Genetic Requirements for Intra-Chromosomal Deletions

Research Thesis

Presented in partial fulfillment of the requirements for graduation with *research distinction* in the undergraduate colleges of The Ohio State University

by

Matthew Thomas McPherson

The Ohio State University

April, 2020

Project Advisors: Assistant Professor Ruben Petreaca, Department of Molecular Genetics

Assistant Professor Renee Bouley, Department of Chemistry

Abstract

Chromosomal deletions are one of the most dangerous types of DNA damage and often arise as a result of inappropriately repaired DNA double strand breaks (DSB). These breaks are usually formed either in an induced manner from exogenous damage such as radiation, or more commonly caused from spontaneous replication errors. If there is a single strand break during replication and it is not repaired properly, as the replication fork progresses it can lead to the formation of a DSB. When there is a DSB present, there is the opportunity for a chromosomal deletion to occur. If the break is in between non-tandem direct repeats, the DNA repair machinery will degrade what is between the direct repeats through a process called Single Strand Annealing (SSA). This massive loss of DNA is what is known as a chromosomal deletion. Using an assay that in *Schizosaccharomyces pombe*, we can detect DSBs and determine DNA repair pathways through a selection screen of yeast cells with inactivated DNA repair genes. We generated an *in vivo* assay that reports exclusively SSA. We validated the assay by showing its dependence on *rad52*$^{+}$ and independence *rad51*$^{+}$. However, we show that earlier events epistatic to *rad52*$^{+}$ and *rad51*$^{+}$ have differential requirements for deletions vs. other forms of repair. Here, we delineate a more detailed epistatic pathway for intrachromosomal deletions.

ACKNOWLEDGEMENTS
I would like to thank Dr. Ruben Petreaca for allowing me to perform undergraduate research in his lab, as well as being my project advisor. Without his support, I would not have been able to make it this far. Since I started working for him, he has always pushed me to do more than I ever thought that I could. Performing research with Dr. Petreaca has been the most beneficial experience of my undergraduate career.

Thank you to Dr. Renee Bouley for agreeing to be on my thesis committee, I appreciate your willingness to help on short notice.

A special thank you to all of my friends and family who have supported me over the years. There have been ups and downs, but I made it through them with your help. I look forward to what the future holds for all of us.

LIST OF FIGURES

TABLE OF CONTENTS

INTRODUCTION

DNA is required for expressing genes that are necessary for life, it is comprised of three parts: A Pentose Sugar, a Nitrogenous Base and a Phosphate Group. This along with lipids, carbohydrates, and proteins form the building blocks of life. DNA is a polymer built from nucleotide monomers which have a secondary structure with the specific bonding pattern two hydrogen bonds between Adenine and Thymine, and three hydrogen bonds between Guanine and Cytosine. This gives DNA its double helix structure of 34 angstroms per turn [1]. Due to this structure, DNA can be replicated semi-conservatively by becoming single stranded and using one copy as a template to create a second strand. This helps protect DNA from being damaged as it must go through two rounds of replication for the damage to be permanent [2].

All transmissible damage that occurs in DNA is in its primary structure from environmental factors and natural processes. This can either be damage to the sugar-phosphate backbone or the nitrogenous bases [3]. DNA damage can be from exogenous factors, such as: chemicals, radiation, or other external factors. However, a majority of the damage comes from endogenous errors, such as stalled replication forks [4-6]. These can be dangerous due to the fact that they can convert a single strand break (SSB) into a double strand break (DSB). If not repaired properly, DSBs can lead to genetic instability and frequently chromosomal deletions [7-9]. This can often result in a loss of heterozygosity, which may cause an inactivation of tumor suppressors. Without at least one functioning copy of a tumor suppressor, the cell cycle will not be properly regulated. If the cell cycle is not maintained, it will lead to uncontrolled cell growth, which causes the formation of cancer cells [10, 11].

Damage to the DNA bases can take many forms (Figure 1), and often affects translated proteins. If this damage is not properly fixed, it will lead to a mutation, which is a change in the genetic code [12]. This code consists of 64 codons which code for amino acids to build proteins, or stop codons to signal the end of a protein. If there is a missense mutation, one base is substituted for another and results in a different amino acid, which may affect the function of a protein. A silent mutation also consists of the substitution of a base, but since the genetic code is degenerate, the substitution codes for the same amino acid. While missense may have negative consequences, they are not the most dangerous type of mutation. A nonsense mutation codes for a premature stop codon that terminates protein synthesis before completion of it.

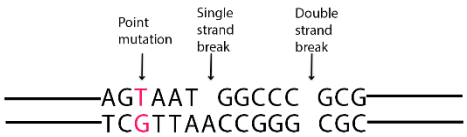

Figure 1. Major types of DNA damage. Three major types of DNA damage can be encountered in a double stranded chromosome. In point mutation the base on one strand has changed such that it violates the rules of base pairing A-T and G-C. In this case C has changed to T. In a single strand break, one of the two strands breaks. This can be easily repaired because the information is preserved on the other strand. In this case a **T** would be inserted opposite **A**. A double strand break constitutes severing of both strands. In this diagram the right part of the chromosome could be lost because it completely dissociates from the left part.

Not all mutations are caused by a base pair substitution, some are a result of an insertion or deletion of a base [13]. This is extremely problematic due to a shift of the three base pair reading frame. This change often leads to the improper coding of multiple amino acids and can potentially create a non-sense mutation. Interestingly, if there are three inserted or deleted bases, the consequences are less severe as the reading frame will not be affected. Another type of damage involving base pairs is thymidine dimers [14]. When UV light shines on the cell, the energy causes two thymine bases to form covalent bonds. This forces a kink in the DNA affecting the three-dimensional structure. In

humans, it is repaired by nucleotide excision repair[15]. This is done by completely removing the damaged section of DNA and synthesizing bases.

Despite the dangers of mutations to DNA bases, breaks between the phosphodiester bonds can be more problematic.

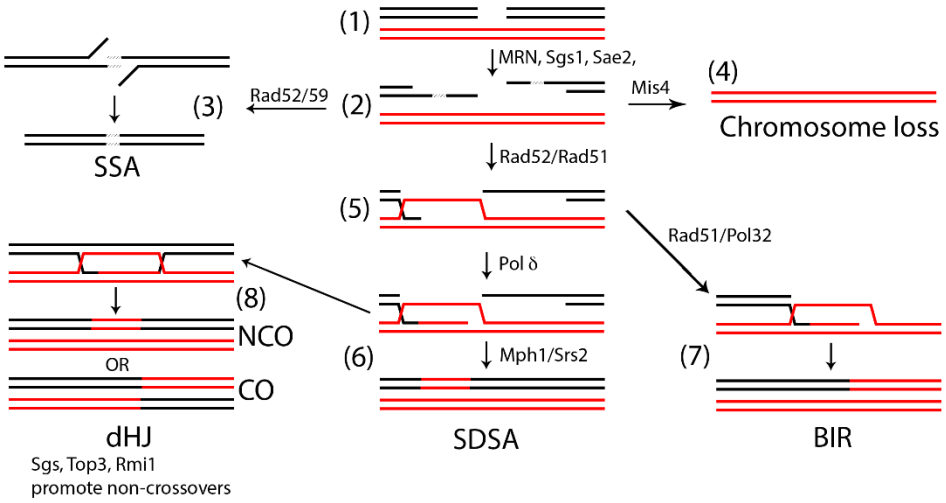

Figure 2. Pathways of repair of DNA double strand breaks. A diploid cell with two homologous chromosomes, black and red, sustains a double strand break (DSB) in the black chromosome (1). The DSB is first resected to expose ssDNA required for invasion of donor regions (2). If direct repeats (shaded areas) exist on the same chromosome, the break may be repaired by single strand annealing (SSA) (3). If repair fails, the chromosome may be lost (4). When homology is found elsewhere or on the other homologue (red), the broken ends may invade this region (the donor sequence) (5). In synthesis-dependent strand annealing (SDSA) (6) the invading strand may copy a small region then release and re-anneal. In break-induced replication (BIR) (7) the invading strand may copy to the end of the red chromosome. In this case the right portion of the broken black chromosome is lost. Occasionally a more complex double Holiday Junction (dJH) may be established (8), the resolution of which can result in crossovers (CO) or non-crossovers (NCO). Note that some of these repair outcomes may lead to loss of heterozygosity meaning that the black sequence has been converted to red. If the red sequence contains a recessive non-functional allele, some of these outcomes will convert the functional black allele to the non-functional red allele resulting in complete inactivation of the gene. Some of the genetic requirements for each pathway are indicated.

Mutations often only affect one gene, but both SSBs and DSBs can affect the entire chromosome. SSBs are less dangerous though as the intact strand can be used as a template to repair the broken strand [3]. However, if not repaired properly, these breaks can become DSBs. This often happens as a result of a replication fork stalling at the break and collapsing [16, 17]. Thankfully, there are several pathways that a cell can utilize to repair this damage (Figure 2)[4, 18-22]. No matter which pathway will be utilized, the first step for repair is to degrade a small amount of DNA on both chromosomes. This is called resection and leads to the creation of "sticky ends," which can be used to reconnect the DNA [23-25]. If the is damage is too significant to repair, the entire chromosome can be lost. Due to this large loss of information, it may be more beneficial for the cell to commit apoptosis, which is programmed cell death. The cell has two known methods of "error prone" break repair. One of them is Non-Homologous End Joining (NHEJ) [26]. This is done by simply connecting the resected chromosomes back together. However, nucleases may degrade a few base pairs, which may result in a frameshift mutation. If there are direct repeats near the DSB, the cell will utilize single strand annealing (SSA) [26, 27]. One unfortunate side effect is that if there is some distance between the repeats, the cell will degrade

the chromosome until they overlap resulting in a deletion. Luckily, there are other pathways that are considered, "error proof" that utilize the homologous chromosome to make repairs [3]. Before any copying of the homologous chromosome takes place, the broken chromosome must first invade it. There are a few different ways that the chromosome can copy from its homologous chromosome. Break induced replication (BIR) starts at the break and copies until the end of the chromosome [28]. While this does prevent chromosome loss, any information after the break will be lost. The closer the break is to the end of the chromosome, the more beneficial this method will be. Perhaps the most accurate form of DSB repair is through Synthesis Dependent Strand Anneal (SDSA). This is accomplished by invading the homologous chromosome and only copying a small portion of the chromosome. This will reduce the chance of having a loss of heterozygosity event as there is less original DNA being replaced. The final pathway is a Double Holiday Junction (DHJ), which can have two results. One can be in the form of a cross over, similar to those found in meiosis. The other is a non-cross over, which has a similar phenotype to SDSA.

***Necessary genetic elements for DNA repair*.**

Resection. Once a DSB has occurred, the cell signals for the MRN complex that consists of Mre11, Rad50, and Nbs1 [29]. The function of this complex is to hold the two halves of the chromosome together, as well as aid in resection [30-32]. After the initial reaction of MRN, Mre11 acts as an endonuclease and exonuclease to create ssDNA on both strands of the broken chromosome. If the resection distance is long, Exo1 and Sae1, which work as more proficient exonucleases [33-36].

Non-Homologous end (NHEJ) joining requires Ku70. For NHEJ to occur, Ku70 must interact with the chromosome to prevent it from excessive degrading [37, 38]. Interestingly, it appears that the MRN complex competes with Ku70 at the break site [30]. MRN favors error proof repair through homologous recombination, while Ku70 favors error prone repair through NHEJ. There is a high probability of NHEJ ending with error as there is no template or proof reading, unlike homologous recombination that is able to utilize the homologous chromosome to prevent errors. Ku70 binds to the broken ends of the chromosome with high affinity. This may be because it is important that the chromosomes are not heavily degraded. Once it has successfully bound to the broken ends of the chromosomes, Ku70 recruits other repair machinery, such as Lig4. This protein is an enzyme that works as a ligase to anneal the two ends back together. Finally, Ku appears to also work in restart or replication forks [39].

Single Strand Annealing (SSA) requires Rad52 and Rad59. Once resection has occurred, the chromosome can be further degraded instead of invading the homologous chromosome. When this happens, the homology is found on the broken chromosome in the form of non-tandem direct repeats, which are interspersed throughout the genome [27]. While the chromosome is being degraded, there are flaps that are formed. Once the direct repeats are overlapping the flaps are degraded, and the chromosome is annealed back together. This results in the DNA between the direct repeats being deleted, which could result in the loss of function if the direct repeats are not located in an intron. This process is mediated by Rad52 and is assisted by Rad59 [40, 41].

Homologous Recombination (HR) is dependent on Rad51 and Rad54. To prevent chromosomal rearrangements, HR allows the broken chromosome to invade the homologous chromosome [42]. However, invasion is only possible in the presence of Rad54 and Rad51 [43, 44]. Rad54 is a chromatin remodeler that removes histones, this loosens the DNA and allows for invasion which is mediated by Rad51 [45]. It is important to note that Rad52 loads Rad51 onto the broken chromosome [46-49]. Once

the DNA invades the homologous chromosome; the broken chromosome can copy the missing segments. This process also inhibits DNA repair through crossovers further preventing translocations [50-53]. It is important to note that once the missing segments have been repaired, the chromosome must be released from the homologous chromosome and the histones must also be replaced in the proper order.

DNA double strand break formation. A common model for DNA DSBs shows the chromosomes as a two-ended break from exogenous damage. However, large amounts of chromosome breaks are also one-ended. These are a result of DNA replication forks stalling at the nick and collapsing leaving one of the daughter strands incomplete (Figure 3) [54-57]. This type of break is often repaired by Break Induced Replication (BIR) [58]. This is because there is no second half to be annealed back together, and it must completely rely on the homologous chromosome to make repairs.

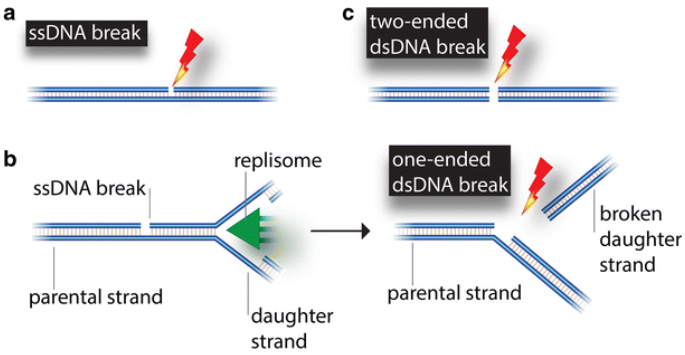

Figure 3. Spontaneous and induced DNA double strand breaks. A. The chromosome receives a nick. **B** A replication fork approaches the nick. Once the replication machinery reaches the nick, it can no longer proceed and causes the fork to stall and collapse. This leaves one of the daughter strains unfinished, since there is only one end for the break, it can be repaired through BIR. **C.** DNA is damaged through an exogenous source such as radiation. The chromosome is cleaved in two similar to **Fig. 2** where it can go through the different mechanisms shown.

Growth of Cancer Cells.

Cancer is defined as the uncontrolled growth of cells, and are frequently a result of chromosomal rearrangements, especially deletions [59-61]. Even though 90% of all mutations occur by the age of 20, it is typically the last 10% of mutations that lead to cancer [62]. The majority of these are a result of endogenous factors such as replication errors. However, exogenous sources of damage, such as radiation will also contribute to these mutations. Thankfully, there are natural processes that help prevent the formation of cancer cells.

Cell cycle checkpoints. To prevent damaged DNA from continuing in the cell cycle, cells have developed checkpoints [63, 64]. At these points in the cell cycle, the cell checks for errors, if there are none, the cycle will continue normally [65-68]. However, if the cell detects problems it must decide whether to make repairs or commit apoptosis, which is programmed cell death. Even though large amounts of apoptosis may be detrimental, it is hypothesized that it is more beneficial than to allow mutated cells to continue to replicate and divide. These checkpoints are regulated by two types of genes, and if problems arise in them, they can lead to the development of cancer cells [69].

Tumor Suppressors slow the cell cycle. These genes work by stopping the cell cycle at the checkpoints and make the decision whether to continue through the cell cycle, repair any damage, or commit apoptosis [9]. Typically, people have two copies of these genes, one from each parent. However, these genes can get mutations that deactivate them, or they can be deleted through repair pathways such as SSA [70]. As long as one copy is functional, the cell cycle will function normally. If someone only inherits one copy they are at a higher risk of cancer, this is because a single mutation can deactivate their tumor suppressor function. People who develop cancer at a young age often only have one copy of a tumor suppressor.

Oncogenes stimulate the cell cycle. When cells are free of errors or damage, oncogenes will signal for the cell to continue through the cell cycle. One problem with these genes is that if they receive a mutation, they can be permanently activated [71]. Unlike tumor suppressors that can be safe with only one wild type gene, if an oncogene is continually expressed it will supersede the normal copy. This will cause the cell cycle to continue, even if there is damage, and will prevent tumor suppressors from stopping the cell cycle. The uncontrolled growth will ultimately lead to cancer.

Chromosomal deletions can lead to cancer cells. When comparing the chromosome sequences of regular cells to cancer cells, there can be many different types of abnormalities including translocations, inversions, duplications, and deletions [72]. Translocations are a result of two non-homologous chromosomes swapping portions of DNA. Inversions involve the change of direction of a piece of the chromosome. Duplications cause part of the chromosome to be repeated. While these abnormalities can be detrimental to the cell, there is little to no loss of DNA. However, chromosomal deletions may remove large segments of DNA that may encode for essential proteins, such as cell cycle regulators [73]. These deletions may cause Loss of Heterozygosity (LOH), which results in a complete inactivation of one copy of a gene. However, as long as one gene is functional, proper cell cycle regulation will occur. If someone inherits only one copy of these cell cycle regulators from their parents, then a chromosome deletion could remove their only function gene. Without proper regulation of cell growth, it becomes more likely that cancer cells will develop. These deletions can occur through DNA repair pathways such as SSA and other related pathway s[26] This is a result of utilizing direct repeats within the broken chromosome as a way to search for homology. Once homology has been found, the segments of DNA in between the direct repeats will be degraded leading to a chromosomal deletion.

Testable hypothesis. To analyze DNA pathways that lead to deletion, we have been using *S. pombe* to create mutations in genes that have been shown to be responsible for resection, homology search, the SSA pathway, the BIR pathway, and NHEJ. These genes have been crossed with an assay to test for rates of deletion that arise from spontaneous chromosomal DSBs. The Petreaca lab has shown that this assay will only report deletion outcomes from repair. Our goal was to delineate the recombination pathways that lead to deletions. **Based on our preliminary data, I hypothesize that the deletions caused from the DNA repair is a result of SSA**.

MATERIALS AND METHODS

Strains. The strains used in this study are listed in **Table 1**. To study chromosomal deletions, we used *S. pombe* as our model organism that contained the *ura4-his3-ura4* cassette, as well as a HO endonuclease to simulate exogenous damage [74]. This assay that utilizes powerful yeast genetics has been published and has been proven to report only deletions for both spontaneous and induced DSBs. However, this project focuses mainly on spontaneous breaks.

Procedure for ura-his-ura assay.

1. First, we streak out colonies frozen in glycerol from a -70^{o}C freezer on an EMM -His plate, to ensure that the cells have not been converted into -his +ura. Colonies should appear after 4-5 days.
2. Once there are single colonies, suspend 10 colonies into 10 Eppendorf tubes, each with 100µL of water.
3. Next, vortex the tubes and count using a hemocytometer. Calculate cells per µL and place 100 cells/µL into test tubes containing 4mL of EMM+ HisUraLeuLysAdeArg.

4. Place in a 32°C incubator for approximately 48 hours, ensuring that the tubes are rotating.
5. After the incubation time, vortex the tubes and count under a hemocytometer.
6. Plate 10^3 cells on YES plates as a control. **7**.Plate 10^5-10^7 on EMM -Ura plates with Phloxin B. This is to prevent false positive as -Ura yeast that cannibalize to survive will turn bright red in the presence of Phloxin B. We used the 150x50mm plates when plating high frequencies on -Ura plates.
7. After plating, place the plates in a 32°C incubator until colonies appear. Then, count the numbers of colonies on each plate and record them.

Table 1. Strains used in this study

Identifier	***Genotype***	***Source***
RCP 270	h- rad32::KanR ura4::ura4-his3-HO-ura4 his3-D1 ura4-D18 leu1-32 ade6-M210	
RCP 271	h+ rad32::KanR ura4::ura4-his3-HO-ura4 his3-D1 ura4-D18 leu1-32 ade6-M210	
RCP 300	h- nbs1::kanMX ura4::ura4-his3-HO-ura4 his3-D1 leu1-32	
RCP 301	h- nbs1::kanMX ura4::ura4-his3-HO-ura4 his3-D1 leu1-32	
RCP 389	cdc27-D1 ura4::ura4-his3-HO-ura4 ura4-D18 his3-D1 ade6-M210	
RCP 390	cdc27-D1 ura4::ura4-his3-HO-ura4 leu1-32 ura4-D18 his3-D1 ade6-M210	
RCP 436	mus81::kanMX ura4::ura4-his3-HO-ura4 leu1-32 his3-D1 ade6-M210	
RCP 471	Δslx4::kanMX4 ura4::ura4-his3-HO-ura4 his3-D1 leu1-32 his3-D1 ade6-M216	
RCP 474	rad50::KanMX6 ura4::ura4-his3-HO-ura4 his3-D1 leu1-32 ade6-M216	
RCP 475	rad50::KanMX6 ura4::ura4-his3-HO-ura4 his3-D1 leu1-32 ade6-M216	
RCP 478-480	ctp1::kanMX6-Bioneer ura4::ura4-his3-HO-ura4 leu1-32 ade6-M216 his3-D1	

Data analysis. To calculate the rate of deletions, the number of colonies on the EMM -Ura plates were devided by the number of colonies on the YES plates devided by 1000. This is to ensure that plating error is accounted for when examining the rate of chromosomal deletions. All calculations were normalized at a frequencey of 10^5.

CRISPR Cas9 make mutations.

1. Streak out a wild type strain from the -70°C freezer, and incubate at 32°C until single colonies appear. Then, take a single colony and place it in a test tube with 4mL of YES liquid solution, and incubator for approximately 24 hours.
2. Then vortex the tube, and transfer 50µL into a 250mL flask containg 50mL of YES liquid solution. Place in a shaker at 250RPM at 32°C for approximately 16 hours.
3. To ensure that the cells are in the log phase of growth, place the flask in ice once it is removed. Pour the liquid into a sterile 50mL centrifuge tube, and centrifuge at 1600RPM at 4°C and discard the supernant. Wash the cells with cold, sterile water, centrifuge again, and discard the supernant. Wash the cells again with a total of 250mL of cold, sterile 1.2M sorbitol, this will take 5 separate washes.
4. Add 500µL of cold, sterile sorbitol to the centrifuge tube and vortex. Add 200µL of cells to an epindorf tube on ice with 2µL of each of the disired primers, 1µL of guide RNA, and 8µL of the CRISPR Cas9 protein.

5. Pipette all 213µL of solution from the epindorf tube into a cold, sterile cuvette, and place it into the electrophorator set to Fungi, SHS. Pulse the cells and immediately add 200µL of cold, sterile sorbitol, and plate onto an EMM -ura plate.
6. Allow it to incubate at 32°C until colonies appear.
7. To verify the mutations, take colonies that have appeared and place them into test tubes containing 4mL of YES liquid solution and incubate at 32°C for approximately 24 with rotation.
8. After this period, perfom DNA extraction and purify them for PCR. Run a DNA gel electrophoresis to ensure that the genes are the proper length.
9. Then, perform a Restriction Length Fragment Polymorphism (RLFP) by treating the potential mutants and wild type with a restriction enzyme. If the bands of DNA differ in length from the wild type, then send the PCR of the gene to sanger sequencing for official results.

RESULTS AND DISCUSSION

Generation and validation of the ura-his-ura cassette to study single strand annealing. To study chromosomal deletions, the Petreaca lab has designed an *in vivo* assay. This consists of two non-functional ura4 genes that share direct repeats with 200 base pairs of overlap. In between these two, there is a functional *his3* gene (Figure 4A) [54]. A deletion can be detected through loss of function of the *his3* gene and gain of function from the *ura4* gene. The deletion may occur spontaneously, or a break can be induced using the *S. cerevisiae* HO endonuclease, a restriction endonuclease with a 50bp restriction site [73, 76, 77]. The long restriction site ensures that this enzyme cuts only once in the yeast genome.

As expected, induced recombinants arise at a much higher rate than spontaneous recombinants (Figure 4B). Spontaneous recombinants are presumed to arise due to random breaks that occur during DNA replication.

To ensure that our assay reports SSA, we used primers that amplify the entire *ura-his-ura* cassette (Figure 4A half arrows). We find that pre-recombinants PCR amplicons are longer than post-recombinants (Figure 4C). The sizes of the PCR products correspond to the size of the *ura-his-ura* cassette or the reconstituted *ura4*$^{+}$ gene. We also showed that no other recombination outcome is possible (e.g. SDSA, BIR, etc) in this system [74] as it is possible in other similar systems [78]. Thus, our system reports exclusively deletions.

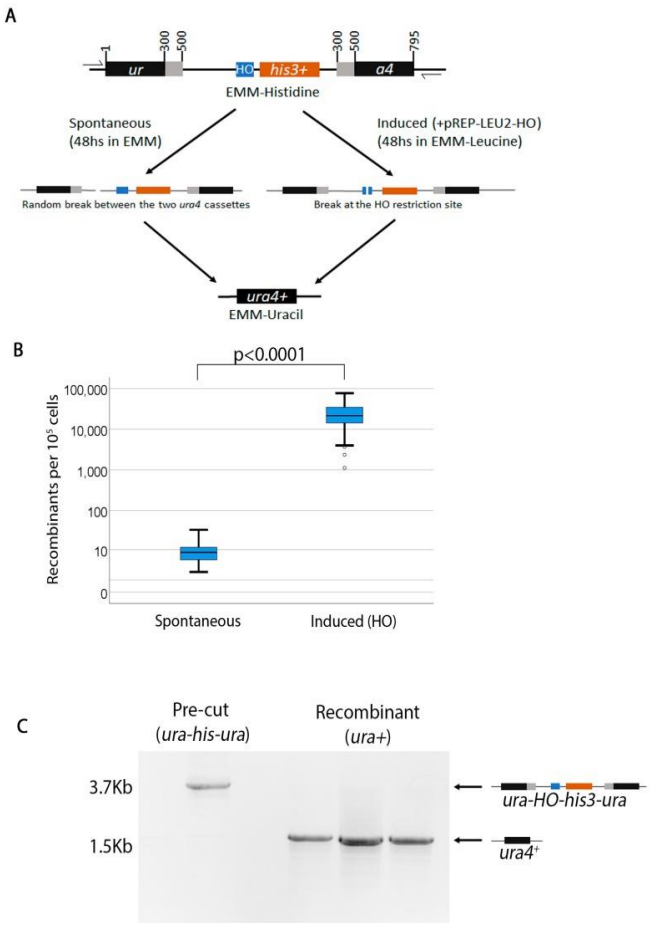

Figure 4. An assay to study spontaneous and induced double-strand breaks at regions of non-tandem repeats. (**A**). The *ura-his-ura* assay. In this assay, two non-functional *ura4* alleles flank a functional *his3*$^{+}$ allele. The *ura4* alleles have 200 bps of identical overlapping sequences, creating two non-tandem repeats (gray areas). The *S. cerevisiae* homothallic endonuclease (*HO*) is cloned just upstream of the *his3*$^{+}$ gene. Spontaneous *ura4*$^{+}$*his3*$^{-}$ recombinants are assayed by growing cells in EMM+UraHisAdeLeu media for 48 h then plating on selective EMM-Uracil. Induced break recombinants are assayed by growing cells for 48 h in media without thiamine to de-repress the *HO* endonuclease, while maintaining selection for the plasmid (EMM-Leucine). Cells are then plated on EMM-Uracil. All experiments were performed at 32 °C. (**B**). Box plot showing the frequency of recombinants for both induced and spontaneous breaks. (**C**). PCR across the *ura-his-ura* cassette in both pre- and post-recombination strains. Half arrowheads in (**A**) show approximate positions of primers.

SSA is dependent on Rad52 but not Rad51 or Ku [27]. To understand whether the deletions reported by our assay occur by SSA, we monitored recombination in strains lacking Rad52, Rad51 and Ku. We conclusively showed that deletions are dependent on Rad52 but not Ku or Rad51 (Figure 5). In fact, Rad51 and Ku inhibit deletions consistent with the fact that SSA is a backup pathway and repair is initially biased through other more efficient pathways (Figure 2). Further, BIR which is known to rescue replication forks is not required for generating deletions in our assay. Cdc27 is a gene central to BIR and a mutation in this gene does not affect recombination outcomes (Table 3). Taken together these data show that our assay reports SSA.

Genetic requirements for SSA.

We next wanted to understand a genetic pathway for SSA besides the role of Rad52. The goal was to understand how the cell makes the choice between SSA and other pathways. We tested several recombination genes for their role in SSA.

<u>*Genes used in this study to understand their role in SSA*</u>. To determine what the genetic requirements are for deletions arising from spontaneous breaks, we specifically chose genes that have been shown to be required for DNA repair (Table 2). One set of genes has been proven to be required for resection, as this is the first step of homologous recombination, we should expect that the rate of deletion will decrease. Another set of genes is required for homology search, invasion of the homologous chromosome, and

copying of the sequence. We also selected genes involved the specific pathways: single strand annealing, break induced replication, and non-homologous end Joining. It is important to note that some of these genes are used in more than one process.

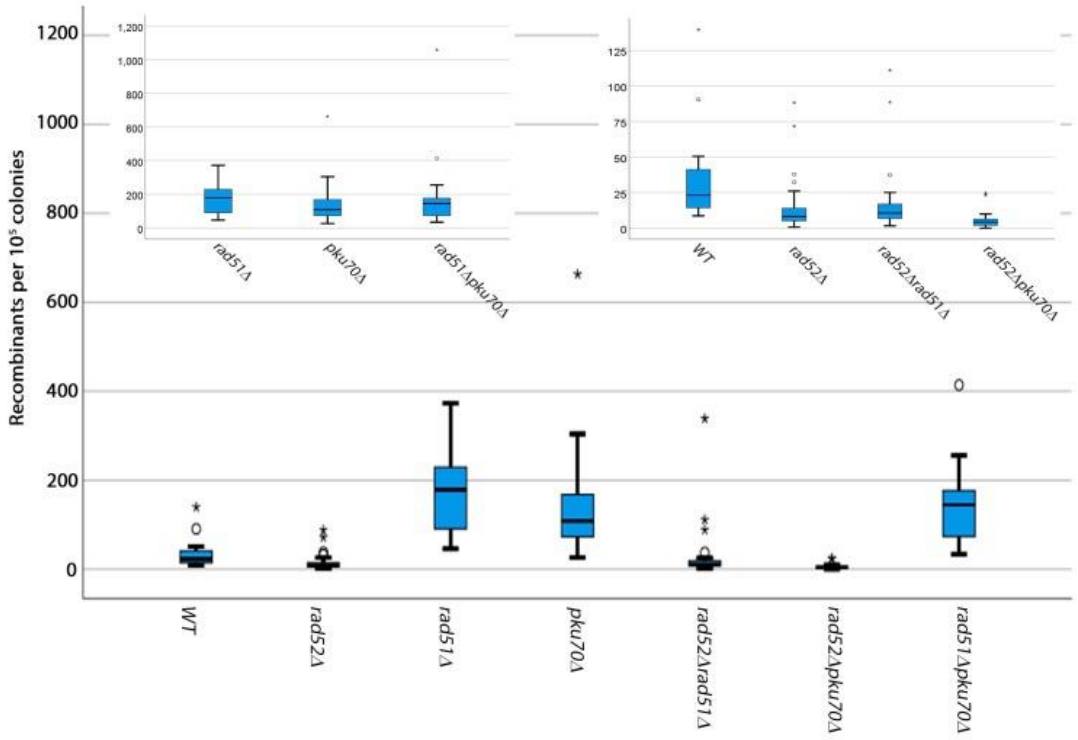

Figure 5. Frequencies of chromosomal deletions from spontaneous break repair in various recombination mutants. All counts are normalized to a rate of 10^5.

<u>*Analyzing resection genes.*</u> As resection is the first step that needs to be completed, we should expect that they inactivation of these genes should reduce the number of deletions, if SSA is the pathway that is leading to the deletions. One of the main factors for resections is the MRN complex, which consists of, Mre11, Rad50, and Nbs1. When these genes are not functional, we saw a statistically significant increase in deletions compared to wild type. This is the opposite of what we would expect to observe. Perhaps even more interesting, Mre11 led to a larger increase in deletions than Rad50 or Nbs1 (Table 3). Further, Ctp1 which has been proposed to function in the same pathway with MRN shows yet another phenotype. These data resection genes may play a role in biasing repair towards the various repair pathways.

<u>*Single strand annealing (SSA) genes.*</u> Our preliminary data using the assay showed that deletions that arise from spontaneous breaks are dependent on rad52, but independent of rad51 (Figure 5). This led us to believe that our assay was following the SSA pathway. However, *slx4* and *mus81* have also been proposed to be required for SSA but the assay used for these studies could not differentiate between deletion through SSA or deletion through other pathways [45]. Surprisingly, the *slx4* mutant led to a high frequency of recombination. This evidence shows that spontaneous deletions do not rely on the function of *slx4*. On the other hand, the function of *mus81* is separable from that of *slx4* because the effect on recombination is smaller.

CONCLUDING REMARKS

Here, we investigated the role of various DNA damage genes in SSA. We found primarily that resection genes may be responsible for biasing repair towards SSA. Resection is required for exposing single stranded DNA which is used for homology search. It is possible that resection has to be more extensive for SDSA and BIR than for SSA. This is because in SSA the repeat is nearby whereas in the other repair models homology has to be found elsewhere. We propose that for SDSA and BIR, a small region is initially resected and homology search ensues. However, if no homology is found, a longer region is resected and so forth. Thus, resection factors may be essential for error-proof repair.

Table 2. Genes that are required for various homologous recombination pathways and processes.

rad50	These genes have been shown to be required for resection Fig. 1(2). When a double strand break occurs, the ends are blunt. In order for recombination to occur the DNA must be made single stranded by resecting one of the strands.
mre11	
nbs1	
srs2	
rqh1	
rad50	These genes have been shown to be required for homologous homology search, invasion of the homologous region, copying of the sequence and resolution Fig. 1(6,8)
rad51	
rad54	
rad55	
rad57	
sgs1	
rad50	These genes have been shown to be required for single strand annealing
rad52	
pol32	These genes have been shown to be required for break induced replication
polδ	
ku	These genes have been shown to be required for nonhomologous end joining
lig4	

Remarkably, the function of the resection factors is separable meaning that not all resection factors function the same. A statistical comparison between the various resection factors shows that their role in SSA is statistically significant (Table 4). Remarkably, analysis of cancer genomes also shows that various mutations in resection factors cause different forms of chromosomal instability (data not shown).

To fully understand the genetic requirements for our assay, further research will need to be conducted. So far, the only gene we have determined to be required for deletions that arise from spontaneous DNA DSBs is rad52. However, the data from slx4 and mus81 has given us insight that it may not be SSA as we previously hypothesized. To determine the differences between mre11, nbs1, and rad50, double mutants will be able to determine their epistatic interactions. Continuing the deletion screen using chromatin remodelers such as Hip1, may help us determine what genes are contributing to these deletions. Since all of the genes selected for this study increased the rate of deletions, it appears that the cell favors error proof repair for spontaneous breaks.

Table 3. Statistics for chromosomal deletions in spontaneous breaks, *ura-his-ura* assay

	N	Mean		Std. Deviation
Gene	**Statistic**	**Statistic**	**Std. Error**	**Statistic**
WT	36	30.70	4.20	25.20
ctp1	30	61.81	17.9	107.94
rad32	34	614.5	153.6	896
nbs1	30	31.9	6.34	34.7
rad50	39	68.3	16.48	102.9
mus81	49	165.97	40.46	283.2
slx4	33	1873.2	566.7	3255.7
cdc27-D1	38	437.4	162.3	1000.4

Table 4. Two-tailed P values for independent samples t-test, *ura-his-ura* assay							
	WT	*ctp1*	*rad32*	*nbs1*	*rad50*	*mus81*	*slx4*
ctp1	P= 0.011						
rad32	P=0.0005	P=0.0014					
nbs1	P=0.0012	P=0.1539	P=0.0007				
rad50	P=0.0028	P=0.8001	P=0.0003	P=0.0679			
mus81	P=0.0035	P=0.0577	P=0.0015	P=0.0119	P=0.0437		
slx4	P=0.0027	P=0.0034	P=0.0335	P=0.003	P=0.0009	P=0.0004	
cdc27-D1	P=0.0225	P=0.045	P=0.4336	P=0.0302	P=0.0247	P=0.0737	p=0.0119

REFERENCES

1. Chaudhry R, Khaddour K: **Biochemistry, DNA Replication**. In: *StatPearls.* Treasure Island (FL); 2020.
2. Snedeker J, Wooten M, Chen X: **The Inherent Asymmetry of DNA Replication**. *Annu Rev Cell Dev Biol* 2017, **33**:291-318.
3. Mehta A, Haber JE: **Sources of DNA double-strand breaks and models of recombinational DNA repair**. *Cold Spring Harb Perspect Biol* 2014, **6**(9):a016428.
4. Ait Saada A, Lambert SAE, Carr AM: **Preserving replication fork integrity and competence via the homologous recombination pathway**. *DNA Repair (Amst)* 2018, **71**:135-147.
5. Nikolov I, Taddei A: **Linking replication stress with heterochromatin formation**. *Chromosoma* 2016, **125**(3):523-533.
6. Cavalier-Smith T: **Origins of the machinery of recombination and sex**. *Heredity (Edinb)* 2002, **88**(2):125-141.

7. Cai Y, Sablina AA: **Cancer-associated chromosomal deletions: Size makes a difference**. *Cell Cycle* 2016, **15**(21):2850-2851.
8. Lasko D, Cavenee W, Nordenskjold M: **Loss of constitutional heterozygosity in human cancer**. *Annu Rev Genet* 1991, **25**:281-314.
9. Stark JM, Pierce AJ, Oh J, Pastink A, Jasin M: **Genetic steps of mammalian homologous repair with distinct mutagenic consequences**. *Mol Cell Biol* 2004, **24**(21):9305-9316.
10. Ortega V, Chaubey A, Mendiola C, Ehman W, Jr., Vadlamudi K, Dupont B, Velagaleti G: **Complex Chromosomal Rearrangements in B-Cell Lymphoma: Evidence of Chromoanagenesis? A Case Report**. *Neoplasia* 2016, **18**(4):223-228.
11. Janssen A, van der Burg M, Szuhai K, Kops GJ, Medema RH: **Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations**. *Science* 2011, **333**(6051):1895-1898.
12. Ovejero S, Bueno A, Sacristan MP: **Working on Genomic Stability: From the S-Phase to Mitosis**. *Genes (Basel)* 2020, **11**(2).
13. Garcia-Diaz M, Kunkel TA: **Mechanism of a genetic glissando: structural biology of indel mutations**. *Trends Biochem Sci* 2006, **31**(4):206-214.
14. Hemminki K, Snellman E: **How fast are UV-dimers repaired in human skin DNA in situ?** *J Invest Dermatol* 2002, **119**(3):699; discussion 700-692.
15. Spivak G: **Nucleotide excision repair in humans**. *DNA Repair (Amst)* 2015, **36**:13-18.
16. Prado F: **Homologous Recombination: To Fork and Beyond**. *Genes (Basel)* 2018, **9**(12).
17. Prado F: **Homologous recombination maintenance of genome integrity during DNA damage tolerance**. *Mol Cell Oncol* 2014, **1**(2):e957039.
18. Aylon Y, Kupiec M: **DSB repair: the yeast paradigm**. *DNA Repair (Amst)* 2004, **3**(8-9):797-815.
19. Tabancay AP, Jr., Forsburg SL: **Eukaryotic DNA replication in a chromatin context**. *Curr Top Dev Biol* 2006, **76**:129-184.
20. Sugawara N, Haber JE: **Repair of DNA double strand breaks: in vivo biochemistry**. *Methods Enzymol* 2006, **408**:416-429.
21. Kowalczykowski SC: **An Overview of the Molecular Mechanisms of Recombinational DNA Repair**. *Cold Spring Harb Perspect Biol* 2015, **7**(11).
22. Ranjha L, Howard SM, Cejka P: **Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes**. *Chromosoma* 2018, **127**(2):187-214.
23. Symington LS: **Mechanism and regulation of DNA end resection in eukaryotes**. *Crit Rev Biochem Mol Biol* 2016, **51**(3):195-212.
24. Symington LS, Gautier J: **Double-strand break end resection and repair pathway choice**. *Annu Rev Genet* 2011, **45**:247-271.
25. Ira G, Pellicioli A, Balijja A, Wang X, Fiorani S, Carotenuto W, Liberi G, Bressan D, Wan L, Hollingsworth NM *et al*: **DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1**. *Nature* 2004, **431**(7011):1011-1017.
26. Chang HHY, Pannunzio NR, Adachi N, Lieber MR: **Non-homologous DNA end joining and alternative pathways to double-strand break repair**. *Nat Rev Mol Cell Biol* 2017, **18**(8):495-506.
27. Bhargava R, Onyango DO, Stark JM: **Regulation of Single-Strand Annealing and its Role in Genome Maintenance**. *Trends Genet* 2016, **32**(9):566-575.
28. Anand RP, Lovett ST, Haber JE: **Break-induced DNA replication**. *Cold Spring Harb Perspect Biol* 2013, **5**(12):a010397.
29. Lisby M, Rothstein R: **DNA repair: keeping it together**. *Curr Biol* 2004, **14**(23):R994-996.
30. Langerak P, Mejia-Ramirez E, Limbo O, Russell P: **Release of Ku and MRN from DNA Ends by Mre11 Nuclease Activity and Ctp1 Is Required for Homologous Recombination Repair of Double-Strand Breaks**. *Plos Genetics* 2011, **7**(9).

31. Langerak P, Mejia-Ramirez E, Limbo O, Russell P: **Release of Ku and MRN from DNA ends by Mre11 nuclease activity and Ctp1 is required for homologous recombination repair of double-strand breaks**. *PLoS Genet* 2011, **7**(9):e1002271.
32. Syed A, Tainer JA: **The MRE11-RAD50-NBS1 Complex Conducts the Orchestration of Damage Signaling and Outcomes to Stress in DNA Replication and Repair**. *Annu Rev Biochem* 2018, **87**:263-294.
33. Tran PT, Erdeniz N, Symington LS, Liskay RM: **EXO1-A multi-tasking eukaryotic nuclease**. *DNA Repair (Amst)* 2004, **3**(12):1549-1559.
34. Makharashvili N, Tubbs AT, Yang SH, Wang H, Barton O, Zhou Y, Deshpande RA, Lee JH, Lobrich M, Sleckman BP *et al*: **Catalytic and noncatalytic roles of the CtIP endonuclease in double-strand break end resection**. *Mol Cell* 2014, **54**(6):1022-1033.
35. Tomita K, Matsuura A, Caspari T, Carr AM, Akamatsu Y, Iwasaki H, Mizuno K, Ohta K, Uritani M, Ushimaru T *et al*: **Competition between the Rad50 complex and the Ku heterodimer reveals a role for Exo1 in processing double-strand breaks but not telomeres**. *Mol Cell Biol* 2003, **23**(15):5186-5197.
36. Pierce AJ, Johnson RD, Thompson LH, Jasin M: **XRCC3 promotes homology-directed repair of DNA damage in mammalian cells**. *Genes Dev* 1999, **13**(20):2633-2638.
37. Clerici M, Mantiero D, Guerini I, Lucchini G, Longhese MP: **The Yku70-Yku80 complex contributes to regulate double-strand break processing and checkpoint activation during the cell cycle**. *EMBO Rep* 2008, **9**(8):810-818.
38. Pastwa E, Blasiak J: **Non-homologous DNA end joining**. *Acta Biochim Pol* 2003, **50**(4):891-908.
39. Teixeira-Silva A, Ait Saada A, Hardy J, Iraqui I, Nocente MC, Freon K, Lambert SAE: **The end-joining factor Ku acts in the end-resection of double strand break-free arrested replication forks**. *Nat Commun* 2017, **8**(1):1982.
40. Mortensen UH, Lisby M, Rothstein R: **Rad52**. *Curr Biol* 2009, **19**(16):R676-677.
41. Brouwer I, Zhang H, Candelli A, Normanno D, Peterman EJG, Wuite GJL, Modesti M: **Human RAD52 Captures and Holds DNA Strands, Increases DNA Flexibility, and Prevents Melting of Duplex DNA: Implications for DNA Recombination**. *Cell Rep* 2017, **18**(12):2845-2853.
42. Osman F, Fortunato EA, Subramani S: **Double-strand break-induced mitotic intrachromosomal recombination in the fission yeast Schizosaccharomyces pombe**. *Genetics* 1996, **142**(2):341-357.
43. Shibata A: **Regulation of repair pathway choice at two-ended DNA double-strand breaks**. *Mutat Res* 2017, **803-805**:51-55.
44. Tarsounas M, Davies AA, West SC: **RAD51 localization and activation following DNA damage**. *Philos Trans R Soc Lond B Biol Sci* 2004, **359**(1441):87-93.
45. Onaka AT, Toyofuku N, Inoue T, Okita AK, Sagawa M, Su J, Shitanda T, Matsuyama R, Zafar F, Takahashi TS *et al*: **Rad51 and Rad54 promote noncrossover recombination between centromere repeats on the same chromatid to prevent isochromosome formation**. *Nucleic Acids Res* 2016, **44**(22):10744-10757.
46. Nakamura K, Okamoto A, Katou Y, Yadani C, Shitanda T, Kaweeteerawat C, Takahashi TS, Itoh T, Shirahige K, Masukata H *et al*: **Rad51 suppresses gross chromosomal rearrangement at centromere in Schizosaccharomyces pombe**. *EMBO J* 2008, **27**(22):3036-3046.
47. Anand R, Beach A, Li K, Haber J: **Rad51-mediated double-strand break repair and mismatch correction of divergent substrates**. *Nature* 2017, **544**(7650):377-380.
48. Thacker J: **The RAD51 gene family, genetic instability and cancer**. *Cancer Lett* 2005, **219**(2):125-135.
49. Doe CL, Osman F, Dixon J, Whitby MC: **DNA repair by a Rad22-Mus81-dependent pathway that is independent of Rhp51**. *Nucleic Acids Res* 2004, **32**(18):5570-5581.

50. Yamaguchi-Iwai Y, Sonoda E, Buerstedde JM, Bezzubova O, Morrison C, Takata M, Shinohara A, Takeda S: **Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52**. *Mol Cell Biol* 1998, **18**(11):6430-6435.
51. Shinohara A, Shinohara M, Ohta T, Matsuda S, Ogawa T: **Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing**. *Genes Cells* 1998, **3**(3):145-156.
52. Mott C, Symington LS: **RAD51-independent inverted-repeat recombination by a strand-annealing mechanism**. *DNA Repair (Amst)* 2011, **10**(4):408-415.
53. McDonald JP, Rothstein R: **Unrepaired heteroduplex DNA in Saccharomyces cerevisiae is decreased in RAD1 RAD52-independent recombination**. *Genetics* 1994, **137**(2):393-405.
54. Li PC, Petreaca RC, Jensen A, Yuan JP, Green MD, Forsburg SL: **Replication fork stability is essential for the maintenance of centromere integrity in the absence of heterochromatin**. *Cell Rep* 2013, **3**(3):638-645.
55. Gupta S, Yeeles JT, Marians KJ: **Regression of replication forks stalled by leading-strand template damage: II. Regression by RecA is inhibited by SSB**. *J Biol Chem* 2014, **289**(41):28388-28398.
56. Gadgil R, Barthelemy J, Lewis T, Leffak M: **Replication stalling and DNA microsatellite instability**. *Biophys Chem* 2017, **225**:38-48.
57. Hastings PJ, Ira G, Lupski JR: **A microhomology-mediated break-induced replication model for the origin of human copy number variation**. *PLoS Genet* 2009, **5**(1):e1000327.
58. Llorente B, Smith CE, Symington LS: **Break-induced replication - What is it and what is it for?** *Cell Cycle* 2008, **7**(7):859-864.
59. Tubbs A, Nussenzweig A: **Endogenous DNA Damage as a Source of Genomic Instability in Cancer**. *Cell* 2017, **168**(4):644-656.
60. Sansregret L, Vanhaesebroeck B, Swanton C: **Determinants and clinical implications of chromosomal instability in cancer**. *Nat Rev Clin Oncol* 2018, **15**(3):139-150.
61. Florl AR, Schulz WA: **Peculiar structure and location of 9p21 homozygous deletion breakpoints in human cancer cells**. *Genes Chromosomes Cancer* 2003, **37**(2):141-148.
62. DeGregori J: **Challenging the axiom: does the occurrence of oncogenic mutations truly limit cancer development with age?** *Oncogene* 2013, **32**(15):1869-1875.
63. Ma M, Rodriguez A, Sugimoto K: **Activation of ATR-related protein kinase upon DNA damage recognition**. *Current Genetics* 2019.
64. Caspari T, Carr AM: **Checkpoints: how to flag up double-strand breaks**. *Curr Biol* 2002, **12**(3):R105-107.
65. Aylon Y, Liefshitz B, Kupiec M: **The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle**. *EMBO J* 2004, **23**(24):4868-4875.
66. Barlow JH, Lisby M, Rothstein R: **Differential regulation of the cellular response to DNA double-strand breaks in G1**. *Mol Cell* 2008, **30**(1):73-85.
67. Caspari T, Murray JM, Carr AM: **Cdc2-cyclin B kinase activity links Crb2 and Rqh1-topoisomerase III**. *Genes Dev* 2002, **16**(10):1195-1208.
68. Du LL, Nakamura TM, Moser BA, Russell P: **Retention but not recruitment of Crb2 at double-strand breaks requires Rad1 and Rad3 complexes**. *Mol Cell Biol* 2003, **23**(17):6150-6158.
69. Lisby M, Barlow JH, Burgess RC, Rothstein R: **Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins**. *Cell* 2004, **118**(6):699-713.
70. Hamid A, Petreaca B, Petreaca R: **Frequent homozygous deletions of the CDKN2A locus in somatic cancer tissues**. *Mutat Res* 2019, **815**:30-40.
71. Bell JC: **Oncogenes**. *Cancer Lett* 1988, **40**(1):1-5.
72. Piazza A, Heyer WD: **Homologous Recombination and the Formation of Complex Genomic Rearrangements**. *Trends Cell Biol* 2019, **29**(2):135-149.

73. Cullen JK, Hussey SP, Walker C, Prudden J, Wee BY, Dave A, Findlay JS, Savory AP, Humphrey TC: **Break-induced loss of heterozygosity in fission yeast: dual roles for homologous recombination in promoting translocations and preventing de novo telomere addition**. *Mol Cell Biol* 2007, **27**(21):7745-7757.
74. Lucas BE, McPherson MT, Hawk TM, Wilson LN, Kroh JM, Hickman KG, Fitzgerald SR, Disbennett WM, Rollins PD, Hylton HM *et al*: **An Assay to Study Intra-Chromosomal Deletions in Yeast**. *Methods Protoc* 2019, **2**(3).
75. Watson AT, Werler P, Carr AM: **Regulation of gene expression at the fission yeast Schizosaccharomyces pombe urg1 locus**. *Gene* 2011, **484**(1-2):76-86.
76. Liddell L, Manthey G, Pannunzio N, Bailis A: **Quantitation and analysis of the formation of HO-endonuclease stimulated chromosomal translocations by single-strand annealing in Saccharomyces cerevisiae**. *J Vis Exp* 2011(55).
77. Fortunato EA, Osman F, Subramani S: **Analysis of spontaneous and double-strand break-induced recombination in rad mutants of S. pombe**. *Mutat Res* 1996, **364**(3):14-60.
78. Hylton HM, Lucas BE, Petreaca RC: **Schizosaccharomyces pombe Assays to Study Mitotic Recombination Outcomes**. *Genes (Basel)* 2020, **11**(1).