Genetic analysis of DNA double-strand break mis-repair mechanisms using the human endogenous *HPRT* gene

DPhil thesis supervised by
Dr. Timothy Humphrey and Dr. Andy Porter

Sara Ahrabi



CRUK MRC Institute for Radiation Oncology
Department of Oncology, University of Oxford

Jesus College

August 2017

Acknowledgements

I would like to start by thanking Dr. Tim Humphrey, who I was really blessed to have as my supervisor. He is a wonderful mentor, with amazing people skills, great sense of humour, and patience. Tim has definitely been fundamental in my transition from being a clinician to a scientist. Without his support and encouragement I would not have achieved anything like this today. He believed in me and gave me the confidence to think independently and critically and to develop my ideas.

Also I would like to extend my special thanks to Dr. Andy Porter, who I had the pleasure to have as my second supervisor. Andy has always taken the time to offer some invaluable insight. I have indeed learnt a lot from Andy, among many more probably the most fundamental one that comes to my mind is that he taught me the basics of mammalian cells' tissue culture techniques!

I am grateful for the funding provided by the Clarendon and CR-UK scholarships.

A big thank you to all the present and past members of the Humphrey lab. Especially Carol, Sovan, Lizzy, Sophia, and Torben who I learnt a lot from. Special thanks to Dr Sovan Sarkar, who helped me extensively with the NAR manuscript. Also I have to thank all wonderful people I met during my DPhil in Oxford institute for radiation oncology. Sorry that I cannot name you all here but you are all in my mind. My Oxford experience would not have been the same without your friendships.

I feel blessed to have my husband, Kambiz, with me during this journey. Without his constant help and motivation, I would have never made it to where I am now. Apart from the emotional support he has always given me, he has been wonderfully taking care of our bundle of joy, Nikaan, during my thesis write up. Also I cannot thank him enough for his expert bioinformatics advice on my work and for critically reading my thesis. Thank you Nikaan for being so patient during my write up!

Also I would like to thank my brother, Saeed, who we grew up together as best friends. Saeed has always been a loving and caring friend to me who I could talk to about anything. He deserves my wholehearted thanks.

Finally I would like to dedicate my thesis to my wonderful parents. I thank them for their endless love, support, and encouragement throughout my life. Thank you both for believing in me and giving me the strength and confidence to chase my dreams. My special thanks go to my father, a truly kind hearted man, who has always been there for me. Also I should thank him for his expert statistical advice on my thesis. I cannot thank my mother enough for her unconditional love and for teaching me to believe and achieve. I owe my inspiration and determination to my mother, this amazing strong woman.

Abstract

DNA double strand breaks (DSBs) are the most lethal radiation-induced lesions in response to which cells employ either the error free homologous recombination (HR) repair pathway or error prone mechanisms, such as non homologous end joining (NHEJ) and microhomology- mediated end joining (MMEJ). While MMEJ is suppressed by C-NHEJ, the relationship between HR and MMEJ is less clear. In this thesis, I have exploited the human endogenous *HPRT* gene to develop a novel genetic technology to detect mutation frequencies and signatures in response to DNA DSBs in different genetic backgrounds. Using a sensitive HPRT assay we describe a role for HR genes in suppressing MMEJ in human cells. By monitoring DSB mis-repair, we found that depletion of HR proteins, including RAD51, BRCA2, BRCA1 or SETD2, resulted in a distinct mutational signature associated with significant increases in break-induced mutation frequencies, deletion lengths and the annealing of short regions of microhomology (2–6 bp) across the break-site. This signature was dependent on CtIP, MRE11, POLQ and PARP, and thus indicative of MMEJ. In contrast to CtIP or MRE11, depletion of BRCA1 resulted in increased partial resection and MMEJ, thus revealing a functional distinction between these early acting HR factors. Together these findings indicate that HR factors suppress mutagenic MMEJ following DSB resection.

Further, I have defined a role for the SETD2 histone methyltransferase in suppressing break-induced mutations, and have shown that CRISPR/Cas9 and I-SceI-induced DSBs resulted in different repair profiles.

Contents

1. Introduction - - -

2.2. Cell lines - - -

4. **A novel role for homologous recombination factors in suppressing microhomology-mediated end joining in human cancer cells** - - - 66

4.1. BRCA2 depletion promotes DSB-induced deletions typical of MMEJ - - - 66

4.2. BRCA1 depletion promotes DSB-induced deletions typical of MMEJ - - - 73

4.3. MMEJ-like DSB repair in HR deficient cells is independent of DNA-PKcs - - - 78

4.4. DSB-induced mutation signatures in HR-deficient cells require MMEJ factors - - 84

4.5. RPA suppresses MMEJ - 95

4.6. Discussion for Chapter 4 - 102

4.6.1. A model for HR repressing MMEJ - 102

4.6.2. Future directions - 104

5. **CRISPR-Cas9 and I-SceI -induced DSBs generate distinct repair profiles** - - - - - - 105

5.1.Targeting the HPRT locus using the CRISPR-Cas9 technology - - - - - - - 105

5.2. Comparison of CRISPR/Cas9- and I-SceI-induced mutation profiles at the HPRT locus - - 109

5.3. Discussion for chapter 5 - 115

6. Developing a novel I-SceI genetic assay to analyse the mechanisms of DSB mis-repair - 116

6.1. Introducing the I-SceI site into the intron 5 of HPRT gene - - - - - - - - - - - 116

6.2. Integrating I-SceI site into intron 5 of HPRT gene in HT1080 cells - - - - - 120

6.3. Characterisation of the Intron assay -- 123

6.4. Discussion for chapter 6 - 126

7. Discussion - 127

7.1. The current understanding of the interplay between different DSB repair pathways 127

7.2. Could the “HR-MMEJ” story explain the underlying mechanism of synthetic lethality between HR and PARP inhibition? - - - - - 134

7.3. The HPRT system links mutation signatures to the driver mutations - - - - 135

7.4. The HPRT assay; capacities, and limitations - 136

7.4.1.The unmet need for an “all-in-one” DSB mutation assay - - - - - - - - - 136

7.4.2. Capacity of the HPRT system - 137

7.4.3. A limitation of the HPRT system - 138

7.5. Distinct mutation signature following I-SceI vs CRISPR/Cas9 induced - - breaks 138

7.6. Relatively low I-SceI-induced mutation frequencies - - - - - - - - - - - - - - - 140

7.7. Future directions - 140

Bibliography - 143

Appendix - 152

Chapter 1

Introduction

1.1. Cancer and genome instability

Cancer is defined as a collection of diseases in which cells acquire the ability to divide and invade to other tissues uncontrollably. All cancers share the underlying principles, also known as the hallmarks of cancer. The hallmarks include the ability to proliferate indefinitely by escaping the cell death/senescence mechanisms, as well as acquiring the capability to invade and metastasise. Importantly, the mechanism underpinning these hallmarks and the most enabling characteristic is “genome instability”, which enables the acquisition of cancer hallmarks by generating random mutations within genome (Hanahan and Weinberg, 2011).

1.2 DNA repair

1.2.1. DNA damage response

Multiple types of internal and external DNA damaging factors constantly threaten the genetic information encoded in DNA. There are different kinds of DNA damage and to overcome each, cells have evolved different repair mechanisms. There are various kinds of DNA damage, each repaired by distinct pathways as described below.

DNA single-stranded breaks (SSBs) and single base damage are caused by ionising radiation (IR) or oxidative stress. These are mainly repaired via base-excision repair (BER) (David et al., 2007).

DNA intrastrand cross-links and bulky-adducts are generated by chemicals or UV

light. These are predominantly repaired by nucleotide excision repair (NER). In addition, some of these lesions are bypassed by translesion synthesis, and are catalysed by special DNA polymerases such as Polη (Roos et al., 2016, Loeb and Monnat, 2008).

DNA mismatch lesions largely form during DNA synthesis as a result of low fidelity DNA replication. These mismatched bases in the newly synthesised DNA strand are removed by the mismatch repair machinery (Jiricny, 2006).

DNA double-stranded breaks (DSBs) are generated by IR, oxidative damage, replication across a SSB, occurrence of two SSBs at close proximity, enzymatic activity, or via collapsed replication forks (Jackson and Bartek, 2009).

Table 1.1 summarises the DNA damage types and the associated repair pathway, as well as the key repair proteins involved in each pathway. Among the different kinds of DNA damage, double strand breaks (DSBs) in which both strands of DNA break simultaneously, are the most toxic (Mehta and Haber, 2014). Thus, if left unrepaired, DSBs can lead to cell death, while if mis-repaired, they can give rise to genomic instability, leading to tumourigenesis (Jeggo and Lavin, 2009).

Table 1.1. Different types of DNA damage, their repair pathways, and the key factors.

DNA damage type	Repair pathway	Key repair factors
Single base damage Single Strand Breaks (SSBs)	Base Excision Repair (BER)	AP endonuclease PARP1/2 XRCC1 Polβ Lig III
Bulky adducts Intrastrand crosslink	Nucleotide excision repair (NER)	XP proteins ERCC1 Polδ Lig I
Base Mismatch	Mismatch repair (MMR)	MSH2/3/6, MLH1 EXO1 Polδ LigI
Double Strand Breaks (DSBs)	HR, NHEJ, Alt-NHEJ, SSA	Described in detail below

To survive such genomic insults and preserve genome integrity, cells possess two main evolutionarily conserved DSB repair mechanisms, namely homologous recombination (HR), and non-homologous end joining (NHEJ)(Symington and Gautier, 2011). Other repair pathways include single strand annealing (SSA) and alternative non-homologous end joining pathways (Alt-NHEJ) (Kabotyanski et al., 1998, Liang and Jasin, 1996, Boulton and Jackson, 1996, Ciccia and Elledge, 2010). Importantly, a subset of alternative non-homologous end joining pathways relies on regions of microhomology on either side of the break, which anneal following limited resection in a process called microhomology-mediated end joining (MMEJ) (Decottignies, 2013, Simsek et al., 2011, Bennardo et al., 2008). The key steps in each of these DSB repair pathways are outlined below.

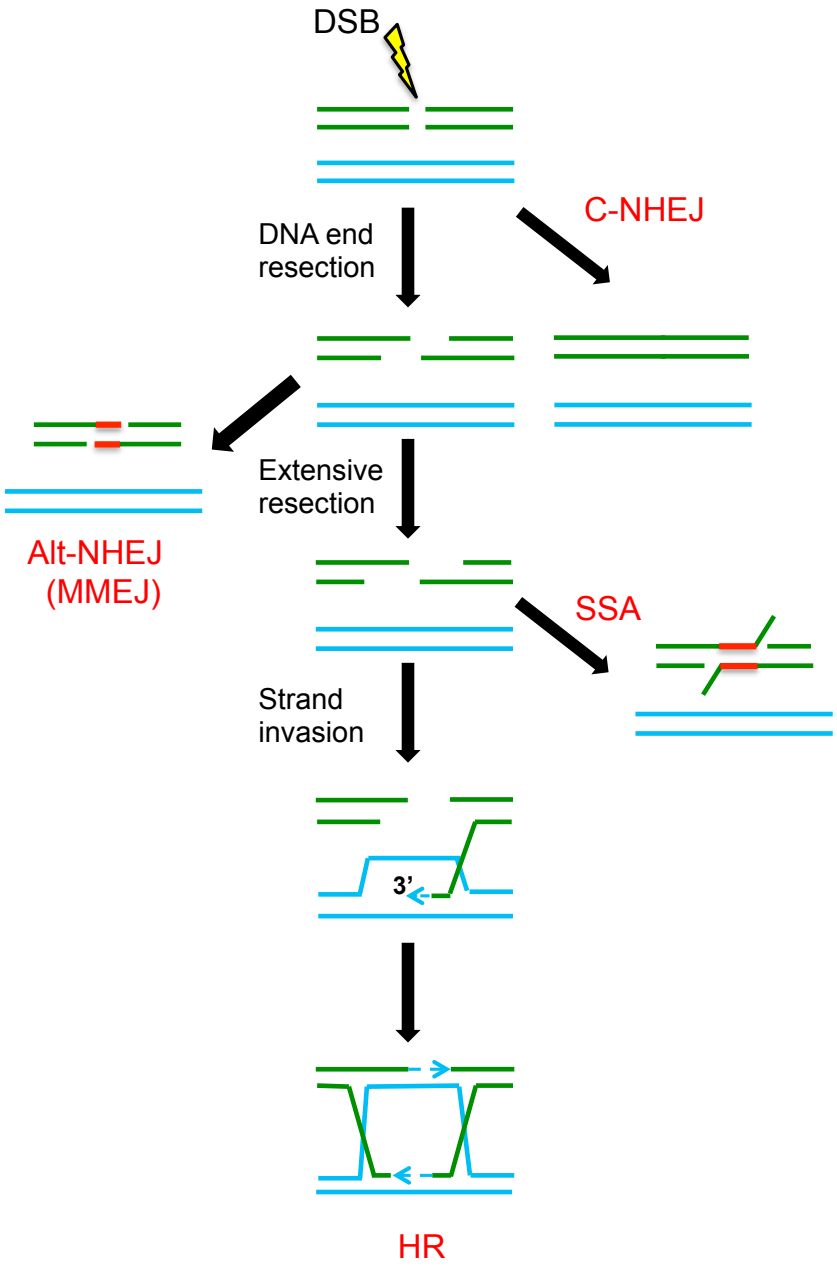


Figure 1.1. Overview of DNA DSB repair pathways. NHEJ does not require resection; MMEJ requires minimal resection; single-stranded annealing (SSA) requires extensive resection; and homologous recombination (HR) requires extensive resection followed by strand invasion.

1.2.2. DNA double strand break (DSB) repair pathways

Homologous recombination

Homologous recombination (HR) is usually active in S and G2 phases of the cell cycle during which the sister chromatids are present to serve as the template for homology-directed synthesis. Therefore, HR is commonly considered to be an error-free DSB repair pathway. HR proceeds through three phases. In mammalian cells the presynaptic phase is triggered by a two-step 5' to 3' end resection that produces 3' single-stranded DNA (ssDNA) overhangs. Resection is initiated by the endonucleolytic activity of the MRE11-RAD50-NBS1 (MRN) complex and the C terminal binding protein interacting protein (CtIP), which exposes short ssDNA tails (Sartori et al., 2007, Takeda et al., 2007). These become substrates for the extensive resection mediators, EXO1, DNA2 and BLM (Gravel et al., 2008, Nimonkar et al., 2008). BRCA1 also facilitates the initial resection step of HR (Stark et al., 2004, Moynahan et al., 1999) in conjunction with MRN (Zhong et al., 1999) and CtIP (Zhong et al., 1999, Yu et al., 1998), where it accelerates the DSB resection rate (Cruz-García et al.). The exposed ssDNA is initially protected by Replication Protein A (RPA)(Wold, 1997), which is then displaced by RAD51, following its recruitment by BRCA2, to form a nucleoprotein filament (Moynahan et al., 2001). The RAD51 nucleofilament promotes strand invasion of the undamaged sister chromatid, which is used as a repair template, resulting in a displacement loop (D-loop). During the synaptic phase of HR, the 3` end is extended by DNA replication, which can subsequently proceed through a number of sub-pathways. During DSB repair, second end capture and annealing results in double Holliday junction (HJ) formation, which can be resolved by GEN1, MUS81/EME1 or SLX1/SLX4 (Ciccia and Elledge, 2010). In the post-synaptic phase of HR, HJ structures can be resolved with or without crossovers, or dissolved, thus preventing crossovers (Wu and Hickson, 2003,

West, 2003). Alternatively, during synthesis-dependent strand annealing (SDSA) that is mediated via RTEL1 helicase (Nassif et al., 1994, Barber et al., 2008), the invading and extended strand is expelled from the D-loop to anneal to the second end which,

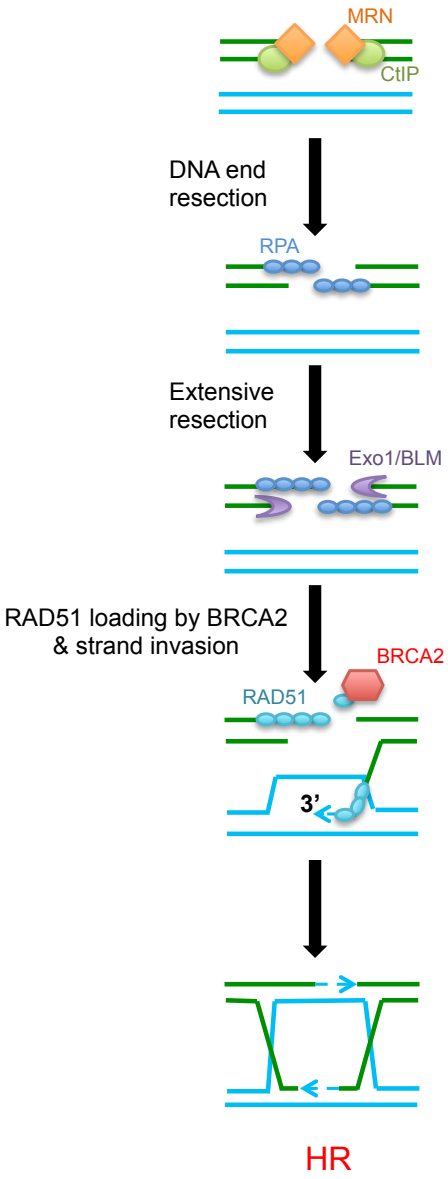


Figure 1.2. The key steps and key proteins involved in HR repair.

following gap filling and ligation, results in error-free repair (Ciccia and Elledge, 2010). The key steps of HR are summarised in a schematic diagram (Fig 1.2).

Non-homologous end joining (NHEJ)

Non-homologous end joining (NHEJ) is active through all phases of cell cycle. Classical NHEJ (C-NHEJ) is triggered by recognition and protection of DNA ends by the Ku70/ Ku80 heterodimer, which forms a ring that encircles duplex DNA. This protects ends from resection and creates a platform to recruit the DNA-PK catalytic subunit (DNA-PKcs) (Deriano and Roth, 2013, Mahaney et al., 2009). Broken ends are then trimmed by Artemis and ligated by DNA Ligase 4 (Lig 4), X-ray repair cross complementing protein 4 (XRCC4) complex, and XRCC4- like factor (XLF), depending on the nature of the damage (Mahaney et al., 2009, Lieber, Meek et al., 2008). Although end-protection by Ku in this pathway minimises resection, thus promoting error-free end joining, this pathway is widely referred to as error-prone as it ligates the ends in a homology-independent fashion potentially leading to small insertions, and/ or deletions (indels) at the DSB sites. From a genome-wide perspective, however, C-NHEJ is not as threatening as alternative NHEJ (Alt-NHEJ) pathways for mammalian genome stability (Deriano and Roth, 2013) and is even considered as a guardian of genome stability (Ferguson et al., 2000).

Importantly, there are a number of NHEJ sub pathways depending on the DNA end structures. For instance blunt ends are recognised by Ku, which then recruits the XRCC4–DNA ligase IV to the DNA ends. XRCC4 associates with DNA ligase IV in a 2:1 ratio and bridges between the two DNA ends. DNA end bridging is independent of DNA-PKcs and the blunt ends are ligated via Ku–XRCC4–DNA ligase IV complex with no DNA end processing.

However most ends need processing due to chemical modifications or mismatching overhangs. In NHEJ, this involves the degradation of short regions of the 5' or 3' overhangs by either exonuclease or endonuclease activity to expose or to generate small regions of microhomology (≤4nucleotides) between the strands that can

facilitate end joining. This differs from the extensive end resection (≥20nucleotides) that occurs to initiate HR repair pathway. When DNA resection is required for NHEJ, DNA-PKcs is recruited in complex with the endonuclease Artemis to trim the ends by limited resection. DNA-PKcs undergoes autophosphorylation that in turn activates Artemis, which then gains the ability to cut DNA. 20-50% of ionizing radiation induced DSBs are reliant on Artemis for repair (Chang et al., 2017). The key steps of C-NHEJ are summarised in a schematic diagram (Fig 1.3).

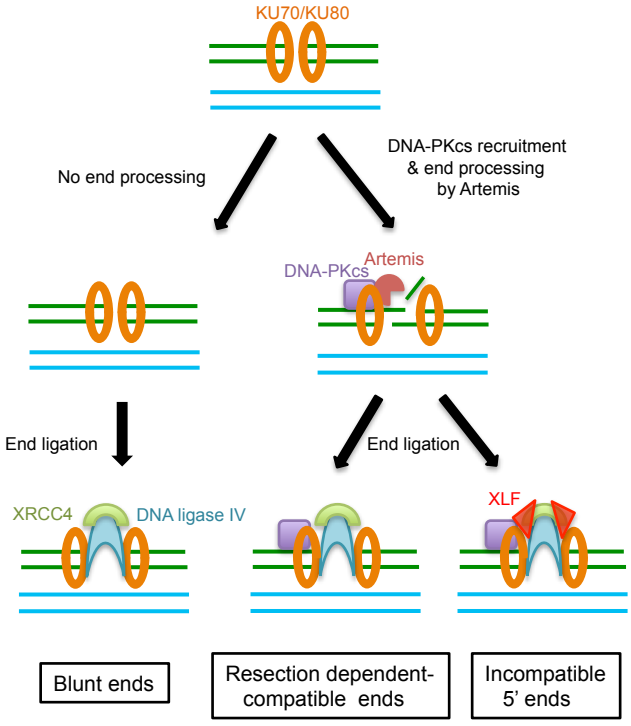


Figure 1.3. The key steps and key proteins involved in C-NHEJ repair.

In brief, following a DSB the Ku heterodimer binds to the DNA end. This in turn recruits DNA-PKcs (This is not required for the blunt ends). Finally, a complex of DNA ligase IV, XRCC4, XLF, and PAXX carries out the re-joining step (Lobrich and Jeggo, 2017). Blunt ends, are preferably rejoined without end processing, while the compatible ends with short stretches of microhomologies and incompatible 5' overhangs undergo end processing by the Artemis-DNA-PKcs complex (Chang et al., 2017).

Alternative non-homologous end joining (Alt-NHEJ)

Alt-NHEJ refers to DSB end joining pathways that are independent of the C-NHEJ factors Ku70/ Ku80, DNA PKcs and DNA Lig4. These are highly mutagenic and are commonly known to be associated with indels and chromosomal rearrangements. However, recently it has been revealed that breakpoint junctions of the chromosomal translocations in human cancers display signature of NHEJ pathway. This is in sharp contrast with the findings in mouse cells in which translocations arise from Alt-NHEJ pathways. Thus, unlike mouse cells, translocations in human cells are generated by C-NHEJ(Ghezraoui et al., 2014).

Importantly, a sub pathway of Alt-NHEJ events termed microhomology mediated end joining (MMEJ) rejoins the ends by base pairing between microhomologous sequences. MMEJ is mediated via CtIP, MRN complex, Poly [ADP-ribose] polymerases1 (PARP-1) and DNA ligase 3 (Lig 3) (Ghezraoui et al., 2014, Jones et al., 2014). In mammalian cells, PARP-1, a key component of the MMEJ pathway (Audebert et al., 2004), initially competes with Ku heterodimer for binding to DNA ends resulting in PAR formation (Wang et al., 2006a, Hochegger et al., 2006). This in turn promotes MRE11 recruitment and the initiation of resection by MRN complex and CtIP (Haince et al., 2008). Subsequently, DNA end ligation is mediated through DNA Lig 3 (Wang et al., 2005, Simsek et al., 2011). Further, recent studies have identified a low fidelity DNA polymerase θ (Polθ also known as POLQ) (Seki et al., 2004) as an important MMEJ factor in mammalian cells. POLQ is recruited to DSB ends by PARP-1 where it facilitates end joining and microhomology annealing (Yousefzadeh et al., 2014). While absent from yeast, the role for POLQ in MMEJ appears to be conserved in flies, worms, mice and humans (Kent et al., 2015). MMEJ generates deletions and translocations at the break points and thus is highly mutagenic (McVey and Lee, 2008). However, there are other types of Alt-NHEJ

events that do not rely on such microhomologies. The ligation step in these pathways is facilitated through DNA ligase 1(Lig1) (Decottignies, 2013). The key steps of MMEJ are summarised in a schematic diagram (Fig 1.4).

Studies in mammalian cells have demonstrated that not only is MMEJ a back up repair mechanism in the absence of C-NHEJ essential factors but it is also frequently used to repair DSBs in normal cycling cells even in the presence of both HR and C-NHEJ(Sfeir and Symington, 2015). It has also been illustrated that MMEJ levels increase when cells enter S phase and that MMEJ has a dynamic role at collapsed replication forks (Truong et al., 2013). However, there is very little known about the contribution of MMEJ in DSB repair in cells lacking the HR core components.

Recently another view on the classification of MMEJ and Alt-NHEJ repair pathways has been discussed, in which MMEJ is also defined as a form of DNA re-joining that uses short microhomology regions. However, because Alt-NHEJ is not involved in the DSB repair in G0/G1 phase, this view argues that probably most MMEJ events in G0/G1 are carried out through Artemis-dependent C-NHEJ. From this perspective, therefore MMEJ is defined as DSB re-joining events that exploit the microhomology regions regardless of the pathway involved (Lobrich and Jeggo, 2017).

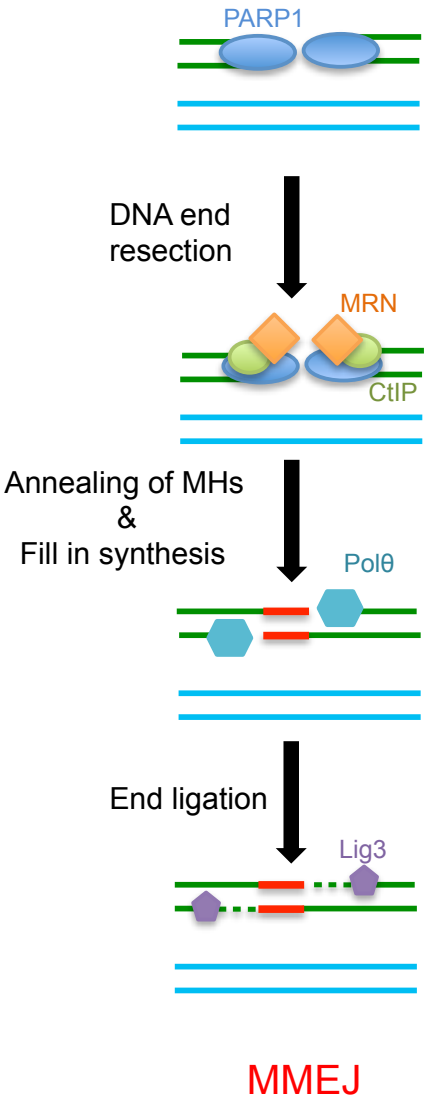


Figure 1.4. The key steps and key proteins involved in MMEJ repair.

Single Strand Annealing (SSA)

Single-strand annealing (SSA) mostly takes place in S and G2 phases of cell cycle. According to the current literature SSA occurs when the DSB resection step occurs at repetitive sequences and this provides complementary single strands that can then anneal (Bhargava et al.). The SSA repair pathway involves repeats flanking the DSB site and is associated with deletion rearrangement between the repeats, and thus is a mutagenic pathway. The critical step of SSA that is also shared with HR and MMEJ is the initiation of resection via CtIP to generate 3′ ssDNA (Fig 1.1). However, the degree of end resection varies between these pathways and is probably a determining factor in the repair pathway choice (Bhargava et al.). While the initial resection step is followed by extensive resection mediated by EXO1/BLM in HR and SSA (Ciccia and Elledge, 2010) MMEJ only requires limited degree of resection. Upon resection in the SSA pathway, RAD52 facilitates the direct annealing of ssDNA ends harbouring homology regions, followed by trimming of the stretches of ssDNA that cannot pair with the homology sequences (Motycka et al., 2004) via ERCC1/XPF endonuclease and filling of gap by DNA polymerase.

Among different DNA DSB repair mechanisms, SSA and Alt-NHEJ result in the most mutagenic outcomes. They share the initial resection step, as well as the annealing step that occurs towards the homology regions at the break junctions. However deletions in SSA result from extensive end-resection followed by annealing of longer (e.g. >100 nt) repeats, whereas MMEJ involves limited end resection resulting in deletions at the site of the DSB. Also MMEJ often involves the annealing of short (e.g. 2–6 nt) sequence homologies from each side of the DSB. Further these repair pathways are mechanistically distinct and possess different molecular requirements.

1.2.3. Interplay between DSB repair pathways

Impairment of a DSB repair pathway is a common feature in human cancers, rendering tumour survival highly dependent on secondary DSB repair pathways (Kennedy and D'Andrea, 2006). Therefore, studies of the relationships between different DSB repair machineries are of interest as they can provide insight into how to exploit these defects for cancer therapy purposes. Notable in this regard, is the fact that HR has an inhibitory impact on C-NHEJ via initiation of resection. This is probably due to the poor binding affinity of the Ku70/80 heterodimer to the resected DNA ends, which results in reduced efficiency of NHEJ (Huertas, 2010).

Conversely, C-NHEJ suppresses HR and Alt-NHEJ by Ku binding thus preventing the resection initiation (Iliakis, 2009, Cheng et al., 2011). However, the exact nature of the interplay between HR and Alt-NHEJ in human cells still remains elusive. A recent study in MEFs (Mouse Embryonic Fibroblast Cells) found elevated levels of PARP-1-dependent translocations following loss of the HR factor RAD54 (Soni et al., 2014). Moreover, a study in *Saccharomyces cerevisiae* (*S. cerevisiae*) showed that RPA impairs MMEJ by removing the secondary structures of ssDNA, which facilitates RAD51 filament assembly leading to HR and thereby suppression of annealing of the terminal microhomologies (Deng et al., 2014). On the other hand, two recent studies in mammalian cells suggested that POLQ-mediated Alt-NHEJ suppresses HR repair mechanisms (Mateos-Gomez et al., 2015, Ceccaldi et al., 2015).

Multiple histone modifications have been described in the context of DNA DSB repair. However for the purpose of this thesis, I focus on H3K36me3 in the DSB repair context.

1.2.4. Histone H3K36 methylation and DSB repair

In humans, H3K36 methylation is catalysed by at least eight methyltransferases. Out of these, seven (NSD1, NSD2, NSD3, SETD2, ASH1L, SETM, and SMYD2) are responsible for H3K36 mono- and/or di-methylation, while the eighth (SETD2) uniquely catalyses the di- to tri-methylation of H3K36 (Sun et al., 2005, Edmunds et al., 2008). In all species, H3K36 is associated with transcription, where SET2 interacts with RNAP II via its C-terminal domain (Kizer et al., 2005, Aymard et al., 2014). In yeast, H3K36 methylation results in decreased chromatin accessibility, DSB resection and enhances NHEJ via recruitment of the ssDNA-dependent and ATP-dependent helicase Ku. In contrast, chromatin accessibility is increased through Gcn5-dependent H3K36 acetylation, thereby increases DSB resection and promotes HR. This suggests that histone H3K36 chromatin switch contributes to DSB repair pathway choice (Pai et al., 2014).

In mammalian cells, dimethylated H3K36 is present within approximately 40% of nucleosomes (Garcia et al., 2008). However, increased H3K36 dimethylation levels are observed within 15 minutes following exposure to IR. This is associated with the appearance of SETMAR at the break site, which dimethylates H3K36. This histone mark is required for the recruitment of Ku70 and NBS1 to a break site, thus enhancing DSB repair by NHEJ (Fnu et al., 2011). H3K36me2 levels are also regulated via demethylation following DNA damage. In response to a DSB, ATM interacts with and phosphorylates the demethylase KDM2A within its PHD domain at T632. While the KDM2A activity is not altered, its phosphorylation abrogates chromatin binding, which leads to increased H3K36 dimethylation at the break sites, and reduced H3K36 monomethylation. ATM-dependent KDM2A phosphorylation is essential for enhanced H3K36me2 in response to DSB inducing agents, and H3K36me2 interacts with NBS1 directly, thus promoting DSB repair through NHEJ and cell survival (Cao et al., 2016).

Histone H3K36 trimethylation is a highly conserved chromatin mark associated with transcriptional elongation, and accumulates across the body of actively transcribed genes (Edmunds et al., 2008). In contrast to H3K36me2, H3K36me3 is not altered following DNA damage, and appears to be pre-established (Aymard et al., 2014, Carvalho et al., 2014, Pfister et al., 2014, Fnu et al., 2011).

DNA repair can be facilitated through chromatin modification by chromatin remodeling complexes, by incorporation of histone variants, and by histone modification (Smeenk and van Attikum, 2013). Roles for H3K36 modification have been identified in DNA repair. For example, H3K36me2 increases at DSB sites and enhances DNA repair by NHEJ through recruiting NBS1 and Ku70 repair factors (Fnu et al., 2011). Furthermore, H3K36me3 has also been shown to regulate DNA mismatch repair by recruiting the mismatch repair protein MSH6 to the chromatin, bringing it closer to the DNA replication machinery (where mismatch lesions first occur) (Li et al., 2013). However, SETD2-deficient cancers manifest a wide range of mutations, including indels, and chromosomal aberrations (Sato et al., 2013, Zhu et al., 2014), which usually stem from mis-repair of DSBs. This suggests an additional role for SETD2 in genome stability, likely through accurate repair of DSBs.

1.2.5. The SETD2 gene and protein and its role in cancer

The *SETD2* gene (also known as Huntingtin-interacting protein B, HYPB) is located on the short arm of chromosome 3 (3p21.31), spanning 147.57 kb across 21 exons and encodes a protein of 2564 amino acids. The SETD2 protein contains a SET domain (flanked by the AWS and Post-SET domain) that is responsible for the catalytic activity of SETD2. The WW domain regulates the interaction between SETD2 and the huntingtin protein, contributing to the pathology of Huntington's disease (Gao et al., 2014). The SRI domain mediates the binding of SETD2 to RNA polymerase II (Li et al., 2005), which is essential for the transcriptional-associated trimethylation of H3K36. There are various processes in cancer that contribute to the

loss of SETD2. Sequence analysis indicates that loss of function could be caused by numerous kinds of mutations, including missense, substitution nonsense and frameshift mutations (Sanger Insititute). Large deletions on chromosome 3p can also result in deletion of the *SETD2* gene.

Mutations in the human *SETD2*, the tumour suppressor gene, have been recently identified in many cancers, including breast (Al Sarakbi et al., 2009), renal carcinomas (Dalgliesh et al., 2010), high-grade gliomas (Fontebasso et al., 2013),) and acute lymphoblastic leukemia (ALL) (Zhang et al., 2012). Also SETD2 has been found to be one of the most frequently mutated genes (Lawrence et al., 2014) Notably , SETD2 mutations have been identified to be correlated with poor prognosis in breast and renal cancers (Al Sarakbi et al., 2009) (Hakimi et al., 2013). Remarkably, H3K36me3 loss is seen in 54% of pediatric high-grade gliomas (Fontebasso et al., 2013). Also, SETD2 mutations has recently been associated with resistance to DNA damaging chemotherapy in leukemia, suggesting a role for this gene in DNA repair (Mar et al., 2017).

1.2.6. SETD2 and DNA repair

SETD2 is a histone methyltransferase that in mammalian cells is exclusively in charge of the trimethylation of the histone H3K36 (Fahey and Davis, 2017) . SETD2 is also a tumour suppressor and is mutated across a range of human cancers including breast, lung, acute lymphoblastic leukemia, clear cell renal cell carcinoma, and glioma (Al Sarakbi et al., 2009, Dalgliesh et al., 2010, Fontebasso et al., 2013, Newbold and Mokbel, 2010, Zhang et al., 2012). SETD2-dependent H3K36 trimethylation has been reported to play a role in DNA mismatch repair (Li et al., 2013). However, because SETD2-deficient cancers harbour various other types of mutations including indels, and chromosomal aberrations, an additional role for this tumour suppressor in promoting genome stability through DSB repair is also speculated (Sato et al., 2013, Zhu et al., 2014).

1.3 I-SceI-based DSB reporter assays

1.3.1 I-SceI endonuclease

I-SceI is an intron-encoded endonuclease, which was initially discovered within the mitochondria of *Saccharomyces cerevisiae. The* meganuclease I-SceI is a homing endonuclease that recognises an 18-base pair sequence (TAGGGATAACAGGGTAAT) and generates a 4 base pair overhang. Statistically this 18 bp sequence will occur at a frequency of 1 in 4^{18} , and does not normally occur in a human or mouse genome. Thus makes an excellent genetic tool to generate site specific DSBs within the genomes of human or mouse cells (Jasin, 1996).

1.3.2. Exogenous DSB reporters

As outlined above, DSBs can be repaired via a number of different repair mechanisms that result in various genetic outcomes. In order to examine distinct repair outcomes upon DSB induction, an array of DSB reporters constructs has been developed that usually contain recognition sites for the rare-cutting endonuclease I-SceI (Pierce et al., 1999, Ogiwara et al., 2011, Bennardo et al., 2008). Typically, an exogenous reporter vector containing the I-SceI recognition site is introduced into the genome of the host cells of interest, and cleaved by transient transfection of an I-SceI expression vector. The reporter constructs are designed such that the reporter gene is expressed only after the I-SceI-induced DSB is repaired by the DSB mechanism of interest. The end point is to detect and quantify cells that have undergone the expected repair mechanism as determined by either by drug-selection (e.g. counting of G418-resitant clones for reporters) or flow cytometric analysis (e.g reporters expressing fluorescent proteins such as GFP). Development of such DSB reporter assays has had significant impacts in the DSB repair research. For example,

by exploiting these assays a number of new determinants of DSB repair have been identified. In this respect, an essential role for BRCA2 in promoting HR was described using the DR-GFP reporter assay (Howard et al., 2015, Moynahan et al., 2001).

Numerous I-SceI-cleavable DSB reporters based on the GFP expression have been developed. Figures below illustrate examples of some of these widely used GFP reporters.

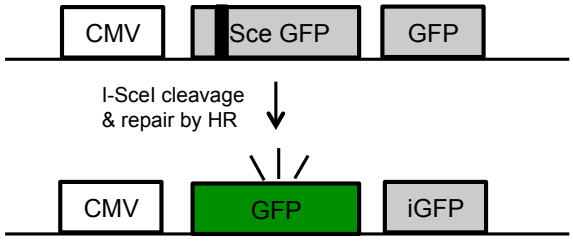


Figure 1.5. HR assay (DR-GFP cassette)

This is a well-characterised GFP-based reporter in which the proportion of GFP positive cells following a DSB is associated with the capacity of cells to undergo HR. This reporter assay contains an I-SceI recognition sequence, which upon I-SceI expression is cleaved to generate a DSB. DSB repair by HR using a direct repeat within the reporter cassette as an HR repair template results in an intact GFP gene (Pierce et al., 1999).

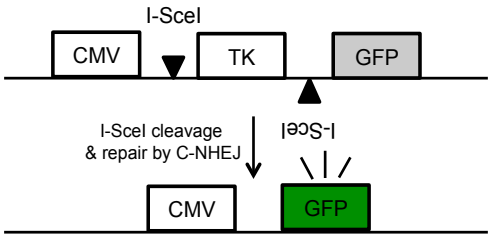


Figure 1.6. NHEJ-GFP assay (IRES-TK-EGFP cassette)

This reporter assay measures the C-NHEJ events following an I-SceI induction.
The NHEJ reporter is designed such that when the two I-SceI sites that are separated by a thymidine kinase gene (TK) are joined by NHEJ, the flanking promoter/enhancer and GFP open reading frame are brought together resulting in GFP expression (Ogiwara et al., 2011).

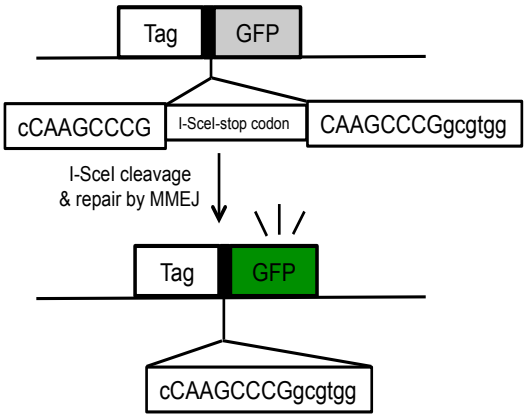


Figure 1.7. MMEJ assay (EJ2-GFP cassette)

In this MMEJ reporter assay (EJ2- GFP) GFP expression is disrupted by a stop codon next to an I-SceI site. An 8nt microhomology region flanks this complex. I-SceI-induced DSB in this context and repair by annealing the DSB ends towards the microhomology regions generates GFP positive cells, indicative of MMEJ levels (Bennardo et al., 2008).

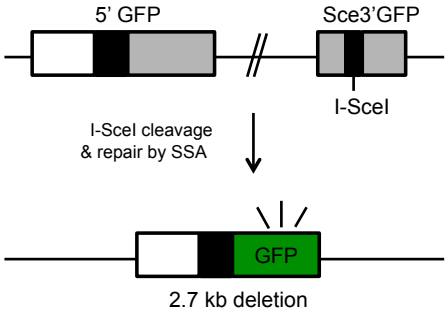


Figure 1.8. SSA-GFP assay (SA-GFP cassette)

The SSA-GFP reporter assay contains a 5' GFP fragment separated 2.7 kb from a 3'GFP fragment that contains an I-SceI recognition site. The two GFP fragments share 266 nt of homology that can bridge the I-SceI-induced DSB during SSA, thereby restoring a functional GFP + cassette (Bennardo et al., 2008).

1.3.3. HPRT as an endogenous reporter gene for assaying DSB repair

1.3.3.1 The *HPRT* gene

The *HPRT1* gene is located on the X chromosome of mammalian cells spanning 60 kb across 9 exons and encodes the enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT). HPRT plays a central role in the purine salvage pathways, where it mediates the recycling of hypoxanthine and guanine into the usable nucleotide pools. Mutations in the *HPRT1* gene cause Lesch-Nyhan syndrome exclusively in males. In this inherited disorder, that is observed in about one in 380,000 live births, cells are unable to recycle the purines, which is due to the deficiency of the HPRT enzyme. Thus patients manifest a spectrum of neurological symptoms, caused by the accumulation of abnormally high levels of uric acid (Seegmiller et al., 1967, Jinnah et al., 2000, Torres et al., 2014). Neurological manifestations include severe action dystonia (twisting and repetitive movements or abnormal fixed postures), cognitive and attention deficiency, and self-injurious behaviour (Torres et al., 2014). *HPRT1* is utilised as a model gene to examine DNA mutations in mammalian cell lines that are caused by a range of DNA mutagenic chemicals. It is an attractive model gene because, being X chromosomal, there is only one active copy per cell, making loss of function mutations easier to detect. Furthermore, drugs can be used to select for cells that have lost or gained a functional *HPRT* gene. Thus, cells with a non-functional or functional *HPRT* gene can be selected in medium containing 6-thioguanine (6-TG) or hypoxanthine, aminopterin and thymidine (HAT), respectively.

The mechanism of HAT selection is that aminopterin blocks de novo purine and pyrimidine biosynthesis, while hypoxanthine and thymidine provide cells with the substrates required to escape this blockade via the purine and pyrimidine salvage pathways, respectively. Cells grown in HAT medium are therefore dependent on these salvage pathways for DNA synthesis and survival and, if their *HPRT* gene is

mutated, they die of purine starvation.

Use of the purine analogue 6-thioguanine (6-TG) forms the basis for selection of cells with a non-functional *HPRT* gene. 6-TG is a purine analogue that gets converted to thioguanosine monophosphate analogue (TGMP) by HPRT, which is then further processed by kinases and reductase activity to make deoxy-TGTP which is then incorporated into DNA in place of dGTP during replication. Further DNA 6-TG undergoes *in situ* methylation (DNA 6-meTG) that preferably matches with thymine during replication. The resemblance of 6-meTG:T base pairs to replication errors, in turn stimulates MMR processing that triggers autophagic and apoptotic programmed cell death (Karran, 2006).Thus cells with a non-functional HPRT survive in the 6-TG-containing medium because they are unable to generate the cytotoxic nucleoside and nucleotide derivatives of 6-TG. In contrast, HPRT-positive cells can generate these poisonous derivatives and are killed (Doetschman et al., 1988, Skandalis and Glickman, 1990, Rothkamm et al., 2008).

1.3.3.2 Development of an I-SceI-cleavable HPRT assay (Exon assay)

In brief, an I-SceI recognition site was silently introduced to the exon 6 of the human *HPRT* gene without disrupting its function. This loss of function HPRT assay was previously validated as a genetic tool to measure DSB mis-repair events. However, this was yet to be fully characterised as a DSB reporter to identify the effects of changes in gene expression or drug treatments on mutagenesis. Importantly, during the course of this thesis, the advent of the CRISPR/Cas9 technology added further options for exploiting the gene as a reporter for DSB repair.

1.4 Clustered regulatory interspaced short palindromic repeat (CRISPR)/Cas-based RNA guided nucleases

1.4.1 CRISPR-Cas system; biology, and mechanisms of action

RNA-guided Cas9 nucleases derived from clustered regularly interspaced short palindromic repeats (CRISPR)/Cas systems have dramatically advanced the genome-editing field(Sander and Joung, 2014).

CRISPR/Cas complex constitutes the adaptive immunity in bacteria and archaea. This defence mechanism delivers immunity through sequence-specific recognition and degradation of exogenous nucleic acid (pathogens). Upon encountering foreign DNA, the CRISPR/Cas system in bacteria and archaea incorporates genetic material from the invading pathogens into CRISPR, which is then transcribed into "spacer" sequences. Each CRISPR locus encodes 'spacers' sequences separated by repeat sequences. Transcription of this locus generates a pre-crRNA, which in turn is processed to produce CRISPR RNAs (crRNAs) comprising of spacer-repeat fragments. This fragment then guides the CRISPR/Cas complexes to cleave the DNA sequences complementary to the spacer. This defence system serves as a genetic memory that assists cells to identify and demolish the previously encountered invader pathogens in future infections (Horvath and Barrangou, 2010).

The bacterial and archaeal CRISPR/Cas systems are categorised into three types, the simplest of which is type II. In this type, invading DNA is cleaved into small pieces and incorporated into the CRISPR locus in amongst a series of short repeat sequences. Transcripts of the loci are then processed to generate crRNAs that are responsible for guiding effector endonucleases, which target invading DNA to cleave the complementary sequences. Interestingly Cas9 (CRISPR associated protein 9) has been demonstrated to play a central role in type II CRISPR system. In this

system, Cas9 contributes to crRNA processing as well as cleaving the target DNA. Site specific DNA targeting and cleavage can only be achieved, if Cas9 forms a complex with the crRNA and the trans-activating crRNA (tracrRNA). TracrRNA is essential for a process called crRNA maturation, during which the tracrRNAs hybridises to repeat regions of the pre-crRNA. This is then followed by the cleavage of the crRNA-tracrRNA hybrid by endogenous RNase III, which in turn generates mature crRNAs that remain associated with both the tracrRNA and Cas9. Finally, the mature complex identifies a target DNA sequence and cleaves both strands (Carter and Wiedenheft, 2015a, Carter and Wiedenheft, 2015b).

The presence of a short conserved DNA sequence (2-6 base pair) known as protospacer-associated motif (PAM), immediately following the DNA sequence targeted by the Cas9 nuclease is an absolute requirement for the endonuclease activity of Cas9. The presence of PAM only in the invading DNA and not in the CRISPR locus ensures that the CRISPR locus itself is not being targeted by Cas9 nuclease (Mojica et al., 2009, Shah et al., 2013).

The simplicity of the type II CRISPR/Cas system, with only three required elements (Cas9, crRNA, and trRNA) makes it an attractive tool amenable to adaptation for genome editing.

1.4.2. CRISPR/Cas9 adaptation to study mis-repair of DSBs

Inspired by the bacterial and archeal defence mechanisms, scientists have transformed the CRISPR/Cas system into a widely used genome-editing tool, called “CRISPR/Cas9”. This genome editing technology possesses two key components; a single guide RNA (sgRNA) that is a substitute for the bacterial CRISPR RNA and guides the system towards the complementary DNA sequence, and a Cas9

endonuclease that generates blunt end DSBs at the locus targeted by sgRNA. The inaccurate repair of these blunt end DSBs via NHEJ machinery can then result in small indels at the break junctions that disrupts the gene (knockout)(Sander and Joung, 2014, Mali et al., 2013).

CRISPR/Cas9 technology not only has advanced the genome-editing field but has also been applied to study targeted chromosomal DSBs and thus has facilitated research in DSB repair field. The ease with which CRISPRs can be developed to cleave almost any target locus of interest, and delivered to a wide range of cell lines without requiring the cloning steps of other endonucleases, has made this technology an appealing tool in the DSB repair research field.

1.5 Aims

In this thesis, I have exploited the human endogenous *HPRT1* gene as a selectable marker in the context of a DSB genetic mutation assay to address the following questions:

1.5.1. Exploring the function of SETD2 in DNA DSB repair

1.5.2. Investigating the impact of interrupting HR repair pathway on the DSB repair outcome in human cells

1.5.3. Comparing the DNA mutational profiles following I-SceI- vs CRISPR/Cas9-induced DSBs

1.5.4. Developing a novel I-SceI genetic assay to analyse the mechanisms of DSB mis-repair

Chapter 2

Materials and Methods

2.1 Tissue culture

2.1.1 Resuscitation of frozen cell lines

Cells stored in liquid nitrogen were rapidly thawed at 37° C for 2 minutes. The cells were washed once with 10ml complete media to remove DMSO from the freezing media and pelleted by centrifugation at 1300 rpm for 3 minutes. The supernatant was aspirated and the cell pellet was resuspended in 5ml complete media and transferred to a T25 tissue culture flask.

2.1.2 Maintenance of cell lines

HT1080 (human fibrosarcoma, HPRT assay) cells were grown in high glucose media (DMEM) supplemented with 1% Sodium Pyruvate (100mM), 1% L-Glutamine (200 mM), 25% MEM (Non-essential amino acids) and HAT medium before the I-SceI transfection, and 6-TG medium 5 days after the I-SceI transfections.

HAT to select for HPRT gain of function

HAT selection was carried out by adding 1X HAT supplement (Invitrogen) directly to the DMEM medium for 10 days.

6-TG to select for HPRT loss of function

6-TG selective medium was made by making up 1000X 6-TG of 15 mg/ml by dissolving the 6-TG powder (sigma) in 1N NaOH and ddH2O. 1000X 6-TG was added to DMEM.

The MMEJ reporter cells (U2OS with integrated EJ2-GFP reporter) were grown in high glucose DMEM supplemented with L-glutamine, 50 ml fetal bovine serum, and 6 ml Pen/Strep solution (10,000 U/ml penicillin, 10,000 mg/ml streptomycin), including 8 mg/ml plasmocin and 2 mg/ml puromycin. All other cell lines used in this thesis, were maintained in complete media consisting of Dulbeccos modified minimal

essential media (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Invitrogen), 100U penicillin and 0.1mg/ml streptomycin (Sigma). All cells were grown at 37 ° C in a humidified incubator with 5% carbon dioxide. Cells were sub-cultured every 3 days, where cells were washed with phosphate buffered saline (PBS) and then incubated with Trypsin-EDTA for 5 minutes at 37 °C. Trypsinised cells were re-suspended in complete media and re-plated at 1:5 dilution. All cell lines were routinely tested for mycoplasma and found to be negative.

2.1.3 Cryopreservation of cells

Log-phase growing cells were trypsinised as described above. $1X10^6$ cells were spun down at 1300 rpm for 3 minutes and re-suspended in 1ml freezing media (FBS containing 10% (v/v) tissue culture ready DMSO (Sigma)). The suspension was transferred to a cryotube and frozen at -80°C overnight before being transferred into liquid nitrogen for long-term storage.

2.2 Cell lines

The cell lines used in this study are described below:

The U2OS cell line was derived from human osteosarcoma and obtained from ATCC (ATCC number HTB-96).

The DR-GFP cell line was derived from U2OS by stably integrating a DR-GFP (HR) reporter construct. It was obtained from the lab of Professor Thomas Helleday.

The H1299-NHEJ cell line was derived from H1299 by stably integrating a NHEJ reporter construct. It was obtained from Dr. Atsushi Shibata and Dr Hideaki Ogiwara.

The EJ2-GFP (MMEJ reporter) cell line was derived from U2OS by stably integrating an MMEJ reporter construct. It was obtained from Dr. Jeremy Stark.

2.3 Generation of stable cell lines

2.3.1 Generation of stable cell lines by electroporation

Linearised plasmids were introduced to HT1080 cells by electroporation as suggested by 'Molecular Cloning: A Laboratory Manual'. Briefly 10^6 log-phase cells were resuspended in 400µl PBS and mixed with 10µg linearised plasmids in a 4mm cuvette. Electroporation was performed using the Gene Pulser Xcell System (Bio- Rad) set up with the following parameters: 800V, infinite resistance and 1.3 ms time constant. After electroporation, cells were left to recover in fresh media for 48 hours, then split at 1/5 dilution in media with appropriate selection. About 20 days after selection, single-cell colonies were isolated and expanded in individual flasks with selection media.

Generation of the I-SceI-cleavable HPRT intron assay

In order to sub clone the I-SceI site into intron 5 of the *HPRT* gene, the PJB2 plasmid (8753 bp) was used, which contains intron 5, exon 6, and intron 6 of human *HPRT* gene (Provided by Dr. Andy Porter, Imperial College London) (Fig 6.1B). The I-SceI oligos (from sigma Aldrich) were annealed in accordance with the manufacturer's instructions using 30 ug of each forward (TCGATAGGGATAACAGGGTAAT) and reverse (TCGAATTACCCTGTTATCCCTA) oligonucleotides, 3 ul of 10 X annealing

buffer and 7 ul of ddH2O in total reaction volume of 30 ul. The annealing reaction was performed at 95º C for 10 min, followed by cooling down at room temperature (RT). Restriction enzymes were supplied by New England Bio labs (NEB) and reactions were performed as per their instructions. Generally, 5 units of enzyme were used for an incubation of 1h at the recommended temperature. Ligations were performed using a 1:3 molar ratio of vector to insert in a 20µL reaction containing 400 units T4 DNA ligase (NEB) and 1x T4 ligase reaction buffer. Ligations were performed at RT for 4h. Competent *Escherichia coli* (*E. coli*) strains (NEB) were used for standard cloning techniques. Transformations were carried out according to the manufacturer's instructions. Bacterial cultures were grown in LB medium or on LB agar plates supplemented with ampicillin (amp) (75µg/mL). Cultures were frozen at -80°C in 15% glycerol volume per volume (v/v) for long-term storage. Plasmid DNA was isolated from bacterial cells using mini preparation kits (Qiagen) as directed by the manufacturers. For genomic DNA, cells were washed in PBS and $2x10^6$ cells were pelleted. Genomic DNA was extracted using the Qiagen Genomic DNA Purification Kit. A NanoDrop, ND-1000 spectrophotometer (Thermo Scientific) was used. DNA concentrations were determined by measuring the optical density (OD) at 260 nm and converted to µg/mL using: 1 OD unit = 50µg/mL DNA. DNA was separated on 1% agarose TAE gels. DNA was stained using 500 ng/ml Ethidium bromide (EtBr). NEB 1 kb ladder and low molecular weight ladders were used, and gels were run at 70 V for 1 hour. The DNA was then observed using a UV transilluminator and an image captured using Quantity-One (Bio Rad).

The I-SceI modified construct (PJB-I-Sce-XhoI) was amplified using Large-Scale Plasmid DNA Purification Kit (Qiagen) according to manufacturer`s instructions. 50 ug of the constructs were linearised using BamHI restriction enzyme (NEB). The digested plasmids were ethanol-precipitated, spanned, washed in 70% ethanol and

air-dried in hood for 10 min under sterile conditions. The vector was re-suspended in 50 ul sterile buffer TE with the final concentration of 1 mg/ml.

2.4 Bacterial transformation and plasmid purification

2.4.1 Bacterial transformation

Chemically competent E. coli cells (produced in-house by Carol Walker) were used for plasmid amplification. 1–5µl of plasmid DNA (1pg–100ng) was added to 50µl of the bacteria cells. The mixture was incubated on ice for 30 minutes, heat-shocked at 42°C for 30 seconds, left on ice for 5 minutes and allowed to recover in 1ml S.O.C (super optimal broth) for 1 hour at 37°C with shaking. 50–100µl of the mixture was plated onto a selection plate and incubated overnight at 37°C. Depending on the plasmid backbone, selection of the plasmid-carrying bacteria clones was achieved by adding either 100µg/ml ampicillin or 50µg/ml kanamycin to the LB agar plates.

2.4.2 Plasmid purification

Plasmid purification from bacterial cultures was performed using QIAprep spin miniprep or maxiprep kits (Qiagen) according to the manufacturer's instructions. The QIAprep columns contain a silica membrane that binds DNA in the presence of a high concentration of chaotropic salt, and allows DNA elution in low-salt buffer or water. In summary, following overnight growth in 5ml of selective LB in a 50ml tube, bacterial cells were harvested by centrifugation at 4000g for 20 minutes. The pellet was re-suspended in 250µl P1 buffer (containing RNase A) and cells were lysed by addition of 250µl P2 buffer and inverting. The lysis buffer was neutralised by adding 350µl of P3 buffer. Cell debris and genomic DNA were precipitated from the solution by centrifugation at 13,000rpm for 10 minutes at room temperature. The supernatant (containing plasmid DNA) was pipetted into a QIAprep spin column and centrifuged for 1 minute at 13,000rpm, and the flow-through was discarded. The

plasmid DNA bound to the silica membrane was washed with 0.75ml PE buffer to remove impurities. The plasmid DNA was eluted by incubating 50µl H2 O in the column for 1 minute before centrifugation into a collection tube.

2.5 Gene silencing

2.5.1 siRNA transfection

All siRNAs (10nM final concentration) were delivered to the cells by Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. Cells were incubated with siRNA-lipofectamine mix in suspension (reverse transfection). The medium was replaced 24 hours after transfection. Cells were analysed or further treated 48 hours post-transfection. Western blotting was performed to determine knockdown efficiency. The sequences of the siRNAs are listed below.

non-targeting (NT) proprietary sequence of supplier (Qiagen),

siBRCA2 (Thermo Scientific):

GGGAAACACUCAGAUUAAAUU, UUUAAUCUGAGUGUUUCCCUU;

siBRCA1(Dharmacon):

AAUGCCAAAGUAGCUAAUGUAUUUU,AAUACAUUAGCUACUUUGGCAUUUU;

siCtIP (Dharmacon):

GAGGUUAUAUUAAGGAAGA,GGAGCUACCUCUAGUAUCA,GAACAGAAUAGGACU
GAGU, GCACGUUGCCCAAAGAUUC;

siMRE11(Dharmacon):GGAGGUACGUCGUUUCAGA,
GGAAAUGAUACGUUUGUAA,CGAAAUGU

CACUACUAAGA, GAAAGGCUCUAUCGAAUGU;

siRPA (Dharmacon): AACUGGUUGACGAAAGUGGUGUU,

CACCACUUUCGUCAACCAGUUUU;

siPOLQ(Ambion,Lifetechnologies):CCGCUUUUGGAGUCAGUAATT,UUACUGACUCC
AAAAGCGT

siSETD2 (Dharmacon): UAAAGGAGGUAUAUCGAAU

2.6 Protein detection by Western blotting

2.6.1 Detection of proteins by Enhanced Chemiluminescence

Cells were lysed in the lysis buffer (50 mM Tris HCl pH7.5, 150 mM NaCl,1% Triton X-100, supplemented with protease and phosphatase inhibitors) for 10 minutes on ice. Lysate was centrifuged for 10 minutes at 13,000 rpm to pellet cellular debris. The supernatant was transferred to fresh tubes and protein concentrations were measured by the Bradford assay. 25µg of protein was mixed with the NuPAGE LDS loading buffer (Invitrogen) and the reducing agent Dithiothreitol (DTT), and denatured by boiling at 95 C for 5 minutes. Samples and molecular weight ladder (Invitrogen) were loaded onto pre-cast NuPAGE gels (Invitrogen) and electrophoresis was performed according to the manufacturer's instructions. Proteins were transferred onto a PVDF membrane (0.45µm pore size) in a transfer unit (Invitrogen) at 30V for 2 hours according to manufacturer's instructions. The membrane was blocked in blocking buffer (PBS containing 5% skimmed milk and 0.2% Tween) for 1 hour at room temperature before incubating with the primary antibody diluted in the blocking buffer at 4 C overnight. After washing three times with PBS-T (PBS containing 0.2% Tween), the membrane was then incubated with a HRP-conjugated secondary antibody diluted 1:5000 in the blocking buffer for 1 hour at room temperature. After washing three times with PBS-T, the proteins on the membrane were visualised with the ECL

chemiluminescence substrates (Thermo Scientific) according to the manufacturer's instructions.

2.7 Antibodies

The antibodies used in this thesis are summarised below:

SETD2 (Abcam, 31358),

BRCA2 (Santa Cruz, sc-28235);

BRCA1 (Calbiochem,OP92);

RAD51 (Santa Cruz, sc-8349);

CtIP (GeneTex 19E8);

MRE11 (Abcam, ab33125);

RPA32 (Abcam, 2175);

Tubulin (Sigma, T5168).

All secondary antibodies were purchased from Invitrogen.

2.8 Inhibitors

All inhibitors were dissolved in Dimethyl sulfoxide (DMSO) and stored at -80°C, and used according to the manufacturer's instructions. The PARP-1 inhibitor, Olaparib (AZD 2281) was used at 5 μM final concentration. The DNA-PKcs inhibitor, NU7441 (Axon Medchem) was used at 5, and 15 μM final concentrations.

2.9 Reporter assay for determining DNA repair pathways

The reporter assays were used to determine the impact of a treatment (inhibitor or siRNA) on the efficiency of repair by either the HR or the NHEJ pathways. I-SceI (intron-encoded endonuclease 1 from *Saccharomyces cerevisiae*) is a rare-cutting endonuclease that has no predicted recognition site in the mammalian genome. The HR and NHEJ reporters each contain an I-SceI recognition site that generates DNA-double-stranded breaks (DSBs) upon I-SceI expression. Successful repair by the pathway in question leads to the expression of the green fluorescent protein (GFP), which can be detected by FACS. The reporter assay was performed in 96-well format and the procedure is described below:

$2X10^4$ reporter cells were seeded in 100µl antibiotics-free medium and reverse-transfected with a siRNA-RNAiMAX mixture according to the manufacturer's instructions. 24 hours after transfection, the medium was refreshed 100µl fresh antibiotic-free media. 48 hours after siRNA transfection,the I-SceI plasmid was delivered to the cells by Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. 18 hours after plasmid transfection, the medium was refreshed with 100µl fresh antibiotic-free medium. Cells were allowed to repair the I-SceI induced DSBs for 48 hours before analysing the percentage of GFP- positive population by FACS.

2.10 qRT-PCR (Quantitative reverse transcription polymerase chain reaction)

RNA was extracted using an RNeasy mini kit (Qiagen). A SuperScript® VILOTM (Invitrogen/Life Technologies) cDNA synthesis kit was used to reverse transcribe cDNA from total RNA according to manufacturer's instructions. Quantitative PCR was performed using 7500 Fast Real-Time PCR detection system (Applied Biosystems). Reactions (25µl each) were prepared in triplicate in a 96-well reaction plate. Each reaction contained 20ng cDNA, 200nM of each primer, 10µl water and

12.5µl Absolute Blue QPCR SYBR low ROX Mix (Thermo Scientific). DNA levels were normalised to the GAPDH calculated using a 2-ΔΔCt method. QPCR settings were as follows: Initialisation at 95^{0}C for 15 min, denaturation at 95^{0}C for 15 seconds, annealing at 60^{0}C for 30 seconds, and extension at 72^{0}C for 30 seconds and repeat for 40 cycles. Primers used for the qRT-PCR are listed below:

Primers used for the qRT-PCR are listed below:

For spliced HPRT mRNA:

RT-PCR-F: CCCTGGCGTCGTGATTAGT

RT-PCR-R: TTCATCACATCTCGAGCAAGAC

For un-spliced HPRT mRNA:

qPCR-F:TGGTGAGAATTACTGTGCTGAA

qPCR-R: TGCGACCTTGACCATCTTTG

2.11 qPCR for HPRT I-SceI cleavage efficiency

Genomic DNA was isolated using the Qiagen kit at 0h and 24h after the I-SceI transfection of the HT1080 (clone 5.2.1) cells. Quantitative PCR was performed using 7500 Fast Real-Time PCR detection system (Applied Biosystems). Reactions (25µl each) were prepared in triplicate in a 96-well reaction plate. Each reaction contained 20ng genomic DNA, 200nM of each primer, 10µl water and 12.5µl Absolute Blue QPCR SYBR low ROX Mix (Thermo Scientific). DNA levels were normalised to the GAPDH calculated using a 2-ΔΔCt method. QPCR settings were as follows: Initialization at 95^{0}C for 15 min, denaturation at 95^{0}C for 15 seconds, annealing at 60^{0}C for 30 seconds, and extension at 72^{0}C for 30 seconds and repeat for 40 cycles. Primers used for qPCR are listed below:

qPCR-F:TGGTGAGAATTACTGTGCTGAA

qPCR-R: TGCGACCTTGACCATCTTTG

2.12 HPRT-based I-SceI-cleavable reporter assay

The assay was based on human fibrosarcoma (HT1080) cells with a functional but I-SceI-cleavable *HPRT* gene (clone 5.2.1), as described previously (Gravells et al., 2015). Use of this assay in conjunction with siRNA knockdowns has been previously described (Pfister et al., 2014). Briefly, an I-SceI cut site was targeted into exon 6 of the endogenous human *HPRT* gene, without disrupting the functionality of the gene. These cells were seeded in 6-well plates and were immediately transfected with siRNAs using RNAiMAX from Invitrogen according to the manufacturer's instructions. The following day, medium was replaced with fresh DMEM. 48 hours after siRNA transfections, cells were lipofected with I-SceI plasmid, which was performed in a ratio of 6µg DNA: 18µl Fugene in a total of 300µl transfection reaction to transfect each well of a 6-well plate using the Fugene transfection reagent (Promega). 24 hours later, medium was replaced with fresh DMEM. 5 days after the I-SceI transfection the cells were seeded in different densities within 100 mm plates (i.e.10^5, $5x10^4$, 10^4, and 10^3 per plate). Cells used to determine the mutation frequencies were then exposed to 6-TG selection, while the cells used for determining plating efficiencies were kept in non-selective media. 6-TG medium was refreshed every 2-3 days. Cells were incubated for 10-12 days to form colonies, which were stained and counted for the purpose for calculating the mutation frequencies. Further, a total of 30 clones from each genetic background were isolated and grown in 48-well plates. These were subsequently expanded into 24-well plates. At full confluency cell for each clone were then trypsinised and collected for extraction of genomic DNA using a Qiagen Flexi gene DNA kit according to the manufacturer's instructions. Genomic

DNA was then quantified and used for PCR amplification across the I-SceI break site and sequencing. The PCR products were sequenced (Sanger Sequencing) using the primers flanking the I-SceI site. High fidelity Phusion master mix (x2) (New England BioLabs) was used for PCR amplification across the I-SceI site using forward (GATCCTGCACCTACAAAATCC) and reverse (GCATAGGTAAGGTGAGGAGGTG) primers (Sigma). PCR settings were as follows:

98 °C for 30 s, 30 cycles of 98°C for 10 s, 67°C for 1 min, and 72°C for 1min.

Products were visualised on 2 % agarose gel with 0.5 µg/ml ethidium bromide.

Mutation frequency was determined as below:

$$\frac{\text{N°6TG}}{\text{P. D} \times \text{P. Eff}}$$

N°6TG=Number of 6-TG resistant colonies (10-12 days following selection in 6-TG-containing DMEM)

P.D= Plating Density (Initial number of plated HT1080 cells)

P.Eff= Plating Efficiency (Number of colonies grown in DMEM with no selection)

2.13 Sequencing and Data analysis

PCR Samples were sent for sequencing at “Source BioScience” (https://www.sourcebioscience.com/)

Sequence alignment was conducted using “EMBOSS Needle_ Pairwise Sequence Alignment” (http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html)

2.14 Graphical display of results and statistical analysis

For all statistical analysis and graphical display, the program GraphPad Prism (www.graphpad.com) was used.

Unpaired two-tailed test was used for comparison of HPRT mutation frequencies in different genetic backgrounds.

To compare the frequency of mis-repair by MMEJ P values were calculated by statistical analysis “difference between proportions”.

2.15 Primers used to generate the HPRT intron assay

2.15.1: Oligonucleotides used to insert an I-SceI site into the XhoI site of pJB2 plasmid

Forward: tcgatagggataacagggtaat

Reverse: tcgaattaccctgttatcccta

2.15.2: Primers used to amplify from the I-SceI site to 1Kb downstream

Forward: gagttgttgagagatgaaacg

Reverse: tcgatagggataacagggta

2.15.3:Primers to amplify genomic DNA (XhoI-I-SceI)

Forward: atgaggcccagcaagctatg

Reverse: cataccttgcgaccttgacc

2.16 **CRISPR-Cas9 assay**

CRISPR-Cas9 constructs targeting the *HPRT* exon 6 locus (with or without the I-SceI site) were generated at the laboratory of Dr.Andy Porter. HT1080 cell line containing an I-SceI site (HPRT exon assay) and wild type HT1080 cells were transfected with CRISPR plasmids (gRNA2 or gRNA4) and gRNA1 respectively. In each transfection a Cas9-expressing plasmid was also delivered into cells. All transfections were performed using Fugene 6 (Promega). The cells were then allowed to repair for 5 days before seeding and selecting for HPRT-negative cells using 6-TG. The mutation frequencies were calculated as described in section 2.12. To determine the mutation sequence, individual single-cell colonies were isolated, and genomic DNA spanning the DSB junction was PCR-amplified using the primers flanking the I-SceI site. Sequence alignments were carried out as described in section 2.13.

Chapter 3

SETD2 promotes genome stability by suppressing break-induced mutations

This chapter presents the characterisation of a new endogenous DSB mutation assay based on the human *HPRT* locus. Using this assay, a novel role for SETD2, a histone methyltransferase and tumour suppressor, in promoting genome stability in response to a DSB is assigned.

3.1 Characterisation of an I-SceI-cleavable HPRT-based exon assay

To study the impact of DSBs on genome stability, while preventing complications arising from the existing exogenous reporter assays, an endogenous DSB mutation assay based on human *HPRT* gene (HPRT exon assay) was developed at the laboratory of Dr.Andy Porter at Imperial College London. The full description of the assay is described in the "materials and methods" section. Briefly, an 18 bp I-SceI recognition sequence, an intron-encoded endonuclease1 from *Saccharomyces cerevisiae* with no predicted recognition site in the mammalian genome, was integrated into the endogenous exon 6 of *HPRT* gene in HT1080 (fibrosarcoma) cells. The I-SceI site maintains the HPRT reading frame and makes only a single amino acid change without disrupting the function of HPRT. Following I-SceI induction, inaccurate DSB repair generates HPRT-negative clones, which can be selected for using 6-thioguanine (6-TG), and indels can be detected by PCR amplification using primers flanking the break site. The HPRT assay is a DSB mutation assay with the capacity to detect a wide range of mutations arising from various DSB mis-repair mechanisms (Fig 3.1A,C)(Gravells et al., 2015).

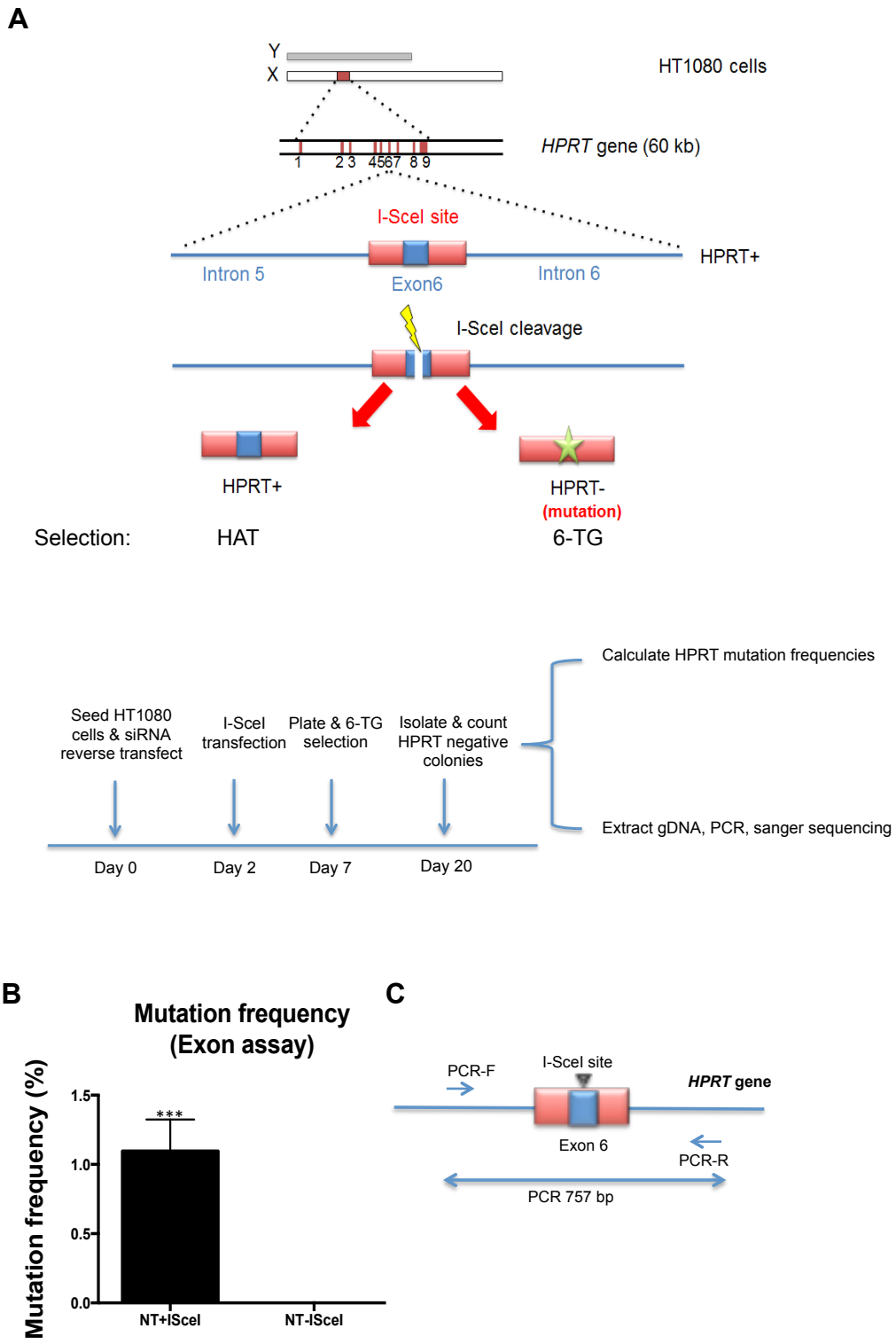


Figure 3.1. Mutation frequency significantly increases following DSB induction

A. Schematic map of the HPRT Exon assay , where an I-SceI recognition site is cloned into the exon 6 of HPRT gene in human HT1080 cells located on X chromosome. I-SceI induction in this context has the capacity to generate cells with functional or non-functional HPRT gene, which then are selectable using HAT or 6-TG medium, respectively. B. HPRT mutation frequencies in the presence or absence of an ISceI-induced break (P=0.0002). C. Schematic map showing the PCR primers utilised to amplify across the ISceI site in the Exon assay.

I-SceI induced DSBs resulted in a significant increase ($1X10^{-2}$) in the mutation frequencies compared to the spontaneous levels (10^{-6}) (P=0.0002) thus representing a ~10,000 fold increase in HPRT loss following I-SceI–induced DSBs (Fig 3.1B). To characterise the expected indels, the exon 6 regions of 40 individual HPRT-negative clones were amplified and sequenced using the primers across the I-SceI site (Fig 3.1C). The majority (35/40) had small or medium-sized deletions involving <10 nt. The rest involved larger deletions (4/40), or insertion (1/40). These results support the use of the I-SceI-cleavable HPRT-based exon assay to detect DSB-inducible indels (Fig 3.2)(Gravells et al., 2015). Further, these findings highlight the significance of a single DSB in threatening genome stability.

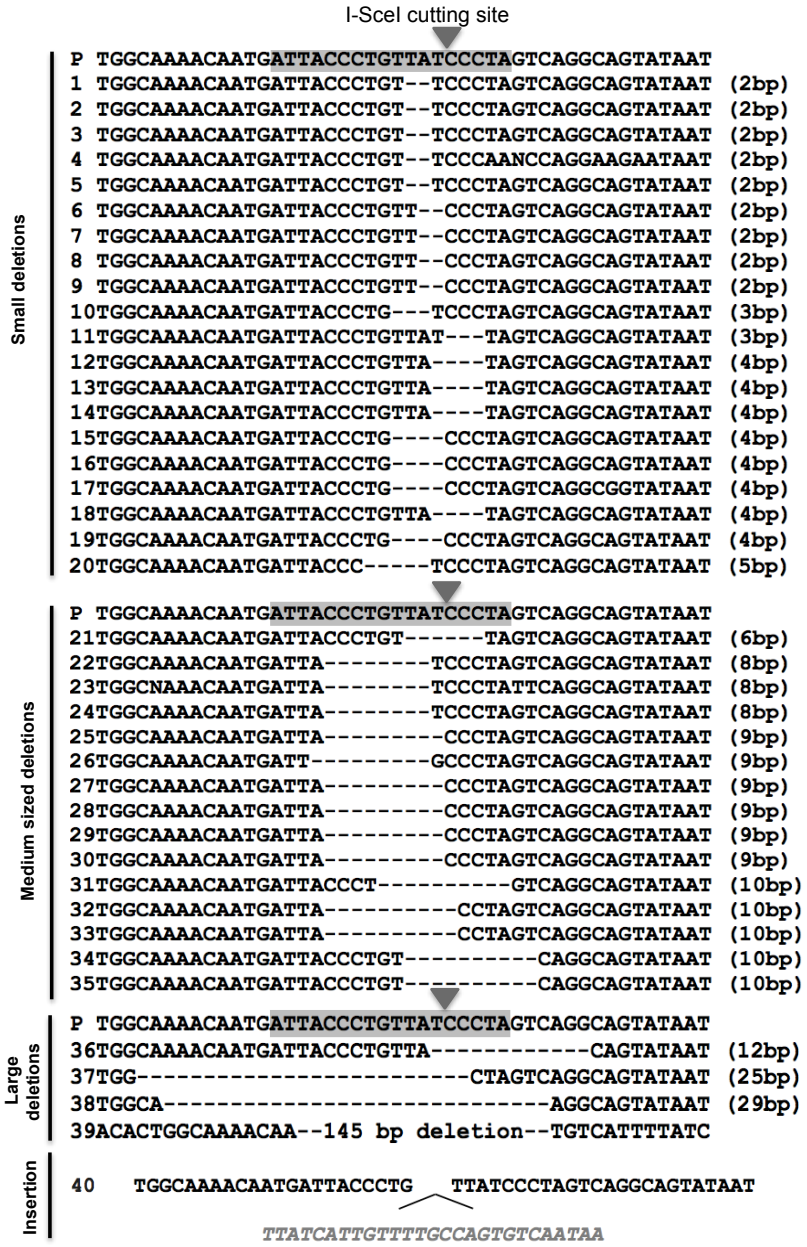


Figure 3.2. Sequence alignments of HPRT mutants in the exon assay

3.2. A common DSB-induced mutation signature for SETD2 and RAD51

Tumour suppressors can suppress carcinogenesis by preventing genome instability, a hallmark of cancer. To investigate the impact of loss of SETD2, a tumour suppressor, on genome stability following DSB induction, I used the HPRT mutation assay. Following small interfering RNA (siRNA) knockdown and I-SceI induction, the frequency of HPRT loss (mutation frequency) in SETD2-depleted cells was 1.65%, significantly higher ($p = 0.0014$) than that of non-targeting controls (1.1%). Also because recently a novel role for SETD2 in promoting HR was described in our laboratory (Pfister et al., 2014), I set out to investigate the impact of depleting other HR factors such as RAD51, a classical HR factor, on the HPRT assay break-induced mutation frequency. Thus I knocked down RAD51 and measured the mutation frequency following a break in the HPRT assay. Interestingly similar to SETD2, depletion of RAD51 also led to a significant increase in break-induced mutation frequencies (1.88%; $p=0.0022$), compared to non-targeting controls ($p = 0.0022$) (Fig 3.3C). To look for possible synergistic effects between RAD51 and SETD2, I then co-depleted SETD2 and RAD51 (siS+SiR), which also significantly increased the mutation frequency to 1.52% ($p = 0.037$). This was similar to the frequency observed following either SETD2 or RAD51 depletion (1.65% and 1.88% respectively) (Fig 3.3C), suggesting that SETD2 and RAD51 are in the same pathway. The efficiency of SETD2 and RAD51 knockdowns were confirmed by Western blot (Fig 3.3A,B). Also I confirmed that I-SceI protein expression and cleavage efficiency were the same in cells transfected with either control siRNA or SETD2 siRNA (Fig 3.3D). Further, I studied the mutation patterns across the break site of 30 individually isolated HPRT-negative clones from each background by PCR amplification and sequencing. Sequencing of HPRT-negative clones from cells treated with non-targeting control siRNA indicated the presence of small-sized deletions of 2-5 bp (Fig 3.4). Sequence analysis indicated that SETD2-depleted cells

exhibited a significant increase in average deletion length (11 bp, $p < 0.0001$) compared to control cells (Fig 3.4,5,8). RAD51-depleted cells exhibited a further significant increase in the average deletion length (22 bp, $p = 0.014$) (Fig 3.4,6,8). RAD51 and SETD2 co-depletion generated deletion lengths (an average of 8 bp) comparable to those observed following SETD2 knockdown, which in turn was significantly higher than the non-targeting control ($p = 0.026$) (Fig 3.7,8). These findings suggest a more upstream role for SETD2 in the HR pathway compared to RAD51. Indeed this was further confirmed in our laboratory using complementary experiments (Pfister et al., 2014).

Next, to investigate the mechanism of mis-repair in cells exhibiting deletions, I examined the sequence alignments obtained from SETD2 and RAD51 depleted cells. Interestingly the sequence analysis revealed the presence of microhomology regions at the break junctions (Fig 3.12). Importantly, the frequency of such microhomology regions across the I-SceI break site was significantly increased in SETD2- or RAD51-depleted cells ($p < 0.05$) (Fig 3.4,9). These regions are reminiscent of microhomology mediated end-joining (MMEJ) pathway, and thus this suggests that HR-deficient cells use MMEJ as an alternative repair mechanism, which gives rise to a mutation signature of sequence loss between regions of microhomology. These results support a role for SETD2 in tumour suppression by maintaining genome stability through preventing deleterious mutations arising in response to DSBs. The similar increase in the proportion of deletions likely to have arisen through MMEJ following SETD2 or RAD51 depletion suggests a similar break-induced mutation signature between SETD2- and RAD51-depleted cells, indicative of an early role for SETD2 in HR repair. Also the resemblance of the HPRT mutation profile in cells co-depleted for SETD2 and RAD51 to SETD2-depleted cells suggests a role for SETD2 in HR possibly upstream of RAD51(Pfister et al., 2014).

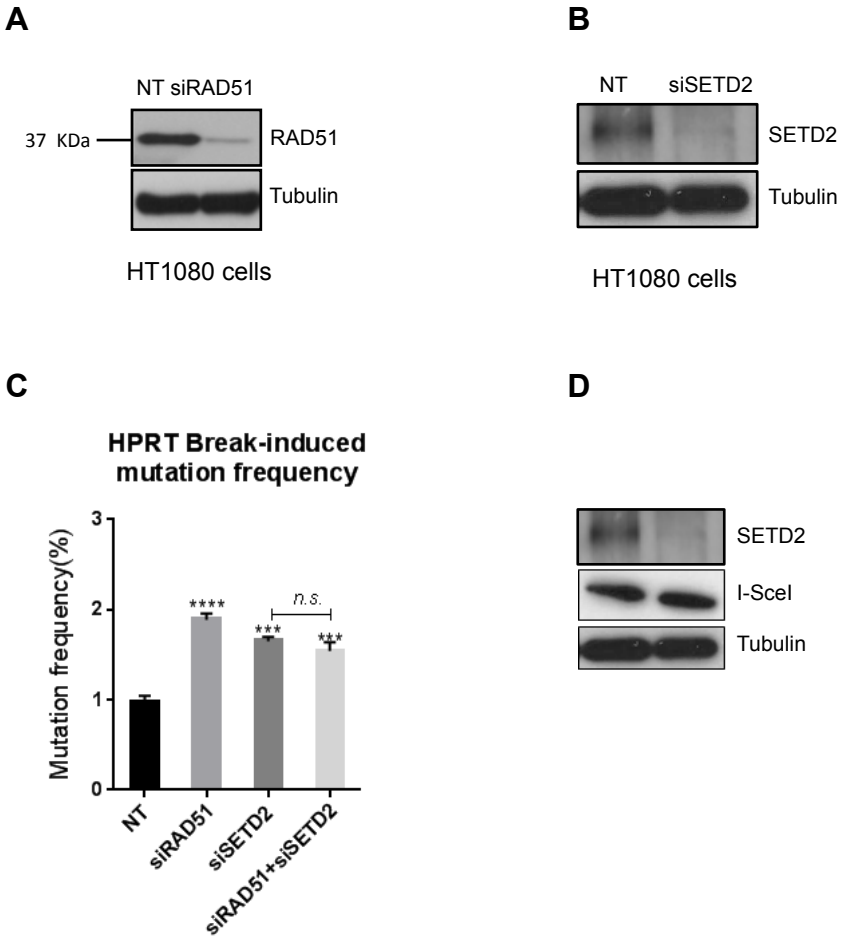


Figure 3.3. Depletion of SETD2 or RAD51 results in enhanced DSB-induced mutation frequencies. A. Western blot showing RAD51 knockdowns in HT1080 cells 48 hours after siRNA transfections. B. Western blot showing SETD2 knockdowns in HT1080 cells 48 hours after siRNA transfections. C.HPRT break-induced mutation frequencies in cells transfected with NT control siRNA, or siRNAs against RAD51, SETD2, or RAD51+SETD2. D. Western blot showing SETD2 knockdown and I-SceI expression in the HPRT assay.

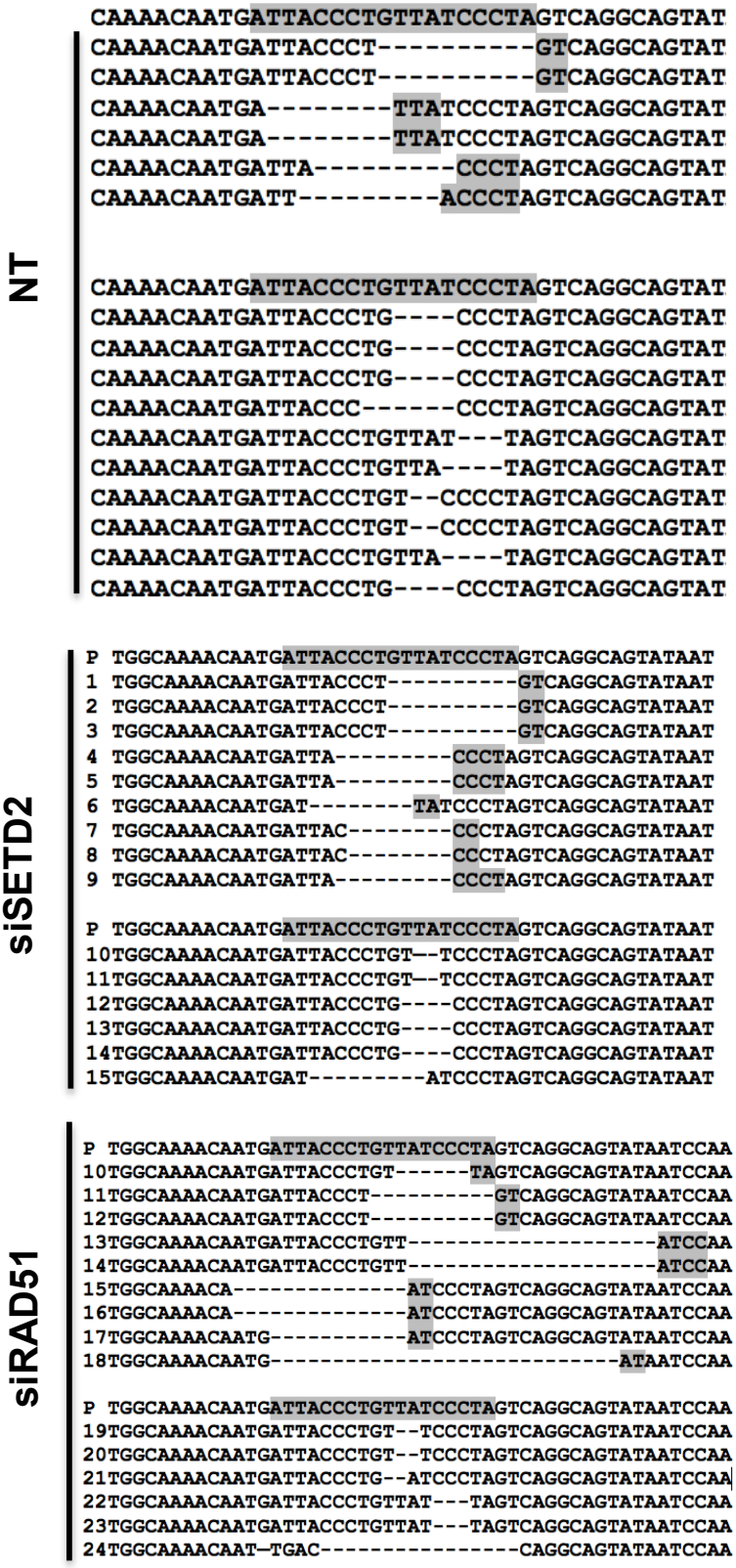


Figure 3.4. A common DSB-induced mutation profile for SETD2 and RAD51
Representative sequence alignments of the PCR products in NT control (NT), and RAD51-depleted (siRAD51) cells isolated from three independent experiments. I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted.

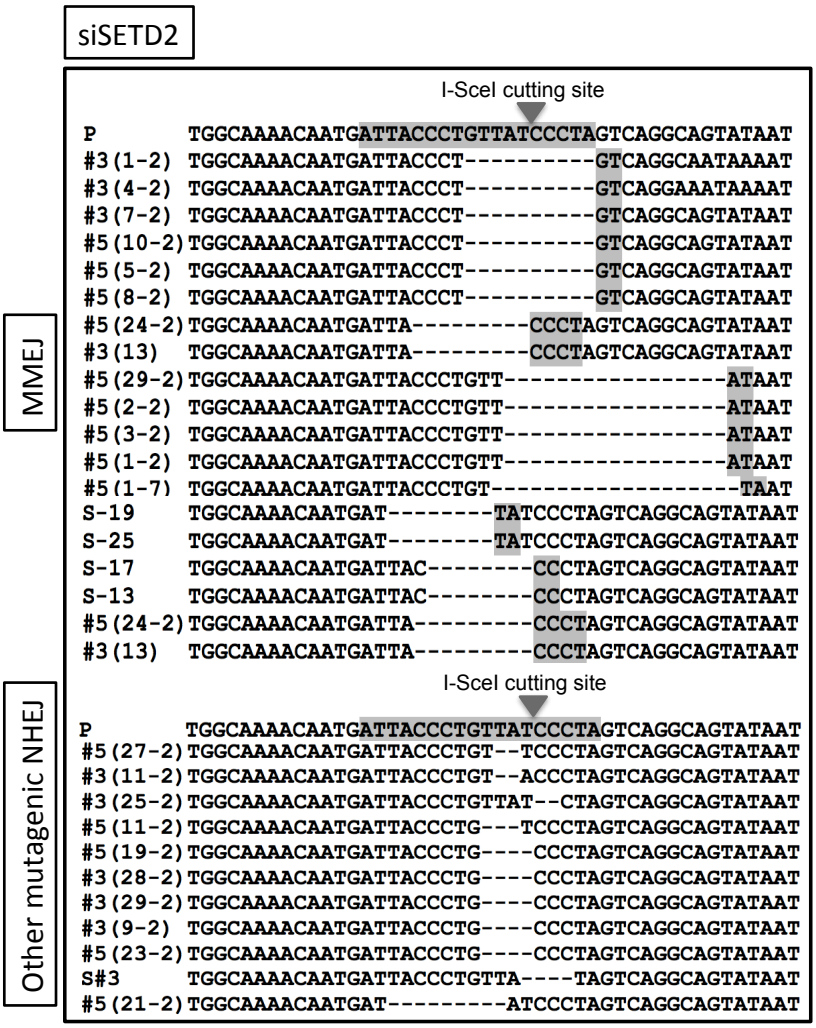


Figure 3.5. Sequence alignments of HPRT negative cells obtained from SETD2-depleted background.

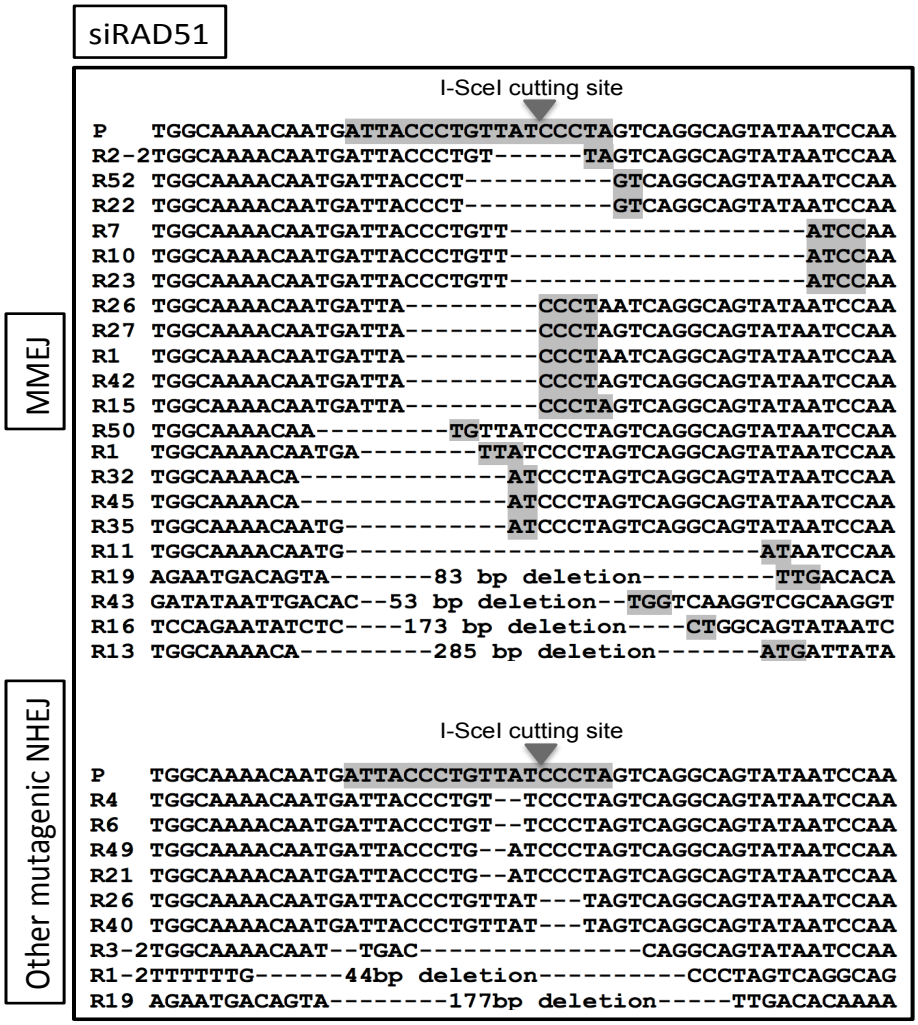


Figure 3.6. Sequence alignments of HPRT negative cells obtained from RAD51-depleted background.

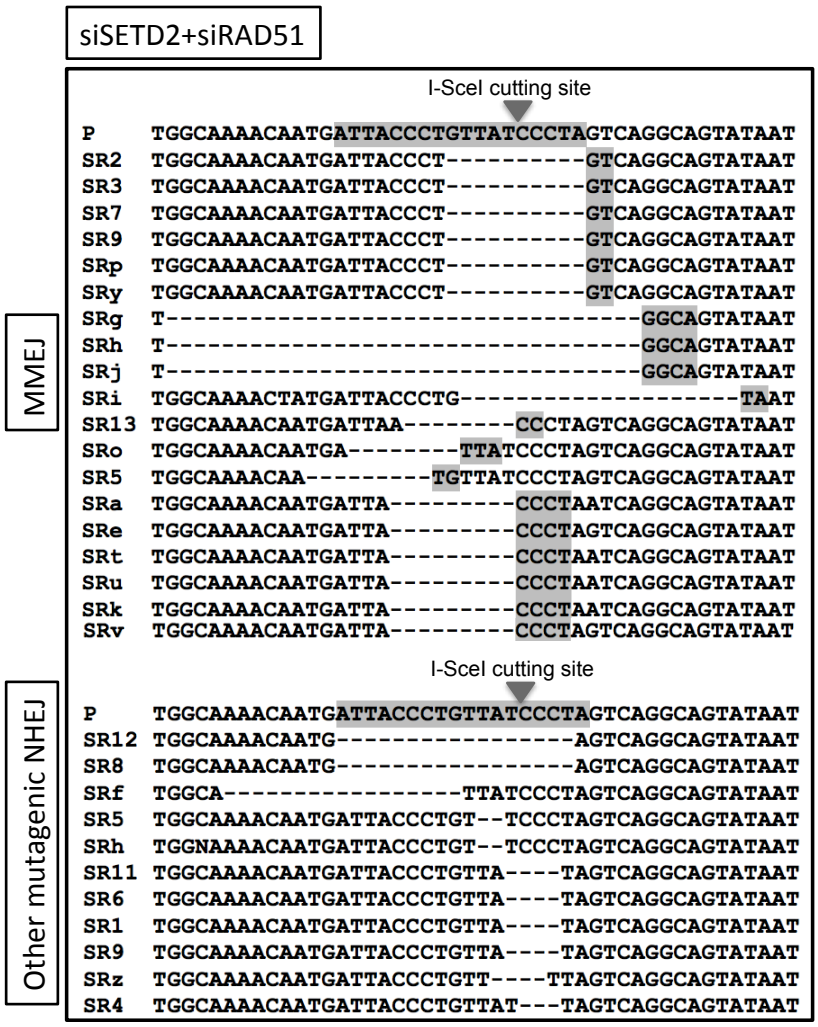


Figure 3.7. Sequence alignments of HPRT negative cells obtained from SETD2/RAD51 co-depleted background.

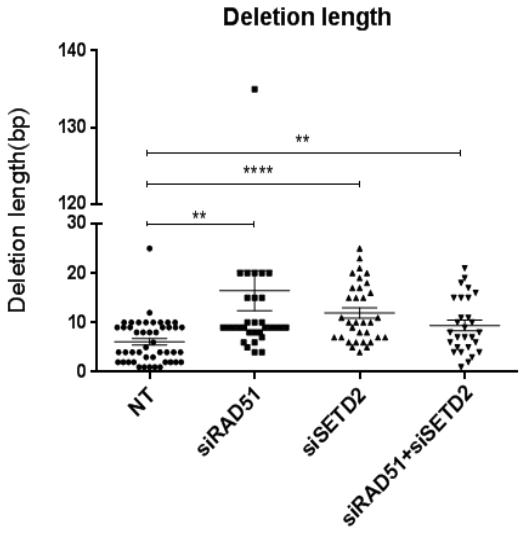


Figure 3.8. Diagram illustrating deletion lengths in cells treated with siRNAs against NT, RAD51, SETD2, or RAD51+SETD2. The solid line in siRAD51 is due to the fact that many clones show a common deletion length of around 10 bp.

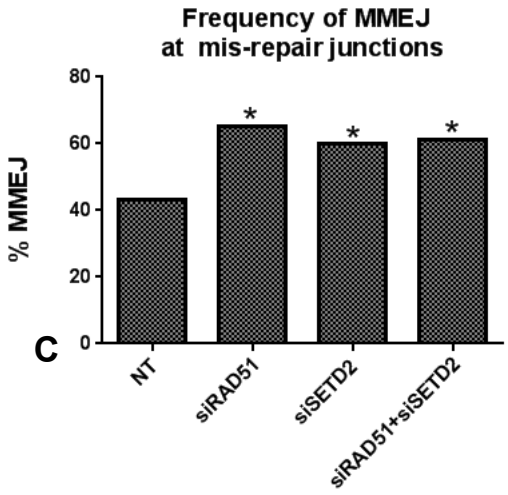


Figure 3.9. Frequency of MMEJ events at HPRT mis-repaired junctions in cells treated with siRNAs against NT, RAD51, SETD2, or RAD51+SETD2.

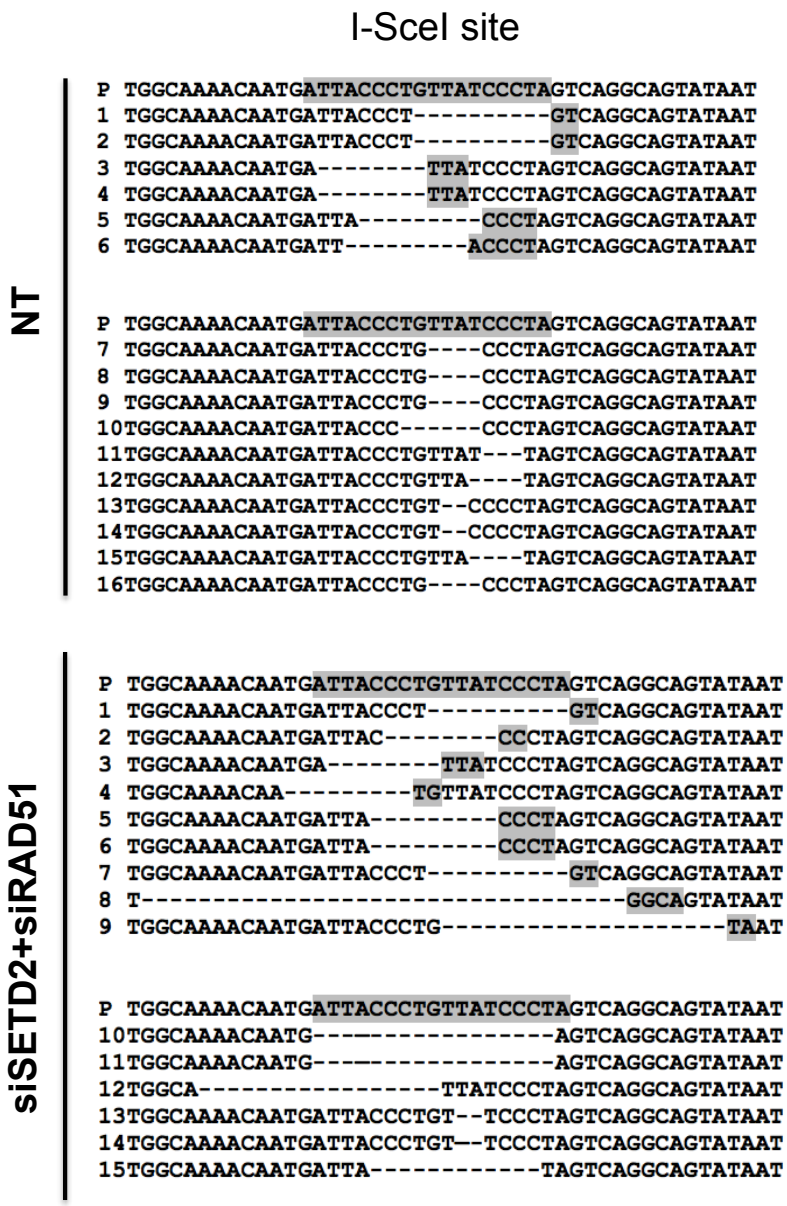


Figure 3.10. Representative sequence alignments of the PCR products in NT control (NT), and SETD2/RAD51-codepleted cells isolated from three independent experiments. I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted.

3.2.1 I-SceI cutting efficiency

To study the dynamics between the I-SceI-induced cleavage and repair, I amplified the I-SceI site using the primers across the break quantitatively using the qPCR approach (Fig 3.13A). I measured the amount of intact PCR product at various time

points before and after I-SceI transfections, and used this as an indication of cleavage efficiency. At 4 hours following I-SceI transfections, the amount of intact PCR product was 40% indicative of cutting efficiency of 60%. Then at 6 hours following I-SceI transfections, the cutting efficiency dropped to 40%. The I-SceI cleavage efficiency was further reduced to 30% at 8 hours time point. And finally at 24 hours post I-SceI transfection, the minimum cutting efficiency of 20% was observed (Fig 3.11).

3.2.2 I-SceI cutting efficiency and *HPRT* gene transcription are not affected by SETD2 knockdown

The increased MMEJ (and reduced C-NHEJ) observed in the HPRT assay following SETD2 depletion could have arisen from an indirect effect of SETD2 depletion on the cutting efficiency. In order to rule out the possibility that SETD2 depletion affected I-SceI expression or activity, I tested for I-SceI protein expression and cleavage efficiency in cells transfected with either NT siRNA or SETD2 siRNA. Western blot analysis showed no difference in the I-SceI protein expression levels between the SETD2-depleted cells and the non-targeting controls (Fig 3.3D). Also I used the similar approach as described above (Fig 3.11) to measure the I-SceI cutting efficiency before and after the break. Comparison of the quantity of PCR products in SETD2-depleted cells and that of non-targeting controls revealed no significant difference. This indicates that cleavage efficiency is the same in cells transfected with either NT siRNA or SETD2 siRNA (Fig 3.13B). In addition, *HPRT* gene transcription was not affected by SETD2 knockdown, as shown by quantitative RT-PCR (qRT-PCR) of spliced and unspliced HPRT mRNA (Fig 3.13C&D). Therefore, the mutation frequency in this system is likely to be directly associated with mis-repair of the DSB at the I-SceI site (Pfister et al., 2014).

Importantly, although PCR amplification across the break junctions in the majority of HPRT- negative clones generated PCR products, a substantial proportion resulted in

no PCR products from the target locus, while other loci were amplified normally. This was the case in all genetic backgrounds examined in this chapter and that of chapter 4. Table 3.1 summarises the proportion of the HPRT-negative/PCR-negative clones to the HPRT-negative/PCR-positive clones in cells treated with different siRNAs/inhibitors in the course of this study.

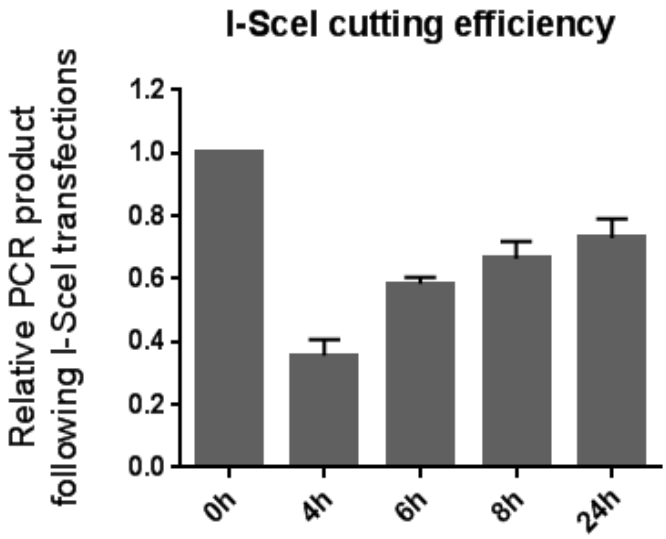


Figure 3.11. I-SceI cutting efficiency at different time points upon I-SceI transfections.

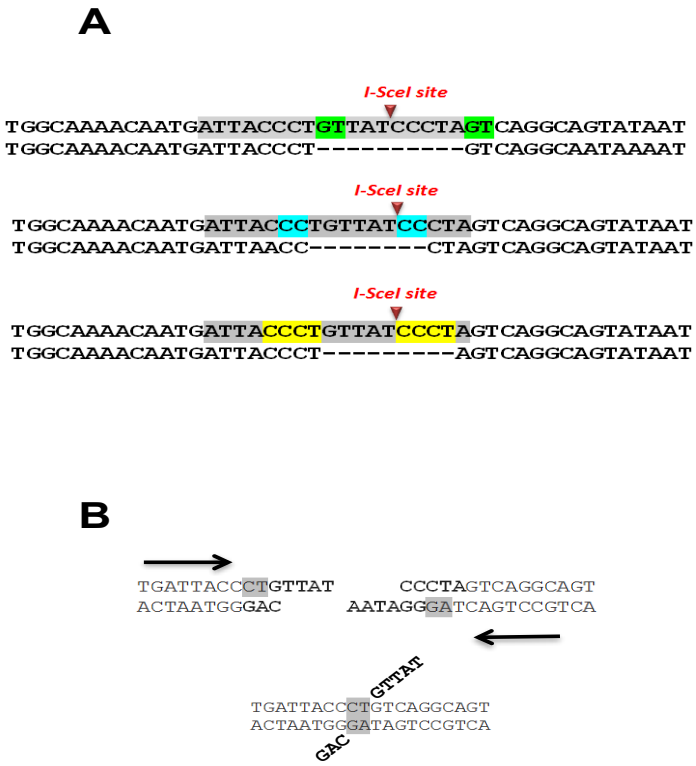


Figure 3.12. A. Examples of MMEJ patterns. The upper sequences represent the parental sequence, and the lower sequences represent the HPRT mutants. The highlighted bits indicate the microhomology regions across the ISceI break site and the deletions are indicated by dashes. B. Cartoon illustrating the formation of the microhomologies at the break site following a break. I-SceI recognition site is shown in bold and microhomologies are highlighted. Annealing of microhomologies results in nucleotide deletions shown in bold black.

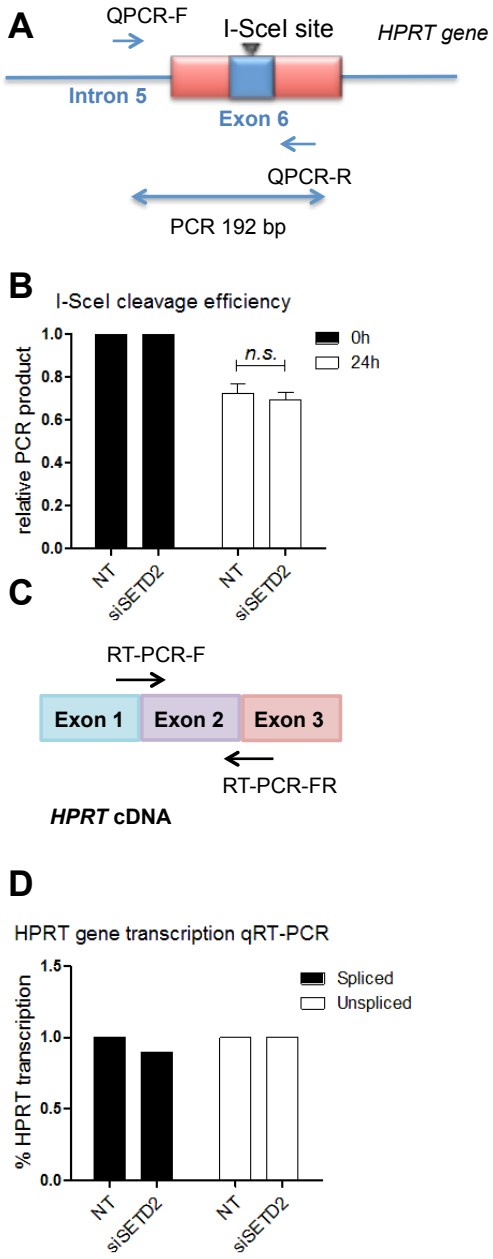


Figure 3.13. I-SceI cutting efficiency and *HPRT* gene transcription is not altered in SETD2-depleted cells. A. Schematic map of the qPCR primers used to amplify across the I-SceI break side to assess I-SceI cutting efficiency. B. Quantification of the I-SceI cleavage efficiency in the NT and SETD2 depleted cells in the HPRT assay. C. Schematic map showing the qRT-PCR exonic primers used to amplify the spliced HPRT mRNA. D. qRT-PCR quantification of HPRT mRNA in the NT and SETD2 depleted cells. Spliced mRNA was amplified by the exonic primers shown in (C) and unspliced mRNA was amplified by the intronic/exonic primers shown in (A).

Table 3.1. The proportion of HPRT-negative PCR-negative clones to the total HPRT-negative clones in cells treated with different siRNAs/inhibitors.

	PCR negative	total	%PCR negative
NT	18	63	28%
SETD2	12	56	21%
BRCA1	13	57	22%
BRCA2	11	40	27%
MRE11	10	38	26%
CtIP	3	38	8%
POLQ	9	37	24%
RAD51	14	43	32%

3.3 Discussion for Chapter 3

The *SETD2* gene is a tumour suppressor and is mutated in an array of cancers (Lawrence et al., 2014). The gene encodes the H3K36 trimethyltransferase and we have recently found it to perform a role in HR (Pfister et al., 2014). We and others have described a role for SETD2 in promoting HR through recruitment of CtIP that facilitates resection (Pfister et al., 2014, Aymard et al., 2014). Here, using the endogenous HPRT DSB mutation assay I demonstrated that depleting SETD2 leads to significantly increased DSB-induced mutation frequencies associated with deletions. These findings identify a role for SETD2, like RAD51, in suppressing break-induced mutations. Further DNA sequence analysis revealed that deletions arising from SETD2 or RAD51 depletion resulted from similarly enhanced levels of mutagenic MMEJ. I therefore propose that MMEJ can arise from failed HR. My data suggests that depletion of SETD2 results in partially resected DNA ends that are not sufficient to promote efficient HR and are no longer substrates for C-NHEJ. Instead, these DNA ends become substrates for the mutagenic MMEJ. Similarly, I propose that following RAD51 loss HR cannot proceed normally. Although in this context resection is pursued normally, these resected ends are also poor substrates for C-NHEJ. Because HR cannot proceed, the ends are instead rejoined by the mutagenic MMEJ. Extensive resection following RAD51 loss facilitates annealing between regions of microhomology further from the break site, thus leading to larger deletions than those observed for SETD2. Together, these findings suggest a role for HR in maintaining genome stability through suppression of the mutagenic MMEJ. Importantly, while the interplay between HR and C-NHEJ is becoming increasingly understood, the relationship between HR and MMEJ is less clear. This chapter thus provides new mechanistic insights into the roles of SETD2, and RAD51 HR repair factors in preserving genome stability through suppressing the mutagenic MMEJ pathway. Therefore, the findings in this chapter provide novel insight into the interplay between the HR and MMEJ repair factors in humans cells.

A novel role for homologous recombination factors in suppressing microhomology-mediated end joining in human cancer cells

The data in the previous chapter suggested roles for SETD2 and RAD51 HR repair factors in suppressing MMEJ. Here, I capitalise on the findings in chapter 3 and investigate the interplay between HR and MMEJ, which is still uncertain, by exploring the interplay between HR and MMEJ repair factors in human cancer cells. This chapter describes a novel role for human HR factors in maintaining genome integrity by suppressing the mutagenic MMEJ following a DSB.

4.1 BRCA2 depletion promotes DSB-induced deletions typical of MMEJ

In the previous chapter I showed that siRNA-mediated depletions of SETD2 or RAD51 had similar effects on I-SceI-induced DSBs in the *HPRT* gene, namely an increased overall frequency of HPRT mutagenesis associated with increases in both deletion lengths and use of microhomologies on either side of the deletions (Pfister et al., 2014). These findings suggested a possible role for HR in suppressing MMEJ. To test this further, I applied the same approach to assess a possible role for other HR factors in suppressing MMEJ. To test the impact of disrupting HR factors on double strand break (DSB) repair outcome, I depleted BRCA2, a key recombination mediator in mammalian cells that facilitates RAD51 loading onto single stranded DNA. BRCA2 depletion in HT1080 cells was confirmed by western blot (Fig 4.1A). I then measured the frequency of HPRT inactivation (mutation frequency) following I-SceI-induced DSBs. Knockdown of BRCA2 resulted in a significantly increased I-

SceI-induced mutation frequency (2.8%; P < 0.0001), compared to cells treated with non-targeting control (NT) siRNA (1.1%) (Fig 4.1B). Sequence analysis of HPRT-negative clones derived from cells treated with NT siRNA (Fig 4.2,3) or BRCA2 siRNA (Fig 4.2,4) indicated average deletion lengths of 5.6 and 49 bp, respectively. Thus BRCA2 depletion resulted in a 9-fold increase in deletion lengths (P <0.0001) (Fig 4.5A). Insertions at the HPRT break site were also detected but their frequencies were not detectably altered by BRCA2-depletion (Fig 4.25-26). Further examination of the sequences at deletion junctions showed that the proportion of deletions appearing to have arisen by MMEJ rose significantly from 43%, for cells treated with NT control siRNA, to 72% for cells treated with BRCA2 siRNA (P< 0.05) (Fig 4.2-4, 5B). To assess the effect of BRCA2 depletion on HR, I used a previously described (Pfister et al., 2014) U2OS cell line carrying a well-characterised GFP-based reporter for HR (DR-GFP). I found that the frequency of HR in BRCA2-depleted cells was significantly reduced by 41% compared to the NT controls (P< 0.0001), consistent with a role for BRCA2 in promoting HR (Moynahan et al., 2001) (Fig 4.1C).

Thus BRCA2 depletion causes both a decrease in HR and an apparent increase in MMEJ suggesting roles for BRCA2 not only in promoting HR but also in suppressing MMEJ (Ahrabi et al., 2016).

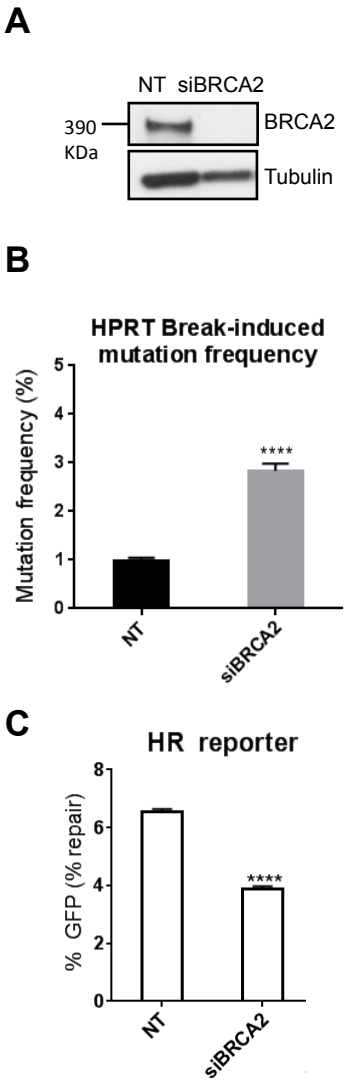


Figure 4.1. BRCA2 depletion promotes DSB-induced deletions typical of MMEJ.A. Western blot showing BRCA2 knockdown in HT1080 cells 48 h following siRNA transfection. B. Break-induced mutation frequency of HPRT reporter cells treated with NT control siRNA (NT), and BRCA2 siRNA (siBRCA2). Error bars represent SEM from three independent experiments. ****P < 0.0001.C. HR repair efficacy of DR-GFP reporter cells treated with NT control siRNA (NT), and BRCA2 siRNA (siBRCA2), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments, ****P < 0.0001.

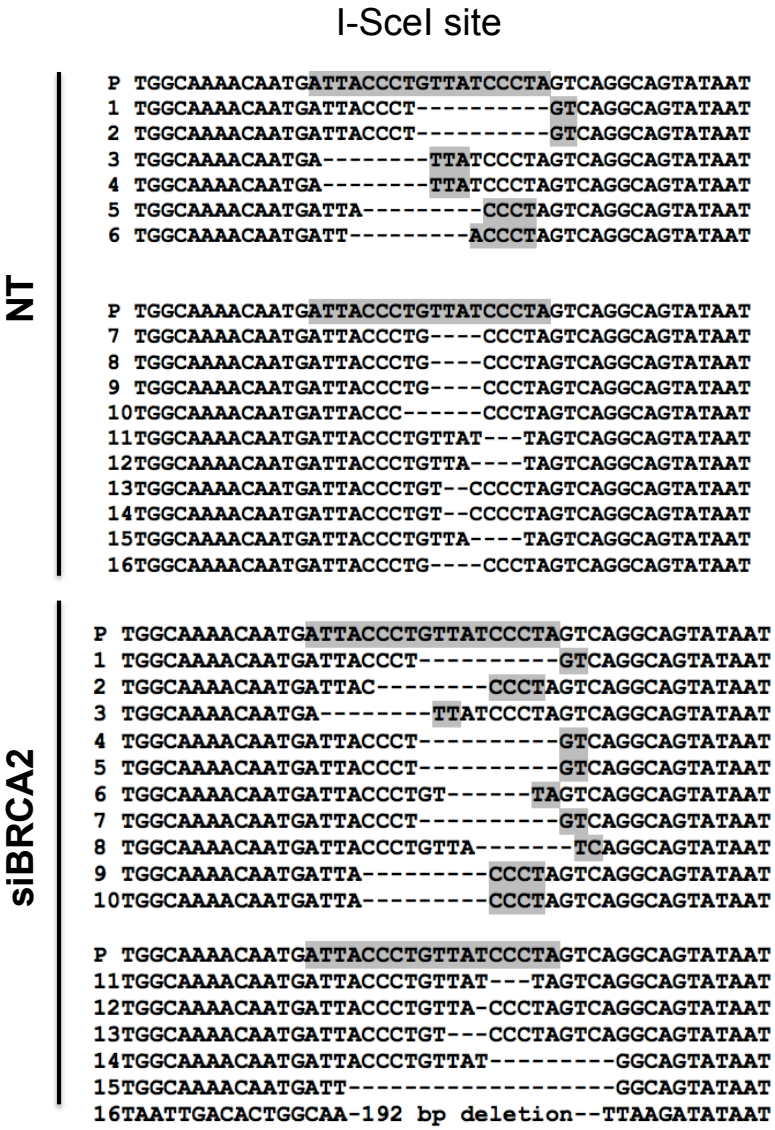


Figure 4.2. Representative sequence alignments of the PCR products in NT control (NT), and BRCA2-depleted (siBRCA2) cells isolated from three independent experiments. I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted.

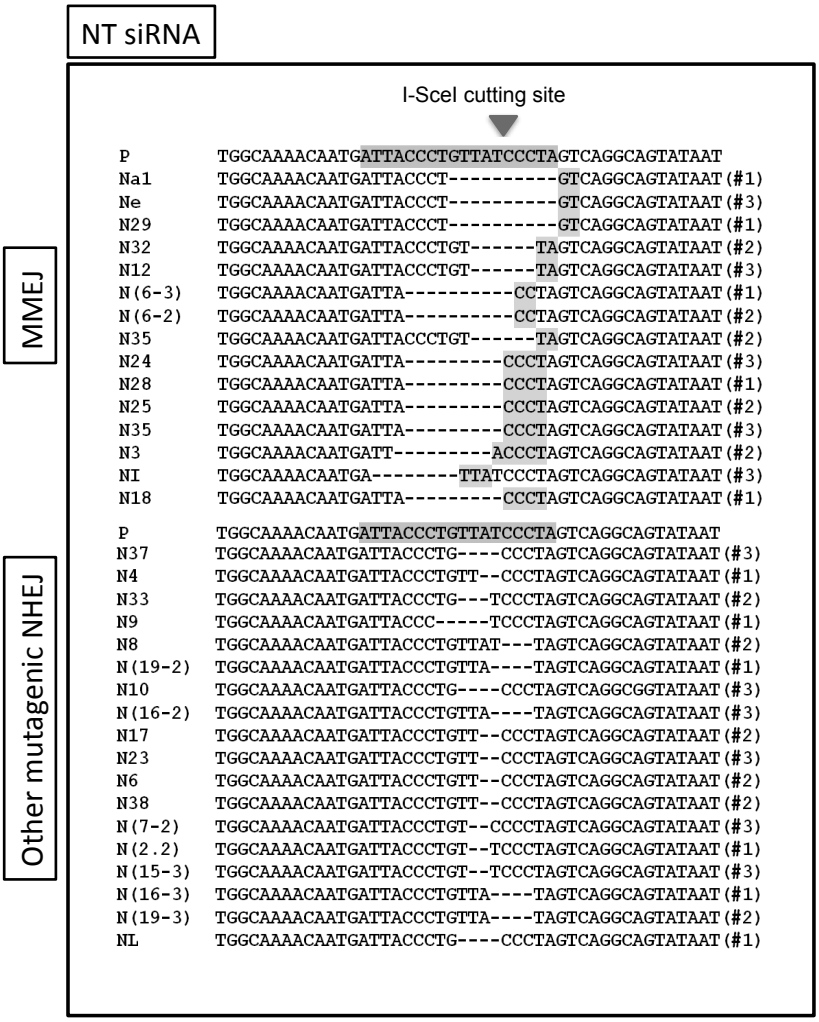


Figure 4.3. Sequence alignments of HPRT negative cells obtained from non-targeting control background.

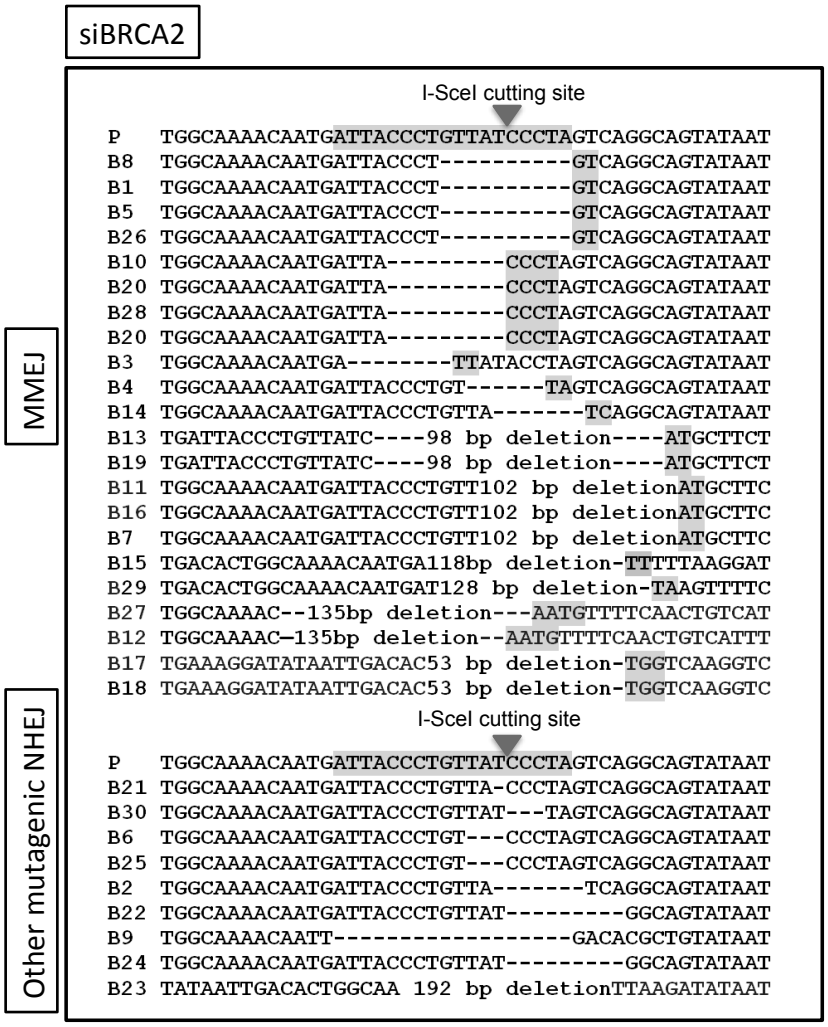


Figure 4.4. Sequence alignments of HPRT negative cells obtained from BRCA2-depleted background.

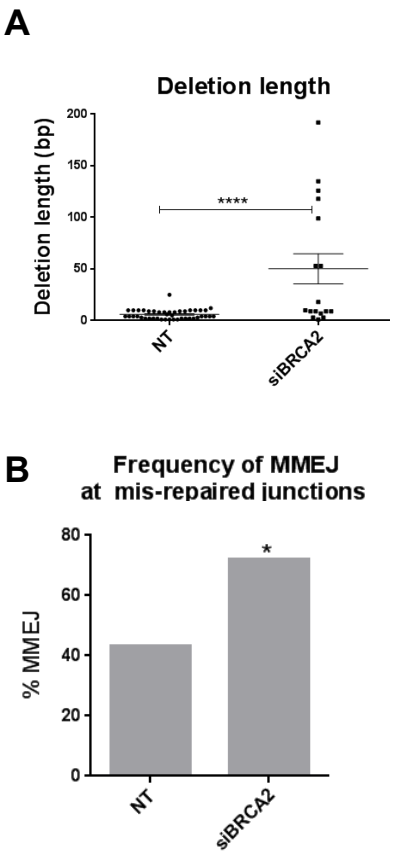


Figure 4.5. A. Average length of deletions (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, **** $P < 0.0001$. B. Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or BRCA2 siRNA (siBRCA2). P values calculated by statistical analysis "difference between proportions", * $P < 0.05$.

4.2 BRCA1 depletion promotes DSB-induced deletions typical of MMEJ

BRCA1 is also an essential factor for HR but has a distinct function in HR compared to BRCA2, where BRCA1 facilitates 5′ to 3′ resection of DSBs to generate 3′ ssDNA tails (Kabotyanski et al., 1998). To investigate a possible role for BRCA1 in suppressing MMEJ, I examined the effect of BRCA1 depletion on DSB mutational signatures in the HPRT reporter system. Like BRCA2 depletion, BRCA1 depletion resulted in a significantly enhanced mutation frequency (3.5%; P=0.001) following I-SceI-induced DSBs compared to that of NT controls (1.1%) (Fig 4.6B). Further, sequencing of individually isolated HPRT mutated clones from a BRCA1-depleted background indicated an average deletion length of 17 bp, significantly (3.4 fold; P= 0.04) greater than that (5 bp) in NT control siRNA-treated cells ((Fig 4.7-9).. Interestingly, this average deletion length of 17 bp in BRCA1-depleted cells was significantly (2.9 fold; P< 0.05) smaller than that in BRCA2-depleted cells (49 bp, Fig 4.5A). This is consistent with the early role for BRCA1 in promoting resection during HR, prior to BRCA2-assisted RAD51 loading. Insertions at the *HPRT* break site were also detected but their frequency was not detectably altered by BRCA1-depletion (Fig 4.26A). Among the sequenced junctions in HPRT negative cells, 66% appeared to have been generated by MMEJ following BRCA1 depletion, a 1.5-fold increase (P< 0.05) over the 43% in cells with normal BRCA1 levels (Fig 4.9B), very similar to the effect of BRCA2 depletion on MMEJ (Fig 4.5B). Therefore, BRCA1 depletion, like BRCA2-depetion, resulted in a significant increase in the proportion of mutagenic NHEJ events appearing to occur by MMEJ (Fig 4.9B). BRCA1 depletion in HT1080 cells was confirmed by Western blot (Fig 4.6A). Consistent with the role of BRCA1 in HR (Moynahan et al., 1999), HR was significantly reduced by 52% (P< 0.0001) in the DR-GFP reporter (U2OS cells) (Pierce et al., 1999) following BRCA1 knockdown (Fig

4.6C). Thus, BRCA1 depletion results in reduced HR repair and significantly increased mutation frequency associated with a larger proportion of MMEJ events (Ahrabi et al., 2016).

Together with the similar effects of BRCA2, RAD51 and SETD2 depletion, these results suggest a common role for HR factors in suppressing MMEJ.

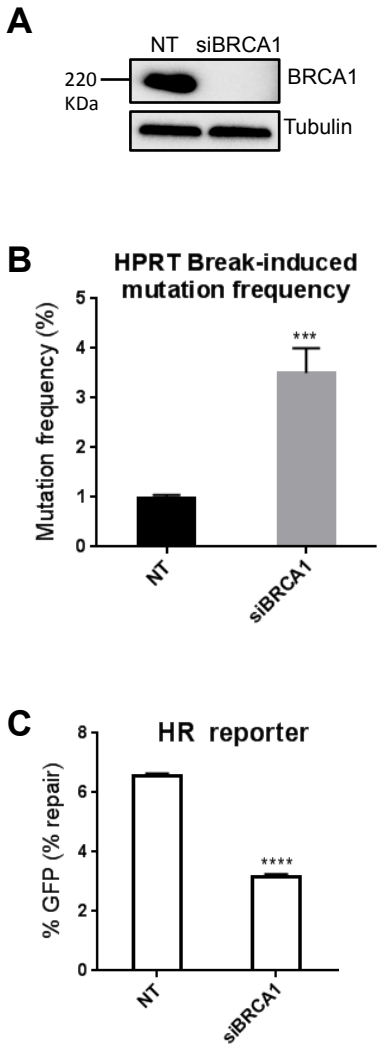


Figure 4.6. BRCA1 depletion promotes DSB-induced deletions typical of MMEJ. A. Western blot showing BRCA1 knockdown in HT1080 cells 48 h following siRNA transfection. B. Break-induced mutation frequency ofHPRT reporter cells treated with NT control siRNA (NT), BRCA1 siRNA (siBRCA1). Error bars represent SEM from three independent experiments. *** $P < 0.001$. C. HR repair efficacy of DR-GFP reporter cells treated with NT control siRNA (NT), and BRCA1 siRNA (siBRCA1), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments, **** $P < 0.0001$.

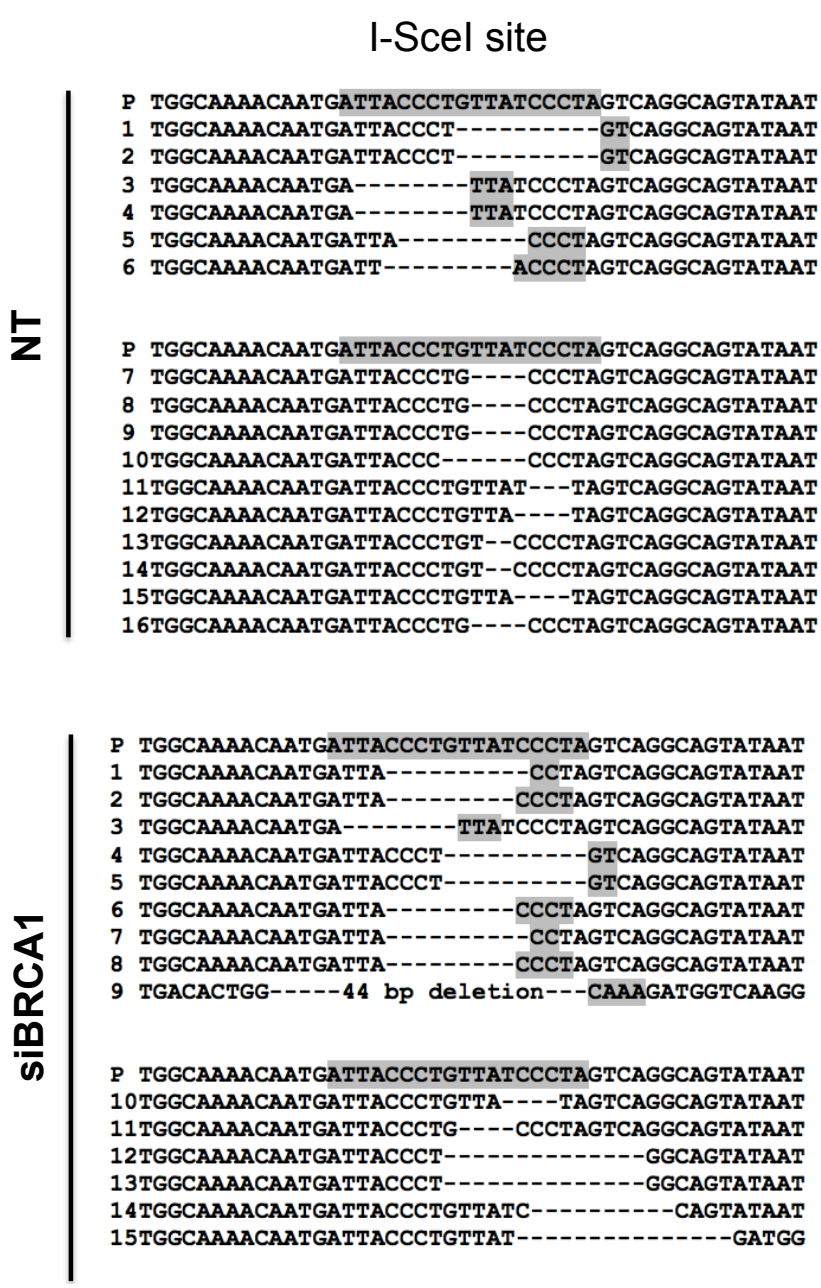


Figure 4.7. Representative sequence alignments of the PCR products in NT control (NT), and BRCA1-depleted (siBRCA1) cells isolated from three independent experiments. I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted.

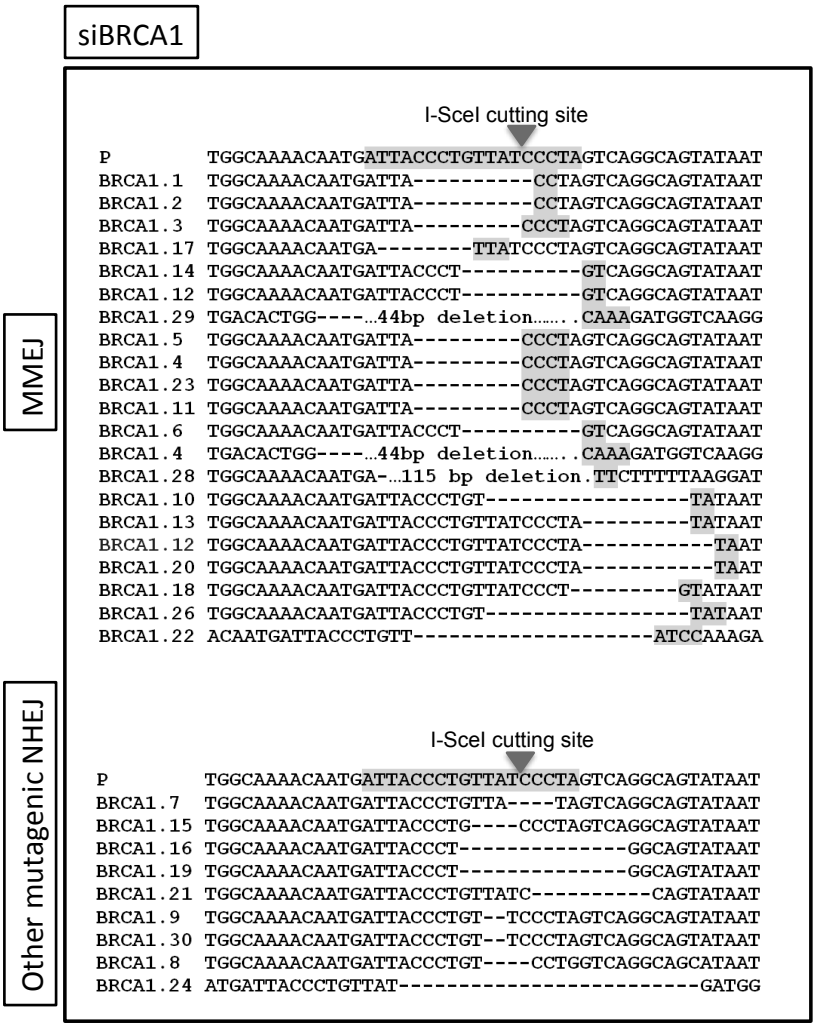
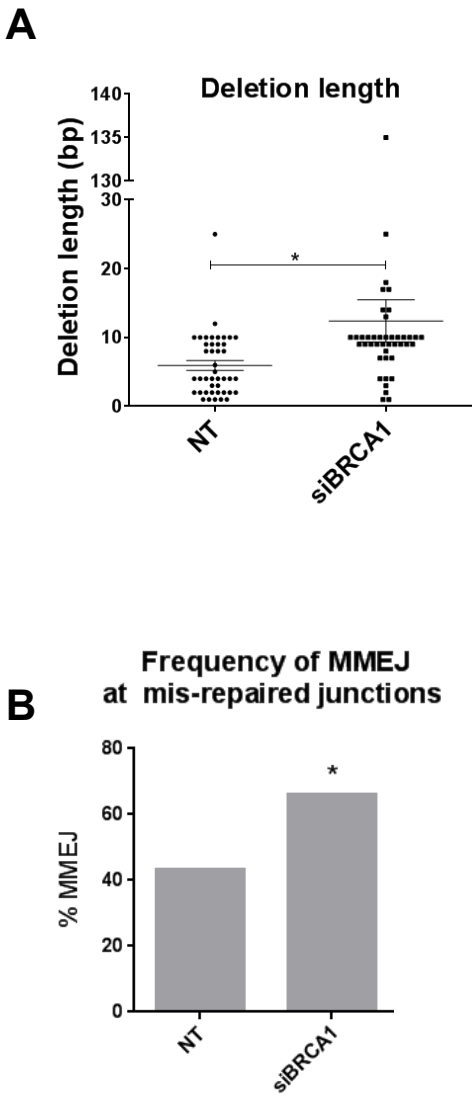


Figure 4.8. Sequence alignments of HPRT negative cells obtained from BRCA1-depleted background.

\



\

Figure 4.9. A. Average length of deletions (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean andSEM, **** P < 0.0001. B. Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or BRCA1 siRNA (siBRCA1). P values calculated by statistical analysis “difference between proportions”, * P < 0.05.

4.3 MMEJ-like DSB repair in HR deficient cells is independent of DNA-PKcs

Although the microhomologies I observed at repair junctions in HR-defective cells were highly suggestive that MMEJ is the key pathway involved, I wanted to test the possible involvement of C-NHEJ. Because C-NHEJ, but not MMEJ, is dependent on DNA-PKcs, I therefore, compared the HPRT mutational profile following treatment of cells with a DNA-PKcs inhibitor (NU7441), or siRNA against RAD51 in the presence and absence of NU7441. This inhibitor has previously been shown to disrupt the C-NHEJ repair pathway (Zhao et al., 2006). I found that treating cells with NU7441 resulted in a significant increase in the deletion length (from 5 to 16 bp, P=0.015) and a significantly increased presence of microhomologies at the break sites (form 43% to 60%, P< 0.05) compared to that of the NT background (Fig 4.10,11A-B). This is consistent with a role for DNA-PKcs in conjunction with Ku in protecting DSB ends and thereby promoting C-NHEJ and counteracting MMEJ (Mahaney et al., 2009, Iliakis, 2009). Further, I found that the RAD51-depleted mutational pattern following DSB induction was largely unaffected by the presence of the DNA-PKcs inhibitor. In particular, similar deletion lengths and use of microhomologies at the HPRT break-sites were observed following RAD51 depletion in the absence or presence of NU7441 (P >0.05) (Fig 4.10,11A-B). No statistical difference was found between the MMEJ levels in cells treated with NU7441, siRNA against RAD51, or siRNA against RAD51 in combination with NU7441 (Fig 11B). These observations suggested that the DSB mutational signature in the absence of HR is largely independent of DNA-PKcs, a core C-NHEJ factor. To confirm this in another system, I used a well-defined GFP-based MMEJ reporter integrated within a U2OS cell line (EJ2-GFP) (Gunn and Stark, 2012) (Fig 12A). Knockdown of BRCA2 in this system significantly increased the GFP positive cells, indicative of MMEJ repair, compared to the NT controls (1.95 fold, P=0.0105) (Fig 12B), confirming that BRCA2 suppresses MMEJ. Also, DNA-

PKcs inhibition by NU7441 in this reporter cell line significantly increased the MMEJ repair events by 2.24 fold (P=0.0129) (Fig 12B), consistent with a well-established role for C-NHEJ factors in suppressing MMEJ (Iliakis, 2009). Further, inhibition of DNA-PKcs by NU7441 in BRCA2-depleted cells led to a slight increase of the GFP-positive cells compared to that of BRCA2-depleted cells (1.62 fold; P=0.0523) or control cells treated with NU7441 (1.4 fold; P>0.05) (Fig 12B). When I repeated these experiments using BRCA1 depletion in place of BRCA2 depletion I obtained very similar results (Fig 12C). These data obtained from the GFP reporter are consistent with those from the HPRT assay; in neither system did treatment with DNA-PKcs inhibitor impair the increase in MMEJ resulting from depletion of an HR protein suggesting that DNA-PKcs is dispensable for the induction of MMEJ in the HR-deficient cells.

Finally, to confirm that NU7441 efficiently inhibits C-NHEJ, I treated the previously described GFP-based C-NHEJ reporter (IRES-TK-EGFP, H1299 cells; (Ogiwara et al., 2011)(Fig 4.13A) with NU7441 and observed a 93.6% (P< 0.0001) fall in C-NHEJ (Fig 4.13B) (Ahrabi et al., 2016).

Collectively these data suggest that MMEJ in HR-impaired cells is independent of DNA-PKcs, a key factor for the C-NHEJ pathway. Also, the GFP-based reporter data suggest that BRCA2 and BRCA1 suppress MMEJ.

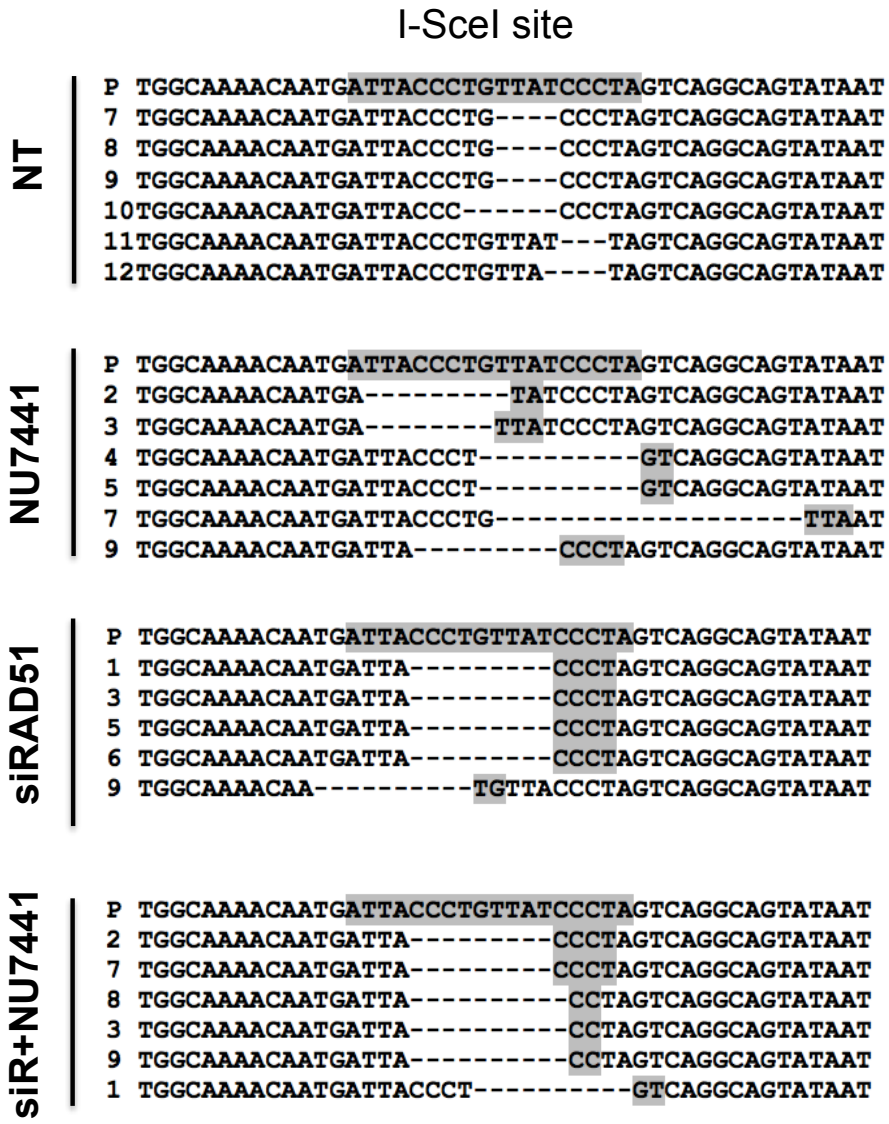


Figure 4.10. MMEJ-like DSB repair in HR-deficient cells is independent of DNA-PKcs. Representative sequence alignments of PCR products obtained from HPRT negative cells treated with NT control siRNA (NT), a DNA-PKcs inhibitor (NU7441), siRNA against RAD51 (siRAD51) or siRNA against RAD51 plus NU7441 (siR+NU7441), from three independent experiments. I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted. The sequences above are only examples of the alignments from each genetic background showing that the MMEJ levels are significantly increased following treatments with DNA-PKcs inhibition, RAD51 depletion or simultaneous depletion of RAD51 and inhibition of DNA-PKcs. Full sequence alignments for the NT background is shown in Fig 4.3.

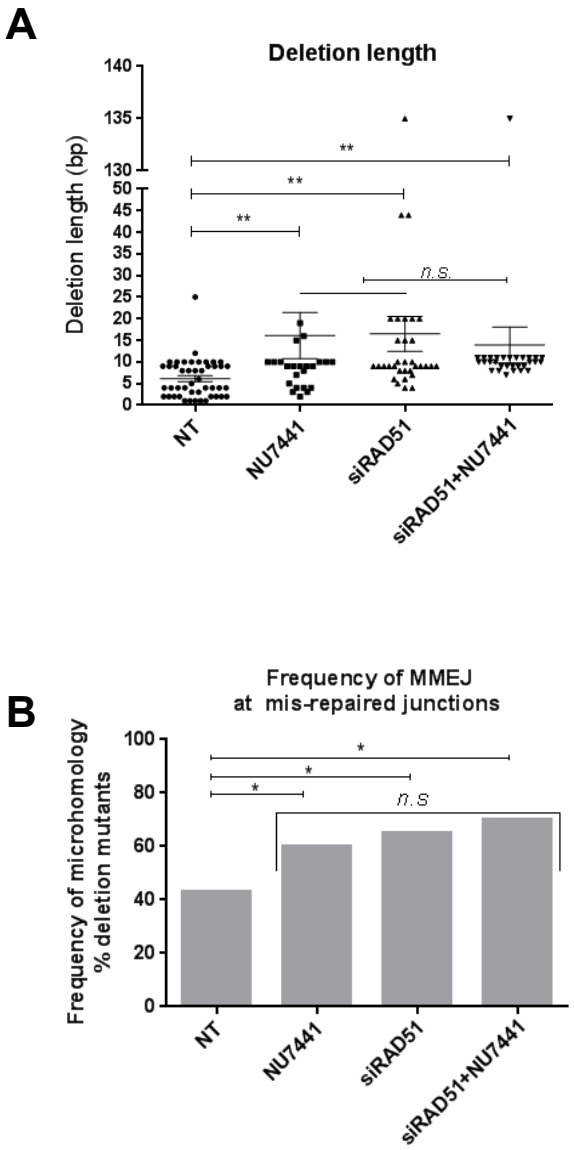


Figure 4.11. A. Average lengths of deletions (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, n.s., not significant, ** P < 0.01. B. Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT), NU7441, RAD51 siRNA (siRAD51) or RAD51 and NU7441 (siRAD51+NU7441). P values calculated by statistical analysis “difference between proportions”, *P < 0.05.

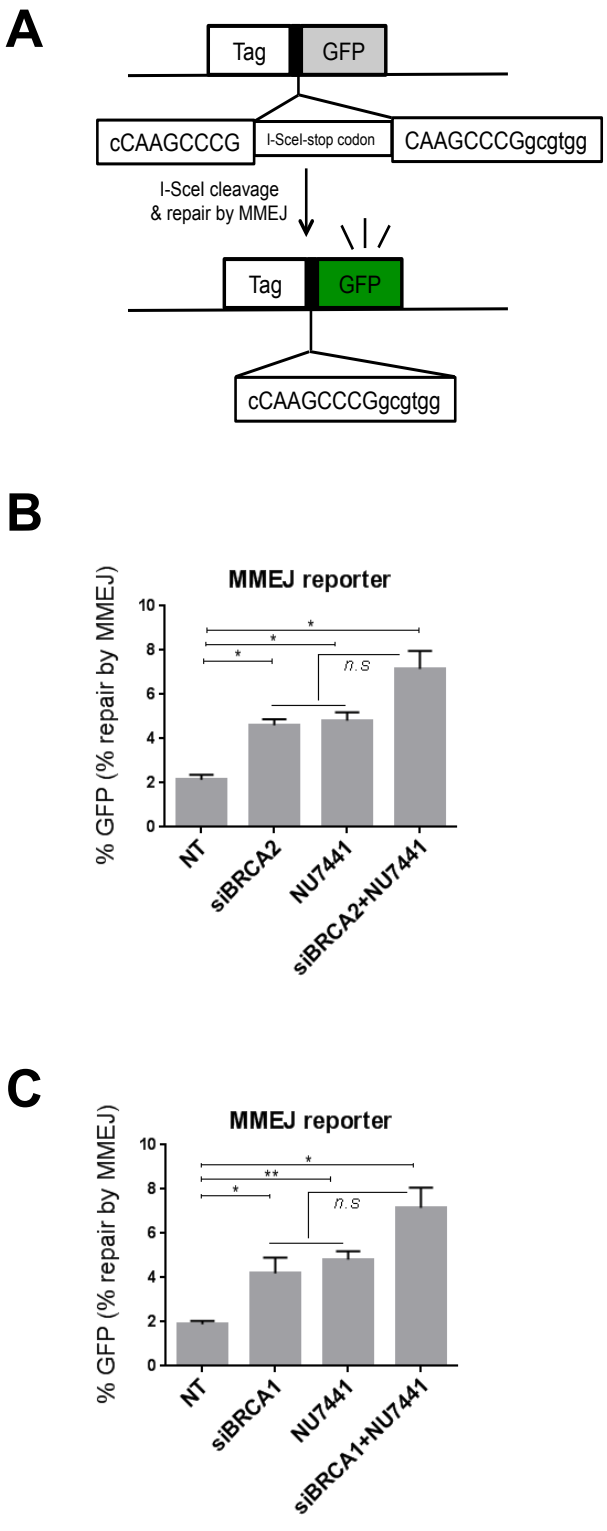


Figure 4.12. A. Schematic map of the EJ2-GFP reporter (53) to assess MMEJ efficacy, where the I-SceI cut site is flanked by 8 nucleotide homologous sequences capable of bridging the I-SceI-induced DSB by MMEJ and thus restoring a functional GFP cassette. B. MMEJ repair efficacy of EJ2-GFP reporter cells treated with NT control siRNA (NT), BRCA2 siRNA (siBRCA2), DNA-PKcs inhibitor (NU7441) or BRCA2 and DNA-PKcs co-depleted cells (siBRCA2+NU7441), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. *P < 0.05. C. MMEJ repair efficacy of EJ2-GFP reporter cells treated with NT control siRNA (NT), BRCA1 siRNA (siBRCA1), DNA-PKcs inhibitor (NU7441) or BRCA1 siRNA and DNA-PKcs (siBRCA1+NU7441), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. *P < 0.05, ** P < 0.01.

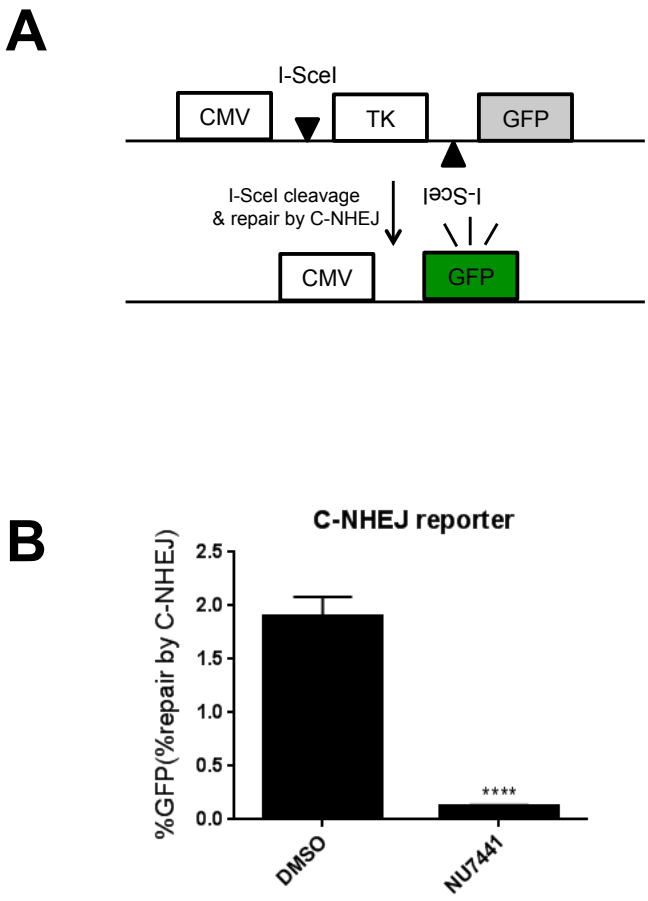


Figure 4.13. A. Schematic map of the IRES-TK-EGFP reporter to assess C-NHEJ efficacy. B. C-NHEJ repair efficacy of IRES-TK-EGFP reporter cells treated with DMSO (control), and DNA-PKcs inhibitor (NU7441) indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. **** $P < 0.0001$.

4.4 DSB-induced mutation signatures in HR-deficient cells require MMEJ factors

The above evidence shows clearly that DSB mis-repaired junctions with microhomologies typical of MMEJ are enhanced when HR is impaired, and that this effect is C-NHEJ-independent. I nevertheless sought direct evidence that MMEJ is indeed the pathway responsible for these mutational signatures. I therefore used the HPRT system to examine the DSB-induced mutational signatures following depletion of essential MMEJ factors. Given that HR and MMEJ share the initial resection step via CtIP and MRE11(Truong et al., 2013, Bennardo et al., 2008, Jasin and Rothstein, 2013), I investigated the impact of depleting these factors on the DSB-induced mutational signature. CtIP and MRE11 were depleted by siRNA-mediated knock down in the HPRT reporter cells and the mutational frequencies following I-SceI-induced break were quantified. I found that, similar to the depletion of RAD51, SETD2, BRCA2 and BRCA1, depletion of either CtIP or MRE11 resulted in a significant increase in the I-SceI-induced mutation frequency (3.5%; P=0.0007; and 3.6%; P < 0.0001 respectively) compared to 1% in NT controls (Fig 4.14B).

The mutational frequencies in these backgrounds were slightly but significantly (P<0.05) greater than those in BRCA2 depleted cells, (Fig 4.1) consistent with roles for CtIP, and MRE11 upstream of BRCA2 during HR. This perhaps reflects an important role for early acting repair/signalling factors in maintaining genome integrity in response to a DSB. Surprisingly, sequence analysis of HPRT-negative clones derived from CtIP- or MRE11-depleted cells identified a distinct mutational signature compared to that of SETD2-, RAD51-, BRCA2-, or BRCA1-depleted cells (Fig 4.15-17). This signature involved deletions with comparable lengths to those in NT controls (mean 5.6 bp), in contrast to the increased deletion lengths seen in BRCA2-

depleted cells (mean 49bp; P<0.0001, and P<0.001 respectively) (Fig 4.19). Further analysis of the mis-repaired junctions of HPRT-negative clones from these backgrounds identified a significant decrease in the proportion of repair junctions associated with microhomologies in both CtIP- (13%; P< 0.05), and MRE11- (31%; P< 0.05) depleted cells, compared to NT controls (43%, Figure 4.5B), or to BRCA2 (72%, Figure 4.5B) and BRCA1-depleted cells (65%, Figure 4.6B). In summary, depletion of CtIP or MRE11, essential MMEJ factors that promote the initiation of resection, significantly reduced the proportion of DSB-induced deletions that were associated with microhomologies, relative to the NT controls (Fig 4.19B). This contrasts sharply with depletion of the HR proteins SETD2, RAD51, BRCA2, or BRCA1, which significantly enhanced the same proportion (Figures 4.5B & 4.6B)(Pfister et al., 2014).

To assess the effect of CtIP and MRE11 depletion on HR, I knocked down CtIP and MRE11 in the DR-GFP reporter (U2OS cells) (Pierce et al., 1999) and found that the frequency of HR was significantly reduced by 92% (P< 0.0001), and 71% (P<0.0001) respectively (Fig 4.14C). CtIP and MRE11 depletions were confirmed by Western blot (Fig 4.14A). Together, these findings are in accordance with roles for CtIP and MRE11 in the initiation of resection during both HR and MMEJ.

PARP-1 is also known to play a key role in MMEJ by competing with Ku to bind DSBs, which in turn leads to the recruitment of MRE11 and CtIP to trigger resection (Beck et al., 2014, Hochegger et al., 2006). Consistent with this, Olaparib has recently been shown to decrease the Alt-NHEJ levels in the MMEJ-GFP reporter assay (Howard et al., 2015). Therefore, I treated the HR-proficient HT1080 cells (HPRT reporter) with a PARP-1 inhibitor, Olaparib (AZD2281) (Zimmermann et al., 2013), and examined the DSB mutational pattern. The effects of Olaparib were similar to those of CtIP or MRE11 depletion. Thus Olaparib significantly increased the mutation frequency (2.6%; P=0.0073; Fig 4.15B) compared to the 1% in NT

controls, generating deletion lengths (mean 7 bp) comparable to the controls (mean 5.6 bp; P>0.05; Figures 4.15 & 4.19A), and significantly shorter than deletion lengths in BRCA2-depleted HPRT mutants (mean 49 bp; P= 0.0005). Similarly, HPRT mutants arising following DSB induction in cells treated with Olaparib showed a significant reduction in the proportion of deletion junctions associated with microhomologies (25%; P <0.05) relative to NT controls (43%) (Fig 4.15,18,19), Collectively, these results suggest that microhomologies at DSB mis-repaired junctions in the HPRT reporter do indeed originate from the MMEJ repair machinery.

To confirm that the MMEJ-like signature in HR-compromised cells arises from MMEJ, I aimed to see whether knocking down a core MMEJ factor could abolish the MMEJ signature in these cells. I therefore, co-depleted BRCA1 and MRE11 in the HPRT system and monitored the DSB mutational signature. Simultaneous knockdown of BRCA1 and MRE11 resulted in a significantly increased mutation frequency of 3.4% compared to that of the NT background (1.1%, P=0.0013) (Fig 4.20A). This increased frequency of HPRT loss is consistent with the roles for BRCA1 and MRE11 in promoting DSB repair (Moynahan et al., 1999, Zhong et al., 1999). Further, the MMEJ signature was absent in cells co-depleted for BRCA1 and MRE11. This was associated with deletion lengths comparable to that of NT cells (mean 5bp), and a significantly reduced proportion of MMEJ (27%) compared to that of NT cells (43%, P<0.05) (Fig 4.20 B, C, 4.21) (Ahrabi et al., 2016).

Together these data confirm that the MMEJ-like signature in the absence of HR factors is arising from the MMEJ pathway.

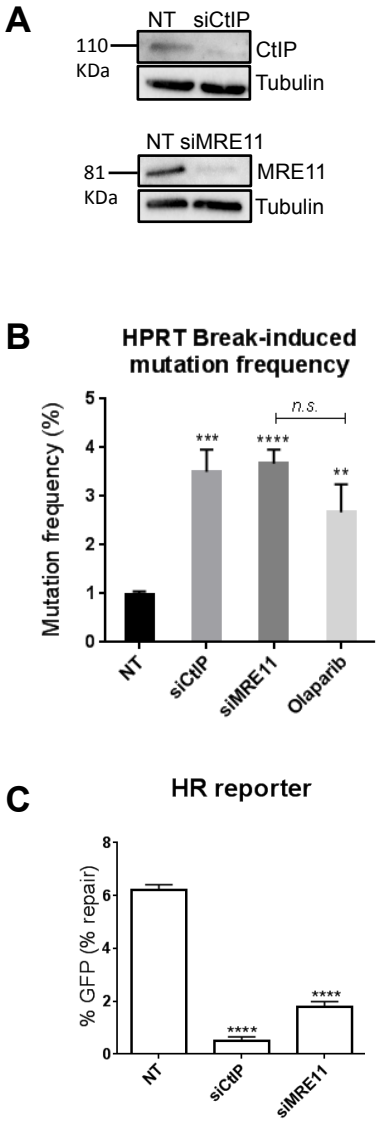


Figure 4.14. DSB-induced mutation signatures in HR-deficient cells require MMEJ factors. A. Western blot showing CtIP and MRE11 knockdowns 48 h following siRNA transfection. B .Break-induced mutation frequency of HPRT reporter cells treated with NT control siRNA (NT), CtIP siRNA (siCtIP), MRE11 siRNA (siMRE11) or PARP-1 inhibitor (Olaparib). Error bars represent SEM from three independent experiments, n.s., not significant, **P < 0.01, ***P < 0.001, ****P < 0.0001. C.HR repair efficacy of DR-GFP reporter cells treated with NT control siRNA (NT), CtIP siRNA (siCtIP), or MRE11 siRNA (siMRE11), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. ****P < 0.0001.

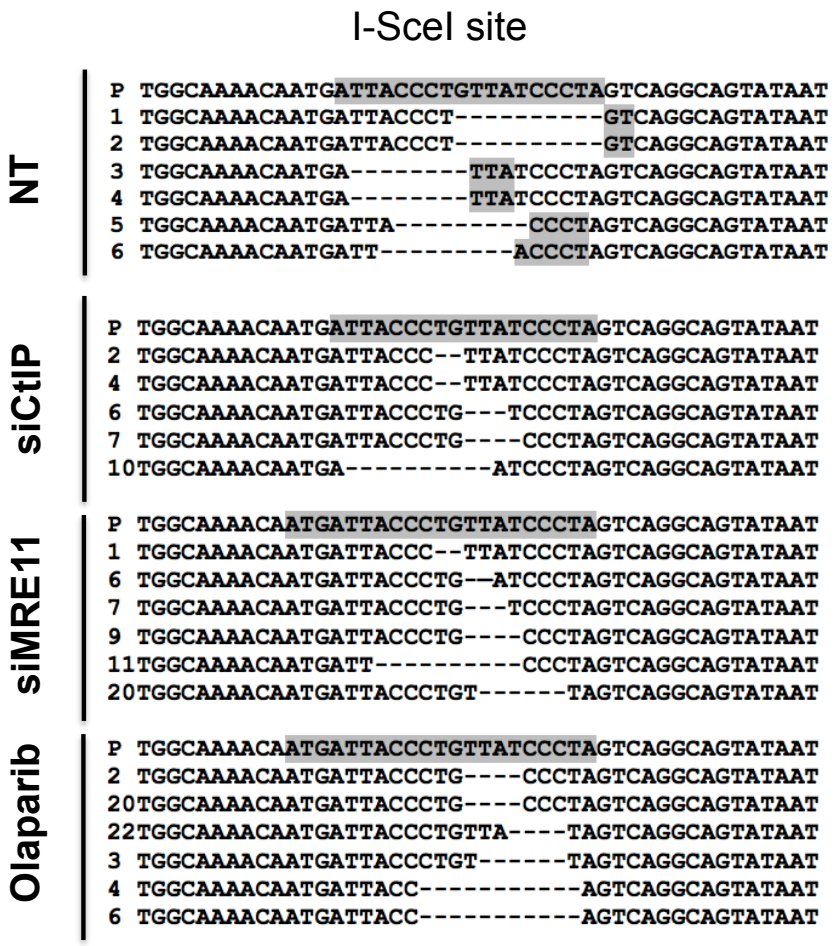


Figure 4.15. Representative sequence alignments of the HPRT negative PCR products in cells treated with NT control siRNA (NT), CtIP siRNA (siCtIP), MRE11 siRNA (siMRE11) or PARP-1 inhibitor (Olaparib), from three independent experiments.I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted.

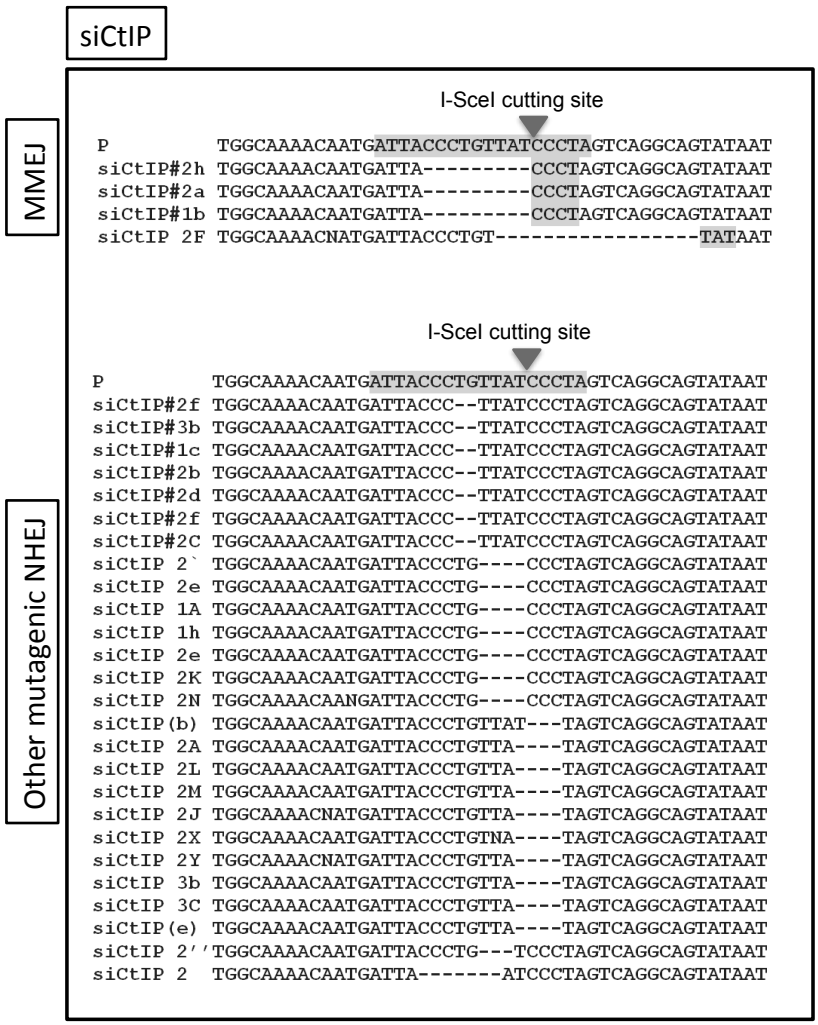


Figure 4.16. Sequence alignments of HPRT negative cells obtained from CtIP-depleted background. HPRT deletion mutants harbouring microhomologies at the breakpoint junctions (consistent with MMEJ) are shown on the top. HPRT deletion mutants with no microhomologies at breakpoint junctions (consistent with other mutagenic NHEJ) are shown on the bottom. Each line represents an individual HPRT negative clone. The parental uncut sequence (P) is shown on the top with the I-SceI recognition site highlighted in gray and the I-SceI cut site indicated with an arrow. Microhomologies are highlighted in gray adjacent to deleted nucleotides. Total of 30 individual cell colonies were isolated, from three independent experiments for the purpose of the corresponding statistical analyses.

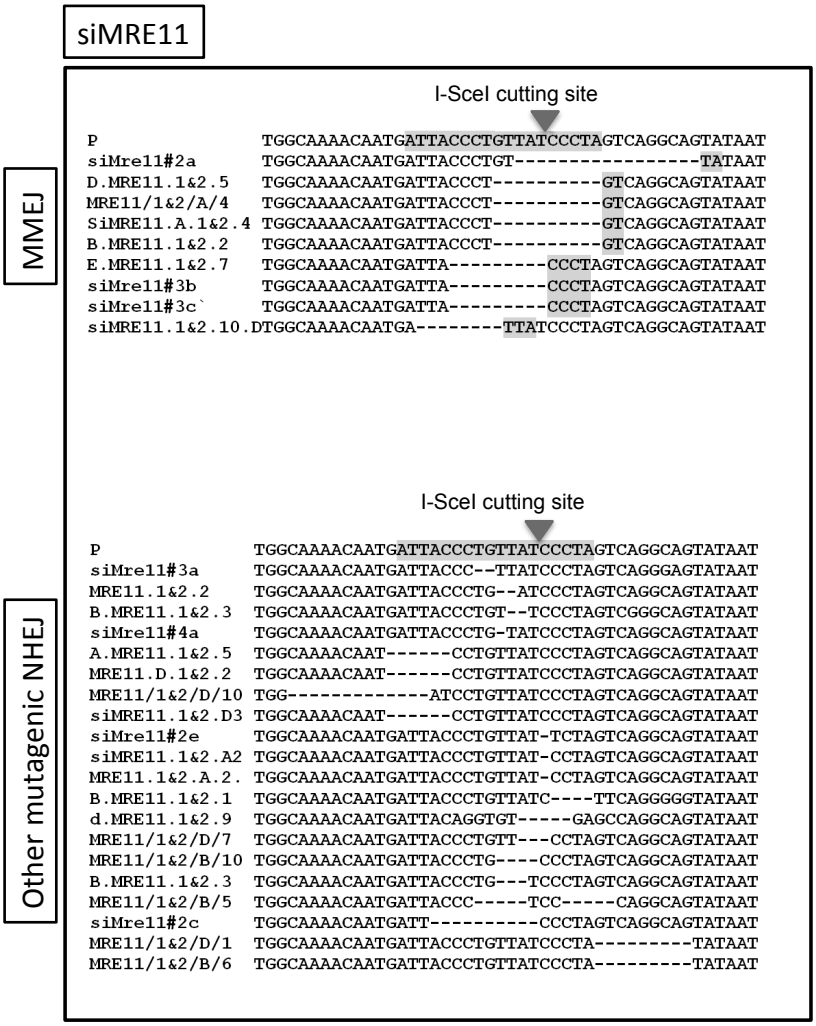


Figure 4.17. Sequence alignments of HPRT negative cells obtained from MRE11-depleted background.

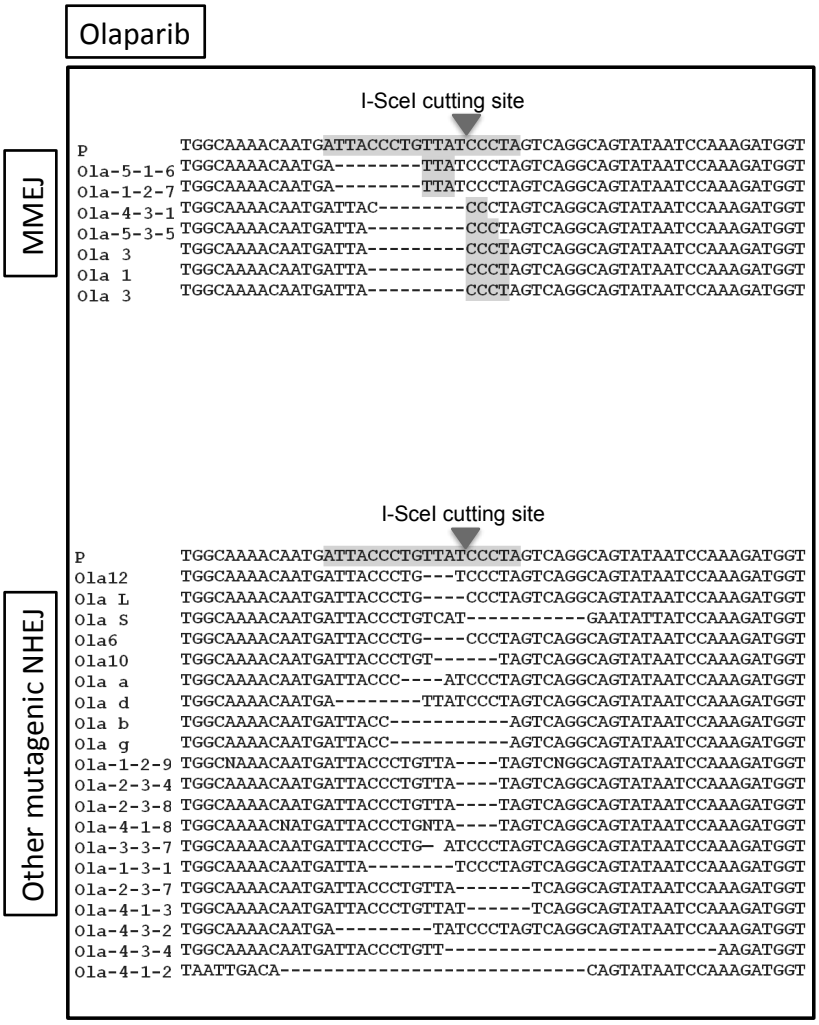


Figure 4.18. Sequence alignments of HPRT negative cells obtained from cells treated with PARP inhibitor (Olaparib)

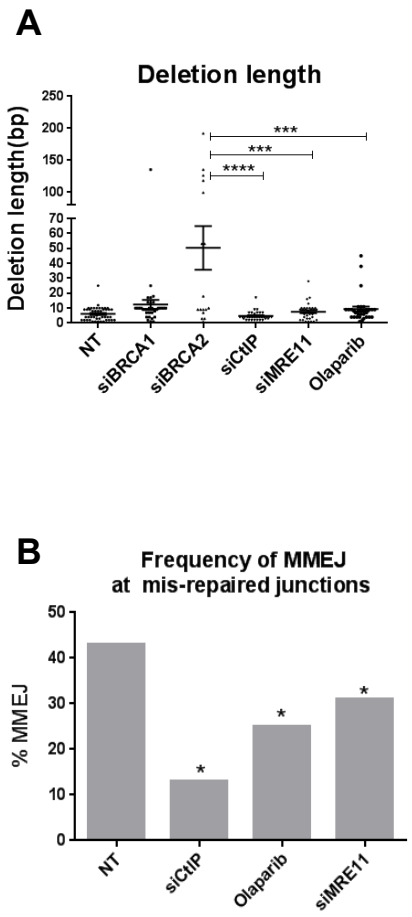


Figure 4.19. A. Average deletion lengths (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, ***P < 0.001, ****P < 0.0001. B. Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or CtIP siRNA (siCtIP), MRE11 siRNA (siMRE11) or Olaparib. P values calculated by statistical analysis “difference between proportions”, *P < 0.05.

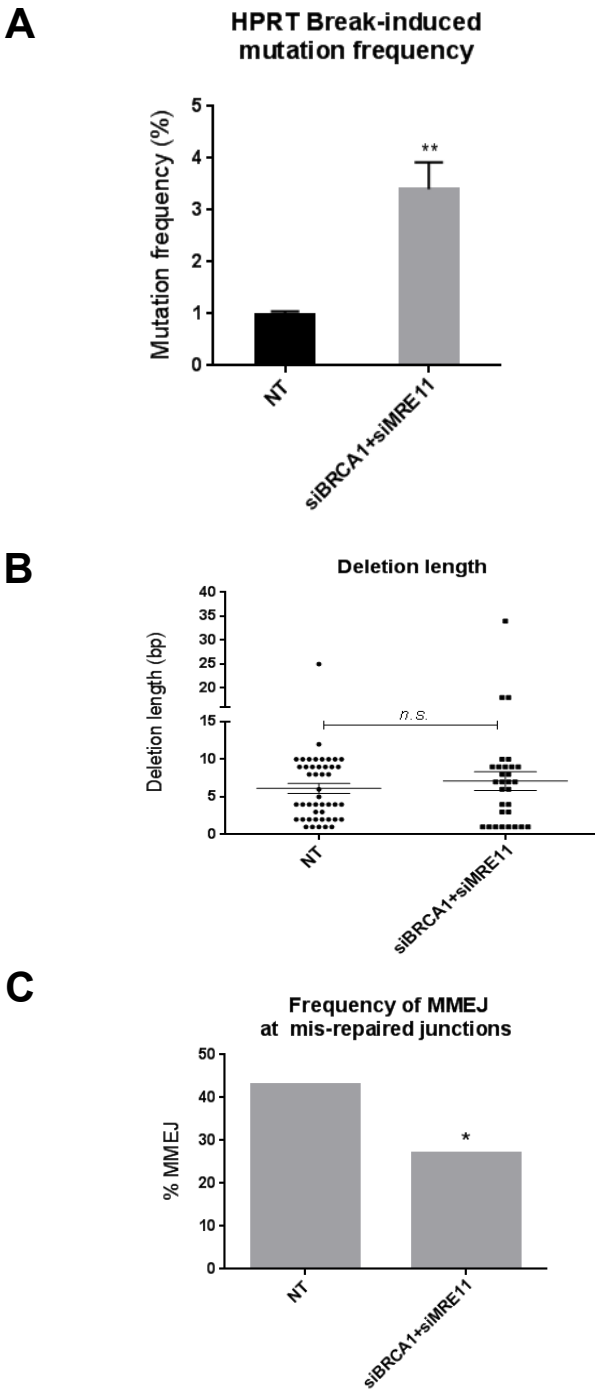


Figure 4.20. A. Break-induced mutation frequency ofHPRT reporter cells treated with NT control siRNA (NT), BRCA1 and MRE11 siRNAs (siBRCA1+siMRE11). Error bars represent SEM from three independent experiments. *** $P < 0.001$. B. Average deletion lengths (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, ***$P < 0.001$, ****$P < 0.0001$.

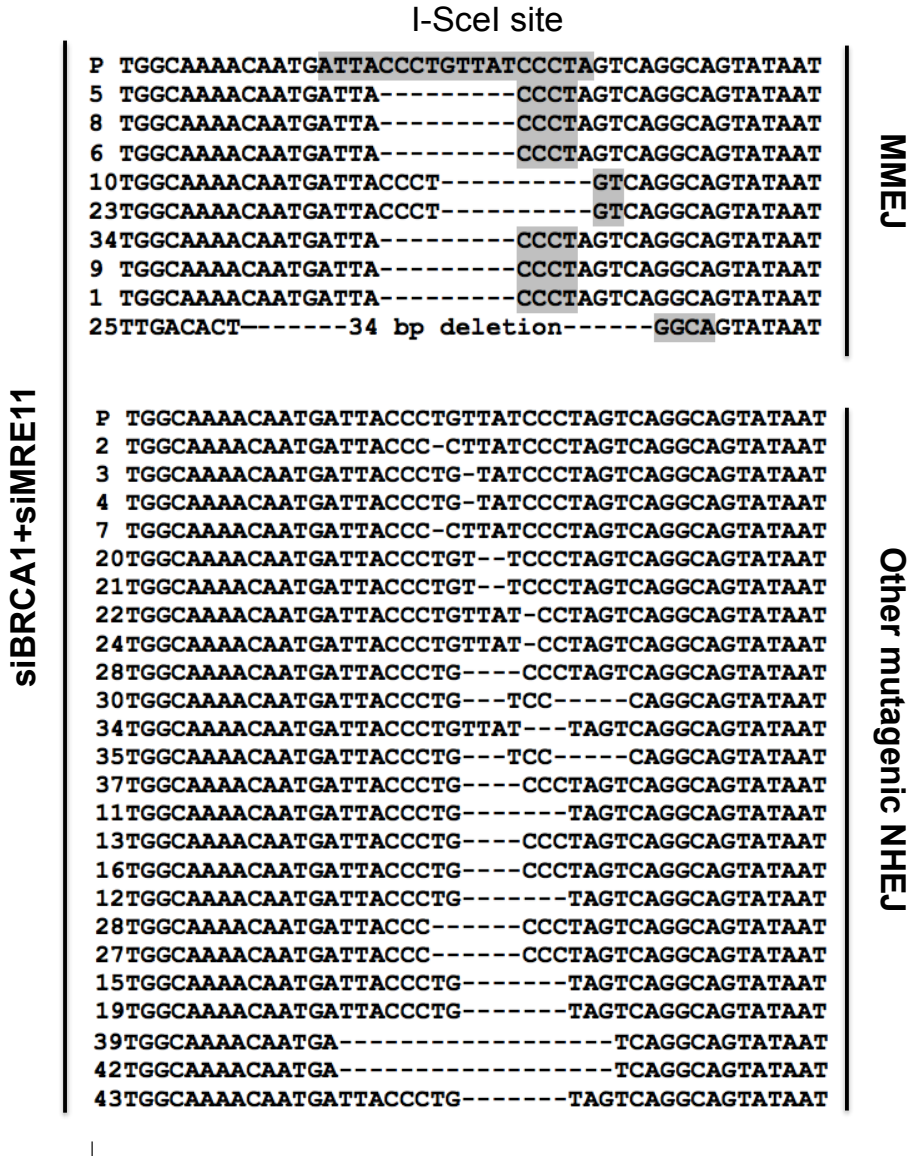


Figure 4.21. Sequence alignments of HPRT negative cells obtained from BRCA1/MRE11 co-depleted background.

4.5 RPA suppresses MMEJ

RPA is a heterotrimeric ssDNA binding protein with roles in DNA replication, recombination, and DSB repair, including HR. RPA displacement by RAD51 is a critical step in HR (Wold, 1997). In this regard, it has been previously shown in *S. cerevisiae* that RPA antagonises MMEJ (Deng et al., 2014). Therefore, to test a possible role for RPA in repressing MMEJ repair in human cells, I used the GFP-based reporter for MMEJ repair (Gunn and Stark, 2012)(Fig 4.12A). RPA knockdown (Fig 4.22A) resulted in significantly enhanced MMEJ levels compared to the NT controls (2.9 fold; P< 0.0001)(Fig 4.22C), suggesting that human RPA also suppresses MMEJ. POLQ has recently been identified as an MMEJ-promoting factor in mammalian cells (Mateos-Gomez et al., 2015, Ceccaldi et al., 2015). Thus siRNA knockdown of POLQ was used as a control for the GFP-based MMEJ reporter, which led to 94% reduction in *POLQ* gene expression levels, as evaluated by qRT-PCR (Fig 4.22B). This was found to significantly reduce MMEJ repair by 58% (P=0.0112), consistent with its role in the mammalian MMEJ pathway (Fig 4.22C). Further, co-depletion of RPA and POLQ in the GFP-based MMEJ reporter led to significantly reduced levels of MMEJ compared to that of RPA-depleted cells (P=0.0007). Surprisingly, however, MMEJ levels in cells co-depleted for POLQ and RPA were still significantly greater than NT controls (P=0.0022) (Fig 4.22C), suggestive of the possible existence of POLQ-dependent and POLQ-independent MMEJ pathways being suppressed by RPA. Finally, to confirm the role for POLQ in promoting MMEJ in the HPRT assay, I depleted POLQ in this system and monitored the DSB mutational pattern. Analysis of the mutational signature in cells treated with siRNA against POLQ showed a significantly decreased MMEJ levels compared to that of the NT background (Fig 4.23,24)(Ahrabi et al., 2016).

These results resemble our findings in the MMEJ-GFP assay, confirming that the HPRT and the GFP reporter assays measure the same process.

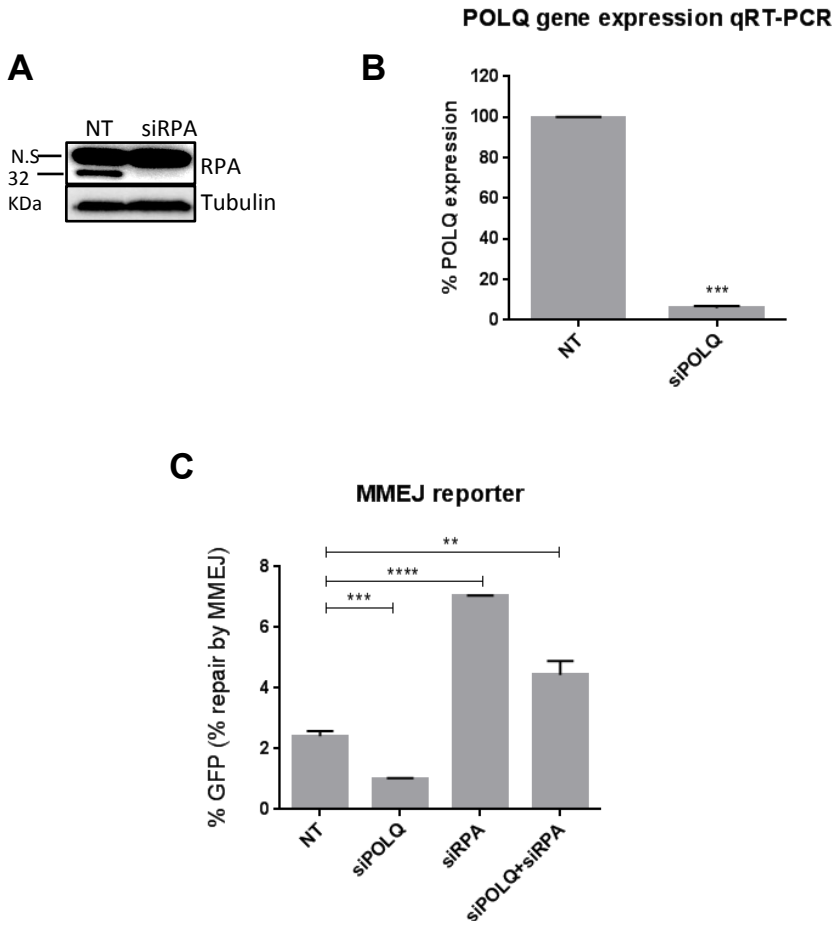


Figure 4.22. RPA suppresses POLQ-dependent and POLQ-independent MMEJ. A. Western blot showing RPA knockdown 48 h following siRNA transfection. B. POLQ gene expression measured by RT-qPCR using the primers indicated in Materials and Methods. C. MMEJ repair efficacy of EJ2-GFP reporter cells treated with NT control siRNA (NT), POLQ siRNA (siPOLQ), RPA siRNA (siRPA) and POLQ plus RPA (siPOLQ+siRPA) indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. **P < 0.01, ***P < 0.001, ****P < 0.0001.

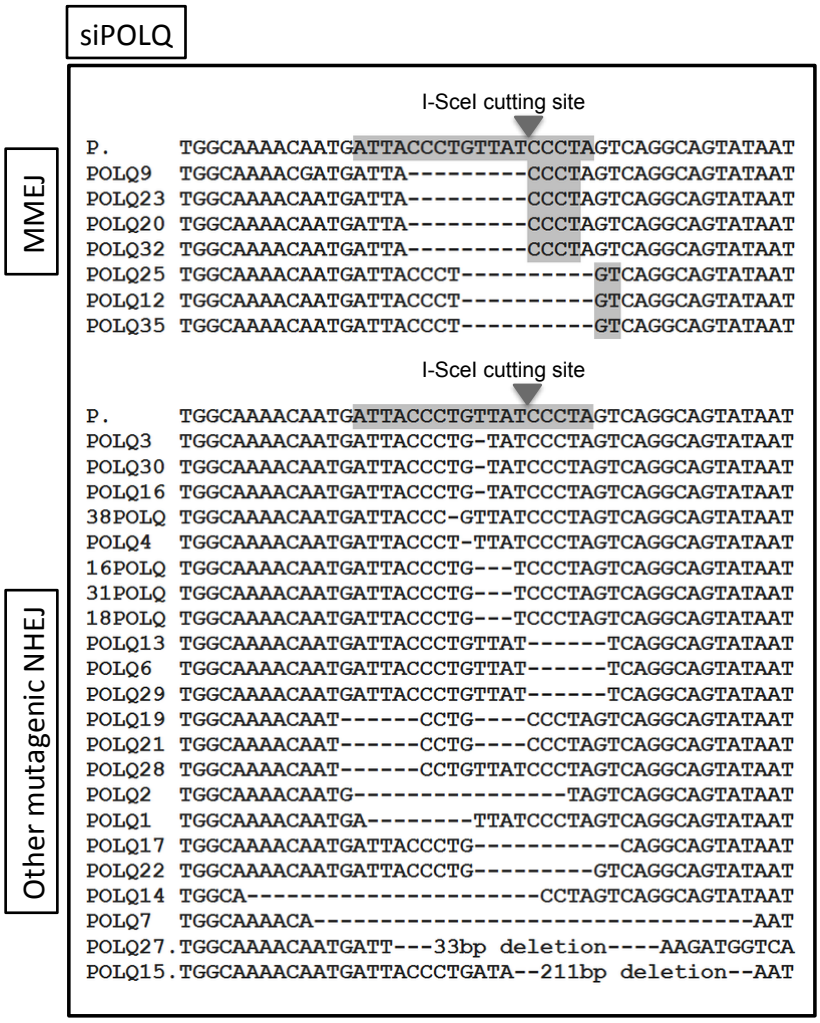


Figure 4.23. Sequence alignments of HPRT negative cells obtained from POLQ-depleted background.

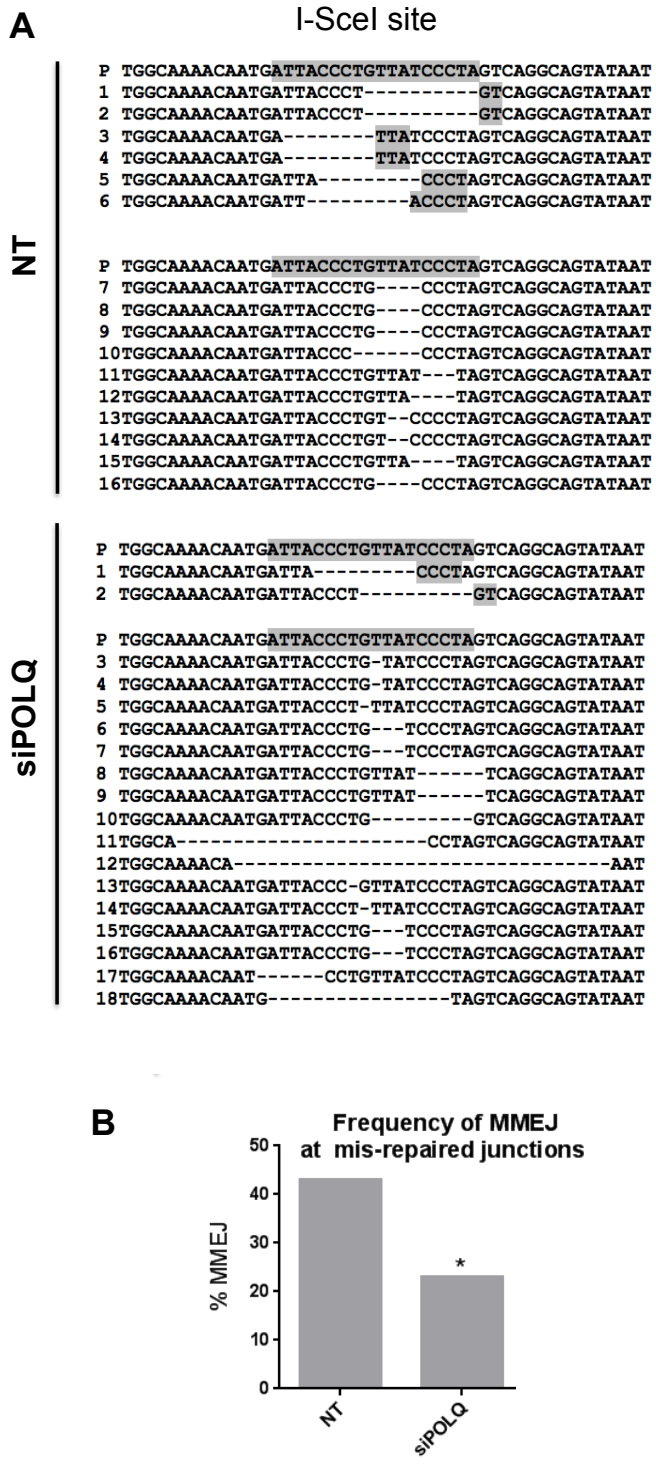


Figure 4.24. A. Representative sequence alignments of the HPRT negative PCR products in cells treated with NT control siRNA (NT), CtIP siRNA (siCtIP), MRE11 siRNA (siMRE11) or PARP-1 inhibitor (Olaparib), from three independent experiments.I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted. B. Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or POLQ siRNA (siPOLQ). P values calculated by statistical analysis “difference between proportions”, *P < 0.05.

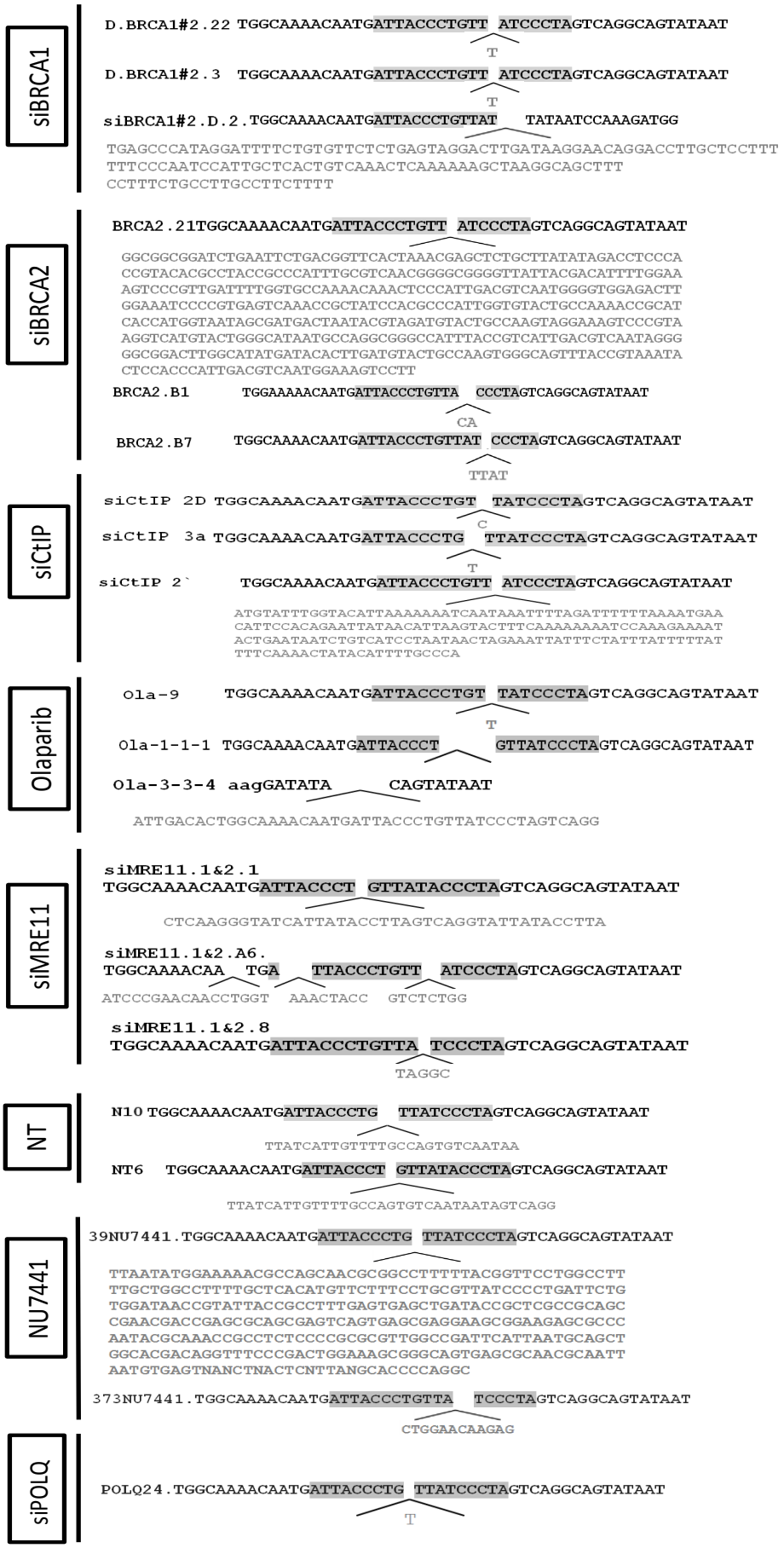
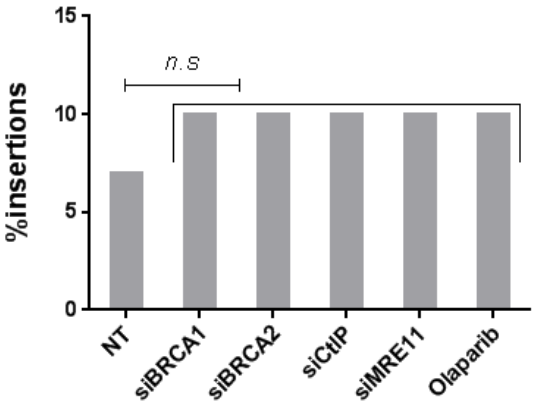


Figure 4.25. Sequence alignments of the inserted sequences at the DSB junctions in different genetic backgrounds.

A



B

Genetic background	Number of inserts	Insert size	Origin
siBRCA1	3/30	138/1/1	Ch 15
siBRCA2	3/30	444/4/2	pCMV6
siCtIP	3/30	180/1/1	12 BAC RP11-377D9
Olaparib	3/30	45/1/1	pBac-IE2-HsCBDBMP4-Puro DNA
siMRE11	3/30	41/33/5	Scaffold PXEA
NT	2/30	27/35	_
NU7441	3/30	336/336/12	_
POLQ	2/30	27/1	_

Figure 4.26. Analysis of breakpoint junctions harboring insertions in different genetic backgrounds. A. Frequencies of insertions at the DSB junctions in different genetic backgrounds. B. Sizes and origins of the inserted sequences at the DSB junctions in cells with different genetic backgrounds.

4.6 Discussion for Chapter 4

4.6.1 A model for HR repressing MMEJ

Mutations arising from aberrant DSB repair are deleterious for cells and can give rise to genomic instability, a hallmark of cancer (Hanahan and Weinberg). Yet how DSB-induced chromosomal rearrangements are suppressed is poorly characterised. Further, there is little mechanistic insight into the relationship between disrupting repair pathways and the mechanisms of DSB mis-repair. Here I have established a role for HR factors in suppressing mutagenic MMEJ following DSB resection. The combined observations in this chapter support a model (Fig 4.27) in which DSB repair in HR-proficient cells, results in a mutational signature with a particular balance between deletions with and without associated microhomologies. When HR is impaired, however, this balance is channeled toward deletions associated with microhomologies as a result of enhanced MMEJ. Mechanistically, I propose that, following DSB induction and initiation of resection, the repair pathway choice can switch from HR to MMEJ, depending on the availability of HR downstream proteins. In the absence of HR factors, HR is blocked, and the resected ends are poor substrates for C-NHEJ. As a result, microhomologous sequences present on ssDNA either side of the resected ends anneal via the mutagenic MMEJ repair mechanism leading to microhomology-mediated deletions.

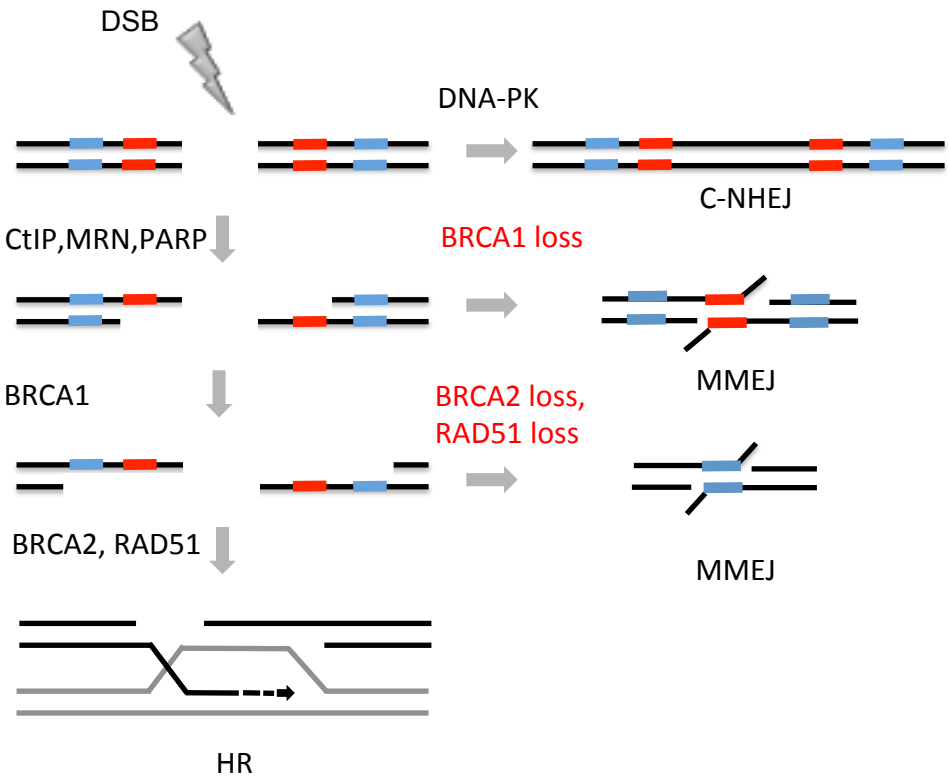


Figure 4.27. Model for the relationships between different DNA DSB repair pathways. DSBs in human cells are mainly repaired by C-NHEJ, however upon initiation of resection by CtIP and MRN complex, the DNA ends are not appropriate substrates for C-NHEJ and repair occurs via HR pathway. If HR is disrupted after initiation of resection, the ends are poor substrates for C-NHEJ and regions of microhomology on either side of the break will anneal, resulting in repair via the mutagenic MMEJ pathway.

4.6.2 Future directions

Our data indicate that while failed HR leads to an increase in MMEJ-induced deletions, surprisingly no effect was observed on insertions, suggesting that these arise through other mutagenic NHEJ events in this context. In this respect, insertions were still observed, albeit at low levels, following inhibition of DNA-PKcs or following knockdown or inhibition of MMEJ factors, suggesting these insertions arise through other end-joining mechanisms. However, this needs to be investigated in more detail. Further, while our data show that MMEJ is enhanced when HR is impaired, the nature of the corresponding decrease in other forms of mutagenic NHEJ is less clear and could indicate inhibition of indel-forming C-NHEJ or other Alt-NHEJ events that are not dependent on junctional microhomologies. Further analyses are required to identify which DSB repair pathways are affected.

Chapter 5

CRISPR-Cas9 and I-SceI -induced DSBs generate distinct repair profiles

In chapters 3 and 4, an I-SceI-cleavable DSB assay (HPRT exon assay) was used to define a novel role for SETD2 and HR mediators in suppressing the mutagenic MMEJ pathway. In order to avoid the necessity of introducing an I-SceI site into the genome and for the universal comparison of DSB repair outcomes in different cell types a CRISPR/Cas9-cleavable DSB assay was developed that targets a common genomic locus with the HPRT assay. By comparing the I-SceI- vs CRISPR/Cas9-induced breaks in the HPRT locus, this chapter describes a new observation: CRISPR/Cas9- and I-SceI-induced DSBs result in distinct repair profiles. This surprising finding suggests that CRISPR/Cas9-induced DSBs are processed differently compared to the I-SceI-induced breaks during repair.

5.1 Targeting the HPRT locus using the CRISPR-Cas9 technology

To develop HPRT-based DSB assays that do not rely on introducing an I-SceI site into the genome, and to compare I-SceI and CRISPR-induced repair pathways, we used three exon 6-specific guide RNAs for the Cas9 nuclease of *Streptococcus pyogenes*: one (gRNA-1) targeting wild-type HPRT alleles and two (gRNA-3 and -4) targeting I-SceI-sensitive HPRT alleles (Fig 5.1). These were designed at the laboratory of Dr. Andy Porter, Imperial College, London.

To test for Cas9-induced indels, wild-type HT1080 cells or the HPRT assay were co-

transfected with expression plasmids for Cas9 and each of the three gRNAs, and HPRT-negative colonies were selected. For a positive control, the I-SceI expression plasmid was transfected in parallel. The results (Fig 5.2) clearly showed the expected gRNA targeting preferences. Thus, high frequencies of HPRT-negative colonies were induced by gRNA-1 in wild-type cells, (5.3% increase compared to that of I-SceI; $p<0.001$) and by gRNA-3 and -4 in clone 5.2.1 cells, (3.3%; $p<0.01$, and 5.8% increase; $p<0.001$) compared to the I-SceI) (Fig 5.3A,B). Conversely, no HPRT-negative colonies were generated in wild-type cells by gRNA-1 or in clone 5.2.1 by gRNA-3 or -4. These results indicate stringent allele-specific exon 6 mutagenesis directed by gRNAs 1, 3 and 4 (Gravells et al., 2015).

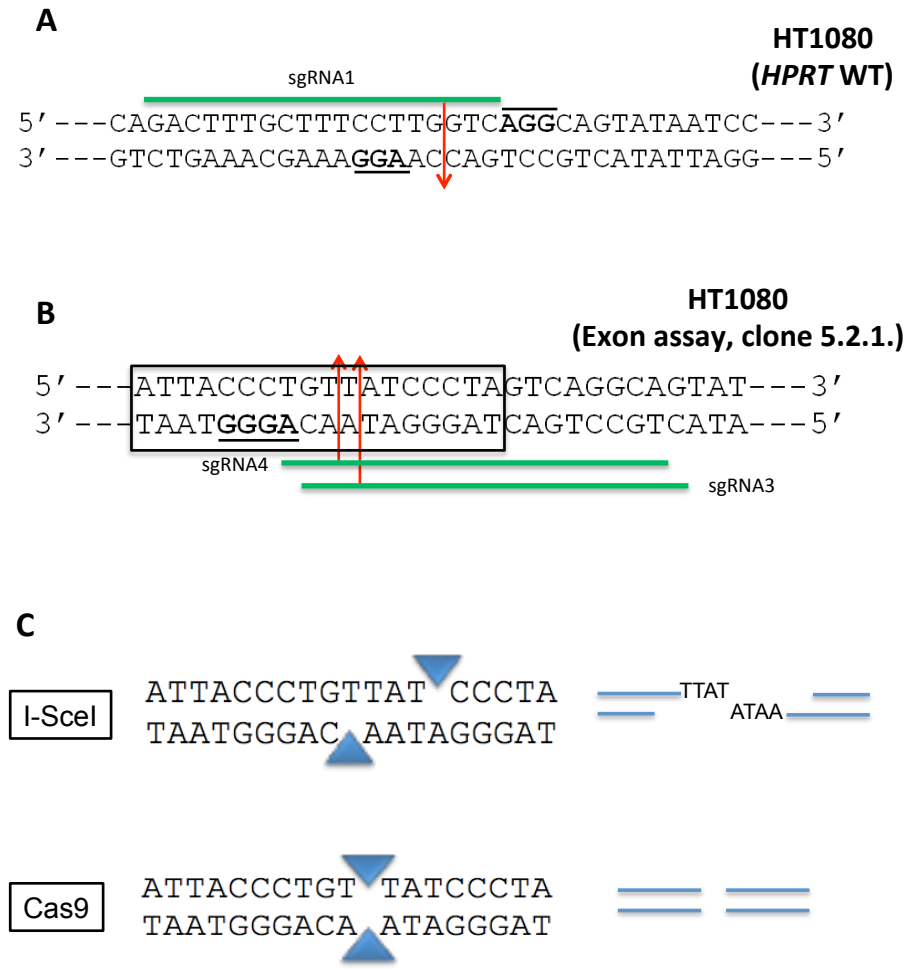


Figure 5.1. CRISPR-Cas9 plasmids targeting the *HPRT* gene. A & B.Design of exon 6-targeted gRNAs. Part of the exon 6 sequence in WT HT1080 cells or clone 5.2.1 is shown. The I-SceI site is boxed with non-WT residues shown in bold and the cleavage site indicated (staggered arrows). Target-specific regions of gRNAs are shown, aligned with their target. For each gRNA, the predicted Cas9 cleavage site is indicated by an arrow, 3 nt upstream of the protospacer-adjacent motif (underlined). C. Schematic illustration of the compatible overhangs generated by I-SceI induced break, and the blunt ends generated by Cas9-induced break.

✓ : Form colonies
× : Negative control
✓: Positive control

A

	gRNA 1	gRNA 3	gRNA 4	I-SceI	GFP
Exon assay (5.2.1)	×	✓	✓	✓	×
HT1080 (WT)	✓	×	×	×	×

B

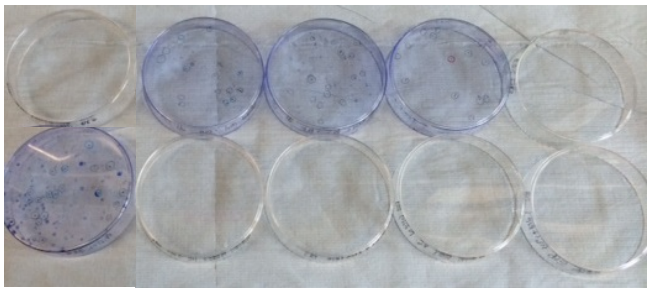


Figure 5.2. Cas9-induced HPRT mutagenesis. A. Table illustrating the colony formation results after transfecting HT1080 cells containing the I-SceI site (Exon assay, clone 5.2.1), and WT HT1080 cells. Cells were co-transfected with expression plasmids for Cas9 and each of the three gRNAs, and HPRT-negative colonies were selected. For a positive control, the I-SceI expression plasmid was transfected in parallel. The results show the expected gRNA targeting preferences. B. Stained petri dishes after selection of transfected HT1080 or clone 5.2.1 cells (10^6 cells per dish) in 6-TG.

5.2. Comparison of CRISPR/Cas9- and I-SceI-induced mutation profiles at the HPRT locus

To gain mechanistic insight into the CRISPR-induced mutational profiles, I then sequenced the exon 6 region in a range of HPRT-negative clones generated following transfecting cells with gRNA1, gRNA3, or gRNA4 alongside the Cas9 construct. I noted a striking difference in the frequency of small insertion mutations generated by I-SceI and Cas9 (Fig 5.4-7). While small insertions represented the least abundant class of I-SceI-generated mutations (1/38; 2.6%), they were the most abundant class of Cas9-generated mutations (20/58; 34.5%) (Fig 5.7A,B). Together these results suggest that CRISPR/Cas9- and I-SceI-generated DSBs are processed differently during DSB repair (Gravells et al., 2015).

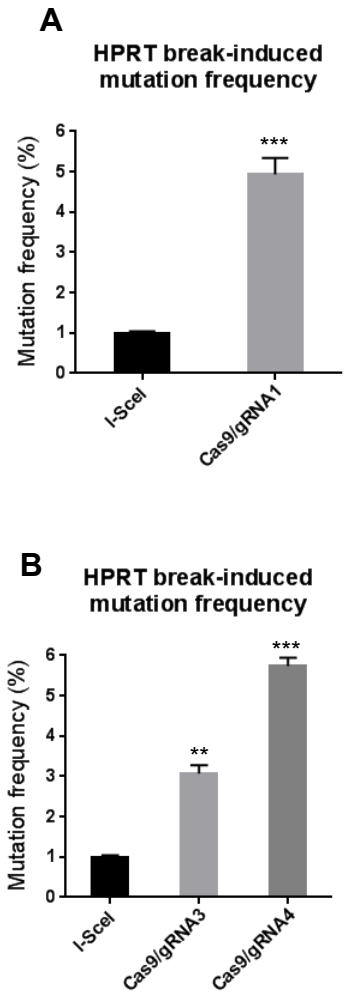


Figure 5.3. HPRT mutation frequencies are significantly higher in CRISPR/Cas9-induced breaks compared to that of the I-SceI-induced breaks. A. HPRT break-induced mutation frequencies following co-transfecting WT HT1080 cells with expression plasmids for Cas9 and gRNA1. Error bars represent SEM from three independent experiments. *** $P < 0.001$. B. HPRT break-induced mutation frequencies following co-transfecting 5.2.1 cells with expression plasmids for Cas9 and gRNA3, and gRNA4 respectively. ** $P < 0.01$.

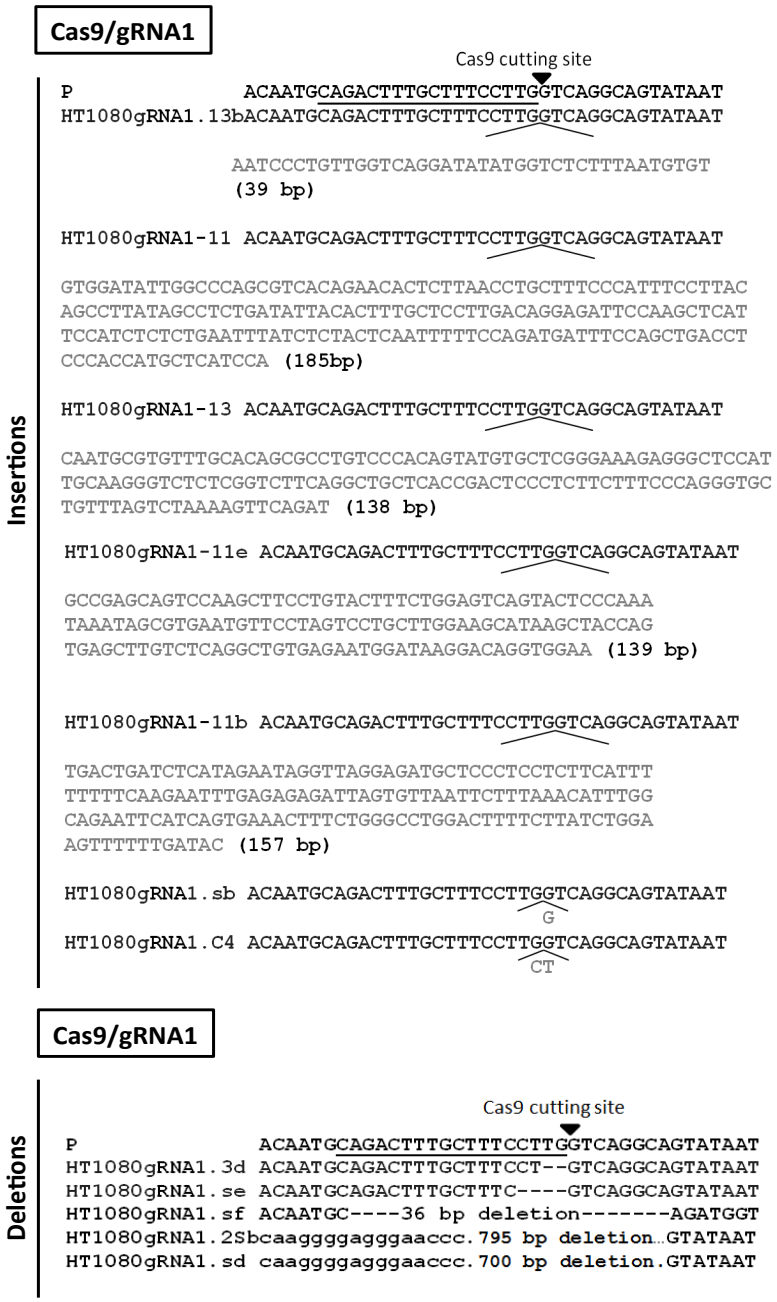


Figure 5.4. Mutation profiles revealed by PCR amplification and sequencing across the Cas9-induced breaks in WT HT1080 cells transfected with gRNA1.

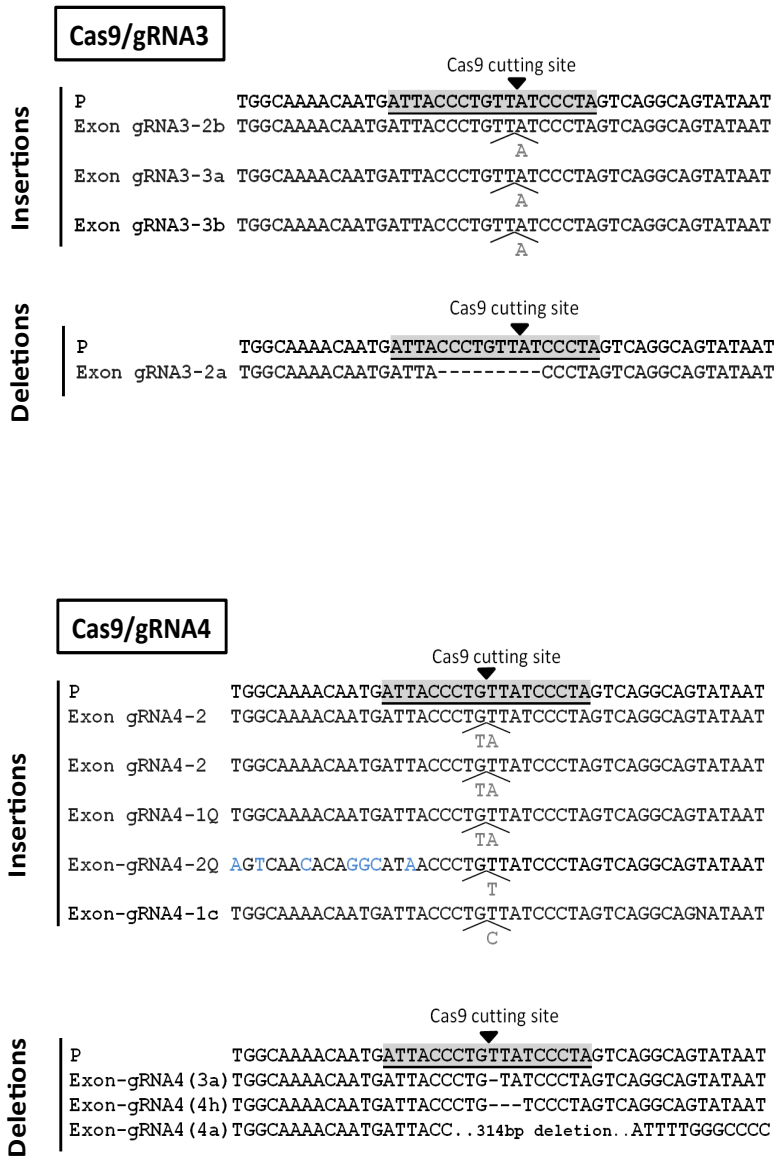


Figure 5.5. Mutation profiles revealed by PCR amplification and sequencing across the Cas9-induced breaks in 5.2.1 cells transfected with gRNA3 & gRNA4 respectively.



Figure 5.6. Mutation profiles revealed by PCR amplification and sequencing across the I-SceI-induced breaks in 5.2.1 cells.

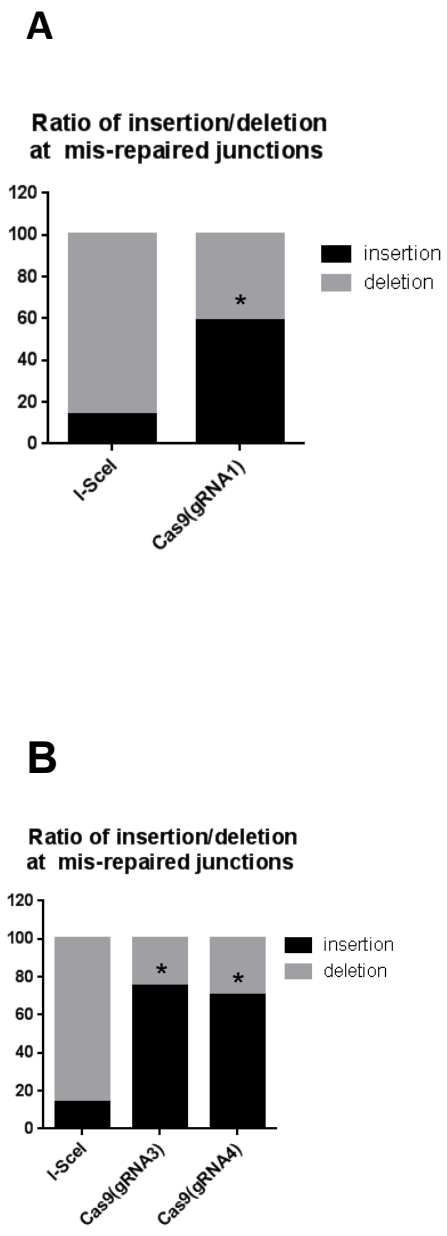


Figure 5.7. A. Relative proportions of different classes of I-SceI- or Cas9-induced mutations (ration of insertion/deletion at mis-repaired junctions) for WT HT1080 cells transfected with gRNA1 construct. B. Relative proportions of different classes of I-SceI- or Cas9-induced mutations (ration of insertion/deletion at mis-repaired junctions) for 5.2.1 cells transfected with gRNA3 & gRNA4 constructs respectively.

5.3 Discussion for chapter 5

The observation that CRISPR/Cas9-induced breaks result in higher mutation frequencies and greater presence of insertions at the break junctions compared to that I-SceI-induced breaks, suggest that the two break types are dealt with differently. These findings can be explained by the fact that DNA ends generated by either of these endonucleases have different characteristics. While the I-SceI-induced breaks generate two complementary cohesive DNA ends (3′ single stranded overhangs of 4 nt), the CRISPR/Cas9-induced breaks generate blunt ends (Fig 5.1C). The complementary cohesive ends can be ligated immediately in an accurate manner, due to the existence of the complementary strand and because no additional end processing is required. Whereas due to the lack of complementary overhangs, the repair of blunt ends is speculated to be inaccurate, thus leading to the randomly insertion of nucleotides at the break junctions in order to ligate the ends.

Chapter 6

Developing a novel I-SceI genetic assay to analyse the mechanisms of DSB mis-repair

The findings described in results chapters 3-5 were obtained by exploiting an exonic HPRT-based DSB mutation assay. In order to specifically identify and quantify the mis-repair events harbouring indels that are greater in length as well as gross chromosomal rearrangements, I generated and characterised the HPRT intron assay.

This chapter describes the generation of a novel assay that can provide mechanistic insights into how DSB repair factors preserve genome integrity by suppressing gross chromosomal rearrangements in response to a DSB induction.

6.1 Introducing the I-SceI site into the intron 5 of HPRT gene

In order to specifically select for the HPRT mutants with larger indels, I generated an intronic reporter assay (HPRT intron assay) using human HT1080 cells. This assay is I-SceI-cleavable and because it harbours an I-SceI site within the intron 5 of its *HPRT* gene, DSB-inducible mis-repair events with indels shorter than 250-300 bp in length are tolerated in this system (Fig 6.1A, 6.6A).

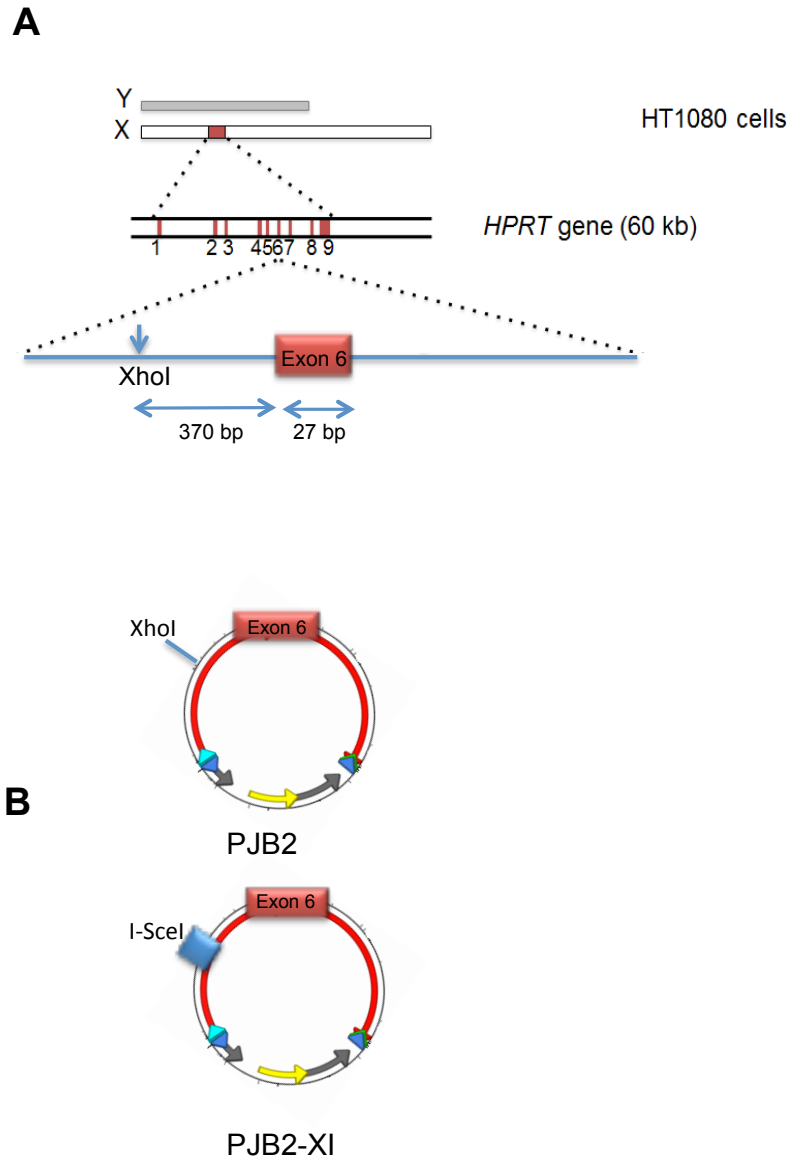


Figure 6.1. A. Schematic map of the human *HPRT* gene, arrows indicate XhoI site. B. Schematic map of the PJB2 plasmid containing human *HPRT* gene (Intron 5, Exon 6, and Intron 6 are shown in red). I-SceI site was substituted for XhoI site by cloning.

Firstly, an 18 bp I-SceI recognition site was cloned into intron 5 of the human *HPRT* gene in pJB2 plasmid (Fig 6.1B). To achieve this, the plasmid was first linearised within the *HPRT* intron 5 (XhoI). Next, annealed oligonucleotides (section 2.15.1, Materials and Methods) encoding the I-SceI site, and with the appropriate complementary 3' overhangs, were inserted into the XhoI restriction site. The resulting plasmid is called pJB2-XI. To identify bacterial colonies containing the desired plasmids, PCR primers were designed (section 2.15.2, Materials and Methods) that amplify from the I-SceI site to a site 1 kb downstream (pJB2-XI) (Fig 6.2A). As shown in figure 6.2A, colonies generating a 1kb PCR product were obtained, suggesting successful cloning. Cloning was confirmed by double digestion of plasmids with I-SceI and SphI (pJB2-XI). As illustrated in figure 6.2B the 600 bp product was obtained, indicative of successful cloning. Finally, sequencing confirmed the insertion of I-SceI site into the intron 5.

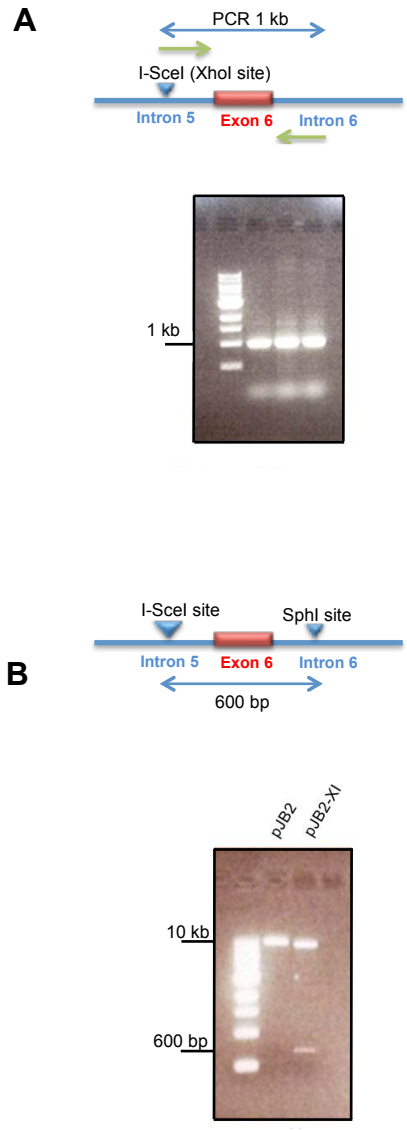


Figure 6.2. Generation of PJB2 plasmids containing ISceI site (PJB2-I-SceI). A.PCR primers shown with green arrows used to amplify across the genomic exon 6, ISceI, and XhoI sites. Agarose gels showing PCR product size. B. I-SceI and SphI enzymes used to digest the PCR products shown in A. Agarose gels showing the PCR product digest sizes.

6.2 Integrating I-SceI site into intron 5 of *HPRT* gene in HT1080 cells

Cloning was followed by linearisation of the constructs that were amplified (by maxiprep approach) (Fig 6.3A) and integration of linearised plasmids into HT1080 cells. The HT1080 cell lines used (clones 10B1 or 3A) had non-functional *HPRT* gene due to a mutation in exon 6. This mutation can be repaired by HR between the *HPRT* gene and the pJB2 derivatives to generate cells with a functional *HPRT* gene that can be selected in HAT medium, as illustrated for pJB2-XI in the schematic of figure 6.3B. Depending on the extent of HR, some of the resulting HAT positive clones should also have taken up the I-SceI site from pJB2-XI into intron 5. Cells were transfected by electroporation or lipofection. Selection for HPRT positive cells was started the day after transfection using HAT supplement and HPRT positive colonies appeared after 10-12 days. The HPRT positive colonies were isolated and expanded for genomic DNA extraction. A total of 6 HPRT positive colonies from electroporation transfection and 22 positive clones from lipofection transfections were isolated. To test for homologous integration of the I-SceI site into the *HPRT* locus, two primer pairs (section 2.15.3, Materials and Methods), each with one oligo outside the homology plasmid arm were used to PCR amplify across the I-SceI sites to generate products of 3kb from intron 5 (used for pJB2-XI transfected clones) (Fig 6.4A,B). These PCR products were then digested with I-SceI to test for the presence of the I-SceI site, which, if present, should generate products of 2.5 and 0.5 kb (from the 3kb product) (Fig 6.4C,D). One HPRT positive clone from the pJB2-XI transfections yielded the expected band size (Fig 6.4D). Therefore, one cell line for use in the *HPRT* intron assay was successfully developed.

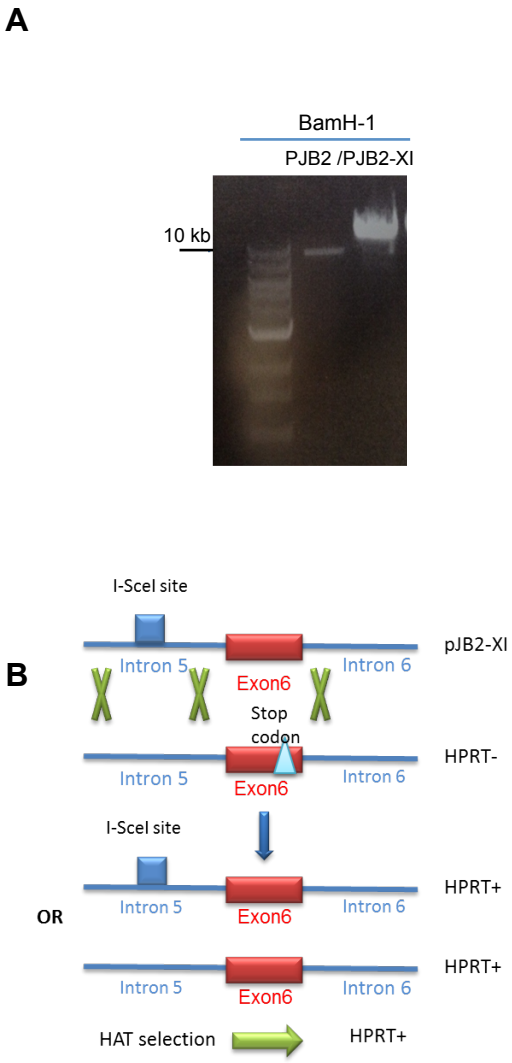


Figure 6.3. Generation of the Intron assay. A. Gel showing the PJB2, and PJB2-XI plasmids amplified by maxiprep and linearised by BamH-1 enzyme. B. Cartoon illustrating HT1080 cells with non-functional *HPRT* gene (HPRT-) transfected with PJB2-XI plasmid. As a result of homologous recombination, HPRT+ cells generated as an outcome of homologous recombination, selected with HAT.

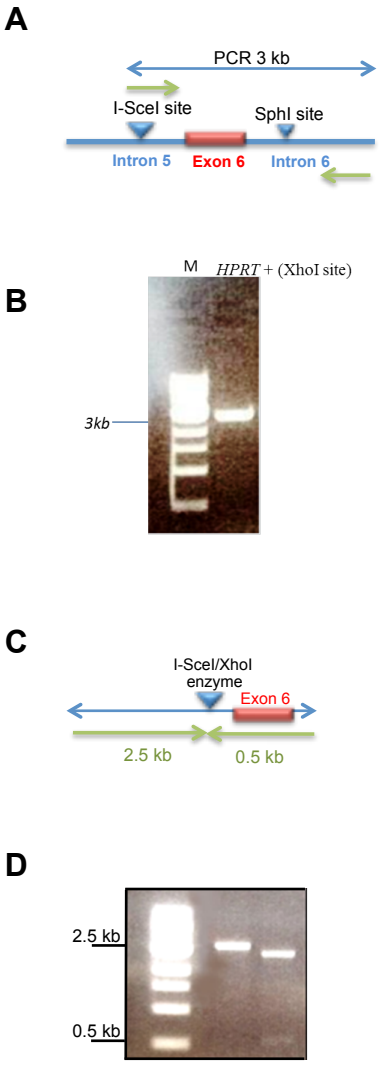


Figure 6.4. Selecting HT1080 cells with an I-SceI site in intron 5. A.PCR primers shown with green arrows, used to amplify across the genomic HPRT gene (XhoI site, exon 6 and the I-SceI site). B. Agarose gels showing PCR product size. C. I-SceI enzyme used to digest the PCR product generated in A. Expected product sizes are shown in green. D. Agarose gels showing the PCR product digest sizes.

6.3 Characterisation of the Intron assay

I-SceI induction in this context led to a significant increase in the mutation frequency, i.e to $5x10^{-4}$ when I-SceI was expressed from spontaneous levels of $5x10^{-5}$ (P=0.0002) (Fig 6.5B). Also, the levels of break-induced mutation were significantly lower (P=0.0002) compared to that of I-SceI-cleavable exon assay, consistent with expected tolerance of this reporter to the smaller indels (Fig 6.6A). This was further confirmed by sequence analysis of HPRT negative clones derived from the intron assay, which confirmed that HPRT loss was associated with significantly greater deletion lengths (P=0.0002) in comparison with that of the exon assay (Fig 6.6B). Moreover, stretches of microhomology sequences were found at the break sites (Fig 6.6C). Altogether these data validate the generation and functionality of the intron reporter assay, which has a significantly lower level of break-induced HPRT loss.

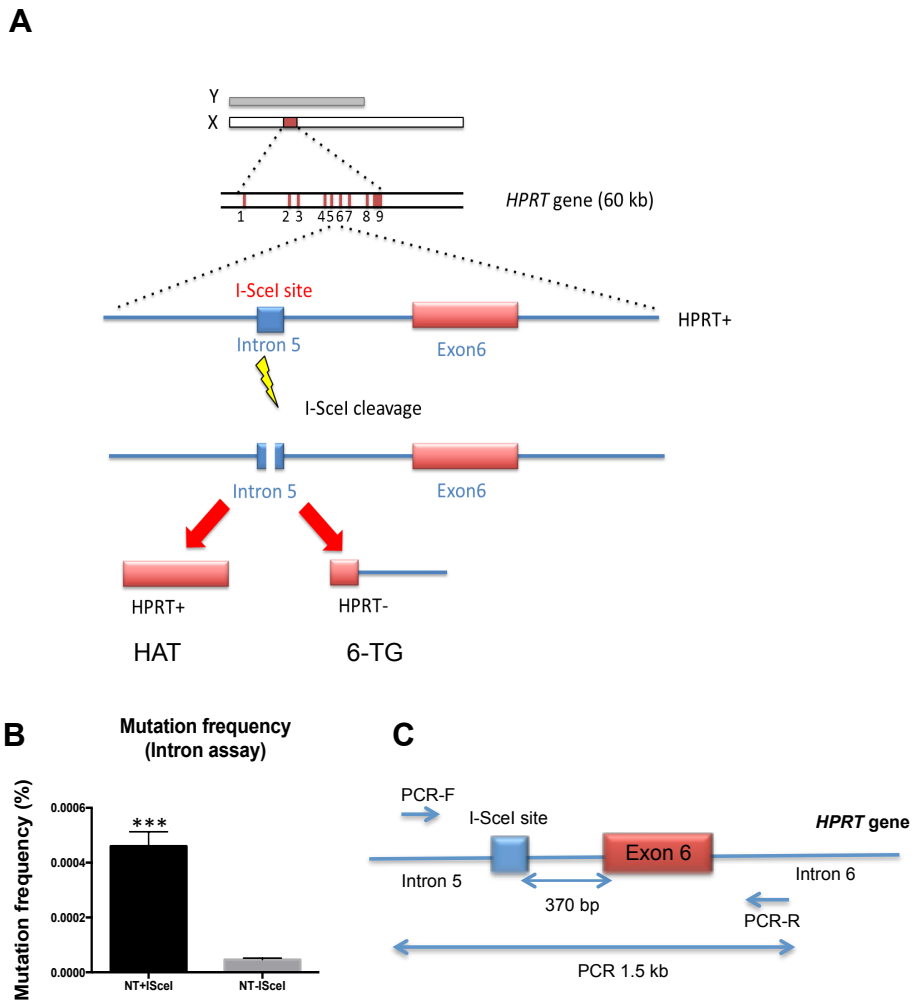
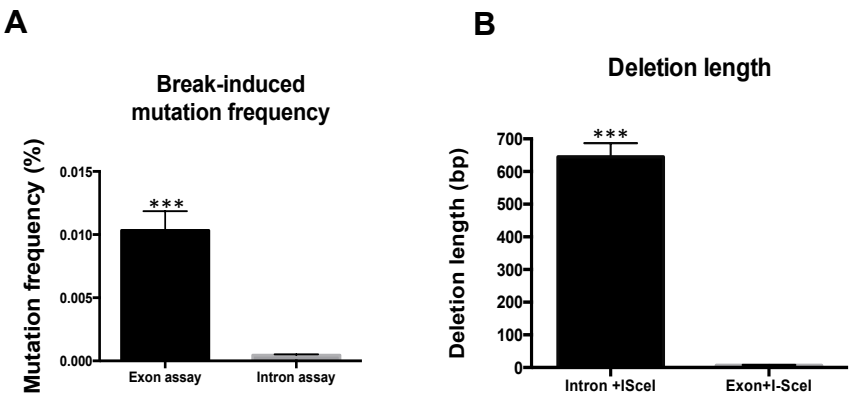


Figure 6.5. Mutation frequency significantly increased following DSB induction. A. Schematic map of the Intron assay, where an I-SceI recognition site is cloned into the intron 5 of *HPRT* gene in human HT1080 cells located on X chromosome. I-SceI induction in this context has the capacity to generate cells with functional or non-functional HPRT gene, which then are selectable using HAT or 6-TG medium respectively. B. HPRT mutation frequencies in the presence or absence of an I-SceI-induced break (P=0.0002). C. Schematic map showing the PCR primers utilised to amplify across the I-SceI site in the Intron assay.



C

```
XhoI22 TAGCGTTATTAATTTGGGTGCCTTTA--778 bp deletion—AAAATTAAGAATAAGGATT
XhoI18 CTGATGGCTCACACTGCTT----788 bp deletions-----TAGACTTAAGATATAATCC
XhoI7  GCAGGAATTTTGAGAGCCATCAG----720 bp deletion-----AAGgtatgtatgacat
XhoI20 TGTTAAAAAGTCTTAAAGTCTTTT--1041 bp deletion---ATCgtttgtcctgaatag
XhoI7  AGGAGGCAGGAATTTTGAGAGCCATCAG----720 bp deletion---AAGgtatgtatga
XhoI10 AATAGTATAAAGCATGTTTCTCG---372 bp deletion—---AATTGACACTGGCAAAAC
XhoI3  TAAAAATAGTATAAAGCATGTTTCTCGA----486 bp deletion----GACACAGAATAT
XhoI2  TTTGGGTGCCTTTAAAAATA---700 bp deletion------------ATTTGGTTTTTC
```

Figure 6.6. A. Break-induced HPRT mutation frequencies in exon and intron assays. B. Average length of deletions (bp) in intron and exon assays. C. Examples of the alignments of HPRT negative clones in the Intron assay. The grey highlighted bits represent the microhomology regions.

6.4 Discussion for chapter 6

The presence of stretches of microhomology associated at increased distances from the DSB site in the context of our I-SceI-cleavable intron assay raises the question as to whether this has arisen through MMEJ or through other pathways. If the former was confirmed it would indicate that MMEJ has the potential to generate very large deletions arising from a single DSB. Also it would be interesting to determine whether this is associated with resection and if the extensive resection factors, such as EXO1, DNA2-BLM play a role. Alternatively other repair pathways such as single strand annealing (SSA), a RAD52 dependent repair pathway that generates deletions between homologous sequences flanking a DSB might be responsible for such large deletions (Lin et al., 1984, Decottignies, 2013). Although MMEJ and SSA are mechanistically similar repair pathways they have distinct genetic requirements. Also the resection length is greater in the latter leading to longer deletions between the flanking microhomologies (McVey and Lee, 2008). Therefore, more in-depth analysis of the intron-derived HPRT mutants is essential to elucidate the underlying mechanisms of the mis-repair events in that context. In this respect depleting factors required for extensive resection, including EXO1, DNA2-BLM, or SSA essential factor, RAD52, in the context of the HPRT intron assay would provide insight into the mechanisms of mis-repair in the HPRT intron assay. This would potentially highlight the importance of alternative repair pathways in response to a single DSB.

Further, the data in chapters 3 and 4 suggested that the deletion length in the HPRT exon assay is associated with the ability of cells to undergo short resection. Thus, the HPRT intron assay can possibly be exploited to identify the new determinants of DSB extensive resection factors.

Chapter 7

Discussion

This chapter discusses the findings in this thesis while comparing those with the current literature. It also highlights the importance of the key findings described in the results chapters while focusing on the following subjects:

7.1. The current understanding of the interplay between different DSB repair pathways.

7.2. Could the “HR-MMEJ” story explain the underlying mechanism of synthetic lethality between HR and PARP inhibition?

7.3. The HPRT system links mutation signatures to the driver mutations

7.4. The HPRT assay; capacities, and limitations.

7.5. Why I-SceI and Cas9 induced breaks generate distinct mutation profiles?

7.6. Relatively low I-SceI-induced mutation frequencies.

7.7. Future directions.

7.1 The current understanding of the interplay between different DSB repair pathways

Mutations arising from aberrant DSB repair are deleterious for cells and can give rise

to genomic instability, the enabling hallmark of cancer (Hanahan and Weinberg). However how DSB-induced chromosomal mutations are suppressed is not fully understood. Further, there is little mechanistic insight into the relationship between disrupting repair pathways and the mechanisms of DSB mis-repair. In this thesis, using a novel genetic approach I have established a role for homologous recombination (HR) factors in suppressing the mutagenic MMEJ following DSB resection. Based on the findings of this study I propose a model in which DSBs are repaired by C-NHEJ that is DNA-PKcs-dependent. However, upon initiation of resection via DNA end-resecting factors the ends are no longer appropriate substrates for C-NHEJ and thus HR will proceed. However, if HR factors are disrupted downstream of resection, HR cannot proceed normally. Also the resected ends are no longer substrates for C-NHEJ. Instead the resected ends anneal towards the terminal microhomology regions and the DNA ends are rejoined via the mutagenic MMEJ. Thus, cells survive a DSB lesion in the expense of inevitable nucleotide loss. Together these findings establish a role for HR in maintaining genome stability through suppression of MMEJ and thus highlight the importance of HR repair genes in maintaining genome stability (Fig 4.27). This study confirms that depleting HR factors (SETD2, RAD51, BRCA2, BRCA1 and RPA) results in significantly elevated levels of a common DSB mutational signature in which DSB-induced deletions are associated with microhomologies of 2–6 bp at the break junctions. These observations are in alignment with the recent studies in human cells indicating that chromosomal rearrangement junctions formed by MMEJ contain 2–6 bp of microhomology regions (Kent et al., 2015).

Because the junctional microhomologies can potentially represent DSB mutational signatures of either C-NHEJ or MMEJ pathways (Gauss and Lieber, 1996, Pannunzio et al., 2014), it is important to know which of these pathways is involved.

However the abrogation of these signatures following depletion or inhibition of CtIP, MRE11, PARP-1, all of which facilitate MMEJ, provides strong genetic evidence that the MMEJ pathway is responsible. The observation that MMEJ levels are decreased following PARP-1 inhibition with Olaparib in the HPRT system is in agreement with a recent study in which a similar finding was found using a GFP-based reporter assay (Howard et al., 2015). Additionally, the persistence of junctional microhomologies in HR-depleted cells treated with DNA-PKcs inhibitor excludes the possibility that C-NHEJ is required. These conclusions are further supported by the significantly greater levels of MMEJ events detected by a GFP-based reporter following depletion of HR factors, even when DNA-PKcs is inhibited (Fig 4.11-12). Our findings therefore indicate that following DSBs HR proteins preserve genomic integrity by suppressing mutations arising through the MMEJ pathway. MMEJ can be considered an important DSB repair pathway in its own right. However our data suggest that its levels can be increased as a result of disrupted HR. In this respect, MMEJ has been proposed to be higher during S/ G2 (Truong et al., 2013), which may arise through incomplete HR. For example, one-ended breaks arising from replication collapse may give rise to both MMEJ and HR.

BRCA1

While the role of CtIP and the MRN complex in resection initiation is well established (Symington, 2014), the exact role of BRCA1 in resection remains uncertain(Nakamura et al., 2010, Reczek et al., 2013, Yun and Hiom, 2009). It has recently been reported that, although the BRCA1-CtIP interaction is not crucial for DNA end-resection, it enhances the speed of CtIP-mediated resection (Cruz-García et al.). It is also known that CtIP and MRE11 are required for both HR and MMEJ repair pathways by triggering resection (Truong et al., 2013). Differing results regarding the contribution of BRCA1 in MMEJ have been reported. Thus, while one

recent study in MEFs found that BRCA1 facilitates MMEJ at uncapped telomeres (Badie et al., 2015), work in DT40 (chicken) B cells suggested that MMEJ is not influenced by BRCA1 (Yun and Hiom, 2009). Various studies of plasmid-based (exogenous) DSB repair substrates in human cells, however, provide evidence that BRCA1 represses Alt-NHEJ events by promoting the fidelity of NHEJ (Jiang et al., 2013, Thompson et al., 2012, Baldeyron et al., 2002, Wang et al., 2006b). In this thesis, using the endogenous HPRT assay together with GFP reporter assays in different human cell lines, I identified a DSB induced mutational signature for BRCA1-depleted cells, which is distinct from that of CtIP-depleted or MRE11-depleted cells. The significantly greater HPRT deletion lengths in BRCA1-depleted cells compared to those of CtIP or MRE11-depleted cells supports the notion that BRCA1 operates downstream of CtIP and MRE11. Further, the significantly smaller HPRT deletion lengths in BRCA1-depleted cells compared to those of BRCA2-depleted cells, suggests that BRCA1 functions downstream of CtIP and MRE11 but upstream of BRCA2 during HR to promote genome stability by facilitating resection and thereby suppressing the mutagenic MMEJ. Moreover, the significantly increased levels of MMEJ, detected by both the HPRT and GFP assays, following BRCA1 depletion suggest that BRCA1 suppresses MMEJ repair. This is in agreement with previous studies using plasmid-based reporter systems (Jiang et al., 2013, Thompson et al., 2012).

SETD2

The *SETD2* gene encoding the H3K36 trimethyltransferase is a tumour suppressor and is mutated in a variety of cancer types (Lawrence et al., 2014). However the mechanism by which SETD2 functions as a tumour suppressor to counteract genome instability observed in cancers with dysfunctional SETD2 is not revealed (Sato et al., 2013, Zhu et al., 2014). SETD2-dependent H3K36 trimethylation has

recently been shown to play a role in DNA mismatch repair (Li et al., 2013). Yet, SETD2-deficient cancers exhibit a wide range of mutations, including indels, and chromosomal aberrations (Sato et al., 2013, Zhu et al., 2014). This suggests an extra role for SETD2 in genome stability. In this thesis using our sensitive HPRT system I have illustrated a common mutation signature following SETD2 or RAD51 loss upon a DSB induction, that is an increased deletion length and enhanced usage of the mutagenic MMEJ pathway. This observation supports a role for SETD2 similar to RAD51 in promoting genome stability in response to a DSB possibly through the error free HR. Our further complementary experiments have confirmed SETD2 as a regulator of HR repair pathway (Pfister et al., 2014).

Also the significantly smaller deletion length in cells depleted for SETD2 compared to that of BRCA1, and BRCA2 suggests a more upstream role for SETD2 in HR compared to the BRCA factors. Complementary experiments carried out in our lab further confirmed a role for SETD2 in HR by facilitating the resection step (Pfister et al., 2014).

POLQ

POLQ has recently been identified as a polymerase that contributes to mammalian MMEJ by promoting DNA end-joining and microhomology annealing (Yousefzadeh et al., 2014, Kent et al., 2015). Further, recent studies have suggested that cancers with defective HR are dependent on a POLQ-mediated MMEJ repair pathway (Mateos-Gomez et al., 2015, Ceccaldi et al., 2015). Moreover, POLQ is up-regulated in a range of cancers, and is indicative of a poor clinical prognosis, especially in breast tumours (Higgins et al., 2010, Kawamura et al., 2004, Lemee et al., 2010). In agreement with these reports, the observations in this thesis support a role for POLQ in facilitating MMEJ in human cancer cells.

RPA

RPA has an established role in repressing MMEJ in *S. cerevisiae* by removing DNA secondary structures, which prevent RAD51 loading and therefore HR (Deng et al., 2014, McVey, 2014). Consistent with this, we demonstrate a striking role for RPA in suppressing mutagenic MMEJ in human cancer cells. This could reflect roles for RPA in early HR, thereby suppressing MMEJ and/ or in binding ssDNA, thereby preventing spontaneous MMEJ, as previously proposed in S. cerevisiae (Deng et al., 2014). The data in this thesis also identify dramatically increased MMEJ levels in cells following simultaneous depletion of POLQ and RPA. This is indicative of a role for RPA in suppressing MMEJ independently of POLQ, which could be explained by the suggested involvement of POLQ in a subset of Alt-NHEJ repair pathways specifically leading to insertions (Koole et al., 2014, Yu and McVey, 2010). The finding that simultaneous depletion of RPA and POLQ resulted in a significant increase in the MMEJ levels (Fig 4.22C) also suggests that there might be POLQ-dependent and POLQ-independent MMEJ pathways in human cancer cells, both of which are suppressed by RPA. Our data do not, however, exclude the possibility that the slight residual activity of POLQ following its siRNA-mediated knockdown, may contribute to the great MMEJ levels in cells co-depleted for RPA and POLQ. In this regard, a more in depth analysis is required.

Model

Our combined observations support a model (Fig 4.27) in which DSB repair in HR-proficient cells, results in a mutational signature with a particular balance between deletions with and without associated microhomologies. When HR is impaired,

however, this balance is channeled toward deletions associated with microhomologies as a result of enhanced MMEJ. Mechanistically, I propose that, following DSB induction and initiation of resection, the repair pathway choice can switch from HR to MMEJ, depending on the availability of HR downstream proteins. In the absence of HR factors, HR is blocked, and the resected ends are poor substrates for C-NHEJ. As a result, microhomologous sequences present on ssDNA either side of the resected ends anneal via the mutagenic MMEJ repair mechanism leading to microhomology-mediated deletions. Also consistent with the findings of this thesis, a recent study using MEFs found increased translocation levels following Rad54 deletion (Soni et al., 2014, Reddy et al., 1997) RAD54 is a key HR factor promoting recombination through interactions with RAD51(Reddy et al., 1997). Although there is no sequence-based evidence for the use of microhomologies in generating translocation junctions, translocations were largely PARP1- and Lig 3-dependent suggesting that they had arisen via the MMEJ pathway (McVey and Lee, 2008, Liang et al., 2008).

Why depletion of MMEJ factors lead to increased HPRT mutation frequencies?

Although it is reasonable to predict that depleting MMEJ factors would result in reduced mutation frequencies, the increased mutation frequencies following depletion of CtIP, MRE11, or PARP could stem from the fact that none of these factors are exclusive to MMEJ pathway. While CtIP and MRE11 are shared between MMEJ and HR, PARP is also communal between MMEJ and BER pathways. Perhaps one way to address the question of whether depletion of MMEJ factors lead to reduced HPRT mutation frequencies, is to knockdown an MMEJ exclusive factor such as POLQ.

7.2. Could the "HR-MMEJ" story explain the underlying mechanism of synthetic lethality between HR and PARP inhibition?

Mutations in DNA DSB repair genes are common features in different human cancer types. Among the DSB repair pathways, MMEJ is necessarily mutagenic and is frequently associated with genomic rearrangements (McVey and Lee, 2008, Ottaviani et al., 2014). Importantly, high-resolution sequencing studies have identified microhomologies as a prevalent mutational signature at rearrangement breakpoints in several cancer types including breast, colorectal and prostate adenocarcinomas (Conrad et al., 2010, Bass et al., 2011). Our observations therefore suggest that the microhomology mutational signature in different types of malignancies may arise from failed HR at a stage after the initiation of resection. This mechanistic understanding of the interplay between HR and MMEJ repair pathways could be exploited to develop therapies for cancer patients deficient in these DSB repair pathways.

In this regard, the synthetic lethal interaction between PARP and HR inhibition (Kennedy and D'Andrea, 2006, Huertas, 2010) could possibly be explained based on the findings in this thesis. Indeed there are different points of view on why the tumours with BRCA1, and BRCA2 mutations are sensitive to PARP inhibition (De Lorenzo et al., 2013). However, on the basis of the model proposed in this study, it is reasonable to speculate that this synthetic lethality could be partially associated with the competitive nature of HR and MMEJ repair pathways. That is, the initiation of resection in S phase cells channels the DSB repair pathway from C-NHEJ to HR (Decottignies, 2013). In this instance if the error-free HR were blocked the resected ends would be biased towards repair via MMEJ. Therefore, it might be predicted that by eliminating the MMEJ repair pathway by using PARP inhibitors in HR mutants,

cells would be unable to repair DSBs, and thereby resulting in cell death or senescence. However, given the fact that the synthetic lethality between PARP and HR inhibition has been exclusively observed upon treating BRCA tumours with PARP inhibitors, and not following siRNA-mediated depletion of PARP, questions the "HR-MMEJ" story as the underlying mechanism of this synthetic lethal interaction. Examining the effect of inhibition/depletion of other MMEJ essential factors in tumours with deficiency in BRCA genes could further test this hypothesis. However, the lack of existence of MMEJ essential factors that are exclusive to this pathway makes the investigation of this hypothesis rather challenging. In this respect, the observation that treatment of cells with Olaparib, a PARP inhibitor, wiped out the terminal microhomology regions while treating cells with NU7441, a DNA-PKcs inhibitor, had no impact on the microhomology regions suggests that the HPRT assay provides a robust tool to advance our understanding of the new MMEJ-essential elements.

7.3. The HPRT system links mutation signatures to the driver mutations

Historically, cancer classification was exclusively based on the organ in which the tumours occur (Munoz and Kurzrock, 2012). Recent advancements in whole-genome sequencing, however, have attributed certain mutation signatures to particular cancer types (Alexandrov et al., 2013). Yet, they lack the power to correlate the driver mutations that give rise to those mutation signatures. In this thesis I have used an "all-in-one" mutation system to determine the effect of single gene disruption on the mutational signature following a DSB. Here I have illustrated that depleting HR factors (SETD2, RAD51, BRCA1, BRCA2) upon a DSB induction gives rise to a distinct mutational signature characterised by significantly increased mutation

frequency, associated with increased deletion length, and enhanced levels of microhomology regions at the break junctions. This is strongly in agreement with a meta-analysis study where out of 20 distinct mutation signatures identified across a wide range of cancers; one signature had a strong correlation with cancers with BRCA1 and BRCA2 mutations. This mutation signature was defined by small deletions of up to 50 bp as well as the presence of microhomology regions at the breakpoints. Interestingly they also found that although cancers with BRCA1 and BRCA2 mutations contributed to a large proportion of this signature, other cancers with no BRCA1 or BRCA2 mutations also exhibited this distinct mutation signature (Lord and Ashworth, 2016). This is in alignment with the findings in this thesis whereby loss of other HR factors, namely RAD51 and SETD2, also gave rise to a mutational signature similar to that of BRCA1 and BRCA2, highlighting the predictive utility of the HPRT system. Thus this "all-in-one" endogenous mutation system allows pinpointing cancer driver genes, as well as providing important insight into how loss of individual genes contributes to somatic mutations and intratumour heterogeneity. In this respect, adopting the "single-cell sequencing" approach that examines the sequence information from individual cells, to the HPRT system might foster the speed and the accuracy by which the mutation profiles are determined.

7.4. The HPRT assay; capacities, and limitations

7.4.1.The unmet need for an "all-in-one" DSB mutation assay

In order to elucidate the underlying mechanism of how DSB mis-repair leads to genome instability in human cells, I aimed to analyze DNA DSB-induced mutations. To achieve this I needed an "all-in-one" DSB reporter that could measure the mis-repair events; while simultaneously reporting for a range of DSB mis-repair

outcomes.

The DNA repair field has benefitted from well established GFP-based DSB reporter assays. These typically report for a specific repair outcome following an I-SceI-induced break. However, they usually fail to simultaneously report for a variety of repair/mis-repair events. Further, the exogenous nature of these reporters and their random integration sites within the genome makes their universal comparisons difficult. Thus I set out to generate “all-in-one” endogenous DSB reporters with the capacity to report for a range of DSB mis-repair events, in a quantitative manner. To achieve this we exploited the endogenous X chromosomal human *HPRT* gene as a selectable marker. Briefly, in our assays the I-SceI recognition site is cloned within the exon 6 or intron 5 of *HPRT* gene. DSB induction in these assays generates cells with functional or non-functional HPRT, depending on the fidelity of repair. Each of these outcomes can then be selected for.

7.4.2. Capacity of the HPRT system

The HPRT system is a sensitive endogenous DSB mutation assay in which unlike the GFP reporter assays the endpoints are not artificially defined. Instead cells can undergo an array of repair/mis-repair outcomes in a non-selective manner. Using 6-TG selective medium, this “all-in-one” DSB mutation assay can simultaneously report for a wide range of mis-repair events (including single point mutaions to indels of hundreds of base pairs) arising from different mutagenic DSB repair pathways. Also due to the endogenous nature of the HPRT assay the mis-repair outcomes in this context are a real reflection of how DSBs are endogenously dealt with in cells.

The HPRT assay provides a clonogenic assay in which the HPRT loss (mutation frequency), which is the percentage of mis-repair events following a DSB, is

quantified. Also the deletion length in the mis-repaired junctions in this context is associated with the capacity of cells to undergo resection upon a DSB induction. Thus the HPRT assay can also serve as a resection assay in which the deletion length is directly correlated with the resection capacity of cells. In this respect, in chapter 4 I have shown that loss of CtIP or MRE11 (DNA end resection factors) gives rise to a significantly reduced deletion length compared to the non-targeting controls, in line with roles for these factors in promoting resection. Whereas, depleting RAD51 or BRCA2 results in a significantly increased deletion length compared to the controls, in alignment with their role in HR downstream of resection.

7.4.3. A limitation of the HPRT system

Despite many advantages of the HPRT assay and its multifaceted characteristics, perhaps the long selection period of the HPRT negative clones in 6-TG-contaning media remains as its main caveat. Importantly this confines the usage of the HPRT assay in high throughput experiments. However substituting the 6-TG selection step with next generation sequencing methods such as HiSeq and MiSeq could possibly improve the speed by which the HPRT negative clones are selected. Consequently this would enhance the applicability of the HPRT system in high throughput experiments.

7.5. Distinct mutation signature following I-SceI vs CRISPR/Cas9 induced breaks

In order to capitalise the use of the HPRT assay across different cell lines we

developed CRISPR/Cas9 constructs that generate endogenous DSBs by targeting DNA at the same locus as the I-SceI. Unexpectedly, we observed significantly increased mutation frequencies in these backgrounds compared to that of the I-SceI-induced breaks. Also interestingly the predominant mutation signature in cells transfected with the CRISPR/Cas9 constructs was the presence of insertions at the break junctions as opposed to the deletions at the break sites in cells transfected with I-SceI constructs.

The common endogenous break site shared between the I-SceI and the Cas9 induced breaks, excludes the possibility that the chromatin context might have had an impact on these distinct outcomes.

Both the higher mutation frequencies upon the Cas9-induced breaks and the highly distinct mutation signatures between the Cas9 and the I-SceI induced breaks could stem from different nature of the breaks generated by either of these nucleases. While the I-SceI-induced breaks produce single stranded overhangs, the Cas9-induced breaks generate blunt ends. The short overhangs generated following an I-SceI induction serve as better resection substrates compared to the blunt ends generated following a Cas9-induced break. Thus the complementary overhangs will either be instantly repaired accurately via the C-NHEJ pathway or will undergo resection in which case either HR or MMEJ are the anticipated outcomes. Whereas the Cas9-induced blunt ends are poor substrates for either accurate repair or resection. Thus, it is expected that the I-SceI-induced complementary overhangs will be predominantly repaired accurately via C-NHEJ or HR (lower mutation frequency) and deletions are the predominant signature in the mutated population of cells. Whereas the Cas9-induced blunt ends are more likely to be repaired inaccurately by random insertion of nucleotides to rejoin the ends due to the lack of complementary template. Therefore, higher mutation frequency and the presence of insertions at the break junctions are anticipated in this population of cells.

7.6. Relatively low I-SceI-induced mutation frequencies

The relatively low break-induced mutation frequencies observed in the HPRT assays raises the question as to if this is due to low I-SceI cutting efficiency/transfection efficiency or highly efficient DSB repair. In this regards, the greater mutation frequencies following a CRISPR/Cas9-induced break compared to that of an I-SceI-induced break (Fig 5.3) suggest the latter. Perhaps the complementary overhangs generated following I-SceI transfections are better substrates for accurate NHEJ repair compared to the blunt ends generated by CRISPR/Cas9 constructs, in which the complementary sequences are absent at DNA ends (Fig 5.1C).

7.7. Future directions

The HPRT assays can be exploited as screening platforms to identify the new factors that promote/suppress genome stability in response to a DSB. Examples of specific screen ideas are listed below:

1. Identification of new suppressors of DSB-induced genome instability
2. Identification of new elements of the mutagenic MMEJ repair pathway
3. Identification of new determinants of DNA-end resection factors

Further, the HPRT assays can be tailored in a number of ways to in order to address specific biological questions. Examples of such modifications are as below:

Recruitment at DNA DSBs

Adopting of ChIP-Seq technique to these assays would allow discovery of protein-DNA interactions such as transcription factor bindings and histone modifications following a DSB at the break site.

Transcription

Transcriptional control of the *HPRT* gene under an inducible promoter would allow comparison of the DSB-induced mutation profiles in the presence and absence of transcription. Using this modification, the study of crosstalk between transcription and DSB repair would be facilitated and the impact of transcriptional factors in response to a DSB on genome stability can be examined.

Chromosomal translocations

Interestingly although PCR amplification across the break junctions in the majority of HPRT- negative clones generated PCR products, a substantial proportion resulted in no PCR products from the target locus, while other loci were amplified normally (Table 3.1). These may reflect simple PCR failure, or perhaps the presence of large indels or chromosomal translocations, leading to loss of one or more PCR primers, in which case using the primers further apart to amplify across a larger genomic region would be enlightening. Interestingly siRNA-mediated knockdown of CtIP, a DNA-end resection factor, led to a striking decrease in the proportion of such PCR-negative clones. Because CtIP is reported to enhance translocations our results suggest that these PCR-negative clones could have undergone translocations upon DSB induction. This is an interesting observation to expand upon which, it would be informative to investigate the role of CtIP and other candidate genes in promoting and suppressing chromosomal translocations. To detect chromosomal rearrangements or transloations the ability of the HPRT assay can be enhanced by addition of fluorescence characteristics. In this regard, fluorescently labeled catalytically dead Cas9 protein assembled with single-guide RNAs could be used to visualise the genomic sequences on each side of the break site with a different colour. Using this modification, translocations would be distinguished from indels by separation of the two colour signals. This approach can subsequently be used to assess a potential role for candidate genes in promoting/suppressing translocations. The translocation break points can then be sequenced using the “inverse PCR”

method by which the amplification of an unknown DNA sequence (the translocated arm) flanked by a known sequence (the HPRT arm) in enabled.

Cell Cycle

DSBs are repaired via distinct pathways in different phases of cell cycle. In this respect, it would be interesting to monitor the DSB-induced mutation profiles following break induction in specific cell cycle phases. This could be achieved by generating a break using I-SceI plasmids that cleave DNA in specific cell cycle phases (Gravells et al., 2015).

The work presented in this thesis exploits the *HPRT* gene to examine the mechanisms of DSB mis-repair in different genetic backgrounds. This has provided a number of insights into the roles of particular genes, and the relationships between DSB repair pathways. As indicated above, the HPRT assay can be further developed to determine the impact of these other factors on DSB repair and genome stability.

Bibliography:

AHRABI, S., SARKAR, S., PFISTER, S. X., PIROVANO, G., HIGGINS, G. S., PORTER, A. C. & HUMPHREY, T. C. 2016. A role for human homologous recombination factors in suppressing microhomology-mediated end joining. *Nucleic Acids Res,* 44**,** 5743-57.

AL SARAKBI, W., SASI, W., JIANG, W., ROBERTS, T., NEWBOLD, R. & MOKBEL, K. 2009. The mRNA expression of SETD2 in human breast cancer: correlation with clinico-pathological parameters. *BMC Cancer,* 9**,** 290.

ALEXANDROV, L. B., NIK-ZAINAL, S., WEDGE, D. C., APARICIO, S. A., BEHJATI, S., BIANKIN, A. V., BIGNELL, G. R., BOLLI, N., BORG, A., BORRESEN-DALE, A. L., BOYAULT, S., BURKHARDT, B., BUTLER, A. P., CALDAS, C., DAVIES, H. R., DESMEDT, C., EILS, R., EYFJORD, J. E., FOEKENS, J. A., GREAVES, M., HOSODA, F., HUTTER, B., ILICIC, T., IMBEAUD, S., IMIELINSKI, M., JAGER, N., JONES, D. T., JONES, D., KNAPPSKOG, S., KOOL, M., LAKHANI, S. R., LOPEZ-OTIN, C., MARTIN, S., MUNSHI, N. C., NAKAMURA, H., NORTHCOTT, P. A., PAJIC, M., PAPAEMMANUIL, E., PARADISO, A., PEARSON, J. V., PUENTE, X. S., RAINE, K., RAMAKRISHNA, M., RICHARDSON, A. L., RICHTER, J., ROSENSTIEL, P., SCHLESNER, M., SCHUMACHER, T. N., SPAN, P. N., TEAGUE, J. W., TOTOKI, Y., TUTT, A. N., VALDES-MAS, R., VAN BUUREN, M. M., VAN 'T VEER, L., VINCENT-SALOMON, A., WADDELL, N., YATES, L. R., ZUCMAN-ROSSI, J., FUTREAL, P. A., MCDERMOTT, U., LICHTER, P., MEYERSON, M., GRIMMOND, S. M., SIEBERT, R., CAMPO, E., SHIBATA, T., PFISTER, S. M., CAMPBELL, P. J. & STRATTON, M. R. 2013. Signatures of mutational processes in human cancer. *Nature,* 500**,** 415-21.

AUDEBERT, M., SALLES, B. & 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**,** 55117-26.

AYMARD, F., BUGLER, B., SCHMIDT, C. K., GUILLOU, E., CARON, P., BRIOIS, S., IACOVONI, J. S., DABURON, V., MILLER, K. M., JACKSON, S. P. & LEGUBE, G. 2014. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol,* 21**,** 366-74.

BADIE, S., CARLOS, A. R., FOLIO, C., OKAMOTO, K., BOUWMAN, P., JONKERS, J. & TARSOUNAS, M. 2015. BRCA1 and CtIP promote alternative non-homologous end-joining at uncapped telomeres. *Embo j,* 34**,** 410-24.

BALDEYRON, C., JACQUEMIN, E., SMITH, J., JACQUEMONT, C., DE OLIVEIRA, I., GAD, S., FEUNTEUN, J., STOPPA-LYONNET, D. & PAPADOPOULO, D. 2002. A single mutated BRCA1 allele leads to impaired fidelity of double strand break end-joining. *Oncogene,* 21**,** 1401-10.

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., TISSENBAUM, H., WEST, S. C., ROSE, A. M. & BOULTON, S. J. 2008. RTEL1 maintains genomic stability by suppressing homologous recombination. *Cell,* 135**,** 261-71.

BASS, A. J., LAWRENCE, M. S., BRACE, L. E., RAMOS, A. H., DRIER, Y., CIBULSKIS, K., SOUGNEZ, C., VOET, D., SAKSENA, G., SIVACHENKO, A., JING, R., PARKIN, M., PUGH, T., VERHAAK, R. G., STRANSKY, N., BOUTIN, A. T., BARRETINA, J., SOLIT, D. B., VAKIANI, E., SHAO, W., MISHINA, Y., WARMUTH, M., JIMENEZ, J., CHIANG, D. Y., SIGNORETTI, S., KAELIN, W. G., SPARDY, N., HAHN, W. C., HOSHIDA, Y., OGINO, S., DEPINHO, R. A., CHIN, L., GARRAWAY, L. A., FUCHS, C. S., BASELGA, J., TABERNERO, J., GABRIEL, S., LANDER, E. S., GETZ, G. & MEYERSON, M. 2011.

Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. *Nat Genet,* 43**,** 964-8.

BECK, C., ROBERT, I., REINA-SAN-MARTIN, B., SCHREIBER, V. & DANTZER, F. 2014. Poly(ADP-ribose) polymerases in double-strand break repair: Focus on PARP1, PARP2 and PARP3. *Experimental Cell Research*.

BENNARDO, N., CHENG, A., HUANG, N. & STARK, J. M. 2008. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet,* 4**,** e1000110.

BHARGAVA, R., ONYANGO, D. O. & STARK, J. M. Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends in Genetics,* 32**,** 566-575.

BOULTON, S. J. & JACKSON, S. P. 1996. Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J,* 15**,** 5093-103.

CAO, L. L., WEI, F., DU, Y., SONG, B., WANG, D., SHEN, C., LU, X., CAO, Z., YANG, Q., GAO, Y., WANG, L., ZHAO, Y., WANG, H., YANG, Y. & ZHU, W. G. 2016. ATM-mediated KDM2A phosphorylation is required for the DNA damage repair. *Oncogene,* 35**,** 301-13.

CARTER, J. & WIEDENHEFT, B. 2015a. CRISPR-RNA Guided Adaptive Immune Systems. *Cell,* 163**,** 260-260.e1.

CARTER, J. & WIEDENHEFT, B. 2015b. SnapShot: CRISPR-RNA-guided adaptive immune systems. *Cell,* 163**,** 260-260.e1.

CARVALHO, S., VITOR, A. C., SRIDHARA, S. C., MARTINS, F. B., RAPOSO, A. C., DESTERRO, J. M., FERREIRA, J. & DE ALMEIDA, S. F. 2014. SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. *Elife,* 3**,** e02482.

CECCALDI, R., LIU, J. C., AMUNUGAMA, R., HAJDU, I., PRIMACK, B., PETALCORIN, M. I., O'CONNOR, K. W., KONSTANTINOPOULOS, P. A., ELLEDGE, S. J., BOULTON, S. J., YUSUFZAI, T. & D'ANDREA, A. D. 2015. Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair. *Nature,* 518**,** 258-62.

CHANG, H. H. Y., PANNUNZIO, N. R., ADACHI, N. & LIEBER, M. R. 2017. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol,* 18**,** 495-506.

CHENG, Q., BARBOULE, N., FRIT, P., GOMEZ, D., BOMBARDE, O., COUDERC, B., REN, G. S., SALLES, B. & CALSOU, P. 2011. Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks. *Nucleic Acids Res,* 39**,** 9605-19.

CICCIA, A. & ELLEDGE, S. J. 2010. The DNA damage response: making it safe to play with knives. *Mol Cell,* 40**,** 179-204.

CONRAD, D. F., BIRD, C., BLACKBURNE, B., LINDSAY, S., MAMANOVA, L., LEE, C., TURNER, D. J. & HURLES, M. E. 2010. Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. *Nat Genet,* 42**,** 385-91.

CRUZ-GARCÍA, A., LÓPEZ-SAAVEDRA, A. & HUERTAS, P. BRCA1 Accelerates CtIP-Mediated DNA-End Resection. *Cell Reports*.

DALGLIESH, G. L., FURGE, K., GREENMAN, C., CHEN, L., BIGNELL, G., BUTLER, A., DAVIES, H., EDKINS, S., HARDY, C., LATIMER, C., TEAGUE, J., ANDREWS, J., BARTHORPE, S., BEARE, D., BUCK, G., CAMPBELL, P. J., FORBES, S., JIA, M., JONES, D., KNOTT, H., KOK, C. Y., LAU, K. W., LEROY, C., LIN, M. L., MCBRIDE, D. J., MADDISON, M., MAGUIRE, S., MCLAY, K., MENZIES, A., MIRONENKO, T., MULDERRIG, L., MUDIE, L., O'MEARA, S., PLEASANCE, E., RAJASINGHAM, A., SHEPHERD, R., SMITH, R., STEBBINGS, L., STEPHENS, P., TANG, G., TARPEY, P. S., TURRELL, K., DYKEMA, K. J., KHOO, S. K., PETILLO, D., WONDERGEM, B., ANEMA, J., KAHNOSKI, R. J., TEH, B. T., STRATTON, M. R. & FUTREAL, P. A. 2010. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature,* 463**,** 360-3.

DAVID, S. S., O'SHEA, V. L. & KUNDU, S. 2007. Base-excision repair of oxidative DNA damage. *Nature,* 447**,** 941-50.

DE LORENZO, S. B., PATEL, A. G., HURLEY, R. M. & KAUFMANN, S. H. 2013. The Elephant and the Blind Men: Making Sense of PARP Inhibitors in Homologous Recombination Deficient Tumor Cells. *Front Oncol,* 3**,** 228.

DECOTTIGNIES, A. 2013. Alternative end-joining mechanisms: a historical perspective. *Front Genet,* 4**,** 48.

DENG, S. K., GIBB, B., DE ALMEIDA, M. J., GREENE, E. C. & SYMINGTON, L. S. 2014. RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nat Struct Mol Biol,* 21**,** 405-12.

DERIANO, L. & ROTH, D. B. 2013. Modernizing the nonhomologous end-joining repertoire: Alternative and classical NHEJ share the stage. *Annual Review of Genetics.*

DOETSCHMAN, T., MAEDA, N. & SMITHIES, O. 1988. Targeted mutation of the Hprt gene in mouse embryonic stem cells. *Proc Natl Acad Sci U S A,* 85**,** 8583-7.

EDMUNDS, J. W., MAHADEVAN, L. C. & CLAYTON, A. L. 2008. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *The EMBO Journal,* 27**,** 406-420.

FAHEY, C. C. & DAVIS, I. J. 2017. SETting the Stage for Cancer Development: SETD2 and the Consequences of Lost Methylation. *Cold Spring Harb Perspect Med,* 7.

FERGUSON, D. O., SEKIGUCHI, J. M., CHANG, S., FRANK, K. M., GAO, Y., DEPINHO, R. A. & ALT, F. W. 2000. The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. *Proc Natl Acad Sci U S A,* 97**,** 6630-3.

FNU, S., WILLIAMSON, E. A., DE HARO, L. P., BRENNEMAN, M., WRAY, J., SHAHEEN, M., RADHAKRISHNAN, K., LEE, S. H., NICKOLOFF, J. A. & HROMAS, R. 2011. Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. *Proc Natl Acad Sci U S A,* 108**,** 540-5.

FONTEBASSO, A. M., SCHWARTZENTRUBER, J., KHUONG-QUANG, D. A., LIU, X. Y., STURM, D., KORSHUNOV, A., JONES, D. T., WITT, H., KOOL, M., ALBRECHT, S., FLEMING, A., HADJADJ, D., BUSCHE, S., LEPAGE, P., MONTPETIT, A., STAFFA, A., GERGES, N., ZAKRZEWSKA, M., ZAKRZEWSKI, K., LIBERSKI, P. P., HAUSER, P., GARAMI, M., KLEKNER, A., BOGNAR, L., ZADEH, G., FAURY, D., PFISTER, S. M., JABADO, N. & MAJEWSKI, J. 2013. Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. *Acta Neuropathol,* 125**,** 659-69.

GAO, Y. G., YANG, H., ZHAO, J., JIANG, Y. J. & HU, H. Y. 2014. Autoinhibitory structure of the WW domain of HYPB/SETD2 regulates its interaction with the proline-rich region of huntingtin. *Structure,* 22**,** 378-86.

GARCIA, B. A., THOMAS, C. E., KELLEHER, N. L. & MIZZEN, C. A. 2008. Tissue-Specific Expression and Posttranslational Modification of Histone H3 Variants. *Journal of proteome research,* 7**,** 4225-4236.

GAUSS, G. H. & LIEBER, M. R. 1996. Mechanistic constraints on diversity in human V(D)J recombination. *Mol Cell Biol,* 16**,** 258-69.

GHEZRAOUI, H., PIGANEAU, M., RENOUF, B., RENAUD, J.-B., SALLMYR, A., RUIS, B., OH, S., TOMKINSON, A. E., HENDRICKSON, ERIC A., GIOVANNANGELI, C., JASIN, M. & BRUNET, E. 2014. Chromosomal Translocations in Human Cells Are Generated by Canonical Nonhomologous End-Joining. *Molecular Cell,* 55**,** 829-842.

GRAVEL, S., CHAPMAN, J. R., MAGILL, C. & JACKSON, S. P. 2008. DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes Dev,* 22**,** 2767-72.

GRAVELLS, P., AHRABI, S., VANGALA, R. K., TOMITA, K., BRASH, J. T., BRUSTLE, L. A., CHUNG, C., HONG, J. M., KALOUDI, A., HUMPHREY, T. C. & PORTER, A. C. 2015. Use of the HPRT gene to study nuclease-induced DNA double-strand break repair. *Hum Mol Genet*.

GUNN, A. & STARK, J. M. 2012. I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. *Methods Mol Biol,* 920**,** 379-91.

HAINCE, J. F., MCDONALD, D., RODRIGUE, A., DERY, U., MASSON, J. Y., HENDZEL, M. J. & POIRIER, G. G. 2008. PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *J Biol Chem,* 283**,** 1197-208.

HAKIMI, A. A., CHEN, Y. B., WREN, J., GONEN, M., ABDEL-WAHAB, O., HEGUY, A., LIU, H., TAKEDA, S., TICKOO, S. K., REUTER, V. E., VOSS, M. H., MOTZER, R. J., COLEMAN, J. A., CHENG, E. H., RUSSO, P. & HSIEH, J. J. 2013. Clinical and pathologic impact of select chromatin-modulating tumor suppressors in clear cell renal cell carcinoma. *Eur Urol,* 63**,** 848-54.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

HIGGINS, G. S., HARRIS, A. L., PREVO, R., HELLEDAY, T., MCKENNA, W. G. & BUFFA, F. M. 2010. Overexpression of POLQ confers a poor prognosis in early breast cancer patients. *Oncotarget,* 1**,** 175-84.

HOCHEGGER, H., DEJSUPHONG, D., FUKUSHIMA, T., MORRISON, C., SONODA, E., SCHREIBER, V., ZHAO, G. Y., SABERI, A., MASUTANI, M., ADACHI, N., KOYAMA, H., DE MURCIA, G. & TAKEDA, S. 2006. Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells. *Embo j,* 25**,** 1305-14.

HORVATH, P. & BARRANGOU, R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science,* 327**,** 167-70.

HOWARD, S. M., YANEZ, D. A. & STARK, J. M. 2015. DNA damage response factors from diverse pathways, including DNA crosslink repair, mediate alternative end joining. *PLoS Genet,* 11**,** e1004943.

HUERTAS, P. 2010. DNA resection in eukaryotes: deciding how to fix the break. *Nat Struct Mol Biol,* 17**,** 11-6.

ILIAKIS, G. 2009. Backup pathways of NHEJ in cells of higher eukaryotes: Cell cycle dependence. *Radiotherapy and Oncology,* 92**,** 310-315.

JACKSON, S. P. & BARTEK, J. 2009. The DNA-damage response in human biology and disease. *Nature,* 461**,** 1071-8.

JASIN, M. 1996. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends in Genetics,* 12**,** 224-228.

JASIN, M. & ROTHSTEIN, R. 2013. Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol,* 5**,** a012740.

JEGGO, P. & LAVIN, M. F. 2009. Cellular radiosensitivity: how much better do we understand it? *Int J Radiat Biol,* 85**,** 1061-81.

JIANG, G., PLO, I., WANG, T., RAHMAN, M., CHO, J. H., YANG, E., LOPEZ, B. S. & XIA, F. 2013. BRCA1-Ku80 Protein Interaction Enhances End-joining Fidelity of Chromosomal Double-strand Breaks in the G(1) Phase of the Cell Cycle. *J Biol Chem,* 288**,** 8966-76.

JINNAH, H. A., DE GREGORIO, L., HARRIS, J. C., NYHAN, W. L. & O'NEILL, J. P. 2000. The spectrum of inherited mutations causing HPRT deficiency: 75 new cases and a review of 196 previously reported cases. *Mutat Res,* 463**,** 309-26.

JIRICNY, J. 2006. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol,* 7**,** 335-46.

JONES, R. E., OH, S., GRIMSTEAD, J. W., ZIMBRIC, J., ROGER, L., HEPPEL, N. H., ASHELFORD, K. E., LIDDIARD, K., HENDRICKSON, E. A. & BAIRD, D. M. 2014. Escape from telomere-driven crisis is DNA ligase III dependent. *Cell Rep,* 8**,** 1063-76.

KABOTYANSKI, E. B., GOMELSKY, L., HAN, J. O., STAMATO, T. D. & ROTH, D. B. 1998. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res,* 26**,** 5333-42.

KARRAN, P. 2006. Thiopurines, DNA damage, DNA repair and therapy-related cancer. *Br Med Bull,* 79-80**,** 153-70.

KAWAMURA, K., BAHAR, R., SEIMIYA, M., CHIYO, M., WADA, A., OKADA, S., HATANO, M., TOKUHISA, T., KIMURA, H., WATANABE, S., HONDA, I., SAKIYAMA, S., TAGAWA, M. & J, O. W. 2004. DNA polymerase theta is preferentially expressed in lymphoid tissues and upregulated in human cancers. *Int J Cancer,* 109**,** 9-16.

KENNEDY, R. D. & D'ANDREA, A. D. 2006. DNA repair pathways in clinical practice: lessons from pediatric cancer susceptibility syndromes. *J Clin Oncol,* 24**,** 3799-808.

KENT, T., CHANDRAMOULY, G., MCDEVITT, S. M., OZDEMIR, A. Y. & POMERANTZ, R. T. 2015. Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat Struct Mol Biol,* 22**,** 230-7.

KIZER, K. O., PHATNANI, H. P., SHIBATA, Y., HALL, H., GREENLEAF, A. L. & STRAHL, B. D. 2005. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol,* 25**,** 3305-16.

KOOLE, W., VAN SCHENDEL, R., KARAMBELAS, A. E., VAN HETEREN, J. T., OKIHARA, K. L. & TIJSTERMAN, M. 2014. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat Commun,* 5**,** 3216.

LAWRENCE, M. S., STOJANOV, P., MERMEL, C. H., ROBINSON, J. T., GARRAWAY, L. A., GOLUB, T. R., MEYERSON, M., GABRIEL, S. B., LANDER, E. S. & GETZ, G. 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature,* 505**,** 495-501.

LEMEE, F., BERGOGLIO, V., FERNANDEZ-VIDAL, A., MACHADO-SILVA, A., PILLAIRE, M. J., BIETH, A., GENTIL, C., BAKER, L., MARTIN, A. L., LEDUC, C., LAM, E., MAGDELEINE, E., FILLERON, T., OUMOUHOU, N., KAINA, B., SEKI, M., GRIMAL, F., LACROIX-TRIKI, M., THOMPSON, A., ROCHE, H., BOURDON, J. C., WOOD, R. D., HOFFMANN, J. S. & CAZAUX, C. 2010. DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. *Proc Natl Acad Sci U S A,* 107**,** 13390-5.

LI, F., MAO, G., TONG, D., HUANG, J., GU, L., YANG, W. & LI, G. M. 2013. The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha. *Cell,* 153**,** 590-600.

LI, M., PHATNANI, H. P., GUAN, Z., SAGE, H., GREENLEAF, A. L. & ZHOU, P. 2005. Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc Natl Acad Sci U S A,* 102**,** 17636-41.

LIANG, F. & JASIN, M. 1996. Ku80-deficient cells exhibit excess degradation of extrachromosomal DNA. *J Biol Chem,* 271**,** 14405-11.

LIANG, L., DENG, L., NGUYEN, S. C., ZHAO, X., MAULION, C. D., SHAO, C. & TISCHFIELD, J. A. 2008. Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Res,* 36**,** 3297-310.

LIEBER, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry.*

LIN, F. L., SPERLE, K. & STERNBERG, N. 1984. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Mol Cell Biol,* 4**,** 1020-34.

LOBRICH, M. & JEGGO, P. 2017. A Process of Resection-Dependent Nonhomologous End Joining Involving the Goddess Artemis. *Trends Biochem Sci*.

LOEB, L. A. & MONNAT, R. J., JR. 2008. DNA polymerases and human disease. *Nat Rev Genet,* 9**,** 594-604.

LORD, C. J. & ASHWORTH, A. 2016. BRCAness revisited. *Nat Rev Cancer,* 16**,** 110-20.

MAHANEY, B. L., MEEK, K. & LEES-MILLER, S. P. 2009. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochem J,* 417**,** 639-50.

MALI, P., ESVELT, K. M. & CHURCH, G. M. 2013. Cas9 as a versatile tool for engineering biology. *Nat Methods,* 10**,** 957-63.

MAR, B. G., CHU, S. H., KAHN, J. D., KRIVSTOV, A. V., KOCHE, R., CASTELLANO, C. A., KOTLIER, J. L., ZON, R. L., MCCONKEY, M. E., CHABON, J., CHAPPELL, R., GRAUMAN, P. V., HSIEH, J. J., ARMSTRONG, S. A. & EBERT, B. L. 2017. SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia. *Blood*.

MATEOS-GOMEZ, P. A., GONG, F., NAIR, N., MILLER, K. M., LAZZERINI-DENCHI, E. & SFEIR, A. 2015. Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature,* 518**,** 254-7.

MCVEY, M. 2014. RPA puts the brakes on MMEJ. *Nature Structural and Molecular Biology,* 21**,** 348-349.

MCVEY, M. & LEE, S. E. 2008. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet,* 24**,** 529-38.

MEEK, K., DANG, V. & LEES-MILLER, S. P. 2008. DNA-PK: the means to justify the ends? *Adv Immunol,* 99**,** 33-58.

MEHTA, A. & HABER, J. E. 2014. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol,* 6**,** a016428.

MOJICA, F. J., DIEZ-VILLASENOR, C., GARCIA-MARTINEZ, J. & ALMENDROS, C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology,* 155**,** 733-40.

MOTYCKA, T. A., BESSHO, T., POST, S. M., SUNG, P. & TOMKINSON, A. E. 2004. Physical and functional interaction between the XPF/ERCC1 endonuclease and hRad52. *J Biol Chem,* 279**,** 13634-9.

MOYNAHAN, M. E., CHIU, J. W., KOLLER, B. H. & JASIN, M. 1999. Brca1 controls homology-directed DNA repair. *Mol Cell,* 4**,** 511-8.

MOYNAHAN, M. E., PIERCE, A. J. & JASIN, M. 2001. BRCA2 Is Required for Homology-Directed Repair of Chromosomal Breaks. *Molecular Cell,* 7**,** 263-272.

MUNOZ, J. & KURZROCK, R. 2012. Targeted therapy in rare cancers--adopting the orphans. *Nat Rev Clin Oncol,* 9**,** 631-42.

NAKAMURA, K., KOGAME, T., OSHIUMI, H., SHINOHARA, A., SUMITOMO, Y., AGAMA, K., POMMIER, Y., TSUTSUI, K. M., TSUTSUI, K., HARTSUIKER, E., OGI, T., TAKEDA, S. & TANIGUCHI, Y. 2010. Collaborative action of Brca1 and CtIP in elimination of covalent modifications from double-strand breaks to facilitate subsequent break repair. *PLoS Genet,* 6**,** e1000828.

NASSIF, N., PENNEY, J., PAL, S., ENGELS, W. R. & GLOOR, G. B. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol,* 14**,** 1613-25.

NEWBOLD, R. F. & MOKBEL, K. 2010. Evidence for a tumour suppressor function of SETD2 in human breast cancer: a new hypothesis. *Anticancer Res,* 30**,** 3309-11.

NIMONKAR, A. V., OZSOY, A. Z., GENSCHEL, J., MODRICH, P. & 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**,** 16906-11.

OGIWARA, H., UI, A., OTSUKA, A., SATOH, H., YOKOMI, I., NAKAJIMA, S., YASUI, A., YOKOTA, J. & KOHNO, T. 2011. Histone acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors. *Oncogene,* 30**,** 2135-46.

OTTAVIANI, D., LECAIN, M. & SHEER, D. 2014. The role of microhomology in genomic structural variation. *Trends Genet,* 30**,** 85-94.

PAI, C. C., DEEGAN, R. S., SUBRAMANIAN, L., GAL, C., SARKAR, S., BLAIKLEY, E. J., WALKER, C., HULME, L., BERNHARD, E., CODLIN, S., BAHLER, J., ALLSHIRE, R.,

WHITEHALL, S. & HUMPHREY, T. C. 2014. A histone H3K36 chromatin switch coordinates DNA double-strand break repair pathway choice. *Nat Commun,* 5**,** 4091.

PANNUNZIO, N. R., LI, S., WATANABE, G. & LIEBER, M. R. 2014. Non-homologous end joining often uses microhomology: implications for alternative end joining. *DNA Repair (Amst),* 17**,** 74-80.

PFISTER, S. X., AHRABI, S., ZALMAS, L. P., SARKAR, S., AYMARD, F., BACHRATI, C. Z., HELLEDAY, T., LEGUBE, G., LA THANGUE, N. B., PORTER, A. C. & HUMPHREY, T. C. 2014. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep,* 7**,** 2006-18.

PIERCE, A. J., JOHNSON, R. D., THOMPSON, L. H. & JASIN, M. 1999. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev,* 13**,** 2633-8.

RECZEK, C. R., SZABOLCS, M., STARK, J. M., LUDWIG, T. & BAER, R. 2013. The interaction between CtIP and BRCA1 is not essential for resection-mediated DNA repair or tumor suppression. *J Cell Biol,* 201**,** 693-707.

REDDY, G., GOLUB, E. I. & RADDING, C. M. 1997. Human Rad52 protein promotes single-strand DNA annealing followed by branch migration. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 377**,** 53-59.

ROOS, W. P., THOMAS, A. D. & KAINA, B. 2016. DNA damage and the balance between survival and death in cancer biology. *Nat Rev Cancer,* 16**,** 20-33.

ROTHKAMM, K., GUNASEKARA, K., WARDA, S. A., KREMPLER, A. & LOBRICH, M. 2008. Radiation-induced HPRT mutations resulting from misrejoined DNA double-strand breaks. *Radiat Res,* 169**,** 639-48.

SANDER, J. D. & JOUNG, J. K. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol,* 32**,** 347-55.

SARTORI, A. A., LUKAS, C., COATES, J., MISTRIK, M., FU, S., BARTEK, J., BAER, R., LUKAS, J. & JACKSON, S. P. 2007. Human CtIP promotes DNA end resection. *Nature,* 450**,** 509-14.

SATO, Y., YOSHIZATO, T., SHIRAISHI, Y., MAEKAWA, S., OKUNO, Y., KAMURA, T., SHIMAMURA, T., SATO-OTSUBO, A., NAGAE, G., SUZUKI, H., NAGATA, Y., YOSHIDA, K., KON, A., SUZUKI, Y., CHIBA, K., TANAKA, H., NIIDA, A., FUJIMOTO, A., TSUNODA, T., MORIKAWA, T., MAEDA, D., KUME, H., SUGANO, S., FUKAYAMA, M., ABURATANI, H., SANADA, M., MIYANO, S., HOMMA, Y. & OGAWA, S. 2013. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat Genet,* 45**,** 860-7.

SEEGMILLER, J. E., ROSENBLOOM, F. M. & KELLEY, W. N. 1967. Enzyme defect associated with a sex-linked human neurological disorder and excessive purine synthesis. *Science,* 155**,** 1682-4.

SEKI, M., MASUTANI, C., YANG, L. W., SCHUFFERT, A., IWAI, S., BAHAR, I. & WOOD, R. D. 2004. High-efficiency bypass of DNA damage by human DNA polymerase Q. *Embo j,* 23**,** 4484-94.

SFEIR, A. & SYMINGTON, L. S. 2015. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends Biochem Sci,* 40**,** 701-14.

SHAH, S. A., ERDMANN, S., MOJICA, F. J. & GARRETT, R. A. 2013. Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol,* 10**,** 891-9.

SIMSEK, D., BRUNET, E., WONG, S. Y., KATYAL, S., GAO, Y., MCKINNON, P. J., LOU, J., ZHANG, L., LI, J., REBAR, E. J., GREGORY, P. D., HOLMES, M. C. & JASIN, M. 2011. DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. *PLoS Genet,* 7**,** e1002080.

SKANDALIS, A. & GLICKMAN, B. W. 1990. Endogenous gene systems for the study of mutational specificity in mammalian cells. *Cancer Cells,* 2**,** 79-83.

SMEENK, G. & VAN ATTIKUM, H. 2013. The chromatin response to DNA breaks: leaving a mark on genome integrity. *Annu Rev Biochem,* 82**,** 55-80.

SONI, A., SIEMANN, M., GRABOS, M., MURMANN, T., PANTELIAS, G. E. & ILIAKIS, G. 2014. Requirement for Parp-1 and DNA ligases 1 or 3 but not of Xrcc1 in chromosomal translocation formation by backup end joining. *Nucleic Acids Res,* 42**,** 6380-92.

STARK, J. M., PIERCE, A. J., OH, J., PASTINK, A. & JASIN, M. 2004. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol Cell Biol,* 24**,** 9305-16.

SUN, X. J., WEI, J., WU, X. Y., HU, M., WANG, L., WANG, H. H., ZHANG, Q. H., CHEN, S. J., HUANG, Q. H. & CHEN, Z. 2005. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J Biol Chem,* 280**,** 35261-71.

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

SYMINGTON, L. S. & GAUTIER, J. 2011. Double-strand break end resection and repair pathway choice. *Annu Rev Genet,* 45**,** 247-71.

TAKEDA, S., NAKAMURA, K., TANIGUCHI, Y. & PAULL, T. T. 2007. Ctp1/CtIP and the MRN complex collaborate in the initial steps of homologous recombination. *Mol Cell,* 28**,** 351-2.

THOMPSON, E. G., FARES, H. & DIXON, K. 2012. BRCA1 requirement for the fidelity of plasmid DNA double-strand break repair in cultured breast epithelial cells. *Environ Mol Mutagen,* 53**,** 32-43.

TORRES, R. J., PRIOR, C., GARCIA, M. G., BELTRAN, L. M. & PUIG, J. G. 2014. HPRT deficiency in Spain: what have we learned in the past 30 years (1984-2013)? *Nucleosides Nucleotides Nucleic Acids,* 33**,** 223-32.

TRUONG, L. N., LI, Y., SHI, L. Z., HWANG, P. Y., HE, J., WANG, H., RAZAVIAN, N., BERNS, M. W. & WU, X. 2013. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc Natl Acad Sci U S A,* 110**,** 7720-5.

WANG, H., ROSIDI, B., PERRAULT, R., WANG, M., ZHANG, L., WINDHOFER, F. & ILIAKIS, G. 2005. DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res,* 65**,** 4020-30.

WANG, H. C., CHOU, W. C., SHIEH, S. Y. & SHEN, C. Y. 2006a. Ataxia telangiectasia mutated and checkpoint kinase 2 regulate BRCA1 to promote the fidelity of DNA end-joining. *Cancer Res,* 66**,** 1391-400.

WANG, M., WU, W., WU, W., ROSIDI, B., ZHANG, L., WANG, H. & ILIAKIS, G. 2006b. PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res,* 34**,** 6170-82.

WEST, S. C. 2003. Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol,* 4**,** 435-45.

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.

WU, L. & HICKSON, I. D. 2003. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature,* 426**,** 870-4.

YOUSEFZADEH, M. J., WYATT, D. W., TAKATA, K., MU, Y., HENSLEY, S. C., TOMIDA, J., BYLUND, G. O., DOUBLIE, S., JOHANSSON, E., RAMSDEN, D. A., MCBRIDE, K. M. & WOOD, R. D. 2014. Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet,* 10**,** e1004654.

YU, A. M. & MCVEY, M. 2010. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res,* 38**,** 5706-17.

YU, X., WU, L. C., BOWCOCK, A. M., ARONHEIM, A. & BAER, R. 1998. The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. *J Biol Chem,* 273**,** 25388-92.

YUN, M. H. & HIOM, K. 2009. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature,* 459**,** 460-3.

ZHANG, J., DING, L., HOLMFELDT, L., WU, G., HEATLEY, S. L., PAYNE-TURNER, D., EASTON, J., CHEN, X., WANG, J., RUSCH, M., LU, C., CHEN, S. C., WEI, L., COLLINS-UNDERWOOD, J. R., MA, J., ROBERTS, K. G., POUNDS, S. B., ULYANOV, A., BECKSFORT, J., GUPTA, P., HUETHER, R., KRIWACKI, R. W., PARKER, M., MCGOLDRICK, D. J., ZHAO, D., ALFORD, D., ESPY, S., BOBBA, K. C., SONG, G., PEI, D., CHENG, C., ROBERTS, S., BARBATO, M. I., CAMPANA, D., COUSTAN-SMITH, E., SHURTLEFF, S. A., RAIMONDI, S. C., KLEPPE, M., COOLS, J., SHIMANO, K. A., HERMISTON, M. L., DOULATOV, S., EPPERT, K., LAURENTI, E., NOTTA, F., DICK, J. E., BASSO, G., HUNGER, S. P., LOH, M. L., DEVIDAS, M., WOOD, B., WINTER, S., DUNSMORE, K. P., FULTON, R. S., FULTON, L. L., HONG, X., HARRIS, C. C., DOOLING, D. J., OCHOA, K., JOHNSON, K. J., OBENAUER, J. C., EVANS, W. E., PUI, C. H., NAEVE, C. W., LEY, T. J., MARDIS, E. R., WILSON, R. K., DOWNING, J. R. & MULLIGHAN, C. G. 2012. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature,* 481**,** 157-63.

ZHAO, Y., THOMAS, H. D., BATEY, M. A., COWELL, I. G., RICHARDSON, C. J., GRIFFIN, R. J., CALVERT, A. H., NEWELL, D. R., SMITH, G. C. & CURTIN, N. J. 2006. Preclinical evaluation of a potent novel DNA-dependent protein kinase inhibitor NU7441. *Cancer Res,* 66**,** 5354-62.

ZHONG, Q., CHEN, C. F., LI, S., CHEN, Y., WANG, C. C., XIAO, J., CHEN, P. L., SHARP, Z. D. & LEE, W. H. 1999. Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science,* 285**,** 747-50.

ZHU, X., HE, F., ZENG, H., LING, S., CHEN, A., WANG, Y., YAN, X., WEI, W., PANG, Y., CHENG, H., HUA, C., ZHANG, Y., YANG, X., LU, X., CAO, L., HAO, L., DONG, L., ZOU, W., WU, J., LI, X., ZHENG, S., YAN, J., ZHOU, J., ZHANG, L., MI, S., WANG, X., ZHANG, L., ZOU, Y., CHEN, Y., GENG, Z., WANG, J., ZHOU, J., LIU, X., WANG, J., YUAN, W., HUANG, G., CHENG, T. & WANG, Q. F. 2014. Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat Genet,* 46**,** 287-93.

ZIMMERMANN, M., LOTTERSBERGER, F., BUONOMO, S. B., SFEIR, A. & DE LANGE, T. 2013. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science,* 339**,** 700-4.

Appendix

The three published papers that were produced as part of my DPhil are included below.

Published online 29 April 2016

Nucleic Acids Research, 2016, Vol. 44, No. 12 ***5743–5757***
doi: 10.1093/nar/gkw326

A role for human homologous recombination factors in suppressing microhomology-mediated end joining

Sara Ahrabi[1], Sovan Sarkar[1], Sophia X. Pfister[1], Giacomo Pirovano[1], Geoff S. Higgins[1], Andrew C.G. Porter[2] and Timothy C. Humphrey[1,*]

[1]CRUK MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Oxford, OX3 7DQ, UK and [2]Gene Targeting Group, Centre for Haematology, Imperial College Faculty of Medicine, London W12 0NN, UK

Received November 14, 2015; Revised April 13, 2016; Accepted April 14, 2016

ABSTRACT

DNA double-strand breaks (DSBs) are toxic lesions, which if improperly repaired can result in cell death or genomic instability. DSB repair is usually facilitated by the classical non-homologous end joining (C-NHEJ), or homologous recombination (HR) pathways. However, a mutagenic alternative NHEJ pathway, microhomology-mediated end joining (MMEJ), can also be deployed. While MMEJ is suppressed by C-NHEJ, the relationship between HR and MMEJ is less clear. Here, we describe a role for HR genes in suppressing MMEJ in human cells. By monitoring DSB mis-repair using a sensitive HPRT assay, we found that depletion of HR proteins, including BRCA2, BRCA1 or RPA, resulted in a distinct mutational signature associated with significant increases in break-induced mutation frequencies, deletion lengths and the annealing of short regions of microhomology (2–6 bp) across the break-site. This signature was dependent on CtIP, MRE11, POLQ and PARP, and thus indicative of MMEJ. In contrast to CtIP or MRE11, depletion of BRCA1 resulted in increased partial resection and MMEJ, thus revealing a functional distinction between these early acting HR factors. Together these findings indicate that HR factors suppress mutagenic MMEJ following DSB resection.

INTRODUCTION

DNA double strand breaks (DSBs) are deleterious lesions that if left unrepaired can lead to cell death, while if mis-repaired can give rise to genomic instability, thus leading to tumorigenesis (1). To survive such lesions and preserve genome integrity, cells possess two main evolutionarily conserved DSB repair mechanisms, namely homologous recombination (HR), and non-homologous end joining (NHEJ) (2). Other repair pathways generally referred to as alternative non-homologous end joining pathways (Alt-NHEJ) (3–5), have been of recent interest. A subset of these repair mechanisms relies on regions of microhomology on either side of the break, which anneal following limited resection in a process called microhomology-mediated end joining (MMEJ) (6–8).

HR is an error-free DSB repair pathway that proceeds through three phases. In mammalian cells the presynaptic phase is triggered by a two-step 5′ to 3′ end resection that produces 3′ single-stranded DNA (ssDNA) overhangs. Resection is initiated by the endonucleolytic activity of the MRE11-RAD50-NBS1 (MRN) complex and the C-terminal binding protein interacting protein (CtIP), which exposes short ssDNA tails (9,10). These become substrates for the extensive resection mediators, Exo1, DNA2 and BLM (11,12). BRCA1 also facilitates the initial resection step of HR (13,14) in conjunction with MRN (15) and CtIP (15,16), where it accelerates the DSB resection rate (17). The exposed ssDNA is initially protected by Replication Protein A (RPA) (18), which is then displaced by RAD51, following its recruitment by BRCA2, to form a nucleoprotein filament (19). The RAD51 nucleofilament promotes strand invasion of the undamaged sister chromatid, which is used as a repair template, resulting in a displacement loop (D-loop). During the synaptic phase of HR, the 3′ end is extended by DNA replication, which can subsequently proceed through a number of sub-pathways. During DSB repair, second end capture and annealing results in double Holliday junction (HJ) formation. In the post-synaptic phase of HR, HJ structures can be resolved with or without crossovers, or dissolved, thus preventing crossovers (20,21). Alternatively, during synthesis-dependent strand annealing (SDSA) (22), the invading and extended strand is expelled from the D-loop to anneal to the second end which, following gap filling and ligation, results in error-free repair (23).

Classical NHEJ (C-NHEJ) is triggered by recognition and protection of DNA ends by the Ku70/Ku80 heterodimer, which forms a ring that encircles duplex DNA.

*To whom correspondence should be addressed. Tel: +44 1865 617327; Fax: +44 1865 617318; Email: timothy.humphrey@oncology.ox.ac.uk

© The Author(s) 2016. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

This protects ends from resection and creates a platform to recruit the DNA-PK catalytic subunit (DNA-PKcs) (24,25). Broken ends are then trimmed by Artemis and ligated by DNA Ligase 4 (Lig 4), X-ray repair cross-complementing protein 4 (XRCC4) complex, and XRCC4-like factor (XLF), depending on the nature of the damage (25–27). Although end-protection by Ku in this pathway minimizes resection, thus promoting error-free end joining, this pathway is widely referred to as error-prone as it ligates the ends in a homology-independent fashion potentially leading to small insertions, and/or deletions (indels) at the DSB sites. From a genome-wide perspective, however, C-NHEJ is not as threatening as alternative NHEJ (Alt-NHEJ) pathways for mammalian genome stability (24) and is even considered as a guardian of genome stability (28).

Alt-NHEJ refers to DSB end joining pathways that are independent of the C-NHEJ factors Ku70/Ku80, DNA-PKcs and DNA Lig4. Unlike C-NHEJ, these pathways are highly mutagenic, always associated with indels and commonly lead to chromosomal rearrangements. Importantly, a sub pathway of Alt-NHEJ events termed microhomology mediated end joining (MMEJ) rejoins the ends by base pairing between microhomologous sequences. MMEJ is mediated via CtIP, MRN complex, Poly [ADP-ribose] polymerases1 (PARP-1) and DNA ligase 3 (Lig 3) (29,30). In mammalian cells, PARP-1, a key component of the MMEJ pathway (31), initially competes with Ku heterodimer for binding to DNA ends resulting in PAR formation (32,33). This in turn promotes MRE11 recruitment and the initiation of resection by MRN complex and CtIP (34). Subsequently, DNA end ligation is mediated through DNA Lig 3 (7,35). Further, recent studies have identified a low fidelity DNA polymerase θ (Polθ also known as POLQ) (36) as an important MMEJ factor in mammalian cells. POLQ is recruited to DSB ends by PARP-1 where it facilitates end joining and microhomology annealing (37). While absent from yeast, the role for POLQ in MMEJ appears to be conserved in flies, worms, mice and humans (38). MMEJ generates deletions and translocations at the break points and thus is highly mutagenic (39). However, there are other types of Alt-NHEJ events that do not rely on such microhomologies. The ligation step in these pathways is facilitated through DNA ligase 1(Lig1) (6).

Impairment of a DSB repair pathway is a common feature in human cancers, rendering their survival highly dependent on secondary DSB repair pathways (40). Therefore, studies of the relationships between different DSB repair machineries are of interest as they can provide insight into how to target these defects for cancer therapy purposes. In this regard, HR has an inhibitory impact on C-NHEJ via initiation of resection (41). Conversely, C-NHEJ suppresses HR and Alt-NHEJ by Ku binding and preventing the resection initiation (42,43). However, the exact nature of the interplay between HR and Alt-NHEJ in human cells still remains elusive. A recent study in MEFs (Mouse Embryonic Fibroblast Cells) found elevated levels of PARP-1-dependent translocations following loss of the HR factor RAD54 (44). Moreover, a study in *Saccharomyces cerevisiae (S. cerevisiae)* showed that RPA impairs MMEJ by removing the secondary structures of ssDNA, which facilitates RAD51 filament assembly leading to HR and thereby suppression of annealing of the terminal microhomologies (45). On the other hand, two recent studies in mammalian cells suggested that POLQ-mediated Alt-NHEJ suppresses HR repair mechanisms (46,47). A role for BRCA1 in promoting NHEJ fidelity has been proposed (48). However, studies in MEFs have also shown that BRCA1 promotes Alt-NHEJ at uncapped telomeres (49).

Here, we have investigated the relationship between HR and MMEJ repair of DSBs. We have used a highly sensitive HPRT-based assay (50,51) together with other GFP-based reporter assays (52–54) to characterize DSB misrepair events observed following depletion or inhibition of HR and NHEJ factors. Following both physical and genetic analyses, we have established a role for HR factors in suppressing MMEJ at break point junctions. Further, our data suggest the existence of POLQ-dependent and POLQ-independent MMEJ pathways, both of which are suppressed by RPA. Last, we define a role for BRCA1 in functioning downstream of CtIP and MRE11 to promote resection, thereby preserving genome integrity by counteracting MMEJ.

MATERIALS AND METHODS

Cell culture

HT1080 (human fibrosarcoma, HPRT assay), U2OS (human osteosarcoma, GFP-based HR reporter (52)), and H1299 (human non-small cell lung carcinoma, GFP-based C-NHEJ reporter (54)) cells were cultured as described previously (51,55). C-NHEJ reporter cells were a kind gift from Atsushi Shibata and Takashi Kohno. MMEJ reporter cells (U2OS with integrated EJ2-GFP reporter) were a kind gift from Jeremy Stark (53) and were grown in high glucose DMEM supplemented with L-glutamine, 10% fetal bovine serum, and 1% Pen/Strep solution (10 000 U/ml penicillin, 10 000 mg/ml streptomycin), including 8 mg/ml plasmocin and 2 mg/ml puromycin.

HAT to select for HPRT gain of function

HAT selection was carried out by adding 1× HAT supplement (Invitrogen) directly to the DMEM medium for 10 days before the I-SceI transfection.

6-TG to select for HPRT loss of function

6-TG selective medium was made by making up 1000 × 6-TG of 15 mg/ml by dissolving the 6-TG powder (sigma) in 1N NaOH and ddH2O. 1000 × 6-TG was added to DMEM and used at a final concentration of 15 mg/ml. 6-TG medium was added 5 days after the I-SceI transfections.

HPRT-based I-SceI-cleavable reporter assay

The assay was based on human fibrosarcoma (HT1080) cells with a functional but I-SceI-cleavable *HPRT* gene (clone 5.2.1), as described previously (50). Use of this assay in conjunction with siRNA knockdowns has been previously described (51). Briefly, an I-SceI cut site was targeted into exon 6 of the endogenous human *HPRT* gene, without disrupting the functionality of the gene. These cells

were seeded in 6-well plates and were immediately transfected with siRNAs using RNAiMAX from Invitrogen according to the manufacturer's instructions. The following day, medium was replaced with fresh DMEM. 48 h after siRNA transfections, cells were lipofected with I-SceI plasmid, which was performed in a ratio of 6 μg DNA: 18 μl Fugene in a total of 300 μl transfection mix in DMEM with no Pen/Strep and fetal bovine serum to transfect each well of a 6-well plate using the Fugene 6 transfection reagent (Promega). 24 h later, medium was replaced with fresh DMEM. Five days after the I-SceI transfection, the cells were seeded in different densities within 100 mm plates (i.e., 10^5, 5×10^4, 10^4, and 10^3 per plate). Cells used to determine the mutation frequencies were then exposed to 6-TG selection, while the cells used for determining plating efficiencies were kept in non-selective media. 6-TG medium was refreshed every 2–3 days. Cells were incubated for 10–12 days to form colonies, which were stained and counted for the purpose for calculating the mutation frequencies. Further, a total of 30 clones from each genetic background were isolated and grown in 48-well plates. These were subsequently expanded into 24-well plates. At full confluency cells for each clone were then trypsinised and collected for extraction of genomic DNA using a Qiagen Flexi gene DNA kit according to the manufacturer's instructions. Genomic DNA was then quantified and used for PCR amplification across the I-SceI break site and sequencing. Primers and PCR conditions to amplify across the I-SceI site have been previously described (51).

Data analysis

Sequence alignment was conducted using DNA Data Bank of Japan (DDBJ) ClustalW program (version 2.1).

Graphical display of results and statistical analysis

For all statistical analysis and graphical display, the program GraphPad Prism (www.graphpad.com) was used.

siRNA transfections

HT1080 or U2OS or H1299 cells were transfected with siRNAs (10 nM final concentration) using RNAiMAX (Invitrogen) according to the manufacturer's instructions. Medium was replaced 24 h after transfection. The sequences of the siRNAs are listed below:

non-targeting (NT) proprietary sequence of supplier (Qiagen),

BRCA2 (Thermo Scientific): GGGAAACACUCAGA UUAAAUU, UUUAAUCUGAGUGUUUCCCUU;

BRCA1 (Dharmacon): AAUGCCAAAGUAGCUAAU GUAUUUU, AAUACAUUAGCUACUUUGGCAUUU U;

CtIP (Dharmacon): GAGGUUAUAUUAAGGAAGA, GGAGCUACCUCUAGUAUCA, GAACAGAAUAGG ACUGAGU, GCACGUUGCCCAAAGAUUC;

MRE11 (Dharmacon): GGAGGUACGUCGUUUC AGA, GGAAAUGAUACGUUUGUAA, CGAAAUGU CACUACUAAGA, GAAAGGCUCUAUCGAAUGU;

RPA (Dharmacon): AACUGGUUGACGAAAGUG GUGUU, CACCACUUUCGUCAACCAGUUUU;

POLQ (Ambion, Lifetechnologies): CCGCUUUUGG AGUCAGUAATT, UUACUGACUCCAAAAGCGGT A.

Western blotting

Whole cell extracts and the protein concentration measurements for western blotting were performed as described previously (51). Equal amounts of protein were separated on NuPAGE Tris-Acetate or Bis-Tris gel (Invitrogen) and transferred to a PVDF membrane (0.45 μm pore size) (Invitrogen). After blocking, the membranes were incubated with appropriate primary antibodies at 4°C overnight followed by secondary antibodies at room temperature for 1 h. Imaging of protein was performed by the BIO-RAD ChemiDoc Imaging sytem.

Primary antibodies used for western blotting are listed below:

BRCA2 (Santa Cruz, sc-28235); BRCA1 (Calbiochem, OP92); RAD51 (Santa Cruz, sc-8349); CtIP (GeneTex 19E8); MRE11 (Abcam, ab33125); RPA32 (Abcam, 2175); Tubulin (Sigma, T5168). All secondary antibodies were purchased from Invitrogen.

Inhibitors

The PARP-1 inhibitor, Olaparib (AZD 2281) was dissolved in DMSO and used at 5 μM final concentration. The DNA-PKcs inhibitor, NU7441 (Axon) was dissolved in DMSO and used at 5, and 15 μM final concentrations.

Quantitative RT-PCR

RNA was extracted using an RNeasy mini kit (Qiagen). A SuperScript® VILOTM (Invitrogen/Life Technologies) cDNA synthesis kit was used to reverse transcribe cDNA from total RNA according to manufacturer's instructions. Quantitative PCR was performed using 7500 Fast Real-Time PCR detection system (Applied Biosystems). Reactions (25 μl each) were prepared in triplicate in a 96-well reaction plate. Each reaction contained 20 ng cDNA, 200 nM of each primer, 10 μl water and 12.5 μl Absolute Blue QPCR SYBR low ROX Mix (Thermo Scientific). DNA levels were normalized to the GAPDH calculated using a 2-ΔΔCt method. QPCR settings were as follows: Initialization at 95°C for 15 min, denaturation at 95°C for 15 s, annealing at 60°C for 30 s, and extension at 72°C for 30 s and repeat for 40 cycles. Primers used for the qRT-PCR are listed below:

qRT-PCR F:5'-AGT CGC ACA CTG CTA CAG GAC GA-3'

qRT-PCR R: 5'-GGC ACA AAG GCA TGT CGC ATG C-3'

GFP-based I-SceI-cleavable reporter assays (HR, C-NHEJ and MMEJ)

To examine the effects of gene loss of function on the HR, C-NHEJ, and MMEJ repair pathways a panel of GFP-based I-SceI-cleavable reporter cell lines was used. The cells were transfected with an I-SceI plasmid using

Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions in 6-well plates 48 h after the siRNA knockdowns. After 2 days, the cells from each well were trypsinised and dispersed in 200 μl medium. The 400 μl suspension was then immediately mixed with 200 μl of 10% formaldehyde, and vortexed for 2–3 s before FACS (FACSCalibur) to quantitate the number of GFP positive cells. For FACS analysis, non-overlapping gates were defined by three populations of cells from the same cell line: normal cells, cells transfected with EGFP or DsRed plasmids. 20 000 cells were analyzed for each sample.

RESULTS

BRCA2 depletion promotes DSB-induced deletions typical of MMEJ

We have previously shown that siRNA-mediated depletions of SETD2 or RAD51 had similar effects on I-SceI-induced DSBs in the *HPRT* gene, namely an increased overall frequency of *HPRT* mutagenesis associated with increases in both deletion lengths and use of microhomologies on either side of the deletions (51). These findings suggested a possible role for HR in suppressing MMEJ.

To test this further, we applied the same approach (Supplementary Figure S1) to assess a possible role for other HR factors in suppressing MMEJ. BRCA2, a key recombination mediator in mammalian cells, facilitates RAD51 loading onto ssDNA. We therefore anticipated that BRCA2 might also suppress MMEJ. We first depleted BRCA2 and measured the frequency of *HPRT* inactivation (mutation frequency) following I-SceI-induced DSBs. Knockdown of BRCA2 resulted in a significantly increased I-SceI-induced mutation frequency (2.8%; $P < 0.0001$), compared to cells treated with non-targeting control (NT) siRNA (1.1%) (Figure 1A). Sequence analysis of HPRT-negative clones derived from cells treated with NT siRNA (Figure 1B, and Supplementary Figure S2A–C) or BRCA2 siRNA (Figure 1B, and Supplementary Figure S3) indicated average deletion lengths of 5.6 and 49 bp, respectively. Thus BRCA2 depletion resulted in a 9-fold increase in deletion lengths ($P < 0.0001$) (Figure 1C). Insertions at the HPRT break site were also detected but their frequencies were not significantly altered by BRCA2-depletion (Supplementary Figure S14).

Further examination of the sequences at deletion junctions showed that the proportion of deletions appearing to have arisen by MMEJ rose significantly from 43%, for cells treated with NT control siRNA, to 72% for cells treated with BRCA2 siRNA ($P < 0.05$) (Figure 1D). BRCA2 depletion in HT1080 cells was confirmed by western blot (Figure 1E). To assess the effect of BRCA2 depletion on HR, we used a previously described (51) U2OS cell line carrying a well-characterized GFP-based reporter for HR (DR-GFP) (52). We found that the frequency of HR in BRCA2-depleted cells was significantly reduced by 41% compared to the NT controls ($P < 0.0001$), consistent with a role for BRCA2 in promoting HR (19) (Figure 1F). Thus BRCA2 depletion causes both a decrease in HR and an apparent increase in MMEJ suggesting roles for BRCA2 not only in promoting HR but also in suppressing MMEJ.

BRCA1 depletion promotes DSB-induced deletions typical of MMEJ

BRCA1 is also an essential factor for HR but has a distinct function in HR compared to BRCA2, where BRCA1 facilitates 5′ to 3′ resection of DSBs to generate 3′ ssDNA tails (56). To investigate a possible role for BRCA1 in suppressing MMEJ, we examined the effect of BRCA1 depletion on DSB mutational signatures in the HPRT reporter system. Like BRCA2 depletion, BRCA1 depletion resulted in a significantly enhanced mutation frequency (3.5%; $P = 0.001$) following I-SceI-induced DSBs compared to that of NT controls (1.1%) (Figure 2A). Further, sequencing of individually isolated HPRT mutated clones from a BRCA1-depleted background (Figure 2B and Supplementary Figure S4) indicated an average deletion length of 17 bp, significantly (3-fold) greater than the deletion lengths in NT control siRNA-treated cells ($P = 0.04$) (Figure 2C). Interestingly, the deletion lengths in BRCA1-depleted cells were significantly (2.8-fold) smaller than to those in BRCA2-depleted cells ($P < 0.05$). This is consistent with the early role for BRCA1 in promoting resection during HR, prior to BRCA2-assisted RAD51 loading. Insertions at the HPRT break site were also detected but their frequency was not significantly altered by BRCA1-depletion (Supplementary Figure S14). Among the sequenced junctions in HPRT negative cells, 66% ($P < 0.05$) appeared to have been generated by MMEJ following BRCA1 depletion, a 1.5-fold increase over the 43% in cells with normal BRCA1 levels (Figure 2D). Therefore, BRCA1 depletion, like BRCA2-depetion, resulted in a significant increase in the proportion of mutagenic NHEJ events appearing to occur by MMEJ (Figure 2D). BRCA1 depletion in HT1080 cells was confirmed by western blot (Figure 2E). Consistent with the role of BRCA1 in HR (14), HR was significantly reduced by 52% ($P < 0.0001$) in the DR-GFP reporter (U2OS cells) (52) following BRCA1 knockdown (Figure 2F). Thus, BRCA1 depletion results in reduced HR repair and significantly increased mutation frequency associated with a larger proportion of MMEJ events. Together with the similar effects of BRCA2, RAD51 and SETD2 depletion, these results suggest a common role for HR factors in suppressing MMEJ.

MMEJ-like DSB repair in HR deficient cells is independent of DNA-PKcs

Although the microhomologies we observed at repair junctions in HR-defective cells were highly suggestive that MMEJ is the key pathway involved, we wanted to test the possible involvement of C-NHEJ. Because C-NHEJ, but not MMEJ, is dependent on DNA-PKcs, we therefore, compared the HPRT mutational profile following treatment of cells with a DNA-PKcs inhibitor (NU7441), or siRNA against RAD51 in the presence and absence of NU7441. This inhibitor has previously been shown to disrupt the C-NHEJ repair pathway (57). We found that treating cells with NU7441 resulted in a significant increase in the deletion length (16 bp, $P = 0.015$) and a significantly increased presence of microhomologies at the break sites (60%, $P < 0.05$) compared to that of the NT background (Figure 3A–C and Supplementary Figure S5). This is consistent with a role for DNA-PKcs in conjunction with Ku in protecting DSB ends

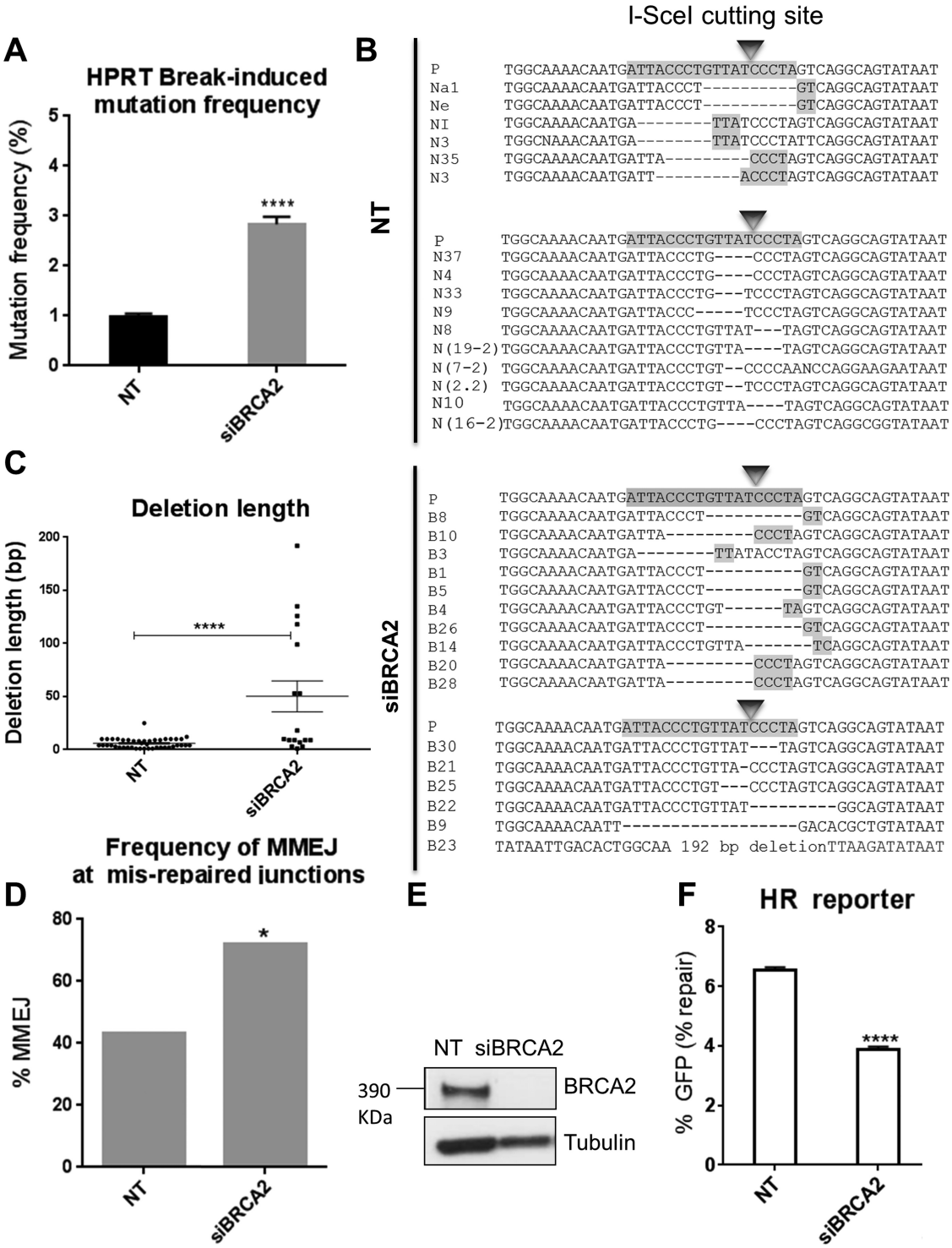


Figure 1. BRCA2 depletion promotes DSB-induced deletions typical of MMEJ. (**A**) Break-induced mutation frequency of HPRT reporter cells treated with NT control siRNA (NT), and BRCA2 siRNA (siBRCA2). Error bars represent SEM from three independent experiments. ****$P < 0.0001$. (**B**) Representative sequence alignments of the PCR products in NT control (NT), and BRCA2-depleted (siBRCA2) cells isolated from three independent experiments (see also Supplementary Figures S1A, S2A–C and S3). I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted. (**C**) Average length of deletions (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, **** $P < 0.0001$. (**D**) Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or BRCA2 siRNA (siBRCA2). P values calculated by statistical analysis "difference between proportions", * $P < 0.05$. (E) Western blot showing BRCA2 knockdown in HT1080 cells 48 h following siRNA transfection. (F) HR repair efficacy of DR-GFP reporter cells treated with NT control siRNA (NT), and BRCA2 siRNA (siBRCA2), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments, **** $P < 0.0001$.

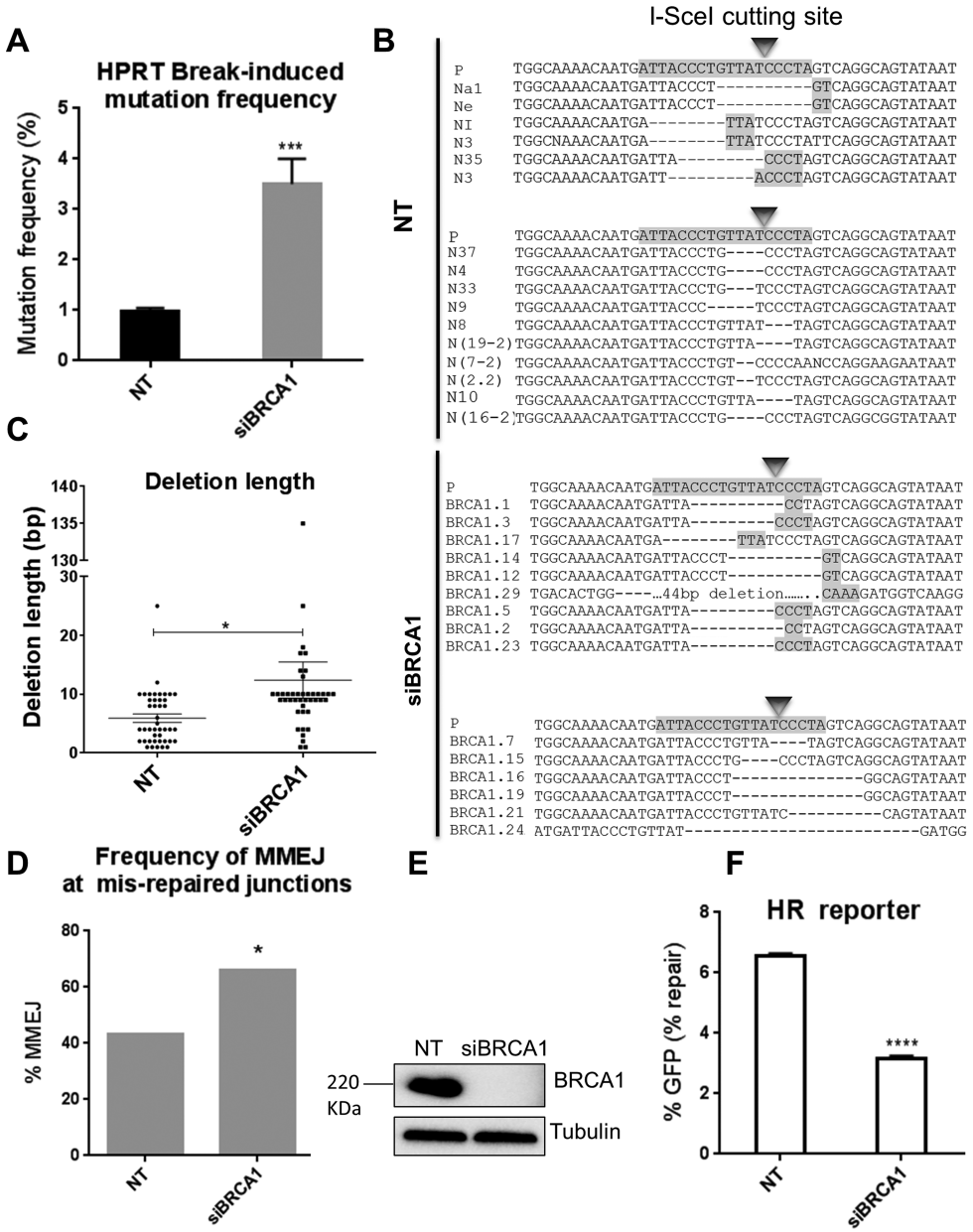


Figure 2. BRCA1 depletion promotes DSB-induced deletions typical of MMEJ. (**A**) Break-induced mutation frequency of HPRT reporter cells treated with NT control siRNA (NT), BRCA1 siRNA (siBRCA1). Error bars represent SEM from three independent experiments. *** $P < 0.001$. (**B**) Representative sequence alignments of the PCR products in NT control (NT), and BRCA1-depleted (siBRCA1) cells from three independent experiments (see also Supplementary Figures S1A, and S4). I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted. (**C**) Average length of deletions (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, *$P < 0.05$, ***$P < 0.001$. (**D**) Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or BRCA1 siRNA (siBRCA1). P values calculated by statistical analysis "difference between proportions", *$P < 0.05$. (**E**) Western blot showing BRCA1 knockdown in HT1080 cells 48 h following siRNA transfection. (**F**) HR repair efficacy of DR-GFP reporter cells treated with NT control siRNA (NT), and BRCA1 siRNA (siBRCA1), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments, **** $P < 0.0001$.

and thereby promoting C-NHEJ and counteracting MMEJ (25,58). Further, we found that the RAD51-depleted mutational pattern following DSB induction was largely unaffected by the presence of the DNA-PKcs inhibitor. In particular, similar deletion lengths and use of microhomologies at the HPRT break-sites were observed following RAD51 depletion in the absence or presence of NU7441 ($P > 0.05$) (Figure 3A–C, and Supplementary Figure S6). No statistical difference was found between the MMEJ levels in cells treated with NU7441, siRNA against RAD51 or siRNA against RAD51 in combination with NU7441 (Figure 3C). These observations suggested that the DSB mutational signature in the absence of HR is largely independent of DNA-PKcs, a core C-NHEJ factor. To confirm this in another system, we used a well-defined GFP-based MMEJ reporter integrated within a U2OS cell line (EJ2-

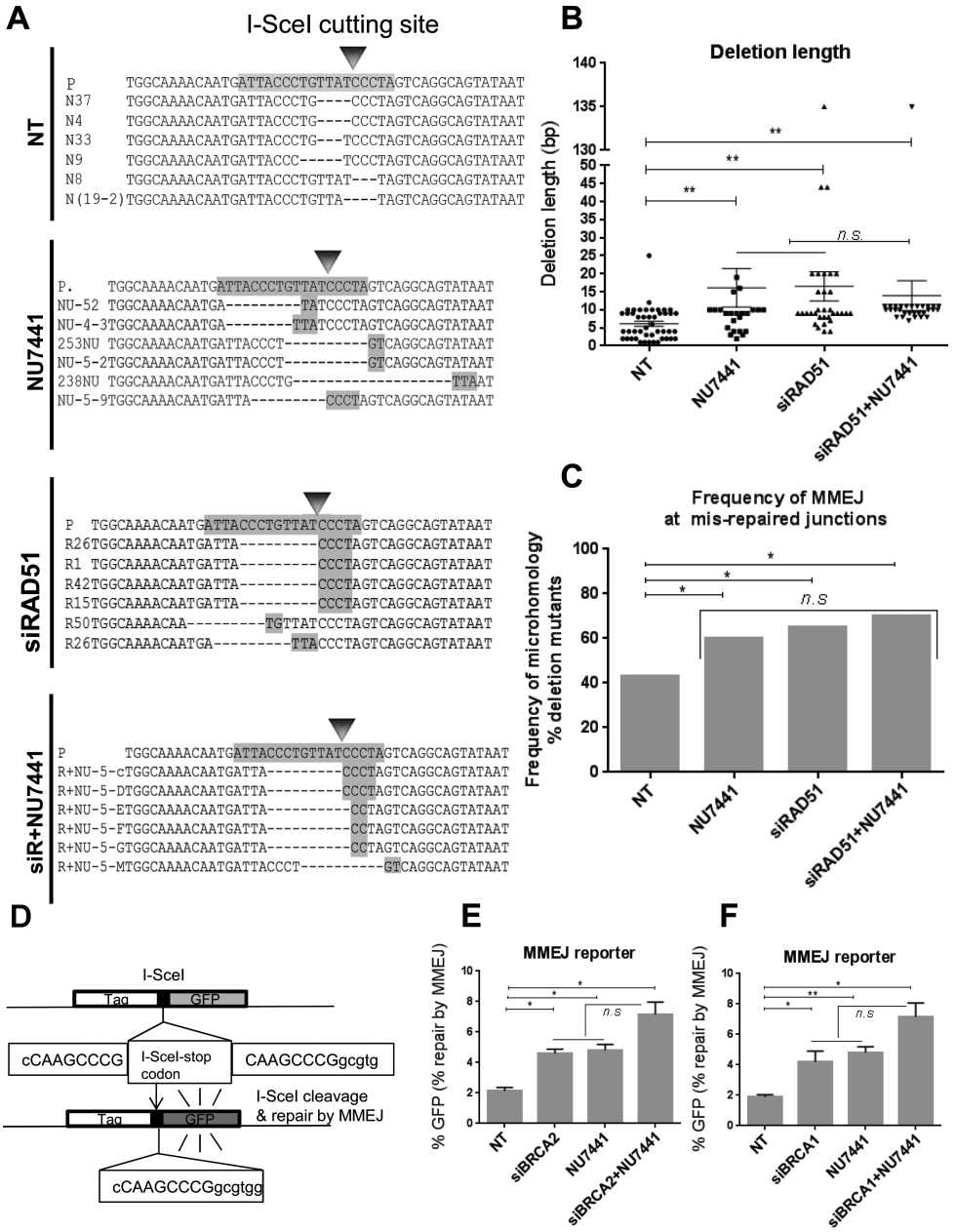


Figure 3. MMEJ-like DSB repair in HR-deficient cells is independent of DNA-PKcs. (**A**) Representative sequence alignments of PCR products obtained from HPRT negative cells treated with NT control siRNA (NT), a DNA-PKcs inhibitor (NU7441), siRNA against RAD51 (siRAD51) or siRNA against RAD51 plus NU7441 (siR+NU7441), from three independent experiments (see also Supplementary Figures S1A, S5-S6). I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted. (**B**) Average length of deletions (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, n.s., not significant, ** $P < 0.01$. (**C**) Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT), NU7441, RAD51 siRNA (siRAD51) or RAD51 and NU7441 (siRAD51+NU7441). P values calculated by statistical analysis "difference between proportions", *$P < 0.05$. (**D**) Schematic map of the EJ2-GFP reporter (53) to assess MMEJ efficacy, where the I-SceI cut site is flanked by 8 nucleotide homologous sequences capable of bridging the I-SceI-induced DSB by MMEJ and thus restoring a functional GFP cassette. (**E**) MMEJ repair efficacy of EJ2-GFP reporter cells treated with NT control siRNA (NT), BRCA2 siRNA (siBRCA2), DNA-PKcs inhibitor (NU7441) or BRCA2 and DNA-PKcs co-depleted cells (siBRCA2+NU7441), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. *$P < 0.05$. (**F**) MMEJ repair efficacy of EJ2-GFP reporter cells treated with NT control siRNA (NT), BRCA1 siRNA (siBRCA1), DNA-PKcs inhibitor (NU7441) or BRCA1 siRNA and DNA-PKcs (siBRCA1+NU7441), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. *$P < 0.05$, ** $P < 0.01$.

GFP) (53) (Figure 3D). Knockdown of BRCA2 in this system significantly increased the GFP-positive cells, indicative of MMEJ repair, compared to the NT controls (1.95-fold, $P = 0.0105$) (Figure 3E), confirming that BRCA2 suppresses MMEJ. Also, DNA-PKcs inhibition by NU7441 in this reporter cell line significantly increased the MMEJ repair events by 2.24-fold ($P = 0.0129$) (Figure 3E), consistent with a well-established role for C-NHEJ factors in suppressing MMEJ (42). Further, inhibition of DNA-PKcs by NU7441 in BRCA2-depleted cells led to a slight increase of the GFP-positive cells compared to that of BRCA2-depleted cells (1.62-fold; $P = 0.0523$) or control cells treated with NU7441 (1.4-fold; $P > 0.05$) (Figure 3E). When we repeated these experiments using BRCA1 depletion in place of BRCA2 depletion, we obtained very similar results (Figure 3F). These data obtained from the GFP reporter are consistent with those from the HPRT assay; in neither system did treatment with DNA-PKcs inhibitor impair the increase in MMEJ resulting from depletion of an HR protein suggesting that DNA-PKcs is dispensable for the induction of MMEJ in the HR-deficient cells.

Finally, to confirm that NU7441 efficiently inhibits C-NHEJ, we treated the previously described GFP-based C-NHEJ reporter (IRES-TK-EGFP, H1299 cells) (54) (Supplementary Figure S7A) with NU7441 and observed a 93.6% ($P < 0.0001$) fall in C-NHEJ (Supplementary Figure S7B). RAD51 depletion was confirmed by western blot (Supplementary Figure S7C). Collectively these data suggest that MMEJ in HR-impaired cells is independent of DNA-PKcs, a key factor for the C-NHEJ pathway. Also, the GFP-based reporter data suggest that BRCA2 and BRCA1 suppress MMEJ.

DSB-induced mutation signatures in HR-deficient cells require MMEJ factors

The above evidence shows clearly that DSB mis-repaired junctions with microhomologies typical of MMEJ are enhanced when HR is impaired, and that this effect is C-NHEJ-independent. We nevertheless sought direct evidence that MMEJ is indeed the pathway responsible for these mutational signatures. We therefore used the HPRT system to examine the DSB-induced mutational signatures following depletion of essential MMEJ factors. Given that HR and MMEJ share the initial resection step via CtIP and MRE11 (8,59,60), we investigated the impact of depleting these factors on the DSB-induced mutational signature. CtIP and MRE11 were depleted by siRNA-mediated knockdown in the HPRT reporter cells and the mutational frequencies following I-SceI-induced break were quantified. We found that, similar to the depletion of RAD51, SETD2, BRCA2 and BRCA1, depletion of either CtIP or MRE11 resulted in a significant increase in the I-SceI-induced mutation frequency (3.5%; $P = 0.0007$; and 3.6%; $P < 0.0001$, respectively) compared to 1.1% in NT controls (Figure 4A). The mutational frequencies in these backgrounds were slightly but significantly greater than those in BRCA2 depleted cells ($P<0.05$), consistent with roles for CtIP, and MRE11 upstream of BRCA2 during HR. Surprisingly, sequence analysis of HPRT-negative clones derived from CtIP- or MRE11-depleted cells identified a distinct mutational signature compared to that of SETD2-, RAD51-, BRCA2- or BRCA1-depleted cells (Figure 4B, and Supplementary Figures S8 and S9). This signature involved deletions with comparable lengths to those in NT controls (mean 5.6 bp), in contrast to the increased deletion lengths seen in BRCA2-depleted cells (mean 49 bp; $P < 0.0001$, and $P < 0.001$, respectively) (Figures 1C and 4C). Further analysis of the mis-repaired junctions of HPRT-negative clones from these backgrounds identified a significant decrease in the proportion of repair junctions associated with microhomologies in both CtIP- (13%; $P < 0.05$), and MRE11- (31%; $P < 0.05$) depleted cells, compared to NT controls (43% Figure 4D), or to BRCA2 72 (Figure 1D) and BRCA1-depleted cells (65%, Figure 2D). In summary, depletion of CtIP or MRE11, essential MMEJ factors that promote the initiation of resection, significantly reduced the proportion of DSB-induced deletions that were associated with microhomologies, relative to the NT controls. This contrasts sharply with depletion of the HR proteins SETD2, RAD51, BRCA2 or BRCA1, which significantly enhanced the same proportion (Figures 1D and 2D) (51).

To assess the effect of CtIP and MRE11 depletion on HR, we knocked down CtIP and MRE11 in the DR-GFP reporter (U2OS cells) (52) and found that the frequency of HR was significantly reduced by 92% ($P < 0.0001$), and 71% ($P < 0.0001$), respectively (Figure 4E). CtIP and MRE11 depletions were confirmed by western blot (Figure 4F). Together, these findings are in accordance with roles for CtIP and MRE11 in the initiation of resection during both HR and MMEJ.

PARP-1 is also known to play a key role in MMEJ by competing with Ku to bind DSBs, which in turn leads to the recruitment of MRE11 and CtIP to trigger resection (33,61). Consistent with this, Olaparib has recently been shown to decrease the Alt-NHEJ levels in the MMEJ-GFP reporter assay (62). Therefore, we treated the HR-proficient HT1080 cells (HPRT reporter) with a PARP-1 inhibitor, Olaparib (AZD2281) (63), and examined the DSB mutational pattern. The effects of Olaparib were similar to those of CtIP or MRE11 depletion. Thus Olaparib significantly increased the mutation frequency (2.6%; $P = 0.0073$; Figure 4A) compared to the 1.1% in NT controls, generating deletion lengths (mean 7 bp) comparable to the controls (mean 5.6 bp; $P > 0.05$; Figure 4B and C), and significantly shorter than deletion lengths in BRCA2-depleted HPRT mutants (mean 49 bp; $P = 0.0005$). Similarly, HPRT mutants arising following DSB induction in cells treated with Olaparib showed a significant reduction in the proportion of deletion junctions associated with microhomologies (25%; $P <0.05$) relative to NT controls (57%) (Figure 4B, D and Supplementary Figure S10), Collectively, these results suggest that microhomologies at DSB mis-repaired junctions in the HPRT reporter do indeed originate from the MMEJ repair machinery.

To confirm that the MMEJ-like signature in HR-compromised cells arises from MMEJ, we aimed to see whether knocking down a core MMEJ factor could abolish the MMEJ signature in these cells. We therefore, co-depleted BRCA1 and MRE11 in the HPRT system and monitored the DSB mutational signature. Simultaneous knockdown of BRCA1 and MRE11 resulted in a signifi-

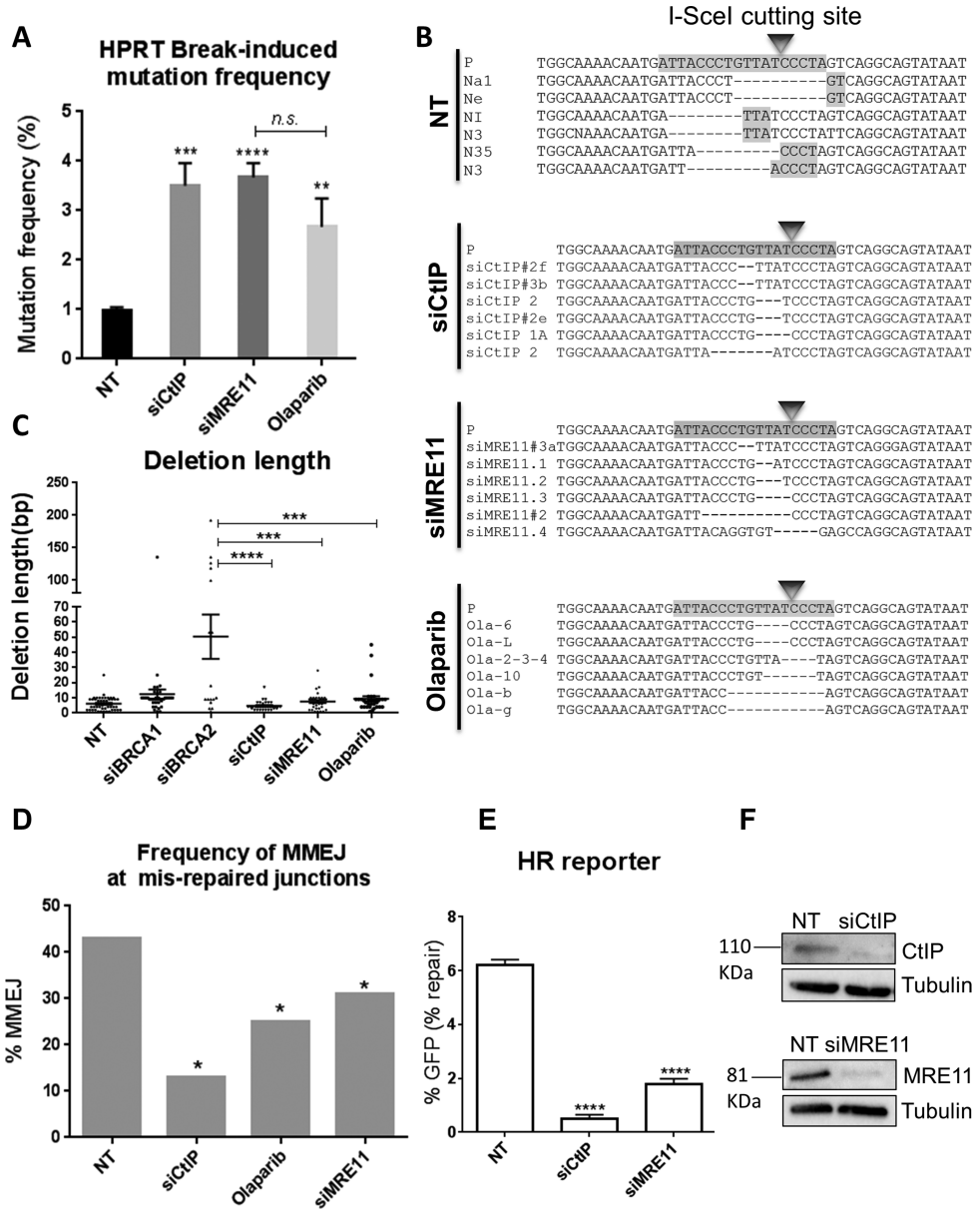


Figure 4. DSB-induced mutation signatures in HR-deficient cells require MMEJ factors. (**A**) Break-induced mutation frequency of HPRT reporter cells treated with NT control siRNA (NT), CtIP siRNA (siCtIP), MRE11 siRNA (siMRE11) or PARP-1 inhibitor (Olaparib). Error bars represent SEM from three independent experiments, n.s., not significant, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. (**B**) Representative sequence alignments of the HPRT negative PCR products in cells treated with NT control siRNA (NT), CtIP siRNA (siCtIP), MRE11 siRNA (siMRE11) or PARP-1 inhibitor (Olaparib), from three independent experiments (see also Supplementary Figures S1A, S8–10). I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted. (**C**) Average deletion lengths (bp) in different genetic backgrounds. Each dot represents an independent clone. The lines represent mean and SEM, $^{***}P < 0.001$, $^{****}P < 0.0001$. (**D**) Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or CtIP siRNA (siCtIP), MRE11 siRNA (siMRE11) or Olaparib. *P* values calculated by statistical analysis "difference between proportions", $^{*}P < 0.05$. (**E**) HR repair efficacy of DR-GFP reporter cells treated with NT control siRNA (NT), CtIP siRNA (siCtIP), or MRE11 siRNA (siMRE11), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. $^{****}P < 0.0001$. (**F**) Western blot showing CtIP and MRE11 knockdowns 48 h following siRNA transfection.

cantly increased mutation frequency of 3.4% compared to that of the NT background (1.1%, $P = 0.0013$) (Supplementary Figure S11A). This increased frequency of HPRT loss is consistent with the roles for BRCA1 and MRE11 in promoting DSB repair (14,15). Further, the MMEJ signature was absent in cells co-depleted for BRCA1 and MRE11. This was associated with deletion lengths comparable to that of NT cells (mean 5.6 bp), and a significantly reduced proportion of MMEJ (27%) compared to that of NT cells (43%, $P < 0.05$) (see Supplementary Figures S11 and S12). Together these data confirm that the MMEJ-like signature in the absence of HR factors is arising from the MMEJ pathway.

RPA suppresses MMEJ

RPA is a heterotrimeric ssDNA binding protein with roles in DNA replication, recombination and DSB repair, including HR. RPA displacement by RAD51 is a critical step in HR (18). In this regard, it has been previously shown in *S. cerevisiae* that RPA antagonizes MMEJ (45). Therefore, to test a possible role for RPA in repressing MMEJ repair in human cells, we used the GFP-based reporter for MMEJ repair (53) (Figure 3D). RPA knockdown (Figure 5C) resulted in significantly enhanced MMEJ levels compared to the NT controls (2.9-fold; $P < 0.0001$) (Figure 5A), suggesting that human RPA also suppresses MMEJ. POLQ has recently been identified as an MMEJ-promoting factor in mammalian cells (46,47). Thus siRNA knockdown of POLQ was used as a control for the GFP-based MMEJ reporter, which led to 94% reduction in *POLQ* gene expression levels, as evaluated by qRT-PCR (Figure 5B). This was found to significantly reduce MMEJ repair by 58% ($P = 0.0112$), consistent with its role in the mammalian MMEJ pathway (Figure 5A). Further, co-depletion of RPA and POLQ in the GFP-based MMEJ reporter led to significantly reduced levels of MMEJ compared to that of RPA-depleted cells ($P = 0.0007$). Surprisingly, however, MMEJ levels in cells co-depleted for POLQ and RPA were still significantly greater than NT controls ($P = 0.0022$) (Figure 5A), suggestive of the possible existence of POLQ-dependent and POLQ-independent MMEJ pathways being suppressed by RPA. Finally, to confirm the role for POLQ in promoting MMEJ in the HPRT assay, we depleted POLQ in this system and monitored the DSB mutational pattern. Analysis of the mutational signature in cells treated with siRNA against POLQ showed a significantly decreased MMEJ levels compared to that of the NT background, (Figure 5D and E and Supplementary Figure S13). These results resemble our findings in the MMEJ-GFP assay, confirming that the HPRT and the GFP reporter assays measure the same process.

DISCUSSION

Mutations arising from aberrant DSB repair are deleterious for cells and can give rise to genomic instability, a hallmark of cancer (64). Yet how DSB-induced chromosomal rearrangements are suppressed is poorly characterized. Further, there is little mechanistic insight into the relationship between disrupting repair pathways and the mechanisms of DSB mis-repair. Here, using an I-SceI-cleavable reporter assay based on the human endogenous *HPRT* gene (50,51), together with an array of I-SceI-cleavable GFP-based reporter assays (52–54), in different cancer cell lines we have established a role for HR factors in suppressing mutagenic MMEJ following DSB resection. In this context, we consider MMEJ to arise as the default pathway following HR inactivation rather than being actively suppressed by HR (see model in Figure 6). Following on from our initial observations in RAD51- and SETD2-depleted cells (51), this study confirms that depletion of other HR factors (BRCA2, BRCA1 and RPA) also results in significantly elevated levels of a common DSB mutational signature in which DSB-induced deletions are associated with microhomologies of 2–6 bp at the break junctions. Our observations are in alignment with the recent studies in human cells indicating that chromosomal rearrangement junctions formed by MMEJ contain 2–6 bp of microhomology regions (38). Further, our data also indicate that while failed HR leads to an increase in MMEJ-induced deletions, surprisingly no effect was observed on insertions, suggesting that these arise through other mutagenic NHEJ events in this context. In this respect, insertions were still observed, albeit at low levels, following inhibition of DNA-PKcs or following knockdown or inhibition of MMEJ factors, suggesting these insertions arise through other end-joining mechanisms. However, this needs to be investigated in more detail.

Because the junctional microhomologies can potentially represent DSB mutational signatures of either C-NHEJ or MMEJ pathways (65,66), it is important to know which of these pathways is involved. However the abrogation of these signatures following depletion or inhibition of CtIP, MRE11, PARP-1, all of which facilitate MMEJ, provides strong genetic evidence that the MMEJ pathway is responsible. The observation that MMEJ levels are decreased following PARP-1 inhibition with Olaparib in the HPRT system, is in agreement with a recent study in which a similar finding was found using a GFP-based reporter assay (62).

Additionally, the persistence of junctional microhomologies in HR-depleted cells treated with DNA-PKcs inhibitor excludes the possibility that C-NHEJ is required. These conclusions are further supported by the significantly greater levels of MMEJ events detected by a GFP-based reporter following depletion of HR factors, even when DNA-PKcs is inhibited (Figure 3E and F). Our findings therefore indicate that, following DSB induction, HR proteins preserve genomic integrity by suppressing mutations arising through the MMEJ pathway. MMEJ can be considered an important DSB repair pathway in its own right. However our data suggest that its levels can be increased as a result of disrupted HR. In this respect, we know MMEJ has been proposed to be higher during S/G2 (59), which may arise through incomplete HR. For example, one-ended breaks arising from replication collapse may give rise to both MMEJ and HR.

While the role of CtIP and the MRN complex in resection initiation is well established (67), the exact role of BRCA1 in resection remains uncertain (68–70). It has recently been reported that, although the BRCA1-CtIP interaction is not crucial for DNA end-resection, it enhances the speed of CtIP-mediated resection (17). It is also known that CtIP and MRE11 are required for both HR and MMEJ repair pathways by triggering resection (59). Differing results regarding the contribution of BRCA1 in MMEJ have been reported. Thus, while one recent study in MEFs found that BRCA1 facilitates MMEJ at uncapped telomeres (49), work in DT40 (chicken) B cells suggested that MMEJ is not influenced by BRCA1 (70). Various studies of plasmid-based (exogenous) DSB repair substrates in human cells, however, provide evidence that BRCA1 represses Alt-NHEJ events by promoting the fidelity of NHEJ. (48,71–73). Here, using an endogenous HPRT assay together with GFP reporter assays in different human cell lines, we identified a DSB-induced mutational signature for BRCA1-depleted cells, which is distinct from that of CtIP-depleted or MRE11-

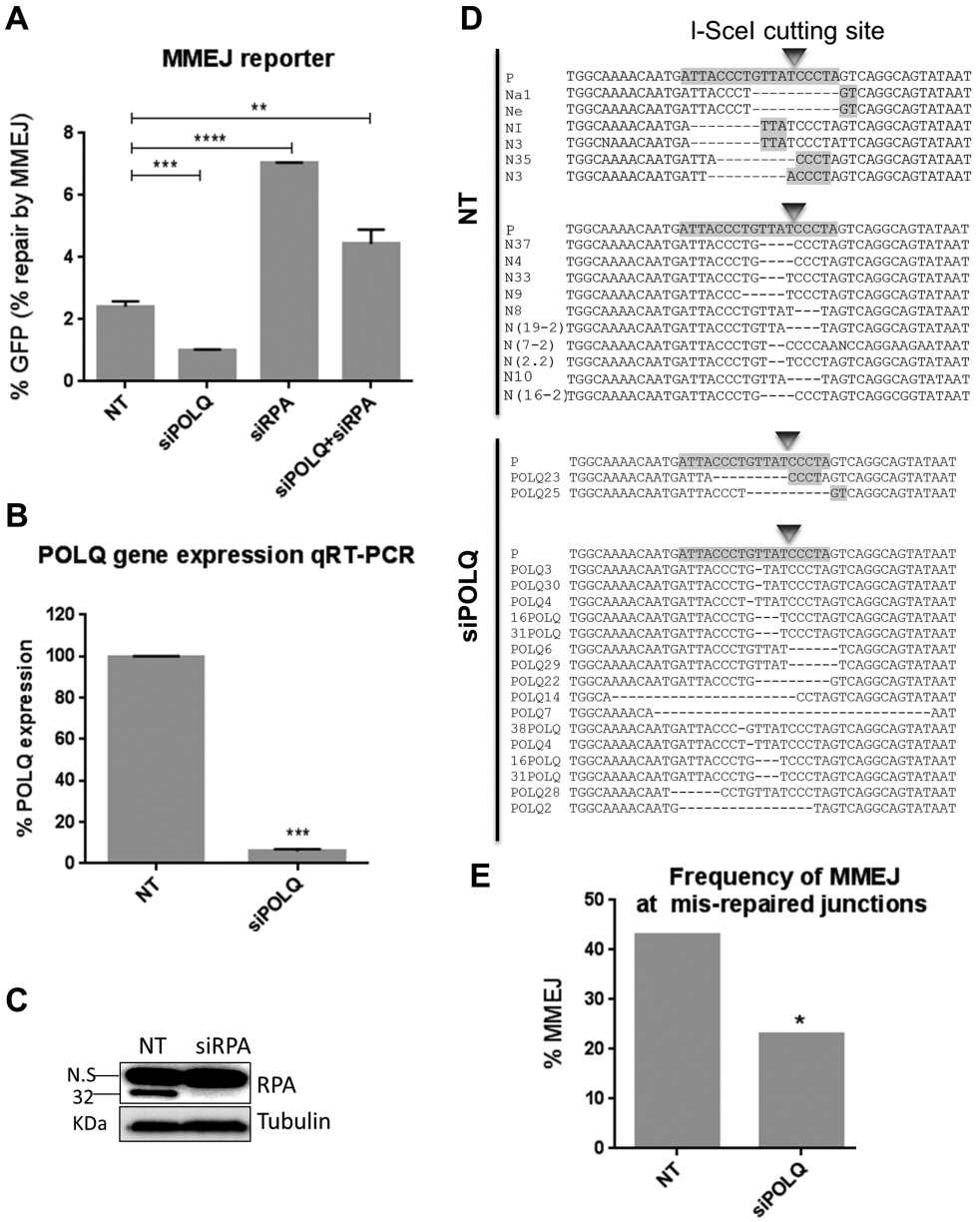


Figure 5. RPA suppresses MMEJ. (**A**) MMEJ repair efficacy of EJ2-GFP reporter cells treated with NT control siRNA (NT), POLQ siRNA (siPOLQ), RPA siRNA (siRPA) and POLQ plus RPA (siPOLQ+siRPA) indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. (**B**) POLQ gene expression measured by RT-qPCR using the primers indicated in Materials and Methods. (**C**) Western blot showing RPA knockdown 48 h following siRNA transfection. (**D**) Representative sequence alignments of the PCR products (see Supplementary Figure S13) in NT control (NT), and POLQ-depleted (siPOLQ) cells. I-SceI recognition sequence and terminal microhomologies at the break sites are highlighted. (**E**) Frequency of MMEJ at mis-repaired junctions in HPRT deletion mutants isolated from cells treated with NT control siRNA (NT) or POLQ siRNA (siPOLQ). *P* values calculated by statistical analysis "difference between proportions", $^{*}P < 0.05$.

depleted cells. The significantly greater HPRT deletion lengths in BRCA1-depleted cells compared to those of CtIP or MRE11-depleted cells supports the notion that BRCA1 operates downstream of CtIP and MRE11. Further, the significantly smaller HPRT deletion lengths in BRCA1-depleted cells compared to those of BRCA2-depleted cells, suggests that BRCA1 functions downstream of CtIP and MRE11 but upstream of BRCA2 during HR to promote genome stability by facilitating resection and thereby suppressing the mutagenic MMEJ. Moreover, the significantly increased levels of MMEJ, detected by both the HPRT and GFP assays, following BRCA1 depletion suggest that BRCA1 suppresses MMEJ repair. This is in agreement with previous studies using plasmid-based reporter systems (48,71).

POLQ has recently been identified as a polymerase that contributes to mammalian MMEJ by promoting DNA end joining and microhomology annealing (37,38). Further, recent studies have suggested that cancer cells with defective HR are dependent on POLQ-mediated MMEJ repair (46,47). Moreover, POLQ is upregulated in a range of human cancers, and indicates a poor clinical prognosis, espe-

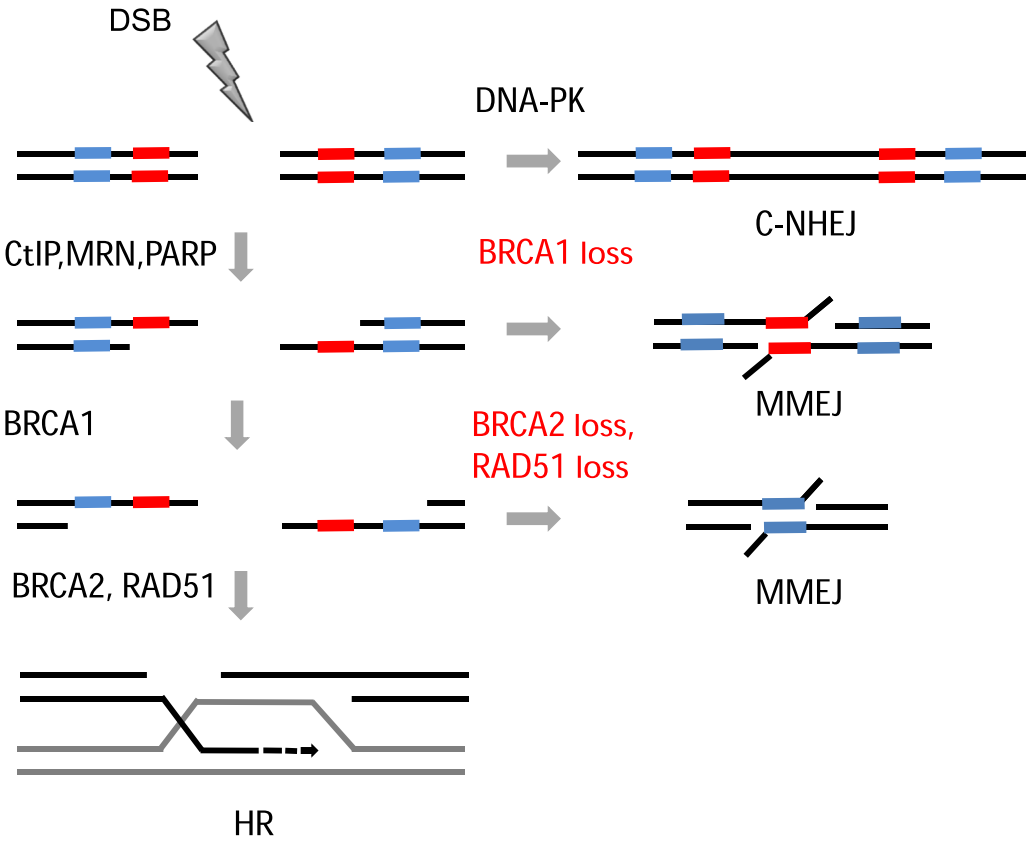


Figure 6. Model. DSBs are mainly repaired by C-NHEJ, however upon resection initiation by CtIP and MRN complex repair occurs via HR pathway. If HR is disrupted after initiation of resection, regions of microhomology on either side of the break will anneal, resulting in repair via the mutagenic MMEJ pathway.

cially in breast tumours (74–76). In agreement with these reports, our observations support a role for POLQ in facilitating MMEJ in human cancer cells. Also RPA has an established role in repressing MMEJ in *S. cerevisiae* by removing DNA secondary structures, which prevent RAD51 loading and therefore HR (45,77). Consistent with this, we demonstrate a striking role for RPA in suppressing mutagenic MMEJ in human cancer cells. This could reflect roles for RPA in early HR, thereby suppressing MMEJ and/or in binding ssDNA, thereby preventing spontaneous MMEJ, as previously proposed in *S. cerevisiae* (45). Our data, however, also identify dramatically increased MMEJ levels in cells following simultaneous depletion of POLQ and RPA. This is indicative of a role for RPA in suppressing MMEJ independently of POLQ, which could be explained by the suggested involvement of POLQ in a subset of Alt-NHEJ repair pathways specifically leading to insertions (78,79). Our results also suggest that there might be POLQ-dependent and POLQ-independent MMEJ pathways in human cancer cells, both of which are suppressed by RPA. Our data do not, however, exclude the possibility that the slight residual activity of POLQ following its siRNA-mediated knockdown, may contribute to the great MMEJ levels in cells co-depleted for RPA and POLQ. In this regard, a more in-depth analysis is required.

While our data show that MMEJ is enhanced when HR is impaired, the nature of the corresponding decrease in other forms of mutagenic NHEJ is less clear and could indicate inhibition of indel-forming C-NHEJ or other Alt-NHEJ events that are not dependent on junctional microhomologies (6). Further analyses are required to identify which DSB repair pathways are affected.

Our combined observations support a model (Figure 6) in which DSB repair in HR-proficient cells, results in a mutational signature with a particular balance between deletions with and without associated microhomologies. When HR is impaired, however, this balance is channeled toward deletions associated with microhomologies as a result of enhanced MMEJ. Mechanistically, we propose that, following DSB induction and initiation of resection, the repair pathway choice can switch from HR to MMEJ, depending on the availability of HR downstream proteins. In the absence of HR factors, HR is blocked, and the resected ends are poor substrates for C-NHEJ. As a result, microhomologous sequences present on ssDNA either side of the resected ends anneal via the mutagenic MMEJ repair mechanism leading to microhomology-mediated deletions. Also consistent with our findings, a recent study using MEFs found increased translocation levels following Rad54 deletion (44). RAD54 is a key HR factor promoting recombination through interactions with RAD51 (80). Although there is no sequence-based evidence for the use of microhomologies in generating translocation junctions, translocations were largely PARP-

1- and Lig 3-dependent suggesting that they had arisen via the MMEJ pathway (39,81).

Mutations in DNA DSB repair genes are common features in different human cancer types. Among the DSB repair pathways, MMEJ is necessarily mutagenic and is frequently associated with genomic rearrangements (39,82). Importantly, high-resolution sequencing studies have identified microhomologies as a prevalent mutational signature at rearrangement breakpoints in several cancer types including breast, colorectal and prostate adenocarcinomas (83,84). Our observations therefore suggest that the microhomology mutational signature in different types of malignancies may arise from failed HR at a stage after the initiation of resection. This mechanistic understanding of the interplay between HR and MMEJ repair pathways could be exploited to develop therapies for cancer patients deficient in these DSB repair pathways.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Jeremy Stark for the MMEJ reporter cell line, Atsushi Shibata and Hideaki Hogiwara for the NHEJ reporter cell line and Fereydoon Ahrabi for his advice on statistical analysis.

FUNDING

Medical Research Council [MC_PC_12003 to S.S. and T.C.H.]; Cancer Research UK [C5255/A15935 to S.A., C38302/A12981 to G.P., C34326/A13092 to G.S.H.]; Clarendon Scholarship [to S.A., S.X.P., G.P]; BBSRC [BB/H003371/1 to A.C.G.P.]. Funding for open access charge: Medical Research Council [MC_PC_12003].
Conflict of interest statement. None declared.

REFERENCES

1. Jeggo,P. and Lavin,M.F. (2009) Cellular radiosensitivity: how much better do we understand it? *Int. J. Radiat. Biol.*, **85**, 1061–1081.
2. Symington,L.S. and Gautier,J. (2011) Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.*, **45**, 247–271.
3. Kabotyanski,E.B., Gomelsky,L., Han,J.O., Stamato,T.D. and Roth,D.B. (1998) Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.*, **26**, 5333–5342.
4. Liang,F. and Jasin,M. (1996) Ku80-deficient cells exhibit excess degradation of extrachromosomal DNA. *J. Biol. Chem.*, **271**, 14405–14411.
5. Boulton,S.J. and Jackson,S.P. (1996) Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J.*, **15**, 5093–5103.
6. Decottignies,A. (2013) Alternative end-joining mechanisms: a historical perspective. *Front. Genet.*, **4**, 48.
7. Simsek,D., Brunet,E., Wong,S.Y., Katyal,S., Gao,Y., McKinnon,P.J., Lou,J., Zhang,L., Li,J., Rebar,E.J. *et al.* (2011) DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. *PLoS Genet.*, **7**, e1002080.
8. Bennardo,N., Cheng,A., Huang,N. and Stark,J.M. (2008) Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet.*, **4**, e1000110.
9. 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.
10. Takeda,S., Nakamura,K., Taniguchi,Y. and Paull,T.T. (2007) Ctp1/CtIP and the MRN complex collaborate in the initial steps of homologous recombination. *Mol. Cell*, **28**, 351–352.
11. 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.
12. 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.
13. Stark,J.M., Pierce,A.J., Oh,J., Pastink,A. and Jasin,M. (2004) Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol. Cell. Biol.*, **24**, 9305–9316.
14. Moynahan,M.E., Chiu,J.W., Koller,B.H. and Jasin,M. (1999) Brca1 controls homology-directed DNA repair. *Mol. Cell*, **4**, 511–518.
15. Zhong,Q., Chen,C.F., Li,S., Chen,Y., Wang,C.C., Xiao,J., Chen,P.L., Sharp,Z.D. and Lee,W.H. (1999) Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science*, **285**, 747–750.
16. Yu,X., Wu,L.C., Bowcock,A.M., Aronheim,A. and Baer,R. (1998) The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. *J. Biol. Chem.*, **273**, 25388–25392.
17. Cruz-Garcia,A., Lopez-Saavedra,A. and Huertas,P. (2014) BRCA1 accelerates CtIP-mediated DNA-end resection. *Cell Reports*, **9**, 451–459.
18. 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.
19. 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.
20. Wu,L. and Hickson,I.D. (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature*, **426**, 870–874.
21. West,S.C. (2003) Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.*, **4**, 435–445.
22. Nassif,N., Penney,J., Pal,S., Engels,W.R. and Gloor,G.B. (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol. Cell. Biol.*, **14**, 1613–1625.
23. Ciccia,A. and Elledge,S.J. (2010) The DNA damage response: making it safe to play with knives. *Mol. Cell*, **40**, 179–204.
24. Deriano,L. and Roth,D.B. (2013) Modernizing the nonhomologus end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.*, **47**, 433–455.
25. Mahaney,B.L., Meek,K. and Lees-Miller,S.P. (2009) Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochem. J.*, **417**, 639–650.
26. 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.
27. Meek,K., Dang,V. and Lees-Miller,S.P. (2008) DNA-PK: the means to justify the ends? *Adv. Immunol.*, **99**, 33–58.
28. Ferguson,D.O., Sekiguchi,J.M., Chang,S., Frank,K.M., Gao,Y., DePinho,R.A. and Alt,F.W. (2000) The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 6630–6633.
29. Ghezraoui,H., Piganeau,M., Renouf,B., Renaud,J.B., Sallmyr,A., Ruis,B., Oh,S., Tomkinson,A.E., Hendrickson,E.A., Giovannangeli,C. *et al.* (2014) Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol. Cell*, **55**, 829–842.
30. Jones,R.E., Oh,S., Grimstead,J.W., Zimbric,J., Roger,L., Heppel,N.H., Ashelford,K.E., Liddiard,K., Hendrickson,E.A. and Baird,D.M. (2014) Escape from telomere-driven crisis is DNA ligase III dependent. *Cell Reports*, **8**, 1063–1076.
31. 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.

32. Wang,M., Wu,W., 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.
33. Hochegger,H., Dejsuphong,D., Fukushima,T., Morrison,C., Sonoda,E., Schreiber,V., Zhao,G.Y., Saberi,A., Masutani,M., Adachi,N. *et al.* (2006) Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells. *EMBO J.*, **25**, 1305–1314.
34. 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.
35. Wang,H., Rosidi,B., Perrault,R., Wang,M., Zhang,L., Windhofer,F. and Iliakis,G. (2005) DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.*, **65**, 4020–4030.
36. Seki,M., Masutani,C., Yang,L.W., Schuffert,A., Iwai,S., Bahar,I. and Wood,R.D. (2004) High-efficiency bypass of DNA damage by human DNA polymerase Q. *EMBO J.*, **23**, 4484–4494.
37. Yousefzadeh,M.J., Wyatt,D.W., Takata,K., Mu,Y., Hensley,S.C., Tomida,J., Bylund,G.O., Doublie,S., Johansson,E., Ramsden,D.A. *et al.* (2014) Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet.*, **10**, e1004654.
38. Kent,T., Chandramouly,G., McDevitt,S.M., Ozdemir,A.Y. and Pomerantz,R.T. (2015) Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat. Struct. Mol. Biol.*, **22**, 230–237.
39. 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.
40. Kennedy,R.D. and D'Andrea,A.D. (2006) DNA repair pathways in clinical practice: lessons from pediatric cancer susceptibility syndromes. *J. Clin. Oncol.*, **24**, 3799–3808.
41. Huertas,P. (2010) DNA resection in eukaryotes: deciding how to fix the break. *Nat. Struct. Mol. Biol.*, **17**, 11–16.
42. Iliakis,G. (2009) Backup pathways of NHEJ in cells of higher eukaryotes: cell cycle dependence. *Radiother. Oncol.*, **92**, 310–315.
43. Cheng,Q., Barboule,N., Frit,P., Gomez,D., Bombarde,O., Couderc,B., Ren,G.S., Salles,B. and Calsou,P. (2011) Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks. *Nucleic Acids Res.*, **39**, 9605–9619.
44. Soni,A., Siemann,M., Grabos,M., Murmann,T., Pantelias,G.E. and Iliakis,G. (2014) Requirement for Parp-1 and DNA ligases 1 or 3 but not of Xrcc1 in chromosomal translocation formation by backup end joining. *Nucleic Acids Res.*, **42**, 6380–6392.
45. Deng,S.K., Gibb,B., de Almeida,M.J., Greene,E.C. and Symington,L.S. (2014) RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nat. Struct. Mol. Biol.*, **21**, 405–412.
46. Mateos-Gomez,P.A., Gong,F., Nair,N., Miller,K.M., Lazzerini-Denchi,E. and Sfeir,A. (2015) Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature*, **518**, 254–257.
47. Ceccaldi,R., Liu,J.C., Amunugama,R., Hajdu,I., Primack,B., Petalcorin,M.I., O'Connor,K.W., Konstantinopoulos,P.A., Elledge,S.J., Boulton,S.J. *et al.* (2015) Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair. *Nature*, **518**, 258–262.
48. Jiang,G., Plo,I., Wang,T., Rahman,M., Cho,J.H., Yang,E., Lopez,B.S. and Xia,F. (2013) BRCA1-Ku80 protein interaction enhances end-joining fidelity of chromosomal double-strand breaks in the G1 phase of the cell cycle. *J. Biol. Chem.*, **288**, 8966–8976.
49. Badie,S., Carlos,A.R., Folio,C., Okamoto,K., Bouwman,P., Jonkers,J. and Tarsounas,M. (2015) BRCA1 and CtIP promote alternative non-homologous end-joining at uncapped telomeres. *EMBO J.*, **34**, 410–424.
50. Gravells,P., Ahrabi,S., Vangala,R.K., Tomita,K., Brash,J.T., Brustle,L.A., Chung,C., Hong,J.M., Kaloudi,A., Humphrey,T.C. *et al.* (2015) Use of the HPRT gene to study nuclease-induced DNA double-strand break repair. *Hum. Mol. Genet.*, **24**, 7097–7110.
51. Pfister,S.X., Ahrabi,S., Zalmas,L.P., Sarkar,S., Aymard,F., Bachrati,C.Z., Helleday,T., Legube,G., La Thangue,N.B., Porter,A.C. *et al.* (2014) SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Reports*, **7**, 2006–2018.
52. 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.
53. Gunn,A. and Stark,J.M. (2012) I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. *Methods Mol. Biol.*, **920**, 379–391.
54. Ogiwara,H., Ui,A., Otsuka,A., Satoh,H., Yokomi,I., Nakajima,S., Yasui,A., Yokota,J. and Kohno,T. (2011) Histone acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors. *Oncogene*, **30**, 2135–2146.
55. Porter,A.C. and Itzhaki,J.E. (1993) Gene targeting in human somatic cells. Complete inactivation of an interferon-inducible gene. *Eur. J. Biochem.*, **218**, 273–281.
56. Prakash,R., Zhang,Y., Feng,W. and Jasin,M. (2015) Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harb. Perspect. Biol.*, **7**, a016600.
57. Zhao,Y., Thomas,H.D., Batey,M.A., Cowell,I.G., Richardson,C.J., Griffin,R.J., Calvert,A.H., Newell,D.R., Smith,G.C. and Curtin,N.J. (2006) Preclinical evaluation of a potent novel DNA-dependent protein kinase inhibitor NU7441. *Cancer Res.*, **66**, 5354–5362.
58. Iliakis,G. (2009) Backup pathways of NHEJ in cells of higher eukaryotes: cell cycle dependence. *Radiother. Oncol.*, **92**, 310–315.
59. Truong,L.N., Li,Y., Shi,L.Z., Hwang,P.Y., He,J., Wang,H., Razavian,N., Berns,M.W. and Wu,X. (2013) Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7720–7725.
60. Jasin,M. and Rothstein,R. (2013) Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.*, **5**, a012740.
61. Beck,C., Robert,I., Reina-San-Martin,B., Schreiber,V. and Dantzer,F. (2014) Poly(ADP-ribose) polymerases in double-strand break repair: Focus on PARP1, PARP2 and PARP3. *Exp. Cell Res.*, **329**, 18–25.
62. Howard,S.M., Yanez,D.A. and Stark,J.M. (2015) DNA damage response factors from diverse pathways, including DNA crosslink repair, mediate alternative end joining. *Plos Genet.*, **11**, e1004943.
63. Zimmermann,M., Lottersberger,F., Buonomo,S.B., Sfeir,A. and de Lange,T. (2013) 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science*, **339**, 700–704.
64. Hanahan,D. and Weinberg,R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
65. Gauss,G.H. and Lieber,M.R. (1996) Mechanistic constraints on diversity in human V(D)J recombination. *Mol. Cell. Biol.*, **16**, 258–269.
66. Pannunzio,N.R., Li,S., Watanabe,G. and Lieber,M.R. (2014) Non-homologous end joining often uses microhomology: implications for alternative end joining. *DNA repair*, **17**, 74–80.
67. Symington,L.S. (2014) End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb. Perspect. Biol.*, **6**, a016436.
68. Nakamura,K., Kogame,T., Oshiumi,H., Shinohara,A., Sumitomo,Y., Agama,K., Pommier,Y., Tsutsui,K.M., Tsutsui,K., Hartsuiker,E. *et al.* (2010) Collaborative action of Brca1 and CtIP in elimination of covalent modifications from double-strand breaks to facilitate subsequent break repair. *PLoS Genet.*, **6**, e1000828.
69. Reczek,C.R., Szabolcs,M., Stark,J.M., Ludwig,T. and Baer,R. (2013) The interaction between CtIP and BRCA1 is not essential for resection-mediated DNA repair or tumor suppression. *J. Cell Biol.*, **201**, 693–707.
70. 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.
71. Thompson,E.G., Fares,H. and Dixon,K. (2012) BRCA1 requirement for the fidelity of plasmid DNA double-strand break repair in cultured breast epithelial cells. *Environ. Mol. Mutagen.*, **53**, 32–43.
72. Baldeyron,C., Jacquemin,E., Smith,J., Jacquemont,C., De Oliveira,I., Gad,S., Feunteun,J., Stoppa-Lyonnet,D. and Papadopoulo,D. (2002) A single mutated BRCA1 allele leads to impaired fidelity of double strand break end-joining. *Oncogene*, **21**, 1401–1410.

73. Wang,H.C., Chou,W.C., Shieh,S.Y. and Shen,C.Y. (2006) Ataxia telangiectasia mutated and checkpoint kinase 2 regulate BRCA1 to promote the fidelity of DNA end-joining. *Cancer Res.*, **66**, 1391–1400.
74. Higgins,G.S., Harris,A.L., Prevo,R., Helleday,T., McKenna,W.G. and Buffa,F.M. (2010) Overexpression of POLQ confers a poor prognosis in early breast cancer patients. *Oncotarget*, **1**, 175–184.
75. Kawamura,K., Bahar,R., Seimiya,M., Chiyo,M., Wada,A., Okada,S., Hatano,M., Tokuhisa,T., Kimura,H., Watanabe,S. *et al.* (2004) DNA polymerase theta is preferentially expressed in lymphoid tissues and upregulated in human cancers. *Int. J. Cancer*, **109**, 9–16.
76. Lemee,F., Bergoglio,V., Fernandez-Vidal,A., Machado-Silva,A., Pillaire,M.J., Bieth,A., Gentil,C., Baker,L., Martin,A.L., Leduc,C. *et al.* (2010) DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 13390–13395.
77. McVey,M. (2014) RPA puts the brakes on MMEJ. *Nat. Struct. Mol. Biol.*, **21**, 348–349.
78. Koole,W., van Schendel,R., Karambelas,A.E., van Heteren,J.T., Okihara,K.L. and Tijsterman,M. (2014) A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.*, **5**, 3216.
79. Yu,A.M. and McVey,M. (2010) Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.*, **38**, 5706–5717.
80. Golub,E.I., Kovalenko,O.V., Gupta,R.C., Ward,D.C. and Radding,C.M. (1997) Interaction of human recombination proteins Rad51 and Rad54. *Nucleic Acids Res.*, **25**, 4106–4110.
81. Liang,L., Deng,L., Nguyen,S.C., Zhao,X., Maulion,C.D., Shao,C. and Tischfield,J.A. (2008) Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Res.*, **36**, 3297–3310.
82. Ottaviani,D., LeCain,M. and Sheer,D. (2014) The role of microhomology in genomic structural variation. *Trends Genet.*, **30**, 85–94.
83. Conrad,D.F., Bird,C., Blackburne,B., Lindsay,S., Mamanova,L., Lee,C., Turner,D.J. and Hurles,M.E. (2010) Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. *Nat. Genet.*, **42**, 385–391.
84. Bass,A.J., Lawrence,M.S., Brace,L.E., Ramos,A.H., Drier,Y., Cibulskis,K., Sougnez,C., Voet,D., Saksena,G., Sivachenko,A. *et al.* (2011) Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. *Nat. Genet.*, **43**, 964–968.

OXFORD

Human Molecular Genetics, 2015, Vol. 24, No. 24 7097–7110

doi: 10.1093/hmg/ddv409
Advance Access Publication Date: 30 September 2015
Original Article

ORIGINAL ARTICLE

Use of the *HPRT* gene to study nuclease-induced DNA double-strand break repair

Polly Gravells[1,†], Sara Ahrabi[2,†], Rajani K. Vangala[1,†], Kazunori Tomita[1], James T. Brash[1], Lena A. Brustle[1], Christopher Chung[1], Julia M. Hong[1], Aikaterini Kaloudi[1], Timothy C. Humphrey[2] and Andrew C.G. Porter[1,*]

[1]Gene Targeting Group, Centre for Haematology, Imperial College Faculty of Medicine, London W120NN, UK and [2]CRUK MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Oxford OX3 7DQ, UK

*To whom correspondence should be addressed at: Gene Targeting Group, Centre for Haematology, Imperial College Faculty of Medicine, Hammersmith Hospital Campus, Du Cane Road, London W12 0NN, UK. Tel: +44 2083838276; Fax: +44 2087409679; Email: andy.porter@imperial.ac.uk

Abstract

Understanding the mechanisms of chromosomal double-strand break repair (DSBR) provides insight into genome instability, oncogenesis and genome engineering, including disease gene correction. Research into DSBR exploits rare-cutting endonucleases to cleave exogenous reporter constructs integrated into the genome. Multiple reporter constructs have been developed to detect various DSBR pathways. Here, using a single endogenous reporter gene, the X-chromosomal disease gene encoding hypoxanthine phosphoribosyltransferase (*HPRT*), we monitor the relative utilization of three DSBR pathways following cleavage by *I-SceI* or CRISPR/Cas9 nucleases. For *I-SceI*, our estimated frequencies of accurate or mutagenic non-homologous end-joining and gene correction by homologous recombination are 4.1, 1.5 and 0.16%, respectively. Unexpectedly, *I-SceI* and Cas9 induced markedly different DSBR profiles. Also, using an *I-SceI*-sensitive *HPRT* minigene, we show that gene correction is more efficient when using long double-stranded DNA than single- or double-stranded oligonucleotides. Finally, using both endogenous *HPRT* and exogenous reporters, we validate novel cell cycle phase-specific *I-SceI* derivatives for investigating cell cycle variations in DSBR. The results obtained using these novel approaches provide new insights into template design for gene correction and the relationships between multiple DSBR pathways at a single endogenous disease gene.

Introduction

Inappropriate repair of chromosomal DNA double-strand breaks (DSBs) can lead to mutagenesis, gross chromosomal instability and genetic disease (1). Knowledge of DSB repair pathways, and their defects in particular cancers, has helped to define oncogenic mechanisms and to develop new therapeutic strategies (2). Such knowledge also underpins powerful genome engineering methods that use customized endonucleases, including clustered regularly interspaced short palindromic repeat (CRISPR)/Cas-based RNA-guided nucleases, to make targeted chromosomal DSBs (3). Therapeutic gene targeting (GT) (4) for the correction of disease-causing mutations is an emerging application of such methods (5).

There are two main DSB repair (DSBR) mechanisms: homologous recombination (HR) (6) and non-homologous end-joining

[†]The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.
Received: August 12, 2015. **Revised and Accepted:** September 23, 2015

© The Author 2015. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

(NHEJ) (7). HR requires a homologous DNA repair template, usually in the form of a sister chromatid, which ensures that the original DNA sequence is restored. Repair templates can be engineered, however, and introduced into cells to make a defined genomic modification via HR (GT). HR requires extensive 5′–3′ end-resection at DSBs to generate single-stranded tails that form a RAD51 nucleofilament capable of invading a duplex DNA template. Requirements for a sister chromatid template and cyclin-dependent kinase activity (8) restrict HR to the S and G2 phases of the cell cycle (9). NHEJ lacks these requirements and can occur throughout the cell cycle. In classical NHEJ, DNA ends are protected by the Ku heterodimer; this minimizes end-resection allowing for accurate end-joining (accNHEJ, also termed precise ligation) (7,10). Here, we use the term accNHEJ to describe the joining of any two DNA ends that have undergone no gain or loss of nucleotides. In some circumstances, such as the absence of Ku heterodimer, NHEJ is more likely to involve limited end-resection, resulting in insertions and/or deletions (indels) at the site of the DSB. Such mutagenic NHEJ (mutNHEJ) often involves the annealing of short (e.g. 2–6 nt) sequence homologies from each side of the DSB (micro-homology-mediated end-joining; MMEJ). In an analogous pathway termed single-strand annealing (SSA), that can also be considered a form of non-conservative RAD51-independent HR, deletions result from extensive end-resection followed by annealing of longer (e.g. >100 nt) repeats (6).

Customized endonucleases can be introduced into cells to generate targeted indels via mutNHEJ or to promote specific modifications via GT. An ability to promote one DSBR pathway at the expense of others would be valuable for facilitating the desired genome engineering outcome. DNA resection, cell cycle phase, DNA end-structure and host cell type are all known to influence DSBR pathway choice (11,12). Further investigations into how these and other variables affect pathway choice will be important to facilitate genome manipulation and improve our understanding of DSBR and genome instability.

Research into DSBR has benefitted greatly from assays for the repair of defined chromosomal DSBs generated by rare-cutting endonucleases. Typically, a homing endonuclease, such as *I-SceI*, is used to cleave an exogenous reporter construct that has been randomly inserted into the genome. Many such constructs have been developed using reporters that fluoresce or confer drug resistance. Most are designed to detect, without recourse to structural analyses of the repaired locus, a single type of DSBR, including intrachromosomal HR (13,14), GT (15), SSA (16), mutNHEJ or accNHEJ (10,17–20) and MMEJ (21) (reviewed in 22). Simultaneous reporting for multiple DSBR pathways can be achieved by integrating multiple constructs into a single clone (20,23,24) or by integrating a single reporter construct that detects multiple types of repair (25,26).

DSBR mechanisms are known to be affected by chromosomal location (27–31) and by epigenetic factors (32–34) that are in turn affected by exogenous DNA sequences (35–37). It is, therefore, important to minimize or eliminate these variables when comparing different exogenous DSBR reporters. Differences in chromosomal location have been avoided by integrating different constructs at the same locus (16) or mitigated by averaging the behaviour of constructs at multiple integration sites (23). These precautions are rarely taken, however, and do not control for any differences between constructs in their effects on the host genome. For these reasons, single reporters that can detect multiple types of DSBR at a single locus are highly desirable. To date, the only reporters that achieve this without the need to sequence breakpoints are traffic light reporter (TLR) systems that elegantly combine multiple fluorescent marker genes to detect mutNHEJ and GT (25) or mutNHEJ, GT and SSA (26).

In the present study, we use an endogenous gene as a reporter to detect three types of nuclease-induced DSBR: mutNHEJ, GT and accNHEJ. Our reporter is the highly conserved X-chromosomal gene encoding hypoxanthine phosphoribosyltransferase (HPRT), an enzyme that is essential for the salvage pathway of purine biosynthesis, and deficiencies of which cause a range of clinical phenotypes, including Lesch Nyan disease (38). We chose *HPRT* not only because it is a model disease gene, naturally residing and expressed in growing cells, but also because it is drug selectable. Thus, cells with inactive or active *HPRT* alleles can be selected, respectively, in medium containing 6-thioguanine (6TG) or hypoxanthine, aminopterin and thymidine (HAT). Although *HPRT* has to be used extensively for studies of random mutagenesis and un-induced GT (39–43), it has not been systematically exploited for monitoring nuclease-induced DSBR. Here, we cleave *HPRT* alleles with *I-SceI* or Cas9 nuclease and estimate the relative frequency of three different DSBR outcomes. We also use an *I-SceI*-sensitive *HPRT* minigene to investigate how the frequency of gene correction is influenced by repair template design. As an additional tool, to facilitate studies of cell cycle control of DSBR, we describe novel *I-SceI* derivatives that are expressed in a cell cycle-dependent fashion.

Results

Silent introduction of an *I-SceI* site into a mouse *HPRT* minigene

The *I-SceI* recognition sequence (ATTACCCTGTTATCCCTA) translates into an amino-acid sequence (ITLLSL) with only one difference to a sequence (QTLLSL) in the normal HPRT enzyme. Assuming this change (Q144I) does not disrupt HPRT function, we introduced the *I-SceI* site into the exon-6-derived region of a mouse *HPRT* minigene (*mHPRT*) to generate *mHP-I-RT* (Fig. 1). As a control, we introduced a stop codon (*) immediately downstream of the *I-SceI* site, to generate *mHP-I*-RT*. When transfected into human *HPRT*$^{-}$ human fibrosarcoma (HT1080) cells, *mHPRT* and *mHP-I-RT* conferred resistance to HAT medium (HAT^{R})

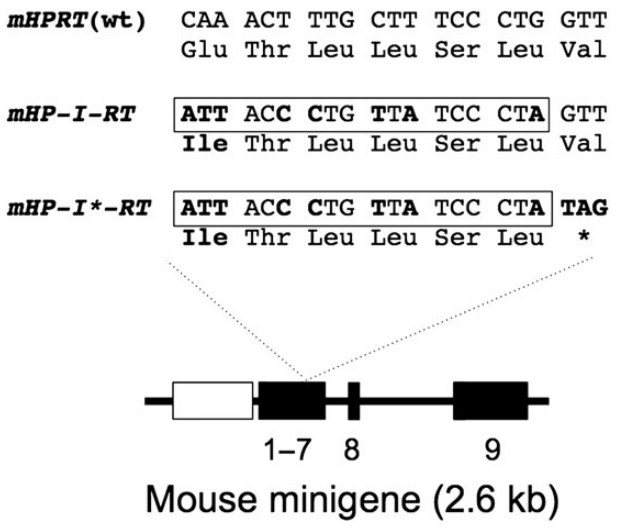


Figure 1. Mouse *HPRT* minigenes used in this study. Numbered black boxes represent exons. The mouse minigene is driven by the PGK promoter (grey box) and all but the last two introns are removed. Partial nucleotide, and amino-acid, sequences from exon 6 are shown for the wild-type (WT) allele and for alleles modified to carry an *I-SceI* site (boxed) with (m*HP-I*-RT*) or without (m*HP-I-RT*) an adjacent stop codon. Altered residues are shown in bold.

equally well, whereas *mHP-I*-RT* conferred no resistance (Table 1). This indicates that the Q145I mutation has little or no effect on HPRT function, which is consistent with structural and evolutionary considerations (Supplementary Material, Fig. S1).

I-SceI-induced correction and mutagenesis of a mouse *HPRT* minigene

To study *I-SceI*-induced gene correction by HR, the *mHP-I*-RT* minigene was stably integrated into an *HPRT*$^{-}$ cell line (HTtetSCE; Supplementary Material, Table S1) carrying a tetracycline-regulated *I-SceI* gene. One of the resulting clones with a single copy of *mHP-I*-RT* (clone m2.1) grew normally with or without tetracycline. Using this clone, we examined the efficiency of *mHPRT* gene correction by various templates: single- or double-stranded oligonucleotides or long double-stranded DNA (Fig. 2A). Each template had the wild-type sequence, an *I-SceI* site, or an *I-SceI* site and adjacent stop codon. Gene correction data are summarized (Fig. 2B) for oligonucleotide templates (left panel) and long templates (right panel). As expected, only templates without the stop codon gave rise to HAT^{R} colonies. Consistent with previous studies (15,44,45), correction by all repair templates was greatly stimulated by *I-SceI* (up to 800-fold). We also made the following observations (see Discussion for interpretations). First, sense and antisense oligonucleotides supported gene correction equally well. Second, single-stranded oligonucleotides were more effective than double-stranded oligonucleotides. Third, unlike correction with single-stranded templates, where increasing the homology length above 40 nt had little impact (46), increasing the homology length of double-stranded templates from 60 bp to

Table 1. Rescue of HPRT^{-} cells with mHP-I-RT

Transfected plasmid	Encoded minigene	Experiment 1[a]	Experiment 2	Experiment 3
pBT/PGK-HPRT (R1)	*mHPRT (wt)*	189	241	266
pBT/PGK-HP-I-RT (R1)	*mHP-I-RT*	198	213	259
pBT/PGK-HP-I*-RT (R1)	*mHP-I*-RT*	0	0	0

[a]In each experiment, plasmids expressing the indicated minigene were individually electroporated into 1 million HPRT^{-} cells which were then selected in HAT medium to generate the indicated numbers of HAT^{R} colonies.

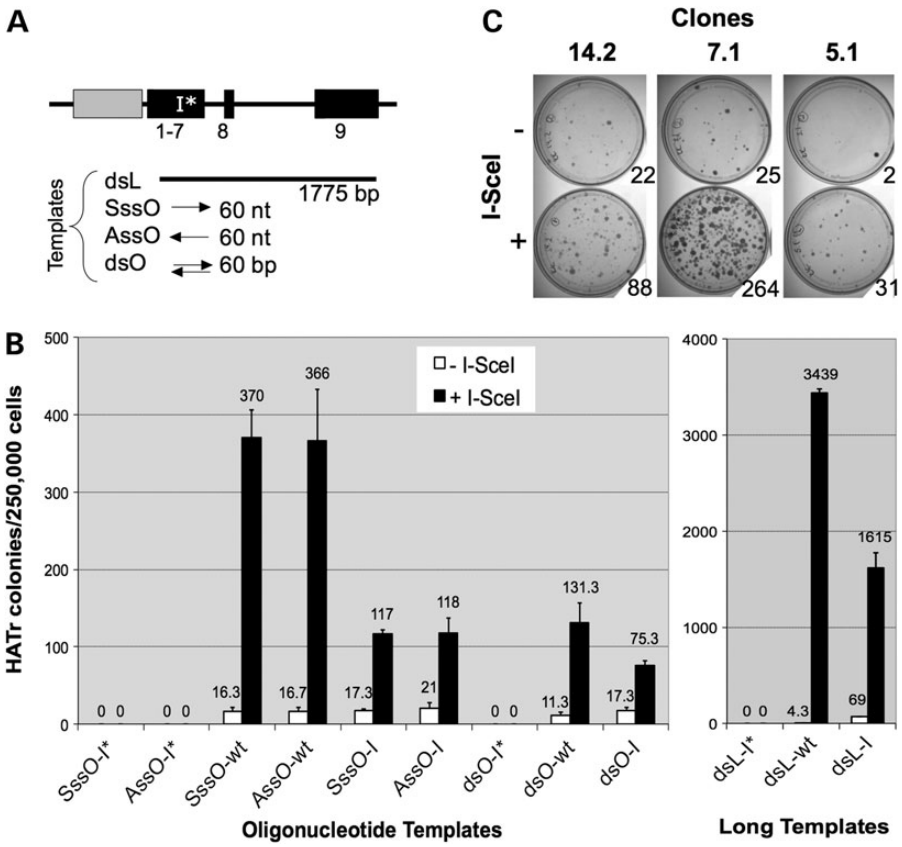


Figure 2. *I-SceI*-induced correction (**A** and **B**) or mutagenesis (**C**) of an integrated *mHPRT* minigene. (A) Gene correction templates are shown schematically, aligned with the target *mHP-I*-RT* minigene. Templates were sense or antisense single-stranded oligonucleotides (SssO, AssO), dsOs or dsL (for oligonucleotide sequences, see Supplementary Material, Table S3). Each of these four template classes was made with the *I-SceI* site and stop codon (-I*), the *I-SceI* site only (-I) or neither (-wt; wild-type). (B) *HPRT*$^{-}$ cells (clone 2.1) carrying the mHP-I*-RT minigene and a tetracycline-regulated *I-SceI* expression cassette were grown with or without tetracycline and transfected with the indicated templates detailed in (A). The frequencies of resulting HAT-resistant colonies are shown for oligonucleotides templates (left panel) and dsL templates (right panel). Error bars show SDs for three experiments. (C) Three HPRT+ clones (14.2, 7.1 and 5.1) carrying the *mHP-I-RT* minigene were transfected with *I-SceI* expression plasmid (+), or a vector control (–), and selected in 6TG. Colonies resulting from selection of 5×10^{5} cells are shown.

1.7 kb resulted in a 21- to 26-fold increase in gene correction [Fig. 2B; compare double-stranded oligonucleotide (dsO) and long double-stranded DNA (dsL)]. Fourth, despite having eight additional mismatches to the target, wild-type repair templates were always more effective than their equivalent templates with *I-SceI* sites. Finally, correction by long double-stranded templates was 9- to 26-fold more efficient than by oligonucleotides.

We next attempted to measure DSB-induced loss of *HPRT* function using clones with integrated *mHP-I-RT* minigenes. Three such HAT^R clones (14.2, 7.1 and 5.1; Supplementary Material, Table S1) were transiently transfected with *I-SceI* expression plasmid (pCMV3xnls-*I-SceI*) and selected in 6TG. For each clone, *I-SceI* expression clearly increased the frequency of 6TG-resistant (6TG^R) colonies (Fig. 2C), but frequencies varied greatly between clones. Such variations, as well as the appreciable frequencies of spontaneous 6TG^R colonies, are most likely result of chromosomal position effects.

Strategy for generating an *I-SceI*-sensitive human *HPRT*$^+$ allele

To avoid chromosomal position effects and establish a standard endogenous locus for analysing DSBR, we generated *I-SceI*-sensitive human *HPRT* alleles. The structure of the endogenous *HPRT* gene and *I-SceI*-sensitive derivatives is shown in Figure 3A. Our strategy for generating these and using them for accNHEJ and mutNHEJ assays is outlined in Figure 3B. First, a targeting construct (pLB-puro) was used to introduce a puromycin-resistance cassette into exon 6 (Fig. 3Ba). This cassette was flanked by *I-SceI* sites, allowing it to be excised by *I-SceI*. If excision is followed by accNHEJ of the two resulting chromosome ends, a functional *HPRT* allele carrying an *I-SceI* site in exon 6 is generated (Fig. 3Bb). The frequency of HAT^R colonies formed in this step, therefore, provides a measure of *I-SceI*-induced accNHEJ. (Like other accNHEJ reporter assays (10,17–20), this scheme involves the deletion of DNA between two *I-SceI* sites, but measures accNHEJ rather than mutNHEJ because the joined ends suffer no gain or loss of nucleotides.) Finally, the resulting HAT^R colonies can then be used to generate 6TG^R colonies as a measure of *I-SceI*-induced mutNHEJ (Fig. 3Bc).

I-SceI-induced accNHEJ at the human *HPRT* locus

HPRT$^+$ HT1080 cells were electroporated with pLB-puro and clones were selected in 6TG and puromycin (Fig. 3Ba). From a total of 20 million cells, 12 clones were obtained, consistent with the low un-induced targeting frequencies in HT1080 cells (47). One such clone (clone 5.2) was then electroporated with *I-SceI* expression plasmid, or mock transfected, and only the former transfection generated HAT^R clones (not shown). Southern analyses of a pool of >200 HAT^R colonies (clone 5.2P), two HAT^R clones (clones 5.2.1 and 5.2.2; both sensitive to puromycin) and parental clone 5.2 confirmed the presence of the *I-SceI* site in exon 6 (Supplementary Material, Fig. S2A and B). To estimate the frequency of accNHEJ, multiple lipofections of clone 5.2 with *I-SceI* expression plasmid were carried out and the frequency of resulting HAT^R colonies was measured at 4.1% [standard deviation (SD): 1.15; $n = 4$]. Among these HAT^R colonies, 4.4% (SD: 0.59; $n = 4$) remained Puro^R, presumably due to reintegration of the excised puromycin cassette at random genomic loci. These results support the use of clone 5.2, or equivalent clones, in assays for accNHEJ as outlined (Fig. 3Bb). Though highly specific, this assay detects only accNHEJ events between two DSBs, excluding the probably more frequent accNHEJ events involving only one DSB (see Discussion).

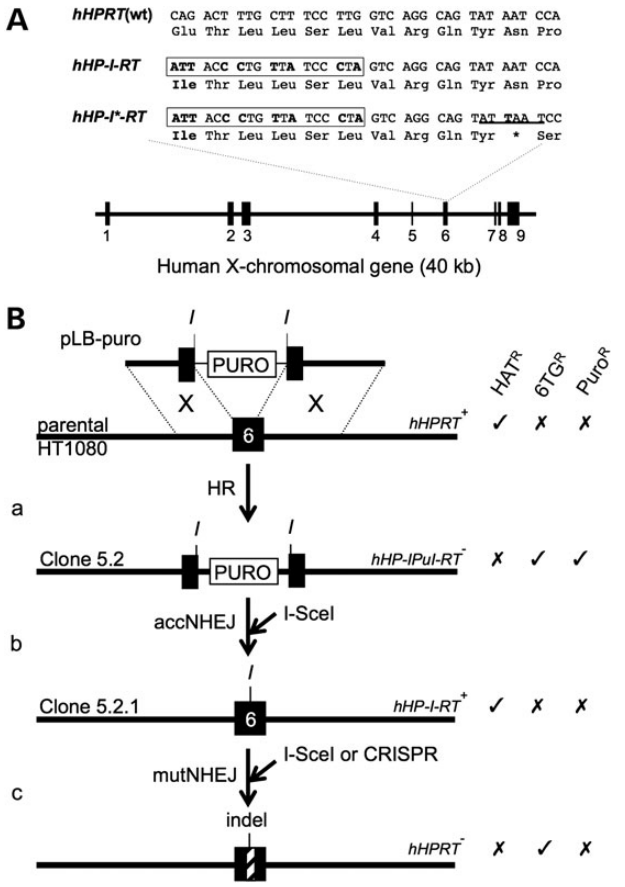
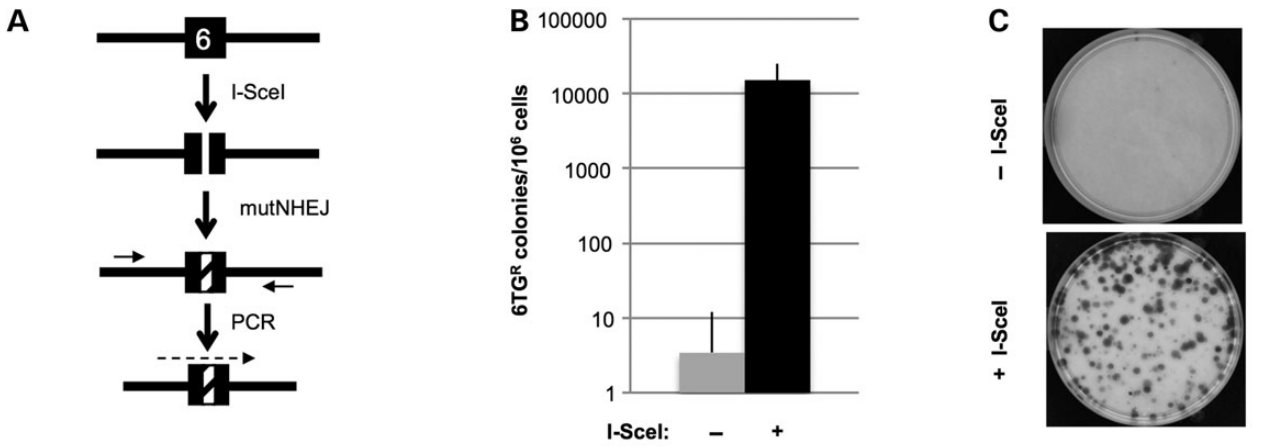


Figure 3. Generating and using *I-SceI*-sensitive *hHPRT* alleles. (**A**) Structure of parental and *I-SceI*-sensitive *hHPRT* alleles. Numbered black boxes represent exons. Partial nucleotide and amino-acid sequences for exon 6 are shown for the WT allele and for alleles modified to carry an *I-SceI* site (boxed), with (*hHP-I***-RT*) or without (*hHP-I-RT*) an adjacent stop codon. Altered residues are shown in bold. An *AseI* site in *hHP-I***-RT* is underlined. (**B**) Strategy for generating *I-SceI*-sensitive alleles and using them for accNHEJ and mutNHEJ assays. (**a**) Disruption of exon 6 by GT with pLB-puro. (**b**) Basis of accNHEJ assay. An *I-SceI*-induced DSB repaired by accNHEJ generates a functional *I-SceI*-sensitive allele. The frequency of resulting HAT^R colonies measures accNHEJ. (**c**) Basis of mutNHEJ assay. An *I-SceI*-induced DSB is repaired by mutNHEJ to generate an indel (hatched box). The frequency of resulting 6TG^R colonies measures mutNHEJ. Exon 6 (black box) and adjacent intronic DNA of the chromosomal *HPRT* locus (long lines) are shown. Targeting construct pLB-puro is shown with its puromycin-resistance cassette (PURO, white box) and its 2.7 and 3.1 kb arms aligned with homologous chromosomal regions (dotted lines) to allow HR (X). Sites for *I-SceI* (*I*) are shown. Alleles are labelled with their names (right) and the names of representative host cell lines (left). Resistance (✓) or sensitivity (✗) to key selective agents is indicated.

I-SceI-induced mutNHEJ at the human *HPRT* locus

When transfected with an *I-SceI* expression plasmid, clone 5.2.1 (*HP-I-RT*$^+$) generated 6TG^R colonies, as expected for DSBR by mutNHEJ resulting in indels (Fig. 4A). In standard lipofection protocols, 6TG^R colonies were generated at frequencies of 1.5% (SD: 0.96, $n = 8$), three to four orders of magnitude higher than the frequency of spontaneous 6TG resistance (Fig. 4B and C). To characterize the expected indels, the exon 6 regions of 28 individual 6TG^R clones were amplified and sequenced (Fig. 4A and D–F). The majority (17/28) had small indels involving <14 nt, all but two of which were simple deletions. The rest involved larger deletions (3/28), insertions (6/28) or both (2/28). These results support the use of the *I-SceI*-sensitive *HPRT*$^+$ allele in clone 5.2.1 (or equivalent clones), in assays to detect DSB-inducible mutNHEJ.



D

```
142                                 152
 T  M  I  T  L   L   S  L  V  R  Q
ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG
TGTTACTAATGGGAC↑AATA-GGGATCAGTCCGTC
```

E

#	Sequence	n	Type	Deletion			Insertion		Effect
				Size	Start	End	Size	Site	
1	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	2	SD	1	–4	–3	-	-	FS
2	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	1	SD	3	–4	–1	-	-	ΔL147
3	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	1	SD	3	0	3	-	-	ΔS148
4	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	1	SD	4	0	4	-	-	FS
5	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	2	SD	8	–8	0	-	-	FS
6	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	5	SD	8	–9	–1	-	-	FS
7	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	1	SD	9	–9	0	-	-	ΔL146-S148
8	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	1	SD	10	–3	7	-	-	FS
9	ACAATGATTACCCTG-TTAT↓CCCTAGTCAGGCAG	1	SD	14	–13	1	-	-	FS
10	ACAATGATTACCCTG-TTAT↓TATATCCCTAGTCAGG	1	SI	0	-	-	5	0	FS
11	CAATGATTACCCTGT-TTAT↓CCCTAGTCAGGCAG	1	SI+SD	12	–17	-5	1	–4	FS

F

#	Sequence		n	Type	Deletion			Insertion		Origin
	Left junction	Right junction			Size	Start	End	Size	Site	
12	tgatatagattccaga	AT↓CCCTAGTCAGGCA	1	LD	187	–181	+6	-	-	-
13	ttttggtgagaattac	caaaattaagaataag	1	LD	481	–150	+331	-	-	-
14	aggaggcaggaatttt	TATAATCCAAAGATGG	1	LD	787	–765	+22	-	-	-
15	na	na	1	LI	0	-	-	642	na	-
16	CCTG-TTATGTATTTTT	GATTAGAC↓CCCTAGTC	1	LI	0	-	-	104	0	Alu LINE
17	CCTG-TTATTAGACTTA	AGAGTTAT↓CCCTAGTC	1	LI	0	-	-	147	–1	Ch15
18	CCTG-TTACCTCAGCTA	AGGACGAT↓CCCTAGTC	1	LI	0	-	-	175	–1	Ch17
19	CCTG-TTATCACCACAG	TGGTTGAT↓CCCTAGTC	1	LI	0	-	-	245	0	pCMV-I-SceI
20	CCTG-TTACCTCAGCCT	AGGACGAT↓CCCTAGTC	1	LI	0	-	-	186	–2	pCMV-I-SceI
21	ACAATGATAATGTTTG	AGATGTTAaatatttt	1	LI+LD	77	–11	+66	90	-	alphoid DNA
22	gaaagGATGGGTGGAT	AGAAGACATTTAAATA	1	LI+LD	224	–37	+187	156	-	Ch7

Figure 4. Assays for *I-SceI*-induced mutNHEJ and DNA sequence analyses of resulting indels in *HPRT* exon 6. (**A**) *HPRT* exon 6 in clone 5.2.1 is cleaved by *I-SceI* and repaired by mutNHEJ to generate 6TG^R clones carrying exon 6 indels (hatched); these are amplified with flanking primers (short arrows) and sequenced (dashed arrow). (**B**) Frequency of 6TG^R clones generated from clone 5.2.1 lipofected with *I-SceI* expression plasmid or a vector control (mean and SD for six experiments are shown). Transfection was by lipofectamine 2000, but similar results were obtained with Fugene. (**C**) Example of 6TG^R colonies forming in the experiment in B. Petri dishes (10 cm diameter) were seeded with 10^5 cells before selecting in 6TG. (**D**) Part of the exon 6 DNA sequence in clone 5.2.1 is shown aligned with encoded amino acids 142–152. The recognition site for *I-SceI* is shown (bold) with positions of staggered nicks it makes (arrows) and un-cleaved phosphodiester bonds opposite each nick (hyphens). (**E**) Small indels. Coding strand DNA sequences of 11 indels identified in seventeen 6TG^R clones are shown with deleted and inserted residues highlighted in grey and black, respectively. The number of clones for each indel is shown (*n*). Positions of indels are summarized with residues numbered relative to the coding strand nick. Micro-homologies likely to have mediated deletion formation are underlined. Indel types: small and large deletions (SD, LD), or insertions (SI, LI). The predicted effects on protein coding are categorized as frame-shifts (FS) or in-frame deletions (Δ). (**F**) Large indels. Sequences of 11 junction regions are shown with DNA represented as in (B). Except for insertions, upper and lower case letters represent exon 6 and intronic residues, respectively. Type and position numbers of indels are summarized as in (B). The origin of the inserted DNA is indicated. na, no sequence available.

I-SceI-induced correction of the human *HPRT* gene

As a model for nuclease-induced gene correction by HR (gc/HR), we used a GT plasmid (pJB1-Ase) to generate clones with an inactive *HPRT* allele (*HP-I*-RT*$^-$) in which the silent *I-SceI* site in exon 6 has an adjacent stop codon and *AseI* site (Figs 3A and 5Aa; Supplementary Material, Fig. S2C). Ten million *HPRT*$^+$ cells electroporated with pJB1-Ase generated five 6TG^R colonies, four of which (clones 3A, 3B, 10B1 and 10B3) had *AseI* fragments indicative of the desired modification (Supplementary Material, Fig. S2C). Southern analyses confirmed the presence of an *I-SceI* site in exon 6 in each of three clones tested (3B, 10B1 and 10B3; Supplementary Material, Fig. S2A). Clone 3B was then used for

gc/HR assays in which a repair construct (pJB2) was co-lipofected with an *I-SceI* expression plasmid (or vector control), followed by selection in HAT (Fig. 5Ab). After normalizing for plating efficiencies, the frequency of HAT^R colonies, and therefore gc/HR, was found to be 0.16% (SD: 0.07; $n = 6$). Representative results for one experiment are shown in Figure 5B. Further experiments (not shown) confirmed that gc/HR requires co-transfected repair template and that clones 3A and 10B1 behaved similarly to clone 3B. These results support the use of the *I-SceI*-sensitive $HPRT^{-}$ allele in clone 3B (or equivalent clones) in assays to detect DSB-inducible gene correction and show that this is ~10-fold less efficient than repair of the same DSB by mutNHEJ.

Cell cycle-specific control of *I-SceI*-induced DSBR

Cell cycle studies typically rely on synchronization methods that use drugs or serum starvation to arrest the cycle or physical separation by centrifugal elutriation or flow cytometric cell sorting. Problems with these methods include the induction of cell stress responses, variable sensitivities to serum starvation and the need for expensive equipment with experienced operators (48). We reasoned that, in the context of DSBR, the expression of nucleases in a cell cycle-restricted fashion might provide a simple alternative to such approaches. We, therefore, used Fucci technology (49) to develop plasmids encoding fluorescent *I-SceI* derivatives that accumulate preferentially in G1/S (pSce-Cy-G1) or S/G2 (pSce-Cy-G2). These encode *I-SceI* fused to AmCyan fluorescent protein (Cy) with C-terminal peptides derived, respectively, from Cdt1 and Geminin (Fig. 6A; Supplementary Material, Table S2). Analyses of DNA content in transiently transfected AmCyan-positive ($AmCyan^+$) cells confirmed that pSce-Cy-G1- and pSce-Cy-G2-transfected cells were enriched in G1/S and S/G2, respectively (Fig. 6B). To investigate whether the constructs have differential effects on previously established DSBR assays, they were transfected into cells stably transfected with the reporters DR-GFP (13) and SA-GFP (16) for measuring, respectively, *I-SceI*-inducible intrachromosomal HR and SSA (Fig. 6C). The DR-GFP signal was ~2-fold greater in cells expressing pSce-Cy-G2 than in cells expressing pSce-Cy-G1, consistent with the known restriction of HR to S/G2. Conversely, the SA-GFP signal was ~2-fold greater in cells expressing pSce-Cy-G1 than in cells expressing pSce-Cy-G2, consistent with studies suggesting that SSA does not occur in G2 (50) and that resection can occur in G1 (51). These results suggested that our Fucci *I-SceI* constructs are able to introduce DSBs in restricted phases of the cell cycle, as intended. To assess how the Fucci *I-SceI* constructs affect mutNHEJ and gc/HR at the *hHPRT* locus, we then transfected them into clones 5.2.1 ($HP\text{-}I\text{-}RT^+$) and 3B ($HP\text{-}I^*\text{-}RT^-$), respectively (Fig. 6D). Similar to the DR-GFP assays, pSce-Cy-G2 induced gc/HR with ~3-fold greater efficiency than pSce-Cy-G1. Conversely, pSce-Cy-G2 induced mutNHEJ ~3-fold less efficiently than pSce-Cy-G1 suggesting that, although NHEJ operates throughout the cell cycle, mutNHEJ occurs preferentially in G1/S in HT1080 cells. Together, these observations support the use of our modified *I-SceI* expression plasmids for studying the influence of cell cycle on any DSB-inducible cell response.

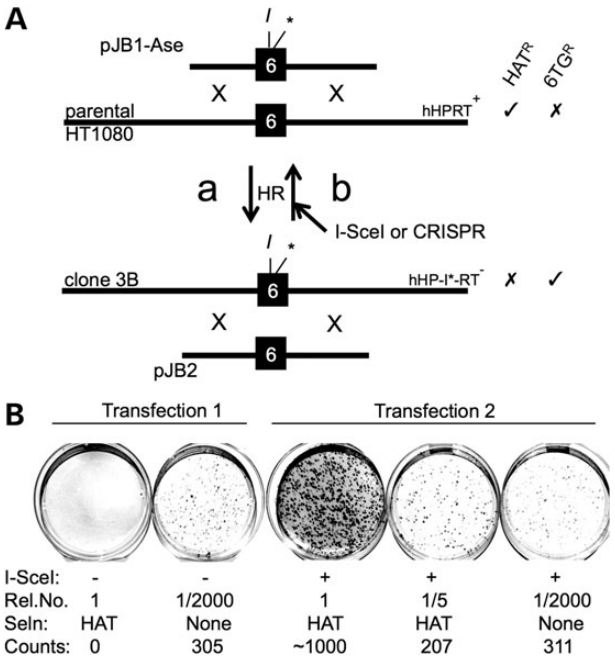


Figure 5. Generating and using *I-SceI*-sensitive $hHPRT^+$ alleles for gc/HR assays. (**A**) DNA is represented as in Figure 3. (**a**) HR between a wt *HPRT* allele and the targeting construct pJB1-*AseI* was used to introduce an *I-SceI* site and adjacent stop codon (*) into exon 6. The resulting cells (e.g. clone 3B) were selected in 6TG. (**b**) Basis of assays for gene correction by HR (gc/HR). Clone 3B cells are transfected with a repair template (pJB2) and nuclease expression plasmid(s). HR with transfected pJB2 regenerates a functional exon 6 and the frequency of resulting HAT^R colonies measures gc/HR. (**B**) Example of one experiment to estimate gc/HR frequencies. Clone 3B cells were co-transfected with pJB2 and either vector DNA (Transfection 1) or *I-SceI* expression plasmid (Transfection 2), and the indicated relative numbers (Rel. No.) of cells were placed in petri dishes to select in HAT for gene correction events or to determine plating efficiencies (no selection). Colonies appearing after ~10 days were counted and used to estimate gc/HR frequencies of <0.00016% (Transfection 1) and 0.17% (Transfection 2).

Comparison of Cas9- and *I-SceI*-induced DSBR at *HPRT* exon 6

To develop *HPRT*-based DSBR assays that do not rely on introducing an *I-SceI* site into the genome, and to compare *I-SceI* and CRISPR-induced DSBR pathways, we designed four exon 6-specific guide RNAs for the Cas9 nuclease of *Streptococcus pyogenes*: two (gRNA-1 and -2) targeting wild-type *HPRT* alleles and two (gRNA-3 and -4) targeting *I-SceI*-sensitive *HPRT* alleles (Fig. 7A). To test for Cas9-induced mutNHEJ, wild-type cells or *I-SceI*-sensitive $HPRT^+$ cells (clone 5.2.1) were co-transfected with expression plasmids for Cas9 and each of the four gRNAs, and $6TG^R$ colonies were selected. For a positive control, the *I-SceI* expression plasmid was transfected in parallel. The results (Fig. 7B, Table 2) clearly showed the expected gRNA targeting preferences. Thus, high frequencies of $6TG^R$ colonies (up to 3.6-fold higher those induced by *I-SceI* in clone 5.2.1) were induced by gRNA-1 and -2 in wild-type cells, and by gRNA-3 and -4 in clone 5.2.1 cells. Conversely, no $6TG^R$ colonies were generated in wild-type cells by gRNA-1 or in clone 5.2.1 by gRNA-3 or -4. These results indicate stringent allele-specific exon 6 mutagenesis directed by gRNAs 1, 3 and 4. Mutagenesis directed by gRNA-2, though highly efficient in clone 5.2.1, was also appreciable in parental cells. This most likely reflects the fact that gRNA-2 has only one mismatch to the wild-type allele and so can direct Cas9-mediated cleavage, albeit less efficiently than at the *I-SceI*-sensitive allele.

We sequenced the exon 6 region in a range of $6TG^R$ clones generated in the experiment of Figure 7B and in similar experiments (Supplementary Material, Fig. S3). For both Cas9 and *I-SceI*, indels occurred at the expected sites. We classified these indels, and those of Figure 4, as small or large deletions or insertions, or other more complex events. We noted a striking difference in

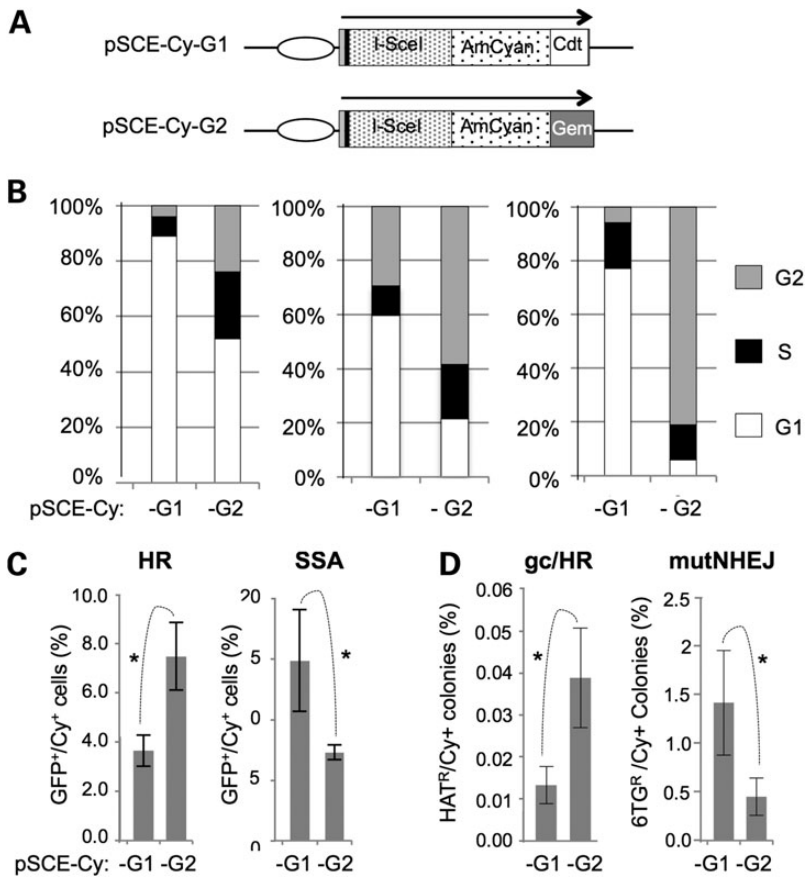


Figure 6. Development and use of cell cycle-specific *I-SceI*. (**A**) Schematic representation of *I-SceI* expression plasmids. Ellipse, cytomegaloviral promoter; black bar, haemagglutinin tag. (**B**) Cell cycle profiles of cells lipofected with the indicated plasmids. Cells were HT1080-WT (left), HT1080-AS5 (Supplementary Material, Table S1; centre) and HeLa (right). (**C**) Differential effects of cell cycle-restricted *I-SceI* on reporters for HR and SSA. The indicated plasmids were transfected into cells (HT-DR-GFP-8 or HT-SSA-GFP; Supplementary Material, Table S1) carrying GFP-based reporters for a HR or SSA. The percentage of AmCyan$^+$ cells that were GFP$^+$ is shown. (**D**) Differential effects of cell cycle-restricted *I-SceI* on mutNHEJ and gene correction at the *HPRT* locus. Assays based on clone 5.2.1 (mutNHEJ) or clone 3B (gene correction) were used with the indicated *I-SceI* expression plasmids. Results were normalized to the amount of AmCyan expression. Mean and SDs for three or more assays are presented with significant differences marked (*$P < 0.05$; Student's *t*-test).

the frequency of small insertion mutations generated by *I-SceI* and Cas9 (Fig. 7C): small insertions represented the least abundant class of *I-SceI*-generated mutations (1/38; 2.6%), but the most abundant class of Cas9-generated mutations (20/58; 34.5%).

Finally, we tested gRNA-1, -3 and -4 for their ability to support Cas9-induced gene correction in *I-SceI*-sensitive *HPRT*$^-$ cells (clone 3B). As expected, gRNA-3 and -4, but not gRNA-1, supported Cas9-mediated gene correction (Table 3). Also as expected, detectable levels of correction induced by gRNA3/Cas9 depended, as for *I-SceI*, on template co-transfection. Interestingly, Cas9-induced gene correction was ~7-fold less efficient than *I-SceI*-induced correction. This contrasted with Cas9-induced mutNHEJ that was on average 1.7-fold more efficient than *I-SceI*-induced mutNHEJ (Table 2). Together with the differences in mutNHEJ mutation signatures (Fig. 7C), these results suggest that Cas9- and *I-SceI*-generated DSBs are processed differently during DSB repair.

Discussion

In this study, we have developed a system for measuring the repair of *I-SceI*- or Cas9-induced DSBs by three different mechanisms, all at the same site in the endogenous *HPRT* gene. We have used these assays to estimate the absolute and relative frequencies of the different DSBR mechanisms, to discover that *I-SceI* and Cas9 induce distinct mutNHEJ mutation signatures and mutNHEJ:GT ratios, and to validate novel *I-SceI* derivatives for studying DSBR in different cell cycle phases. In addition, we have used an *I-SceI*-sensitive *HPRT* minigene to compare the efficiencies of different DNA repair templates.

Advantages and limitations of DSBR assays using the endogenous *HPRT* locus

A key attraction of the endogenous *HPRT* assay system is its ability to measure three DSBR pathways (GT, mutNHEJ and accNHEJ) at the same site. As outlined in the Introduction, the comparison of multiple DSBR pathways at a single locus is preferable to the comparison of single pathways at multiple loci. This is because the interactions between a reporter and its chromosomal environment, and any associated effects on DSBR, vary unpredictably between loci. Several studies have compared two DSBR pathways at a single locus but, to our knowledge, detection of three pathways at a single locus has been described only once, for a different combination of pathways (GT, mutNHEJ and SSA) using a TLR construct (26).

For mutNHEJ, the endogenous *HPRT* system has the advantage of being able to detect a full spectrum of indels. Thus, by

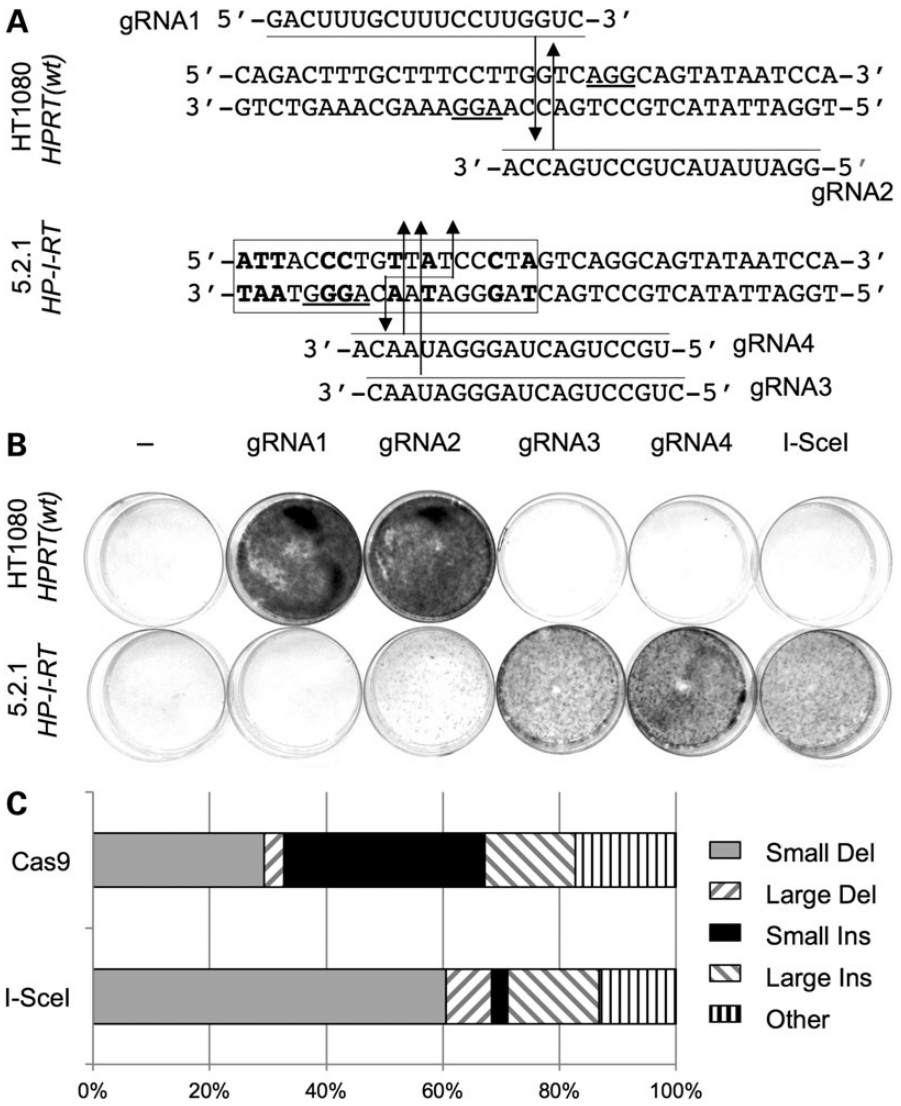


Figure 7. Cas9-induced *HPRT* mutagenesis. (**A**) Design of exon 6-targeted gRNAs. Part of the exon 6 sequence in WT cells or clone 5.2.1 is shown. The *I-SceI* site is boxed with non-WT residues shown in bold and the cleavage site indicated (staggered arrows). Target-specific regions of gRNAs are shown, aligned with their target. For each gRNA, the predicted Cas9 cleavage site is indicated by an arrow, 3 nt upstream of the protospacer-adjacent motif (underlined). (**B**) Stained petri dishes after selection of transfected HT1080 or clone 5.2.1 cells (10^6 cells per dish) in 6TG. Cells were lipofected with DNA encoding no nuclease (column 1), Cas9 and the indicated gRNA (columns 2–5), or *I-SceI* (column 6). Cells from the same transfection shown in (B) were plated at lower dilutions and selected in 6TG and used to determine frequencies (Materials and Methods; Table 3, Experiment 1). (**C**) Relative proportions of different classes of *I-SceI*- or Cas9-induced mutations (data from Fig. 4D and E; Supplementary Material, Fig. S3).

Table 2. Relative frequencies of *HPRT* mutagenesis by Cas9 and *I-SceI*[a]

Expt.	Host cells	*HPRT* genotype	Nuclease None	gRNA1/Cas9	gRNA2/Cas9	gRNA3/Cas9	gRNA4/Cas9	*I-SceI*
1	WT	*HPRT*$^+$	UD	1.5	1.7	UD	UD	UD
1	5.2.1	*HP-I-RT*$^+$	UD	UD	0.05	1	3.6	1
2	5.2.1	*HP-I-RT*$^+$	nt	UD	nt	1.3	nt	1
3	5.2.1	*HP-I-RT*$^+$	nt	nt	nt	0.5	1.4	1
4	WT	*HPRT*$^+$	UD	2.6	nt	UD	UD	UD
4	5.2.1	*HP-I-RT*$^+$	UD	UD	UD	1.6	2.9	1

UD, undetectable ($<10^{-5}$); nt, not tested.
[a]Figures show the frequencies of $6TG^R$ colonies generated by the indicated nucleases relative to frequencies generated by *I-SceI*.

cleaving an essential *HPRT* coding region, we detected single nucleotide indels and in-frame deletions of just a single codon (Fig. 4D). Furthermore, we detected indels of over 600 bp (Fig. 4F) closer in size the *HPRT* mutations commonly detected after ionizing irradiation (43). Most mutNHEJ assays use gain-of-function reporters whose reading frame must be restored so

Table 3. Frequencies of gc/HR induced by different HPRT-targeted nucleases[a]

Nuclease	*I-SceI*	*I-SceI*	gRNA1/ Cas9	gRNA3/ Cas9	gRNA3/ Cas9	gRNA4/ Cas9
Template	–	+	+	–	+	+
Experiment 1	UD	0.093	UD	UD	0.013	nt
Experiment 2	nt	0.12	nt	nt	0.016	0.019

UD, undetectable; nt, not tested.
[a]Figures show % of colony forming units that became HAT^R after lipofection of clone 3B with the indicated nucleases with or without repair template (pJB2).

that, at best, only one-third of all possible indels can be detected. Furthermore, many large indels are expected to prevent reporter cassette activation, for example, by deleting into its promoter or coding sequences, or by the inserting exogenous sequences that impair expression. In a loss-of-function mutNHEJ assay based on the endogenous thymidine kinase gene, *I-SceI* cleaved a non-coding region and so was unable to detect all indels (52).

Like other reporters, the endogenous *HPRT* system requires two *I-SceI* sites to detect accNHEJ, but it is unusually selective because, as shown in our mutNHEJ assays, even the smallest indels are incompatible with HAT resistance. This contrasts with other reporter systems where small indels were compatible with reporter gene expression, so that DNA sequence analyses were required to distinguish accNHEJ from mutNHEJ (10,17–20). In common with most other accNHEJ reporters, however, our *HPRT*-based assay underestimates the true level of accNHEJ, partly because multiple cycles of accNHEJ and re-cleavage increase the chance that DSBR will ultimately occur by mutNHEJ. This issue was addressed in another study by use of an exogenous reporter in which accNHEJ of two mutant *I-SceI* generates a site resistant to further cleavage (10), a feature that could readily be incorporated into the *HPRT* assay. Nevertheless, another cause for accNHEJ underestimation, lack of reporter activation by accNHEJ at only one of the two *I-SceI* sites, remains to be addressed in any accNHEJ assay.

Assaying for GT at the cleaved *HPRT* locus is, like the mutNHEJ and accNHEJ assays, highly selective. *HPRT* can, therefore, report for these different DSBR mechanisms without the need for DNA sequencing of multiple clones, an advantage shared with TLRs. Although selection and scoring of drug-resistant colonies is less convenient than analyses of fluorescent reporters, the *HPRT* system does not require a multicolour flow cytometer and conveniently delivers clones ready for any further analyses that might be required. Furthermore, the *HPRT* assay is highly sensitive, easily capable of unambiguously detecting one event among a million or more cells.

The fact that *HPRT* is an endogenous housekeeping gene also has attractions. First, use of *HPRT* avoids the need to integrate an exogenous reporter into the genome and to demonstrate single copy integration. Indeed, measurements of Cas9-induced mutNHEJ can be carried out without any preparatory modifications of the host cells. Although *HPRT* modifications, such as the introduction of an *I-SceI* site, are required for measuring accNHEJ and GT, the reagents and selection steps we have described should be applicable in most cell lines. Second, as outlined above, the fixed genomic location of *HPRT* is an advantage if one wishes to compare DSBR in different cell types. Similarly, the conservation of *HPRT* allows for well-controlled comparisons of DSBR between cells from different species. Although a fixed location precludes studies of chromosome position effects on DSBR, for this purpose, an *HPRT* minigene or bacterial artificial chromosome carrying the whole *HPRT* gene can be ectopically integrated, with the endogenous locus serving as a reference. Third, the endogenous *HPRT* system allows DSBR to be studied at a natural gene with features such as multiple introns and repeat sequences intact and unmodified by exogenous sequences that may affect DSBR.

Relative frequencies of different DSBR pathways

Our estimated efficiencies of *I-SceI*-induced mutNHEJ and GT the endogenous *HPRT* locus in HT1080 (human fibrosarcoma) cells were 1.5 and 0.16%, respectively, giving a mutNHEJ:GT ratio of ~9:1. In a previous study, where different chromosome positions of the mutNHEJ and GT reporters were taken into account, the *I-SceI*-induced mutNHEJ:GT ratio was estimated at random loci in HCA-2 (human colon cancer) cells to be 6:1 (23). In another study, where chromosomal position effects were avoided by use of a TLR, the average mutNHEJ:GT ratios were estimated at transcribed and untranscribed loci in HEK293 (human embryonic kidney) cells to be 4:1 and 25:1, respectively (26). For comparison with our estimate, these estimates must be increased 3-fold (because their mutNHEJ reporters detect only in-frame indels), to give ratios of 18:1, 12:1 and 75:1, respectively. Thus, despite the differences in variables such as cell type and DNA delivery method, our estimate of 9:1 is in good agreement with the value of 12:1 corresponding to a pool of transcribed loci derived from the study of Kuhar *et al*. This suggests that the endogenous *HPRT* locus is representative of an average transcribed locus, when measuring DSBR.

Our estimated efficiencies of *I-SceI*-induced accNHEJ and mutNHEJ at the *HPRT* locus in HT1080 cells were 4.1 and 1.5%, respectively, giving an accNHEJ:mutNHEJ ratio of ~3:1. As discussed above, although our accNHEJ and mutNHEJ assays are particularly specific, they are also likely, in common with other assays, to underestimate this ratio, partly because of re-cleavage of accurately repaired DSBs. Thus, our estimate is a robust only as a minimum estimate; true values require further investigation and are likely to be much higher. Perhaps the best estimate to date (~50:1) used an accNHEJ substrate that cannot be re-cleaved by *I-SceI* after accNHEJ (10). Further estimates are required that not only compensate for re-cleavage by nuclease but also detect accNHEJ at a single DSB.

Differences between *I-SceI*- and Cas9-induced DSBR

It was surprising that, compared with *I-SceI*, Cas9 induced higher frequencies of mutNHEJ and lower frequencies of GT, even though both nucleases targeted the same regions in *HPRT* exon 6. It was also striking that Cas9, unlike *I-SceI*, had a marked propensity to generate small (mostly 1 nt) insertions. A high proportion of single nucleotide insertions among Cas9-induced indels has been noted previously (53,54), but its significance is unclear. It is possible that these differences between Cas9- and *I-SceI*-induced DSBR reflect the different DNA ends they generate. Thus, the recessed 5′ ends generated by *I-SceI* may be more prone than blunt Cas9-generated ends to undergo further 5′ end-resection, favouring RAD51 filament formation and GT at the expense of mutNHEJ. This interpretation is also consistent with recent data showing that a customized nuclease that generates protruding 5′ ends stimulated GT less efficiently than cleavage of the same reporter by *I-SceI* (26). The high proportion of single base insertions among Cas9-induced indels is accompanied by a decrease in the proportion of small deletions (Fig. 7C). This could be explained if blunt ends are less susceptible to nuclease

digestion than 3′-overhangs, and so more susceptible to nucleotide addition. Consistent with this, chromosomal translocations induced by wild-type Cas9 involve deletions less frequently than natural translocations or translocations induced by paired Cas9 nickases (55). Further investigations of how DSBR pathways are affected by different nuclease and DSB end-structures will be valuable for elucidating natural mutagenesis and optimizing genome engineering methods.

The ease with which CRISPRs can be developed will facilitate further analyses of DSBR at various regions in the *HPRT* locus. The ability to place a silent *I-SceI* site in exon 6 will nevertheless remain a valuable option for studies of DSBR using *HPRT* as a reporter. For example, it will allow comparisons of DSBR at the *HPRT* locus in different species using a single defined nuclease, or studies of the influence of effector proteins, targeted to the *HPRT* locus by fusion to a nuclease-inactive Cas9, on *I-SceI*-induced DSBR.

Cell cycle regulated *I-SceI*

Fusion of Geminin- or Cdt1-derived peptides to AmCyan-tagged *I-SceI* allowed us to induce DSBs preferentially in S/G2 and G1/S, respectively. This approach offers a relatively simple and non-invasive way to study how any DSB-induced cellular response varies with cell cycle and adds to previous modifications of *I-SceI* designed to extend its versatility as a tool in the study of DSBR (24,56,57). So far, we have used our approach to demonstrate cell cycle regulation of intrachromosomal HR, SSA, gc/HR and mutNHEJ, but it will be interesting to extend this to other DSBR pathways, including accNHEJ. It should be noted that DSBR assays preferentially activated by a phase-specific *I-SceI*, while indicative of phase-specific DSBR, do not always indicate that DSBR occurs in that phase. In cells with defects in NHEJ and cell cycle checkpoints, for example, DNA damage inflicted in one cell cycle phase may be repaired in a subsequent phase (58).

A similar approach to that described here may be applicable to CRISPRs with potential implications for genome engineering. For example, the use of S/G2-restricted nucleases may provide a way to enhance the proportion of any desired HR-mediated genomic modifications at the expense of unwanted NHEJ-mediated events. In our study, AmCyan$^+$ cells were enriched in G1/S or G2/S phases, but the absolute numbers of AmCyan$^+$ phase-enriched cells were limited. It will, therefore, be important to optimize levels of *I-SceI*-fusion protein expression if these reagents are to be fully exploited.

Use of a *HPRT* minigene reporter

Although our initial experiments involved an ectopic mouse *HP-I*-RT* minigene and were not used to measure absolute frequencies of gene correction, they did generate several interesting results concerning the mechanisms of GT efficiencies using different templates. First, we measured similar frequencies of DSB-induced gene correction for sense and antisense single-stranded oligonucleotides (ssOs). This could not have been predicted from a similar result for un-induced ssO-directed modifications (59), which are thought to involve ssO annealing at replication forks (60,61) rather than at resected DSBs, as is envisaged for nuclease-induced DSBR (62). Our results, therefore, suggest that resection and oligonucleotide annealing on each side of the DSB are equally efficient. Second, the higher efficiency of ssO- relative to dsO-templated correction suggests that ssO annealing at resected DSB ends is preferred over invasion of the dsO by the resected ends, as is envisaged for long double-stranded templates (see 63). Third, the >20-fold increase in GT on increasing the ds template length from 60 to ~1700 bp is of note. Un-induced GT efficiencies increase exponentially with template length (64), probably because the chance that a random chromosome DSB occurs within the homology region increases with template length. For nuclease-induced GT, however, this explanation does not apply. It may be, therefore, that increasing homology favours homology searching and/or strand invasion, or that long double-stranded templates are less susceptible to degradation than double-stranded oligonucleotides. Fourth, the greater effectiveness of wild-type templates compared with equivalent templates with *I-SceI* sites was observed for both double- and single-stranded templates. This suggests that re-cleavage of the repaired chromosome by *I-SceI* is more of a barrier to gene correction than cleavage by *I-SceI* of double-stranded templates prior to HR. Finally, the greater efficiency of GT with long double-stranded templates than with oligonucleotides contrasts with other studies where ssOs supported nuclease-induced gene modifications with similar or greater efficiencies than plasmid templates (46,65,66). The reason for this difference is unclear, but it may reflect differences in optimal experimental conditions required for different loci and cell types. For example, although our oligonucleotide concentrations were similar to those used by others (46), they were not necessarily optimal. Indeed, increasing oligonucleotide concentrations enhanced correction frequencies up to 4-fold (Supplementary Material, Fig. S3).

In summary, we have demonstrated the versatility of using *HPRT* to investigate nuclease-induced DSBR at an endogenous locus. Although nucleases have previously been used to induce DSBR at endogenous selectable genes (52,67–71), our study uniquely combines the analysis of three different types of DSBR at an endogenous locus, the use of *I-SceI* to cleave a functional exon (allowing a fuller range of DSBR events to be selected) and the comparison of *I-SceI*- and Cas9-induced DSBR. The use of an exogenous *HPRT* minigene to study gene correction and of modified *I-SceI* to study cell cycle-specific DSBR adds further novelty and utility to this study. These systems will be applicable in a wide range of investigations into DSBR.

Materials and Methods

Cell culture and transfection

HT1080 fibrosarcoma cells and derivatives and HeLa cells were grown as previously described (72). Supplementary Material, Table S1 summarizes key properties of various HT1080 derivatives used. For electroporations, a Gene Pulser (BioRad) was used as previously described (72). For lipofections, lipofectamine 2000 (LF; Life Technologies) or Fugene 6 (Promega) was used according to the manufacturers' instructions, in multiwall well formats with, per cm^2, 0.5–1 million cells, a total of 0.4 µg (LF) or 0.6 µg (Fugene) DNA and 1 µl (LF) or 1.8 µl (Fugene) cationic lipid reagent. Under these conditions, 70–90% of HT1080 cells lipofected with a green fluorescent protein (GFP) expression plasmid became GFP$^+$ 24–48 h post-lipofection. For nucleofection, an Amaxa Nucleofector I (Lonza) was used according to the manufacturer's recommendations (program L005, transfection solution T).

Drug selections and colony counts

To select for stably transfected/modified clones or populations, the following drug selections were used: hygromycin B (Invitrogen; 100 µg/ml), zeocin (Invivogen 200 µg/ml), puromycin (Invivogen; 0.4 µg/ml), HAT supplement (Life technologies) and 6TG (Sigma; 15 µg/ml). After transfection, but prior to selecting in

6TG, cells were maintained without selection for 5 days, passaging as necessary. Other selective agents were added to transfected cells 48 h post-transfection. Because 6TG and HAT selection are initiated 48 h or more after transfection, by which time most DSB is complete (14), any influence of HAT or 6TG on DSBR is minimized. To determine frequencies of HAT^R or $6TG^R$ cells, colonies were selected in 9 cm diameter dishes over a range of cell dilutions. To determine plating efficiencies, a high dilution was plated without selection. Colonies were stained with crystal violet as described (72). Dilutions generating appreciable but limited numbers of HAT^R or $6TG^R$ colonies (e.g. 50–300) were chosen for colony counting. To allow for plating efficiency variations, the number of selected colonies in a given dish was divided by the number of colony forming units in that dish, as calculated from the number of colonies that formed without selection at a known dilution of the same culture.

DNA manipulations and plasmids

Key features of the plasmids used in this study are summarized in Supplementary Material, Table S2. Standard recombinant DNA methods were used for plasmid construction. Correct assembly was confirmed by restriction enzyme mapping and/or DNA sequencing. Details of construction and/or plasmid maps are available on request. The following kind gifts were made—*mHPRT* minigene vector pBT/PGK-HPRT (73): David Melton (Edinburgh University); *I-SceI* expression plasmid pCMV3xnls-*I-SceI* (74) and reporter constructs DR-GFP (13) and SA-GFP (16): Maria Jasin (Memorial Sloan-Kettering Cancer Center) and expression constructs for gRNA (Addgene 41824) and Cas9 (Addgene 41815): George Church (Harvard Medical School, via Addgene). Cloning of gRNA into 41824 was as described (75).

Oligonucleotides, polymerase chain reaction and DNA sequencing

Oligonucleotides used as *mHP-I*-RT* repair templates are described in Supplementary Material, Table S3. To characterize indels, *hHPRT* exon 6 regions were amplified as follows. A 1433 bp region was amplified with primers P1 (5′-AGGGAACCCTTCTGTGTGTG-3′) and P2 (5′-GGACAATTCCCTATGCCTCA-3′) under the following conditions: 120 min at 94°C, then 30 cycles of 10 min at 94°C, 30 min at 60°C, 80 min at 72°C, then 120 min at 72°C. Alternatively, a 507 bp region was amplified with primers P3 (5′-GGCATTCT TACTGCTTGCTG-3′) and P4 (5′-TCTGCCATGCTATTCAGGAC-3′) under same conditions except with 30 min at 72°C instead of 80 min. Polymerase chain reaction products were purified and sequenced with primer P3 or P4.

Southern analyses

Use of genomic DNA (preparation, digestion, electrophoresis and transfer to nylon membranes) and probes (^{32}P-labelling and hybridizations to membranes) were essentially as described previously (72). The *hHPRT* probe was a 654 bp *Xba*I–*Bam*H1 fragment of *hHPRT* genomic DNA including exon 5 and part of intron 5. The puromycin probe was a 1.2 kb *Xba*I fragment from pBL-Puro/R (76).

Flow cytometry

Measurements of GFP fluorescence alone were made in a Facscalibur (Beckton Dickinson; 488 nm laser). Measurements of AmCyan fluorescence in combination with propidium iodide (PI) or GFP fluorescence were made in an LSRII flow cytometer (Becton Dickinson; 405 and 488 nm lasers). To determine cell cycle profiles, cells were washed in phosphate buffered saline A (PBSA) and fixed by drop-wise addition of 500 µl cold (–20°C) 70% (v/v) ethanol, with vortexing, and stored at –20°C. Fixed cells were washed in PBSA at 40°C and incubated in RNaseA (10 µl; 1 mg/ml) for 30 min on ice, then resuspended in PI (200 µl, 50 µg/ml). For transfected cells, PI fluorescence was measured in populations gated for AmCyan fluorescence.

mHPRT-based gene correction assays

Clone 2.1 cells were grown with (*I-SceI* off) or without (*I-SceI* on) tetracycline (Sigma; 1 µg/ml) for 48 h prior to nucleofection. Unless stated otherwise, each nucleofection used 250 000 cells with 5 µg of oligonucleotide or plasmid template. After nucleofection, cells were passaged for 5 days, with or without tetracycline. The entire population was then selected in HAT for 10–14 days and the resulting HAT^R colonies were counted.

hHPRT-based DSBR assays

To measure *I-SceI*-induced accNHEJ, clone 5.2 cells were lipofected with pCMV3xnls-*I-SceI* and the frequency of resulting HAT^R colonies was determined. To measure I-SceI-induced mutNHEJ, clone 5.2.1 was lipofected with pCMV3xnls-*I-SceI* and the frequency of resulting $6TG^R$ colonies was determined. The same protocol was used to measure CRISPR-induced mutNHEJ in clone 5.2.1 cell or wt HT1080 cells, except that equal weights of gRNA- and Cas9-expression plasmids were co-transfected. To measure *I-SceI*- or CRISPR-induced gc/HR, clone 3B was lipofected with equal amounts of pCMV3xnls-*I-SceI* and repair template (pJB2) or equal amounts of plasmids encoding gRNA, Cas9 and *I-SceI*. The frequency of resulting HAT^R colonies was then determined.

Assays for *I-SceI*-AmCyan-induced DSBR

For mutNHEJ and gc/HR assays induced by *I-SceI*-AmCyan fusions, part of the transfected culture was analysed 48 h post-lipofection to determine the proportion of $AmCyan^+$ cells (typically <1%). Frequencies of mutNHEJ and gc/HR were then divided by their corresponding fraction of $AmCyan^+$ cells. Cell lines HT-DRGFP-8 and HT-SAGFP-6 lipofected with *I-SceI*-AmCyan expression plasmids were analysed flow cytometrically 48 h post-lipofection. After gating on 50 000–100 000 AmCyan-expressing cells, the proportion that also expressed GFP was measured.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

This work would not have been possible without the research project component of B.Sc. or M.Sc. courses taken at Imperial College (J.T.B., L.A.B., A.K. and C.C.). We are grateful to James Elliot for his expert assistance with flow cytometric analyses.

Conflict of Interest statement. None declared.

Funding

This work was supported in part by the Biotechnology and Biological Research Council (BB/H003371/1 to A.C.G.P.), the Medical Research Council (MC_PC_12003 to T.C.H.), Cancer Research UK (C5255/A15935 to S.A.) and University of Oxford

(Clarendon Scholarship to S.A.). Funding to pay the Open Access publication charges for this article was provided by the Research Councils UK open access fund.

References

1. Kasparek, T.R. and Humphrey, T.C. (2011) DNA double-strand break repair pathways, chromosomal rearrangements and cancer. *Semin. Cell Dev. Biol.*, **22**, 886–897.
2. Liu, C., Srihari, S., Cao, K.-A.L., Chenevix-Trench, G., Simpson, P.T., Ragan, M.A. and Khanna, K.K. (2014) A fine-scale dissection of the DNA double-strand break repair machinery and its implications for breast cancer therapy. *Nucleic Acids Res.*, **42**, 6106–6127.
3. Mali, P., Esvelt, K.M. and Church, G.M. (2013) Cas9 as a versatile tool for engineering biology. *Nat. Methods*, **10**, 957–963.
4. Yáñez, R.J. and Porter, A.C. (1998) Therapeutic gene targeting. *Gene Ther.*, **5**, 149–159.
5. Li, H.L., Nakano, T. and Hotta, A. (2014) Genetic correction using engineered nucleases for gene therapy applications. *Dev. Growth Differ.*, **56**, 63–77.
6. Jasin, M. and Rothstein, R. (2013) Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.*, **5**, a012740.
7. Chiruvella, K.K., Liang, Z. and Wilson, T.E. (2013) Repair of double-strand breaks by end joining. *Cold Spring Harb. Perspect. Biol.*, **5**, a012757.
8. Yata, K. and Esashi, F. (2009) Dual role of CDKs in DNA repair: to be, or not to be. *DNA Repair (Amst.)*, **8**, 6–18.
9. Rothkamm, K., Krüger, I., Thompson, L.H. and Löbrich, M. (2003) Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell. Biol.*, **23**, 5706–5715.
10. Lin, W.Y., Wilson, J.H. and Lin, Y. (2013) Repair of chromosomal double-strand breaks by precise ligation in human cells. *DNA Repair (Amst.)*, **12**, 480–487.
11. Symington, L.S. and Gautier, J. (2011) Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.*, **45**, 247–271.
12. Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y. and Takeda, S. (2006) Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair (Amst.)*, **5**, 1021–1029.
13. 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.
14. Mao, Z., Bozzella, M., Seluanov, A. and Gorbunova, V. (2008) DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle*, **7**, 2902–2906.
15. Radecke, F., Peter, I., Radecke, S., Gellhaus, K., Schwarz, K. and Cathomen, T. (2006) Targeted chromosomal gene modification in human cells by single-stranded oligodeoxynucleotides in the presence of a DNA double-strand break. *Mol. Ther.*, **14**, 798–808.
16. Stark, J.M., Pierce, A.J., Oh, J., Pastink, A. and Jasin, M. (2004) Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol. Cell. Biol.*, **24**, 9305–9316.
17. Lin, Y., Waldman, B.C. and Waldman, A.S. (2003) Suppression of high-fidelity double-strand break repair in mammalian chromosomes by pifithrin-alpha, a chemical inhibitor of p53. *DNA Repair (Amst.)*, **2**, 1–11.
18. Lin, Y., Lukacsovich, T. and Waldman, A.S. (1999) Multiple pathways for repair of DNA double-strand breaks in mammalian chromosomes. *Mol. Cell. Biol.*, **19**, 8353–8360.
19. Dahm-Daphi, J., Hubbe, P., Horvath, F., El-Awady, R.A., Bouffard, K.E., Powell, S.N. and Willers, H. (2005) Nonhomologous end-joining of site-specific but not of radiation-induced DNA double-strand breaks is reduced in the presence of wild-type p53. *Oncogene*, **24**, 1663–1672.
20. Guirouilh-Barbat, J., Huck, S., Bertrand, P., Pirzio, L., Desmaze, C., Sabatier, L. and Lopez, B.S. (2004) Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Mol. Cell*, **14**, 611–623.
21. Wang, H., Shao, Z., Shi, L.Z., Hwang, P.Y.-H., Truong, L.N., Berns, M.W., Chen, D.J. and Wu, X. (2012) CtIP protein dimerization is critical for its recruitment to chromosomal DNA double-stranded breaks. *J. Biol. Chem.*, **287**, 21471–21480.
22. Gunn, A. and Stark, J.M. (2012) I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. *Methods Mol. Biol.*, **920**, 379–391.
23. Mao, Z., Bozzella, M., Seluanov, A. and Gorbunova, V. (2008) Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst.)*, **7**, 1765–1771.
24. Bindra, R.S., Goglia, A.G., Jasin, M. and Powell, S.N. (2013) Development of an assay to measure mutagenic non-homologous end-joining repair activity in mammalian cells. *Nucleic Acids Res.*, **41**, e115.
25. Certo, M.T., Ryu, B.Y., Annis, J.E., Garibov, M., Jarjour, J., Rawlings, D.J. and Scharenberg, A.M. (2011) Tracking genome engineering outcome at individual DNA breakpoints. *Nat. Methods*, **8**, 671–676.
26. Kuhar, R., Gwiazda, K.S., Humbert, O., Mandt, T., Pangallo, J., Brault, M., Khan, I., Maizels, N., Rawlings, D.J., Scharenberg, A.M. *et al.* (2014) Novel fluorescent genome editing reporters for monitoring DNA repair pathway utilization at endonuclease-induced breaks. *Nucleic Acids Res.*, **42**, e4.
27. Yáñez, R.J. and Porter, A.C.G. (2002) A chromosomal position effect on gene targeting in human cells. *Nucleic Acids Res.*, **30**, 4892–4901.
28. Cornea, A.M. and Russell, D.W. (2010) Chromosomal position effects on AAV-mediated gene targeting. *Nucleic Acids Res.*, **38**, 3582–3594.
29. Aymard, F., Bugler, B., Schmidt, C.K., Guillou, E., Caron, P., Briois, S., Iacovoni, J.S., Daburon, V., Miller, K.M., Jackson, S.P. *et al.* (2014) Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat. Struct. Mol. Biol.*, **21**, 366–374.
30. Kulkarni, A., Zschenker, O., Reynolds, G., Miller, D. and Murnane, J.P. (2010) Effect of telomere proximity on telomere position effect, chromosome healing, and sensitivity to DNA double-strand breaks in a human tumor cell line. *Mol. Cell. Biol.*, **30**, 578–589.
31. Lemaître, C., Grabarz, A., Tsouroula, K., Andronov, L., Furst, A., Pankotai, T., Heyer, V., Rogier, M., Attwood, K.M., Kessler, P. *et al.* (2014) Nuclear position dictates DNA repair pathway choice. *Genes Dev.*, **28**, 2450–2463.
32. Jeggo, P.A. and Downs, J.A. (2014) Roles of chromatin remodellers in DNA double strand break repair. *Exp. Cell Res.*, **329**, 69–77.
33. Pfister, S.X., Ahrabi, S., Zalmas, L.-P., Sarkar, S., Aymard, F., Bachrati, C.Z., Helleday, T., Legube, G., La Thangue, N.B., Porter, A.C.G. *et al.* (2014) SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep.*, **7**, 2006–2018.
34. Clouaire, T. and Legube, G. (2015) DNA double strand break repair pathway choice: a chromatin based decision? *Nucleus*, **6**, 107–113.
35. Weber, S., Hofmann, A., Herms, S., Hoffmann, P. and Doerfler, W. (2015) Destabilization of the human epigenome:

consequences of foreign DNA insertions. *Epigenomics*, doi:10.2217/epi.15.40.
36. Kong, Q.R., Wu, M.L., Zhang, L., Wang, F., Yin, Z., Mu, Y.S. and Liu, Z.H. (2011) Transgene insertion affects transcription and epigenetic modification of flanking host sequence in transgenic pigs. *Cell. Mol. Biol. (Noisy-le-Grand)*, **57** (suppl.), OL1505-12.
37. Würtele, H., Little, K.C.E. and Chartrand, P. (2003) Illegitimate DNA integration in mammalian cells. *Gene Ther.*, **10**, 1791–1799.
38. Jinnah, H.A., De Gregorio, L., Harris, J.C., Nyhan, W.L. and O'Neill, J.P. (2000) The spectrum of inherited mutations causing HPRT deficiency: 75 new cases and a review of 196 previously reported cases. *Mutat. Res.*, **463**, 309–326.
39. Doetschman, T., Maeda, N. and Smithies, O. (1988) Targeted mutation of the Hprt gene in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **85**, 8583–8857.
40. Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S. and Smithies, O. (1987) Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature*, **330**, 576–578.
41. Skandalis, A. and Glickman, B.W. (1990) Endogenous gene systems for the study of mutational specificity in mammalian cells. *Cancer Cells*, **2**, 79–83.
42. McLachlan, J., Fernandez, S., Helleday, T. and Bryant, H.E. (2009) Specific targeted gene repair using single-stranded DNA oligonucleotides at an endogenous locus in mammalian cells uses homologous recombination. *DNA Repair (Amst.)*, **8**, 1424–1433.
43. Rothkamm, K., Gunasekara, K., Warda, S.A., Krempler, A. and Löbrich, M. (2008) Radiation-induced HPRT mutations resulting from misrejoined DNA double-strand breaks. *Radiat. Res.*, **169**, 639–648.
44. Rouet, P., Smih, F. and Jasin, M. (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.*, **14**, 8096–8106.
45. Wang, Z., Zhou, Z.-J., Liu, D.-P. and Huang, J.-D. (2008) Double-stranded break can be repaired by single-stranded oligonucleotides via the ATM/ATR pathway in mammalian cells. *Oligonucleotides*, **18**, 21–32.
46. Chen, F., Pruett-Miller, S.M., Huang, Y., Gjoka, M., Duda, K., Taunton, J., Collingwood, T.N., Frodin, M. and Davis, G.D. (2011) High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods*, **8**, 753–755.
47. Yáñez, R.J. and Porter, A.C. (1999) Gene targeting is enhanced in human cells overexpressing hRAD51. *Gene Ther.*, **6**, 1282–1290.
48. Rosner, M., Schipany, K. and Hengstschläger, M. (2013) Merging high-quality biochemical fractionation with a refined flow cytometry approach to monitor nucleocytoplasmic protein expression throughout the unperturbed mammalian cell cycle. *Nat. Protoc.*, **8**, 602–626.
49. Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H. *et al.* (2008) Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*, **132**, 487–498.
50. Bryant, P.E., Mozdarani, H. and Marr, C. (2008) G2-phase chromatid break kinetics in irradiated DNA repair mutant hamster cell lines using calyculin-induced PCC and colcemid-block. *Mutat. Res.*, **657**, 8–12.
51. Averbeck, N.B., Ringel, O., Herrlitz, M., Jakob, B., Durante, M. and Taucher-Scholz, G. (2014) DNA end resection is needed for the repair of complex lesions in G1-phase human cells. *Cell Cycle*, **13**, 2509–2516.
52. Honma, M., Izumi, M., Sakuraba, M., Tadokoro, S., Sakamoto, H., Wang, W., Yatagai, F. and Hayashi, M. (2003) Deletion, rearrangement, and gene conversion; genetic consequences of chromosomal double-strand breaks in human cells. *Environ. Mol. Mutagen.*, **42**, 288–298.
53. Cradick, T.J., Fine, E.J., Antico, C.J. and Bao, G. (2013) CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.*, **41**, 9584–9592.
54. Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.-L., Wang, Z., Zhang, Z., Zheng, R., Yang, L. *et al.* (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **111**, 4632–4637.
55. Ghezraoui, H., Piganeau, M., Renouf, B., Renaud, J.-B., Sallmyr, A., Ruis, B., Oh, S., Tomkinson, A.E., Hendrickson, E.A., Giovannangeli, C. *et al.* (2014) Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol. Cell*, **55**, 829–842.
56. Hartlerode, A., Odate, S., Shim, I., Brown, J. and Scully, R. (2011) Cell cycle-dependent induction of homologous recombination by a tightly regulated I-SceI fusion protein. *PLoS One*, **6**, e16501.
57. Chan, S.-H., Stoddard, B.L. and Xu, S.-Y. (2011) Natural and engineered nicking endonucleases—from cleavage mechanism to engineering of strand-specificity. *Nucleic Acids Res.*, **39**, 1–18.
58. Delacôte, F. and Lopez, B.S. (2008) Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model. *Cell Cycle*, **7**, 33–38.
59. Aarts, M. and te Riele, H. (2010) Subtle gene modification in mouse ES cells: evidence for incorporation of unmodified oligonucleotides without induction of DNA damage. *Nucleic Acids Res.*, **38**, 6956–6967.
60. Aarts, M. and te Riele, H. (2010) Parameters of oligonucleotide-mediated gene modification in mouse ES cells. *J. Cell. Mol. Med.*, **14**, 1657–1667.
61. Aarts, M. and te Riele, H. (2011) Progress and prospects: oligonucleotide-directed gene modification in mouse embryonic stem cells: a route to therapeutic application. *Gene Ther.*, **18**, 213–219.
62. Storici, F., Snipe, J.R., Chan, G.K., Gordenin, D.A. and Resnick, M.A. (2006) Conservative repair of a chromosomal double-strand break by single-strand DNA through two steps of annealing. *Mol. Cell. Biol.*, **26**, 7645–7657.
63. Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M.H. and Chandrasegaran, S. (2005) Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.*, **33**, 5978–5990.
64. Deng, C. and Capecchi, M.R. (1992) Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell. Biol.*, **12**, 3365–3371.
65. Ding, Q., Lee, Y.-K., Schaefer, E.A.K., Peters, D.T., Veres, A., Kim, K., Kuperwasser, N., Motola, D.L., Meissner, T.B., Hendriks, W.T. *et al.* (2013) A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*, **12**, 238–251.
66. Merkert, S., Wunderlich, S., Bednarski, C., Beier, J., Haase, A., Dreyer, A.-K., Schwanke, K., Meyer, J., Göhring, G., Cathomen, T. *et al.* (2014) Efficient designer nuclease-based homologous recombination enables direct PCR screening for footprintless targeted human pluripotent stem cells. *Stem Cell Rep.*, **2**, 107–118.
67. Sargent, R.G., Brenneman, M.A. and Wilson, J.H. (1997) Repair of site-specific double-strand breaks in a mammalian

chromosome by homologous and illegitimate recombination. *Mol. Cell. Biol.*, **17**, 267–277.
68. Majumdar, A., Muniandy, P.A., Liu, J., Liu, J., Liu, S., Cuenoud, B. and Seidman, M.M. (2008) Targeted gene knock in and sequence modulation mediated by a psoralen-linked triplex-forming oligonucleotide. *J. Biol. Chem.*, **283**, 11244–11252.
69. Smurnyy, Y., Cai, M., Wu, H., McWhinnie, E., Tallarico, J.A., Yang, Y. and Feng, Y. (2014) DNA sequencing and CRISPR-Cas9 gene editing for target validation in mammalian cells. *Nat. Chem. Biol.*, doi:10.1038/nchembio.1550.
70. Wu, X., Blackburn, P.R., Tschumper, R.C., Ekker, S.C. and Jelinek, D.F. (2014) TALEN-mediated genetic tailoring as a tool to analyze the function of acquired mutations in multiple myeloma cells. *Blood Cancer J.*, **4**, e210.
71. Kan, Y., Ruis, B., Lin, S. and Hendrickson, E.A. (2014) The mechanism of gene targeting in human somatic cells. *PLoS Genet.*, **10**, e1004251.
72. Porter, A.C. and Itzhaki, J.E. (1993) Gene targeting in human somatic cells. Complete inactivation of an interferon-inducible gene. *Eur. J. Biochem.*, **218**, 273–281.
73. Magin, T.M., McEwan, C., Milne, M., Pow, A.M., Selfridge, J. and Melton, D.W. (1992) A position- and orientation-dependent element in the first intron is required for expression of the mouse hprt gene in embryonic stem cells. *Gene*, **122**, 289–296.
74. Johnson, R.D., Liu, N. and Jasin, M. (1999) Mammalian XRCC2 promotes the repair of DNA double-strand breaks by homologous recombination. *Nature*, **401**, 397–399.
75. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
76. Yun, S., Lie-A-Cheong, C. and Porter, A.C.G. (2004) Discriminatory suppression of homologous recombination by p53. *Nucleic Acids Res.*, **32**, 6479–6489.

OPEN ACCESS
CellPress

Cell Reports
Article

SETD2-Dependent Histone H3K36 Trimethylation Is Required for Homologous Recombination Repair and Genome Stability

Sophia X. Pfister,[1] Sara Ahrabi,[1] Lykourgos-Panagiotis Zalmas,[2] Sovan Sarkar,[1] François Aymard,[3,4] Csanád Z. Bachrati,[5] Thomas Helleday,[6] Gaëlle Legube,[3,4] Nicholas B. La Thangue,[2] Andrew C.G. Porter,[7] and Timothy C. Humphrey[1,*]
[1]CRUK MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Oxford OX3 7DQ, UK
[2]Laboratory of Cancer Biology, Department of Oncology, University of Oxford, Oxford OX3 7DQ, UK
[3]Laboratoire de Biologie Cellulaire et Moléculaire du Contrôle de la Prolifération, Université de Toulouse, Université Paul Sabatier, 31062 Toulouse, France
[4]CNRS, Laboratoire de Biologie Cellulaire et Moléculaire du Contrôle de la Prolifération, 31062 Toulouse, France
[5]School of Life Sciences, University of Lincoln, Lincoln LN6 7TS, UK
[6]Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 21 Stockholm, Sweden
[7]Gene Targeting Group, Centre for Haematology, Imperial College Faculty of Medicine, London W12 0NN, UK
*Correspondence: timothy.humphrey@oncology.ox.ac.uk
http://dx.doi.org/10.1016/j.celrep.2014.05.026
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

SUMMARY

Modulating chromatin through histone methylation orchestrates numerous cellular processes. SETD2-dependent trimethylation of histone H3K36 is associated with active transcription. Here, we define a role for H3K36 trimethylation in homologous recombination (HR) repair in human cells. We find that depleting SETD2 generates a mutation signature resembling RAD51 depletion at I-SceI-induced DNA double-strand break (DSB) sites, with significantly increased deletions arising through microhomology-mediated end-joining. We establish a presynaptic role for SETD2 methyltransferase in HR, where it facilitates the recruitment of C-terminal binding protein interacting protein (CtIP) and promotes DSB resection, allowing Replication Protein A (RPA) and RAD51 binding to DNA damage sites. Furthermore, reducing H3K36me3 levels by overexpressing KDM4A/JMJD2A, an oncogene and H3K36me3/2 demethylase, or an H3.3K36M transgene also reduces HR repair events. We propose that error-free HR repair within H3K36me3-decorated transcriptionally active genomic regions promotes cell homeostasis. Moreover, these findings provide insights as to why oncogenic mutations cluster within the H3K36me3 axis.

INTRODUCTION

DNA double-stranded breaks (DSBs) are potentially lethal lesions if unrepaired, and their misrepair can give rise to genome instability, a hallmark of cancer (Jeggo and Lavin, 2009). To maintain genome stability in response to such lesions, cells employ either homologous recombination (HR) or nonhomologous end-joining (NHEJ) pathways to repair DSBs. HR is initiated by 5′ end resection to generate a 3′ single-stranded DNA (ssDNA) overhang. Resection is a two-step process initiated by removing a short oligonucleotide through the activities of the Mre11-Rad50-Nbs1 (MRN) complex and CtIP. Extensive resection is performed by Exo1 or DNA2-BLM in conjunction with Replication Protein A (RPA), which binds ssDNA and removes secondary structures (Sugiyama et al., 1998; Symington and Gautier, 2011). This, in turn, leads to RAD51 nucleofilament formation, which promotes strand invasion of a homologous chromatid, leading to accurate repair (Heyer et al., 2010). During classical NHEJ (C-NHEJ), the broken ends are rapidly bound and protected by the Ku70/Ku80 heterodimer, which acts as a platform to recruit the DNA-PK catalytic subunit (DNA-PKcs). Damaged ends are then processed and subsequently joined by the Ligase 4 (Lig4), XRCC4, XLF complex in a template-independent manner, which can lead to inaccurate repair (Lieber, 2010). DSBs may also be repaired through alternative end-joining pathways, such as microhomology-mediated end-joining (MMEJ), which do not require Ku or Lig4. Like HR, MMEJ is initiated by resection, and end-joining is mediated through annealing of short direct repeats of microhomology. MMEJ leads to deletions and is frequently associated with chromosomal rearrangements (McVey and Lee, 2008).

DSB repair is further facilitated through chromatin modification by chromatin remodeling complexes, by incorporation of histone variants, and by histone modification (Smeenk and van Attikum, 2013). Histone H3K36 dimethylation has recently been proposed to facilitate NHEJ, where Metnase (SETMAR) directly mediates H3K36 dimethylation near the break site, leading to recruitment and stabilization of NBS1 and Ku70 (Fnu et al., 2011). Histone H3K36 is also trimethylated, which in mammalian cells is performed uniquely by the SETD2/HYPB methyltransferase (Edmunds et al., 2008). H3K36me3 is associated with transcriptional elongation and is found in gene coding regions,

peaking at 3′ ends (Edmunds et al., 2008). SETD2-dependent H3K36 trimethylation facilitates a number of processes within the cell, including splicing, repression of intragenic transcripts, and chromatin accessibility (Li et al., 2013; Wagner and Carpenter, 2012). *SETD2* is also mutated in a number of cancer types, including breast, lung, acute lymphoblastic leukemia, clear cell renal cell carcinoma, and glioma, supporting its role as a tumor suppressor (Al Sarakbi et al., 2009; Dalgliesh et al., 2010; Fontebasso et al., 2013; Newbold and Mokbel, 2010; Zhang et al., 2012). SETD2-dependent H3K36 trimethylation has recently been shown to regulate DNA mismatch repair (Li et al., 2013). However, SETD2-deficient cancers exhibit a wide range of mutations, including insertions, deletions (indels), and chromosomal aberrations (Sato et al., 2013; Zhu et al., 2014), suggesting an additional role for SETD2 in genome stability.

Additionally, H3K36 methylation is regulated by the KDM4/JMJD2 family of histone demethylases. These contain JmjN-JmjC and tandem-Tudor domains that specifically remove the tri- and dimethyl forms of both H3K9 and H3K36, with preference for the trimethyl form being observed in the case of KDM4A/JMJD2A (Couture et al., 2007; Klose et al., 2006; Whetstine et al., 2006). KDM4 family proteins are frequently overexpressed in cancers and are associated with poor patient survival (Berry and Janknecht, 2013; Black et al., 2013).

Here, we investigate the role of SETD2-dependent H3K36me3 in maintaining genome stability. We establish a role for H3K36me3 in HR repair by facilitating resection. In addition, we define a role for SETD2 and RAD51 in maintaining genome stability through suppressing MMEJ at break sites.

RESULTS

SETD2 Suppresses Break-Induced Mutations

Tumor suppressors can suppress carcinogenesis by preventing genome instability, a hallmark of cancer. To investigate the impact of loss of SETD2, a tumor suppressor, on genome stability following DSB induction, we developed an I-SceI-induced loss of function assay (HPRT$^+$:I-SceI). Intron-encoded endonuclease 1 from *Saccharomyces cerevisiae* (I-SceI) is a rare-cutting endonuclease that has no predicted recognition site in the mammalian genome (Jasin, 1996). An I-SceI recognition sequence was inserted into the endogenous *HPRT* exon 6 in HT1080 (fibrosarcoma) cells. The I-SceI site maintains the HPRT reading frame and makes only a single amino acid change that does not impair the function of HPRT (Figure 1A). Following I-SceI induction, inaccurate DSB repair generates HPRT-negative mutants, which can be selected for using 6-thioguanine (6-TG), and indels can be detected by PCR amplification using primers flanking the break site (Figure 1A). Following small interfering RNA (siRNA) knockdown and I-SceI induction, the frequency of HPRT loss (mutation frequency) in SETD2-depleted cells was 1.65%, significantly higher ($p = 0.0014$) than that of nontargeting controls (1.1%). These findings resembled those in RAD51-depleted cells in which the frequency of HPRT loss was also significantly increased to 1.88% ($p = 0.0022$) compared to nontargeting controls (Figure 1B). SETD2 and RAD51 codepletion (siS+SiR) also significantly increased the mutation frequency to 1.52% ($p = 0.037$), which was similar to the frequency observed following either SETD2 or RAD51 depletion (Figure 1B). We confirmed that I-SceI protein expression and cleavage efficiency were the same in cells transfected with either control siRNA or SETD2 siRNA (Figures S1A–S1C). In addition, *HPRT* gene transcription was not affected by SETD2 knockdown, as shown by quantitative RT-PCR (qRT-PCR) of spliced and unspliced HPRT mRNA (Figures S1D and S1E). Therefore, the mutation frequency in this system is likely to be directly associated with misrepair of the DSB at the I-SceI site.

Further, we studied the mutation patterns across the break site of 30 individually isolated HPRT-negative clones from each background by PCR amplification and sequencing. HPRT-negative clones from cells treated with nontargeting control siRNA indicated the presence of microdeletions of 2–5 bp, consistent with DSB repair by C-NHEJ. Larger microdeletions were also observed that were associated with regions of microhomology located either side of the break site, consistent with DSB repair through MMEJ (Figures 1C and S1F).

Sequence analysis indicated that SETD2-depleted cells exhibited a significant increase in average deletion length (11 bp, $p < 0.0001$) compared to control cells (5 bp) (Figures 1C and 1D). RAD51-depleted cells exhibited a further significant increase in the average deletion length (22 bp, $p = 0.014$) (Figures 1C and 1D). RAD51 and SETD2 codepletion generated deletion lengths (an average of 8 bp) comparable to those observed following SETD2 knockdown, which in turn was significantly higher than the nontargeting control ($p = 0.026$) (Figures 1C and 1D). Insertions at the HPRT break site were also detected. However, the frequency and size of the insertions was not significantly altered between SETD2-depleted cells, SETD2/RAD51 codepleted and nontargeting controls.

We next examined the mechanism for misrepair in cells exhibiting deletions. We found that the frequency of microhomology-mediated end-joining (MMEJ) was significantly increased in SETD2 and RAD51-depleted cells ($p < 0.05$) (Figures 1E and S1G). This suggested that HR-deficient cells use MMEJ as an alternative repair mechanism, which gives rise to a mutation signature of sequence loss between regions of microhomology. These results support a role for SETD2 in tumor suppression by maintaining genome stability through preventing deleterious mutations arising in response to DSBs. These findings further identify a similar break-induced mutation signature between SETD2- and RAD51-depleted cells, suggesting an early role for SETD2 in HR repair.

SETD2 Is Required for Homologous Recombination Repair

To test the role for SETD2 in DSB repair, we depleted SETD2 in U2OS and HeLa cells using two independent siRNAs and measured the clonogenic survival following exposure to DNA-damaging agents. SETD2 knockdown significantly reduced survival after treatment with mitomycin C (MMC), camptothecin (CPT), or ionizing radiation (IR) in U2OS cells and HeLa cells compared to nontargeting controls (Figures 2A and S2A). These results identify a role for SETD2 in the cellular response to DNA damage and are consistent with a role for SETD2 in DSB repair.

SETD2 knockdown was validated by western blotting (Figure 2B). Western blotting also showed that SETD2 knockdown

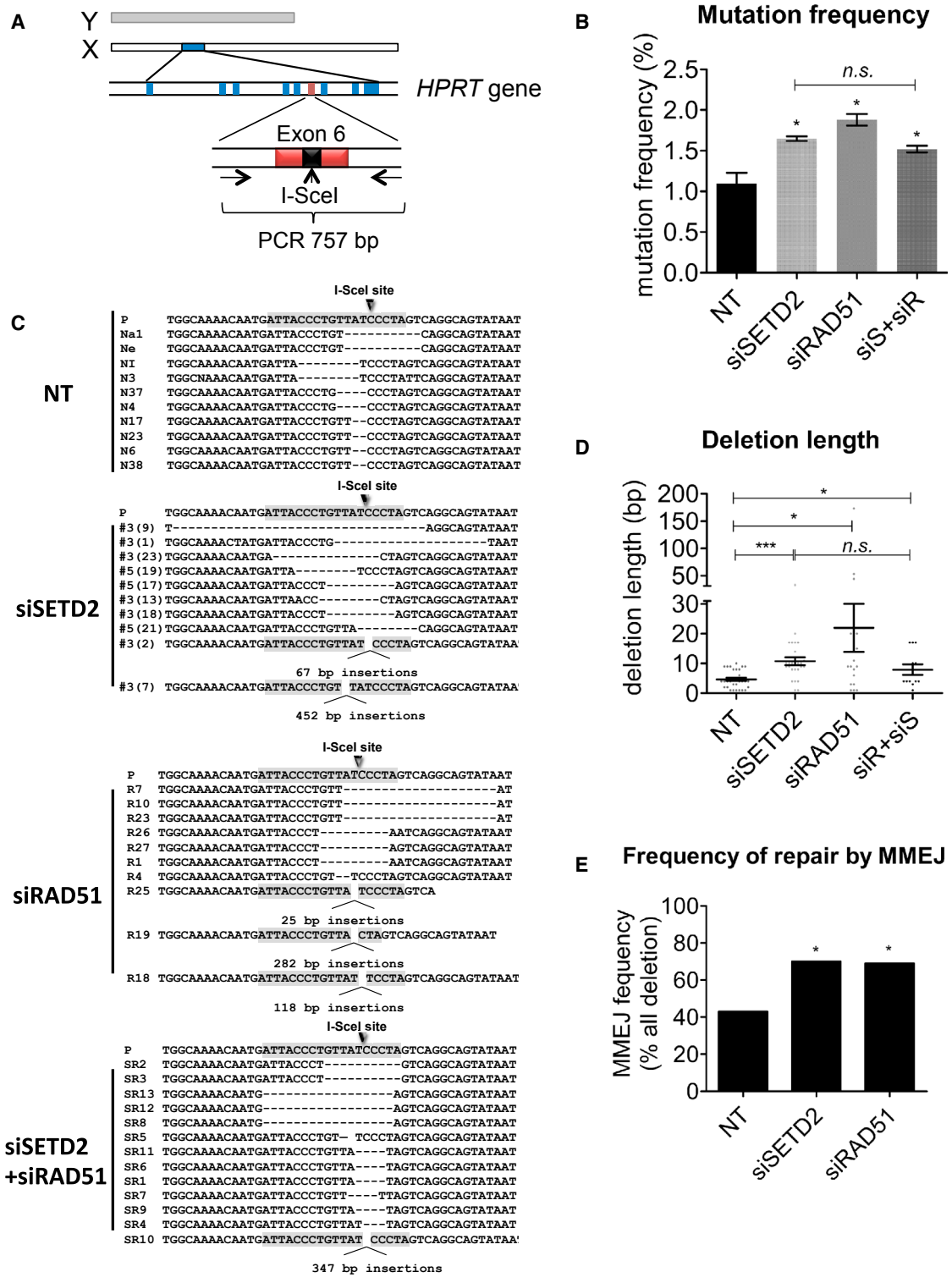


Figure 1. SETD2- and RAD51-Depleted Cells Exhibit a Common Break-Induced Mutation Signature

(A) Schematic map of the HPRT$^+$:I-SceI assay, where arrows indicate PCR primers used to amplify genomic DNA to allow sequencing across the break site (see Supplemental Experimental Procedures).

(B) Mutation frequency of nontargeting control (NT), SETD2-depleted, RAD51-depleted, and SETD2/RAD51-codepleted cells (siS+siR) after I-SceI induced DSB. Error bars represent SEM and *$p < 0.05$; n.s., not significant.

(legend continued on next page)

significantly reduced global levels of H3K36me3, confirming SETD2 as the major H3K36me3 methyltransferase in human cells (Figure 2B). Because RAD51 is more likely to be affected among DNA repair genes by siRNA off-target effects (Adamson et al., 2012), we also confirmed that SETD2 knockdown by siRNA did not affect the RAD51 protein levels (Figure 2B). We confirmed that SETD2 knockdown had no effect on the expression of major HR proteins, including BRCA2, RAD50, CtIP, MRE11, and RPA and the NHEJ protein KU80 (Figure S2B). We also observed no changes to the cell cycle after SETD2 knockdown (Figure S2C).

Both MMC and CPT induce S phase-specific DSBs, which are predominantly repaired by HR (Arnaudeau et al., 2001; Moynahan et al., 2001). To test a possible role for SETD2 in HR repair, we used a well-characterized GFP-based reporter for HR (DR-GFP). The HR reporter contains an I-SceI recognition sequence, which upon I-SceI expression is cleaved to generate a DSB. DSB repair by HR using a direct repeat within the reporter cassette as a template results in an intact GFP gene (Figure 2C)(Pierce et al., 1999). SETD2 knockdown by two independent siRNAs significantly reduced HR repair of an I-SceI induced DSB by 67%–77% compared to nontargeting control cells ($p < 0.0003$) (Figure 2D). As a control for the HR reporter, knockdown of RAD51 reduced HR activity by 97% ($p < 0.0001$), whereas DNA-PKcs inhibitor NU7441 had no effect on HR (Figure 2D). In contrast, SETD2 knockdown had no effect on repair of an I-SceI induced DSB in an NHEJ reporter (IRES-TK-EGFP) in which joining of two I-SceI sites separated by a thymidine kinase gene (TK) results in GFP expression (Figures 2E and 2F)(Ogiwara et al., 2011). As a control, RAD51 had no effect on the same NHEJ reporter, whereas the DNA-PK inhibitor NU7441 reduced NHEJ activity by 98% ($p < 0.0001$) (Figure 2F). These results identify a role for SETD2 in HR repair.

An Early Role for SETD2 in Homologous Recombination

To further investigate the role of SETD2 in DSB repair, we examined the rate of repair following IR. We used γH2AX foci as a marker of DNA damage and found that following exposure to 5 Gy IR, SETD2 knockdown significantly delayed removal of γH2AX foci at 48 hr ($p < 0.003$) compared to nontargeting controls, supporting a role for SETD2 in efficient DSB repair (Figure 3A). We next investigated the effect of SETD2 depletion on the recruitment of HR repair proteins following DNA damage. SETD2 knockdown resulted in significantly reduced RAD51 foci formation 4 and 8 hr after exposure to IR compared to nontargeting controls ($p < 0.0013$) (Figure 3B). SETD2 knockdown also significantly reduced RAD51 foci formation after exposure to CPT ($p < 0.05$) (Figure S2D).

We also examined the effect of SETD2 depletion on RPA recruitment, where we induced DSBs in S phase cells using CPT or MMC. A significant reduction in RPA foci formation was observed after CPT or MMC treatment following SETD2 knockdown compared to nontargeting controls ($p < 0.023$) (Figures 3C and S2E). SETD2 knockdown leads to reduced H3K36me3 in the nuclei as confirmed by costaining the cells with antibodies against H3K36me3 and RPA (Figure S2F). Together, these findings support a role for SETD2 in promoting HR through facilitating recruitment of RPA and RAD51 to DNA damage sites.

To further confirm the impact of SETD2 loss on the recruitment of HR factors to the site of DSBs, we performed chromatin immunoprecipitation (ChIP) near the I-SceI site in the DR-GFP cassette (Figure S3A). The I-SceI site is located within the DR-GFP cassette under the control of the highly active CMV promoter, and thus H3K36me3, which is associated with active transcription, is predicted to be present throughout the cassette. After induction of DSBs by I-SceI, no change in H3K36me3 or SETD2 levels was detected proximal or distal to the break site after 8, 24, or 48 hr following I-SceI transfection (Figure S3B). The results were also quantified by ChIP-qPCR (Figure S3C). Although other groups also observed no changes to H3K36me3 level after DSB, globally (Fnu et al., 2011) or locally at the break site (Pei et al., 2011; Aymard et al., 2014), we were surprised by the lack of reduction in H3K36me3 or SETD2 knowing that regional transcriptional silencing has been observed at sites of DSBs (Shanbhag et al., 2010). We speculate that the unexpected maintenance of H3K36me3 at the DSB site by SETD2 may serve as an anchor to mediate HR repair.

We monitored recruitment of major HR factors to the DSB site in the ChIP system. RPA and RAD51 were recruited to the DSB site, resulting in increased binding to the "up" site 18 hr after I-SceI transfection, compared to uncut controls (Figure 3D). SETD2 depletion substantially reduced the recruitment of RPA and RAD51 after the I-SceI cut (Figure 3D). SETD2 protein depletion was validated by western blotting (Figure S3D), and was further confirmed by the absence of SETD2 or H3K36me3 at the break site (Figure 3D). These findings are consistent with the RPA and RAD51 foci data, indicating that SETD2 is required to recruit RPA and RAD51 to DSB sites. Together, these findings confirm a presynaptic role for SETD2 in facilitating HR repair of DSBs.

H3K36me3 Is Required for Homologous Recombination Repair

To investigate if the methyltransferase activity of SETD2 is required for its role in HR, we generated two doxycycline-inducible U2OS cell lines, where either a wild-type *SETD2* cDNA or a methyltransferase-dead *SETD2* cDNA was inserted behind a Tet operator in the genome of U2OS cells. Upon addition of doxycycline to the cell culture medium, the Tet operator was derepressed and the wild-type or mutant SETD2 was expressed. Both cDNA constructs were mutated so that they are refractory to SETD2 siRNA#3. The methyltransferase-dead SETD2 was

(C) Representative sequence alignments of the above-mentioned PCR products in nontargeting control (NT), SETD2-depleted, RAD51-depleted, and SETD2/RAD51-codepleted backgrounds, respectively.

(D) Average length of deletions (bp) in different backgrounds, where each dot represents an independent clone. The lines represent mean and SEM, *$p < 0.05$, ***$p < 0.001$.

(E) Frequency of repair by MMEJ in deletion mutants isolated from different backgrounds. p values calculated by statistical analysis "difference between proportions," *$p < 0.05$. See also Figure S1.

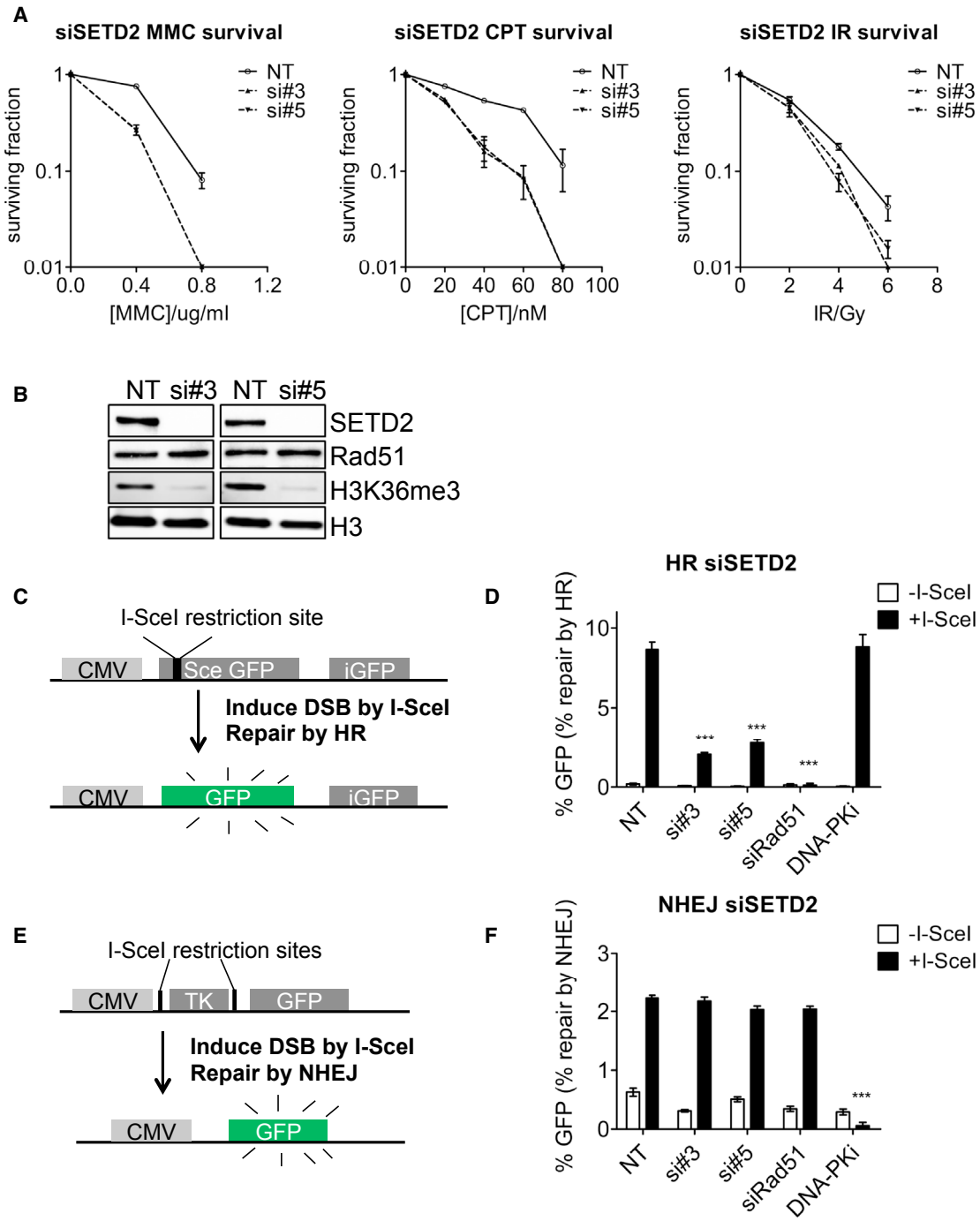


Figure 2. SETD2 Is Required for Homologous Recombination

(A) Clonogenic survival of SETD2 knockdown (si#3 and si#5) or nontargeting control (NT) U2OS cells treated with indicated concentrations of MMC, CPT, and IR. Error bars show SEM from three independent experiments.

(B) Western blots showing levels of SETD2, RAD51, and H3K36me3 72 hr after siRNA transfection.

(C) Schematic map of the DR-GFP cassette for assessing HR efficacy.

(D) HR repair efficacy of reporter cells treated with nontargeting control siRNA (NT), SETD2 siRNAs (si#3 and si#5), RAD51 siRNA (siRad51), or DNA-PK inhibitor NU7441 (Axon), indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. ***$p < 0.001$.

(E) Schematic map of the NHEJ cassette.

(F) NHEJ repair efficacy of reporter cells treated with nontargeting control siRNA (NT), SETD2 siRNAs (si#3 and si#5), RAD51 siRNA (siRAD51), or DNA-PK inhibitor NU7441, indicated by the percentage of GFP-positive cells. Error bars show SEM from three independent experiments. ***$p < 0.001$. See also Figure S2.

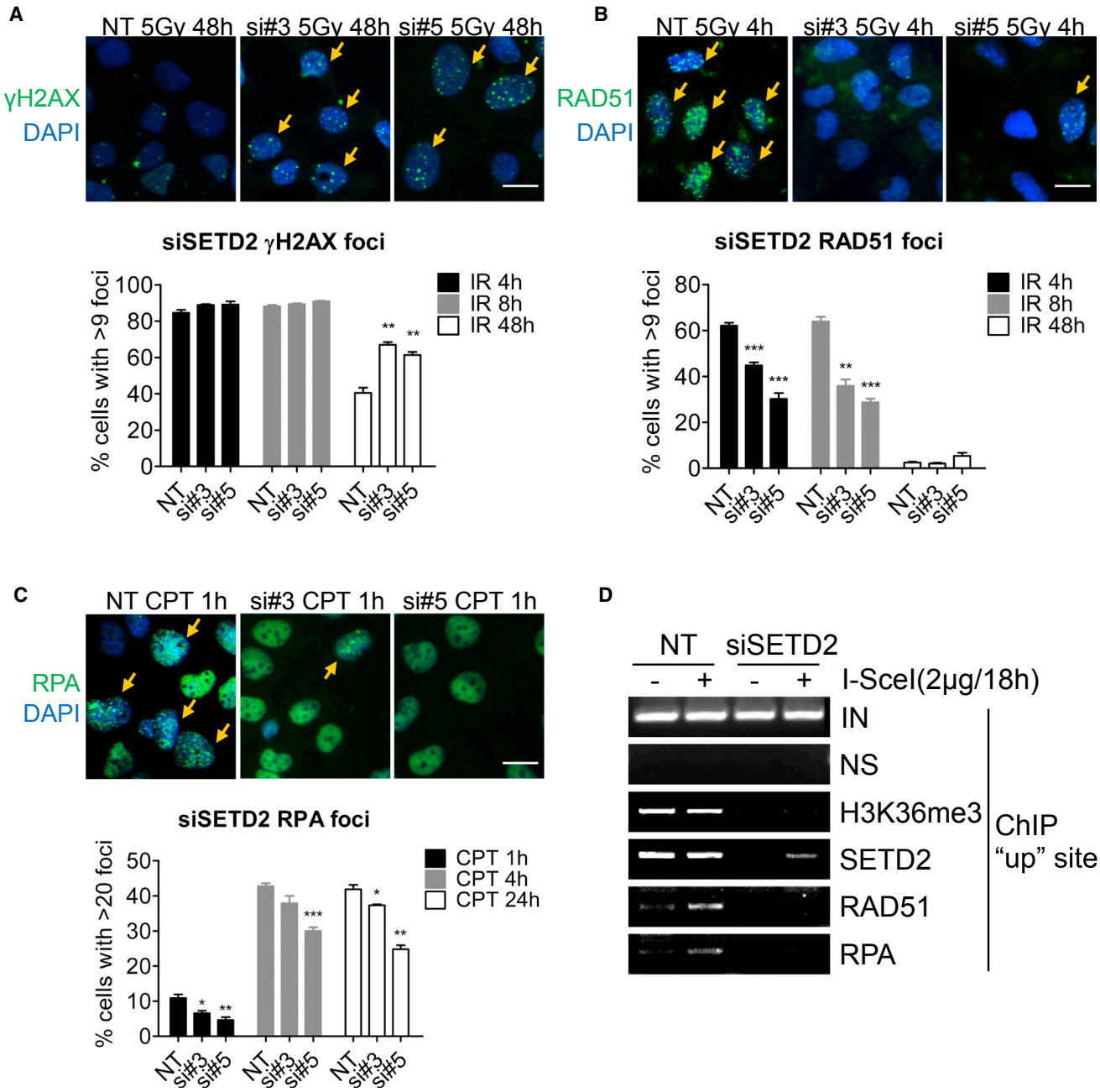


Figure 3. SETD2 Is Required for the Recruitment of RPA and RAD51

(A and B) γH2AX (A) or RAD51 (B) foci formation at indicated times after IR (5 Gy) in U2OS cells treated with nontargeting control siRNA (NT) or SETD2 siRNAs (si#3 and si#5). Error bars represent SEM from three independent experiments.

(C) RPA32 foci formation at indicated times after treatment with CPT (10 μM) in U2OS cells transfected with control siRNA (NT) or SETD2 siRNAs (si#3 and si#5). Representative fluorescent images are shown; scale bar, 20 μm. For each condition, 45 fields and at least 500 cells were examined by Incell (GE Healthcare). Error bars show SEM from three independent experiments. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

(D) ChIP analysis on DR-GFP U2OS cells transfected with nontargeting control (NT) or SETD2 siRNA for 72 hr, followed by transfection of either vector or I-SceI plasmid for a further 18 hr, as indicated. ChIP on the "up" DNA site of the DR-GFP cassette was performed on the lysate using antibodies against nonspecific Ig (NS), H3K36me3 (Abcam), SETD2 (Abcam), RAD51 (Santa Cruz Biotechnology), or RPA (Millipore); inputs (IN) are also indicated.

See also Figures S2 and S3.

mutated at two amino acid residues located within the catalytic site of the SET domain, which are conserved with yeast Set2. The corresponding mutations in yeast were shown to abolish the methyltransferase activity (Rea et al., 2000;Strahl et al., 2002) (Figure 4A). Both arginine (R) and cysteine (C) residues were mutated in the methyltransferase-dead (MD) cDNA by

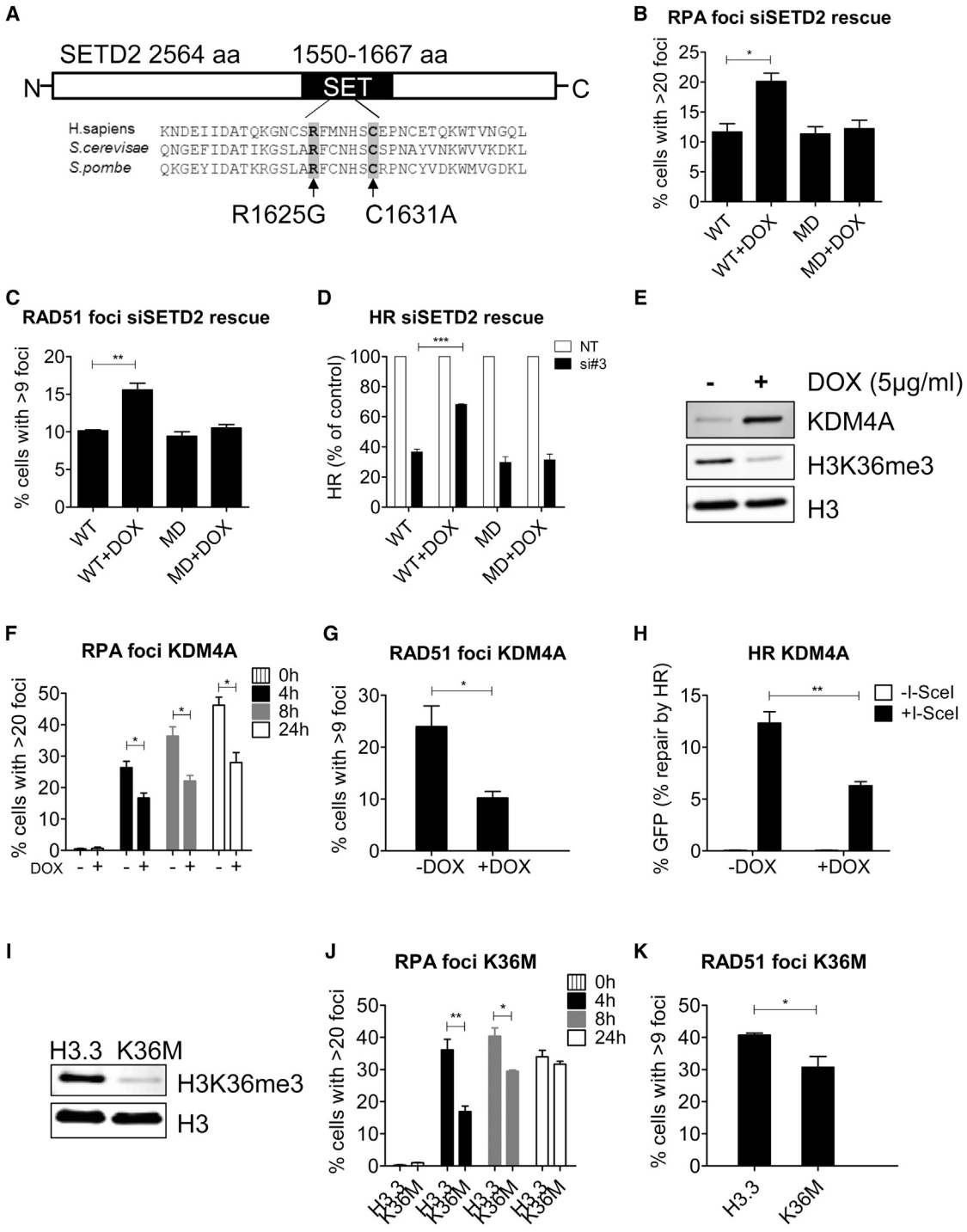


Figure 4. H3K36me3 Is Required for Efficient HR Repair

(A) Schematic map showing site-directed mutagenesis for abolishing the methyltransferase activity of SETD2, with both R and C residues mutated simultaneously. Sequence alignment of human SETD2 and *Saccharomyces cerevisiae* Set2 and *Schizosacchromyces pombe* Set2 amino acid sequences reveals evolutionarily conserved residues that reside in the SET domain.

(B) RPA32 foci formation at 2 hr after treatment with CPT (10 μM) in cells transfected with SETD2 siRNA. WT: T-REx U2OS clone (see Experimental Procedures) with wild-type *SETD2* cDNA integrated but not expressing exogenous SETD2. WT+DOX: the same WT clone expressing exogenous wild-type SETD2. MD: T-REx U2OS clone with methyltransferase-dead mutant *SETD2* cDNA integrated but not expressing exogenous SETD2. MD+DOX: the same MD clone expressing exogenous mutant SETD2. Error bars show SEM from three independent experiments, *p < 0.05.

(legend continued on next page)

site-directed mutagenesis (see Supplemental Experimental Procedures). Western blot analysis confirmed that doxycycline-induced expression of wild-type SETD2 could restore the level of H3K36me3 after siSETD2 treatment, whereas expression of methyltransferase-dead SETD2 could not restore the level of H3K36me3 (Figure S4A).

We depleted endogenous SETD2 with siRNA#3 and examined the rescue of the phenotype by the two constructs. Although the wild-type construct increased foci formation of RPA and RAD51 after DNA damage in siSETD2 cells, the methyltransferase-dead mutant did not (Figures 4B and 4C). This is consistent with the observation that the wild-type construct partially restored HR efficiency whereas the methyltransferase-dead (MD) mutant did not (Figure 4D). Five to six clones containing each construct were independently isolated and tested. All WT clones were able to restore HR efficiency, but none of the MD clones showed rescue (Figures S4B and S4C).

The observation that the methyltransferase-dead SETD2 mutant could not rescue HR efficiency suggested that H3K36me3 is required for efficient HR. To test this further, we used SETD2-independent approaches to disrupt H3K36me3 levels. First, we overexpressed KDM4A, an oncogene and a demethylase specific for tri- and dimethylated histone H3K36 and H3K9, predicting that overexpression of KDM4A would reduce the level of H3K36me3 and thus impair HR. Therefore, a doxycycline-inducible KDM4A-expression U2OS cell line was generated. Upon addition of doxycycline, KDM4A was overexpressed, which led to a global reduction in H3K36me3 levels (Figure 4E). We then compared the recruitment of the HR factors in this cell line with or without KDM4A overexpression. We found that KDM4A overexpression inhibited the formation of RPA and RAD51 foci after DNA damage (Figures 4F and 4G). Accordingly, KDM4A overexpression also reduced HR efficiency as assessed by the HR reporter ($p = 0.0078$) (Figure 4H). Using the same cell line without *KDM4A* cDNA integration, we found that doxycycline itself had no effect on HR efficiency of U2OS cells (Figure S4D). These findings indicate that overexpression of the oncogene KDM4A significantly reduces levels of H3K36me3 and HR.

Second, we utilized a dominant-negative mutation in the histone *H3.3* gene (*H3.3K36M*) that was recently shown to reduce H3K36me3 specifically without affecting other histone methylations (Lewis et al., 2013). Accordingly, in cells expressing the mutant *H3.3K36M* transgene, global levels of H3K36me3 were depleted compared to the wild-type H3.3 control (Figure 4I). Cells expressing H3.3K36M also exhibited delayed RPA and RAD51 foci formation following DNA damage (Figures 4J and 4K), consistent with similar delays seen in KDM4A overexpressing cells or SETD2-depleted cells. Neither KDM4A overexpression nor H3.3K36M expression affected cell cycle progression as assessed by BrdU incorporation (Figures S4E and S4F), thus excluding the possibility that the reduction in RPA, RAD51 foci formation and HR efficacy was due to cell cycle effects. The common feature of all three different systems (SETD2 knockdown and rescue, KDM4A overexpression, and H3.3K36M expression) was the reduction in the recruitment of HR proteins and HR efficiency associated with reduced H3K36me3. Together these findings confirm a role for H3K36me3 in HR.

SETD2 Promotes DSB End Resection

Lens epithelium-derived growth factor p75 (LEDGF) binds to H3K36me3 constitutively through its PWWP domain. Upon DNA damage LEDGF recruits CtIP, which facilitates the resection step during HR repair (Daugaard et al., 2012). Accordingly, we were able to coprecipitate LEDGF and H3K36me3, independently of DNA damage in U2OS cells (Figure 5A). LEDGF binding to chromatin was reduced upon SETD2 depletion, consistent with preferential binding of LEDGF to H3K36me3 (Eidahl et al., 2013), and was DSB independent (Figure 5B). We predicted that the reduced LEDGF binding following SETD2 depletion would reduce recruitment of CtIP upon damage. We therefore used micro-irradiation to induce localized DSBs and studied the recruitment of CtIP using a U2OS cell line expressing GFP-tagged CtIP (Sartori et al., 2007). We found CtIP to be localized to sites of DNA damage marked by γH2AX (Figure 5C). In contrast, SETD2 knockdown resulted in reduced recruitment of CtIP to the DNA damage site as marked by γH2AX (Figure 5C). If SETD2 facilitates HR through LEDGF binding and CtIP recruitment, we would expect SETD2 to be epistatic with CtIP. Indeed, codepletion of SETD2 and CtIP (siC+siS) showed the same level of reduction in HR as CtIP-depletion alone (Figures 5D and S5A).

(C) RAD51 foci formation at 4 hr after IR (5 Gy) in cells transfected with SETD2 siRNA. The clones used are the same as in (B). Error bars show SEM from three independent experiments, $^{**}p < 0.01$.

(D) HR efficacy in cells transfected with NT or SETD2 siRNA. The clones used are the same as in (B); error bars show SEM from three independent experiments, $^{***}p < 0.001$.

(E) Western blot showing levels of KDM4A, H3K36me3 and H3 following KDM4A induction with 5 μg/ml doxycycline (DOX) for 72 hr in the KDM4A T-REx U2OS cells (see Experimental Procedures).

(F) RPA32 foci formation at indicated time after 2 hr CPT (10 μM) treatment in KDM4A T-REx cells treated with or without 5 μg/ml doxycycline (DOX). Error bars show SEM from three independent experiments, $^{*}p < 0.05$.

(G) RAD51 foci formation 8 hr after IR (5 Gy) in KDM4A T-REx cells treated with or without 5 μg/ml doxycycline (DOX). Error bars show SEM from three independent experiments. $^{*}p < 0.05$.

(H) HR repair efficacy in KDM4A T-REx DR-GFP U2OS cells treated with or without 5 μg/ml doxycycline (DOX). Error bars show SEM for three independent experiments. $^{**}p < 0.01$.

(I) Western blot showing levels of H3K36me3 and H3 in U2OS cells with stably integrated *H3.3* (control) or *H3.3K36M* lentiviral construct (see Supplemental Experimental Procedures).

(J) RPA32 foci formation at indicated time after 2 hr CPT (10 μM) treatment in H3.3 or H3.3K36M stable expression cells. Error bars show SEM from three independent experiments.

(K) RAD51 foci formation 8 hr after IR (5 Gy) in H3.3 or H3.3K36M stable expression cells. Error bars show SEM for three independent experiments, $^{*}p < 0.05$. See also Figure S4.

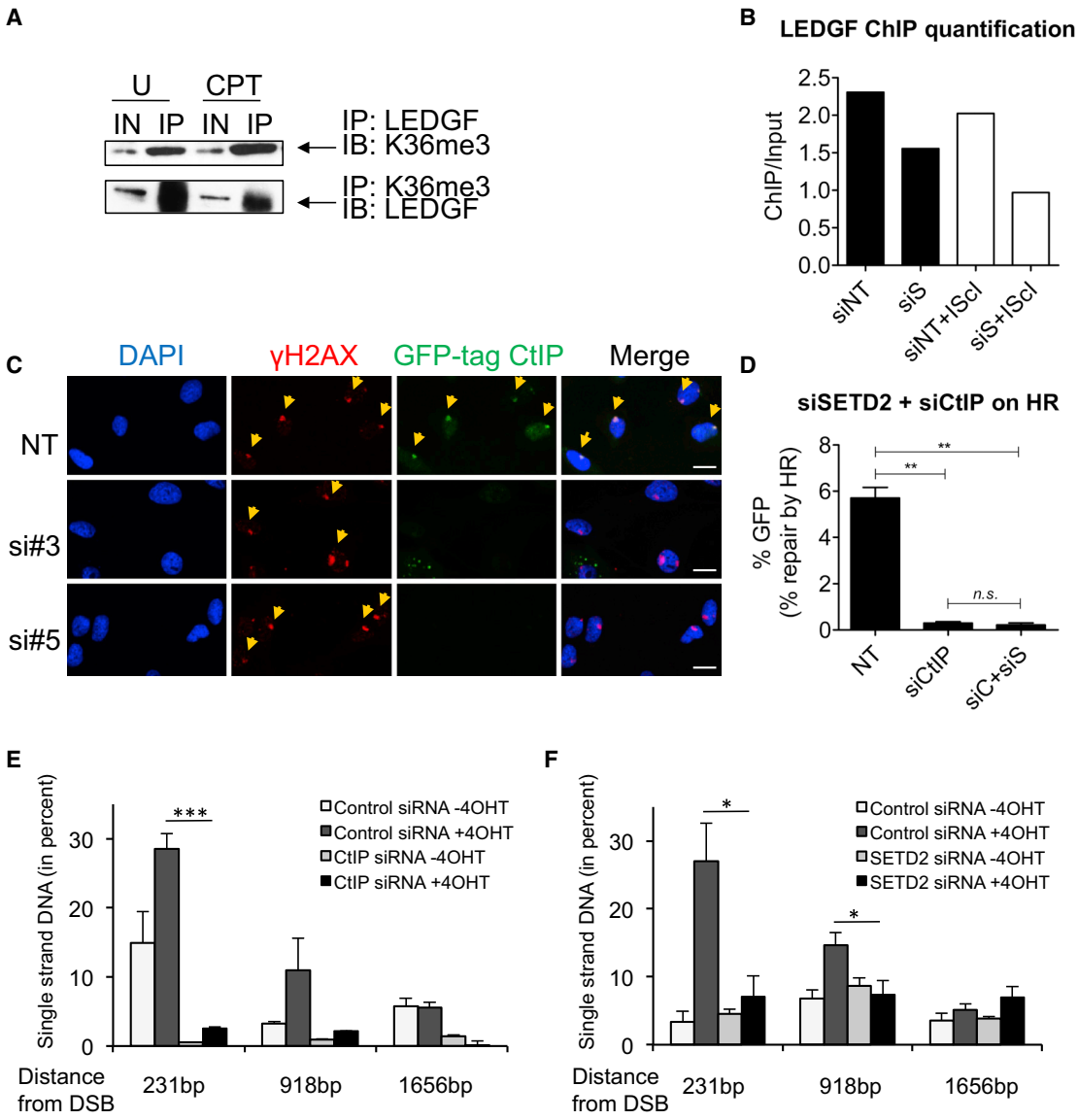


Figure 5. SETD2 Promotes HR through LEDGF/CtIP-Facilitated Resection

(A) Coimmunoprecipitation of LEDGF and H3K36me3 in the presence or absence of DNA damage. U2OS cells were left untreated (U) or treated with CPT (15 μM for 4 hr). Reciprocal co-IPs are shown.

(B) ChIP analysis of LEDGF binding to the DSB site before and after I-SceI induction in NT or SETD2 knockdown DR-GFP cells. siNT, control siRNA-treated cells before cut; siS, SETD2 siRNA-treated cells before cut; siNT+ISceI, control siRNA-treated cells after cut; siS+ISceI: SETD2 siRNA-treated cells after cut. Numbers were quantified from one ChIP-PCR experiment and show LEDGF-ChIP over input.

(C) Microirradiation showing CtIP recruitment to the damage site in U2OS cells treated with nontargeting control siRNA (NT) or SETD2 siRNAs (si#3 and si#5). Fluorescent images were acquired using a confocal microscope (Zeiss); scale bars, 20 μm.

(D) HR efficacy (as measured by GFP reporter assay as in Figure 2D) in cells transfected with control siRNA (NT) or CtIP siRNA (siCtIP) or both CtIP and SETD2 siRNA (siC+siS). Error bars show SEM from three independent experiments, $^{**}p < 0.01$, n.s. not significant.

(E) CtIP depletion impedes resection at an AsiSI induced DSB. DNA was extracted from 4OHT-treated or untreated DIvA cells, transfected with control or CtIP siRNA, as indicated. Resection was analyzed, using a protocol detailed in Zhou et al. (2014), at an AsiSI-induced DSB reported to be repaired by a RAD51-dependent pathway (DSB-II in Aymard et al., 2014). The mean and SEM (n = 4, technical replicate) of a representative experiment are presented.

(F) SETD2 depletion impedes resection at an AsiSI induced DSB. DNA was extracted from 4OHT-treated or untreated DIvA cells, transfected with control or SETD2 siRNA, as indicated. Same as in (E) except that cells were transfected with a siRNA directed against SETD2. The mean and SEM (n = 4, technical replicate) of a representative experiment are presented. See also Figure S5.

These findings are consistent with the hypothesis that SETD2-dependent H3K36me3 promotes recruitment of CtIP by LEDGF to break sites.

Reduced CtIP recruitment, together with the reduced RPA foci formation following SETD2 depletion, suggested that SETD2-dependent H3K36 trimethylation might promote DSB resection.

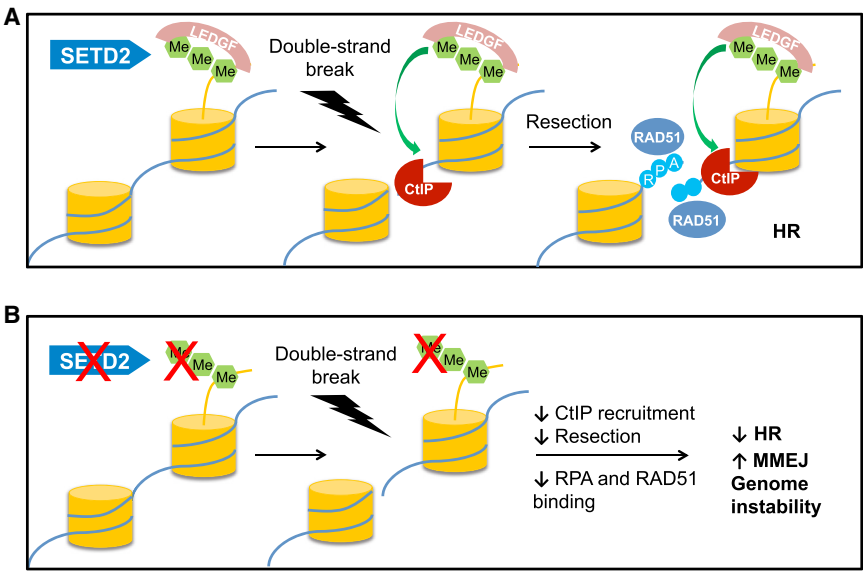


Figure 6. Model for the Role of SETD2-Dependent H3K36me3 in HR and Genome Stability

(A) SETD2-dependent H3K36 trimethylation constitutively recruits LEDGF to chromatin. Following DNA double-strand break, LEDGF recruits CtIP, thereby promoting resection, facilitating RPA and RAD51 recruitment.

(B) In the absence of H3K36 trimethylation, lack of LEDGF chromatin association results in reduced recruitment of CtIP after double-strand breaks, impairing resection and HR, leading to elevated levels of MMEJ and genome instability.

To directly measure resection, we used an in vivo resection assay (Zhou et al., 2014) in which levels of ssDNA were assessed at increasing distances from a specific DSB induced by the AsiSI restriction enzyme, which was found to recruit RAD51 (Iacovoni et al., 2010; Aymard et al., 2014). Following DSB induction (+4OHT), ssDNA was readily detected 231 bp from the AsiSI break site (DSB-II in Aymard et al., 2014) and at reduced levels further from the break site. Following CtIP depletion, ssDNA levels were significantly reduced at 231 bp ($p = 2.2 \times 10^{-5}$) and 918 bp ($p = 0.057$) from the break site compared to control siRNAs (Figure 5E). Following SETD2 depletion, ssDNA levels were also significantly reduced at 231 bp ($p = 0.021$) and 918 bp ($p = 0.037$) from the break site compared to control siRNAs (Figure 5F), and SETD2 depletion did not affect DSB induction at this site as assayed by XRCC4 ChIP (Figure S5B), thus confirming a role for SETD2 in facilitating DSB resection.

DISCUSSION

The *SETD2* gene encoding the H3K36 trimethyltransferase is among the most mutated in human cancers (Lawrence et al., 2014). Yet, how SETD2 functions to suppress the genome instability frequently observed in SETD2-deficient cancer cells is unknown (Sato et al., 2013; Zhu et al., 2014). Here, we describe a role for SETD2-dependent H3K36 trimethylation in facilitating error-free DSB repair and genome stability through efficient HR. Our data support a model in which LEDGF is constitutively anchored to H3K36me3-decorated chromatin through its PWWP domain. Upon DNA damage, chromatin-bound LEDGF recruits CtIP, which promotes resection—an essential step in HR repair (Figure 6A). In the absence of SETD2-dependent H3K36me3, the chromatin association of LEDGF is compromised, and DNA damage-induced CtIP recruitment is impaired, leading to reduced resection and HR and an increase in error-prone MMEJ and subsequent genome instability (Figure 6B). Our data do not, however, exclude the possibility that other H3K36me3 binding factors may also contribute to efficient HR.

The involvement of H3K36me3 in HR raises the question as to whether this histone mark is DSB inducible or preset to determine the choice of repair pathways. We did not detect any appreciable increase in H3K36me3 levels at the break site following I-SceI induction. These findings are consistent with previous observations indicating that H3K36me3 levels do not change following DNA damage, either globally (Fnu et al., 2011) or locally (Aymard et al., 2014; Pei et al., 2011). Given that SETD2-dependent H3K36 trimethylation strongly correlates with actively transcribed regions of the genome (Edmunds et al., 2008; Yoh et al., 2008), our findings thus support a spatial link between HR and transcriptionally active, H3K36me3 decorated regions of the genome. Indeed, because HR repair is usually error free, this would be expected to protect genome integrity following DSB induction or replication fork collapse at such sites, which would be important in maintaining cell homeostasis.

Genome-wide sequencing has suggested certain mutation signatures are associated with specific cancer types (Alexandrov et al., 2013) but lacks the power to identify driver mutations that give rise to that mutation signature. Here, we use a system to determine the impact of individual gene dysfunction on mutation patterns arising during the repair of a single DSB. We show that, by depleting SETD2 or RAD51 HR proteins, DSB induction leads to significantly increased mutation frequencies associated with deletions within six cell cycles (5 days). These findings identify a role for SETD2, like RAD51, in suppressing break-induced mutations. Further mechanistic analysis indicated that deletions arising from SETD2 or RAD51 depletion resulted from a significant increase in error-prone MMEJ. We propose that MMEJ arises from failed HR. SETD2 loss results in partial resection, short ssDNA 3′ ends and reduced RPA foci. These partially resected ends are insufficient to promote efficient HR and are no longer substrates for C-NHEJ. Instead, these partially resected ends become substrates for MMEJ. Similarly, following RAD51 depletion, whereas resection proceeds normally, HR is blocked presynaptically, resulting in the accumulation of longer unrepaired resected ends. Such ends are no longer substrates for C-NHEJ, and, because HR cannot proceed, the ends are instead

rejoined by MMEJ. Extensive resection following RAD51 loss facilitates annealing between regions of microhomology further from the break site, thus leading to larger deletions than those observed for SETD2. The observations that HR was significantly reduced whereas C-NHEJ was unchanged following SETD2 or RAD51 depletion are consistent with this model. Together, these findings suggest a role for HR in maintaining genome stability through suppression of MMEJ. This "real-time" mutation study can therefore provide important insights into how loss of these genes contributes to somatic mutation and intratumor heterogeneity, and it highlights the importance of HR repair genes in maintaining genome stability.

SETD2 is a tumor suppressor and its loss is an important step in the progression of a number of cancer types (Dalgliesh et al., 2010; Fontebasso et al., 2013; Gerlinger et al., 2012). Our findings provide mechanistic insights into how SETD2 functions as a tumor suppressor. Moreover, all three Tudor domain-containing KDM4/JMJD2 proteins (A, B, and C) are frequently overexpressed in various cancers, including breast, colorectal, lung, and prostate (Berry and Janknecht, 2013). Our findings predict that overexpression of these KDM4 family members causes genome instability and tumorigenesis through loss of H3K36me3-mediated HR repair. The identification of a role for H3K36me3 in facilitating HR repair, in addition to its recently described role in mismatch repair (Li et al., 2013), establishes the H3K36me3 chromatin mark as an important functional node for maintaining genome stability.

EXPERIMENTAL PROCEDURES

I-SceI-Inducible HPRT Mutation Assay

The reporter HT1080 cell line with a functional but I-SceI-cleavable *HPRT* gene was generated by inserting an 18 bp I-SceI recognition site into exon 6 of the *HPRT* gene using gene targeting. Reporter cells were transfected with the *I-SceI* plasmid 48 hr after siRNA knockdowns. The cells were allowed to repair in a nonselective medium for 5 days before seeding and selecting for HPRT-negative cells using 6-TG (15 μg/ml).

The mutation frequency was calculated as the number of HPRT-negative colonies divided by number of cells plated, after correcting for plating efficiency. To determine the mutation sequence, individual single-cell colonies were isolated, and genomic DNA spanning the DSB junction was PCR-amplified using the primers flanking the I-SceI site. Sequence alignment of the PCR products was conducted using the DNA Data Bank of Japan (DDBJ) ClustalW program.

In Vivo DNA End Resection Assay

DIvA cells (DSB Inducible via AsiSI [Aymard et al., 2014]) were transfected with siRNA using the Cell Line Nucleofactor kit V (Amaxa) according to the manufacturer's instructions. Forty-eight hours after siRNA transfection, cells were treated or not with 300 nM of 4-hydroxytamoxifen (4OHT) (Sigma; H7904) for 4 hr. DNA was purified using QIAGEN DNeasy kit. Single-stranded DNA generated at an AsiSI-induced DSB was analyzed using the procedure described in Zhou et al. (2014), with the modifications described in Supplemental Experimental Procedures. The percentage of ssDNA was calculated with the following equation: ssDNA% = $1/(2^{(\Delta Ct-1)} + 0.5) \times 100$, where ΔCt is calculated by subtracting the Ct obtained from mock-digested sample from the BanI-digested sample.

Generation of Inducible KDM4A Overexpression Cell Lines

U2OS-Flp-In/T-REx cells express the Tet repressor and contain a single Flp-In site (Gordon et al., 2009). U2OS-Flp-In/T-REx/DRGFP cells were generated by integrating the pDR-GFP expression plasmid (Pierce et al., 1999) into the U2OS-Flp-In/T-REx cells. The *KDM4A* cDNA (Genecopoeia) was cloned into the pDESTfrtto destination vector by a Gateway LR reaction (Life Technologies). U2OS-Flp-In/T-REx or U2OS-Flp-In/T-REx/DRGFP cells were cotransfected with the resulting *KDM4A* destination plasmid and pOG44 (Invitrogen). Stable clones that express the KDM4A protein under the control of the T-REx system were subjected to the treatments indicated.

Data Analysis

Data points from individual assays represent mean ± SEM. Statistical significance between two conditions was assessed by a two-tailed unpaired t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and n.s. represents $p \geq 0.05$.

Full details of all other experimental procedures are given in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and five figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.05.026.

AUTHOR CONTRIBUTIONS

S.X.P., S.A., L.-P.Z., S.S., and F.A. designed and performed experiments. The HPRT$^+$:I-SceI assay was developed by A.C.G.P. The FLIP-IN cells and plasmids were developed by C.Z.B., T.C.H., A.C.G.P., T.H., N.B.L.T., and G.L. supervised the work. T.C.H. and S.X.P. wrote the manuscript with discussion and input from all authors. Experiments in Figure 1 were performed by S.A., with technology developed by A.C.G.P.; Figure 2 was performed by S.X.P. and S.A.; Figure 3 was performed by S.X.P. and L.-P.Z.; Figure 4 was performed by S.S. and S.X.P., with technology developed by C.Z.B.; Figure 5 was performed by S.S., L.-P.Z., S.X.P., and F.A.; Figure 6 was produced by S.X.P. Experiments in Figure S1 were performed by S.A. and S.S.; Figure S2 was performed by S.X.P. and S.A.; Figure S3 was performed by L.-P.Z.; Figure S4 was performed by S.X.P. and S.S.; Figure S5 was performed by F.A. and S.X.P.

ACKNOWLEDGMENTS

We thank David Allis for the *H3.3* and *H3.3K36M* lentiviral vectors; Catherine Millar for the U2OS-Flp-In/T-REx cell line; Atsushi Shibata and Hideaki Hogiwara for the NHEJ reporter cell line; Wojciech Niedzwiedz for the siRNAs against CtIP; and Annika Baude and Jiri Lukas for CtIP-GFP U2OS cells. We thank Mick Woodcock for helping with FACS analysis and Fereydoon Ahrabi for statistical analysis. We thank Peter McHugh and Valentine Macaulay for helpful discussions. This research was supported by the Medical Research Council (C.Z.B., T.H., S.S., T.C.H. grant R066538 and R19583), (L.-P.Z., N.B.L.T. grant ANRUBT00); Cancer Research UK (S.A. grant C5255/A15935) (L.-P.Z., N.L.T. grant 300/A13058); the Clarendon Scholarship (S.A., S.X.P.); the BBSRC (A.C.G.P. grant BB/H003371/1 and C.Z.B. grant BB/K019597/1); the Agence Nationale pour la Recherche (G.L. grant ANR-09-JCJC-0138); Association Contre le Cancer (G.L., F.A.); and the Torsten and Söderberg Foundation (T.H.).

Received: December 20, 2013
Revised: April 16, 2014
Accepted: May 12, 2014
Published: June 12, 2014

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.

Al Sarakbi, W., Sasi, W., Jiang, W.G., Roberts, T., Newbold, R.F., and Mokbel, K. (2009). The mRNA expression of SETD2 in human breast cancer: correlation with clinico-pathological parameters. BMC Cancer *9*, 290.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.L., et al.; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Arnaudeau, C., Lundin, C., and Helleday, T. (2001). DNA double-strand breaks associated with replication forks are predominantly repaired by homologous recombination involving an exchange mechanism in mammalian cells. J. Mol. Biol. *307*, 1235–1245.

Aymard, F., Bugler, B., Schmidt, C.K., Guillou, E., Caron, P., Briois, S., Iacovoni, J.S., Daburon, V., Miller, K.M., Jackson, S.P., and Legube, G. (2014). Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. Nat. Struct. Mol. Biol. *21*, 366–374.

Berry, W.L., and Janknecht, R. (2013). KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. Cancer Res. *73*, 2936–2942.

Black, J.C., Manning, A.L., Van Rechem, C., Kim, J., Ladd, B., Cho, J., Pineda, C.M., Murphy, N., Daniels, D.L., Montagna, C., et al. (2013). KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors. Cell *154*, 541–555.

Couture, J.F., Collazo, E., Ortiz-Tello, P.A., Brunzelle, J.S., and Trievel, R.C. (2007). Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. Nat. Struct. Mol. Biol. *14*, 689–695.

Dalgliesh, G.L., Furge, K., Greenman, C., Chen, L., Bignell, G., Butler, A., Davies, H., Edkins, S., Hardy, C., Latimer, C., et al. (2010). Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature *463*, 360–363.

Daugaard, M., Baude, A., Fugger, K., Povlsen, L.K., Beck, H., Sørensen, C.S., Petersen, N.H., Sorensen, P.H., Lukas, C., Bartek, J., et al. (2012). LEDGF (p75) promotes DNA-end resection and homologous recombination. Nat. Struct. Mol. Biol. *19*, 803–810.

Edmunds, J.W., Mahadevan, L.C., and Clayton, A.L. (2008). Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. EMBO J. *27*, 406–420.

Eidahl, J.O., Crowe, B.L., North, J.A., McKee, C.J., Shkriabai, N., Feng, L., Plumb, M., Graham, R.L., Gorelick, R.J., Hess, S., et al. (2013). Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. Nucleic Acids Res. *41*, 3924–3936.

Fnu, S., Williamson, E.A., De Haro, L.P., Brenneman, M., Wray, J., Shaheen, M., Radhakrishnan, K., Lee, S.H., Nickoloff, J.A., and Hromas, R. (2011). Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. Proc. Natl. Acad. Sci. USA *108*, 540–545.

Fontebasso, A.M., Schwartzentruber, J., Khuong-Quang, D.A., Liu, X.Y., Sturm, D., Korshunov, A., Jones, D.T., Witt, H., Kool, M., Albrecht, S., et al. (2013). Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. Acta Neuropathol. *125*, 659–669.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N. Engl. J. Med. *366*, 883–892.

Gordon, W.R., Vardar-Ulu, D., L'Heureux, S., Ashworth, T., Malecki, M.J., Sanchez-Irizarry, C., McArthur, D.G., Histen, G., Mitchell, J.L., Aster, J.C., and Blacklow, S.C. (2009). Effects of S1 cleavage on the structure, surface export, and signaling activity of human Notch1 and Notch2. PLoS ONE *4*, e6613.

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

Iacovoni, J.S., Caron, P., Lassadi, I., Nicolas, E., Massip, L., Trouche, D., and Legube, G. (2010). High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. EMBO J. *29*, 1446–1457.

Jasin, M. (1996). Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet. *12*, 224–228.

Jeggo, P., and Lavin, M.F. (2009). Cellular radiosensitivity: how much better do we understand it? Int. J. Radiat. Biol. *85*, 1061–1081.

Klose, R.J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature *442*, 312–316.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495–501.

Lewis, P.W., Müller, M.M., Koletsky, M.S., Cordero, F., Lin, S., Banaszynski, L.A., Garcia, B.A., Muir, T.W., Becher, O.J., and Allis, C.D. (2013). Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science *340*, 857–861.

Li, F., Mao, G., Tong, D., Huang, J., Gu, L., Yang, W., and Li, G.M. (2013). The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSα. Cell *153*, 590–600.

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.

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.

Moynahan, M.E., Cui, T.Y., and Jasin, M. (2001). Homology-directed dna repair, mitomycin-c resistance, and chromosome stability is restored with correction of a Brca1 mutation. Cancer Res. *61*, 4842–4850.

Newbold, R.F., and Mokbel, K. (2010). Evidence for a tumour suppressor function of SETD2 in human breast cancer: a new hypothesis. Anticancer Res. *30*, 3309–3311.

Ogiwara, H., Ui, A., Otsuka, A., Satoh, H., Yokomi, I., Nakajima, S., Yasui, A., Yokota, J., and Kohno, T. (2011). Histone acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors. Oncogene *30*, 2135–2146.

Pei, H., Zhang, L., Luo, K., Qin, Y., Chesi, M., Fei, F., Bergsagel, P.L., Wang, L., You, Z., and Lou, Z. (2011). MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. Nature *470*, 124–128.

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.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593–599.

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.

Sato, Y., Yoshizato, T., Shiraishi, Y., Maekawa, S., Okuno, Y., Kamura, T., Shimamura, T., Sato-Otsubo, A., Nagae, G., Suzuki, H., et al. (2013). Integrated molecular analysis of clear-cell renal cell carcinoma. Nat. Genet. *45*, 860–867.

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.

Smeenk, G., and van Attikum, H. (2013). The chromatin response to DNA breaks: leaving a mark on genome integrity. Annu. Rev. Biochem. *82*, 55–80.

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol. Cell. Biol. *22*, 1298–1306.

Sugiyama, T., New, J.H., and Kowalczykowski, S.C. (1998). DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. Proc. Natl. Acad. Sci. USA *95*, 6049–6054.

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

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nat. Rev. Mol. Cell Biol. *13*, 115–126.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467–481.

Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes Dev. *22*, 3422–3434.

Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S.L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M., et al. (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature *481*, 157–163.

Zhou, Y., Caron, P., Legube, G., and Paull, T.T. (2014). Quantitation of DNA double-strand break resection intermediates in human cells. Nucleic Acids Res. *42*, e19.

Zhu, X., He, F., Zeng, H., Ling, S., Chen, A., Wang, Y., Yan, X., Wei, W., Pang, Y., Cheng, H., et al. (2014). Identification of functional cooperative mutations of SETD2 in human acute leukemia. Nat. Genet. *46*, 287–293.