Humans are variously and continuously exposed to a wide range of different DNA-damaging agents, some of which are classed as carcinogens. DNA damage can arise from exposure to exogenous agents, but damage from endogenous processes is probably far more prevalent. That said, epidemiological studies of migrant populations from regions of low cancer risk to high cancer risk countries point to a role for environmental and/or lifestyle factors playing a pivotal part in cancer aetiology. One might reasonably surmise from this that carcinogens found in our environment or diet are culpable. Exposure to carcinogens is associated with various forms of DNA damage such as single-stand breaks, double-strand breaks, covalently bound chemical DNA adducts, oxidative-induced lesions and DNA–DNA or DNA–protein cross-links. This review predominantly concentrates on DNA damage induced by the following carcinogens: polycyclic aromatic hydrocarbons, heterocyclic aromatic amines, mycotoxins, ultraviolet light, ionising radiation, aristolochic acid, nitrosamines and particulate matter. Additionally, we allude to some of the cancer types where there is molecular epidemiological evidence that these agents are aetiological risk factors. The complex role that carcinogens play in the pathophysiology of cancer development remains obscure, but DNA damage remains pivotal to this process.

Introduction

DNA damage occurs through exogenous and endogenous processes. Carcinogens, irrespective of their origin, have the ability to evoke the development of DNA damage through a variety of mechanisms. This includes, for instance, covalent binding of carcinogen with DNA or DNA double-strand breaks (DSBs) formed as a result of ionising radiation (IR)-induced free radical generation [1,2]. Carcinogens are categorised as being chemical or physical agents [3], causing DNA damage attributable to their physico-chemical properties, such as DNA molecule distortion or DNA cross-linking [3–6]. Table 1 shows a small subset of environmental and/or dietary carcinogens; however, there are multiple other examples to which humans are potentially exposed (Table 1).

Exposure to carcinogens can either directly [7] or indirectly [1,8] induce DNA damage. Subsequent repair mechanisms may result in alterations in DNA sequences, i.e. mutations [2,9]. Induced mutations may be initiating events in cancer causation, when the damage is fixed within oncogenes or tumour suppressor genes [10]. Such risk may also be directly influenced by individual susceptibility and genetic instability [11]. For example, in the inherited genetic disorder Xeroderma Pigmentosum (XP), mutations in the XP proteins disrupt DNA repair resulting in the build-up of sunlight-induced lesions in skin DNA and a high rate of skin cancer [12].

Types of DNA damage: direct- and indirect-acting carcinogens

Carcinogens may fall into two categories: activation-dependent and activation-independent. Those which do not require metabolic activation or any molecular modification in order to induce DNA damage are termed direct-acting carcinogens and examples include nitrosamines, ultraviolet (UV), IR
Ultraviolet (UV) Direct- and indirect-acting genotoxic cancer-causing agent, primarily absorbed by epidermal

Nitrosamines Metabolism of nitrosamines subsequently induces alkylating DNA damage via the formation of

Aristolochic acid (AA) Naturally derived acids from Aristolochiaceae plants. Ingestion of these carcinogens shown to

Nanoparticles (NPs) Nanotechnology engineering has seen increasing usage of nanoparticles in medical, cosmetics

Polycyclic aromatic hydrocarbons (PAHs) Combustion of organic matter results in the generation of PAHs. These are the most abundant

and alkylating agents [5,7,26,35]. These agents have the capability to interact directly with DNA and other cellular components due to their electrophilic groups. These electrophilic groups exhibit an inherent reactivity, allowing them to interact with nitrogen and oxygen atoms within negatively charged cellular macromolecules to induce molecular modifications and distortion [3]. Alteration of DNA bases causes a disarrangement of the genetic material and formation of DNA adducts depending on the type of carcinogen. Failure within DNA repair mechanisms allows DNA lesions to be inherited by daughter cells [7], eventually leading to the accumulation of DNA damage and potentially the development of cancer.
Indirect-acting carcinogens are relatively unreactive parent compounds that include polycyclic aromatic hydrocarbons (PAHs), heterocyclic aromatic amines (HAAs), N-nitrosamines, mycotoxins and aristolochic acid (AA). These typically require bioactivation in host cells to transform them into carcinogenic metabolites or reactive intermediates that are capable of exerting genotoxic effects [1,8,36]. This is often mediated by phase I and/or II metabolic reactions. Phase I reactions include oxidation, reduction or hydrolysis, mainly involving cytochrome P450 (CYP) mixed function oxidase isoforms, commonly referred to as CYPs. These enzymes have the ability to activate carcinogens independently or in conjugation with phase II enzymes such as sulfotransferases and N-acetyltransferase [8,37,38]. A classic example is the bioactivation process of benzo[a]pyrene (B[a]P), which undergoes a multi-step process involving CYP1A1 and epoxide hydrolase-mediated conversion to r7, t8-dihydroxy-t-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene (BPDE) (Figure 1). Bulky chemical adducts are commonly seen as a result of the interaction between activated carcinogens and DNA, e.g., reactive nitrenium ions formed through the reduction and hydrolysis of AA, yield bulky purine DNA adducts at the exocyclic amino group of purines [39]. Nitrosamines encompass a large diverse group of compounds formed by various combinations of amines and nitrogen functional groups. Some nitrosamines are known to be direct-acting carcinogens such as those formed in foodstuffs and are implicated in oesophageal cancer or stomach cancer, while others such as the tobacco-specific lung pro-carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone are bioactivated [40]. Other work shows that nitrosamines can be locally activated within the urothelium [41]. In contrast, α-nitrosoamino aldehydes are highly reactive compounds which are direct-acting mutagens [42]. We see the relevance of this in the high incidence of gastric cancers in certain regions associated with consumption of particular foodstuffs [43].

Although indirect carcinogens are reliant on activation, a few have the ability to enhance bioactivation by inducing changes in gene expression. PAHs such as B[a]P increase the expression in members of CYP450 family by acting as exogenous ligands of the cytosolic aryl hydrocarbon receptor (AhR)–aromatic receptor nuclear translocator complex [36,44]. Such enzymes are also involved in bioactivation of HAAs, PAHs, AA and aflatoxins, therefore potentially increasing the metabolism and subsequent exposure of DNA to reactive intermediates [1]. Expression of such enzymes has been investigated in tissues possessing the capability of bioactivating carcinogens to reactive electrophiles [45,46].

**UV-induced damage**

UV-induced lesions promote chemical modification and structural distortion of DNA by forming photoproducts and oxidative stress. Production of photoproducts, such as cyclobutane pyrimidine dimers (CPDs), pyrimidine-(6–4)-pyrimidone photoproducts and their dewar isomers, is achieved through the direct absorption of UVB (290–320 nm) incident photons by DNA bases and methylation of cysteine bases [47,48]. CPDs account for 75% of the mutations, which are induced by UV [47,49]. T-C and C-C CPD lesions are predominant in the tumour suppressor TP53 and in patients with skin cancer. T-T CPDs are less persistent as...
eradication of these dimers is induced by insertion of adenine bases by DNA repair mechanisms [17]. Lesions that are difficult to remove result in the following: stalling of DNA and RNA polymerase, reduction in DNA replication, protein synthesis and mRNA synthesis [49].

UVA (320–400 nm), a poorly absorbed radiation by DNA [50] with an unknown mutagenic effect, is suggested to be associated with promoting DNA damage by oxidative stress through an activation-independent route [17,50–55]. UVA photons absorbed by photosensitisers promote photo-oxidation reactions largely giving rise to single oxygen molecules or highly reactive electrons, which subsequently target guanine bases for hydration and deprotonation. Hydration of guanine bases promotes production of 8-oxo,7,8-dihydroguanyln (8-oxodGuo) radicals which are considered to be a miscoding lesion (a lesion capable of base pairing with either a cytosine or/and adenine residue) in DNA and a marker for oxidative stress [5,51]. UV-derived 8-oxodGuo radicals have not been shown to promote G:T transversions in mammalian cells which are a common hallmark for 8-oxodGuo-induced mutations [51,52]. However, these lesions are known to cause molecular distortion by changing the structure of purine bases within DNA, but other mutagenic factors are likely to be linked with UVA-induced damage [48].

**Chemical-induced bulky DNA lesions**

Bulky chemical DNA adducts are formed when a reactive electrophilic carcinogen, formed by the metabolism of an indirect carcinogen, binds to a particular nucleophilic moiety in DNA. Nucleophilic targets of the reactive carcinogen include nitrogen and oxygen atoms within the bases and phosphodiester backbone of DNA. The binding of the electrophile to a nucleophile is dependent on the electrophilic strength of the carcinogen [56–58]. Common target sites on DNA include N and O of guanine or N of adenine [56] (see Table 1). Because generation of such adducts leaves the replication process prone to error, the presence of DNA adducts results in replication arrest to facilitate repair mechanisms being recruited to remove the covalently bonded chemical [2]. Unsuccessful repair of the damage often results in transversion or transition mutations. An example of this is seen with a reactive intermediate of the mycotoxin, aflatoxin B₁ (AFB₁). AFB₁–8,9-epoxide interacts with guanine bases in hepatocyte DNA to form adducts [27,28,59]. A common result of replication errors induced by AFB₁ is a transversion mutation at codon 249 in exon 7 in TP53, where guanine is substituted by thymine [59].

**Oxidative damage**

Oxidative-induced DNA damage is formed due to either exogenous or endogenous factors such as UVA, IR or endogenously generated oxygen molecules, which induce intracellular oxidative stress. The most prevalent sources of primarily induced oxidative stress include reactive oxygen species (ROS), such as hydroxyl radicals (OH·), singlet oxygen or reactive nitrogen species such as peroxynitrite [60]. These mutagenic species are known to interact with macromolecules causing defects in DNA synthesis and repair mechanisms, as well as inactivating various key proteins and repair enzymes. Guanine bases are the main target for these species, especially ROS [60–63]. ROS-induced damage forms modified bases, apurinic/apyrimidinic (AP) sites and single-strand breaks (SSBs). The addition of OH· at position C8 within the guanine ring generates the oxidative product, 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxodG) [60–62]. Similarly, the addition of OH· at position C8 of deoxyadenosine generates the oxidative product 8-oxo-7,8-dihydro-2′-deoxyadenosine (8-oxoA). These radicals are capable of further reduction or oxidation forming 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyGua) or 8-oxo-7,8-dihydroguanine (8-oxoG), in deoxyguanosine or 4,6-diamino-5-formamidopyrimidine (FapyA) or 7,8-dihydro-8-oxoadenine (8-oxoA) in deoxyadenosine. These are non-coding mutagenic DNA bases (cannot be copied by the constitutive replication machinery) [60,63]. Another prevalent oxidative product is thymine glycol, produced by the insertion of OH· at position C5 of thymine rings [63]. Similarly, another oxidation product of cytosine is cytosine glycol, which upon deamination leads to the formation of uracil glycol. These bases are removed by DNA glycosylase enzyme through base excision repair (BER) [60,61,63]. Accumulation of these modified bases over time enhances genomic structure defects and instability. For instance, thymine glycol-induced conformational alterations modify telomeres [64], with 8-oxoDG also playing a role.

**Cross-linking damage**

Not only do many chemotherapeutic agents (e.g. cisplatin) form DNA cross-links, but also cross-links can result from the endogenous generation of malondialdehyde and acetaldehyde, which can form either in
combination or in isolation [4,65]. Cross-linking agents are often clastogens, their exposure resulting in chromosomal aberrations in the form of broken fragments of chromosomes [66]. They generate what can be pre-mutagenic lesions via the formation of covalent bonds between two nucleotide residues (e.g. two exocyclic guanine amino groups) in either the same strand (known as an intrastrand link) or opposite strands (known as interstrand links). Intrastrand links are often repaired via nucleotide excision repair (NER) and homologous recombination (HR). Intrastrand and interstrand links are generated when the cross-linking agent possesses two independent reactive groups, which then react with the DNA base groupings on either the same or the opposing strands. Common targets of the nucleotides for interstrand links include N7 of guanine or the exocyclic N2 amino group of guanine on opposite strands [4]. Interstrand links are toxic DNA lesions that inhibit replication and transcription of the affected DNA, due to the physical obstruction of strand separation [67]. If interstrand link lesions remain unrepaired in the mammalian genome, they can result in catastrophic cell breakdown [4,66,67].

**Single- and double-strand breaks**

SSBs occur as a consequence of endonuclease enzyme activity during base excision repair [68]. These can be induced by numerous different exogenous and endogenous elements such as UV, IR, B[a]P, mycotoxins and a vast array of intracellular reactions activating the formation of radicals or enzymes including nucleases [68–71]. The most common endogenous source associated with the formation of SSBs is the presence of OH· within cells [68,70]. OH· radicals are induced during oxidative stress and through Fenton’s reaction, which involves the intracellular reduction in hydrogen peroxide (H2O2) by a transition ion, mostly iron²⁺ (Fe²⁺) [68,70]. OH· radicals interact with hydrogen atoms within DNA backbone causing catalysis of phosphodiester bonds resulting in the formation of phosphoglycolate and DNA lesions [70]. Carcinogen-generated radicals also trigger oxidative stress, including a series of events leading to the production of miscoding DNA bases such as thymine glycerol and activation of nucleases [68]. Nuclease activation results in a scenario that resembles the process of apoptosis because of cleavage of the DNA backbone, creating DNA strand lesions [68].

DSBs, if left unrepaired or misrepaired (or misreplicated), can result in cell death, genetic instability and carcinogenesis [72]. The formation of DSBs can be caused endogenously, e.g. during meiosis I DSBs are intentionally induced to ensure chromosomal segregation; they can also be exogenously induced by IR or chemical carcinogens [72,73]. It should be noted that other types of damage can lead to DSBs, and these are often associated with the action of exogenous agents [74]. IR is a significant external agent that induces DSBs directly and indirectly, primarily damage mediated by ROS generated by radiolysis of water. Direct induction of damage occurs when a high-energy particle collides with the phosphodiester backbone of the DNA strands, causing cleavage [56,73]. IR-mediated damage via ROS generation can be targeted by BER, thus generating SSBs. Repair of clustered damage in both strands of DNA can result in closely opposing SSBs, which then present in the form of DSBs [72]. Subsequent repair and processing of DSBs can lead to mutations, loss of heterozygosity and chromosomal translocations resulting in cell death [73].

**Human studies of cancers associated with carcinogen-induced DNA damage**

There are many human studies implicating environmental- and/or dietary-associated carcinogen exposures in the aetiology of cancer. Carcinogen-DNA damage formed post-combustion of tobacco implicated in the aetiology of lung cancer is a prime example [75]. Studies of migrant populations that demonstrate a generational change in susceptibility to cancers such as breast, prostate, colorectal and stomach cancer also document exposure to DNA-damaging carcinogens [76]. For many of these cancers, there can be compelling evidence, through molecular epidemiology studies, that carcinogen exposure is a pivotal causative factor [77,78]. While the pathophysiology of cancer is undoubtedly complex, many different agents can generate specific forms of DNA damage. Within the scope of this review, only a few typical examples will be highlighted.

**Lung cancer and tobacco smoking**

Combustion of a cigarette during smoking generates thousands of agents, many of which have the potential to be DNA damaging [79]. It is estimated, in countries where tobacco smoking is common, that 90% of lung cancer cases are directly attributed to smoking [80]; however, it is also well established that smoking-derived carcinogens induce cancers at other tissue sites [81]. Carcinogenic PAHs and tobacco-specific nitrosamines are
the major carcinogens found in tobacco smoke, generated in the 900°C combustion environment at a lit cigarette tip induced when a smoker puffs. Following bioactivation of PAHs, primarily via the CYP mixed function oxidase system, their electrophilic metabolites covalently bind to nucleophilic sites on macromolecules, including DNA bases, to form bulky chemical DNA adducts. This is shown in Figure 1 for the pro-carcinogen \( \text{BaP} \) and its major DNA adduct \( r7,t8,t9\)-triacehodixy-c-10-(\( N^2 \)-deoxyguanosyl)-7,8,9,10-tetrahydrobenzo[\( a \)]pyrene (BPdG). Formation of PAH-DNA adducts can result in the induction of G-T transversions in \( TP53 \), due to DNA replication of unrepaired DNA, which produces mutations at sites of DNA adduct formation. At codon 157 in \( TP53 \) (a hotspot for mutation induction), G-T transversions are frequently seen in smokers’ lung cancers, but not in never-smokers [59].

**Prostate cancer and carcinogenic PAH exposure**

Residents of countries such as India, China and Japan have typically been at a lower risk of prostate cancer compared with UK residents; however, risk within such ethnic groupings rises in the grandchildren of migrants from India, China or Japan to North/Western Europe, implicating environmental and/or dietary exposures as causative factors [45,46]. Normal human prostate has been shown to possess the extra-hepatic metabolic capacity of phase I and II enzymes able to bioactivate many pro-carcinogens [45]. In fact, some such as CYP1B1 are expressed at a higher level in the cancer-susceptible peripheral zone of human prostate compared with the transition zone [82]. A small cohort study of 12 UK- versus 14 India-resident individuals was undertaken to

---

**Figure 2. PAH-DNA immunostaining of human prostate.**

(A) Representative example of PAH-DNA immunostaining in a UK prostate sample stained for carcinogenic PAH-DNA adducts: (left panel) specific PAH-DNA adduct staining is shown by nuclei stained pink and indicated by arrows; (middle panel) the corresponding absorbed serum control shows the same area with no staining and (right panel) haematoxylin staining of the same area shows localisation of nuclei. (B) Values for PAH-DNA adducts/10^8 nucleotides, for 10 prostate samples from the U.K. and 13 samples from India, were obtained from IHC using ACIS OD/nucleus values (with absorbed serum subtracted) by calculation from a standard curve [85,86].
explore the level of expression of various proteins in benign prostate tissue, including CYP1B1, oestrogen receptor-alpha (ERα) and oestrogen receptor-beta (ERβ). Expression of CYP1B1, an enzyme potentially involved in the metabolic bioactivation of PAHs or HAAs, was markedly higher in UK individuals. Increased ERβ expression correlated with under-expression of CYP1B1, especially in Indian residents [46]. Figure 2 shows an immunostaining of carcinogenic PAH-DNA adducts in the peripheral zone of cancer-free prostates from the U.K. and India stained with antiserum specific for a family of carcinogenic PAHs bound to DNA, and semi-quantified by the Automated Cellular Imaging System (ACIS) [83]. The prostate tissues from India were obtained at a time (2008–2010) when radical retropubic prostatectomy was a rare procedure in this region [84]. While Figure 2B shows a marked interindividual variation in adduct levels, there is a relatively similar range of PAH-DNA adduct levels in the prostates from both groups. One might argue that these groups were very small, and may not be completely representative. Furthermore, India has, in recent years, seen an increasing Westernisation of lifestyle and these tissues would likely have been sourced from individuals from a high socio-economic class. Changes in prostate cancer incidence remain to be seen [76].

**Urothelial carcinoma and AA**

Balkan endemic nephropathy (BEN) and aristolochic acid nephropathy (AAN) [formerly known as Chinese herb nephropathy (CHN)] are conditions with significant evidence of a carcinogen causing DNA damage. BEN and CHN are associated with a high incidence of upper tract urothelial carcinoma and renal failure, both caused by ingestion of AA. Exposure and subsequent metabolism (bioactivation) of AA lead to the formation of aristolactam AL-DNA adducts in urothelial tissue [87,88]. These bulky chemical DNA adducts were shown to be directly linked to A:T to T:A transversion mutations in TP53 in a study conducted by Grollman and colleagues [20,89]. There is significant evidence that AA is both a powerful nephrotoxic and carcinogenic agent with an extremely short latency period, not only in animals but also in humans [90]. In a typical human subject presenting with a urothelial malignancy 6 years post-presentation with AA-associated nephropathy, mutation analysis showed AAG → TAG mutations in codon 139 (Lys → Stop) of exon 5 of TP53 [91].

**Measurements of DNA damage**

Given the differing forms of DNA damage, a range of techniques measuring different endpoints have been developed. Antisera elicited against DNA adducts or carcinogen-modified DNA samples have been used to detect adducts of specific classes by immunoassay or immunohistochemistry [86]. For bulky chemical DNA adducts, where chemical characterisation is not required, the 32P-postlabelling method based on multi-dimensional thin layer chromatography has been commonly used [83,85,86]. To determine and quantify levels of DNA SSBs or DSBs, alkaline or neutral versions, respectively, of the single cell-gel electrophoresis ('comet') assay can be used [92,93]. Post-lysis incorporation of enzymes [formamidopyrimidine DNA glycosylase (Fpg) or 8-oxoGua DNA glycosylase (OGG1)] to measure 8-oxoGua in DNA] to cleave bulky lesions or oxidative damage into SSBs can be employed to discriminate an agent's mechanism of DNA damage induction or to enhance the sensitivity of the alkaline version of the comet assay [94]. The cytokinesis-block micronucleus assay determines levels of chromosomal damage, primarily clastogenic or aneuploidy effects [95,96]. Endpoints of oxidative damage such 8-hydroxy-2'-deoxyguanosine (8-OHdG) can be determined using competitive enzyme-linked immunosorbent assay (ELISA) [97]. Highly sensitive variations in the chemical-specific mass spectrometry-based methodologies have been developed and used very successfully to obtain precise characterisation of DNA adducts in human tissues [98,99].

**Repair mechanisms**

DNA repair mechanisms [BER, NER, HR and non-homologous end-joining (NHEJ)] maintain genomic stability by eradicating DNA damage induced prior to replication completion [100]. BER is activated upon spontaneous depurination, deamination, methylation and oxidation of DNA bases. It is initiated by hydrolysis of N-glycosyl bond between deoxyribose sugar and DNA base by glycosylase enzyme, creating an abasic site (i.e. AP site) [71,100]. The AP site is cleaved by two enzymes: 5'-AP endonuclease and deoxyribose phosphodiesterase, inducing a nucleotide gap [71,100]. The gap is recovered by DNA polymerase β, which uses a template strand to introduce a new nucleotide, and DNA ligase, which stabilises and seals the phosphate-sugar backbone. An example of BER activation includes the substitution of normal DNA base by oxidative-induced base such as thymine glycol or 8-oxoguanine [67].
NER is a highly conserved process initiated by a multi-subunit complex promoting eradication of UV-induced lesions such as CPDs and bulky chemical-induced DNA adducts (see Figure 1) [71,101]. The mechanism is divided into five stages: recognition, incision, excision, DNA synthesis and DNA ligation [71]. Recognition of DNA damage occurs either due to stalling of RNA polymerase during transcription or by a repair complex, XPC-HHR23B [71,100,101]. Other repair components are recruited such as TFIIH, XPB and XPD, which perform helicase activity and unwind the DNA developing a bubble containing 24–30 nucleotides, followed by the recruitment of pre-incision components (XPA, RPA and XPG). XPF–ERCC1 incision complex removes the oligonucleotide leaving behind a gap within the DNA strand, which is synthesised and sealed by DNA polymerase δ/ε and DNA ligase [71,100,101].

There are two subtypes of NHEJ: classical (C-NHEJ) and alternative (A-NHEJ), both of which require no template for repair of DSB lesions [9]. C-NHEJ encompasses four steps: DNA end recognition, bridging and stabilising of ends, processing and ligation. This process requires heterodimer Ku, containing Ku70/80 subunits. Ku has a high affinity to DNA ends enabling it to localise and bind to the phosphate backbone situated at the break. Once bound, Ku acts as a scaffold recruiting other complexes to allow bridging and to create ligatable ends. DNA ligase IV is activated and stabilised by XRCC4—allowing ligation of the broken ends to repair the DSB [9,102]. Mutation or inhibition of C-NHEJ initiates A-NHEJ, which induces complex indels (insertion/deletion) in the repair junctions affecting genome integrity [9].

HR is predominantly involved in the repair of DSBs and also interstrand cross-links in conjunction with NER. HR utilises sister chromatids as a template ensuring genetic information is retained. Owing to the use of similar/identical nucleotide sequences as a template, the repair can only occur in the S-/G2-phases. HR follows NER. HR utilises sister chromatids as a template ensuring genetic information and providing support for DNA replication in the case of broken or stalled replication forks [103,104]. The distinction between error-prone and error-free DNA repair appears to be an interplay between the DNA repair mechanism and the lesion being repaired; for instance, NHEJ has been regarded as error-prone, but this might have been overestimated, whereas HR, which is typically described as error-free, is increasingly being considered to be highly mutagenic [9].

**Conclusion**

Cancer is a complex multi-stage process that likely starts with an initiating mutation post-exposure to a DNA-damaging agent, which is followed by mechanisms such as inflammation [105]. However, it may be that not all mutagens are carcinogens. For instance, there are agents that test positive for mutagenicity in *Salmonella typhimurium* (i.e. the Ames test), but appear to be non-carcinogenic in rodents, e.g. benzene amines and substituted benzene amines [106]; the validity of these observations remains a subject of debate even after several decades. Exposure to carcinogens leads to various forms of DNA damage through indirect and direct pathways. DNA damage can also be implicated in other pathologies, such as neurodegenerative disease [107]. Identification of mutation spectra resulting from carcinogen exposures could give rise to intervention studies resulting in reduced cancer risk in certain cases [108]. For many cancers that may have a dietary and/or lifestyle component [109], there remain enormous gaps in our knowledge regarding candidate causative agents, the interaction between metabolic bioactivation to DNA-damaging species and subsequent repair of the DNA lesion, and the following processes that lead to cancer. Understanding this complex interplay is critical towards understanding the aetiology of this disease.

**Abbreviations**

A-NHEJ, alternative NHEJ; AA, aristolochic acid; AAN, aristolochic acid nephropathy; AFB₁, aflatoxin B₁; AP, apurinic/apyrimidinic; BEN, Balkan endemic nephropathy; BER, base excision repair; B[a]P, benzo[a]pyrene; BPDE, r7,t8-dihydroxy-t-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene; BPDG, r7,t8,t9-trihydroxy-c-10-(N²-deoxyguanosyl)-7,8,9,10-tetrahydrobenzo[a]pyrene; C-NHEJ, classical NHEJ; CHN, Chinese herb nephropathy; CPDs, cyclobutane pyrimidine dimers; CYP, cytochrome P450; DSBs, double-strand breaks; ERα, oestrogen receptor-alpha; ERβ, oestrogen receptor-beta; HAAs, heterocyclic aromatic amines; HR, homologous recombination; IR, ionising radiation; NER, nucleotide excision repair; NHEJ, non-homologous end-joining; NP, nanoparticles; PAHs, polycyclic aromatic hydrocarbons; ROS, reactive oxygen species; SSBs, single-strand breaks; UV, ultraviolet; XP, Xeroderma pigmentosum
Acknowledgements
A bursary from UK Environmental Mutagen Society (UKEMS) funded placements for J.L.B. and M.Z. F.L.M. acknowledges invitation to present at the “Hydrogen Bonds & DNA: Commemoration of the 70th anniversary of the discovery by J.M. Creeth and colleagues at Nottingham in 1947” that took place at the University of Nottingham on 10 November 2017, a Biochemical Society and RSC Chemistry Biology Interface Division Focused Meeting. The prostate immunohistochemistry studies were supported by the Intramural Research Program of the Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD, U.S.A.

Competing Interests
The Authors declare that there are no competing interests associated with the manuscript.

References

© 2018 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY-NC-ND).