

Regulation of error-prone DNA double-strand break repair and its impact on genome evolution

**Terrence Hanscom and Mitch McVey
Tufts University**

Corresponding author:
Mitch McVey
Dept. of Biology
200 Boston Avenue, Suite 4700
Medford, MA 02421
mitch.mcvey@tufts.edu

Keywords

alt-EJ
polymerase theta
microhomology-mediated end joining
chromosome rearrangements
resection

Abstract

Double strand breaks are one of the most deleterious DNA lesions. Their repair via error-prone mechanisms can promote mutagenesis, loss of genetic information, and deregulation of the genome. These detrimental outcomes are significant drivers of human diseases, including many cancers. Mutagenic double-strand break repair also facilitates heritable genetic changes that drive organismal adaptation and evolution. In this review, we discuss the mechanisms of various error-prone DNA double strand break repair processes and the cellular conditions that regulate them, with a focus on alternative end joining. We provide examples that illustrate how mutagenic double strand break repair drives genome diversity and evolution. Finally, we discuss how error-prone break repair can be crucial to the induction and progression of diseases such as cancer.

1 **Introduction**

2 DNA double-strand breaks (DSBs) are highly dangerous lesions that arise with surprising
3 frequency. By one estimate, up to ten DSBs per cell per day occur in humans [1]. To deal with
4 these breaks, cells have evolved a variety of robust and conserved repair mechanisms. In many
5 cases, these can faithfully restore the genetic information at the break site. However, each repair
6 mechanism also brings with it a certain risk of mutagenesis, which varies depending on the specific
7 type of repair and the context in which it occurs.

8
9 We are beginning to understand some of the factors that dictate repair pathway choice and how
10 these may impact rates of mutagenesis. For example, certain DSB repair pathways are more active
11 during particular cell cycle phases. Classical or canonical non-homologous end joining (c-NHEJ),
12 which is typically accurate or results in small insertions and deletions, operates throughout
13 interphase, while high-fidelity homologous recombination (HR) is active mainly during the S and
14 G2 phases of the cell cycle [2]. Interestingly, a highly inaccurate form of DSB repair named
15 alternative end joining (alt-EJ) also occurs mainly during S/G2. This cell cycle preference can be
16 explained by the observation that both HR and alt-EJ require resection to form single-stranded
17 DNA, a process which is upregulated S/G2 [3]. However, the way that cells choose between HR
18 and alt-EJ is less clear. Importantly, alt-EJ can operate in cells that are competent for both c-NHEJ
19 and HR. Alt-EJ is associated with the generation of genomic diversity at sites prone to DSB
20 formation and can facilitate cellular and organismal adaptation [4]. In addition, it is a driver of
21 human disease, and is associated with poor cancer prognosis [5-7]. Thus, a better understanding
22 of the factors that promote the decision to use alt-EJ and other types of error-prone repair is needed.

23
24 In this review, we begin by cataloging the various types of error-prone DSB repair and examining
25 their similarities and differences. Our main focus is alternative end joining, as it has been most
26 closely tied to cellular transformation and genome evolution. We then delve into the various factors
27 that drive cells to prefer one type of repair over another. We conclude by providing selected
28 examples of how error-prone DNA repair promotes genetic diversity in different contexts and may
29 act to drive genome evolutionary processes at both the cellular and organismal levels.

Types of Inaccurate Double-Strand Break Repair

Double-strand break repair mechanisms can be roughly grouped into two main categories: those that involve synapsis of broken ends and subsequent ligation, including cNHEJ, alt-EJ, and single-strand annealing (SSA), and those that involve templated DNA synthesis using a homologous sequence, including synthesis-dependent strand annealing (SDSA) and break-induced replication (BIR). While many of these repair mechanisms have the potential to faithfully restore the original DNA sequence, all can also result in error-prone repair. Some, like alt-EJ, are always mutagenic. Error-prone DSB repair can result in a variety of inaccurate repair events, including deletions, insertions (both templated and non-templated), deletions coupled with insertions, and chromosome rearrangements. In the next sections, we describe each of these repair mechanisms, highlighting their mutagenic potential and discussing contexts which favor particular repair outcomes.

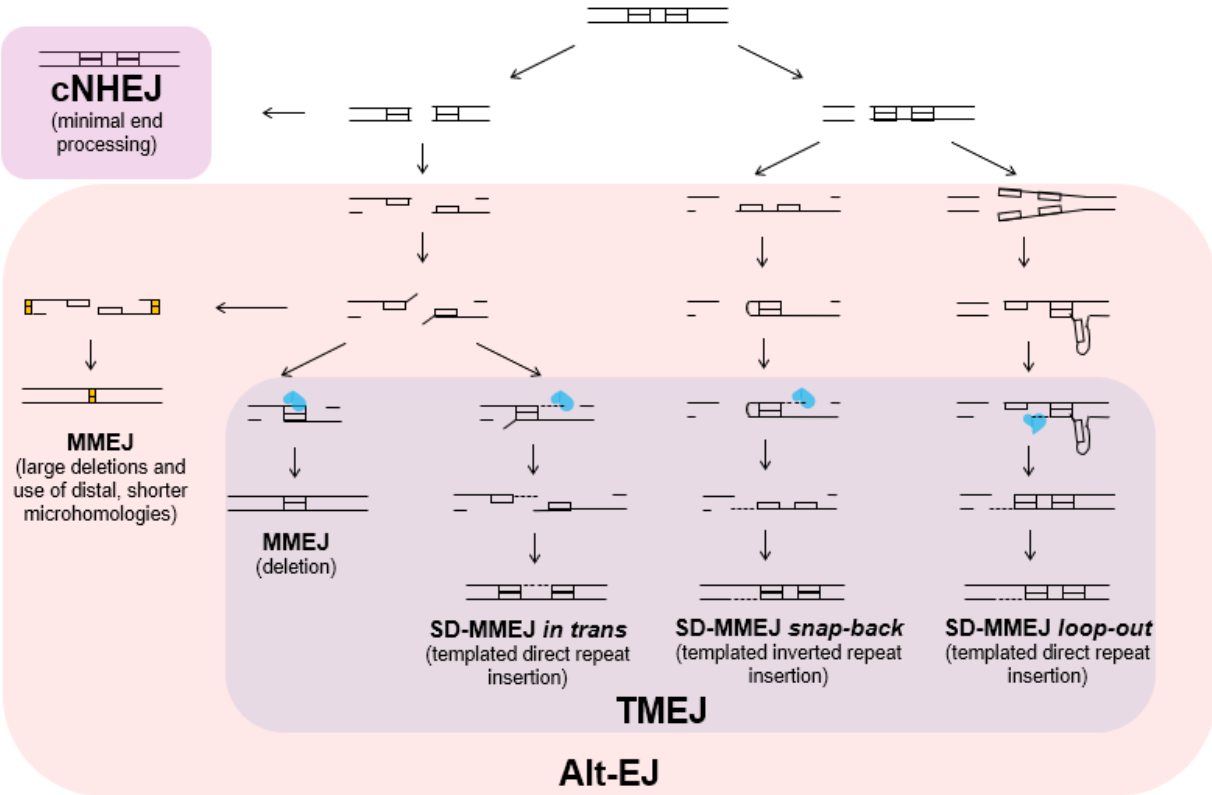


Figure 1. Categories of end-joining repair. The various forms of end joining are shown according to their mechanistic and genetic requirements. Alternative end joining (alt-EJ) is genetically distinct from canonical NHEJ (cNHEJ). Three models for synthesis-dependent microhomology-mediated end joining (SD-MMEJ) have been proposed, all of which can utilize DNA polymerase theta. In addition, simple MMEJ in metazoans (but not in yeast) requires polymerase theta. These types of MMEJ are grouped together as theta-mediated end joining (TMEJ). In the absence of polymerase theta, breaks can be repaired by a genetically undefined form of MMEJ that results in extremely large deletions.

Classical Nonhomologous End-Joining (cNHEJ): A First Responder

cNHEJ is a rapid repair pathway that is commonly defined by the protein components that promote it. It is active throughout interphase and is the preferred DSB repair pathway in vertebrates [1, 8, 9]. cNHEJ begins with binding of the Ku70/80 heterodimer to blunt DNA ends or those with less than 5 nucleotides (nt) of single stranded DNA. This binding acts as an initial barrier to resection and end processing. Upon binding, Ku70/80 recruits and activates the DNA-PKcs catalytic subunit which together form the DNA-dependent protein kinase (DNA-PK). For DSB ends that are directly ligatable, the XRCC4-XLF complex forms a sleeve-like structure around the duplex, stabilizing it for ligation via DNA ligase 4 [1, 8]. This outcome usually results in error-free repair. In cases where the ends cannot be directly joined, cNHEJ proceeds using a variety of end-processing factors (reviewed in [8]). Depending on the organism, these can include the Artemis nuclease, tyrosyl DNA phosphodiesterase (TDP), DNA polymerases μ and λ , and the Mre11-Rad50-Xrs2/Nbs1 (MRX/MRN) complex [1].

cNHEJ is typically accurate or minimally mutagenic, causing 1-5 base pair (bp) deletions [10], 1-4 bp microhomology joins, and small insertions [11] (Figure 1). Sometimes, these minor genomic changes can be advantageous, such as during class switch recombination and V(D)J recombination that promotes antibody diversity in the vertebrate immune system. In other contexts, cNHEJ clearly contributes to disease proliferation [8, 12, 13]. For example, in human cells cNHEJ is responsible for a majority of chromosomal translocations resulting from ionizing radiation and nuclease-induced breaks [14].

Alternative End- Joining: A Quick and Dirty Fix

Alt-EJ was originally used to describe any end joining that occurs in the absence of c-NHEJ, but we now know that it also happens in cNHEJ-competent cells. Often, alt-EJ is used as a general term to describe a variety of pathways that vary based on mechanistic and genetic differences but have significant overlap. Alt-EJ is slower than cNHEJ [1] and tends to be used less frequently. For example, in studies of mammalian cells alt-EJ occurred at 10% the frequency of cNHEJ [15] and at 10-20% the frequency of HR [16]. Alt-EJ is used to repair 0.5-1% of all DSBs in mammalian cells, but only 0.006% of all DSBs in yeast [8, 17]. The lower rate of alt-EJ in yeast may be due to

the absence of DNA polymerase theta, which has been shown to contribute to a majority of alt-EJ in organisms that have it (see below, and [10, 18-21]). While the original view of alt-EJ was that it acts as a fail-safe mechanism to fix persistent chromosomal breaks, the prevailing model now is that it is a highly regulated, independent repair pathway that exists as a viable option even when both cNHEJ and HR are available. [8, 16].

Alt-EJ can be defined by repair outcomes

One way that alt-EJ can be classified is according to the properties of the repair junctions that are created. The most common form of alt-EJ involves resection to create single-stranded DNA (ssDNA) at the break site and annealing of microhomologous sequences within the ssDNA. After annealing, non-homologous 3' tails are removed and repair is completed by fill-in synthesis and ligation. This process, termed microhomology-mediated end joining (MMEJ), results in deletions of varying sizes (Figure 1) [22].

In budding yeast, MMEJ-induced deletions range from a few base pairs to several kilobases of DNA [23, 24], while in mammalian cells MMEJ typically results in shorter deletions. In human osteosarcoma cells, MMEJ requires limited 5'-3' resection of less than 20 nt [16], while in mouse embryonic fibroblasts MMEJ can occur with longer, 70 nt ssDNA overhangs [10]. Typically, mammalian MMEJ uses short microhomologies of 1-8 bp for rejoining. While longer microhomologies may exist further from the break site, they are infrequently used. Notably, this tendency for minimal end processing and the use of 1-4 bp microhomologies can make it difficult to determine a priori whether a given repair event is generated through cNHEJ or MMEJ [25].

Alt-EJ can also produce insertions at the break site, with or without an accompanying deletion. In many cases, the insertions appear to be templated from DNA flanking the break. In order to explain the etiology of these types of junctions, the synthesis-dependent microhomology-mediated end-joining (SD-MMEJ) model was proposed [26]. SD-MMEJ is similar to MMEJ but has an additional intermediate DNA synthesis step to create de novo microhomologies that are then used in an MMEJ-like process. One key aspect of SD-MMEJ is the transient formation of secondary DNA structures (loops or hairpins) that occur via pairing of 1-4 bp direct or inverted repeats found near the break. These secondary structures serve as primers for limited DNA synthesis, likely by

polymerase theta (see below). The nascent DNA then dissociates, either spontaneously or via helicase activity, and the newly synthesized DNA anneals with microhomologous sequences on the other side of the break. A similar form of SD-MMEJ, in which the initial synthesis of microhomology occurs following transient base pairing across the break (trans SD-MMEJ), has also been described [27].

The signature of SD-MMEJ is the creation of a characteristic repeat motif, which distinguishes it from simple MMEJ (Figure 1). The repeat motif can be direct or inverted and includes both the secondary-structure forming repeat and the microhomology used for annealing. While SD-MMEJ is most easily identified by the presence of templated insertions, the process can also produce simple deletions and apparent blunt-end joins. Insertion events with apparently untemplated nucleotides, which are commonly seen in alt-EJ, may represent multiple templated synthesis events, or they could be created through SD-MMEJ using an error-prone polymerase [26, 27].

Alt-EJ can be defined by genetic requirements

Another way to classify alt-EJ is by the proteins that promote it. In budding yeast, MMEJ was originally characterized as a Ku-independent repair process that also occurs without the Rad52 protein, which distinguishes it from SSA [23]. Budding yeast MMEJ is promoted by Sae2 (CtIP) and Tel1 (ATM) and mutations in these genes confer an increase in blunt ended ligation via cNHEJ [24, 28]. In yeast, MMEJ also depends on Mre11, Rad50, and Rad1, and is modulated by Pol32, a non-essential subunit of polymerase delta, and CDC9, the ligase 1 catalytic subunit [23].

In vitro experiments using DT40 chicken B cells initially revealed roles for DNA ligases 1 and 3 and poly-ADP ribose polymerase (PARP) in an alternative end-joining mechanism that occurred in the absence of cNHEJ [29, 30]. Many of the proteins required for alt-EJ were found to be the same as those required for the primary steps of HR, suggesting that these two repair processes may share a common initial mechanism.

In 2010, a major role for DNA polymerase theta in alt-EJ was described in *Drosophila* [31]. Subsequently, polymerase theta was shown to be critical for alt-EJ repair of DSBs following replication fork collapse in *C. elegans* [20]. Since then, polymerase theta has been implicated in

alt-EJ in Arabidopsis and many vertebrates [18, 21, 32, 33], Because of the central role of polymerase theta in alt-EJ, many groups refer to alt-EJ as theta-mediated end joining (TMEJ) [10, 34, 35].

The Mechanism of TMEJ

Pol theta contains an A family polymerase domain at its C terminus which is essential for TMEJ. It facilitates end joining through a unique insertion loop in the polymerase domain that contacts the 3' terminal phosphate of the primer DNA and promotes its extension using a short, microhomologous template. Pol theta can use as little as 1-2 bp of microhomology to prime synthesis of 1-25 nt of nascent DNA, with an average length of 3-6 nt [36]. Evidence from studies using mouse embryonic fibroblasts with Cas9-generated DSBs suggest pol theta can sample all microhomologous sequences for annealing within the 3' terminal 25 nucleotides of a pair of ends [10]. Iterative rounds of the annealing, synthesis, and dissociation steps can lead to insertions templated from regions around the DSB [20, 32, 37].

Pol theta also contains an ATPase-dependent SF2 helicase domain at its N terminus. This domain has been shown to unwind short pieces of duplex DNA [38] and remove replication protein A (RPA) from ssDNA [39]. The helicase domain can also anneal complementary ssDNA ends [19], consistent with its role in promoting TMEJ. Further investigation is needed to determine if the helicase domain can also unwind newly synthesized microhomology intermediates in SD-MMEJ prior to the final annealing step.

Recently, the importance of the pol theta helicase domain was demonstrated *in vivo*. In *Drosophila*, pol theta ATPase-dead and null mutants are equally sensitive to interstrand crosslinking agents. In addition, ATPase-dead mutants produce significantly fewer insertions during alt-EJ and use shorter microhomologies at repair junctions [40]. In human cells, pol theta ATPase activity is required for survival in an HR deficient background [41]. Thus, the helicase domain is vitally important for TMEJ, a property that could be exploited in HR-deficient cancers [42].

Pol theta plays an important role in the repair of one-ended DSBs and breaks with damaged bases or protein adducts that cannot be readily ligated [43]. It promotes mutagenesis via deletion-prone

repair and by its high nucleotide misincorporation rate [21]. In the absence of pol theta, another type of cNHEJ-independent end joining mechanism still operates, producing very large deletions at sites of DSBs [4, 10, 20, 31] (Figure 1). In the *C. elegans* and *Drosophila* germlines, loss of polymerase theta is associated with deletions of multiple kilobases of DNA, suggesting that the mutational cost of TMEJ is small compared to the genomic catastrophe that results in its absence [4, 10, 34].

When homologous recombination ‘goes off the rails’

HR repairs double strand breaks by copying missing or damaged information from a homologous template, such as a sister chromatid, and typically ensures faithful reproduction of the original sequence. HR, like alt-EJ, requires an initial resection step that is carried out by the MRN complex and phosphorylated CtIP. Continued extensive resection, catalyzed by the BLM helicase and DNA2 endonuclease or the Exo1 exonuclease, differentiates HR from alt-EJ. The resulting 3' ssDNA ends are coated with RPA, which is then exchanged for the Rad51 recombinase. The DNA-Rad51 nucleoprotein filament initiates the search for a suitable template through the formation of a displacement loop (D-loop). The paired 3' single-stranded DNA is then extended via polymerase activity [16, 44]. In most instances, the D-loop dissociates after synthesis of several hundred nucleotides and the nascent strand pairs with single-stranded DNA on the opposite side of the break in a process called synthesis-dependent strand annealing (SDSA, Figure 2) [44]. Alternatively, the second end of the DSB can be captured prior to dissociation of the D-loop, resulting in a double Holliday junction structure that must be cleaved to finish repair. This mechanism is commonly referred to as the double strand break repair (DSBR) model [45].

Although HR is typically error-free, there are contexts where it can result in genome instability. For example, invasion of the RAD51-coated ssDNA into a similar but not identical sequence can result in non-allelic, or ectopic homologous recombination. Non-allelic HR (NAHR) can lead to copy number variants of short sequence motifs or entire genes (Figure 2). In a study investigating 17 human cancer genomes, NAHR was estimated to be responsible for 22% of all structural variations [12, 46]. Nearly 45% of the human genome is made up of repetitive elements that can contribute to ectopic recombination during DSB repair, leading to chromosomal rearrangements associated with human disease [47, 48]. For example, Alu repeats are ~300 bp

highly homologous sequences that represent ~11% of the overall DNA in the human genome. Strikingly, 0.3% of all new human genetic diseases are caused by ectopic recombination between Alu repeats [49, 50]. As one example, 41.5% of the intronic sequences of BRCA1 are made up of Alu repeats that range in size from 0.5 kb - 23.8 kb, and erroneous recombination between them is one of the most frequent causes of mutation in BRCA1 deficient cancers [49, 50].

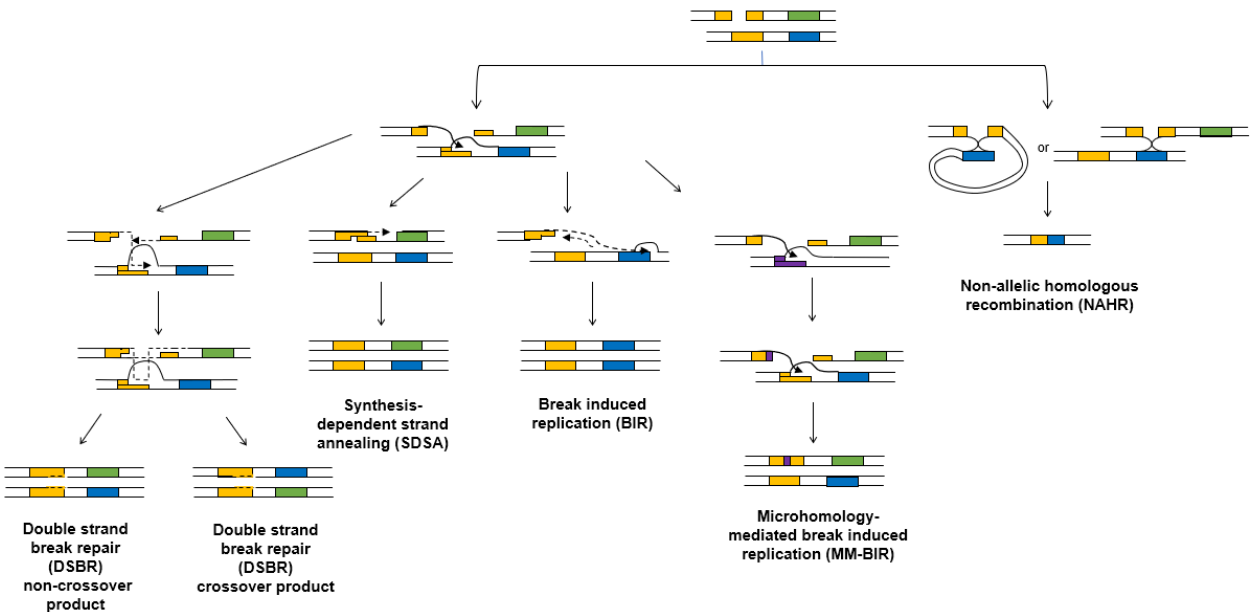


Figure 2. Mutagenic outcomes of homologous recombination. All types of homologous recombination begin with resection to form single-stranded DNA, followed by RAD51-mediated strand invasion. SDSA and DSBR usually use the sister chromatid as a template but use of the homologous chromosome can result in loss of heterozygosity (SDSA) or mitotic crossovers (DSBR). BIR involves the formation of a mobile D-loop and produces long, single-stranded DNA tracts. BIR can result in loss of heterozygosity and mutations caused by damage in persistent ssDNA. MM-BIR involves multiple rounds of strand invasion, which can cause insertions of heterologous sequences. NAHR occurs when single-stranded DNA invades into a non-allelic template. NAHR can be intrachromosomal or interchromosomal and results in deletions (shown), duplications, and inversions.

Break-induced replication

BIR is a highly mutagenic HR pathway that was first described in yeast at one-ended double strand breaks, but has since been shown to operate in human cells at persistently collapsed replication forks and during the repair of eroded telomeres in both yeast and human cells [51, 52]. BIR requires the resection proteins involved in SDSA and is Rad51- and Rad52-dependent in budding yeast but

is sometimes Rad51 independent in mammalian cells [51, 53]. DNA synthesis in BIR is conservative and is mainly carried out by DNA polymerase delta, with small contributions from polymerases epsilon and alpha [51]. BIR differs from SDSA in that BIR generates extremely long DNA synthesis tracts of up to 10kb in mammalian cells and >100kb in yeast [54] (Figure 2). While one-ended DSBs are a preferred substrate for BIR, limited homology of less than 150 base pairs at the non-invading end of a two-ended DSB causes inefficient second end capture and can also bias repair toward BIR [55, 56].

BIR can cause extensive mutagenicity in the form of loss of heterozygosity, translocations, dicentric chromosomes, and copy number variations. In addition, it is 1000-fold more mutagenic than replication in yeast due to the ssDNA that accumulates due to asynchronous synthesis of the leading and lagging strands. This ssDNA is highly susceptible to base damage and is a prime contributor to BIR mutagenicity [57]. Lastly, BIR at damaged replication forks in mammals has been reported to result in genomic duplications seen in many cancers [55].

At the extreme end of mutagenic HR is microhomology-mediated break induced replication (MM-BIR). During MM-BIR, ssDNA tails arising from the processing of broken replication forks temporarily anneal with microhomologous sequences at other locations in the genome (Figure 2). MM-BIR initiates with low processivity synthesis and frequently involves multiple template switches, generating complex rearrangements that are occasionally seen in mammalian cancers and other human diseases. These template switches can involve microhomologies of only 2-5bp [58]. This use of short microhomologies distinguishes the mechanism of MM-BIR from the initial strand invasion steps of traditional BIR [59]. Overall, the insights garnered from recent studies of MM-BIR demonstrate that rampant template switching can be a potent driver of genome instability and evolution.

Single-strand annealing: not just extreme MMEJ

SSA is a conserved DSB repair mechanism that was initially classified as a form of HR because both repair processes require Rad52 in yeast cells. Like HR, SSA requires resection and formation of 3' ssDNA ends. However, instead of RAD51-mediated strand invasion, SSA repair proceeds by annealing of homologous sequences, clipping of 3' single-stranded tails, and ligation [60]. Like

MMEJ, SSA results in deletions of the region between the homologous sequences that initially anneal to each other. However, SSA differs from MMEJ in that it uses longer microhomologies than MMEJ (>15 bp) and does not require DNA pol theta. Extensive resection may favor SSA over alt-EJ due to an increased probability of exposing longer microhomologies, especially in the context of vertebrate genomes with abundant transposons and long repeats [60].

Factors that affect repair pathway choice

Many factors are implicated in the DSB repair pathway decision, including the complexity and number of DSBs, stage of cell cycle, genome size, genome organization, and the presence of repetitive DNA. In addition, the availability of repair proteins and their post-translational modification status can influence what repair mechanisms are used. In the next sections, we highlight some of the most important factors for DSB repair pathway choice, with an emphasis on those that drive mutagenesis and genome diversity (Figure 3).

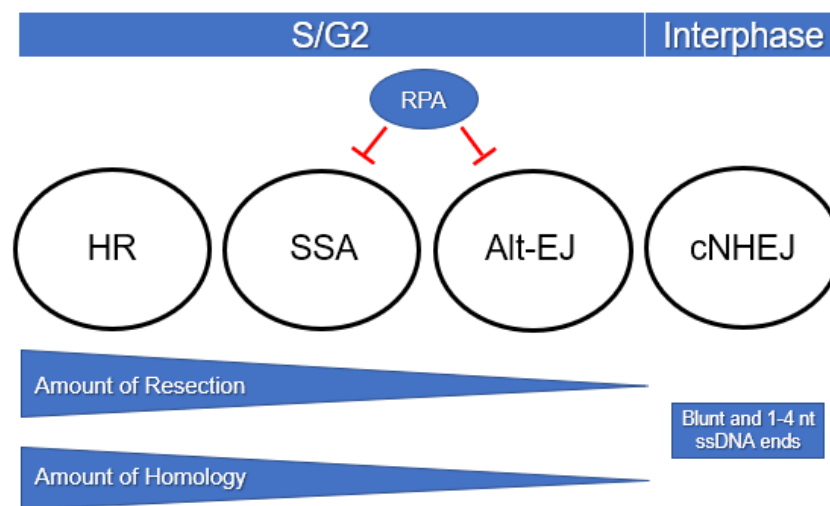


Figure 3. Factors that influence DSB repair pathway choice. cNHEJ operates throughout the cell cycle. HR, SSA, and alt-EJ require DNA resection and are therefore utilized primarily in S/G2. The required amounts of resection and homology are highest for HR and decrease for SSA and alt-EJ. RPA inhibits SSA and alt-EJ by preventing annealing of complementary sequences in single-stranded DNA.

Cell cycle position

In many organisms, the decision between HR and cNHEJ is tightly regulated by cell cycle dependent phosphorylation of DNA repair proteins [9]. In budding yeast, the cyclin dependent kinases CDK1 and CDK2 are upregulated during S and G2, when they downregulate cNHEJ by opposing Ku and DNA ligase 4 binding and phosphorylate Sae2 to promote resection [61]. In human cells CDKs also phosphorylate the Sae2 ortholog, CtIP, at multiple residues to promote resection [62].

While the preference for HR or cNHEJ varies widely between different organisms, the mechanisms that determine preference are similar. In both yeast and mammals Mre11 responds to DSBs throughout the cell cycle and is involved in both HR and cNHEJ [61]. In yeast during G1, the Ku heterodimer promotes cNHEJ and limits the endonuclease activity of Mre11 to inhibit resection. During G2, HR is the dominant DSB repair pathway and yeast will only perform cNHEJ if HR is perturbed [63]. In contrast, the dominant DSB repair pathway in human cells throughout interphase is cNHEJ, with HR and alt-EJ serving important roles during S and G2. For example, following exposure to ionizing radiation, approximately 80% of IR-induced DSBs in human fibroblasts are repaired by end joining. While most repair is likely carried out by cNHEJ, it is likely that some of the breaks are repaired by alt-EJ, as repair junctions with small insertions or deletions are characteristic of both pathways [61, 64, 65]. Interestingly, clusters of IR-induced DSBs require alt-EJ for their repair, suggesting that some radiation damage may inhibit repair by cNHEJ [65]. During G1, HR is repressed and cNHEJ performs an even greater proportion of DSB repair. The reason why mammals rely so heavily on cNHEJ over HR in comparison to yeast may have to do with the size and composition of the genome and the potential for deleterious rearrangements via HR in the repeat-rich mammalian genome.

Until recently it was thought that both HR and cNHEJ were inhibited during mitosis due to partial inactivation of the DNA damage response (DDR), which is a network of proteins that sense, signal, and repair DNA lesions [9]. However, recent studies have shown that DSB repair enzymes from both HR and cNHEJ respond to ionizing radiation-induced DSBs in metaphase and that the completion of repair is delayed until G1 [66]. Proteins such as the MRN complex, the complete cNHEJ complex, as well as early HR factors were found to be enriched at mitotic DSB sites while

other proteins involved in DSB repair were not recruited [67, 68]. Currently, it is thought that early repair factors like MRN may keep DNA ends tethered together until G1, when repair can be completed [66]. Tethering DSB ends together is important for proper chromosome segregation during cell division. Untethered fragments may become micronuclei that when reincorporated can induce chromothripsis, which is a devastating series of genomic rearrangements frequently associated with cancer (expanded upon below).

Resection

End resection involves a multitude of proteins and is highly regulated. In budding yeast, the MRX complex is thought to initially direct repair towards cNHEJ. Following activation of the endonuclease activity of Sae2 via CDK phosphorylation, the MRX-Sae2 complex cleaves the 5' strand internally. Similarly, in mammalian cells, CDK phosphorylation of threonine 847 in CtIP promotes resection, while phosphorylation of serine 327 promotes CtIP interaction with BRCA1, initiating the commitment to HR [64, 65]. These activities also remove barriers to resection including protein adducts such as stalled topoisomerases and the Spo11 meiotic recombination protein, along with DNA secondary structures and lesions [62, 65]. In yeast, resection is further potentiated by the recruitment of Tel1 by the MRX complex and Mec1 by RPA coated ssDNA. RPA-ssDNA promotes long-range resection by the Exo1 exonuclease and the Sgs1 helicase/Dna2 endonuclease.

In metazoans, additional regulators of resection and repair pathway choice exist. One important protein is BRCA1, which promotes resection during G2 and whose loss impairs Rad51-mediated recombination. 53BP1 promotes cNHEJ throughout interphase and is one of the main barriers to resection. At DSBs in G1, 53BP1 is phosphorylated by ATM, stimulating interaction with its effector, RIF1, to inhibit BRCA1 mediated resection [69]. During S/G2, BRCA1 blocks the interaction of RIF1 and 53BP1, allowing resection to occur. Mutations in BRCA1 result in resection defects that can be rescued by loss of 53BP1 [62]. Therefore, BRCA1 is not essential for resection but its role is rather to surmount the energetic barrier to resection by antagonizing 53BP1 [9, 62, 70]. Recently, 53BP1 and RIF1 were also shown to recruit the shieldin complex, composed of SHLD1, SHLD2, SHLD3, and REV7, to DSBs in G1 [71]. Thus, the combined action of 53BP1-

RIF1-shieldin serves to block resection in G1, ensuring that it occurs only when a sister chromatid template is present.

End resection is one of the main determinants of DSB repair pathway choice (Figure 3). The initiation of end resection is inhibitory for cNHEJ. In contrast, small amounts of resection expose microhomologies close to the break site and provide opportunities for alt-EJ. Continued resection that creates long stretches of single-stranded DNA is necessary for HR and SSA [72]. After long-range resection occurs, if a homologous template is not available for HR repair, then SSA becomes more likely.

Although resection occurs mainly in S and G2, limited resection has also been reported during G1 in human cells. This limited resection appears to be mediated by MRN and phosphorylated CtIP and is enhanced by the 53BP1 protein [73]. Rather than the CDKs, the Polo-like-kinase 3 is responsible for phosphorylating resection machinery in G1 [74]. This suppresses cNHEJ and promotes alt-EJ [73]. While alt-EJ that occurs in G1 is slower and more mutagenic than cNHEJ, it may be necessary for the repair of complex breaks that occur in G1 [75].

DNA packaging can facilitate or inhibit DSB repair by altering enzyme accessibility to damage. Therefore, a highly controlled chromatin response to a DSB is important to promote repair fidelity [76]. Often, chromatin remodeling enzymes are needed for the movement of nucleosomes to allow for efficient resection. One example is the budding yeast Fun30 chromatin remodeler that is an important mediator of both Sgs1/Dna2- and Exo1-dependent resection [77]. The human ortholog SMARCAD1 performs a similar function and its perturbation impairs end resection, HR, and sensitizes cells to PARP inhibitors [78].

RPA

Replication protein A is a heterotrimeric ssDNA binding protein involved in all ssDNA transactions in the nucleus. RPA has both low- and high-affinity binding modes and undergoes constant local dissociation and exchange via these modes [79]. It is required for replication, recombination, preventing spontaneous annealing of single-stranded DNA, and protecting ssDNA from degradation by nucleases. RPA prevents mutagenesis by reducing the formation of hairpins

and other secondary structures at repeated sequences, including trinucleotide repeats and palindromes. These structures have the potential to create dicentric chromosomes and tandem duplications that are often seen in cancer genomes [80].

RPA promotes the initiation of recombination-based repair pathways through several different mechanisms. In budding yeast, RPA bound to ssDNA enhances the endonuclease activity of MRX-Sae2 to promote HR in S/G2 [81]. RPA is also required for Sgs1/Dna2-mediated long-range resection. Following Sgs1 unwinding of DNA at a break, RPA binds to and protects the 3' strand and directs Dna2 to cleave the 5' strand. While RPA is not required for Exo1 recruitment to resected DSBs, the Exo1-mediated resection pathway is dysfunctional without RPA [80, 82].

In contrast, RPA inhibits MMEJ, which relies upon secondary structure formation and annealing of ssDNA. RPA likely also inhibits SD-MMEJ, although this has yet to be formally demonstrated. Overexpression of yeast Rfa1, the largest RPA subunit, suppresses MMEJ by inhibiting spontaneous annealing. In contrast, expression of a mutated form of Rfa1 that is defective in DNA binding confers a 350-fold increase in MMEJ [83]. These properties of RPA mutants are likely relevant to cancer cells where replication stress may lead to abundant ssDNA. Physiological levels of RPA may be insufficient in this context, leading to increased alt-EJ [84, 85].

Microhomology

During alt-EJ, the extent of microhomologous sequences flanking the break site and the microhomologies themselves can affect the efficiency of repair. One important consideration is the size of the microhomology. In budding yeast, the frequency of MMEJ is proportional to the length and melting temperature of the microhomologies, as the frequency of MMEJ increases 10-fold for every additional base pair of microhomology between 12-17 bp [86]. Similarly, in fission yeast even a single heterologous base at the 3' end reduces MMEJ efficiency two-fold [17, 87]. Microhomologies used in MMEJ in budding yeast are longer than in metazoans, requiring at least 8 complementary base pairs in one study [88]. In contrast, the microhomologies used in metazoan alt-EJ can be as short as 1 bp, likely due to the annealing action of polymerase theta.

A second factor relevant to alt-EJ is the proximity of the microhomologies to the break. In one study using mammalian cells, the efficiency of TMEJ was high when 2-4 bp microhomologies were directly adjacent to a DSB but dropped significantly when the distance from the break was increased to 10-30 bp. Increasing the length of microhomology to 6 bp rescued this distance-dependent decrease, illustrating that successful alt-EJ depends on a balance between microhomology size and distance from the break [37, 86].

The DNA context adjacent to the DSB also affects the proficiency of alt-EJ. In a study using mouse embryonic fibroblasts (MEFs), GC-rich flanking sequences produced more stable microhomologies and favored MMEJ while AT-rich flanking sequences were utilized 75% less often during MMEJ. This study also found that doubling an AT-rich microhomology from 3 bp to 6 bp increased the efficiency of MMEJ, highlighting a balance between homology length and AT/GC content. For alt-EJ outcomes that involved insertions, flanking sequences with 80% AT content produced more insertions than sequences with 80% GC content [37]. This bias may be due to the propensity of AT-rich microhomologies to dissociate following the initial MMEJ repair attempt, leading to insertion formation.

As previously described, a defining feature of SD-MMEJ is the formation of secondary structures (due to annealing of short repeats) that can be used for the initial DNA synthesis step, creating de novo microhomologies. Three parameters related to these repeats are important for SD-MMEJ outcomes: repeat length, distance from the repeat to the break, and distance between the repeats. As might be expected, the most frequently used repeats are those that are closest to the break site. As structure-forming repeats approach a distance of 30 bp from the break site, their utilization decreases to near zero [27]. Similarly, the overall efficiency of SD-MMEJ decreases as the distance between the repeats decreases. When the distance between a commonly used GGCC repeat increases from 5 to 15 base pairs, its usage during SD-MMEJ declines in a stepwise manner (our unpublished data).

Perhaps unexpected is the finding that increased length of structure-forming repeats does not correlate with a higher efficiency of SD-MMEJ. Studies in *Drosophila* using sequence variants flanking I-SceI-induced breaks have shown that the preferred length of these repeats is only 2-4

base pairs, while repeats of > 8 base pairs are utilized with low efficiency [27]. This suggests that the formation and dissociation of secondary structures is a highly dynamic (and perhaps regulated) process during SD-MMEJ.

Homology length also influences outcomes of HR repair pathways (Figure 3). Both SDSA and BIR are regulated by the amount of homology between the two DSB ends and the donor template. In yeast, if homology at the second end is $\leq 150\text{bp}$, second end capture is inefficient and repair shifts from gene conversion to BIR [56], but if significant homology is present SDSA outcompetes BIR. Shorter homologies between donor and recipient result in decreased viability and rates of repair [56]. Any heterology between the two sequences also decreases their recombination. For example, 3 nt of nonhomologous 3' DNA is enough to significantly reduce recombination and promote heteroduplex rejection and unwinding of mismatched DNA [89].

How error-prone repair relates to genome evolution

Evolution occurs via mutation of DNA and natural selection for mutations that increase reproductive fitness. Error-prone DNA repair can therefore serve as a driver of genome evolution. To ensure high-fidelity repair of DSBs during germ cell formation, break repair during meiosis is preferentially completed by HR. While cNHEJ is the major DSB repair pathway in somatic tissues in metazoans, it is suppressed during meiosis in *Drosophila* and *C. elegans* [90-92]. Still, mutational signatures common to different types of error-prone DSB repair have been found in the germline of organisms from roundworms to humans, suggesting these mechanisms can contribute to evolution.

The extent to which different organisms display a bias for one type of repair over another contributes to evolutionary diversification. As one example, a study investigating genome size change in *Arabidopsis thaliana* and barley determined the *Arabidopsis* genome has been shrinking over time due to frequent genome deletions, while the barley genome has generally remained constant in size, with a greater incidence of genome insertions. These trends could result from differential reliance on error-prone DSB repair pathways, different tolerances to genome loss and gain, and different genome structural features that may influence these changes [93]. To elucidate which of these factors are most important in various genome contexts, further investigation is

needed. Still, there are some published studies that provide some initial clues. We describe these in the next section and point out remaining gaps in our understanding.

Alt-EJ promotes Genome Evolution

There is substantial evidence that alt-EJ can act in the germline and thereby plays an important role in genome evolution. Whole genome sequencing studies of *C. elegans* passaged over multiple generations have demonstrated that genome changes with signatures of polymerase theta accumulate over time. These include deletions with 1 or more nucleotides of microhomology and small templated insertions [4]. These alt-EJ genome scars are also observed in wild *C. elegans*. Notably, >95% of error prone DSB repair following transposon mobilization or Cas9 breakage in the *C. elegans* germline is pol theta dependent and mutations in HR or cNHEJ do not affect the frequency or pattern of these alterations [4].

Several organisms, including larvacean species like *O. dioica*, have completely lost the ability to repair DSBs by cNHEJ. In these cases, a form of alt-EJ that produces deletions with 4 bp microhomologies and templated insertions appears to be a prominent repair option [94]. Unlike in *C. elegans*, it is less clear how the genome of *O. dioica* has been influenced by this reliance on more error-prone end joining. It has been suggested that a preference for alt-EJ may be transformative for genome organization, as more genome rearrangements are observed in the absence of the cNHEJ protein Ku [2, 94].

African trypanosomes, a type of parasitic flagellate protozoa, are also reported to lack cNHEJ machinery and rely on MMEJ for end-joining repair. A comprehensive study of DSB repair junctions in trypanosomes reveal microhomologies between 5-20 bp with a mean deletion size of 54 bp and a tolerance for mismatches within long stretches of microhomology [95]. Consistent with MMEJ observed in other organisms, proximity to the break, number and proportion of matched bases, and GC content all influence MMEJ in this organism. It is unknown whether trypanosomes have an enzyme similar to polymerase theta, but indications from other protozoan parasites suggest that it may and the DSB repair signatures are consistent with TMEJ [96]. Interestingly, a type of microhomology-mediated HR that leads to genomic rearrangements was also reported in trypanosomes [95].

Genome Compaction

The DSB repair processes that drive genome compaction, while incompletely understood, likely involve a variety of error-prone mechanisms, including cNHEJ, alt-EJ, and non-allelic homologous recombination. One process that is associated with a reduction in genome size is intron loss. A study of *Saccharomycetaceae* showed that intron loss almost always results in the complete loss of an intron without deletion of flanking coding sequence, consistent with HR replacement of the intron-containing genomic locus with a reverse transcribed copy of the corresponding mature mRNA [97, 98]. Less than 3% of intron loss events in these yeasts is associated with loss or gain of < 3 codons, suggesting that cNHEJ and MMEJ play a minor role in this process. In *Drosophila*, intron loss has also been associated with alt-EJ involving microhomology annealing within a gene [99].

Transposable elements are mobile sequences whose movement can restructure genomes. Many transposons contain long direct terminal repeats that can participate in SSA, resulting in the deletion of most of the transposon. Such events leave behind a solo LTR and have been well documented in *Arabidopsis thaliana* [93]. Similarly, the genome of *Oryza brachyantha*, a type of wild grass in the rice genus, has contracted greatly during evolution via a combination of SSA involving transposon LTRs and also through cNHEJ [100]. In mammals, long interspersed nuclear elements (LINEs) are a type of transposable element that make up approximately 17% of the genome. DSBs created in LINEs in murine cells are repaired via cNHEJ approximately 95% of the time, although non-allelic HR between homeologous LINE sequences occurs with appreciable frequency [101]. Comparison of the human and chimpanzee genomes suggests that both non-allelic HR and cNHEJ involving LINEs has resulted in deletions comprising about 450 kb during human evolution [102].

Genome Expansion

Genomes can also increase in size through several mechanisms, some of which are associated with inaccurate DSB repair. A study of the rice genome suggested that inaccurate DSB repair likely causes insertions between 10-1000 bp. These can involve tandem or partially tandem duplications, short duplications from ectopic donors, and non-tandem duplications with insertions

corresponding to DNA sequences located within 200 bp of a break, often from multiple locations. The investigators hypothesized that these types of duplications resulted from the formation of DSBs by synchronous, asymmetrical nicks during base excision repair, followed by end joining [103].

Genome expansion can also occur through the endosymbiotic gene transfer of organelle genes to the nuclear genome. Foreign genetic elements in the nucleus are most often degraded upon transfer but are occasionally integrated downstream of a strong nuclear promoter via error-prone double strand break repair mechanisms [104]. In *Arabidopsis*, up to 18% of nuclear genes appear to be derived from chloroplasts and their ancestor organelle [105]. An even greater percentage of yeast genes appear to have been transferred from mitochondria [106]. Similarly, many insertions in the *Drosophila* genome have occurred via end joining repair with mitochondrial DNA [99]. In humans, fragments of DNA that result from oxidative damage in the mitochondria can be inserted into the nuclear genome via cNHEJ. In addition, the presence of microhomology and templated additions up to 11 bp suggest that alt-EJ also mediates these insertion events [107, 108].

Most genome evolution occurs via a combination of both expansion and contraction events, the most well-known example of which can be found in *Saccharomyces cerevisiae*. Its genome is believed to have evolved via a whole genome duplication event followed by loss of nearly 90% of the duplicated genes via deletion-prone DSB repair [109]. The remaining duplicated genes evolved novel but often related functions as paralogs. A second example of rapid genome evolution was demonstrated using a highly novel genome restructuring method, in which yeast and *Arabidopsis* genomes were evolved following induction of multiple DSBs by a heat-inducible endonuclease. In this study, the evolved strains were found to harbor a number of rearrangements, copy number variations, and translocations involving retrotransposons that likely occurred through both cNHEJ and alt-EJ [110].

Error-prone repair contributes to genome evolution in cancer

Cancer cells experience severe replication stress that can lead to DSBs; this replication stress is exacerbated by exposure to chemotherapeutic agents that damage DNA. To deal with this stress, cancer cells often increase their reliance on error-prone DSB repair pathways, which contributes

to mutagenesis [111, 112]. Indeed, the mutational signatures observed in many different types of cancers often provide clues to the types of repair favored in each context [113]. Several groups have identified a ‘BRCAness’ signature that is associated with HR-deficient cancers, commonly referred to as Signature 3 in the Catalog of Somatic Mutations in Cancer (COSMIC). Signature 3 is characterized by an elevated number of base substitutions, deletions, and insertions >3 bp with microhomologies located at the suspected breakpoint junctions, characteristic of alt-EJ repair [114].

Mutations in the BRCA1 or BRCA2 genes result in a defect in HR repair, which cause cancer cells with these mutations to become ‘addicted’ to error-prone alt-EJ pathways for their survival. Indeed, perturbation of alt-EJ repair in HR-deficient backgrounds can result in synthetic lethality. One illustration of this is the sensitization of BRCA1- or BRCA2-deficient cancers to PARP inhibitors. Because PARP1 promotes alt-EJ, its impairment in an HR-defective background leaves cells with no repair options. Furthermore, many PARP inhibitors trap PARP1 at ssDNA breaks, which obstructs DNA replication and leads to DSBs that would normally be repaired by HR [5]. In addition to cancer cells lacking BRCA1/2, many others with HR deficiencies are reliant on alt-EJ for their growth, including those with mutations in RAD51, CHD1, PALB2, and FANCD2 (Table 1).

A second example of alt-EJ addiction is the synthetic lethality observed when polymerase theta is removed or inhibited in HR-deficient backgrounds [115]. Pol theta promotes DSB repair with junctions characteristic of those found in HR-deficient cells. Many HR deficient cancers overexpress pol theta, relying upon it for survival, and higher levels of pol theta correlate with increased cancer severity and poorer outcomes [7, 116]. For these reasons, pol theta shows great promise as a cancer therapeutic target [42].

Cancer/Disease	Affected Repair Pathway	Genetic Background	Drugable Targets	Effects on Genome	References
Breast, ovarian, melanoma, prostate, pancreatic	HR	BRCA1 deficient	POLQ, PARP1	Translocations, TDs, LOH, point mutations	[117]
Breast, ovarian	HR	BRCA2 deficient	POLQ, PARP1	Translocations, LOH, TDs, point mutations	[41]
Epithelial ovarian cancers	HR	FANCD2 deficient, Increased expression of POLQ	PARP1, FANCD2, POLQ	Chromosomal aberrations, nonsynonymous mutations	[118]
Breast, ovarian, Fanconi anemia	HR	PALB2 deficient	PARP1	Translocations, LOH, TDs, point mutations	[117]
Breast, stomach, prostate	HR	CHD1 deficient, Deficient in CtIP recruitment to DSBs	PARP1, PTEN	Translocations, LOH, TDs, point mutations	[119]
Breast, ovarian, Fanconi anemia	HR	RAD51C deficient	PARP1	Genomic instability, aneuploidy, chromosome aberrations	[120]
Breast, ovarian, Fanconi anemia	HR	RAD51D deficient	PARP1	Large deletions, genomic instability	[121]
Chronic myeloid leukemia	HR	BCR-ABL (constitutively active tyrosine kinase), increased expression of LIG3a, PARP1, and WRN	PARP1 combined with ligase inhibitors	Translocations, LOH, TDs, point mutations	[122, 123]
Lynch Syndrome, colorectal, endometrial, ovarian, gastric, urinary, small bowel, pancreatic, prostate	MMR, HR	MLH1, MSH2, MSH6, PMS2 deficient	PARP1	Point mutations, microsatellite instability	[124]
High grade bladder tumors, colon	cNHEJ	KU, DNA-PK, or XRCC4 deficient	Unknown	Deletions (<125bp on average) and microhomology at repair junctions	[125]
Multiple myeloma, leukemia, pro B-cell lymphoma	cNHEJ	KU and P53 deficient	Unknown	Rearrangements and gene amplifications, nonreciprocal translocations, increased microhomologies at repair junctions	[126-128]
MCF7 breast cancer	cNHEJ	Reduced levels of DNA LIG4 Increased levels of DNA LIG3a and PARP1	PARP1 combined with ligase inhibitors	Large deletions, translocations	[129]

Acute myeloid leukemia	cNHEJ	Decreased expression of KU proteins Increased expression of DNA LIG3a	Unknown	TDs, microhomology at repair junctions, deletions	[130]
Neuroblastoma	cNHEJ	LIG4 and Artemis deficient Increased expression of DNA LIG3, LIG1, and PARP1	PARP1	Translocations, TDs	[131]
Non-BRCA1/2 breast cancer	cNHEJ	XRCC4 deficient	Unknown	Translocations	[132]

Table 1. Cancers reliant upon error-prone DSB repair. For each type of cancer/disease, the repair pathway that is impaired, genetic background, and the types of genomic mutations observed are listed. The druggable targets column indicates therapeutic approaches that may be used to exploit a synthetic lethal interaction. TD = tandem duplications, LOH = loss of heterozygosity

Genome rearrangements

Genome rearrangements, including chromosome translocations, copy number variations, and telomere fusions, are frequently created through aberrant DSB repair. The repair pathways involved in their genesis differ widely between organisms. For example, translocations in budding yeast typically involve 2-20 bp microhomologies and are elevated in cNHEJ-deficient cells [86]. Mutations in the cNHEJ antagonist Sae2 also decrease translocations [133]. These findings suggest that in yeast, cNHEJ protects against translocations and a polymerase theta-independent form of alt-EJ mediates translocations.

In metazoans such as *C. elegans* that possess pol theta, TMEJ has been shown to generate most chromosome rearrangements in the germline [4]. Similarly, alt-EJ has a dominant role in translocation formation in mice, to a far greater degree than cNHEJ [86, 111, 134, 135]. Loss of Ku or DNA ligase 4 causes a 3-fold increase in translocation frequency in MEFs [10]. Similarly, MEFs deficient in polymerase theta have reduced translocations in non-telomeric sequences and the remaining translocations display less use of microhomology [18, 39]. In contrast, translocations in human cells are largely caused by cNHEJ. Translocation frequency decreases in DNA ligase 4 mutants [14, 136]. While 31% of balanced translocations and inversions in human cells use

microhomology, alt-EJ seem to be less important for their formation [12]. The reasons behind these differences in mice and humans are not clear.

Genome rearrangements are prevalent in cancer genomes. Distributions of rearrangements are highly variable and depending on the type of cancer, can be relatively even or clustered in hot spots within the genome. Common rearrangements seen in many cancers involve tandem duplications (TDs) between 3 kb and 1 Mb. These TDs often have microhomology at their ends, suggesting that they may result from alt-EJ repair [137]. These rearrangements cause structural changes in DNA that alter proteins, drive aberrant transcriptional patterns, and modify the chromatin environment. While many rearrangements are detrimental to cells, some can confer a fitness advantage, driving cancer progression [137].

Translocations can also occur at degraded telomere ends. Over time, shortened telomeres can cause genome instability via end-to-end fusions and large-scale genomic rearrangements, predisposing to cancer. In human cells, telomere fusions can be mediated through both DNA ligase 3 (alt-EJ) and DNA ligase 4 (cNHEJ), but cells that escape rearrangement crisis preferentially utilize cNHEJ [138].

Chromothripsis

Chromothripsis is a highly mutagenic process whereby the genome acquires dozens to thousands of simultaneous rearrangements. It was first observed in a chronic lymphocytic leukemia that contained 42 genomic rearrangements, all within the long arm of chromosome 4 [139]. Cancer cells frequently contain massive rearrangements that are consistent with a chromothripsis mechanism; several studies have shown that the genomes of more than 50% of tumors from selected human cancers show evidence of chromothripsis [140-142]. Similar to translocations seen in human cells, chromothripsis repair events are often mediated by cNHEJ and to a lesser degree by alt-EJ and MM-BIR [140-142].

Repair pathway choice in the context of genome editing

The use of CRISPR-Cas systems has revolutionized molecular biology and provided a diverse tool kit for genome editing, therapeutics, and diagnostics. To fully realize the potential of this

technology, a complete understanding of how Cas9-induced DSBs are repaired is needed. Specifically, the ability to predict CRISPR outcomes, including the relative balance of cNHEJ and alt-EJ, in genome editing experiments would be ideal. A study performed in mammalian cells with DSBs induced by SpCas9 and more than 40,000 different single guide RNAs (sgRNAs) illustrated that editing outcomes mostly depend on the targeted sequence itself. Strikingly, 11% of all sgRNAs tested had a dominant outcome that composed >40% of all repair events for that sgRNA [143]. Analysis of repair outcomes showed that 58% of events involved a deletion of at least 3bp and half of those used microhomology, indicative of alt-EJ repair (Figure 4). Three percent of all repair events were insertions of ≥ 1 bp and 99% of these were an insertion of the PAM distal nucleotide. These events were likely due to fill-in of Cas9-induced DSBs with single nucleotide 5' overhangs. In another study of Cas9-induced breaks, 5-10% were repaired by TMEJ under normal physiological conditions [37]. Thus, it appears clear that the ability to predict alt-EJ repair products will be critical to reach the goal of understanding potential outcomes of Cas9 genome editing experiments.

The type of DNA ends can also impact the outcome of genome editing experiments (Figure 4). In mammalian cells, the use of Cas9 nickase variants can produce DSBs with either 3' or 5' overhangs. In one study, repair of DSBs with 3' overhangs frequently produced insertions apparently templated from sequences adjacent to the break. These insertions were longer than those formed at 5' ends, occurred more frequently, and use microhomologies indicative of TMEJ [144]. In contrast, DSBs with 5' overhangs promoted HR more efficiently than 3' overhangs and typically led to larger deletions. These deletions likely arose from processing of the 5' overhang into a 3' overhang prior to repair by HR or alt-EJ [144].

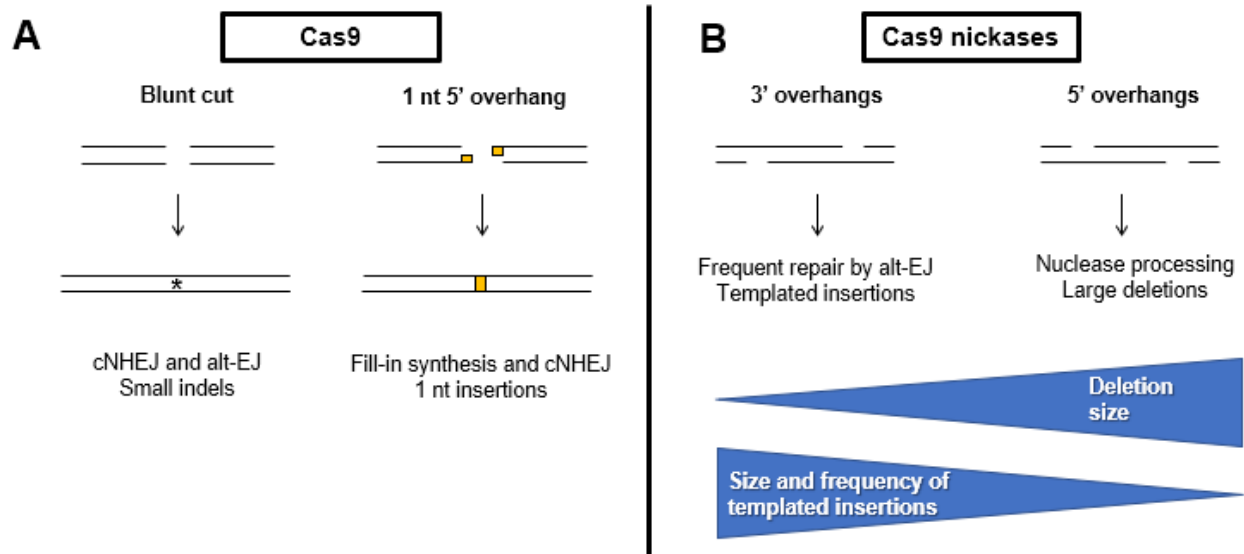


Figure 4. Mutagenic outcomes of Cas9-mediated genome editing. (A) Cas9-mediated double-strand breaks are usually blunt-ended and can be repaired inaccurately by cNHEJ and alt-EJ. Occasionally, Cas9 leaves 5' single-stranded overhangs that are filled in, resulting in a single nucleotide insertion. (B) 3' overhangs produced by Cas9 nickases are often repaired by alt-EJ, creating large, templated insertions. In contrast, 5' overhangs are processed by nucleases to produce substrates suitable for alt-EJ and HR, sometimes resulting in large deletions.

Conclusions and Remaining Questions

Genomes are constantly assaulted by agents that cause DNA double-strand breaks. To deal with this, organisms rely on both end-joining and recombination-based repair pathways. These are highly regulated and usually result in accurate restoration of the genome. However, it is becoming clear that in spite of its mutagenic potential, error-prone DSB repair occurs with surprising frequency. Importantly, error-prone DSB repair is not simply a backup mechanism utilized when accurate repair is not possible. Multiple factors influence its usage, including enzymatic processes related to cell cycle stage, genome structure, and local sequence context. Error-prone repair is also purposeful—in its absence, more serious genome catastrophe can occur.

A survey of the current literature demonstrates that error-prone break repair serves as a driving force of genome evolution, both in the germline and in somatic tissues in the context of cancer. It is increasingly apparent that a detailed understanding of how alt-EJ and other mutagenic repair

mechanisms are regulated will have important therapeutic implications, in the contexts of disease treatment, aging, and genome editing.

In order to further this understanding, there are many questions that remain to be answered. For example, how does local sequence variation and chromatin context affect alt-EJ repair, particularly TMEJ? What is the mechanism behind the rare but detrimental deletion-prone end joining that is observed in the absence of both cNHEJ and TMEJ? With what frequency does error-prone repair operate during mitosis and what are the consequences of this type of repair for genome stability? Why do different end-joining mechanisms preferentially mediate translocation formation in different organisms and how does this preference impact large scale genome evolution events like chromothripsis? Continued investigation of these questions and others will certainly provide additional surprises into the ways that error-prone repair is leveraged to drive genome evolution.

Acknowledgements

The authors wish to thank Justin Blanch and members of the McVey lab for discussion and feedback on the manuscript. We apologize to investigators whose work could not be discussed or cited in this review due to space constraints. Research in the McVey lab is supported by grants MCB1716039 from the National Science Foundation and P01GM10547 and R01GM125827 from the National Institutes of Health.

Author Contributions

Writing – Original Draft Preparation, T.H. and M.M.; Writing – Review & Editing, T.H. and M.M.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair*. Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.
2. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway*. Annu Rev Biochem, 2010. **79**: p. 181-211.
3. Huertas, P., *DNA resection in eukaryotes: deciding how to fix the break*. Nat Struct Mol Biol, 2010. **17**(1): p. 11-6.
4. van Schendel, R., et al., *Polymerase Theta is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis*. Nat Commun, 2015. **6**: p. 7394.
5. Helleday, T., *The underlying mechanism for the PARP and BRCA synthetic lethality: clearing up the misunderstandings*. Mol Oncol, 2011. **5**(4): p. 387-93.
6. Leeman, J.E., et al., *Human papillomavirus 16 promotes microhomology-mediated end-joining*. Proc Natl Acad Sci U S A, 2019. **116**(43): p. 21573-21579.
7. Higgins, G.S., et al., *Overexpression of POLQ confers a poor prognosis in early breast cancer patients*. Oncotarget, 2010. **1**(3): p. 175-84.
8. Deriano, L. and D.B. Roth, *Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage*. Annu Rev Genet, 2013. **47**: p. 433-55.
9. Krenning, L., J. van den Berg, and R.H. Medema, *Life or Death after a Break: What Determines the Choice?* Mol Cell, 2019. **76**(2): p. 346-358.
10. Wyatt, D.W., et al., *Essential Roles for Polymerase theta-Mediated End Joining in the Repair of Chromosome Breaks*. Mol Cell, 2016. **63**(4): p. 662-673.
11. Lieber, M.R., et al., *Nonhomologous DNA end joining (NHEJ) and chromosomal translocations in humans*. Subcell Biochem, 2010. **50**: p. 279-96.
12. Ottaviani, D., M. LeCain, and D. Sheer, *The role of microhomology in genomic structural variation*. Trends Genet, 2014. **30**(3): p. 85-94.
13. Sishc, B.J. and A.J. Davis, *The Role of the Core Non-Homologous End Joining Factors in Carcinogenesis and Cancer*. Cancers (Basel), 2017. **9**(7).
14. Ghezraoui, H., et al., *Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining*. Mol Cell, 2014. **55**(6): p. 829-842.
15. Corneo, B., et al., *Rag mutations reveal robust alternative end joining*. Nature, 2007. **449**(7161): p. 483-6.
16. Truong, L.N., et al., *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, 2013. **110**(19): p. 7720-5.
17. Sfeir, A. and L.S. Symington, *Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway?* Trends Biochem Sci, 2015. **40**(11): p. 701-714.
18. Mateos-Gomez, P.A., et al., *Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination*. Nature, 2015. **518**(7538): p. 254-7.
19. Kent, T., et al., *Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta*. Nat Struct Mol Biol, 2015. **22**(3): p. 230-7.
20. Roerink, S.F., R. van Schendel, and M. Tijsterman, *Polymerase theta-mediated end joining of replication-associated DNA breaks in C. elegans*. Genome Res, 2014. **24**(6): p. 954-62.

21. Yousefzadeh, M.J. and R.D. Wood, *DNA polymerase POLQ and cellular defense against DNA damage*. DNA Repair (Amst), 2013. **12**(1): p. 1-9.
22. McVey, M. and S.E. Lee, *MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings*. Trends Genet, 2008. **24**(11): p. 529-38.
23. Ma, J.L., et al., *Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences*. Mol Cell Biol, 2003. **23**(23): p. 8820-8.
24. Lee, K. and S.E. Lee, *Saccharomyces cerevisiae Sae2- and Tel1-dependent single-strand DNA formation at DNA break promotes microhomology-mediated end joining*. Genetics, 2007. **176**(4): p. 2003-14.
25. Pannunzio, N.R., et al., *Non-homologous end joining often uses microhomology: implications for alternative end joining*. DNA Repair (Amst), 2014. **17**: p. 74-80.
26. Yu, A.M. and M. McVey, *Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions*. Nucleic Acids Res, 2010. **38**(17): p. 5706-17.
27. Khodaverdian, V.Y., et al., *Secondary structure forming sequences drive SD-MMEJ repair of DNA double-strand breaks*. Nucleic Acids Res, 2017. **45**(22): p. 12848-12861.
28. Tsabar, M. and J.E. Haber, *Chromatin modifications and chromatin remodeling during DNA repair in budding yeast*. Curr Opin Genet Dev, 2013. **23**(2): p. 166-73.
29. Iliakis, G., et al., *DNA damage checkpoint control in cells exposed to ionizing radiation*. Oncogene, 2003. **22**(37): p. 5834-47.
30. Paul, K., et al., *DNA ligases I and III cooperate in alternative non-homologous end-joining in vertebrates*. PLoS One, 2013. **8**(3): p. e59505.
31. Chan, S.H., A.M. Yu, and M. McVey, *Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila*. PLoS Genet, 2010. **6**(7): p. e1001005.
32. Zahn, K.E., et al., *Human DNA polymerase theta grasps the primer terminus to mediate DNA repair*. Nat Struct Mol Biol, 2015. **22**(4): p. 304-11.
33. Inagaki, S., et al., *Arabidopsis TEBICHI, with helicase and DNA polymerase domains, is required for regulated cell division and differentiation in meristems*. Plant Cell, 2006. **18**(4): p. 879-92.
34. Schimmel, J., et al., *Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells*. EMBO J, 2017. **36**(24): p. 3634-3649.
35. Zelensky, A.N., et al., *Inactivation of Pol theta and C-NHEJ eliminates off-target integration of exogenous DNA*. Nat Commun, 2017. **8**(1): p. 66.
36. He, P. and W. Yang, *Template and primer requirements for DNA Pol theta-mediated end joining*. Proc Natl Acad Sci U S A, 2018. **115**(30): p. 7747-7752.
37. Carvajal-Garcia, J., et al., *Mechanistic basis for microhomology identification and genome scarring by polymerase theta*. Proc Natl Acad Sci U S A, 2020. **117**(15): p. 8476-8485.
38. Ozdemir, A.Y., et al., *Polymerase theta-helicase efficiently unwinds DNA and RNA-DNA hybrids*. J Biol Chem, 2018. **293**(14): p. 5259-5269.
39. Mateos-Gomez, P.A., et al., *The helicase domain of Poltheta counteracts RPA to promote alt-NHEJ*. Nat Struct Mol Biol, 2017. **24**(12): p. 1116-1123.

40. Beagan, K., et al., *Drosophila DNA polymerase theta utilizes both helicase-like and polymerase domains during microhomology-mediated end joining and interstrand crosslink repair*. PLoS Genet, 2017. **13**(5): p. e1006813.
41. Ceccaldi, R., et al., *Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair*. Nature, 2015. **518**(7538): p. 258-62.
42. Zhou, J., et al., *Polymerase Theta Inhibition Kills Homologous Recombination Deficient Tumors*. bioRxiv, 2020: p. 2020.05.23.111658.
43. Howard, S.M., D.A. Yanez, and J.M. Stark, *DNA damage response factors from diverse pathways, including DNA crosslink repair, mediate alternative end joining*. PLoS Genet, 2015. **11**(1): p. e1004943.
44. Wright, W.D., S.S. Shah, and W.D. Heyer, *Homologous recombination and the repair of DNA double-strand breaks*. J Biol Chem, 2018. **293**(27): p. 10524-10535.
45. Rodgers, K. and M. McVey, *Error-Prone Repair of DNA Double-Strand Breaks*. J Cell Physiol, 2016. **231**(1): p. 15-24.
46. Kidd, J.M., et al., *A human genome structural variation sequencing resource reveals insights into mutational mechanisms*. Cell, 2010. **143**(5): p. 837-47.
47. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
48. Zepeda-Mendoza, C.J., et al., *Identical repeated backbone of the human genome*. BMC Genomics, 2010. **11**: p. 60.
49. White, T.B., M.E. Morales, and P.L. Deininger, *Alu elements and DNA double-strand break repair*. Mob Genet Elements, 2015. **5**(6): p. 81-85.
50. Montagna, M., et al., *Identification of a 3 kb Alu-mediated BRCA1 gene rearrangement in two breast/ovarian cancer families*. Oncogene, 1999. **18**(28): p. 4160-5.
51. Kramara, J., B. Osia, and A. Malkova, *Break-Induced Replication: The Where, The Why, and The How*. Trends Genet, 2018. **34**(7): p. 518-531.
52. Malkova, A., E.L. Ivanov, and J.E. Haber, *Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication*. Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7131-6.
53. Signon, L., et al., *Genetic requirements for RAD51- and RAD54-independent break-induced replication repair of a chromosomal double-strand break*. Mol Cell Biol, 2001. **21**(6): p. 2048-56.
54. Saini, N., et al., *Migrating bubble during break-induced replication drives conservative DNA synthesis*. Nature, 2013. **502**(7471): p. 389-92.
55. Costantino, L., et al., *Break-induced replication repair of damaged forks induces genomic duplications in human cells*. Science, 2014. **343**(6166): p. 88-91.
56. Mehta, A., A. Beach, and J.E. Haber, *Homology Requirements and Competition between Gene Conversion and Break-Induced Replication during Double-Strand Break Repair*. Mol Cell, 2017. **65**(3): p. 515-526 e3.
57. Chan, K., et al., *Base damage within single-strand DNA underlies in vivo hypermutability induced by a ubiquitous environmental agent*. PLoS Genet, 2012. **8**(12): p. e1003149.
58. Zhang, F., et al., *The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans*. Nat Genet, 2009. **41**(7): p. 849-53.

- 870 59. Sakofsky, C.J. and A. Malkova, *Break induced replication in eukaryotes: mechanisms,*
871 *functions, and consequences.* Crit Rev Biochem Mol Biol, 2017. **52**(4): p. 395-413.
- 872 60. Bhargava, R., D.O. Onyango, and J.M. Stark, *Regulation of Single-Strand Annealing and its*
873 *Role in Genome Maintenance.* Trends Genet, 2016. **32**(9): p. 566-575.
- 874 61. Zhang, Y., et al., *Regulation of repair choice: Cdk1 suppresses recruitment of end joining*
875 *factors at DNA breaks.* DNA Repair (Amst), 2009. **8**(10): p. 1235-41.
- 876 62. Escribano-Diaz, C., et al., *A cell cycle-dependent regulatory circuit composed of 53BP1-*
877 *RIF1 and BRCA1-CtIP controls DNA repair pathway choice.* Mol Cell, 2013. **49**(5): p. 872-
878 83.
- 879 63. Emerson, C.H. and A.A. Bertuch, *Consider the workhorse: Nonhomologous end-joining in*
880 *budding yeast.* Biochem Cell Biol, 2016. **94**(5): p. 396-406.
- 881 64. Her, J. and S.F. Bunting, *How cells ensure correct repair of DNA double-strand breaks.* J
882 Biol Chem, 2018. **293**(27): p. 10502-10511.
- 883 65. Iliakis, G., E. Mladenov, and V. Mladenova, *Necessities in the Processing of DNA Double*
884 *Strand Breaks and Their Effects on Genomic Instability and Cancer.* Cancers (Basel), 2019.
885 **11**(11).
- 886 66. Blackford, A.N. and M. Stucki, *How Cells Respond to DNA Breaks in Mitosis.* Trends
887 Biochem Sci, 2020. **45**(4): p. 321-331.
- 888 67. Giunta, S., R. Belotserkovskaya, and S.P. Jackson, *DNA damage signaling in response to*
889 *double-strand breaks during mitosis.* J Cell Biol, 2010. **190**(2): p. 197-207.
- 890 68. Ayoub, N., et al., *The carboxyl terminus of Brca2 links the disassembly of Rad51 complexes*
891 *to mitotic entry.* Curr Biol, 2009. **19**(13): p. 1075-85.
- 892 69. Chapman, J.R., et al., *RIF1 is essential for 53BP1-dependent nonhomologous end joining*
893 *and suppression of DNA double-strand break resection.* Mol Cell, 2013. **49**(5): p. 858-71.
- 894 70. Bunting, S.F., et al., *53BP1 inhibits homologous recombination in Brca1-deficient cells by*
895 *blocking resection of DNA breaks.* Cell, 2010. **141**(2): p. 243-54.
- 896 71. Noordermeer, S.M., et al., *The shieldin complex mediates 53BP1-dependent DNA repair.*
897 Nature, 2018. **560**(7716): p. 117-121.
- 898 72. Daley, J.M., et al., *Biochemical mechanism of DSB end resection and its regulation.* DNA
899 Repair (Amst), 2015. **32**: p. 66-74.
- 900 73. Xiong, X., et al., *53BP1 promotes microhomology-mediated end-joining in G1-phase cells.*
901 Nucleic Acids Res, 2015. **43**(3): p. 1659-70.
- 902 74. Barton, O., et al., *Polo-like kinase 3 regulates CtIP during DNA double-strand break repair*
903 *in G1.* J Cell Biol, 2014. **206**(7): p. 877-94.
- 904 75. Reginato, G., E. Cannavo, and P. Cejka, *Physiological protein blocks direct the Mre11-*
905 *Rad50-Xrs2 and Sae2 nuclease complex to initiate DNA end resection.* Genes Dev, 2017.
906 **31**(23-24): p. 2325-2330.
- 907 76. Du, Q., et al., *Replication timing and epigenome remodelling are associated with the*
908 *nature of chromosomal rearrangements in cancer.* Nat Commun, 2019. **10**(1): p. 416.
- 909 77. Eapen, V.V., et al., *The Saccharomyces cerevisiae chromatin remodeler Fun30 regulates*
910 *DNA end resection and checkpoint deactivation.* Mol Cell Biol, 2012. **32**(22): p. 4727-40.
- 911 78. Costelloe, T., et al., *The yeast Fun30 and human SMARCAD1 chromatin remodellers*
912 *promote DNA end resection.* Nature, 2012. **489**(7417): p. 581-4.

79. Yates, L.A., et al., *A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA*. Nat Commun, 2018. **9**(1): p. 5447.
80. Chen, H., M. Lisby, and L.S. Symington, *RPA coordinates DNA end resection and prevents formation of DNA hairpins*. Mol Cell, 2013. **50**(4): p. 589-600.
81. Symington, L.S., *Mechanism and regulation of DNA end resection in eukaryotes*. Crit Rev Biochem Mol Biol, 2016. **51**(3): p. 195-212.
82. Cejka, P., et al., *DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2*. Nature, 2010. **467**(7311): p. 112-6.
83. Deng, S.K., et al., *RPA antagonizes microhomology-mediated repair of DNA double-strand breaks*. Nat Struct Mol Biol, 2014. **21**(4): p. 405-12.
84. Lecona, E. and O. Fernandez-Capetillo, *Replication stress and cancer: it takes two to tango*. Exp Cell Res, 2014. **329**(1): p. 26-34.
85. Toledo, L.I., et al., *ATR prohibits replication catastrophe by preventing global exhaustion of RPA*. Cell, 2013. **155**(5): p. 1088-103.
86. Villarreal, D.D., et al., *Microhomology directs diverse DNA break repair pathways and chromosomal translocations*. PLoS Genet, 2012. **8**(11): p. e1003026.
87. Decottignies, A., *Microhomology-mediated end joining in fission yeast is repressed by pku70 and relies on genes involved in homologous recombination*. Genetics, 2007. **176**(3): p. 1403-15.
88. Lee, K., et al., *Microhomology Selection for Microhomology Mediated End Joining in Saccharomyces cerevisiae*. Genes (Basel), 2019. **10**(4).
89. Anand, R., et al., *Rad51-mediated double-strand break repair and mismatch correction of divergent substrates*. Nature, 2017. **544**(7650): p. 377-380.
90. Clejan, I., J. Boerckel, and S. Ahmed, *Developmental modulation of nonhomologous end joining in Caenorhabditis elegans*. Genetics, 2006. **173**(3): p. 1301-17.
91. Joyce, E.F., et al., *Multiple barriers to nonhomologous DNA end joining during meiosis in Drosophila*. Genetics, 2012. **191**(3): p. 739-46.
92. Manheim, E.A., et al., *Cytoplasmic localization and evolutionary conservation of MEI-218, a protein required for meiotic crossing-over in Drosophila*. Mol Biol Cell, 2002. **13**(1): p. 84-95.
93. Vu, G.T.H., et al., *Deletion-bias in DNA double-strand break repair differentially contributes to plant genome shrinkage*. New Phytol, 2017. **214**(4): p. 1712-1721.
94. Deng, W., S. Henriët, and D. Chourrout, *Prevalence of Mutation-Prone Microhomology-Mediated End Joining in a Chordate Lacking the c-NHEJ DNA Repair Pathway*. Curr Biol, 2018. **28**(20): p. 3337-3341 e4.
95. Glover, L., J. Jun, and D. Horn, *Microhomology-mediated deletion and gene conversion in African trypanosomes*. Nucleic Acids Res, 2011. **39**(4): p. 1372-80.
96. Fernandez-Orgilero, A., et al., *A putative Leishmania DNA polymerase theta protects the parasite against oxidative damage*. Nucleic Acids Res, 2016. **44**(10): p. 4855-70.
97. Hooks, K.B., D. Delneri, and S. Griffiths-Jones, *Intron evolution in Saccharomycetaceae*. Genome Biol Evol, 2014. **6**(9): p. 2543-56.
98. Hu, K., *Intron exclusion and the mystery of intron loss*. FEBS Lett, 2006. **580**(27): p. 6361-5.

99. Yenerall, P., B. Krupa, and L. Zhou, *Mechanisms of intron gain and loss in Drosophila*. BMC Evol Biol, 2011. **11**: p. 364.
100. Chen, J., et al., *Whole-genome sequencing of Oryza brachyantha reveals mechanisms underlying Oryza genome evolution*. Nat Commun, 2013. **4**: p. 1595.
101. Tremblay, A., M. Jasin, and P. Chartrand, *A double-strand break in a chromosomal LINE element can be repaired by gene conversion with various endogenous LINE elements in mouse cells*. Mol Cell Biol, 2000. **20**(1): p. 54-60.
102. Han, K., et al., *L1 recombination-associated deletions generate human genomic variation*. Proc Natl Acad Sci U S A, 2008. **105**(49): p. 19366-71.
103. Vaughn, J.N. and J.L. Bennetzen, *Natural insertions in rice commonly form tandem duplications indicative of patch-mediated double-strand break induction and repair*. Proc Natl Acad Sci U S A, 2014. **111**(18): p. 6684-9.
104. Lloyd, A.H. and J.N. Timmis, *The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome*. Mol Biol Evol, 2011. **28**(7): p. 2019-28.
105. Martin, W., et al., *Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus*. Proc Natl Acad Sci U S A, 2002. **99**(19): p. 12246-51.
106. Esser, C., et al., *A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes*. Mol Biol Evol, 2004. **21**(9): p. 1643-60.
107. Hazkani-Covo, E. and S. Covo, *Numt-mediated double-strand break repair mitigates deletions during primate genome evolution*. PLoS Genet, 2008. **4**(10): p. e1000237.
108. Puertas, M. and M. Gonzalez-Sanchez, *Insertions of Mitochondrial DNA into the Nucleus. Effects and Role in Cell Evolution*. Genome, 2020.
109. Kellis, M., B.W. Birren, and E.S. Lander, *Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae*. Nature, 2004. **428**(6983): p. 617-24.
110. Muramoto, N., et al., *Phenotypic diversification by enhanced genome restructuring after induction of multiple DNA double-strand breaks*. Nat Commun, 2018. **9**(1): p. 1995.
111. Sallmyr, A. and A.E. Tomkinson, *Repair of DNA double-strand breaks by mammalian alternative end-joining pathways*. J Biol Chem, 2018. **293**(27): p. 10536-10546.
112. Trenner, A. and A.A. Sartori, *Harnessing DNA Double-Strand Break Repair for Cancer Treatment*. Front Oncol, 2019. **9**: p. 1388.
113. Volkova, N.V., et al., *Mutational signatures are jointly shaped by DNA damage and repair*. Nat Commun, 2020. **11**(1): p. 2169.
114. Zamborszky, J., et al., *Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions*. Oncogene, 2017. **36**(35): p. 5085-5086.
115. Feng, W., et al., *Genetic determinants of cellular addiction to DNA polymerase theta*. Nat Commun, 2019. **10**(1): p. 4286.
116. Lemee, F., et al., *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, 2010. **107**(30): p. 13390-5.

998 117. Polak, P., et al., *A mutational signature reveals alterations underlying deficient*
999 *homologous recombination repair in breast cancer*. Nat Genet, 2017. **49**(10): p. 1476-
1000 1486.

1001 118. McCabe, N., et al., *Deficiency in the repair of DNA damage by homologous recombination*
1002 *and sensitivity to poly(ADP-ribose) polymerase inhibition*. Cancer Res, 2006. **66**(16): p.
1003 8109-15.

1004 119. Zhou, J., et al., *Human CHD1 is required for early DNA-damage signaling and is uniquely*
1005 *regulated by its N terminus*. Nucleic Acids Res, 2018. **46**(8): p. 3891-3905.

1006 120. Kuznetsov, S., et al., *RAD51C deficiency in mice results in early prophase I arrest in males*
1007 *and sister chromatid separation at metaphase II in females*. J Cell Biol, 2007. **176**(5): p.
1008 581-92.

1009 121. Reh, W.A., et al., *The homologous recombination protein RAD51D protects the genome*
1010 *from large deletions*. Nucleic Acids Res, 2017. **45**(4): p. 1835-1847.

1011 122. Tobin, L.A., et al., *Targeting abnormal DNA double-strand break repair in tyrosine kinase*
1012 *inhibitor-resistant chronic myeloid leukemias*. Oncogene, 2013. **32**(14): p. 1784-93.

1013 123. Sallmyr, A., A.E. Tomkinson, and F.V. Rassool, *Up-regulation of WRN and DNA ligase*
1014 *IIIalpha in chronic myeloid leukemia: consequences for the repair of DNA double-strand*
1015 *breaks*. Blood, 2008. **112**(4): p. 1413-23.

1016 124. Duraturo, F., et al., *Genetics, diagnosis and treatment of Lynch syndrome: Old lessons and*
1017 *current challenges*. Oncol Lett, 2019. **17**(3): p. 3048-3054.

1018 125. Bentley, J., et al., *DNA double strand break repair in human bladder cancer is error prone*
1019 *and involves microhomology-associated end-joining*. Nucleic Acids Res, 2004. **32**(17): p.
1020 5249-59.

1021 126. Difilippantonio, M.J., et al., *Evidence for replicative repair of DNA double-strand breaks*
1022 *leading to oncogenic translocation and gene amplification*. J Exp Med, 2002. **196**(4): p.
1023 469-80.

1024 127. Zhu, C., et al., *Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene*
1025 *amplification subsequent to translocations*. Cell, 2002. **109**(7): p. 811-21.

1026 128. Wang, J.H., et al., *Mechanisms promoting translocations in editing and switching*
1027 *peripheral B cells*. Nature, 2009. **460**(7252): p. 231-6.

1028 129. Tobin, L.A., et al., *Targeting abnormal DNA repair in therapy-resistant breast cancers*. Mol
1029 Cancer Res, 2012. **10**(1): p. 96-107.

1030 130. Fan, J., et al., *Cells expressing FLT3/ITD mutations exhibit elevated repair errors generated*
1031 *through alternative NHEJ pathways: implications for genomic instability and therapy*.
1032 Blood, 2010. **116**(24): p. 5298-305.

1033 131. Newman, E.A., et al., *Alternative NHEJ Pathway Components Are Therapeutic Targets in*
1034 *High-Risk Neuroblastoma*. Mol Cancer Res, 2015. **13**(3): p. 470-82.

1035 132. He, M., et al., *A recessive variant of XRCC4 predisposes to non- BRCA1/2 breast cancer in*
1036 *chinese women and impairs the DNA damage response via dysregulated nuclear*
1037 *localization*. Oncotarget, 2014. **5**(23): p. 12218-32.

1038 133. Lee, K., Y. Zhang, and S.E. Lee, *Saccharomyces cerevisiae ATM orthologue suppresses*
1039 *break-induced chromosome translocations*. Nature, 2008. **454**(7203): p. 543-6.

1040 134. Javadekar, S.M. and S.C. Raghavan, *Snaps and mends: DNA breaks and chromosomal*
1041 *translocations*. FEBS J, 2015. **282**(14): p. 2627-45.

135. Meyer, D., B.X. Fu, and W.D. Heyer, *DNA polymerases delta and lambda cooperate in repairing double-strand breaks by microhomology-mediated end-joining in Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A, 2015. **112**(50): p. E6907-16.
136. Brunet, E., et al., *Chromosomal translocations induced at specified loci in human stem cells*. Proc Natl Acad Sci U S A, 2009. **106**(26): p. 10620-5.
137. Stephens, P.J., et al., *Complex landscapes of somatic rearrangement in human breast cancer genomes*. Nature, 2009. **462**(7276): p. 1005-10.
138. Jones, R.E., et al., *Escape from telomere-driven crisis is DNA ligase III dependent*. Cell Rep, 2014. **8**(4): p. 1063-76.
139. Stephens, P.J., et al., *Massive genomic rearrangement acquired in a single catastrophic event during cancer development*. Cell, 2011. **144**(1): p. 27-40.
140. Zhang, C.Z., et al., *Chromothripsis from DNA damage in micronuclei*. Nature, 2015. **522**(7555): p. 179-84.
141. Kloosterman, W.P., et al., *Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms*. Cell Rep, 2012. **1**(6): p. 648-55.
142. Cortes-Ciriano, I., et al., *Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing*. Nat Genet, 2020. **52**(3): p. 331-341.
143. Allen, F., et al., *Predicting the mutations generated by repair of Cas9-induced double-strand breaks*. Nat Biotechnol, 2018.
144. Bothmer, A., et al., *Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus*. Nat Commun, 2017. **8**: p. 13905.