Genome maintenance mechanisms for preventing cancer

Jan H. J. Hoeijmakers

MGC Department of Cell Biology and Genetics, Centre for Biomedical Genetics, Erasmus University, PO Box 1738, 3000DR Rotterdam, The Netherlands (e-mail: Hoeijmakers@gen.fgg.eur.nl)

The early notion that cancer is caused by mutations in genes critical for the control of cell growth implied that genome stability is important for preventing oncogenesis. During the past decade, knowledge about the mechanisms by which genes erode and the molecular machinery designed to counteract this time-dependent genetic degeneration has increased markedly. At the same time, it has become apparent that inherited or acquired deficiencies in genome maintenance systems contribute significantly to the onset of cancer. This review summarizes the main DNA caretaking systems and their impact on genome stability and carcinogenesis.

Cancer is a disease of our genes. Over time, DNA accumulates changes that activate proto-oncogenes and inactivate tumour-suppressor genes. The genetic instability driving tumorigenesis is fuelled by DNA damage and by errors made by the DNA machinery. However, 'spontaneous' mutations are insufficient to explain the lifetime cancer risk[1]. Indeed, numerous links have been identified between oncogenesis and acquired or inherited faulty genome guardians that cause a 'mutator' phenotype, highlighting the key role of DNA protection systems in tumour prevention. Here I focus on the main DNA maintenance mechanisms operating in mammals — nucleotide- and base-excision repair, homologous recombination, end joining, mismatch repair and telomere metabolism — and their relevance for cancer.

A plethora of damages in DNA

The physicochemical constitution of our genes does not guarantee life-long stability or proper function. A perplexing diversity of lesions arises in DNA from three main causes. First, environmental agents such as the ultraviolet (UV) component of sunlight, ionizing radiation and numerous genotoxic chemicals cause alterations in DNA structure, which, if left unrepaired, may lead to mutations that enhance cancer risk. A pronounced example is exposure to genotoxic compounds in cigarette smoke, which are responsible for the most frequent cancer in Western men. Second, (by)products of normal cellular metabolism constitute a permanent enemy to DNA integrity from within. These include reactive oxygen species (superoxide anions, hydroxyl radicals and hydrogen peroxide) derived from oxidative respiration and products of lipid peroxidation. Over 100 oxidative modifications have been identified in DNA[2]. Evolution has invested significantly in reducing the price of its own metabolism by implementing an intricate antioxidant defence system composed of enzymatic (superoxide dismutase, catalase, glutathione peroxidase and peroxyredoxins) and low-molecular-mass scavengers (such as glutathione)[3]. Finally, some chemical bonds in DNA tend to spontaneously disintegrate under physiological conditions. Hydrolysis of nucleotide residues leaves non-instructive abasic sites. Spontaneous or induced deamination of cytosine, adenine, guanine or 5-methylcytosine converts these bases to the miscoding uracil, hypoxanthine, xanthine and thymine, respectively[4]. Figure 1a summarizes some of the most common types of DNA damage and their sources.

The consequences of DNA injury

The outcome of DNA damage is diverse and generally adverse (Fig. 1b). Acute effects arise from disturbed DNA metabolism, triggering cell-cycle arrest or cell death. Long-term effects result from irreversible mutations contributing to oncogenesis.

Many lesions block transcription, which in effect inactivates every gene containing damage on the transcribed strand — an outcome directly related to gene length. This has elicited the development of a dedicated repair system, transcription-coupled repair (TCR), which displaces or removes the stalled RNA polymerase and assures high-priority repair. Transcriptional stress, arising from persistent blockage of RNA synthesis, constitutes an efficient trigger for p53-dependent apoptosis (see ref. 5 and the article in this issue by Evan and Vousden, pages 342–348), which may be a significant anti-cancer mechanism.

Lesions also interfere with DNA replication. Recently, a growing class of DNA polymerases, numbered ζ to κ, was discovered which seems devoted specifically to overcoming damage-induced replicational stress[6,7]. These special polymerases take over temporarily from the blocked replicative DNA polymerase-δ/ε (polδ/ε), and possibly from polα (Fig. 2, follow upper strand). They have more flexible base-pairing properties permitting translesion synthesis, with each polymerase probably designed for a specific category of injury. The number of polymerases preferring damaged templates currently exceeds that for undamaged DNA, which illustrates the magnitude of the problem. But this solution generally comes at the expense of a higher error rate. In fact, this process is responsible for most of damage-induced point mutations[8] and is thus particularly relevant for oncogenesis. Nevertheless, translesion polymerases still protect the genome. For instance, inherited defects in pol-η, which specializes in relatively error-free bypassing of UV-induced cyclobutane pyrimidine dimers, cause the variant form of the skin cancer-prone disorder xeroderma pigmentosum[9,10]. In the yeast *Saccharomyces cerevisiae*, a second, probably even more important pathway exists that allows error-free bypass of lesions[8]. This mechanism is based on reinitiation of DNA

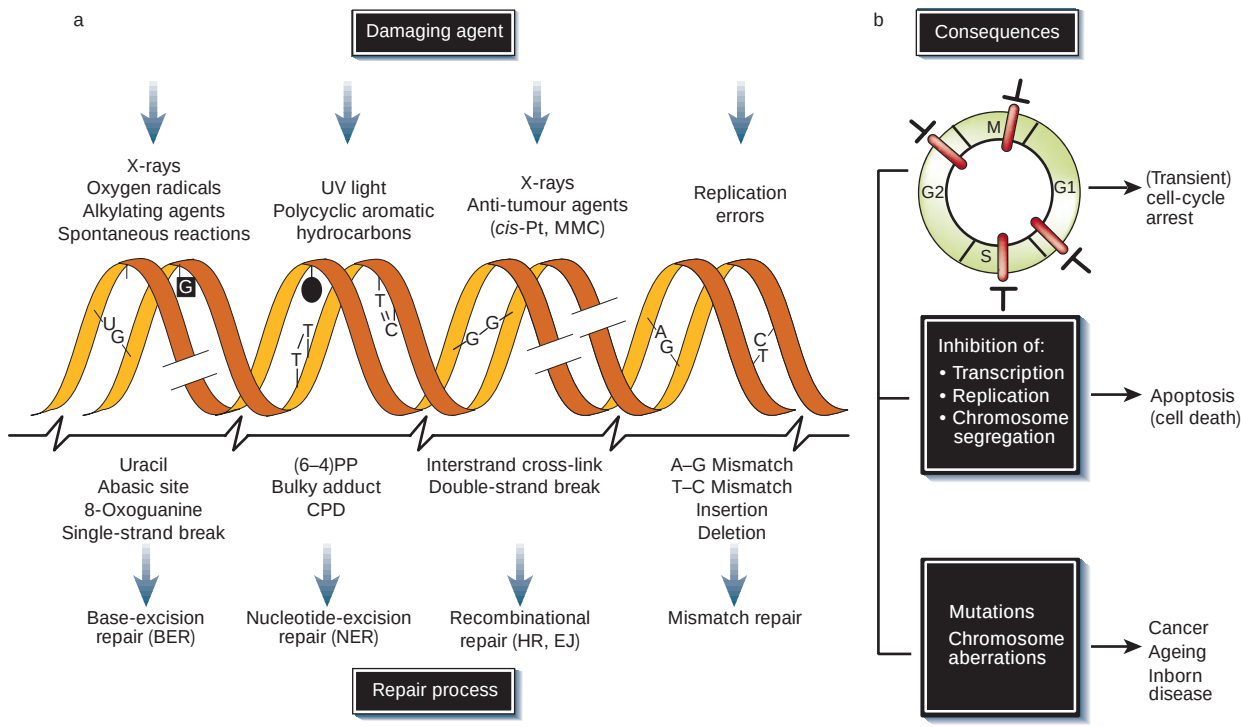


Figure 1 DNA damage, repair mechanisms and consequences. **a**, Common DNA damaging agents (top); examples of DNA lesions induced by these agents (middle); and most relevant DNA repair mechanism responsible for the removal of the lesions (bottom). **b**, Acute effects of DNA damage on cell-cycle progression, leading to transient arrest in the G1, S, G2 and M phases (top), and on DNA metabolism (middle). Long-term consequences of DNA injury (bottom) include permanent changes in the DNA sequence (point mutations affecting single genes or chromosome aberrations which may involve multiple genes) and their biological effects. Abbreviations: *cis*-Pt and MMC, cisplatin and mitomycin C, respectively (both DNA-crosslinking agents); (6–4)PP and CPD, 6–4 photoproduct and cyclobutane pyrimidine dimer, respectively (both induced by UV light); BER and NER, base- and nucleotide-excision repair, respectively; HR, homologous recombination; EJ, end joining.

replication downstream of the blocking injury. The resulting gap is filled in by recombinational replication, using the newly synthesized complementary strand as a template and ignoring the original lesion-containing one (Fig. 2, follow lower strand). Yeast proteins implicated in this process, such as the Ubc13/Mms2 complex, are conserved all the way to mammals. Thus, this largely unexplored system undoubtedly exists in humans and may be important in carcinogenesis. The endpoint of both of these pathways is that damage persists and — when unrepaired — will cause similar problems in subsequent rounds of replication. This is particularly relevant for damage that is not efficiently recognized by any mammalian repair process, such as cyclobutane pyrimidine dimers.

Double-strand DNA breaks (DSBs) induced by X-rays, chemicals or during replication of single-strand breaks (SSBs) and presumably during repair of interstrand crosslinks are particularly relevant for the recombination machinery. Cells with specialized DNA recombination activities, such as B- and T-cells, may be very sensitive to DSBs when they are rearranging their immunoglobin or T-cell-receptor genes. This explains the frequent involvement of these genetic loci in oncogenic translocations in leukaemia and lymphomas and the preferential induction of these cancers by ionizing irradiation. DSBs also pose problems during mitosis, as intact chromosomes are a prerequisite for proper chromosome segregation during cell division. Thus, these lesions frequently induce various sorts of chromosomal aberrations, including aneuploidy, deletions (loss of heterozygosity) and chromosomal translocations — events which are all intimately associated with carcinogenesis.

The cell-cycle machinery somehow senses genome injury and arrests at specific checkpoints in G1, S, G2 and M to allow repair of lesions before they are converted into permanent mutations (reviewed in ref. 11). Lesion detection may occur by blocked transcription, replication or specialized sensors. When damage is too significant, a cell may opt for the ultimate mode of rescue by initiating apoptosis at the expense of a whole cell (see review by Evan and Vousden, pages 342–348).

DNA damage repair systems

In view of the plethora of types of lesions, no single repair process can cope with all kinds of damage. Instead, evolution has moulded a tapestry of sophisticated, interwoven DNA repair systems that as a whole cover most (but not all) of the insults inflicted on a cell's vital genetic information. Inherited defects in any of these pathways in general predisposes to malignancy (Table 1). Because the problem of DNA damage has existed *ab initio*, DNA repair systems must have arisen early in evolution. This explains why all known repair pathways are highly conserved (usually across the pro/eukaryotic evolutionary border). At least four main, partly overlapping damage repair pathways operate in mammals — nucleotide-excision repair (NER), base-excision repair (BER), homologous recombination and end joining[12,13]. The division of tasks between them can be roughly defined as follows (see also Fig. 1a).

NER deals with the wide class of helix-distorting lesions that interfere with base pairing and generally obstruct transcription and normal replication. Small chemical alterations of bases are targeted by BER. These lesions may or may not impede transcription and replication, although they frequently miscode. BER is therefore particularly relevant for preventing mutagenesis. Most NER lesions arise from exogenous sources (except for some oxidative lesions), whereas BER is mostly, but not exclusively, concerned with damage of endogenous origin. Lesions for these two repair processes affect only

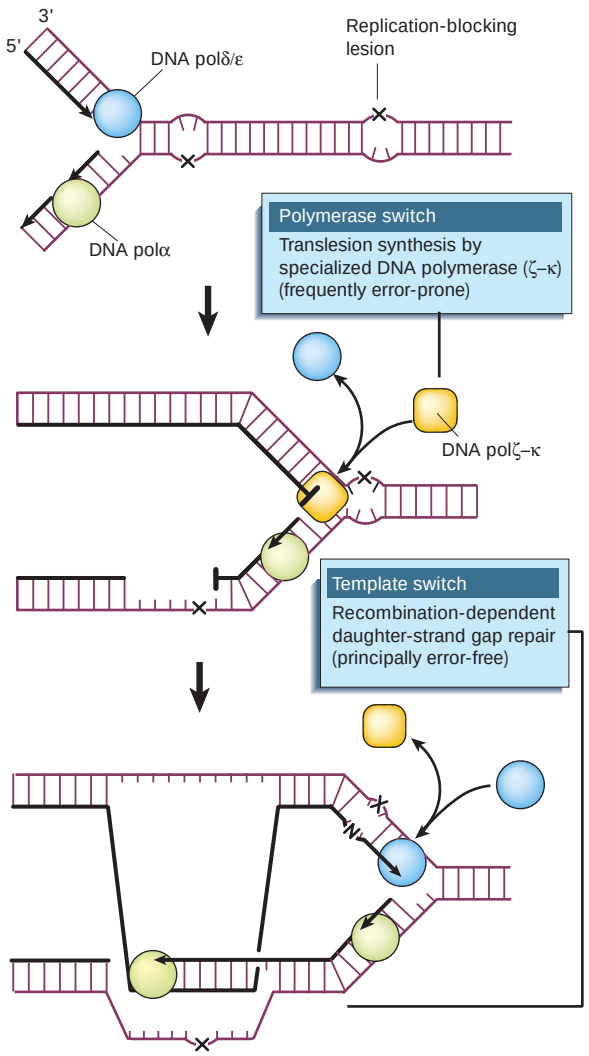


Figure 2 Mechanisms of replicational bypass of DNA lesions. Lesions in the DNA template (indicated by an 'X') may be bypassed by the replication apparatus in two different ways: DNA polymerase switch (upper strand) and template switch (lower strand). In the DNA polymerase switch, the regular DNA polymerase (in this case polδ/ε, carrying out leading-strand synthesis) is arrested at the site of the damage. A specific translesion polymerase (polζ–κ), or a combination of these polymerases, takes over synthesis to bypass the injured site, after which the regular polymerase continues. This process can be highly error-prone. In the template switch (model), the regular DNA polymerase (in this case polα, responsible for lagging-strand synthesis) is arrested at a damaged site. The resulting gap in the newly synthesized strand is filled in using the undamaged, newly synthesized leading strand via recombinational strand exchange (or alternatively by fork regression and annealing of the new strand, not shown). This mechanism may involve specific factors as well as members of the RAD52 family implicated in homologous recombination repair. In principle, this mode of lesion bypass is error-free. Note that in both of these processes the lesion remains and that the two scenarios may apply to both strands.

one of the DNA strands. In a 'cut-and-patch'-type reaction, the injury (with or without some flanking sequences) is taken out and the resulting single-stranded gap is filled in using the intact complementary strand as template.

DSBs are more problematic, as both strands are affected. To properly heal such breaks the cell has to know which ends belong together, a difficult task given the size of the mammalian genome. Two pathways, homologous recombination and end joining (and presumably additional back-up systems), were developed for solving

Table 1 Human syndromes with defective genome maintenance

Syndrome	Affected maintenance mechanism	Main type of genome instability	Major cancer predisposition
Xeroderma pigmentosum	NER (±TCR)	Point mutations	UV-induced skin cancer
Cockayne syndrome	TCR	Point mutations	None*
Trichothiodystrophy	NER / TCR	Point mutations	None*
Ataxia telangiectasia (AT)	DSB response/repair	Chromosome aberrations	Lymphomas
AT-like disorder	DSB response/repair	Chromosome aberrations	Lymphomas
Nijmegen breakage syndrome	DSB response/repair	Chromosome aberrations	Lymphomas
BRCA 1/BRCA2	HR	Chromosome aberrations	Breast (ovarian) cancer
Werner syndrome	HR?/TLS?	Chromosome aberrations	Various cancers
Bloom syndrome	HR?	Chromosome aberrations (SCE↑)	Leukaemia, lymphoma, others
Rothmund–Thomson syndrome	HR?	Chromosome aberrations	Osteosarcoma
Ligase IV deficiency†	EJ	Recombination fidelity	Leukaemia(?)
HNPCC	MMR	Point mutations	Colorectal cancer
Xeroderma pigmentosum variant	TLS‡	Point mutations	UV-induced skin cancer

*Defect in transcription-coupled repair triggers apoptosis, which may protect against UV-induced cancer.
†One patient with leukaemia and radiosensitivity described with active-site mutation in ligase IV.
‡Specific defect in relatively error-free bypass replication of UV-induced cyclobutane pyrimidine dimers.
Abbreviations: BER, base-excision repair; DSB, double-strand break; HNPCC, hereditary non-polyposis colorectal cancer; HR, homologous recombination; MMR, mismatch repair; NER, nucleotide-excision repair; SCE, sister-chromatid exchange; TCR, transcription-coupled repair; TLS, translesion synthesis.

the DSB problem. Homologous recombination seems to dominate in S and G2 when the DNA is replicated, providing a pristine second copy of the sequence (sister chromatid) for aligning the breaks. In contrast, the less-accurate end joining is most relevant in the G1 phase of the cell cycle, when a second copy is not available[14].

Finally, some single repair proteins directly revert certain injuries, such as O^6-methylguanine methyltransferase, which removes O^6-methyl guanine. This highly mutagenic lesion permits base pairing with both C or T and is capable of fooling the mismatch repair system into triggering futile rounds of mismatch removal and subsequent reincorporation of the erroneous base by repair replication. The dedicated methyl transferase specifically removes the non-native methyl group from the guanine residue and transfers it to an internal cysteine. However, in doing so, the protein irreversibly inactivates itself[13]. This illustrates how in some situations an entire protein may be sacrificed for the repair of a single damaged base. Below I describe the four main multi-step damage repair processes in mammals and their relevance for preventing cancer.

Nucleotide-excision repair and transcription-coupled repair

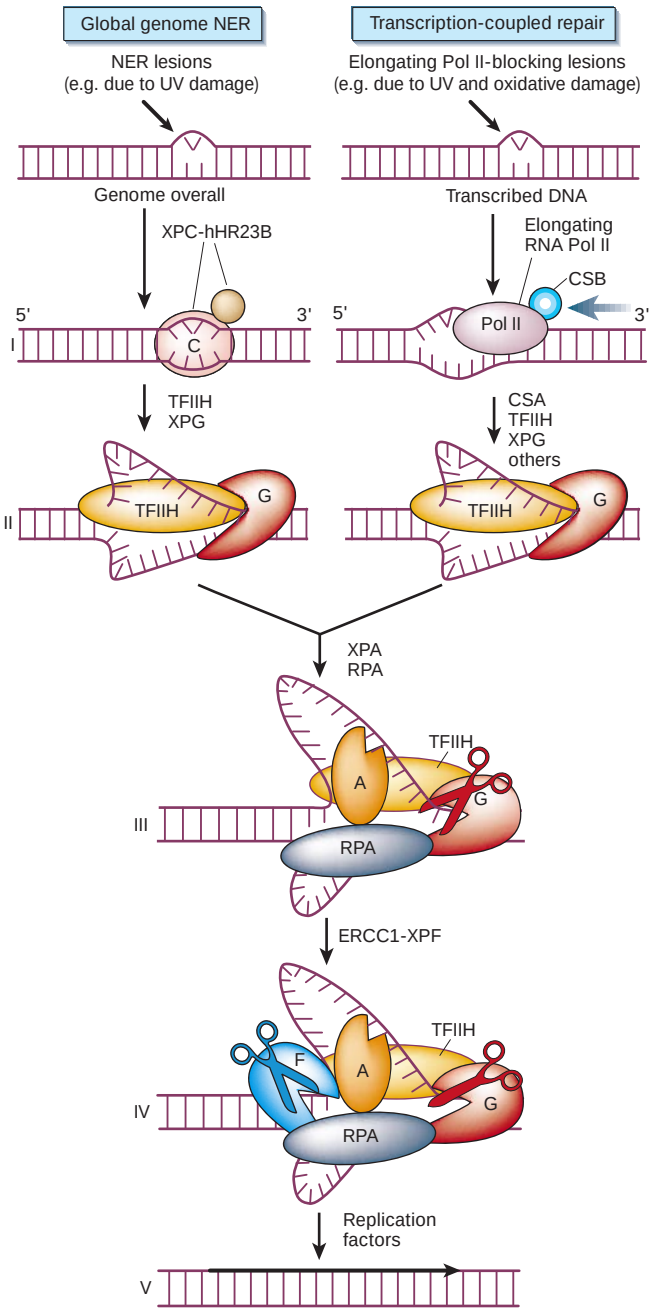
Of all repair systems, NER is the most versatile in terms of lesion recognition. Two NER subpathways exist with partly distinct substrate specificity: global genome NER (GG-NER) surveys the entire genome for distorting injury, and transcription-coupled repair (TCR) focuses on damage that blocks elongating RNA polymerases[15]. Box 1 presents the most likely mechanisms of action for these pathways (and see refs 16, 17).

NER, TCR and cancer

At least three syndromes are associated with inborn defects in NER (Table 1): xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy (TTD), all characterized by exquisite sun sensitivity[18,19]. The prototype repair disorder, xeroderma pigmentosum,

Box 1
Model for mechanism of global genome nucleotide-excision repair and transcription-coupled repair

The GG-NER-specific complex XPC-hHR23B screens first on the basis of disrupted base pairing[53], instead of lesions per se. This explains why mildly distorting injury such as cyclobutane pyrimidine dimers are poorly repaired[54]. In TCR, the ability of a lesion (whether of the NER- or BER-type) to block RNA polymerase seems critical (stage I in the figure opposite). The stalled polymerase must be displaced to make the injury accessible for repair[55], and this requires at least two TCR-specific factors: CSB and CSA. The subsequent stages of GG-NER and TCR may be identical. The XPB and XPD helicases of the multi-subunit transcription factor TFIIH open ~30 base pairs of DNA around the damage (II). XPA probably confirms the presence of damage by probing for abnormal backbone structure[56], and when absent aborts NER[53]. The single-stranded-binding protein RPA (replication protein A) stabilizes the open intermediate by binding to the undamaged strand (III). The use of subsequent factors, each with limited capacity for lesion detection *in toto*, still allows very high damage specificity[57]. The endonuclease duo of the NER team, XPG and ERCC1/XPF, respectively cleave 3′ and 5′ of the borders of the opened stretch only in the damaged strand, generating a 24–32-base oligonucleotide containing the injury (IV). The regular DNA replication machinery then completes the repair by filling the gap (V). In total, 25 or more proteins participate in NER. *In vivo* studies indicate that the NER machinery is assembled in a step-wise fashion from individual components at the site of a lesion. After a single repair event (which takes several minutes) the entire complex is disassembled again[58].

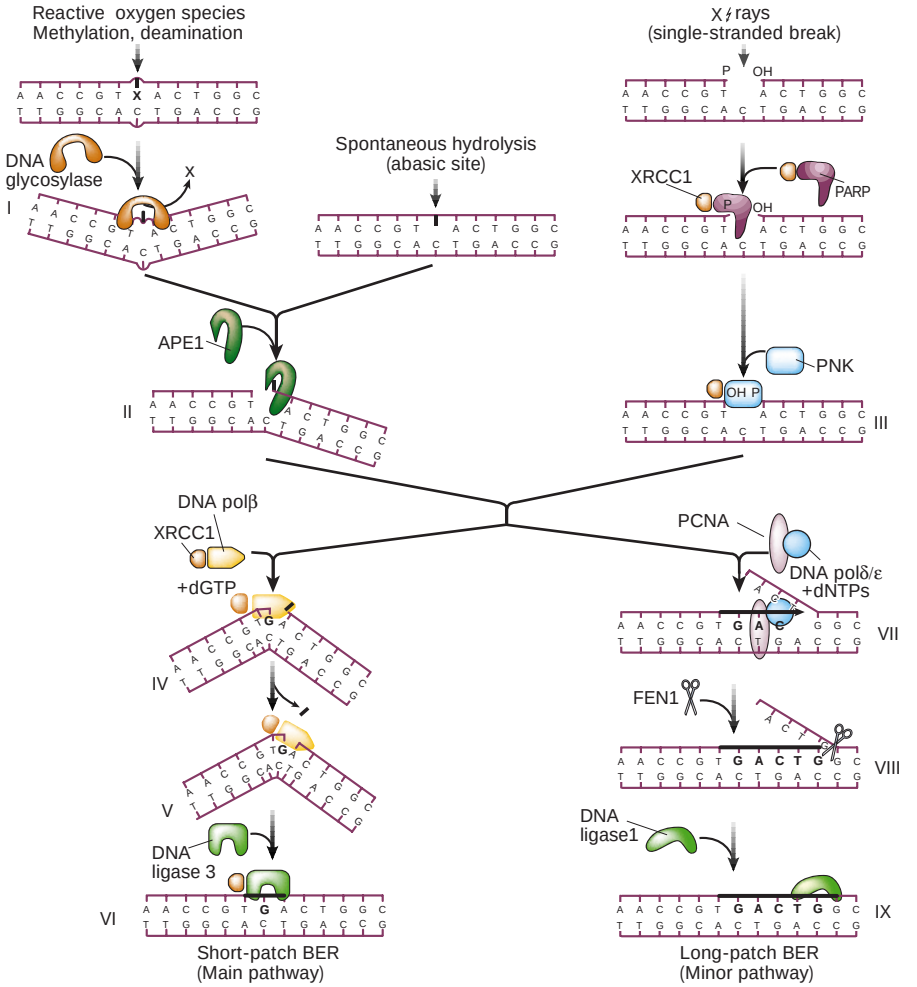


exhibits a dramatic >1000-fold incidence of sun-induced skin cancer. Frequency of internal tumours is modestly elevated and accelerated neurodegeneration is often noted. The disorder arises from mutations in one of seven genes (*XPA–XPG*). Cockayne syndrome, caused by mutation in the *CSA* or *CSB* genes, is a TCR-specific disorder that is remarkably dissimilar from xeroderma pigmentosum. No predisposition to cancer is observed, which may be explained by the fact that the TCR defect causes Cockayne syndrome cells to be particularly sensitive to lesion-induced apoptosis, thereby protecting against tumorigenesis. Physical and neurological development are impaired, resulting in dwarphism and dysmyelination. The syndrome includes features of premature ageing, which may be related to the increased trigger for apoptosis induced by transcriptional arrest from endogenous lesions in combination with the TCR defect . TTD is a condition sharing many symptoms with Cockayne syndrome, but with the additional hallmarks of brittle hair, nails and scaly skin. Mutations in the *XPD* or *XPB* genes can give rise to all three diseases. This puzzle is explained by the fact that, as subunits of TFIIH, XPB and XPD have dual functions: NER and transcription initiation. Mutations may not only compromise NER, but also affect transcription, causing developmental delay and reduced expression of the matrix proteins that causes brittle hair and scaly skin[20].

For almost all NER factors, mouse mutants have been generated[21]. Overall, the NER defect is accurately preserved, although cancer predisposition is more pronounced and neurological complications are milder in mice. Moreover, mice exhibit features of premature ageing.

Box 2
Mechanism for base-excision repair

A battery of glycosylases, each dealing with a relatively narrow, partially overlapping spectrum of lesions, feeds into a core reaction. Glycosylases flip the suspected base out of the helix by DNA backbone compression to accommodate it in an internal cavity of the protein. Inside the protein, the damaged base is cleaved from the sugar-phosphate backbone (stage I in the figure). The resulting abasic site can also occur spontaneously by hydrolysis. The core BER reaction is initiated by strand incision at the abasic site by the APE1 endonuclease (II). Poly(ADP-ribose) polymerase (PARP), which binds to and is activated by DNA strand breaks, and the recently identified polynucleotide kinase (PNK)[59] may be important when BER is initiated from a SSB to protect and trim the ends for repair synthesis (III). In mammals, the so-called short-patch repair is the dominant mode for the remainder of the reaction. DNA polβ performs a one-nucleotide gap-filling reaction (IV) and removes the 5′-terminal baseless sugar residue via its lyase activity (V); this is then followed by sealing of the remaining nick by the XRCC1–ligase3 complex (VI). The XRCC1 scaffold protein interacts with most of the above BER core components and may therefore be instrumental in protein exchange. The long-patch repair mode involves DNA polβ, polδ/ε and proliferating cell nuclear antigen (PCNA) for repair synthesis (2–10 bases) as well as the FEN1 endonuclease to remove the displaced DNA flap and DNA ligase 1 for sealing (VII–IX). The above BER reaction operates across the genome. However, some BER lesions block transcription, and in this case the problem is dealt with by the TCR pathway described above, including TFIIH, XPG (which also stimulates some of the glycosylases) and probably the remainder of the core NER apparatus.



Base-excision repair

BER is the main guardian against damage due to cellular metabolism, including that resulting from reactive oxygen species, methylation, deamination and hydroxylation. The molecular mechanism[13] has been resolved to the tertiary structure of all core components[22–24] and is explained in Box 2.

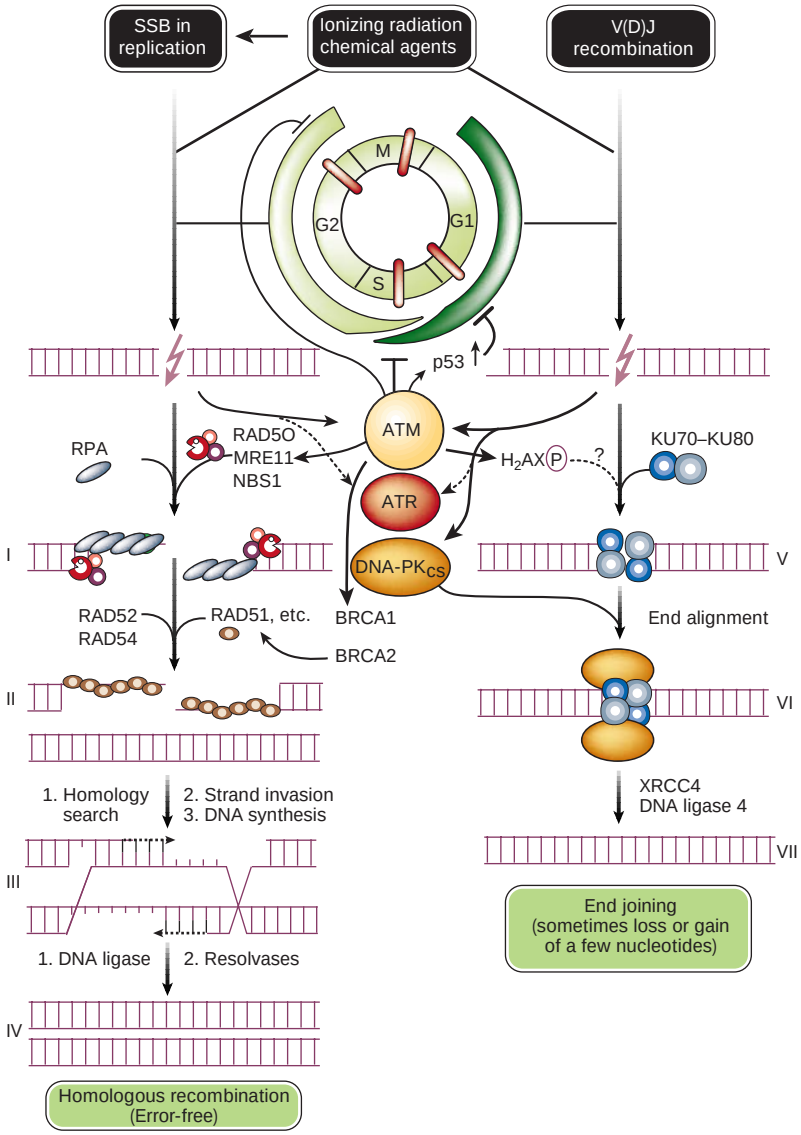
BER and cancer

No human disorders caused by inherited BER deficiencies have been identified. Mouse models generated in recent years may provide an explanation: knockout of individual glycosylases does not cause an overt phenotype, which is explained by partial redundancy between different glycosylases[13,25] and overlap with TCR. In fact, even a number of double mutants show only mild phenotypes, although mutagenesis and cancer susceptibility are probably increased. But inactivation of BER core proteins induces embryonic lethality, highlighting the vital importance of the process as a whole. This might be due to the contribution of spontaneously occurring abasic sites and SSBs that directly feed into the BER core reaction (Box 2) and/or to the generation of reaction intermediates by the glycosylases that cannot be further processed[13,25]. Interestingly, specific polymorphisms in XRCC1 seem associated with lung and other cancers[26].

DSB repair: homologous recombination and end joining

DSBs arise from ionizing radiation or X-rays, free radicals, chemicals and during replication of a SSB. After DSB detection, a complex cascade of reactions is triggered aimed at halting the cell-cycle machinery and recruiting repair factors[11,27] (Fig. 5). One of the early initiators is the ataxia telangiectasia mutated (ATM) protein kinase, which is defective in the cancer-prone, X-ray-sensitive syndrome ataxia telangiectasia[28]. Arrest in G1 is mediated via p53. Another early event, which depends on the giant protein-kinases ATM, ATR (ataxia telangiectasia related) and DNA-PK$_{cs}$, is phosphorylation of histone H2AX in the DNA domain next to the DSB over a megadalton distance[29]. This may provide a local chromatin state required for the complex repair reactions or for recruiting repair proteins. Homologous recombination and end joining are the main repair modes. When, after replication, a second identical DNA copy is available, homologous recombination seems to be preferred; otherwise cells

Box 3
Mechanism of homologous recombination and end joining



A tentative scenario for the homologous-recombination reaction is depicted in the left panel of the figure. To promote strand invasion into homologous sequences, the 5′–3′ exonuclease activity of the RAD50/MRE11/NBS1 complex (also a substrate for ATM phosphorylation) exposes both 3′ ends[30] (I). RPA facilitates assembly of a RAD51 nucleoprotein filament that probably includes RAD51-related proteins XRCC2, XRCC3, RAD51B, C and D. RAD52 stimulates filament assembly (II). RAD51 has, like its *Escherichia coli* RecA counterpart, the ability to exchange the single strand with the same sequence from a double-stranded DNA molecule. Correct positioning of the sister chromatids by cohesins probably facilitates the identification of a homologous sequence. A candidate for the complex chromatin transactions associated with these DNA gymnastics is RAD54, a member of the SWI/SNF family of DNA-dependent ATPases. After identification of the identical sister chromatid sequence, the intact double-stranded copy is used as a template to properly heal the broken ends by DNA synthesis (III). Finally, the so-called Holliday-junctions are resolved by resolvases[27,33,60] (IV). Homologous recombination involves the simultaneous action of large numbers of the same molecules, which are found to be concentrated in radiation-induced nuclear foci. These depend on, and also include, the BRCA1 and BRCA2 proteins[36]. Recent evidence implicates BRCA2 directly or indirectly in nuclear translocation of RAD51 (ref. 61).

Cells in G1 have only the homologous chromosome for recombination repair. However, this may be difficult to find in the complex genome. Moreover, it is potentially dangerous as a template for repair as it may lead to homozygosity for recessive mutations. As an alternative, the end-joining reaction simply links ends of a DSB together, without any template, using the end-binding KU70/80 complex and DNA-PK$_{cs}$, followed by ligation by XRCC4–ligase4 (reviewed by 27,33; see the right panel of the figure, stages V–VII). The function of KU70/80 might involve end protection and approximating the ends, in addition to a signalling function by DNA-PK$_{cs}$. End joining may be further facilitated when the ends are still held together through nucleosomes or other structures. End joining is sometimes associated with gain or loss of a few nucleotides if internal microhomologies are used for annealing before sealing. This implies the involvement of DNA polymerases and/or nucleases. Note that the KU complex is also involved in telomere metabolism[27,62].

rely on end joining, which is more error-prone. Their presumed mechanisms are explained in Box 3.

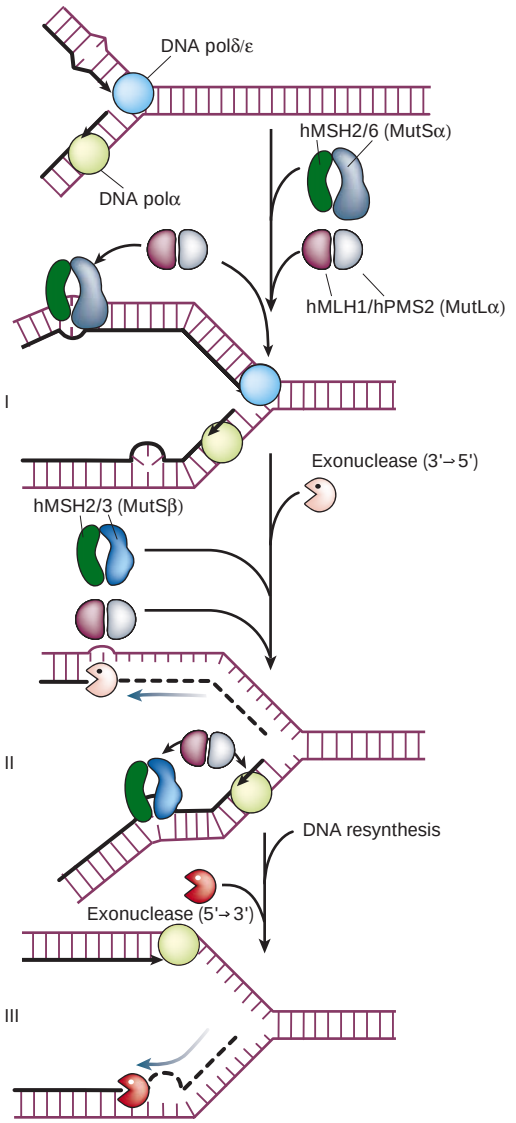
DSB repair and cancer

Besides ataxia telangiectasia, mutations in *MRE11* give rise to an ataxia telangiectasia-like disorder, whereas defects in *NBS1* are associated with the Nijmegen breakage syndrome (NBS)[30] (Table 1). All three conditions display cancer predisposition (particularly lymphomas), immunodeficiency, hypersensitivity to X-rays and chromosomal instability. Ataxia telangiectasia is additionally characterized by ataxia, cerebellar degeneration and ocular telangiectasia, whereas the cardinal symptoms of NBS are microcephaly and growth retardation[28,31]. Inherited defects in *BRCA1* and *BRCA2* strongly predispose to breast cancer. In addition, cancer-prone chromosomal-instability conditions such as Werner, Bloom and Rothmund Thomson syndrome, which all involve RecQ-like helicases, might carry defects in homologous recombination (Table 1). Inborn defects in the ligase IV component of end joining have been described for a single patient with leukaemia[32].

Except for ATM, mice with null mutations in the above homologous-recombination factors tend to suffer from early embryonic lethality or in some cases display a mild phenotype (Rad52, Rad54), presumably because of functional redundancy[33]. Lethality is preceded by gross chromosomal rearrangements, perhaps because endogenous lesions such as SSBs are converted to DSBs upon replication. The viable phenotype of mice and patients with ataxia telangiectasia may be due to partial functional overlap with ATR and DNA-PK$_{cs}$. Recently, double-mutant mice of *ATM* and *DNAPK$_{cs}$* were indeed

Box 4
Model for mismatch repair

Mammalian MMR involves multi-member families of the *E. coli* prototype factors MutS and MutL[63,64]. Heterodimers of hMSH2/6 (called hMutSα) focus on mismatches and single-base loops (stage I in the figure below, upper strand), whereas hMSH2/3 dimers (hMutSβ) recognize insertion/deletion loops (II, lower strand). Heterodimeric complexes of the hMutL-like proteins hMLH1/hPMS2 (hMutLα) and hMLH1/hPMS1 (hMutLβ) interact with MSH complexes and replication factors. Strand discrimination may be based on contact with the nearby replication machinery. A number of proteins are implicated in the excision of the new strand past the mismatch and resynthesis steps, including polδ/ε, RPA, PCNA, RFC, exonuclease 1, and endonuclease FEN1 (II, III). MMR components also interact functionally with NER and recombination. Recent crystallographic studies have revealed that a MutS dimer detects the structural instability of a heteroduplex by kinking the DNA at the site of the mismatch, which is facilitated when base pairing is affected[65,66]. However, DNA damage with similar characteristics, such as that caused by alkylating agents and intercalators, may fool MutS, triggering erroneous or futile MMR. Intact MMR thus confers sensitivity, and as several of these agents are used in chemotherapy, tumours may become resistant to them on the basis of selection for defective MMR, so confounding therapeutic strategies[67]



found to be lethal[34]. Inactivation of *ATR* by itself is inviable already at the blastocyst stage. Inactivation of *BRCA1* and *BRCA2* in mice is also embryonically lethal; cell lines display defects in homologous recombination[35–37].

The severe phenotype of the mouse mutants and the highly cancer-prone human syndromes highlight the importance of homologous recombination. Mouse KU mutants display sensitivity to agents that lead to breaks in DNA, and have immunological problems because the KU proteins are involved in V(D)J recombination of antibody gene sequences. In addition, these mutants display poor development, several features of premature ageing and increased apoptosis of postmitotic neurons in the developing brain. Mice with defects in DNA-PK$_{cs}$ (SCID mice) display a similar but generally milder phenotype. In contrast, XRCC4- and ligase IV-knockout mice seem more severe, with late embryonic lethality resulting from massive ATM- and p53-dependent neuronal apoptosis[33,38].

Mismatch repair

Specific sequence motifs comprised of dinucleotide repeats are unstable in some human cancers[39]. This phenotype of 'microsatellite instability' is caused by defects in MMR in the hereditary non-polyposis colorectal cancer (HNPCC) and in a variety of sporadic cancers. MMR removes nucleotides mispaired by DNA polymerases and insertion/deletion loops (ranging from one to ten or more bases) that result from slippage during replication of repetitive sequences or during recombination. Defects in this system dramatically increase mutation rates, fuelling the process of oncogenesis. Four principal steps in MMR can be delineated: (1) mismatch recognition; (2) recruitment of additional MMR factors; (3) search for a signal that identifies the wrong (newly synthesized) strand, followed by degradation past the mismatch; and (4) resynthesis of the excised tract. A tentative model is depicted in Box 4.

MMR and cancer

Germline mutations in *hMLH1*and *hMSH2* together account for approximately half of all HNPCC patients, with *hMLH1* being responsible for most (~60%) of these cases. Defects in *hMSH6* cause late-onset atypical HNPCC. No *hMSH3* mutations have been reported. This is consistent with the notion that loss of hMLH1 and hMSH2 is associated with complete inactivation of MMR, whereas defects in the other proteins causes only a partial MMR deficiency. Mutations in *hPMS2* and *hPMS1* have been reported only in very few cases[40], implying that other factors have still to be identified. The reason why these MMR defects cause predominantly cancers of the colon, endometrium and ovary is still unclear.

Surprisingly, homozygous MMR deficiencies in mice are compatible with normal (albeit cancer-prone) development[41]. Mutants exhibit the expected molecular defects in terms of mutagenesis based on the role of the corresponding protein in MMR. Null mutations in the key genes *Mlh1* and *Msh2* predispose the mice mainly to lymphomas, although gastrointestinal tumorigenesis is also enhanced. This phenotype is similar to the combined *Msh3/6* defect, whereas a single *Msh3* or *Msh6* mutation induces cancer at a later age. *Pms2*$^{-/-}$ mice display mainly haematological malignancies, but no intestinal neoplasias. In addition, *Mlh1*$^{-/-}$ males are sterile owing to the occurrence of apoptosis during meiosis; this occurs secondary to the premature separation of chromosomes, which suggests a role of MLH in meiotic recombination. Null alleles of the MutS homologues *Msh4* and *Msh5* display infertility for both sexes, indicating unique functions of these genes in gametogenesis.

The telomere-division limitator

Telomeres constitute the caps of chromosome ends, and function as a buffer to prevent loss of important genomic sequence during replication. In humans they consist of a 5–15-kilobase repeated array of the sequence TTAGGG bound by a specific set of proteins. DNA replication proceeding in the 5′→3′ direction needs an RNA primer

before it can initiate. Therefore, it leaves a terminal stretch of unreplicated DNA at the 5′-end of linear molecules. This leads to loss of a number of the terminal telomeric repeats with every S-phase, shortening telomeres by about 100 base pairs per cell division. In the germ line and in some specific tissues, telomere length is maintained by a specialized reverse transcriptase, called telomerase, adding new repeats using a tightly associated RNA template to compensate for the loss (reviewed in ref. 42). However, in many human cells and tissues telomerase activity is low or absent[43], leading to gradual telomere attrition with each cell division. This limits the replicative capacity of a cell but also prevents the outgrowth of a transformed cell to a full-blown tumour[44].

Telomeres in human tumours are often shorter compared with the tissue from which they derive. A large proportion (>90%) of cancer cells has reacquired telomerase activity at a shorter set length[43], demonstrating the need for an active telomere metabolism to sustain tumour growth[45,46]. Furthermore, late-generation mice lacking functional telomerase seem to be resistant to skin carcinogenesis[47], indicating that the telomere-division limitator is relevant for preventing this type of epithelial carcinomas. However, it has been proposed that when telomeres become too short, a transient period of genomic havoc is induced[48,49]. This stage of chromosomal instability could fuel the degeneration of the tumour to a more malignant state, for example via a loss of checkpoint functions leading to a scenario in which chromosomal damage is tolerated and actually drives tumour evolution. This stage of massive genomic instability could contribute to reacquisition of the telomerase activity[48] or to invention of alternative mechanisms that can solve the telomere problem[50]. This may explain the increase in the incidence of spontaneous tumours in highly proliferative cell types such as lymphomas and teratocarcinomas, which are apparent in late-generation telomerase-deficient mouse mutants[51]. Thus, the consequences of the telomere-division limitation may have both advantages and disadvantages in terms of carcinogenesis[48,49].

Concluding remarks and perspectives

Research over the past few years has provided ample evidence that genome instability is one of the main forces driving the onset and progression of carcinogenesis. Genetic degeneration is linked intimately with all aspects of maintenance of DNA integrity and gene function and is fuelled by the continuous erosion of the genome by environmental and endogenous genotoxic agents. The outlines of key systems involved are rapidly emerging, although we may still be missing other important mechanisms, including error-free damage tolerance, additional intricacies of homologous recombination, and chromatin-modification mechanisms. For instance, *de novo* hypermethylation of CpG-rich islands nested in gene promoters seems to be a common means of silencing tumour-suppressor genes in cancer[52]. The era of postgenomics will enable the delineation of the complex response of cells, tissues and intact organisms against DNA injury, disclosing the intricate interactions between DNA repair, replication, transcription, chromatin dynamics, cell-cycle progression and apoptosis. The study of DNA maintenance mechanisms will not only reveal the biological impact of the havoc time wreaks on the genome, including oncogenesis and age-related diseases, but should also uncover new paradigms for prevention, genetic susceptibility, diagnosis and rational therapy. □

1. Loeb, L. A. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.* **51,** 3075–3079 (1991).
2. Cadet, J., Berger, M., Douki, T. & Ravanat, J. L. Oxidative damage to DNA: formation, measurement, and biological significance. *Rev. Physiol. Biochem. Pharmacol.* **131,** 1–87 (1997).
3. Finkel, T. & Holbrook, N. J. Oxidants, oxidative stress and the biology of ageing. *Nature* **408,** 239–247 (2000).
4. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362,** 709–715 (1993).
5. Yamaizumi, M. & Sugano, T. UV-induced nuclear accumulation of p53 is evoked through DNA damage of actively transcribed genes independent of the cell cycle. *Oncogene* **9,** 2775–2784 (1994).
6. Goodman, M. F. & Tippin, B. Sloppier copier DNA polymerases involved in genome repair. *Curr. Opin. Genet. Dev.* **10,** 162–168 (2000).
7. Kunkel, T. A. & Bebenek, K. DNA replication fidelity. *Annu. Rev. Biochem.* **69,** 497–529 (2000).
8. Lawrence, C. The Rad6 DNA repair pathway in *Saccharomyces cerevisiae*: what does it do and how does it do it? *BioEssays* **16,** 253–258 (1994).
9. Masutani, C. *et al.* The *XPV* (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. *Nature* **399,** 700–704 (1999).
10. Johnson, R. E., Kondratick, C. M., Prakash, S. & Prakash, L. hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science* **285,** 263–265 (1999).
11. Zhou, B. B. & Elledge, S. J. The DNA damage response: putting checkpoints in perspective. *Nature* **408,** 433–439 (2000).
12. Friedberg, E. C., Walker, G. C. & Siede, W. *DNA Repair and Mutagenesis* (ASM Press, Washington DC, 1995).
13. Lindahl, T. & Wood, R. D. Quality control by DNA repair. *Science* **286,** 1897–1905 (1999).
14. Takata, M. *et al.* Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* **17,** 5497–5508 (1998).
15. Tornaletti, S. & Hanawalt, P. C. Effect of DNA lesions on transcription elongation. *Biochimie* **81,** 139–148 (1999).
16. de Laat, W. L., Jaspers, N. G. & Hoeijmakers, J. H. Molecular mechanism of nucleotide excision repair. *Genes Dev.* **13,** 768–785 (1999).
17. Batty, D. P. & Wood, R. D. Damage recognition in nucleotide excision repair of DNA. *Gene* **241,** 193–204 (2000).
18. Bootsma, D., Kraemer, K. H., Cleaver, J. & Hoeijmakers, J. H. J. in *The Metabolic and Molecular Basis of Inherited Disease* Vol. 1 (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 677–703 (McGraw-Hill, New York, 2001).
19. Lehmann, A. R. The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev.* **15,** 15–23 (2001).
20. Vermeulen, W. *et al.* A temperature-sensitive disorder in basal transcription and DNA repair in humans. *Nature Genet.* **27,** 299–303 (2001).
21. de Boer, J. & Hoeijmakers, J. H. Cancer from the outside, aging from the inside: mouse models to study the consequences of defective nucleotide excision repair. *Biochimie* **81,** 127–137 (1999).
22. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E. & Samson, L. D. 3-methyladenine DNA glycosylases: structure, function, and biological importance. *BioEssays* **21,** 668–676 (1999).
23. Mol, C. D., Parikh, S. S., Putnam, C. D., Lo, T. P. & Tainer, J. A. DNA repair mechanisms for the recognition and removal of damaged DNA bases. *Annu. Rev. Biophys. Biomol. Struct.* **28,** 101–128 (1999).
24. Bruner, S. D., Norman, D. P. & Verdine, G. L. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature* **403,** 859–866 (2000).
25. Wilson, D. M. & Thompson, L. H. Life without DNA repair. *Proc. Natl Acad. Sci. USA* **94,** 12754–12757 (1997).
26. Divine, K. K. *et al.* The XRCC1 399 glutamine allele is a risk factor for adenocarcinoma of the lung. *Mutat. Res.* **461,** 273–278 (2001).
27. Khanna, K. K. & Jackson, S. P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genet.* **27,** 247–254 (2001).
28. Rotman, G. & Shiloh, Y. ATM: from gene to function. *Hum. Mol. Genet.* **7,** 1555–1563 (1998).
29. Paull, T. T. *et al.* A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr. Biol.* **10,** 886–895 (2000).
30. Petrini, J. H. The Mre11 complex and ATM: collaborating to navigate S phase. *Curr. Opin. Cell Biol.* **12,** 293–296 (2000).
31. Digweed, M., Reis, A. & Sperling, K. Nijmegen breakage syndrome: consequences of defective DNA double strand break repair. *BioEssays* **21,** 649–656 (1999).
32. Riballo, E. *et al.* Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. *Curr. Biol.* **9,** 699–702 (1999).
33. Van Gent, D. C., Hoeijmakers, J. H. J. & Kanaar, R. Chromosomal stability and the DNA double-stranded break connection. *Nature Rev. Genet.* **2,** 196–206 (2001).
34. Gurley, K. E. & Kemp, C. J. Synthetic lethality between mutation in Atm and DNA-PKcs during murine embryogenesis. *Curr. Biol.* **11,** 191–194 (2001).
35. Jasin, M. Chromosome breaks and genomic instability. *Cancer Invest.* **18,** 78–86 (2000).
36. Scully, R. & Livingston, D. M. In search of the tumour-suppressor functions of BRCA1 and BRCA2. *Nature* **408,** 429–432 (2000).
37. Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7,** 263–272 (2001).
38. Gao, Y. *et al.* Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature* **404,** 897–900 (2000).
39. Jiricny, J. Replication errors: cha(lle)nging the genome. *EMBO J.* **17,** 6427–6436 (1998).
40. Jiricny, J. & Nystrom-Lahti, M. Mismatch repair defects in cancer. *Curr. Opin. Genet. Dev.* **10,** 157–161 (2000).
41. Heyer, J., Yang, K., Lipkin, M., Edelmann, W. & Kucherlapati, R. Mouse models for colorectal cancer. *Oncogene* **18,** 5325–5333 (1999).
42. McEachern, M. J., Krauskopf, A. & Blackburn, E. H. Telomeres and their control. *Annu. Rev. Genet.* **34,** 331–358 (2000).
43. Shay, J. W. & Bacchetti, S. A survey of telomerase activity in human cancer. *Eur. J. Cancer* **33,** 787–791 (1997).
44. Counter, C. M. *et al.* Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* **11,** 1921–1929 (1992).
45. Greider, C. W. Telomerase activity, cell proliferation, and cancer. *Proc. Natl Acad. Sci. USA* **95,** 90–92 (1998).
46. Tang, R., Cheng, A. J., Wang, J. Y. & Wang, T. C. Close correlation between telomerase expression and adenomatous polyp progression in multistep colorectal carcinogenesis. *Cancer Res.* **58,** 4052–4054 (1998).
47. Gonzalez-Suarez, E., Samper, E., Flores, J. M. & Blasco, M. A. Telomerase-deficient mice with short telomeres are resistant to skin tumorigenesis. *Nature Genet.* **26,** 114–117 (2000).
48. DePinho, R. A. The age of cancer. *Nature* **408,** 248–254 (2000).
49. de Lange, T. & Jacks, T. For better or worse? Telomerase inhibition and cancer. *Cell* **98,** 273–275 (1999).
50. Bryan, T. M., Englezou, A., Dalla-Pozza, L., Dunham, M. A. & Reddel, R. R. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nature Med.* **3,** 1271–1274 (1997).

51. Rudolph, K. L. *et al.* Longevity, stress response, and cancer in aging telomerase-deficient mice. *Cell* **96,** 701–712 (1999).
52. Jones, P. A. & Laird, P. W. Cancer epigenetics comes of age. *Nature Genet.* **21,** 163–167 (1999).
53. Sugasawa, K. *et al.* A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.* **15,** 507–521 (2001).
54. Tang, J. Y., Hwang, B. J., Ford, J. M., Hanawalt, P. C. & Chu, G. Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis. *Mol. Cell* **5,** 737–744 (2000).
55. Le Page, F. *et al.* Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIIH, and CSB and implications for Cockayne syndrome. *Cell* **101,** 159–171 (2000).
56. Buschta-Hedayat, N., Buterin, T., Hess, M. T., Missura, M. & Naegeli, H. Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA. *Proc. Natl Acad. Sci. USA* **96,** 6090–6095 (1999).
57. Sugasawa, K. *et al.* Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol. Cell* **2,** 223–232 (1998).
58. Houtsmuller, A. B. *et al.* Action of DNA repair endonuclease ERCC1/XPF in living cells. *Science* **284,** 958–961 (1999).
59. Whitehouse, C. J. *et al.* XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair. *Cell* **104,** 107–117 (2001).
60. Haber, J. E. Partners and pathways repairing a double-strand break. *Trends Genet.* **16,** 259–264 (2000).
61. Davies, A. A. *et al.* Role of BRCA2 in control of RAD51 recombination and DNA repair protein. *Mol. Cell* **7,** 273–282 (2001).
62. Hsu, H. L. *et al.* Ku acts in a unique way at the mammalian telomere to prevent end joining. *Genes Dev.* **14,** 2807–2812 (2000).
63. Kolodner, R. D. & Marsischky, G. T. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.* **9,** 89–96 (1999).
64. Harfe, B. D. & Jinks-Robertson, S. DNA mismatch repair and genetic instability. *Annu. Rev. Genet.* **34,** 359–399 (2000).
65. Obmolova, G., Ban, C., Hsieh, P. & Yang, W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **407,** 703–710 (2000).
66. Lamers, M. H. *et al.* The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch. *Nature* **407,** 711–717 (2000).
67. Karran, P. & Bignami, M. DNA damage tolerance, mismatch repair and genome instability. *BioEssays* **16,** 833–839 (1994).

Acknowledgements
I apologize to my colleagues for being able to cite only recent original papers and for the use of reviews for earlier relevant work, owing to space and reference limitations. I thank L. Niedernhofer for valuable help with the manuscript. J.H.J.H. is supported by the Dutch Cancer Society, the Dutch Science Organization (medical and chemical divisions), the EC, NIH and IACR.