, O.A., Rogakou, E.P., Panyutin, I.G., and Bonner, W.M. (2002). Quantitative detection of (125)IdU-induced DNA double-strand breaks with gamma-H2AX antibody. Radiat Res *158*, 486-492.

Shakya, R., Reid, L.J., Reczek, C.R., Cole, F., Egli, D., Lin, C.S., deRooij, D.G., Hirsch, S., Ravi, K., Hicks, J.B.*, et al.* (2011). BRCA1 tumor suppression depends on BRCT phosphoprotein binding, but not its E3 ligase activity. Science *334*, 525-528.

Shakya, R., Szabolcs, M., McCarthy, E., Ospina, E., Basso, K., Nandula, S., Murty, V., Baer, R., and Ludwig, T. (2008). The basal-like mammary carcinomas induced by Brca1 or Bard1 inactivation implicate the BRCA1/BARD1 heterodimer in tumor suppression. Proc Natl Acad Sci U S A *105*, 7040-7045.

Shanbhag, N.M., Rafalska-Metcalf, I.U., Balane-Bolivar, C., Janicki, S.M., and Greenberg, R.A. (2010). ATM-dependent chromatin changes silence transcription in cis to DNA double-strand breaks. Cell *141*, 970-981.

Shao, G., Patterson-Fortin, J., Messick, T.E., Feng, D., Shanbhag, N., Wang, Y., and Greenberg, R.A. (2009). MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks. Genes Dev *23*, 740-754.

Sharan, S.K., Morimatsu, M., Albrecht, U., Lim, D.S., Regel, E., Dinh, C., Sands, A., Eichele, G., Hasty, P., and Bradley, A. (1997). Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. Nature *386*, 804-810.

Shen, Z., Cloud, K.G., Chen, D.J., and Park, M.S. (1996). Specific interactions between the human RAD51 and RAD52 proteins. J Biol Chem *271*, 148-152.

Shin, C., Nam, J.W., Farh, K.K., Chiang, H.R., Shkumatava, A., and Bartel, D.P. (2010). Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell *38*, 789-802.

Shin, K.H., Kang, M.K., Kim, R.H., Christensen, R., and Park, N.H. (2006). Heterogeneous nuclear ribonucleoprotein G shows tumor suppressive effect against oral squamous cell carcinoma cells. Clin Cancer Res *12*, 3222-3228.

Shin, K.H., Kim, R.H., Kang, M.K., Kim, S.G., Lim, P.K., Yochim, J.M., Baluda, M.A., and Park, N.H. (2007). p53 promotes the fidelity of DNA end-joining activity by, in part, enhancing the expression of heterogeneous nuclear ribonucleoprotein G. DNA Repair (Amst) *6*, 830-840.

Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K., and Ogawa, T. (1993). Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA. Nat Genet *4*, 239-243.

Shinohara, A., Ogawa, H., and Ogawa, T. (1992). Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. Cell *69*, 457-470.

Shinohara, A., and Ogawa, T. (1998). Stimulation by Rad52 of yeast Rad51-mediated recombination. Nature *391*, 404-407.

Sigoillot, F.D., and King, R.W. (2011). Vigilance and validation: Keys to success in RNAi screening. ACS Chem Biol *6*, 47-60.

Sigoillot, F.D., Lyman, S., Huckins, J.F., Adamson, B., Chung, E., Quattrochi, B., and King, R.W. (2011). A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. Nat Methods *9*, 363-366.

Sigurdsson, S., Trujillo, K., Song, B., Stratton, S., and Sung, P. (2001). Basis for avid homologous DNA strand exchange by human Rad51 and RPA. J Biol Chem *276*, 8798-8806.

Silva, J.M., Marran, K., Parker, J.S., Silva, J., Golding, M., Schlabach, M.R., Elledge, S.J., Hannon, G.J., and Chang, K. (2008). Profiling essential genes in human mammary cells by multiplex RNAi screening. Science *319*, 617-620.

Singh, T.R., Ali, A.M., Busygina, V., Raynard, S., Fan, Q., Du, C.H., Andreassen, P.R., Sung, P., and Meetei, A.R. (2008). BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvasome. Genes Dev *22*, 2856-2868.

Singleton, M.R., Wentzell, L.M., Liu, Y., West, S.C., and Wigley, D.B. (2002). Structure of the single-strand annealing domain of human RAD52 protein. Proc Natl Acad Sci U S A *99*, 13492-13497.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol Cell *42*, 794-805.

Slabicki, M., Theis, M., Krastev, D.B., Samsonov, S., Mundwiller, E., Junqueira, M., Paszkowski-Rogacz, M., Teyra, J., Heninger, A.K., Poser, I.*, et al.* (2010). A genome-scale DNA repair RNAi screen identifies SPG48 as a novel gene associated with hereditary spastic paraplegia. PLoS Biol *8*, e1000408.

Smerdon, M.J. (1991). DNA repair and the role of chromatin structure. Curr Opin Cell Biol *3*, 422-428.

Smith, J., Tho, L.M., Xu, N., and Gillespie, D.A. (2010). The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. Adv Cancer Res *108*, 73-112.

Smogorzewska, A., Desetty, R., Saito, T.T., Schlabach, M., Lach, F.P., Sowa, M.E., Clark, A.B., Kunkel, T.A., Harper, J.W., Colaiacovo, M.P.*, et al.* (2010). A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. Mol Cell *39*, 36-47.

Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D.*, et al.* (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell *129*, 289-301.

Sobhian, B., Shao, G., Lilli, D.R., Culhane, A.C., Moreau, L.A., Xia, B., Livingston, D.M., and Greenberg, R.A. (2007). RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. Science *316*, 1198-1202.

Sobol, R.W., Prasad, R., Evenski, A., Baker, A., Yang, X.P., Horton, J.K., and Wilson, S.H. (2000). The lyase activity of the DNA repair protein beta-polymerase protects from DNA-damage-induced cytotoxicity. Nature *405*, 807-810.

Sorensen, C.S., Syljuasen, R.G., Falck, J., Schroeder, T., Ronnstrand, L., Khanna, K.K., Zhou, B.B., Bartek, J., and Lukas, J. (2003). Chk1 regulates the S phase checkpoint by coupling the physiological turnover and ionizing radiation-induced accelerated proteolysis of Cdc25A. Cancer Cell *3*, 247-258.

Soulas-Sprauel, P., Le Guyader, G., Rivera-Munoz, P., Abramowski, V., Olivier-Martin, C., Goujet-Zalc, C., Charneau, P., and de Villartay, J.P. (2007). Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. J Exp Med *204*, 1717-1727.

Soutoglou, E., Dorn, J.F., Sengupta, K., Jasin, M., Nussenzweig, A., Ried, T., Danuser, G., and Misteli, T. (2007). Positional stability of single double-strand breaks in mammalian cells. Nat Cell Biol *9*, 675-682.

Soutoglou, E., and Misteli, T. (2008). Activation of the cellular DNA damage response in the absence of DNA lesions. Science *320*, 1507-1510.

St Onge, R.P., Besley, B.D., Pelley, J.L., and Davey, S. (2003). A role for the phosphorylation of hRad9 in checkpoint signaling. J Biol Chem *278*, 26620-26628.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A., and Soreq, H. (2005). Function of alternative splicing. Gene *344*, 1-20.

Stasiak, A.Z., Larquet, E., Stasiak, A., Muller, S., Engel, A., Van Dyck, E., West, S.C., and Egelman, E.H. (2000). The human Rad52 protein exists as a heptameric ring. Curr Biol *10*, 337-340.

Stewart, G.S., Panier, S., Townsend, K., Al-Hakim, A.K., Kolas, N.K., Miller, E.S., Nakada, S., Ylanko, J., Olivarius, S., Mendez, M.*, et al.* (2009). The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. Cell *136*, 420-434.

Stewart, G.S., Wang, B., Bignell, C.R., Taylor, A.M., and Elledge, S.J. (2003). MDC1 is a mediator of the mammalian DNA damage checkpoint. Nature *421*, 961-966.

Stirling, P.C., Bloom, M.S., Solanki-Patil, T., Smith, S., Sipahimalani, P., Li, Z., Kofoed, M., Ben-Aroya, S., Myung, K., and Hieter, P. (2011a). The complete spectrum of yeast chromosome instability genes identifies candidate CIN cancer genes and functional roles for ASTRA complex components. PLoS Genet *7*, e1002057.

Stirling, P.C., Chan, Y.A., Minaker, S.W., Aristizabal, M.J., Barrett, I., Sipahimalani, P., Kobor, M.S., and Hieter, P. (2011b). R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. Genes Dev *26*, 163-175.

Stucki, M., Clapperton, J.A., Mohammad, D., Yaffe, M.B., Smerdon, S.J., and Jackson, S.P. (2005). MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. Cell *123*, 1213-1226.

Sudbery, I., Enright, A.J., Fraser, A.G., and Dunham, I. (2010). Systematic analysis of off-target effects in an RNAi screen reveals microRNAs affecting sensitivity to TRAIL-induced apoptosis. BMC Genomics *11*, 175.

Sun, Y., Jiang, X., Chen, S., Fernandes, N., and Price, B.D. (2005). A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. Proc Natl Acad Sci U S A *102*, 13182-13187.

Sung, P. (1994). Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. Science *265*, 1241-1243.

Suraweera, A., Becherel, O.J., Chen, P., Rundle, N., Woods, R., Nakamura, J., Gatei, M., Criscuolo, C., Filla, A., Chessa, L.*, et al.* (2007). Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. J Cell Biol *177*, 969-979.

Svendsen, J.M., and Harper, J.W. (2010). GEN1/Yen1 and the SLX4 complex: Solutions to the problem of Holliday junction resolution. Genes Dev *24*, 521-536.

Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. Cell *138*, 63-77.

Sy, S.M., Huen, M.S., and Chen, J. (2009). PALB2 is an integral component of the BRCA complex required for homologous recombination repair. Proc Natl Acad Sci U S A *106*, 7155-7160.

Symington, L.S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. Annu Rev Genet *45*, 247-271.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. (1983). The double-strand-break repair model for recombination. Cell *33*, 25-35.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell *116*, 51-61.

Tavtigian, S.V., Simard, J., Rommens, J., Couch, F., Shattuck-Eidens, D., Neuhausen, S., Merajver, S., Thorlacius, S., Offit, K., Stoppa-Lyonnet, D.*, et al.* (1996). The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nat Genet *12*, 333-337.

Thorslund, T., McIlwraith, M.J., Compton, S.A., Lekomtsev, S., Petronczki, M., Griffith, J.D., and West, S.C. (2010). The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. Nat Struct Mol Biol *17*, 1263-1265.

Tomita, K., Matsuura, A., Caspari, T., Carr, A.M., Akamatsu, Y., Iwasaki, H., Mizuno, K., Ohta, K., Uritani, M., Ushimaru, T.*, et al.* (2003). Competition between the Rad50 complex and the Ku heterodimer reveals a role for Exo1 in processing double-strand breaks but not telomeres. Mol Cell Biol *23*, 5186-5197.

Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A.D., Filipowicz, W., Ramos, A., Gherzi, R., and Rosenfeld, M.G. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature *459*, 1010-1014.

Trujillo, K.M., Yuan, S.S., Lee, E.Y., and Sung, P. (1998). Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95. J Biol Chem *273*, 21447-21450.

Tsui, M., Xie, T., Orth, J.D., Carpenter, A.E., Rudnicki, S., Kim, S., Shamu, C.E., and Mitchison, T.J. (2009). An intermittent live cell imaging screen for siRNA enhancers and suppressors of a kinesin-5 inhibitor. PLoS One *4*, e7339.

van Attikum, H., and Gasser, S.M. (2009). Crosstalk between histone modifications during the DNA damage response. Trends Cell Biol *19*, 207-217.

van der Heijden, T., Seidel, R., Modesti, M., Kanaar, R., Wyman, C., and Dekker, C. (2007). Real-time assembly and disassembly of human RAD51 filaments on individual DNA molecules. Nucleic Acids Res *35*, 5646-5657.

van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N., and Stuitje, A.R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell *2*, 291-299.

Van Dyck, E., Stasiak, A.Z., Stasiak, A., and West, S.C. (1999). Binding of double-strand breaks in DNA by human Rad52 protein. Nature *398*, 728-731.

Van Dyck, E., Stasiak, A.Z., Stasiak, A., and West, S.C. (2001). Visualization of recombination intermediates produced by RAD52-mediated single-strand annealing. EMBO Rep *2*, 905-909.

van Mameren, J., Modesti, M., Kanaar, R., Wyman, C., Peterman, E.J., and Wuite, G.J. (2009). Counting RAD51 proteins disassembling from nucleoprotein filaments under tension. Nature *457*, 745-748.

Venkitaraman, A.R. (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell *108*, 171-182.

Wang, B., and Elledge, S.J. (2007). Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage. Proc Natl Acad Sci U S A *104*, 20759-20763.

Wang, B., Hurov, K., Hofmann, K., and Elledge, S.J. (2009). NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. Genes Dev *23*, 729-739.

Wang, B., Matsuoka, S., Ballif, B.A., Zhang, D., Smogorzewska, A., Gygi, S.P., and Elledge, S.J. (2007). Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. Science *316*, 1194-1198.

Wang, B., Matsuoka, S., Carpenter, P.B., and Elledge, S.J. (2002). 53BP1, a mediator of the DNA damage checkpoint. Science *298*, 1435-1438.

Wang, J., Gao, Q.S., Wang, Y., Lafyatis, R., Stamm, S., and Andreadis, A. (2004). Tau exon 10, whose missplicing causes frontotemporal dementia, is regulated by an intricate interplay of cis elements and trans factors. J Neurochem *88*, 1078-1090.

Wang, M., Wu, W., Rosidi, B., Zhang, L., Wang, H., and Iliakis, G. (2006). PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. Nucleic Acids Res *34*, 6170-6182.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-738.

West, S.C. (2009). The search for a human Holliday junction resolvase. Biochem Soc Trans *37*, 519-526.

Westhorpe, F.G., Diez, M.A., Gurden, M.D., Tighe, A., and Taylor, S.S. (2010). Re-evaluating the role of Tao1 in the spindle checkpoint. Chromosoma *119*, 371-379.

Wold, M.S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu Rev Biochem *66*, 61-92.

Wong, A.K., Pero, R., Ormonde, P.A., Tavtigian, S.V., and Bartel, P.L. (1997). RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. J Biol Chem *272*, 31941-31944.

Wu, L., Davies, S.L., North, P.S., Goulaouic, H., Riou, J.F., Turley, H., Gatter, K.C., and Hickson, I.D. (2000). The Bloom's syndrome gene product interacts with topoisomerase III. J Biol Chem *275*, 9636-9644.

Wu, L., and Hickson, I.D. (2003). The Bloom's syndrome helicase suppresses crossing over during homologous recombination. Nature *426*, 870-874.

Wu, L.C., Wang, Z.W., Tsan, J.T., Spillman, M.A., Phung, A., Xu, X.L., Yang, M.C., Hwang, L.Y., Bowcock, A.M., and Baer, R. (1996). Identification of a RING protein that can interact in vivo with the BRCA1 gene product. Nat Genet *14*, 430-440.

Xia, B., Sheng, Q., Nakanishi, K., Ohashi, A., Wu, J., Christ, N., Liu, X., Jasin, M., Couch, F.J., and Livingston, D.M. (2006). Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. Mol Cell *22*, 719-729.

Xu, D., Guo, R., Sobeck, A., Bachrati, C.Z., Yang, J., Enomoto, T., Brown, G.W., Hoatlin, M.E., Hickson, I.D., and Wang, W. (2008). RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. Genes Dev *22*, 2843-2855.

Yamaguchi-Iwai, Y., Sonoda, E., Buerstedde, J.M., Bezzubova, O., Morrison, C., Takata, M., Shinohara, A., and Takeda, S. (1998). Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. Mol Cell Biol *18*, 6430-6435.

Yan, C.T., Boboila, C., Souza, E.K., Franco, S., Hickernell, T.R., Murphy, M., Gumaste, S., Geyer, M., Zarrin, A.A., Manis, J.P.*, et al.* (2007a). IgH class switching and translocations use a robust non-classical end-joining pathway. Nature *449*, 478-482.

Yan, J., Kim, Y.S., Yang, X.P., Li, L.P., Liao, G., Xia, F., and Jetten, A.M. (2007b). The ubiquitin-interacting motif containing protein RAP80 interacts with BRCA1 and functions in DNA damage repair response. Cancer Res *67*, 6647-6656.

Yang, H., Jeffrey, P.D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N.H., Zheng, N., Chen, P.L., Lee, W.H., and Pavletich, N.P. (2002). BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. Science *297*, 1837-1848.

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

Yazdi, P.T., Wang, Y., Zhao, S., Patel, N., Lee, E.Y., and Qin, J. (2002). SMC1 is a downstream effector in the ATM/NBS1 branch of the human S-phase checkpoint. Genes Dev *16*, 571-582.

Yin, J., Sobeck, A., Xu, C., Meetei, A.R., Hoatlin, M., Li, L., and Wang, W. (2005). BLAP75, an essential component of Bloom's syndrome protein complexes that maintain genome integrity. EMBO J *24*, 1465-1476.

Yoo, S., and Dynan, W.S. (1998). Characterization of the RNA binding properties of Ku protein. Biochemistry *37*, 1336-1343.

Yoshimura, Y., Morita, T., Yamamoto, A., and Matsushiro, A. (1993). Cloning and sequence of the human RecA-like gene cDNA. Nucleic Acids Res *21*, 1665.

Youds, J.L., Mets, D.G., McIlwraith, M.J., Martin, J.S., Ward, J.D., NJ, O.N., Rose, A.M., West, S.C., Meyer, B.J., and Boulton, S.J. (2010). RTEL-1 enforces meiotic crossover interference and homeostasis. Science *327*, 1254-1258.

Yu, X., and Baer, R. (2000). Nuclear localization and cell cycle-specific expression of CtIP, a protein that associates with the BRCA1 tumor suppressor. J Biol Chem *275*, 18541-18549.

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

Yu, X., Chini, C.C., He, M., Mer, G., and Chen, J. (2003). The BRCT domain is a phospho-protein binding domain. Science *302*, 639-642.

Yun, M.H., and Hiom, K. (2009). CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. Nature *459*, 460-463.

Zhang, F., Ma, J., Wu, J., Ye, L., Cai, H., Xia, B., and Yu, X. (2009). PALB2 links BRCA1 and BRCA2 in the DNA-damage response. Curr Biol *19*, 524-529.

Zhang, R., Poustovoitov, M.V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L.*, et al.* (2005). Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Dev Cell *8*, 19-30.

Zhang, X., Wan, G., Berger, F.G., He, X., and Lu, X. (2011). The ATM kinase induces microRNA biogenesis in the DNA damage response. Mol Cell *41*, 371-383.

Zhou, B.B., and Elledge, S.J. (2000). The DNA damage response: putting checkpoints in perspective. Nature *408*, 433-439.

Zhu, J., Petersen, S., Tessarollo, L., and Nussenzweig, A. (2001). Targeted disruption of the Nijmegen breakage syndrome gene NBS1 leads to early embryonic lethality in mice. Curr Biol *11*, 105-109.

Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell *134*, 981-994.

Zlotnick, A., Mitchell, R.S., Steed, R.K., and Brenner, S.L. (1993). Analysis of two distinct single-stranded DNA binding sites on the recA nucleoprotein filament. J Biol Chem *268*, 22525-22530.

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