Chapter 38

Conserved and Divergent Features of DNA Repair: Future Perspectives in Genome Instability Research

I. Kovalchuk
University of Lethbridge, Lethbridge, AB, Canada

Chapter Outline

1. AN OVERVIEW AND COMPARISON OF DNA-REPAIR PATHWAYS IN DIFFERENT ORGANISMS

Since mid-1980s, a significant progress has been made in discovering and describing in detail various mechanisms of DNA repair in different species, including bacteria and humans. The conservation of various DNA-repair mechanisms is obvious, and the discovery of a third domain of life, the "archaea" in 1977, allowed to further cement the knowledge of DNA-repair processes.

Animals have evolved many parallel and often overlapping repair pathways to be able to deal with various types of DNA damage, including chemical modifications of the base, nucleotide misincorporation, DNA and protein cross-links, and even DNA strand breaks (Fig. 38.1) [1]. Among these pathways are mismatch repair (MMR), base excision repair (BER), and nucleotide excision repair (NER), as well as global genome repair (GGR) and transcription-coupled repair (TCR), translesion synthesis (TLS), homologous recombination (HR), and nonhomologous end joining (NHEJ) (Fig. 38.1) [1].

It would be difficult to compare all types of DNA repair in organisms, and it is beyond the scope of this chapter. We only compare major repair pathways and note major differences among them.

1.1 Direct Reversal of DNA Damage

The direct reversal of DNA damage involves repair mechanisms that restore an original molecular structure of DNA without new DNA synthesis. Three major direct reversal-repair pathways are: photolyases that repair UV-induced DNA damage,

Genome Stability. http://dx.doi.org/10.1016/B978-0-12-803309-8.00038-0

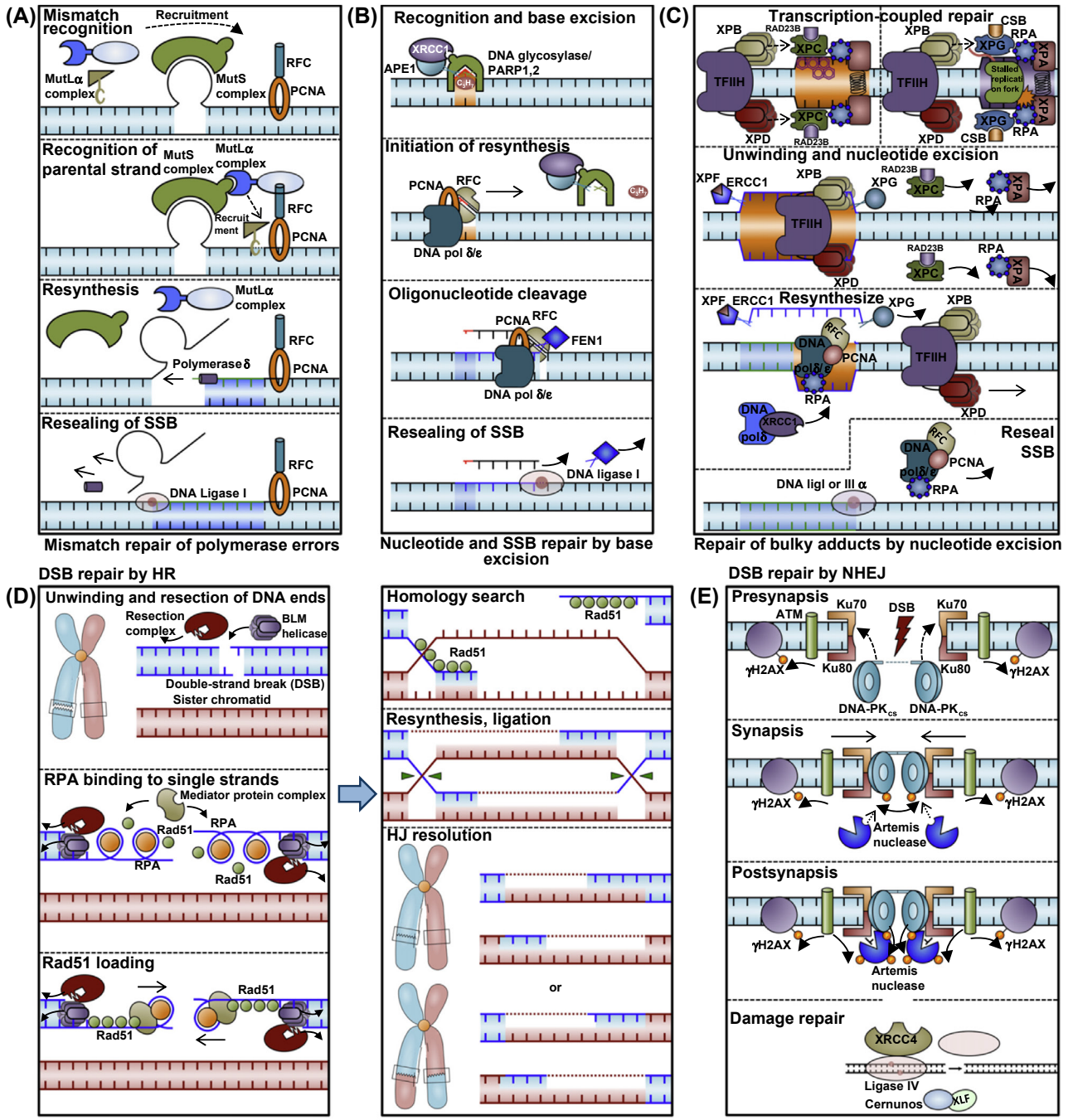

FIGURE 38.1 **Mammalian DNA damage–repair pathways.** Distinct DNA-repair pathways are traditionally thought to operate independently in parallel to clear different types of DNA lesions occurring in distinct cell-cycle phases. (A) Erroneous misincorporations, insertions, and deletions of nucleotides are handled by mismatch repair (MMR). (B) Single-strand breaks (SSBs) and nonhelix-distorting base modifications are repaired by base excision repair (BER). (C) Bulky helix-distorting lesions are cleared by nucleotide excision repair (NER) which can be initiated either as global-genome NER (GG-NER) or as transcription-coupled NER (TC-NER). (D) Double-strand breaks (DSBs) can be repaired through homologous recombination (HR) or (E) nonhomologous end joining (NHEJ), depending on the cell-cycle phase. In late stages of the cell cycle (S and G2) when the genomic DNA has been replicated and a homologous sister chromatid is available, DSBs are repaired by HR (D). (E) In G1 cells where no intact template is available for homology-mediated repair, the error-prone NHEJ pathway is preferred. *Reproduced from Dietlein F, Thelen L, Reinhardt HC. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. Trends Genet 2014;30(8):326–39; with permission.*

and two mechanisms for repairing alkylated bases, O^6-alkylguanine alkyl transferases (AGTs) and the AlkB-family dioxygenases [2]. These are ancient repair pathways that are highly active in bacteria (see details in Chapter 4). Most of them have retained their function in many eukaryotes.

Photolyases are proteins that reverse the UV damage and create a kind of a bridge between pyrimidine nucleotides. This type of repair process exists in prokaryotes and eukaryotes, including plants and animals, although mammals seem to lack such activity [3]. These enzymes are activated by light; they bind to a strand opposite to the damaged strand and cleave the bridge (like the dimer cyclobutane bridge formed between two thymines).

AGT enzymes function in both pro- and eukaryotes. These enzymes transfer the alkyl groups to their cysteine residues, rendering themselves inactive. A human AGT homolog is encoded by the O^6-methylguanine DNA methyltransferase (MGMT) gene. In *Escherichia coli*, the alkylation damage to phosphodiester DNA backbone, the S_p-methylphosphotriester, is repaired by the N-terminal domain of the Ada protein (N-Ada). There is no homolog of N-Ada found in eukaryotes, and it is unclear whether methylphosphotriester can be repaired in eukaryotes [4].

Oxidative dealkylation is performed by the AlkB protein. It activates dioxygen to oxidize the methyl group. The resulting oxidized product decomposes spontaneously and restores the base while releasing formaldehyde. Humans have eight homologs of AlkB, hABH1 to hABH8, with hABH2 and hABH3 being functional homologs [5]. The hABH1 protein has mitochondrial localization, with only a fraction of it being in the nucleus [6]. The authors demonstrated that hABH1 is a functional mitochondrial AlkB homolog that repairs 3-methylcytosine in single-stranded DNA and RNA.

1.2 Base Excision Repair

BER corrects DNA lesions that do not distort the DNA double helix, involving DNA damage caused by oxidation, deamination, and alkylation (Fig. 38.1). The importance of BER was recognized by the Nobel Prize Committee who in 2015 awarded Tomas Lindahl the Nobel Prize in Chemistry for his discovery of BER and detailed description of its components [7]. See Chapters 4 and 17 for more details on BER in prokaryotes and eukaryotes.

One of the initial critical steps of BER is performed by DNA glycosylases, enzymes that recognize and flip out the damaged base. DNA glycosylases are conserved among different species, although the conservation is mainly limited to the enzymatic core domain. Some exceptions also occur; for example, a prokaryote homolog of mammalian methyl purine DNA glycosylase (MPG) has been identified [8]. In addition, mammalian DNA glycosylases have extensions at the N- and C-ends of the protein that are not found in prokaryotic counterparts [9].

BER is conserved among bacteria and archaea, although the archaeal components of BER and their molecular biology are more similar to eukaryotes than to bacteria (Chapter 4). The BER pathways in some archaea have novel features, while other pathways use an additional mechanism to prevent mutations due to genomic uracil. *Ferroplasma acidarmanus* encodes a novel AGT protein (AGTendoV) that has an O^6-methyltransferase domain fused to an endonuclease V domain [10]. This bifunctional enzyme has been found in other archaeal genomes suggesting that it may be a general adaptation to their harsh environments. In addition, archaea have a unique BER ability; they use uracil scanning DNA polymerases. While bacterial polymerases replicate past uracil by inserting an adenine, archaeal replicative polymerases stall before misplaced uracils, representing a read-ahead proofreading function not found in bacteria or eukaryotes (Chapter 4).

While bacteria contain six DNA glycosylases, mammals have 11 glycosylases (Chapter 4). In mammals, BER functions in repair of both the nuclear and mitochondrial genomes, and both genomes undergo short- and long-patch repair [8]. It should be noted, however, that the ligation step of BER is performed by LIG1 upon DNA repair in the nucleus and by LIG3 in mitochondria; LIG3 seems to function only in vertebrates.

In *E. coli*, misincorporated uracil is removed mostly by the uracil-DNA-glycosylase UNG (UDG). The UDG superfamily consists of six subfamilies in prokaryotes and three members (subfamilies I, II, and III) in higher eukaryotes [11]. Only one member, the UNG protein, has been found in the human mitochondria [12].

1.3 Nucleotide Excision Repair

NER is an excision-repair pathway for repairing bulky DNA lesions (Fig. 38.1). NER mechanisms are highly conserved across all three domains of life; the major steps in GGR and TCR are conserved. At the same time, some significant differences exist.

The NER pathways in bacteria and archaea are functionally similar; however, despite the ancient nature of archaea, some aspects of their NER pathways are more similar to eukaryotic versions than to bacterial versions, and they may or may not have *uvr* homologs (Chapter 4). The presence of clear *uvr* homologs seems to coincide with their lifestyle: mesophilic

archaea tend to have *uvr* genes, while hyperthermophilic archaea do not have them [13]. A universal feature, however, seems to be the presence of homologs of eukaryotic factors. Therefore, archaea such as *Methanosarcina mazei* clearly have both bacterial and eukaryotic homologs of NER.

NER is substantially more complex in eukaryotes as compared to prokaryotes. While bacteria contain three proteins involved in NER commonly known as the UvrABC system, up to 50 different proteins may be involved in NER in animals (see Chapter 17). In addition to obvious differences between the names of proteins involved and their sequence conservation, large differences exist in many steps of the process. While prokaryotes require only three proteins, eukaryotes employ many more of them with a versatile and sometimes unique function (see Chapters 4 and 17). The size of the excised DNA fragment is also substantially different: whereas in bacteria, a 12–13 nt–long fragment is removed, in eukaryotes it is more likely to be 27–29 nt. Since NER is functionally less complex in bacteria, details of the biochemistry of NER in bacteria are better understood than in eukaryotes. Please refer to Table 4.1 for a direct comparison of enzymes involved in NER in prokaryotes vs. eukaryotes.

1.4 Mismatch Repair

MMR is a complex repair process that recognizes and replaces mismatched nucleotides. It is present in all species and bacteria, excluding some *Archaea* (eg, *Actinobacteria*, *Mollicutes*, *Mycobacterium tuberculosis*, and *Helicobacter pylori*) that likely lost this repair pathway during the evolution [14]. In *E. coli*, the MMR pathway has three main functional proteins: MutS, MutL, and MutH (Chapter 4).

Mismatches in bacteria are recognized by MutS homodimer. MutS protein subunits A and B encircle mismatch-containing oligonucleotide duplexes by binding in an asymmetric manner—only one subunit contacts the mismatch [15]. In the process of evolution, eukaryotic organisms further developed this asymmetry by developing the binding through a combination of two different proteins—MutS homologs (MSHs) that form two heterodimers. Eight MSHs homologs are known to be capable of forming various dimers. MSH1, MSH6, MSH7, and MSH8 contain a special motif for the recognition of a mismatch—the GxFxE motif. While the biochemical activity of MSH8 is unknown, MSH6 forms heterodimers with MSH2, which leads to the formation of the MutSα complex, and MSH7 interacts with MSH2 to form the MutSγ complex [15]. In *Arabidopsis*, the MutSγ heterodimer binds efficiently to C/G and G/A mismatches but lacks its activity on single-nucleotide insertion-deletion loops (IDLs) [16]. In yeasts, larger IDLs are recognized by the MSH2/MSH3 heterodimer known as MutSβ [17].

In the second step, a MutL dimer binds the MutS–DNA complex, thereby stabilizing it and activating the MutH restriction endonuclease. In *E. coli*, MutL functions as a homodimer. Eukaryotes contain several MutL homologs (MLH proteins). In humans, functional MLH homologs form a heterodimer with MLH1 and PMS2 called MutLα. Two other heterodimers are known in human cells: MLH1/PMS1, referred to as MutLβ, with an unknown function [18], and MutLγ (MLH1/MLH3) which can partially substitute the activity of MutLα in vitro [19]. In *E. coli*, MutL was considered as an intermediate molecule linking MutS and MutH proteins and mediating the interaction between the mismatch recognition machinery and the strand-discrimination and excision machineries [15]. MutL in *E. coli* has no endonucleolytic activity. Such activity has likely evolved in eukaryotes. It has been discovered that MutLα possesses the endonuclease activity [20]. Kadyrov et al. [20] showed that human MutL alpha is a latent endonuclease activated in a mismatch-, MutSα-, RFC-, PCNA-, and ATP-dependent manner. See more details about the function of MMR in mammals in Chapter 18.

In the third step, MutH nicks the strand containing the incorrectly incorporated base; this protein is present in *E. coli* and *Salmonella*. MutH has no homologs in eukaryotes, and it is proposed that its activity is substituted by MutL homologs (MLH proteins).

Archaea substantially differ from bacteria in MMR. For example, in mesothermophilic archaea, the MMR pathways mirror the canonical bacterial pathways [13]. In contrast, hyperthermophilic archaea lack MutS and MutL homologs; these species also lack Uvr proteins. Nevertheless, genome replication is accurate in these organisms [13]. In fact, the error rate per genome replication is lower in the archaeon *Sulfolobus acidocaldarius* than in nearly all other microbial genomes analyzed similarly [13]. Therefore, the lack of increased mutagenesis suggests that hyperthermophiles have the same mechanism that accomplishes the same net result as MMR (see details in Chapter 4).

1.5 Double-Strand Break Repair

Double-strand break (DSB) repair is broadly divided into NHEJ and HR, and there is a great diversity and variability in these repair pathways in all three domains of life. DSB repair in mammals is presented in Fig. 38.1.

1.5.1 Nonhomologous End Joining

E. coli do not have the canonical NHEJ pathway. One hypothesis suggests that since bacteria are often in the replicative stage, they constantly possess homologous chromosomes for the efficient HR repair, thus making NHEJ nonessential [21]. Those bacteria that do have NHEJ are frequently found to be in the stationary stage when no chromosome homologs are present [22]. A 2010 work, however, demonstrated that *E. coli* strains do possess an end-joining mechanism, now called alternative end joining (A-EJ) [23]. This novel pathway does not share conserved factors with the canonical NHEJ pathways. It depends on bidirectional strand resection, the frequent use of microhomology, and nontemplated DNA synthesis. In archaea, a functional repair pathway was identified only in 2013 in a mesophilic archaeon *Methanocella paludicola*, although conserved components of NHEJ were previously identified [24,25]. The scientists described the complete NHEJ complex in archaea consisting of DNA ligase (Lig), polymerase (Pol), phosphoesterase (PE), and the Ku protein [26]. The A-EJ or A-NHEJ pathway has also been well described in mammals (see Chapter 19).

Unlike the eukaryotic NHEJ ligase IV, LigD is a large multidomain protein that contains three components within a single polypeptide: a polymerase (POL) domain, a phosphoesterase (PE) domain, and a ligase (LIG) domain [21]. Therefore, it was initially believed that bacterial Ku and LigD are sufficient to repair all DSBs generated in vivo [27]. Later on, a faithful NHEJ pathway operating specifically for 3′ overhang DSB repair and functioning independently of Ku and LigD proteins was reported [28]. It was also shown that the structure of broken ends determines the pathway and the outcome of DSB repair. In wild-type cells, this pathway joins nearly one-third of the 3′ overhangs with 100% fidelity [28].

In NHEJ occurring in vertebrates, a complex of Artemis:DNA-PKcs processes DNA ends with different configurations. The *DNA-PKcs* gene has not been found in *Arabidopsis* [29] or in other plant genomes. Instead, in yeast, plants, and vertebrates, the same job is likely performed by the MRX complex [30]. In mammals, Pol μ can add nucleotides in a template-independent manner under physiological conditions [31], whereas Pol λ can do this only when Mg^{2+} is replaced by Mn^{2+} [32]. Similarly in bacteria, the polymerase can also add nucleotides in a template-independent manner [21], likely reflecting convergent evolution.

NHEJ appears to be mechanistically different in *Saccharomyces cerevisiae* and mammals since in yeast, blunt-end ligation is inefficient, and at least one base pair of terminal microhomology may be necessary for efficient NHEJ to occur [30]. In *S. cerevisiae*, NHEJ can actually occur in the absence of Lig IV, but it is at least 10-fold less efficient than with Lig IV [33]. In contrast, Ku-deficient yeast cells repair DSBs as efficiently as wild-type yeast cells.

In mammals, there are no microhomology requirements for end joining, although occasionally single nucleotides and longer microhomologies can be found. This is a default situation for most genomic positions (excluding repetitive elements). The use of more extensive microhomology is possible in mammals when such homology is provided artificially or when some NHEJ components are lacking. In the latter case, NHEJ occurs at a slower pace with more frequent resections and alignments that use more extensive (2–3 bp) microhomology [30].

1.5.2 Homologous Recombination

The early stage of the HR mechanism is well studied and appears to be conserved in three life domains [34]. All steps of DSB repair are conserved between archaea and bacteria, although protein conservation is not as obvious as it may seem. Similarly to other repair pathways, archaea contain protein homologs from bacteria and eukarya. For example, proteins such as Mre11, Rad50, and RadA are conserved between archaea and eukarya.

In bacteria, DSBs are repaired through one of two overlapping repair pathways, one relying on the multifunctional helicase/nuclease complex RecBCD (the RecBCD pathway) and the other involving the RecF protein (the RecF pathway) [35]. In the RecBCD pathway, resection is catalyzed by the RecBCD complex, whereas in eukarya, it is done with the Mre11 and Rad50 protein complex [36]. In archaea, Rad50 and Mre11 cooperate with the HerA helicase and NurA nuclease to catalyze 3′-end resection for HR [37].

Next, the single-stranded DNA-binding protein (SSB) (known as replication protein A, RPA, in eukaryotes) binds to the 3′-ssDNA ends generated by the action of RecBCD or Mre11/Rad50, or Mre11/Rad50/HerA/NurA. SSBs and RPA function as dimers, trimers, and tetramers. In bacteria, they occur as homodimers and homotetramers, whereas in archaea, they function as monomers, dimers, and trimers [38]. In eukaryotes, RPA acts as a homodimer and heterotrimer. The DNA recombinase of the RecA family is conserved in all three domains, but it has a different name in each domain: the RecA recombinase in bacteria, RadA in archaea, and Rad51 and its paralogs in eukaryotes. RadA is more similar to Rad51 rather than to RecA. In animals, Rad51 is an essential protein that causes embryonic lethality when disrupted. Similarly, the disruption of RecA in bacteria or Rad51 in yeast is also highly deleterious to the cell, although it is not lethal [39]. Eukaryotes contain multiple Rad51 paralogs, including Rad51B, Rad51C, Rad51D, DMC1, XRCC2, and XRCC3 that cooperate with Rad51 in strand exchange. Archaea also have several RadA paralogs, including RadB, RadC2, Sso2452, and Rad55B.

Proteins involved in the late stage of HR, representing the RuvABC–Holliday junction (HJ) complex that processes recombination intermediates, have been well described in bacteria [40]. Eukaryotic cells possess two mechanisms for HJ processing: the first mechanism known as HJ dissolution relies on the activity of the BTR complex (BLM helicase–topoisomerase IIIα–RMI1–RMI2), whereas the second one requires structure-selective nucleolytic endonucleases such as MUS81–EME1 and GEN1 [41]. While the outcome of HJ dissolution is exclusively the formation of non-crossovers (NCOs), nucleolytic resolution results in the formation of NCOs as well as crossover events (COs). Mitotic cells preferentially use NCOs to prevent the loss of heterozygosity [42]; this correlates well with the fact that the HJ resolvases function late in the cell cycle to ensure a proper chromosome segregation.

Yeast cells are able to separate the recombination resulting in mitotic and meiotic DSB repair leading to CO and NCO formation [41]. In mitosis, the majority of double Holiday junctions (dHJs) are processed at early stages of the cell cycle by STR-mediated dissolution leading to NCOs. In *sgs1* mutants, however, DNA joint molecules formed between the damaged and sister chromatids persist to a later stage of the cell cycle where they are processed by Mus81–Mms4 or Yen1, resulting in a mixture of COs and NCOs as an outcome [41]. As of 2016, the temporal separation of CO and NCO formation in human cells has not been demonstrated. However, human cells deficient in BLM (Bloom's syndrome cells) have the increased frequency of CO formation [43]. CO formation in these cells occurs through the activities of MUS81–EME1, SLX1–SLX4, and GEN1 [44].

2. RECENT ADVANCES AND FUTURE DIRECTIONS IN DNA REPAIR

Despite a substantial progress in understanding the mechanisms of functioning of various DNA-repair pathways, many questions remain unanswered. We only discuss some of them.

2.1 The Remaining Questions in MMR

MMR is one of the critical repair mechanisms involved in fixing replication errors as well as regulating other repair mechanisms such as HR and DNA-damage signaling. Despite many advances in the field, several questions remain to be answered (for review, see [15]).These are:

1. Where does the MSH/MLH complex assemble? Is it assembled at the mismatch or at the strand-discrimination signal?
2. If the complex does indeed assemble at the mismatch and translocate, is the direction of its movement random, or does it depend on its initial loading on the heteroduplex?
3. What directs the process exonuclease degradation toward the mismatch rather than away from it?
4. Are all proteins involved in MMR described? Is there another exonuclease involved in MMR?

Some additional questions in MMR research were asked by Yang and Hsieh in Chapter 18. In particular, it remains to be established how the MMR machinery is recruited to newly replicated DNAs and how it is spatially positioned in relation to the advancing replisome. Other unanswered questions are: When do MutS and MutL proteins interact, and when do they function separately? How exactly is the MutL endonuclease activity targeted to the newly synthesized strand, and what is its biological scope? What is the mechanism of recruitment of replicative polymerases to single-strand gaps? In what contexts are error-prone polymerases employed instead and what are the consequences? How is MMR influenced by higher-order chromatin arrangements and nuclear architecture? Finally, how can the knowledge of MMR mechanisms improve clinical diagnostics and therapeutic outcomes?

Although the answers to the most of the abovementioned questions remain unknown, several models summarized by Li et al. [45] may have merit, and future research will show whether they hold true.

2.1.1 The Role of a Strand-Discrimination Signal in MMR

In MMR, the protein machinery has to recognize the mismatch on the strand that is opposite to the strand with the mismatch and to perform a strand-specific nick on the strand with the mismatch. One of the unanswered questions concerns the mechanism by which MMR proteins communicate between two physically separated DNA sites: the mismatch and the strand-discrimination signal [45]. The strand-discrimination signal represents a strand-specific nick that occurs in a similar manner in all three domains of life. It is not clear, however, what the source of the nicking activity in eukaryotic cells is. Several alternative models describing this process have been proposed; they can be classified into "*cis*-" or "moving" and "*trans*-" or "stationary" models (Fig. 38.2). The stationary model suggests that interactions among MMR proteins induce DNA looping that allows to bring two distant sites together; MutS or MSH heterodimers, MutSα and MutSβ remain bound

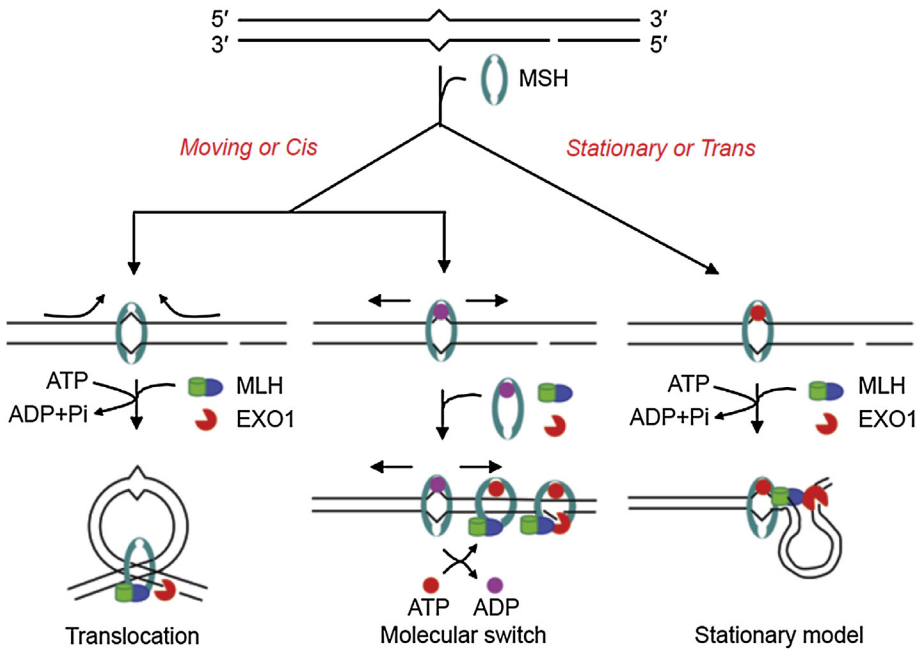

FIGURE 38.2 **Models for signaling downstream MMR events following mismatch recognition.** A schematic diagram for signaling between the mismatch and the strand discrimination signal is shown. Here, a 5′-nick is the strand discrimination signal. Similar models apply for 3′-nick-directed MMR. The "stationary" or "*trans*" model (*right*) emphasizes that MutS or its homolog (MSH) proteins remain bound at the mismatch. It is the protein–protein interactions that induce DNA bending or looping that brings the two distant sites together. The two DNA sites can cooperate in a "*trans*"-configuration. In two "*cis*" or "moving" models, one called the "translocation" model (*left*) and the other called the "molecular switch" or "sliding clamp" model (*middle*), it is hypothesized that the MSH proteins bind to the mismatch and then move away from the site to search for the strand discrimination signal. The translocation model suggests that ATP hydrolysis drives unidirectional movement of MSH proteins, resulting in the formation of an α-like loop. In the molecular switch model (*center*), binding of an MSH protein (in its ADP-bound state) to the mismatch triggers an ADP to ATP exchange that promotes bidirectional sliding of the protein away from the mismatch, thereby emptying the mismatch site for an incoming MSH protein. Mismatch excision begins when an MSH protein reaches the strand break. *Reproduced from Li GM. Mechanisms and functions of DNA mismatch repair. Cell Res 2008;18(1):85–98; with permission.*

at the mismatch [46]. The ATPase activity of MSH heterodimers is required for proofreading and verification of mismatch binding and downstream excision [47]. In support of this model, it was shown that MutS recognizes a mismatch, activating MutH to cleave the GATC site located on a separate DNA molecule that does not contain a mismatch [47]. Similarly, it was also shown that DNA excision activated by the presence of a mismatch could be initiated when a biotin–streptavidin blockade was placed between the mismatch and the preexisting nick [48].

The second model, known as a "*cis*-" or "moving" model, suggests that the MutS–MutL (or MutSα/β-MutLα) complexes bind to the mismatched site and then move away from the site to search for the strand break where exonucleases can be recruited to initiate excision [45]. There have been known two moving models: the "translocation" model and the "molecular switch" model (also known as the "sliding clamp" model) (Fig. 38.2). In the translocation model, ATP reduces the mismatch-binding affinity of MutS, and ATP hydrolysis drives unidirectional translocation of MutS proteins along the DNA helix [49]. DNA is threaded through the protein complex until the complex reaches a strand-discrimination signal in either orientation; this process forms a DNA loop (Fig. 38.2). In the molecular switch model, MutS binds to the mismatched DNA in an ADP-bound state [45]. MutS binds the mismatch, thus triggering a conformational change in these proteins and allowing an ADP to ATP exchange. ATP binding promotes a second conformational change that allows MutS to form a sliding clamp [50,51]. ATP binding also signals downstream events such as the formation of the ternary complex with MutL or MLH heterodimers in eukaryotes and sliding of the ternary complex from the mismatch to the strand break [50,51].

2.1.2 The Role of MMR in DNA-Damage Signaling

MMR plays an essential role in DNA-damage signaling. Two models have been proposed to describe details of the signaling. In the "futile DNA-repair cycle" model (Fig. 38.3), strand-specific MMR which only interacts with replicated DNA encounters DNA lesions in the template strand and engages in a futile DNA-repair cycle by activating DNA damage–signaling pathways and inducing the cell-cycle arrest and apoptosis [52]. This model was synthesized from several experiments conducted in vivo and in vitro. Exposure to MNNG was shown to induce DNA breaks/gaps, the cell-cycle arrest, and

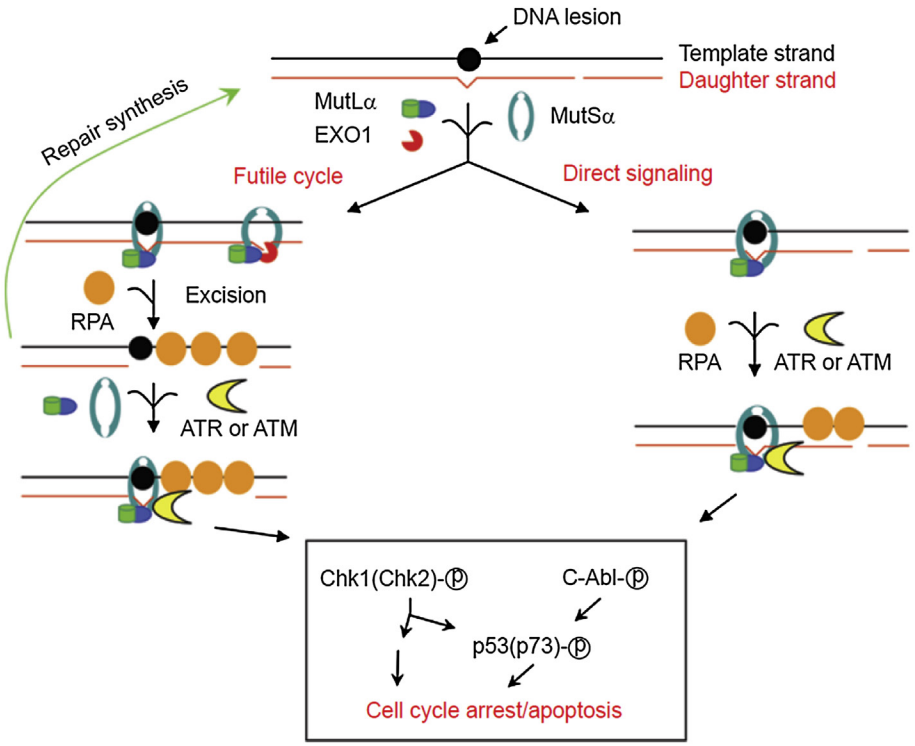

FIGURE 38.3 Models for MMR-dependent DNA-damage signaling. The "futile DNA-repair cycle" model (*left*) suggests that DNA adducts (*solid black circle*) induce misincorporation which triggers the strand-specific MMR reaction. Since MMR only targets the newly synthesized strand for repair, the offending adduct in the template strand cannot be removed, and it will provoke a new cycle of MMR upon repair resynthesis. Such a futile repair cycle persists and activates the ATR and/or ATM damage–signaling network to promote cell-cycle arrest and/or programmed cell death. The direct signaling model proposes that recognition of DNA adducts by MSH–MLH complexes allows the proteins to recruit ATR and/or ATM to the site, activating the downstream damage signaling. *Reproduced from Li GM. Mechanisms and functions of DNA mismatch repair. Cell Res 2008;18(1):85–98; with permission.*

persistent nuclear foci at DNA-damage sites [53]. Repair foci associated with DNA damage contain damage signaling and DNA-repair proteins such as ATR, γ-H2AX, and RPA(45). In addition, it was shown that the nicked circular heteroduplex plasmid DNA containing a single O^6-methylguanine (O^6-meG)-thymine (T) mispair can only be repaired by MMR when the lesion (O^6-meG) and the nick are on the same strand; MMR cannot repair the mismatch when the lesion and the nick are on the opposite strands, suggesting a futile repair process [54]. The direct signaling model (Fig. 38.3) proposes that hMutSα/hMutLα heterodimer triggers DNA-damage signaling directly by recruiting ATM or ATR/ARTIP to the lesion and activating a checkpoint response [45]. The functionality of this model is supported by experiments showing that ATR and ATRIP form a complex with MutSα/MutLα in the presence of O^6-meG/T which activates the ATR kinase and phosphorylates Chk1 [55] (see also Chapter 18). This model is consistent with the fact that the MutSα/MutLα complex functions as a sensor for DNA damage in mammalian cells.

2.2 The Remaining Questions in DSBs Repair

Organisms vary dramatically in the frequency of using NHEJ or HR in strand-break repair. As it has been mentioned earlier, bacteria, for example, are much less efficient in canonical NHEJ, but are very proficient in HR. Similarly, yeast also uses HR more frequently than NHEJ repair. In contrast, higher eukaryotes appear to predominantly utilize NHEJ. Such selective processing of strand breaks in yeast and higher eukaryotes can be in part explained by their genome content. The yeast genome is rather small and consists predominantly of coding sequences, and the content of noncoding and repetitive elements is relatively low; whereas higher eukaryotes have much larger genomes with a higher ratio of noncoding to coding sequences. It can be hypothesized that being the dominant repair mechanism, HR in yeasts prevents the generation of random changes in coding DNA. On the contrary, relatively small-scale deletions and insertions resulting from NHEJ repair can be tolerated by higher eukaryotes due to their large-size genomes and a low ratio of coding to noncoding sequences [56]. Another important factor to consider is the amount of time and energy required for HR to occur. In large genomes, locating a homologous template in somatic, nondividing cells requires a substantial amount of time, and instead, it can be very beneficial to use NHEJ for strand-break repair.

NHEJ is the predominant mechanism in higher eukaryotes including plants, and HR is more active during S and G2 phases as compared to other phases of the cell cycle. Now there is no doubt that genome sizes are the consequences of species evolution. The question to ask would be: how do cells decide on the preferred mechanism of DNA repair? Since mutation frequency varies among different cell types and during different developmental stages of an

organism, there definitely exists some type of chromatin-based regulation of genome stability. See Chapter 28 on epigenetic regulation of genome stability in cancer for more details.

2.2.1 Nonhomologous End Joining

A good summary of the remaining questions in NHEJ research was made by Lieber [30].

2.2.1.1 Are the Two DNA Ends Held in Proximity During NHEJ or Is There Synapsis?

Depending on the size of the microhomology, there can be different requirements on Ku proteins that hold DNA ends together. In vitro experiments show that when 4 bp of terminal microhomology is used, the XRCC4:DNA ligase IV complex does not require a protein helper to bring and hold two ends of DNA together. In contrast, at 2 bp or less of terminal microhomology, the addition of Ku improves XRCC4:DNA ligase IV ligation. This is likely due to the ability of Ku to stabilize XRCC4:DNA ligase IV rather than to bring DNA ends together. Under certain salt conditions (30 mM monovalent salt or lower), DNA-PKcs appear to help bring DNA ends together. Therefore, there is an ambiguity as to the requirement of helper proteins to hold DNA ends together during NHEJ repair. Another relevant question is whether the two DNA ends generated at a single DSB are joined more efficiently than two DNA ends that arise at a distance (as in a chromosomal translocation) [30].

2.2.1.2 How Do DNA-Damage Response Proteins Interact With NHEJ Enzymes?

NHEJ is a rapid and efficient repair process, and it may repair DNA damage without the activation of DNA-damage response (DDR) proteins, ATM, the RAD50:MRE11:NBS1 complex, γ-H2AX, and 53BP1. The kinetics of the NHEJ machinery and DDR is not entirely clear and is still an area of intensive research. Many in vitro and in vivo biochemical experiments are conducted in such a way that many breaks are present simultaneously. Such conditions may artificially favor the activation of DDR pathways, therefore skewing the enzymology of NHEJ. Obviously, a competition between HR and NHEJ components may be one of the aspects that arises. In addition, systems in which multiple DSBs or SSBs are introduced trigger the activation of DDR, NHEJ, and HR pathways by altering the kinetics of the association with DNA damage, the chromatin, and nuclear scaffold, as well as with each other. Cells in which so many breaks occur simultaneously are not likely to survive in nature, and therefore, in reality, what we observe in the experimental system may not occur in nature.

2.2.1.3 Are All Components of NHEJ Accounted for?

Do we now know all proteins involved in NHEJ? It is unlikely, considering how many components in various organisms have been developed through the convergent evolution. Most obvious genetic components have been identified through experiments that show the sensitivity of cells or organisms to DNA-damaging agents such as ionizing radiation (IR) or cisplatin. Novel components may be identified by using specific experimental conditions, such as a specific mutant background, a specific cell type or/and the use of specific DNA-damaging agents.

The Werner's 3′-exonuclease/helicase enzyme is likely to be involved in DSB repair through NHEJ because WRN interacts with Ku and PARP1. However, since the mutant is not sensitive to IR and the interaction with Ku and PARP1 may simply reflect the role of WRN in replication fork repair, it is still unclear whether the WRN protein is one of the NHEJ components [57,58]. Another potential candidate can be metnase that has been proposed as a possible nuclease and helicase functioning in NHEJ [59]. Metnase appears to be present only in humans; it seems to be absent in all vertebrates tested, and it is also absent in yeast. In addition, the absence of genetic knockouts to demonstrate IR sensitivity and the absence of experiments with purified protein are also a problem. Future experiments will show whether or not Metnase is a unique NHEJ protein in humans.

2.2.2 Homologous Recombination

Many questions still remain unanswered as to the recognition of DSBs, the choice of the pathway, dynamics and kinetics of repair, and so on. How do cells choose which DNA-repair pathway to use to fix DSB? Is the choice different in different species, and what are the components responsible for the choice? What is the kinetics of HR repair in various organisms? What is the chromatin dynamics at various steps of HR repair? Specifically, what impact does chromatin have on the late stages of HR, including the formation and resolution of joint molecules? These and other questions are still to be addressed.

One of the 2015 reports clarified one step in the chromatin modification associated with DSB repair. DSBs can trigger nonproteolytic ubiquitylation of adjacent chromatin areas exposing the DNA to DNA-binding proteins. The E3 ubiquitin ligase RNF168 catalyzes ubiquitylation of H2A histones, leading to recruitment of the 53BP1 repair protein. It was

demonstrated that ubiquitylation of the H1 linker histone rather than one of the core histones by RNF8 and UBC13 ubiquitin ligases is also important for DSB repair [60].

Another unresolved question is the site specificity of V(D)J cleavage. V(D)J genes are surrounded by recombination signal sequences (RSS), and the recombination is initiated by RAG2 proteins. The human genome contains a great number of RSS and RAG2-binding sites, so it is not clear how V(D)J is mediated and carried out at correct sites.

There is also an unanswered question regarding the use of HR during TLS repair. HR repair is also involved in the DNA-damage tolerance (DDT) repair pathway that promotes ssDNA-lesion bypass. This HR repair is known as the salvage HR-repair pathway. During replication, this pathway generates hyper-recombinogenic intermediates that lead to chromosome rearrangements and genome instability [61]. The salvage HR-repair pathway functions independently of the Rad5-dependent error-free tTLS-repair pathway and is repressed during the S phase [62]. It is unclear till 2016 why it is actually repressed during the S phase. It was suggested that both error-free and error-prone pathways are initiated by Rad51-mediated strand invasion, but they are different in the mechanism forming sister chromatid junctions, with Rad5 being dependent on PCNA polyubiquitination, whereas HR pathways being dependent on Rad51-mediated strand exchange that is presumably prevented by SUMOylated PCNA and Srs2 during the S phase. It only becomes derepressed when the level of sumoylation of PCNA declines in the end of the S phase [62].

Research on mammals demonstrates that meiotic HR occurs at DNA DSB hot spots and is initiated by programmed DSBs [63]. Since recombination repair occurs between homologous chromosomes, the sequence from the unbroken (undamaged) chromosome is copied. One prevailing hypothesis based on data from research on mammals suggests that over time DSB hot spots should be replaced with normal (cold) regions that should undergo the normal recombination frequency [64]. In consistence with this hypothesis, patterns of hot spots in mammals exhibit a higher divergence between species and individuals [65]. Research articles in 2015 in *Science* (vol 350, issue 6263) challenge this hypothesis, demonstrating that yeast and birds have a remarkable evolutionary stability of meiotic recombination patterns [66–68].

3. FUTURE DIRECTIONS IN RESEARCH ON DNA REPAIR, GENOME STABILITY, AND CANCER

DNA-repair capacity is often altered in cancer cells. This leads to the accumulation of many genomic alterations that allow cancer cells to prevail in the competition for resources and to survive for much longer than normal cells. Enormous variations in cancers do not allow for a complete understanding of the role that deficiencies in DNA repair and genome instability play in the development of cancers. However, advances in next-generation sequencing allowed to sequence multiple cancer cell lines and to obtain details of genomic alterations associated with these lines (Fig. 38.4). These studies showed that mutations in different cancers accumulate in distinct signaling pathways [69–72]. It was shown that inactivating mutations in the affected pathways were cancer specific (Fig. 38.4). For example, it was found that mutations in HR-repair genes were enriched in breast and ovarian cancer, as well as in chronic lymphocytic leukemia (CLL). In contrast, colorectal cancers are altered in MMR and HR-repair pathways, whereas prostate cancers are enriched in mutations inactivating HR and NER pathways (Fig. 38.4A–C,E,F) [72–74]. At the same time, cancers developing in endometrial tissues had no apparent clustering (Fig. 38.2D) [75]. Because of such heterogeneity and specificity in targeting particular DNA-repair pathways in specific cancers, it can be an arduous task to identify the subcohort of DNA repair–defective tumors for each cancer entity. Therefore, it may be more efficient to identify genetic alterations in the pathways that are predominantly altered in the respective tumor types [1].

Another area of research that might be interesting to explore in the future is the identification of proteins that play a critical role in several DNA-repair pathways. Targeting such genes in cancers may appear to be more effective. Proteins involved in distinct repair pathways may appear to be hubs of DNA-damage recognition and repair signaling [1]. The best example is the MuTSβ complex involved in MMR and HR repair. Under normal conditions, the Msh2/Msh3 heterodimer complex binds to intact dsDNA with a low efficiency, whereas when dsDNA has a 30 ssDNA overhang or contains large (~8 nt) insertion loops, the affinity increases dramatically [1]. Other enzymes, such as BRCA1, BRCA2, ERCC1, RAD50, PARP1, ATM, PTIP, XRCC1, PCNA, RFC, and RPA, share similar dual properties with Msh3.

For example, the PARP1 protein functions in both BER and A-NHEJ-repair pathways (Fig. 38.1B,E) [1]. Similarly, BRCA1 and BRCA2 proteins are involved in several DNA-repair pathways; besides being involved in HR repair, they are also important for interstrand cross-link (ICL) repair [76,77] (Fig. 38.5). Moreover, BRCA proteins might be important for the protection of replication forks, and in this role, they may be assisted by Rad51 protein [78,79]. Specifically, BRCA2-dependent fork stabilization requires the interaction between Rad51 and BRCA2 through the conserved C-terminal domain in BRCA2, and this domain is different from the BRCT domain that interacts with Rad51 for HR repair.

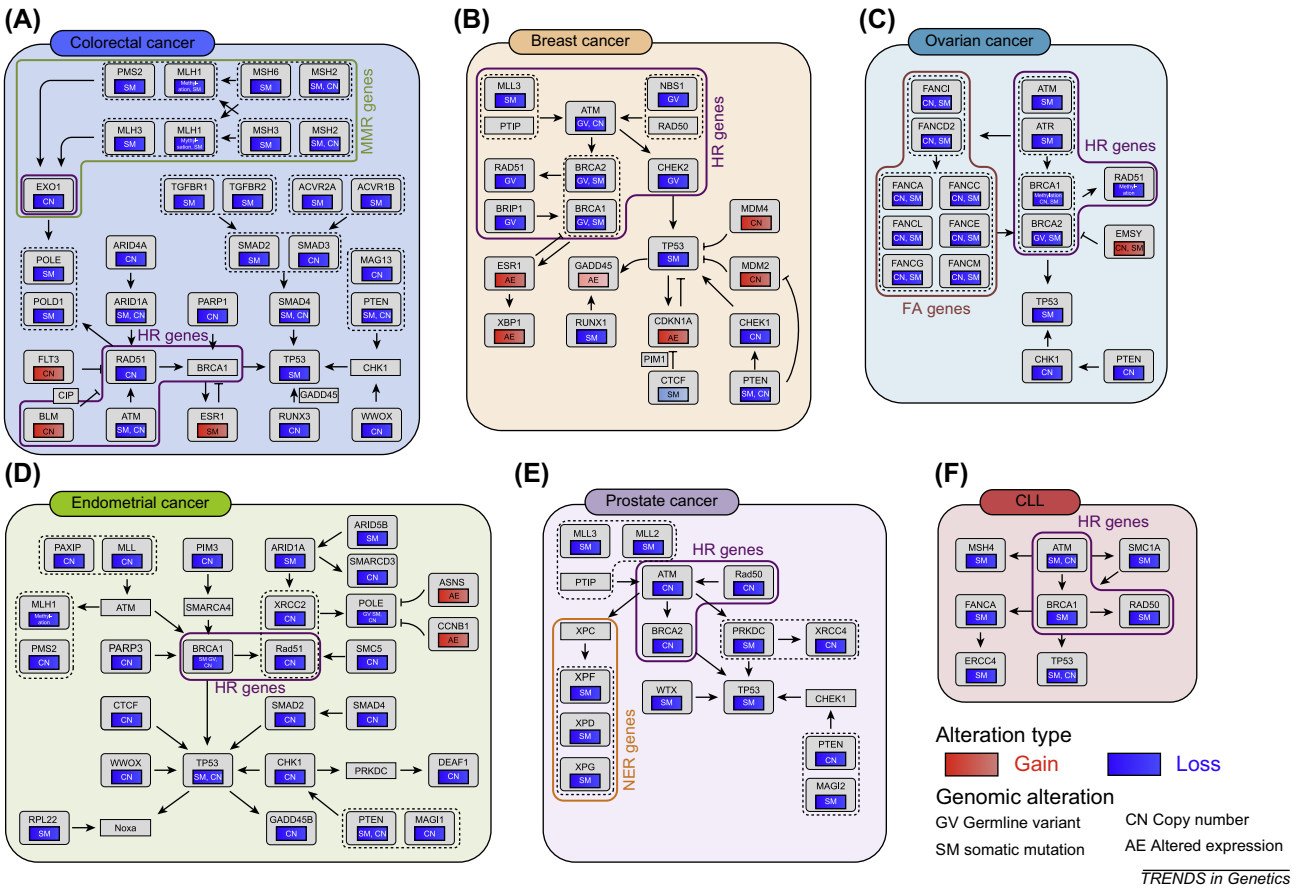

FIGURE 38.4 **Genomic aberrations in DNA-repair pathways are common in cancer.** Genes involved in different DNA damage–repair pathways are frequently altered by genomic alterations in cancer. Germline variants (GV), somatic mutations (SM), copy-number alterations (CN), and altered expression patterns (AE) in damage-repair genes are schematically represented across (A) colorectal, (B) breast, (C) ovarian, (D) endometrial, and (E) prostate cancer, as well as in (F) chronic lymphocytic leukemia (CLL). All genomic alterations were scored as significant in the respective cancer entity either by MutSig (mutations) or GISTIC (copy-number alterations) algorithms. Genomic gains are colored in red and functional losses are colored in blue. To derive entity-specific signatures of affected damage-repair pathways, genes were clustered according to their association to different pathways (*encircled*). Broken lines encircle genes encoding proteins that act together in a protein complex. *Reproduced from Dietlein F, Thelen L, Reinhardt HC. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. Trends Genet 2014;30(8):326–39; with permission.*

The knowledge about the involvement of DNA-repair proteins in multiple signaling and DNA-repair hubs is critically important because such proteins are frequently affected in cancers, including genomic and epigenomic alterations [69–72]. The prominent example is the classification of colorectal cancers (CRCs) deficient in MSH3; for therapeutic purposes, CRCs are classified as microsatellite unstable with MMR defects [70]. Such stratification, however, completely neglects the fact that MSH3 is also involved in the HR-repair pathway, and thus CRCs can be considered as HR-defective tumors that can be treated using PARP1 and/or DNA-PKcs inhibitors [80,81] (Fig. 38.5). Therefore, the identification of molecular hubs such as BRCA1, BRCA2, RPA, or MSH3 involved in multiple repair pathways will allow targeting cancers by disabling the ability of cells to deal with different types of DNA damage. For example, knowing that BRCA1 is involved in multiple DNA-repair pathways, such as HR, ICL (the Fanconi anemia–repair pathway), and mechanisms of replication fork protection would allow to treat BRCA1-deficient tumors with a combination of PARP inhibitors that target the HR pathway and platinum-based drugs that introduce cross-links. It is expected that the surrounding normal tissues will still be able to cope with these drugs by activating multiple repair pathways that function normally (such as NER and FA pathways) to cope with cisplatin damage, and the HR-repair pathway to overcome PARP inhibition [1].

It would be important to learn manipulating DNA-repair processes specifically in cancer cells. This would allow targeting them more efficiently for elimination by exposing them to several DNA-damaging agents and activating multiple repair pathways. Detailed information about DNA-repair genes that are altered in various cancers may allow a more efficient stratification of patients and designing new drugs targeting DNA repair–deficient tumors.

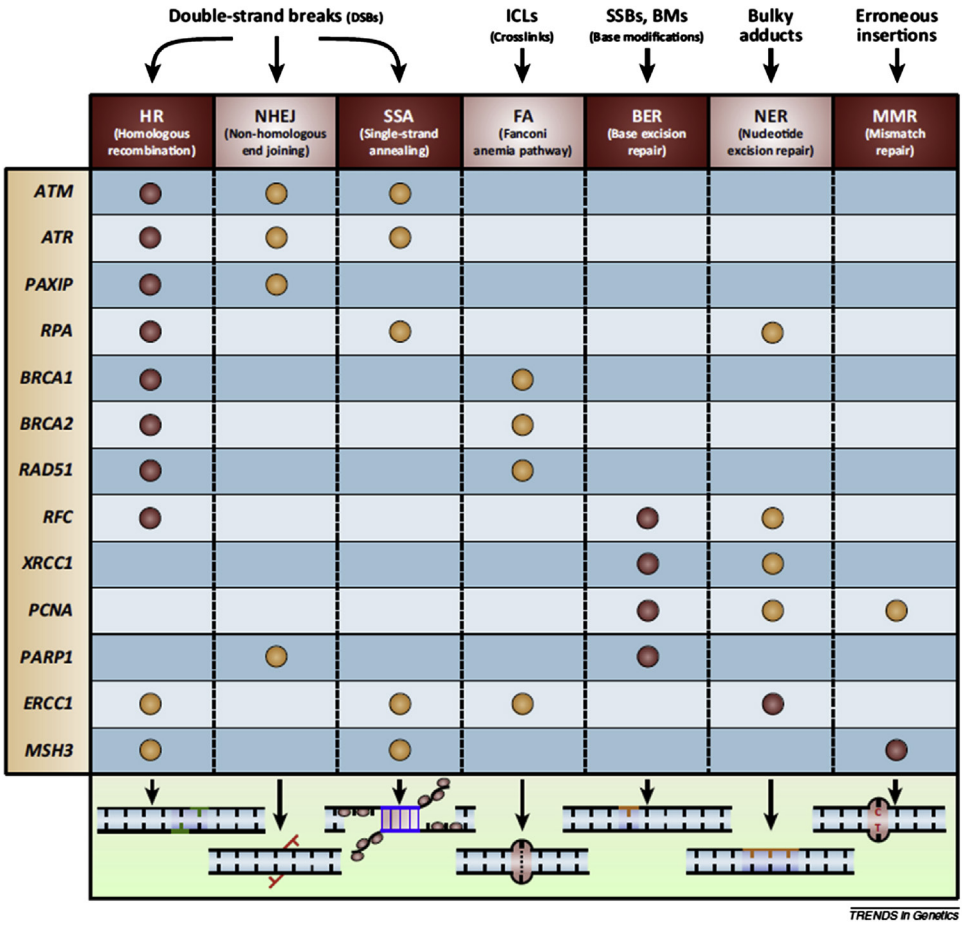

FIGURE 38.5 **Functional overlaps in DNA-repair pathways define hubs of signaling collaterals in cancer.** Multifunctional roles of proteins, which have been traditionally associated with a single DNA damage–repair pathway, are shown for six pathways. The pathway with which the gene has been primarily associated in the literature is marked by a red dot. Orange dots refer to additional roles of the protein in other pathways, which have been confirmed either by functional or biochemical experiments. These proteins are frequently affected by genomic alterations in cancer or have emerged as therapeutically amenable targets of specific kinase inhibitors. Such overlap suggests that novel druggable targets may emerge from this enhanced classification to treat DNA damage–deficient neoplastic diseases. *Reproduced from Dietlein F, Thelen L, Jokic M, Jachimowicz RD, Ivan L, Knittel G, et al. A functional cancer genomics screen identifies a druggable synthetic lethal interaction between MSH3 and PRKDC. Cancer Discov 2014;4(5):592–605; with permission.*

4. FUTURE PERSPECTIVES IN DNA-EDITING TECHNOLOGIES

The ability to delete, insert, or replace DNA sequences in the genome is critically important for both analyzing the function of various genes and reconstituting phenotypes. Gene targeting is a technique that allows sequence replacement (or deletion) by using HR. Unfortunately, for a long time, gene targeting had a low efficiency due to a low frequency of HR repair in most eukaryotes, including plants and animals. The efficiency of gene targeting can be increased dramatically when sequence-specific nicks are generated in the genome. The advent of novel engineered nucleases such as meganucleases, zinc finger nucleases (ZFNs), transcription activator–like effector-based nucleases (TALENs), and the CRISPR-Cas9 system that are able to generate such nicks in the genome allowed to develop a new field of research and technology, genome editing. See details about the CRISPRs-Cas9 system in Chapter 6 and details of using TALENs and CRISPRs for crop improvement in Chapter 13.

Some incredible uses of genome-editing technologies have been demonstrated since 2013. In fact, CRISPR was called the most important discovery of 2015 by *Science* magazine. It was used to edit the CCR5 gene encoding a cell-surface protein called C–C chemokine receptor type 5 (CCR5) in human lymphocytes; HIV is able to enter immune cells by latching onto a receptor expressed on the surface of T-helper cells and macrophages. Gene editing in a subpopulation of helper T cells followed by subculturing these cells and reintroducing them into patients allowed to achieve a significant success in HIV therapy [82].

Gantz and Bier [83] used CRISPR-Cas9 to generate a "mutagenic chain reaction" (MCR) where a mutation in one chromosome copies itself to another chromosome. The authors used *Drosophila* to show that MCR mutations efficiently spread from their chromosome of origin to the homologous chromosome, converting heterozygous mutations to homozygous mutations in most somatic and germline cells. The authors suggested that MCR technology should have broad applications; this technology can perform F1 screens for mutations in newly generated organisms, it can accelerate genetic manipulations and genome engineering due to the rapid achievement of homozygosity, it can be used to deliver transgenes to disease vectors and pest species, and it can also be used for gene therapy applications [83].

Genome editing can soon be used for the elimination of some of the carriers of dangerous diseases. For example, TEP1 gene editing by TALENs makes mosquitoes sensitive to malaria parasites, thus resulting in the elimination of infected mosquito populations in the laboratory settings [84]. It remains to be shown how these mosquitoes with edited genomes thrive in the wild.

The genome-editing technology is rather young (barely over 10 years old), with CRISPR being only a 4-year-old tool for genome editing. Hence, it is likely that many new genome-editing technologies will be discovered and developed in the nearest future. For example, a new tool based on the pentatricopeptide repeat (PPR) protein (initially suggested as a transcriptome-editing tool) [85] was proposed in 2015 to be used for genome editing [86].

The use of these editing technologies on human embryos is still debatable, so it may take years to produce the first genome-edited human. Nowadays, a great success has been achieved in genome editing, but if the CRISPR-Cas9 technology is used for in vitro or in vivo targeting in the treatment of humans, gene-targeting rates should be increased and its target specificity should be confirmed beyond doubt.

GLOSSARY

A-EJ or A-NHEJ Alternative end-joining repair of DSBs.
CO Crossover outcome of nucleolytic dissolution.
DDT The DNA damage–tolerance pathway; the translesion synthesis pathway that allows to insert the preset nucleotides base pairing with damaged nucleotides during the replication process. Usually requires switching from regular to translesion polymerases.
Genome editing Precise genetic engineering based on the ability to generate a sequence-specific nick in the genome, with the subsequent integration of desired sequences or the replacement of target sequences or the generation of random insertion/deletion (by the activity of NHEJ repair) inactivating the target gene.
HJ dissolution The mechanism of HJ processing through the BTR complex (BLM helicase–topoisomerase IIIα–RMI1–RMI2).
ICLs Interstrand cross-links; cross-links between DNA strands. Patients with Fanconi anemia are sensitive to such cross-links.
IDLs Insertion-deletion loops; DNA loops formed upon insertion/deletion of several nucleotides.
MCR Mutagenic chain reaction, the CRISPR/Cas-mediated process upon which a mutation in one chromosome copies itself to another chromosome.
NCO Non-crossover outcomes of HJ dissolution or nucleolytic dissolution.
RSS Recombination signal sequences; RSS are required for the recognition of regions involved in V(D)J recombination.

LIST OF ABBREVIATIONS

A-EJ Alternative end joining
BER Base excision repair
CLL Chronic lymphocytic leukemia
CRCs Colorectal cancers
DDT DNA-damage tolerance
DSB Double-strand break
GG-NER global-genome NER
HR Homologous recombination
ICLs Interstrand cross-links
IDLs Insertion-deletion loops
IR Ionizing radiation
MCR Mutagenic chain reaction
MMR Mismatch repair
NER Nucleotide excision repair
NHEJ Nonhomologous end joining
RSS Recombination signal sequences
SSB Single-strand break
TC-NER Transcription-coupled NER

REFERENCES

[1] Dietlein F, Thelen L, Reinhardt HC. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. Trends Genet 2014;30(8):326–39.
[2] Yi C, He C. DNA repair by reversal of DNA damage. Cold Spring Harb Perspect Biol 2013;5(1):a012575.
[3] Selby CP, Sancar A. A cryptochrome/photolyase class of enzymes with single-stranded DNA-specific photolyase activity. Proc Natl Acad Sci USA 2006;103(47):17696–700.
[4] Mishina Y, Duguid EM, He C. Direct reversal of DNA alkylation damage. Chem Rev 2006;106(2):215–32.
[5] Shankaracharya DS, Vidyarthi AS. Homology modeling and function prediction of hABH1, involving in repair of alkylation damaged DNA. Interdiscip Sci 2011;3(3):175–81.
[6] Westbye MP, Feyzi E, Aas PA, Vagbo CB, Talstad VA, Kavli B, et al. Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. J Biol Chem 2008;283(36):25046–56.
[7] Lindahl T. An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. Proc Natl Acad Sci USA 1974;71(9):3649–53.
[8] Krokan HE, Bjoras M. Base excision repair. Cold Spring Harb Perspect Biol 2013;5(4):a012583.
[9] Hegde ML, Hazra TK, Mitra S. Functions of disordered regions in mammalian early base excision repair proteins. Cell Mol Life Sci 2010;67(21):3573–87.
[10] Kanugula S, Pauly GT, Moschel RC, Pegg AE. A bifunctional DNA repair protein from *Ferroplasma acidarmanus* exhibits O^6-alkylguanine-DNA alkyltransferase and endonuclease V activities. Proc Natl Acad Sci USA 2005;102(10):3617–22.
[11] Prakash A, Doublie S. Base excision repair in the mitochondria. J Cell Biochem 2015;116(8):1490–9.
[12] Schormann N, Ricciardi R, Chattopadhyay D. Uracil-DNA glycosylases-structural and functional perspectives on an essential family of DNA repair enzymes. Protein Sci 2014;23(12):1667–85.
[13] Grogan DW. Understanding DNA repair in hyperthermophilic archaea: persistent gaps and other reasons to focus on the fork. Archaea 2015;2015:942605.
[14] Sachadyn P. Conservation and diversity of MutS proteins. Mutat Res 2010;694(1–2):20–30.
[15] Jiricny J. Postreplicative mismatch repair. Cold Spring Harb Perspect Biol 2013;5(4):a012633.
[16] Wu SY, Culligan K, Lamers M, Hays J. Dissimilar mispair-recognition spectra of *Arabidopsis* DNA-mismatch-repair proteins MSH2*MSH6 (MutSalpha) and MSH2*MSH7 (MutSgamma). Nucleic Acids Res 2003;31(20):6027–34.
[17] Harrington JM, Kolodner RD. *Saccharomyces cerevisiae* Msh2-Msh3 acts in repair of base-base mispairs. Mol Cell Biol 2007;27(18):6546–54.
[18] Raschle M, Marra G, Nystrom-Lahti M, Schar P, Jiricny J. Identification of hMutLbeta, a heterodimer of hMLH1 and hPMS1. J Biol Chem 1999;274(45):32368–75.
[19] Cannavo E, Gerrits B, Marra G, Schlapbach R, Jiricny J. Characterization of the interactome of the human MutL homologues MLH1, PMS1, and PMS2. J Biol Chem 2007;282(5):2976–86.
[20] Kadyrov FA, Dzantiev L, Constantin N, Modrich P. Endonucleolytic function of MutLalpha in human mismatch repair. Cell 2006;126(2):297–308.
[21] Shuman S, Glickman MS. Bacterial DNA repair by non-homologous end joining. Nat Rev Microbiol 2007;5(11):852–61.
[22] Pitcher RS, Brissett NC, Doherty AJ. Nonhomologous end-joining in bacteria: a microbial perspective. Annu Rev Microbiol 2007;61:259–82.
[23] Chayot R, Montagne B, Mazel D, Ricchetti M. An end-joining repair mechanism in *Escherichia coli*. Proc Natl Acad Sci USA 2010;107(5):2141–6.
[24] Weller GR, Kysela B, Roy R, Tonkin LM, Scanlan E, Della M, et al. Identification of a DNA nonhomologous end-joining complex in bacteria. Science 2002;297(5587):1686–9.
[25] Aravind L, Koonin EV. Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. Genome Res 2001;11(8):1365–74.
[26] Bartlett EJ, Brissett NC, Doherty AJ. Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates. Proc Natl Acad Sci USA 2013;110(22):E1984–91.
[27] Gu J, Lieber MR. Mechanistic flexibility as a conserved theme across 3 billion years of nonhomologous DNA end-joining. Genes Dev 2008;22(4):411–5.
[28] Aniukwu J, Glickman MS, Shuman S. The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. Genes Dev 2008;22(4):512–27.
[29] Tamura K, Adachi Y, Chiba K, Oguchi K, Takahashi H. Identification of Ku70 and Ku80 homologues in *Arabidopsis thaliana*: evidence for a role in the repair of DNA double-strand breaks. Plant J 2002;29(6):771–81.
[30] Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 2010;79:181–211.
[31] Sonoda E, Hochegger H, Saberi A, Taniguchi Y, Takeda S. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. DNA Repair (Amst) 2006;5(9–10):1021–9.
[32] Ramadan K, Maga G, Shevelev IV, Villani G, Blanco L, Hubscher U. Human DNA polymerase lambda possesses terminal deoxyribonucleotidyl transferase activity and can elongate RNA primers: implications for novel functions. J Mol Biol 2003;328(1):63–72.
[33] Daley JM, Palmbos PL, Wu D, Wilson TE. Nonhomologous end joining in yeast. Annu Rev Genet 2005;39:431–51.
[34] Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, Rehrauer WM. Biochemistry of homologous recombination in *Escherichia coli*. Microbiol Rev 1994;58(3):401–65.
[35] Hiom KDN. A repair: common approaches to fixing double-strand breaks. Curr Biol 2009;19(13):R523–5.

[36] D'Amours D, Jackson SP. The Mre11 complex: at the crossroads of DNA repair and checkpoint signalling. Nat Rev Mol Cell Biol 2002;3(5):317–27.
[37] Hopkins BB, Paull TT. The *P. furiosus* mre11/rad50 complex promotes 5′ strand resection at a DNA double-strand break. Cell 2008;135(2):250–60.
[38] Lin Y, Lin LJ, Sriratana P, Coleman K, Ha T, Spies M, et al. Engineering of functional replication protein a homologs based on insights into the evolution of oligonucleotide/oligosaccharide-binding folds. J Bacteriol 2008;190(17):5766–80.
[39] White MF. Homologous recombination in the archaea: the means justify the ends. Biochem Soc Trans 2011;39(1):15–9.
[40] Constantinou A, Davies AA, West SC. Branch migration and Holliday junction resolution catalyzed by activities from mammalian cells. Cell 2001;104(2):259–68.
[41] Matos J, West SC. Holliday junction resolution: regulation in space and time. DNA Repair (Amst) 2014;19:176–81.
[42] Wu L, Hickson ID. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. Nature 2003;426(6968):870–4.
[43] Ray JH, German J. Bloom's syndrome and EM9 cells in BrdU-containing medium exhibit similarly elevated frequencies of sister chromatid exchange but dissimilar amounts of cellular proliferation and chromosome disruption. Chromosoma 1984;90(5):383–8.
[44] Wechsler T, Newman S, West SC. Aberrant chromosome morphology in human cells defective for Holliday junction resolution. Nature 2011;471(7340):642–6.
[45] Li GM. Mechanisms and functions of DNA mismatch repair. Cell Res 2008;18(1):85–98.
[46] Guarne A, Ramon-Maiques S, Wolff EM, Ghirlando R, Hu X, Miller JH, et al. Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. EMBO J 2004;23(21):4134–45.
[47] Junop MS, Obmolova G, Rausch K, Hsieh P, Yang W. Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. Mol Cell 2001;7(1):1–12.
[48] Wang H, Hays JB. Signaling from DNA mispairs to mismatch-repair excision sites despite intervening blockades. EMBO J 2004;23(10):2126–33.
[49] Allen DJ, Makhov A, Grilley M, Taylor J, Thresher R, Modrich P, et al. MutS mediates heteroduplex loop formation by a translocation mechanism. EMBO J 1997;16(14):4467–76.
[50] Fishel R. Mismatch repair, molecular switches, and signal transduction. Genes Dev 1998;12(14):2096–101.
[51] Jiang J, Bai L, Surtees JA, Gemici Z, Wang MD, Alani E. Detection of high-affinity and sliding clamp modes for MSH2-MSH6 by single-molecule unzipping force analysis. Mol Cell 2005;20(5):771–81.
[52] Li GM. The role of mismatch repair in DNA damage-induced apoptosis. Oncol Res 1999;11(9):393–400.
[53] Stojic L, Mojas N, Cejka P, Di Pietro M, Ferrari S, Marra G, et al. Mismatch repair-dependent G2 checkpoint induced by low doses of SN1 type methylating agents requires the ATR kinase. Genes Dev 2004;18(11):1331–44.
[54] York SJ, Modrich P. Mismatch repair-dependent iterative excision at irreparable O^6-methylguanine lesions in human nuclear extracts. J Biol Chem 2006;281(32):22674–83.
[55] Yoshioka K, Yoshioka Y, Hsieh P. ATR kinase activation mediated by MutSalpha and MutLalpha in response to cytotoxic O^6-methylguanine adducts. Mol Cell 2006;22(4):501–10.
[56] Gorbunova VV, Levy AA. How plants make ends meet: DNA double-strand break repair. Trends Plant Sci 1999;4(7):263–9.
[57] Yannone SM, Roy S, Chan DW, Murphy MB, Huang S, Campisi J, et al. Werner syndrome protein is regulated and phosphorylated by DNA-dependent protein kinase. J Biol Chem 2001;276(41):38242–8.
[58] Li B, Comai L. Displacement of DNA-PKcs from DNA ends by the Werner syndrome protein. Nucleic Acids Res 2002;30(17):3653–61.
[59] Hromas R, Wray J, Lee SH, Martinez L, Farrington J, Corwin LK, et al. The human set and transposase domain protein Metnase interacts with DNA Ligase IV and enhances the efficiency and accuracy of non-homologous end-joining. DNA Repair (Amst) 2008;7(12):1927–37.
[60] Thorslund T, Ripplinger A, Hoffmann S, Wild T, Uckelmann M, Villumsen B, et al. Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. Nature 2015;527(7578):389–93.
[61] Bi X. Mechanism of DNA damage tolerance. World J Biol Chem 2015;6(3):48–56.
[62] Moldovan GL, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. Cell 2007;129(4):665–79.
[63] de Massy B. Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. Annu Rev Genet 2013;47:563–99.
[64] Nicolas A, Treco D, Schultes NP, Szostak JW. An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*. Nature 1989;338(6210):35–9.
[65] Baudat F, Imai Y, de Massy B. Meiotic recombination in mammals: localization and regulation. Nat Rev Genet 2013;14(11):794–806.
[66] Lichten M. Molecular biology. Putting the breaks on meiosis. Science 2015;350(6263):913.
[67] Lam I, Keeney S. Nonparadoxical evolutionary stability of the recombination initiation landscape in yeast. Science 2015;350(6263):932–7.
[68] Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, et al. Stable recombination hotspots in birds. Science 2015;350(6263):928–32.
[69] Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490(7418):61–70.
[70] Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013;499(7456):43–9.
[71] Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature 2011;474(7353):609–15.
[72] Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. Nature 2012;487(7406):239–43.
[73] Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet 2012;44(1):47–52.
[74] Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature 2011;475(7354):101–5.

[75] Cancer Genome Atlas Research Network, Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, et al. Integrated genomic characterization of endometrial carcinoma. Nature 2013;497(7447):67–73.
[76] Bunting SF, Callen E, Kozak ML, Kim JM, Wong N, Lopez-Contreras AJ, et al. BRCA1 functions independently of homologous recombination in DNA interstrand crosslink repair. Mol Cell 2012;46(2):125–35.
[77] Schlacher K, Wu H, Jasin M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. Cancer Cell 2012;22(1):106–16.
[78] Schlacher K, Christ N, Siaud N, Egashira A, Wu H, Jasin M. Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. Cell 2011;145(4):529–42.
[79] Hashimoto Y, Ray Chaudhuri A, Lopes M, Costanzo V. Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis. Nat Struct Mol Biol 2010;17(11):1305–11.
[80] Lord CJ, Ashworth A. Targeted therapy for cancer using PARP inhibitors. Curr Opin Pharmacol 2008;8(4):363–9.
[81] Dietlein F, Thelen L, Jokic M, Jachimowicz RD, Ivan L, Knittel G, et al. A functional cancer genomics screen identifies a druggable synthetic lethal interaction between MSH3 and PRKDC. Cancer Discov 2014;4(5):592–605.
[82] Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 2014;370(10):901–10.
[83] Gantz VM, Bier E. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. Science 2015;348(6233):442–4.
[84] Smidler AL, Terenzi O, Soichot J, Levashina EA, Marois E. Targeted mutagenesis in the malaria mosquito using TALE nucleases. PLoS One 2013;8(8):e74511.
[85] Manna S. An overview of pentatricopeptide repeat proteins and their applications. Biochimie 2015;113:93–9.
[86] Yagi Y, Shirakawa M, Nakamura T. The challenges faced by EditForce Inc., to go beyond genome editing. Nature 2015:12. Sponsor feature.