**Author for correspondence:**
Linda Chelico
e-mail: linda.chelico@usask.ca

# The interesting relationship between APOBEC3 deoxycytidine deaminases and cancer: a long road ahead

Milaid Granadillo Rodríguez, Ben Flath and Linda Chelico

Department of Microbiology and Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

MGR, 0000-0001-5399-4836; LC, 0000-0003-4673-8978

Cancer is considered a group of diseases characterized by uncontrolled growth and spread of abnormal cells and is propelled by somatic mutations. Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3) family of enzymes are endogenous sources of somatic mutations found in multiple human cancers. While these enzymes normally act as an intrinsic immune defence against viruses, they can also catalyse 'off-target' cytidine deamination in genomic single-stranded DNA intermediates. The deamination of cytosine forms uracil, which is promutagenic in DNA. Key factors to trigger the APOBEC 'off-target' activity are overexpression in a non-normal cell type, nuclear localization and replication stress. The resulting uracil-induced mutations contribute to genomic variation, which may result in neutral, beneficial or harmful consequences for the cancer. This review summarizes the functional and biochemical basis of the APOBEC3 enzyme activity and highlights their relationship with the most well-studied cancers in this particular context such as breast, lung, bladder, and human papillomavirus-associated cancers. We focus on APOBEC3A, APOBEC3B and APOBEC3H haplotype I because they are the leading candidates as sources of somatic mutations in these and other cancers. Also, we discuss the prognostic value of the APOBEC3 expression in drug resistance and response to therapies.

## 1. Introduction

Cancer is a generic term used to describe more than 100 diseases in which cells grow out of control in any part of the body. It has long been known that cancer has a basis in somatic mutations that alter a diversity of cellular functions resulting in sustained proliferative signalling, evasion of growth suppressors and genome instability [1]. Mutations contribute to genomic variation, and may result in neutral, beneficial or harmful consequences for an organism. Cancer genomic sequencing studies have identified mutational signatures that reflect the corresponding causes of these mutations. Mutagenesis originates from exogenous sources and endogenous sources that reside intracellularly [2–5]. Exogenous sources include radiation and chemical damage [2,6,7] and endogenous sources are DNA replication errors, the inability to repair the DNA damage after it has been triggered, and agents that impair DNA directly [8].

The first evidence revealing the apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3) family of enzymes as endogenous sources of somatic mutations found in human cancer were provided in 2012 by Nik-Zainal *et al.* [9] and Roberts *et al.* [10]. Then, in 2013, the extensive resources generated by The Cancer Genome Atlas (TCGA) revealed APOBEC mutagenesis in multiple cancer types [11–13]. There are seven APOBEC3 (A3) enzymes in humans (A3A–H, excluding E) that are capable of inducing DNA mutations through the deamination of cytosine to form promutagenic uracil on single-stranded (ss) DNA. This is the main mechanism by which these enzymes

restrict replication of endogenous retroelements, DNA viruses and RNA viruses [14–18]. Some of the A3 enzymes can also deaminate RNA, although the physiological function is yet to be determined [19–22].

The studies of A3 enzymes have been predominantly focused on understanding their antiviral activities, biochemical properties or retroactive analysis of mutated human genomic DNA sequences, but still little is known about the step-by-step process of how these enzymes have a propensity to act in an 'off-target' fashion in human genomes. This new concept in cancer involves the overexpression of these enzymes in a 'wrong cell' or at the 'wrong time' and catalysing the 'off-target' deaminations in human ssDNA intermediates, with implications in somatic mutagenesis. Those somatic mutations are found in approximately 15% of sequenced human tumours [3,5,11,12,23] with A3A, A3B and A3H haplotype I (Hap I) as leading candidates [3–5,13,24].

This review addresses the interesting relationship between A3 enzymes and cancer. There are several excellent reviews about this topic [25–31]; however, in this review, we bring together for the first time the clinical, molecular, genetic and biochemical perspectives regarding A3A, A3B and A3H Hap I for four main cancers where there is the most published information (breast (BRCA), lung, bladder (BLCA) and human papillomavirus (HPV)-associated cancers), although A3-induced mutations are involved in at least 16 cancer types [3,5,11,12,24,32–38]. This review brings ideas together from multiple disciplines and enables novel conclusions, hypotheses and gaps of knowledge in the field to be identified. First, we explain the physiological functions of these enzymes and the biochemical basis of the somatic mutagenesis to provide a better understanding on how these enzymes are involved in cancer. Second, using clinical, molecular and genetic information, this review discusses the relevance of these enzymes for the prognosis or treatment of cancer.

## 2. A3 physiological functions

A3 enzymes belong to a larger cytidine deaminase family, which in humans also includes activation-induced cytidine deaminase (AID), APOBEC1, APOBEC2 and APOBEC4 [15]. The human A3 family was identified in 2002 and is composed of seven members (A3A, A3B, A3C, A3D, A3F, A3G and A3H) that are encoded by genes located on chromosome 22 [39]. The expression of A3 enzymes is tissue and cell type-specific [40,41] and they are expressed at different levels in immune cell populations such as $CD4^+$ (naive and memory subsets) and myeloid cells. Peripheral blood leucocytes express transcripts for all A3 enzymes, with A3A and A3G being the most represented [40]. The expression of A3A is specific to cells of the myeloid lineage, whereas A3G is highly expressed in $CD4^+$ T lymphocytes [40,41]. The A3 enzymes are also present in non-immune tissues (e.g. epithelial, lung, ovary and adipose tissue), sometimes constitutively or after upregulation due to viral infection [40,42–44].

The A3 enzymes deaminate cytosine in ssDNA which forms uracil, which is not a natural base in DNA and is treated as a promutagenic lesion [45,46]. The intrinsic deaminase activity of these enzymes is mitigated by DNA repair processes, which usually restores the original DNA sequence in an error-free manner through base excision repair (BER) that can remove this lesion from DNA through the action of uracil DNA glycosylase (UNG) [45,46]. However, A3-mediated deamination of cytosines to uracils can also lead to C-to-T mutations directly through DNA replication using uracil as a template or other mutations by translesion synthesis (TLS) polymerases that insert incorrect bases opposite abasic sites after uracil removal. According to yeast experiments, the observed C-to-G transversions that are linked to A3 deamination activity may be caused by TLS bypass over an abasic site by REV1 and DNA polymerase ζ after uracil base removal by UNG2 [47,48]. The yeast studies also observed that a smaller number of C-to-A transversions occurred as a consequence of the generation of abasic sites after removal of uracils, but by a still unclear mechanism [48]. The loss of Rev1 catalytic activity does not have any apparent effect on the ability to insert A opposite abasic sites, indicating that polymerase activities besides that of Rev1 are entirely responsible for A insertion events [49].

The subcellular localization of A3 enzymes is important for their biological functions. A3 enzymes can have different subcellular localizations: A3G, A3F and A3D are cytoplasmic, whereas A3A, A3C and A3H display pan-cellular localization and A3B is localized predominantly in the nucleus [50]. The localization differences (cytoplasmic or nuclear) place these enzymes in the best position to inhibit different viral pathogens, e.g. human immunodeficiency virus (HIV)-1 (cytoplasm) or retroelements (nucleus) and the activities of A3s in the same cellular compartment are redundant [51–53]. Primarily, the A3 enzymes constitute an innate barrier to retroviruses including HIV-1, endogenous retroelements, DNA viruses (e.g. hepatitis B virus (HBV), adeno-associated virus, herpes simplex virus 1, HPV, Epstein–Barr virus (EBV)) and RNA viruses such as human coronavirus [14–18].

The restriction of the replication of these viruses occurs primarily through the deaminase-dependent activity of A3 enzymes which results in hypermutated and inactivated viral genomes. A3G was the first A3 enzyme demonstrated to have restriction activity against HIV infection through G-to-A mutations in the sense DNA strand creating non-infectious virions when uracils in the anti-sense DNA were used as a template in DNA synthesis. Also, uracil containing viral DNA can be degraded by host DNA repair enzymes [54,55]. Deamination-independent antiretroviral mechanisms for A3G have also been described, such as inhibition of HIV reverse transcriptase [56]. Similar to A3G, A3F and A3H can function as HIV restriction factors through encapsidation into budding virions to exert their antiviral activity in newly infected cells, whereas A3A can restrict infection directly in the target cells where it is endogenously expressed [57]. However, this only occurs if A3 enzymes bypass the HIV viral infectivity factor (Vif), which facilitates their ubiquitination and degradation [58,59]. Initially, it was thought that this intricate interplay of virus and host interactions was specific to HIV and A3s. However, it has been shown that even for DNA viruses, such as EBV, there exists a viral protein, in this case BORF2, that binds to and inhibits A3B [60]. Otherwise, A3B would suppress viral replication through deamination of cytosines. There may be other viruses that express an A3 antagonist protein, but currently, other viruses that A3s can restrict, such as coronavirus [18] and HBV [61], are not known to have a counteraction mechanism. In addition to the role of A3 enzymes in intrinsic antiviral responses, these enzymes are also involved in innate and adaptive immunity. It was documented that A3G has a role

in the recognition of HIV-infected cells by NK cells and CD8⁺ cytotoxic T lymphocytes [62,63], whereas A3A is able to edit the transcripts of several genes associated with viral pathogenesis, in monocytes and macrophages [19].

Another important physiological function of A3 enzymes is the restriction of endogenous retroelements. These retroelements including those containing long terminal repeats (LTR) like endogenous retroviruses, as well as non-LTR elements like long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs), could have provided the evolutionary pressure necessary for the maintained expansion of the A3 locus in primates. Non-LTR elements, including LINE1 and Alu, are restricted by human A3 enzymes and this mechanism appears to be deamination-independent, except for A3A, in contrast with the restriction mechanism for LTR containing endogenous retroviruses [64–68]. Endogenous retroelements and their mobility are believed to have played a central role early in shaping the human genome during speciation [66,69]. Ultimately, cells have devised strategies to defend and preserve genomic integrity and the evolution of the A3 family has likely played a prominent role in this defence and in diversifying the retroelements to make them more useful for the host species [66].

At a population level, there is a cost to this good defence. When the enzymes localize to the nucleus and have access to the genomic DNA, there is the risk of 'off-target' activity against host genomic DNA and the potential for mutagenesis through error-prone DNA repair systems. In general, these mutations occur randomly across the genome during our lifetime and sometimes the 'wrong combination' of somatic mutations can transform a normal cell into a tumoural cell.

# 3. Biochemical basis of A3 enzyme activity

Four members of A3 enzymes contain two zinc-coordinating domains (A3B, A3D, A3F and A3G) and three members contain one zinc-coordinating domain (A3A, A3C and A3H) with the consensus sequence His-$X$-Glu-$X_{23-28}$-Pro-Cys-$X_{2-4}$-Cys (figure 1$a$) [73,74]. For A3 enzymes with two Z-domains, only the C-terminal domain is catalytically active, although both domains coordinate zinc (figure 1$b$) [75–78]. Among all these enzymes, A3A has the highest catalytic activity [79]. The APOBEC enzymes induce mutations in a sequence-specific manner and the majority of A3 family members preferentially deaminate the central cytidine in 5′HT$\underline{C}$W trinucleotide motifs (where H = A, C or T and W = A or T; the deaminated based is underlined) within ssDNA substrate, except A3G (5′CC$\underline{C}$A) that deaminates cytidines in a different sequence motif [80].

The A3 enzymes can bind cellular RNA and this can have an impact on their enzymatic activity. A3D, A3G and A3F all bind cellular RNA and form a ribonucleoprotein high molecular weight molecule that is catalytically inactive *in vitro* unless treated with RNaseA [81–83]. Some studies have suggested that RNA binding could be a mechanism to inhibit A3 activity on human genomic DNA and promote cytoplasmic localization [84,85]. In the case of A3H, there is a clear association of RNA binding and cytoplasmic localization, suggesting that RNA is involved as a regulatory mechanism. A3H uses RNA in cells to dimerize (figure 1$c$), which promotes enzyme activity [86]. A3H has seven major haplotypes (Hap I–VII) with A3H Hap II mostly localized to the cytoplasm

and A3H Hap I mainly localized to the nucleus [87], and only A3H Hap I is involved in somatic mutagenesis [5]. By contrast, A3B is strongly inhibited by RNA indicating that A3B likely requires activation for activity in cells [24], similar to what was reported for AID [88,89] and A3A does not bind cellular RNA [24,83]. A3H Hap I has minimal inhibition by RNA, consistent with primarily but not exclusively nuclear localization, suggesting an intermediate phenotype between A3B and A3A [84,86,87]. The negative regulatory role of RNA in suppressing the DNA deaminase activity is likely to be relevant to preventing the accumulation of somatic mutations in development, ageing and cancer [84].

# 4. Biochemical basis of A3-induced somatic mutagenesis

The A3 enzymes that induce somatic mutagenesis must not only be able to localize to the nucleus but must also be able to deaminate transiently available ssDNA created during dynamic processes such as transcription, replication or double-strand break (DSB) repair. These enzymes can cause mutations in the cellular genome at replication forks (figure 2$a$) or within transcription bubbles (figure 2$b$), depending on both the physiological state of the cell and the phase of the cell cycle during which they are expressed [44]. The deoxycytidine deaminase activity of A3A, A3B and A3H Hap I, as leading candidates, has been implicated in cancer and tumour evolution by providing the cells with a diverse pool of mutations [25,90]. Previous experiments of the capability of the A3s to affect cell cycle progression suggest that also A3D might play a role in genomic mutation [91]. Interestingly, DNA damage can cause cytoplasmic A3G to enter the nucleus, but evidence suggests that A3G cytidine deamination promotes DNA repair [92]. Previous studies have shown that A3 enzymes primarily deaminate genomic DNA during replication and favour the lagging strand due to the greater abundance of ssDNA from discontinuous synthesis [93–97]. Single-stranded substrates of A3 deamination include single-stranded intermediates at replication forks (figure 2$a$), DSBs (figure 2$c$) and those generated during break-induced replication (BIR) (figure 2$d$) [98]. Phosphorylation of the Ser-139 residue of the histone variant H2AX, forming $\gamma$H2AX, is an early cellular response to the induction of DSBs. Detection of this phosphorylation event has emerged as a highly specific and sensitive molecular marker for monitoring DNA damage initiation and resolution. The potential for A3 enzymes to cause DSBs has been discussed due to increased $\gamma$H2AX in response to ectopic A3 expression in cell lines [3,99]. The replication stress induced by hydroxyurea treatment of cancer cells expressing A3A led to increased $\gamma$H2AX [100] and A3B expression also caused induction of a DNA damage response characterized by $\gamma$H2AX and ssDNA-binding protein (RPA) phosphorylation [101]. A model was suggested in which BIR was observed to provide a substrate for clustered mutations during DNA replication [102,103]. During BIR, only one end of a DSB can be repaired and this uncouples the replication bubble so that the lagging strand is delayed behind the leading strand and ssDNA accumulates, providing a substrate for clustered mutations.

For enzymes that modify the DNA, processivity has always been thought of as essential for enabling deamination of multiple cytidines in a single enzyme–substrate encounter.
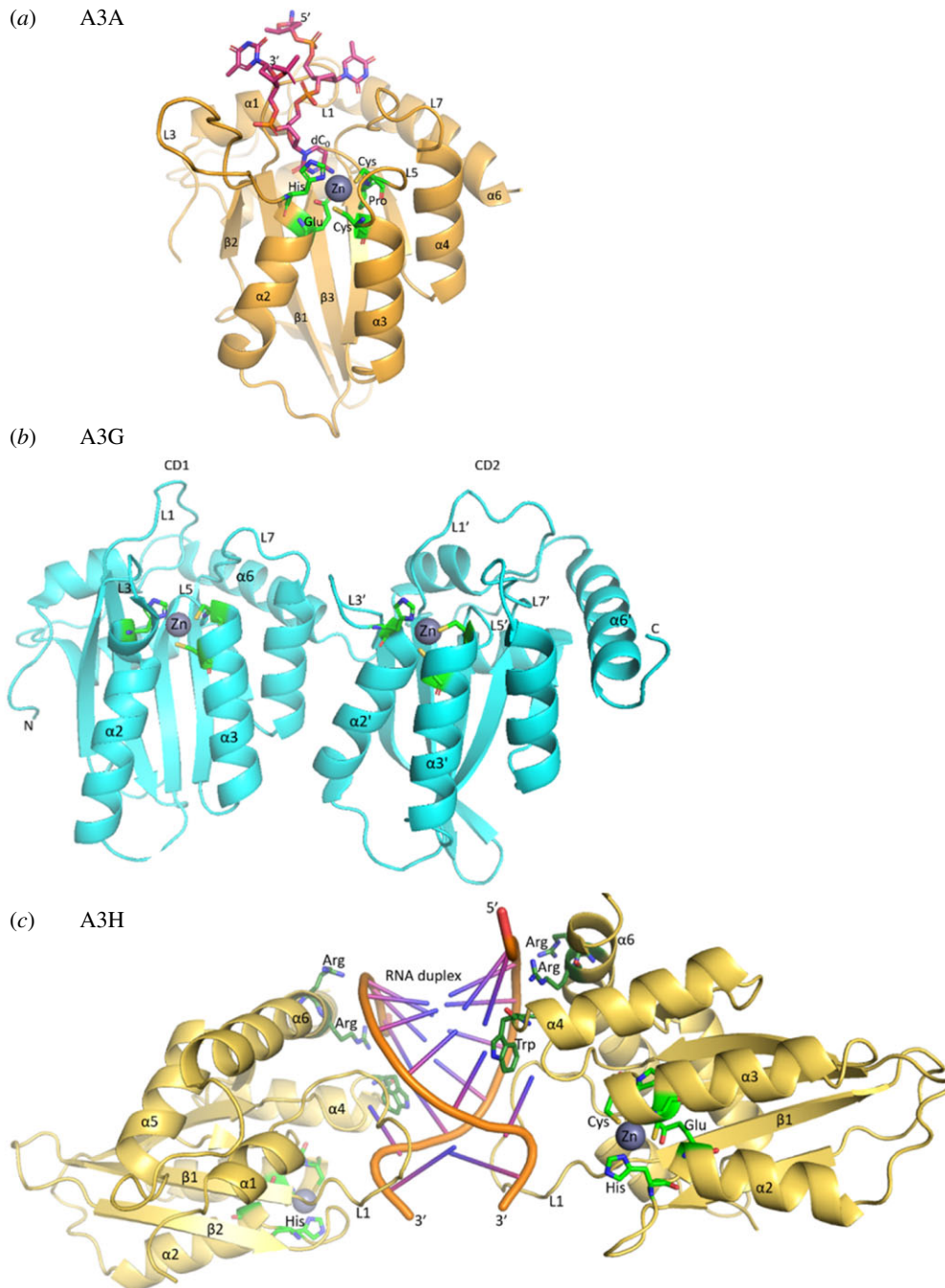
**Figure 1.** A3 structures. (*a*) The core domain and active site of A3s, showing the single-domain A3A as an example. PDB 5KEG shown with modification to include the catalytic glutamate on α-helix 2 (α2) instead of alanine [70]. The consensus sequence His-*X*-Glu-*X*$_{23-28}$-Pro-Cys-*X*$_{2-4}$-Cys is represented as green residues coordinating a Zn atom. Pink carbon atoms represent the four nucleotide DNA (5′- dT2, dT1, dC0, dT-1 -3′) substrate. Blue, nitrogen; red, oxygen; yellow, sulfur and orange, phosphorus. (*b*) The double-domain A3G (PDB: 6P40) with loop 7′ (L7′) being the principal determinant of sequence specificity in the C-terminal domain (CD2) and loop 7 (L7) being a principle determinant for processivity and oligomerization (not shown) in the N-terminal domain (CD1) [71]. In both the CD1 and CD2, loops 1 (L1, L1′), 3 (L3, L3′) and 5 (L5, L5′) also contribute to DNA binding. (*c*) A3H RNA-mediated dimer (PDB 5W3 V) [72]. The catalytic residues (light green) coordinate the zinc atom. Key residues lying on α-helix 6 (α6) and loop 7 (L7) electrostatically mediate the protein DNA interface (dark green). The highlighted Trp residue lies on L7 forming a critical stacking interaction with nucleotide 3 of the RNA duplex.

The loop 7 and/or helix 6 structures have been shown for all A3 enzymes to be important for processivity on and binding to ssDNA, with loop 7 being the principal determinant of sequence specificity, but loops 1, 3 and 5 also contributing to DNA binding in both single domain (e.g. A3A; figure 1*a*) and double domain (e.g. A3G; figure 1*b*) A3s [104–112]. The enzymes that do not use an energy source to move on DNA, such as APOBECs, can search for their DNA-specific target motif using a mechanism called 'facilitated diffusion' [113,114]. This mechanism includes sliding along DNA, microscopic dissociation–reassociation events between closely spaced sites (jumping or hopping) and intersegment transfer [113]. The 'facilitated diffusion' enables deamination of multiple closely spaced targets on DNA [115,116], but if an enzyme undergoes macroscopic dissociation from DNA after one turnover, it is considered to be non-processive or distributive. The processivity of several members of the AID/APOBEC family members on ssDNA has been reported and
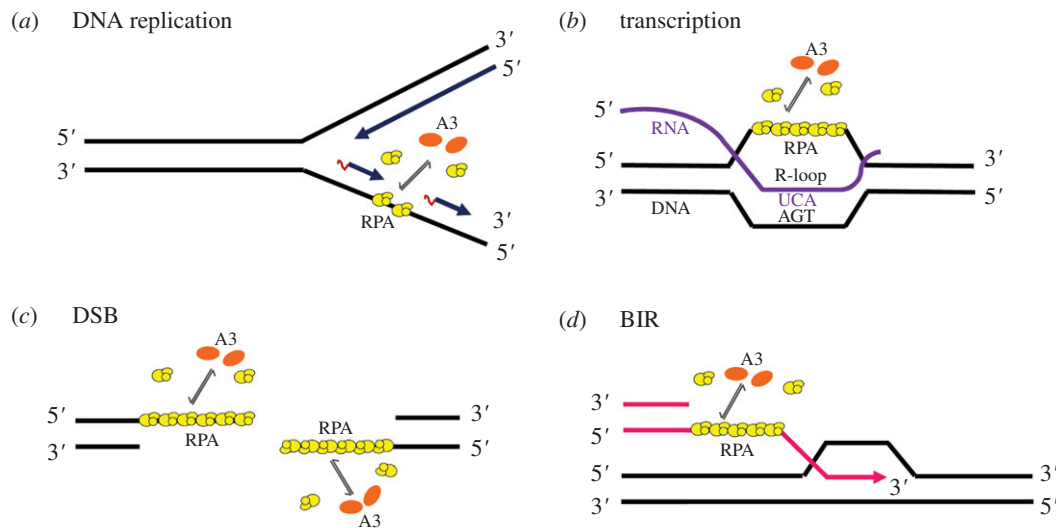
**Figure 2.** A3 deamination of cellular ssDNA templates. Single-stranded substrates of A3 deamination include ssDNA intermediates generated at replication forks (a), during transcription (b), at DSBs (c) and those generated during BIR (d). RPA (yellow) is bound to ssDNA and for A3 enzymes (orange) to access the ssDNA, they must be able to compete with RPA. The ssDNA mutagenesis during transcription can take place only if the lesions in the non-transcribed strand persist until DNA replication, then they can be fixed into mutations by TLS. By contrast, resolution of the R-loops provides the undamaged template for excision repair and prevents mutagenesis. For the mutagenesis associated with both DSB and BIR, there is no template for accurate excision repair. An A3 monomer is shown, although A3B forms larger oligomers and A3H is a dimer. Only A3A is monomeric.

the results have implications for the finding that mutations associated with AID or A3s are often clustered [117,118].

Adolph *et al.* in 2017 [119] reported the first biochemical characterization of A3B and compared properties of A3B, A3A and A3H Hap I on substrates relevant to catalysing cytidine deaminations in genomic DNA. They found, using *in vitro* transcription and replication model systems that surprisingly, A3A, a non-processive enzyme, was equally effective in deamination during replication as A3B and A3H Hap I that were processive enzymes. A3B and A3H Hap I cycled between ssDNA substrates, but also maintained processivity while bound to a substrate, whereas A3A simply cycled on and off substrates rapidly. The ability to cycle between ssDNA substrates was important to not only quickly access the transiently ssDNA being replicated but also ssDNA bound by RPA, suggesting that the ability to compete for ssDNA rather than processivity was most important for A3s to induce 'off-target' deaminations. In order to access the protected ssDNA, A3 enzymes would also have to displace RPA by a mechanism known as 'facilitated dissociation' [120,121], similar to how excess RPA or RAD51 was previously found to be able to exchange with bound RPA on the ssDNA. The *in vitro* study of A3 enzymes competing with RPA-saturated ssDNA for DNA binding showed that rapid cycling was required for RPA facilitated dissociation by A3A, A3B and A3H Hap I. By contrast, A3G, which is one of the most efficient enzymes for HIV-1 restriction, has a decreased frequency of cycling between ssDNA substrates and was greatly inhibited in the presence of RPA with a 10-fold decrease in specific activity, in contrast with A3A, A3B and A3H Hap I that had at most a twofold decrease in specific activity [119]. These results demonstrate that during replication stress where larger amounts of ssDNA accumulate, the protective RPA barrier is less effective against A3A, A3B and A3H Hap I.

The inherent properties of the A3s also determine if the enzymes deaminate the ssDNA during replication or

transcription, with transcription being more selective [119]. Although A3A, A3B and A3H Hap I could deaminate during DNA replication or on ssDNA bound by RPA, the amount of deamination correlated with the ability of the enzymes to cycle between ssDNA substrates [119]. The results obtained using *in vitro* transcription and replication model systems were in good agreement with previous studies done in yeast. For A3A, there was an association of mutations in a yeast system with the non-transcribed strand, but the majority of mutations correlated with the lagging strand of replication [96]. In the case of A3B, the mutations correlated with the lagging strand of replication, not the ssDNA on the non-transcribed strand generated during transcription [96], probably due to A3B oligomerization exceeding the size limitation of the transcription bubble [119,122]. On the basis of the available data, A3B would be unable to deaminate ssDNA generated during transcription, unless stable R-loops were to form [119]. The ssDNA mutagenesis during transcription can take place if the lesions in the non-transcribed strand persist until DNA replication, at which time they can be fixed into mutations by TLS and this would result in a mutation bias favouring the non-transcribed strand [98]. No cellular studies examining A3H Hap I in this regard are available, but a study using a bioinformatics approach has suggested that A3H Hap I could act early in lung cancer mutations and possibly contribute to the APOBEC signature in A3B-null BRCA [5]. This is in contrast with A3B-induced mutations that were absent early in tumour formation but then suddenly arose and were maintained at a high level. One limitation of the precedent studies in which transcription has not been observed as a dominant form of ssDNA substrate was that the end-point analysis from tumours or after months of expression of A3 enzymes in yeast takes into account only accumulated but not temporal mutations [93–97]. Whether A3s enzymes play a role in driving carcinogenesis and/or tumour progression needs further specific studies in each cancer type.

Determining how and when (during replication, transcription or repair) A3 enzymes can deaminate genomic ssDNA will allow us to identify if they cause cell transformation, contribute to a mutator phenotype or both.

# 5. A3 enzymes and cancer

Cancer is considered a group of diseases that involve dynamic changes in the genome. Genomic instability is one of the hallmarks of cancer that cause both aberrant chromosomal architecture and mutational changes at the single nucleotide level [1]. The diversity created by the above-mentioned processes provides the substrate for selection within tumours. Although an elevated mutation rate is not necessarily a requirement for the initiation of a tumour, it will likely contribute faster to tumour evolution and adaptation [123]. In addition to other endogenous mutational factors, now it is well known that some members of the A3 enzymes are an endogenous source of somatic mutations found in approximately 15% of sequenced human tumours such as, BRCA, bladder, cervix, lung (adenocarcinoma and squamous cell carcinoma), head and neck, myeloma, renal cell carcinoma, stomach and thyroid [3,5,11,12,24,32–38]. All cancers are caused by somatic mutations that are the aggregate outcome of one or more mutational processes operative through the life of the cancer patient [9,124]. Different mutational processes often generate different combinations of mutation types, termed 'signatures'. Each mutational process leaves a characteristic mutational signature determined by the mechanisms of DNA damage and repair. The APOBEC family of cytidine deaminases generates particular genome-wide mutational signatures and a signature of localized hypermutation called 'kataegis' or 'mutation clusters' [9,10,125,126]. Two signatures characterized by C-to-T and/or C-to-G mutations at TpCpX trinucleotides were identified (the underlined base is the mutated base and X can be any base) in several cancer types and are among the most common mutational signatures found in human cancer [9,11]. These signatures have been designated Signatures 2 and 13 [11]. Signature 2 is composed predominantly of C-to-T transitions with fewer C-to-G transversions and Signature 13 is dominated by C-to-G transversions at a TpCpX sequence context and due to error-prone repair of APOBEC-induced uracils [9,11].

APOBEC deoxycytidine deaminases are considered the second-most prominent source of mutagenesis in sequenced tumours, next to mutations caused by ageing [11–13,24]. The association between A3 enzymes and carcinogenesis is evidenced from multi-dimensional observations. Human A3s are mainly studied using primary or immortalized cells of human or non-human origin, and the results are often cell line-specific. Also, biochemical and bioinformatic studies have provided information about the role of the A3 enzymes in the cancer field. One important observation regarding the studies of these enzymes in cancer was that the *in vivo* evaluation using a complex mammalian system was lacking in the field. Previously, no authentic animal model for the studies of individual A3 genes and proteins was reported because rodents have a single A3 gene, while humans have seven. However, the role of the transgenic expression of human A3 enzymes in two models for tumourigenesis in mice was reported recently [127], providing some *in vivo* validation to the current knowledge. In developing these two transgenic murine tumour models, it was demonstrated *in vivo* that the human A3A enzyme catalyses mutagenesis and promotes tumourigenesis in colorectal and hepatocellular carcinoma [127].

The 'off-target' mutations in the host genome produced by members of the A3 family, that are the basis of its relationship to cancer, has been associated with cancer development, progression, metastasis and drug resistance [31]. Most cancer cells and tumours show overexpression (20- to 60-fold) of A3B, A3A or A3H Hap I mRNA [3,5,11,12,128,129]. The efforts in characterizing the potential mutagenic activity of APOBECs have focused primarily on APOBEC mRNA expression levels, but sometimes the reported enrichment of the mutagenesis associated with a particular A3 enzyme (e.g. A3A) does not correlate with mRNA levels in tumours from the TCGA database [4] raising the possibility that expression levels alone do not determine A3 activity [29]. A key factor limiting comparisons between mutational burdens and mRNA abundance are potential differences in APOBEC expression levels at the time of mutagenesis and at the time of RNA sampling. Mutations captured in cancer genomes could have been generated by APOBEC deaminases over the lifetime of a cell lineage, whereas mRNA captures expression at the single time point of sample acquisition and not necessarily at the time of active mutagenesis [130]. Also, there is a possibility that signals of APOBEC expression in tumours originate from infiltrating immune cells with naturally higher APOBEC levels [131]. This review summarizes specific information on the leading A3s candidates (A3A, A3B and A3H Hap I) causing mutation signatures in different human cancers [3–5,13,24,132–135] taking into account both original ideas that led to the identification of APOBEC activity in cancers and the current views on A3 expression levels and levels of mutagenesis.

## 5.1. Breast cancer

BRCA is a leading cancer burden in females and the primary cause of cancer-associated deaths among women worldwide [136]. The probability of developing BRCA is modulated by the interaction of many factors such as lifestyle, environmental and genetic factors. Mutations are thought to be the key drivers of recurrence, metastasis and therapeutic resistance of cancer. The studies on the molecular origins of mutations in BRCA have implicated several mechanisms, including both spontaneous and enzyme catalysed deamination of DNA cytidine [3,9,11–13]. The former process correlates with ageing and is mostly due to hydrolytic conversion of 5-methyl cytosine (mC) bases within 5′ NmCG (N = A, C, G, or T) motifs into thymines, which escape BER and are converted into C-to-T transition mutations by DNA replication. The latter process is attributable to ssDNA deamination catalysed by one or more members of the A3 family of enzymes, characterized by C-to-T transitions and C-to-G transversions in 5′HTCW motifs (H = A, C, or T and W = A or T). The extensive resources generated by the TCGA database enabled a deep analysis of somatic mutagenesis in breast tumours [137]. Based on the patterns of mutations, at least 12 somatic mutational signatures have been annotated in TCGA breast tumours [138], two of these have been attributed to the activity of APOBEC family of proteins [11,12]. In particular, A3A and A3B have been considered the main mutagenic enzymes that generate APOBEC-signature mutations in breast and other tumour types because overexpression of these enzymes triggers DNA damage responses
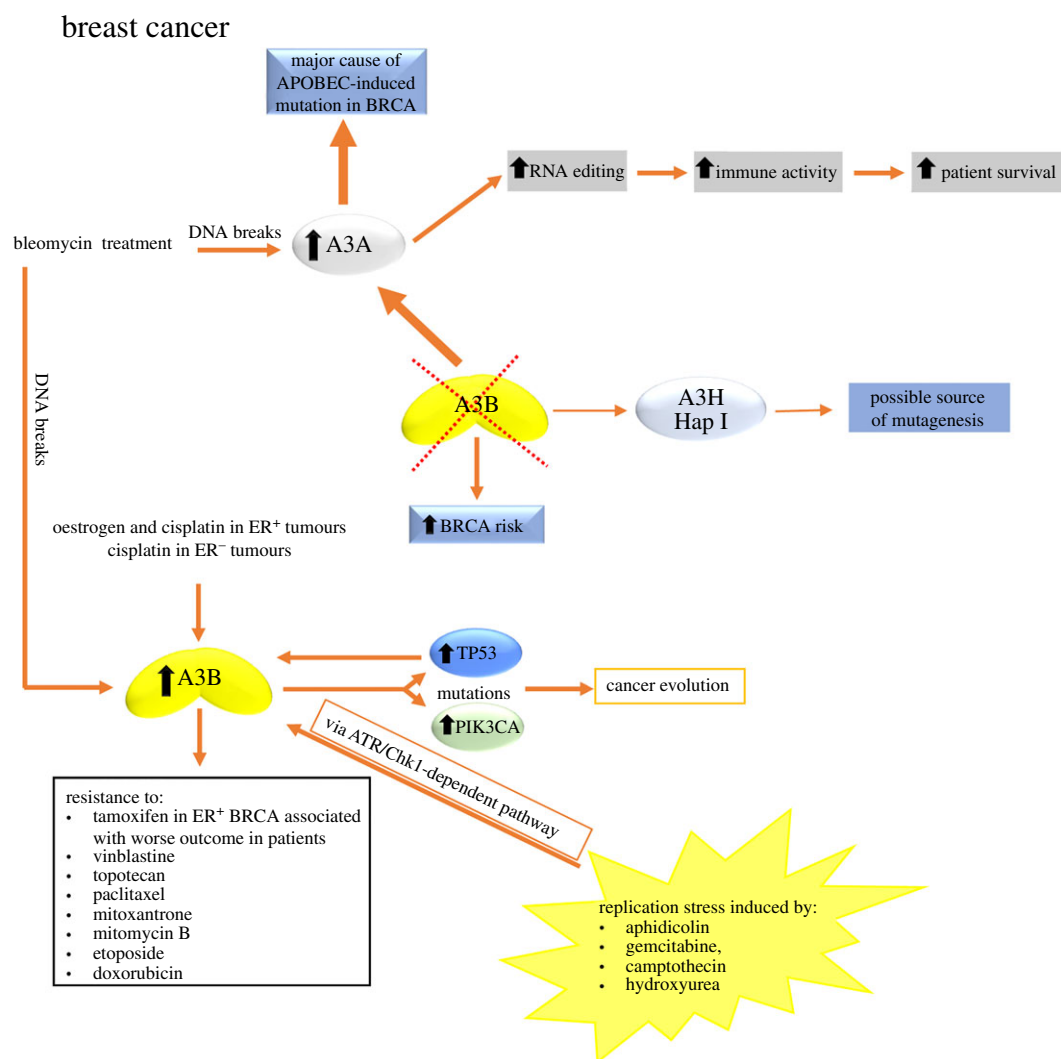
breast cancer



**Figure 3.** Causes and effects of A3 upregulation in BRCA. Causes of A3B upregulation could include the action of hormones (oestrogen), DNA-damaging drugs (cisplatin, bleomycin) and replication stress. The induction of A3B was found to be stronger in the presence of mutant TP53 compared to wild-type TP53. A3B upregulation contributes to increase mutations in TP53 and PIK3CA genes driving to cancer evolution and A3B upregulation is also associated with worse outcomes in cancer patients as well as with resistance to many anti-cancer drugs. In the absence of A3B, there is an increased risk to develop BRCA. A3A has emerged as a major cause of APOBEC-induced mutations in BRCA. Also, in the absence of A3B, A3H Hap I could be a source of mutagenesis. Bleomycin treatment can also upregulate A3A, although the effect is more robust for A3B. The upregulation of A3A is also associated with an increased RNA-editing activity that positively correlated with patient survival. For more details, see §5.1.

and inflicts chromosomal mutations in hallmark trinucleotide contexts [3,4,9,134,139]. When introducing A3A and A3B into yeast, genome-wide mutation patterns of A3A- and A3B-mediated deamination show strong similarity to mutation signatures found in BRCA, which strengthens the proposed role of A3A and A3B in BRCA hypermutation [118]. The analysis of the APOBEC-signature mutation load in cancer exons showed that it is statistically correlated with A3A and A3B transcript abundance [24]. Although A3B mRNA abundance tends to be greater than that of A3A in BRCA cancer samples, A3A is a more potent inducer of DNA damage [24]. Due to the nonlinear relationship between mRNA abundance and activity levels and different studies relying on often one or the other as a measure of A3 activity, the relative contributions of A3A and A3B to mutagenesis in BRCA cancer have been extensively debated in the literature. Collectively, the data point to both A3A and A3B being involved in BRCA, but with distinct mechanisms and effects.

Multiple lines of evidence suggested that A3B, the only constitutively nuclear ssDNA deaminase [91], was the primary source of the mutations found in BRCA [3,11–13]. The reasoning behind these suggestions included the following observations: A3B is overexpressed in greater than 50% of breast tumours; in more than 75% of BRCA cell lines, and it was the only detectable DNA deaminase activity in BRCA cell extracts [3,13]; increased A3B levels correspond positively with overall cytosine mutational loads [3] and A3B expression associates with worse clinical outcomes of hormone therapy resistance in BRCA [140,141]. The analysis of cell line and tumour datasets showing that A3B gene expression is upregulated in malignant versus normal tissues and epithelial cell lines have shown correlations between A3B expression and the presence of certain somatic mutations, particularly in TP53 [3,142] and phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PIK3CA) (figure 3) [143]. These observations form the basis for a model where A3B expression contributes to the accumulation of somatic alterations during the process of carcinogenesis and subsequent evolution, and it has been suggested that inhibition of this activity could represent a strategy for

cancer prevention or an adjuvant to other therapies [3,144]. The report that A3B expression is associated with adverse outcomes in oestrogen receptor-positive (ER⁺) BRCA would be consistent with this hypothesis (figure 3) [141] as well as the analysis of 30 human cell lines, including BRCA cell lines, from which expression levels of the A3B gene were associated with resistance to anti-cancer drugs such as vin-blastine, topotecan, paclitaxel, mitoxantrone, mitomycin C, etoposide and doxorubicin (figure 3) [145].

High A3B expression in ER⁺ BRCA showed short progression-free time with tamoxifen treatment and suppression of endogenous levels of A3B enhanced tamoxifen benefit [140]. The mitogenic effect of oestrogen was supported by observations of increased cellular proliferation induced by A3B expression in BRCA cells [146]. Recently, Udquim et al. [147] provided some elements on the understanding of the aetiology and clinical outcomes of BRCA supporting the mitogenic hypothesis of oestrogen action independent of somatic mutagenesis (mutagenic hypothesis). Their analysis in BRCA cell lines and TCGA breast tumours suggest that A3B expression is induced by oestradiol in an ER-dependent way. They proposed that A3B expression is unlikely to account for APOBEC-mediated mutagenesis in breast tumours but might contribute to cancer development based on the possible mitogenic effect of A3B, a deamination-independent effect. The promotion of cancer progression by oestrogen-induced A3B expression affecting the tumour microenvironment in ER⁺ cells would be consistent with the proposed mitogenic effect of oestrogen [148] that is also supported by observations of increased cellular proliferation induced by A3B expression in BRCA cells [146]. The expression of A3A in the BRCA cell lines evaluated was not detectable, suggesting that A3B is an oestrogen-responsive gene and A3A is not. The expression of A3B was also induced by a DNA-damaging chemotherapy drug, cisplatin, regardless of the ER status (figure 3). Accordingly, treatment with cisplatin in women with high levels of endogenous oestrogen or receiving hormone replacement therapy may result in a further increase of A3B expression (figure 3). However, in women with oestrogen receptor-negative (ER⁻) breast tumours, endogenous oestrogen or hormone replacement therapy may not affect A3B expression, even during cisplatin therapy. p53 controls the A3B expression [149,150] and since ER⁻ tumours are enriched for inactivating p53 mutations [137,151], this could contribute to elevated levels of A3B expression in this type of tumour that is significantly higher compared with ER⁺ breast tumours. The study of Udquim et al. [147] has not addressed the direct causal connection between exposure to endogenous and exogenous oestrogen, A3B expression in human samples and BRCA risk. To address these relationships, it would be necessary to conduct epidemiological studies with large cohorts of BRCA patients with tumours of different subtypes taking into consideration all relevant covariates such as environmental exposures, germline and somatic variants, and clinical outcomes.

An interesting concept is that DNA damage/replication stress can induce expression of A3 genes. Expression of A3s in BRCA cell lines is inducible in response to various environmental exposures (figure 3) [95,134,147], as addressed in the study of Udquim et al. [147] demonstrating the A3B overexpression in response to cisplatin treatment. In addition, the issue of DNA damage/replication stress was addressed by

Middlebrooks et al. [134] by treating three BRCA cell lines (MCF-7, MDA-MB-231 and T-47D) with bleomycin, a DNA-damaging drug known to induce DNA breaks. The study found that both A3A and A3B were induced in most cell lines, with the effect being more robust for A3B (figure 3). The induction of A3B was found to be stronger in cell lines harbouring mutant TP53 than in those with wild-type TP53 (figure 3) [134]. Following the same principle, Kanu et al. [95] demonstrated that chemical and cytotoxic induction of replication stress, through aphidicolin, gemcitabine, camptothecin or hydroxyurea exposure activates transcription of A3B via an ATR/Chk1-dependent pathway in vitro (figure 3). The role of DNA replication stress in mediating genomic instability could link the high level of somatic copy number aberrations and single nucleotide diversity caused by A3 activity that are both observed in human epidermal growth factor receptor 2 (HER2)⁺ tumours.

Within breast carcinomas, the HER2⁺ subtype has been shown to display evidence of A3-mediated mutagenesis and is associated with high levels of somatic copy number aberrations [12]. This subtype of breast tumour displays evidence of elevated levels of replication stress-associated DNA damage in vivo. The oncogenic signalling, cytotoxic drugs and genetic modulators of replication stress are all able to modulate A3 activity, although it has not been explored if there is a mechanistic connection between the underlying causes of chromosomal copy number aberrations and the generation of A3 mutagenesis in HER2⁺ BRCA [95]. These findings implicate the ability of therapeutics that either attenuate oncogenic signalling or exacerbate DNA replication stress to alter cancer's mutagenic landscape and evolutionary potential by multiple mechanisms.

Despite all aforementioned findings, the importance of A3B in cancer has been questioned with the observation that APOBEC-signature mutations are still clearly evident in A3B-null breast tumours (figure 3) [152]. A 29.5 kb deletion that removes the entire A3B coding sequence and fuses the 3′ untranslated regions (UTR) of A3A and A3B forms a hybrid gene that is predicted to produce a transcript which is predominantly constituted of A3A sequence but replaces the A3A 3′UTR with the A3B 3′ UTR and encodes a protein that has an identical amino acid sequence to A3A [153]. This polymorphism occurs at different frequencies in different populations around the world. The deletion allele has a frequency of approximately 8% in European populations [153,154], 37% in East Asians and 93% in Oceania [153]. Some studies showed that the A3B deletion increases the risk of BRCA [155,156] and increases tumour mutational burden [152]. A genome-wide association study in the Chinese population demonstrated the A3B deletion is associated with BRCA (odds ratio (OR) 1.3 one-copy, 1.8 two-copy deletion, $p = 2.0 \times 10^{-24}$) [155], which was replicated in a European population (OR 1.2 one-copy, 2.3 two-copy deletion, $p = 0.005$) [156]. However, including familial BRCA for the first time, a later study showed a lack of association of the A3B deletion with BRCA risk, which was independently validated in three European cohorts (in total: 2972 cases and 3682 controls) [157]. This study provided direct evidence for the generation of the transcriptionally active hybrid gene A3A/A3B from the allele with the A3B deletion and confirmed the suggested structure of A3A/A3B transcript, which enabled A3A, A3B and A3A/A3B expression levels to be distinguished. The knowledge of the exact structure of

the hybrid transcript is vital for the design of comprehensive tests for analysis of the influence of the A3B deletion genotype on the expression of A3B, A3A and the A3A/A3B hybrid gene. A recent study showed that the germline A3B deletion influenced the APOBEC mutational signature, neoantigen loads and relative immune cell compositions in BRCA [36]. This study [36] and recent studies demonstrating that APOBEC plays an important role in promoting programmed cell death protein-1 (PD-1) expression [133,146,158,159], as well as immune activation in multiple cancer types, highlight the importance of the APOBEC genes in immunogenicity and cancer immunotherapy.

As the homozygous carriers of the A3B deletion allele are predicted not to make any A3B protein, other APOBEC enzymes must contribute to APOBEC-signature mutations during tumour development. In this regard, one study of APOBEC-induced mutations from A3B deleted BRCA tumours revealed that the only tumours displaying the APOBEC mutation signature also contained the nuclear A3H Hap I, providing correlative evidence that this protein may be the additional source of mutagenesis (figure 3) [5]. As well as for A3H Hap I [5], some evidence has implicated A3A in A3B-null BRCA (figure 3). The A3A/A3B hybrid mRNA can be expressed in the A3B-null BRCA cell line SKBR3 and this finding has suggested that A3A can contribute to a mutator phenotype in cancer (figure 3) [4,160]. The initial studies regarding A3A were done in cell lines, yeast cells or by retroactively analysing cancer genome databases for mutations without specifically examining the A3 expression patterns in A3B-null tumour cells [4,118,152,160]. Another study found the A3A mutational footprints in tumours, but no corresponding A3A expression and suggested that A3A is upregulated early, but later inactivated, perhaps due to being the most active deaminase that could cause cell death through its activity over time [4,99,161]. In addition, A3A and A3B can be differentiated by their different preferred tetranucleotide motifs, 5′YTCA and 5′RTCA (Y = T or C and R = G or A), respectively, when inducing mutations in a yeast model system [4]. However, over-representation of mutations in the 5′YTCA motif predominates in a variety of cancers [4] as well as among mutations actively acquired in BRCA cell lines [130], suggesting A3A may likewise contribute to cancer mutagenesis. A3A was often reported to have undetectable expression in BRCA lines [3,5], but more recently, some evidence indicates that A3A may be a major cause of APOBEC-induced mutation in BRCA and account for the majority of cytidine deaminase activity in extracts from multiple BRCA cell lines, despite higher A3B expression [24]. Consistent with previous reports [3,12,13,162], this study found that A3B was expressed at high levels compared to the other A3s. However, this high A3B expression level existed in both APOBEC-mutagenized and non-APOBEC-mutagenized cell lines [24], indicating that elevated A3B mRNA levels may not be directly responsible for APOBEC-induced mutagenesis as previously thought [3,12,13]. By contrast, a median 13.1-fold higher A3A mRNA expression level was observed in the APOBEC-mutated BRCA lines compared to non-APOBEC-mutated lines and the overall abundance of APOBEC-induced mutations linearly correlated with A3A expression [24]. Also, they demonstrated that A3A is the primary source of cytidine deamination activity in A3B-null AU565 and SKBR3 cell lines. A3A is more active biochemically than the

next most potent somatic mutators A3B and A3H Hap I [24,79,119,163,164] and RNA binding is known to inhibit the activity of A3B and partially inhibit the activity of A3H, but not A3A [24,84,86], suggesting that A3A may be a better candidate than A3H Hap I in causing the APOBEC mutation signature in A3B null BRCA [24]. In the presence of cellular RNA, A3A contributes significantly to cytidine deaminase activity, even from extracts of BRCA cell lines with elevated A3B expression and containing A3H Hap I, indicating that in some cellular contexts, A3A is the dominant active APOBEC present in the cell. The mechanism for relieving RNA inhibition of A3B or A3H Hap I in cells is currently unknown and the A3H Hap I deaminase activity in A3B-null cells, AU565 or SKBR3, (that have at least one A3H Hap I allele) was not detected [24].

In addition to mutagenesis linked to deamination of ssDNA, A3A as well as A3B have been reported to be involved in RNA editing [19,165]. Like its preference for DNA stem-loops, A3A also recognizes RNA stem-loops in a sequence-specific manner [19]. Since A3A is much more catalytically active than A3B, but it is typically expressed at lower levels than A3B in tumours, both A3A protein and A3A mRNA are difficult to quantify. This makes it particularly challenging to predict the levels of currently ongoing A3A activity in tumours. Recently, Jalili *et al.* [166] developed a strategy using hotspot APOBEC-signature mutations in RNA stem-loops identified from A3A-positive tumours and droplet digital PCR to quantify the ongoing activity of A3A in tumours. They found that A3A expression and A3A-mediated DNA mutagenesis in tumours, but not those of A3B, correlate with APOBEC-signature mutations in RNA stem-loops. Interestingly, those RNA mutations are not present in their DNA templates, suggesting that they are directly generated by A3A. Because of the labile, transient nature of RNA, they concluded that the RNA-editing activity of A3A accurately reflects the currently ongoing activity of A3A and that the RNA mutation-based A3A assay is superior to A3A protein- and mRNA-based assays in predicting the currently ongoing A3A activity on DNA. Finally, they show that the RNA mutation-based A3A assay can be applied to clinical samples from cancer patients, providing a new opportunity to investigate the role of A3A in tumour evolution and to target A3A-induced vulnerabilities in cancer therapy [166].

In line with this RNA-editing activity of the A3 enzymes, a recent bioinformatic study identified that A3-mediated RNA editing occurs in breast tumours and is positively associated with elevated immune activity and improved survival (figure 3) [22]. Interestingly, the RNA-editing scores had the best correlation with A3A gene expression [22]. The findings of this study imply that A3 enzymes are relevant in BRCA not only because of their DNA mutagenicity but also their RNA-editing activity, and they highlight the pertinence of finer dissection of such editing in further studies. While they could detect C-to-U RNA-editing events in the tumours, the cellular origin of such events remains unclear. Overrepresentation of edited sites among immune-related genes, enriched expression of such genes in editing-high tumours, and the fact that A3 gene expression is much higher among immune compared to epithelial cells suggests that the editing occurs in the tumour microenvironment. Further investigations implementing methods such as single-cell sequencing and isolation of sub-populations of cells from tumours are needed to definitively know if the editing
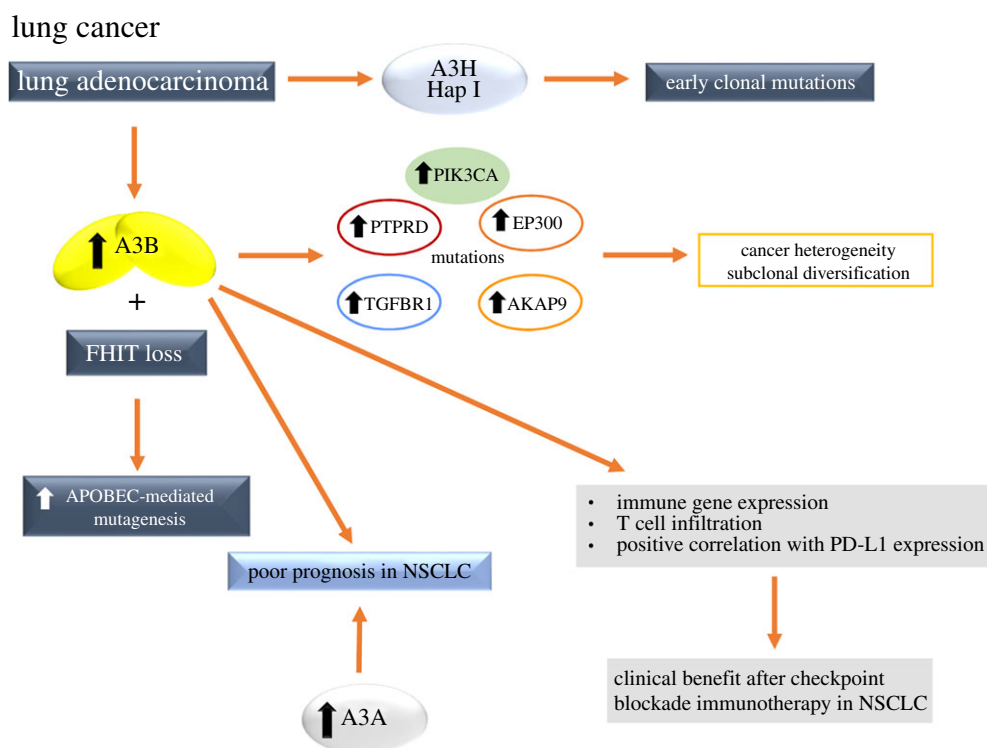
**Figure 4.** Causes and effects of A3 upregulation in lung cancer. Mutations within an A3B context were found in driver genes such as PTPRD, PIK3CA, EP300, TGFBR1 and AKAP9, fuelling cancer heterogeneity and subclonal diversification. In lung adenocarcinoma, A3H Hap I expression was associated with early clonal mutations and increased A3B expression and the loss of FHIT protein expression was associated with higher levels of APOBEC-mediated mutagenesis. The overexpression of both A3B and A3A is associated with poor clinical outcomes in NSCLC patients, but A3B overexpression also predicts clinical benefit after checkpoint blockade immunotherapy in patients. For more details, see §5.2.

occurs in cancerous epithelial or immune cells of breast tumours. The biological consequences of the editing events on cancer development, progression and immune response also remain unknown.

Based on the literature, the existing data have led to an unclear understanding as to the relative contributions of individual APOBECs to mutagenesis in BRCA. A side-by-side comparison among these studies is difficult because they were not done using the same cell lines or under the same conditions. Also, it is important to elucidate which of the enzymes A3A, A3B and A3H Hap I are involved earlier in promoting cell transformation or later in promoting BRCA progression. In this regard, mechanistic studies to determine if those enzymes can deaminate the ssDNA during replication, transcription and/or DNA repair will allow us to understand better their role in BRCA.

## 5.2. Lung cancer

Lung cancer is the most common cancer in men and the third most commonly occurring cancer in women, with 2 million new cases in 2018 [167]. There are two main types: small cell lung cancer (SCLC) and non-SCLC (NSCLC). It is estimated that about 80–85% of lung cancers are NSCLC, and about 10–15% are SCLC. Also, there are three main subtypes of NSCLC: adenocarcinoma, squamous cell carcinoma and large-cell carcinoma [168]. In addition to smoking, one major cause of the heavy mutation load of NSCLC [169], the expression of APOBEC family members, especially A3B, was reported as a key source of mutations specifically in two subtypes of NSCLC: adenocarcinoma and squamous cell carcinoma [13].

A3 mutational signatures may occur at different stages in different types of cancer, contributing to later subclones in lung adenocarcinoma as the tumours evolve [170]. In an analysis of intratumour heterogeneity in early-stage NSCLC, multiregion sampling allowed the timing of mutational processes during tumour evolution to be deciphered, establishing the temporal dynamics of APOBEC mutational processes [170,171]. The enrichment of the APOBEC signature was seen in the branches of tumour evolutionary trees relative to the early clonal truncal mutations, particularly prominent in adenocarcinomas of the lung [170,171]. According to this endogenous mutational process driving subclonal expansions, mutations within an A3B context were found in driver genes such as PTPRD, PIK3CA, EP300, TGFBR1 and AKAP9 (figure 4) [170]. Also, there was evidence for spatial heterogeneity in APOBEC activity; in one adeno squamous tumour, the APOBEC signature was found enriched in the adenocarcinoma branch, harbouring driver mutations in PTPRD and TGFBR1 within an APOBEC context, but not the squamous carcinoma branch [170]. Notably, in lung adenocarcinoma and lung squamous cell carcinoma, over 85% of subclonal mutations in PIK3CA occurred in an APOBEC context [172]. Most of these subclonal mutations were found in the PIK3CA helical domain (E545 K) that have been previously linked to APOBEC-mediated mutagenesis in cervical and head/neck tumours [143]. These data highlight that both the genome instability processes and the APOBEC mutagenic process can be spatially and temporally heterogeneous during the disease course. The importance of APOBEC later in tumour evolution is highlighted by the observation that this mutational process in lung adenocarcinoma and squamous carcinoma was found to be the major source of subclonal cancer gene mutations

(figure 4) [172] relative to clonal driver gene mutations, suggesting APOBEC is a mutagenic source, fuelling cancer heterogeneity and subclonal diversification. In addition to A3B, a bioinformatics approach significantly associated the cytidine deaminase A3H Hap I with clonal APOBEC-signature mutations in lung adenocarcinoma (figure 4) [5]. Later, a computational study supported this idea and identified the association of SNP rs139298, that is correlated with lung cancer and creates a K121E mutation in A3H Hap I [173]. As a follow-up to this study, the effect of the K121E mutation was assessed. First, the ability of A3H Hap I to induce DNA damage in lung cells was for the first time directly demonstrated by observing A3H Hap I-induced $\gamma$H2AX foci [174]. However, the K121E mutation was shown to destabilize A3H Hap I in cells and supported the conclusion that the loss of A3H Hap I activity through the K121E variant may benefit the cancer and be detrimental to the host, suggesting that A3H Hap I deamination activity can induce tumour cell death or immune recognition [174]. These data emphasize that it is important to use additional genetic or clinical data for determining the beneficial or detrimental effects of A3-induced mutations.

APOBEC-mediated mutagenesis may also be increased in the case of a reduced expression or the loss of protein activity of the tumour suppressor fragile histidine triad protein (FHIT), and higher levels of APOBEC-mediated mutagenesis were observed from TCGA lung adenocarcinoma tumours that had both increased A3B expression and the loss of FHIT protein expression (figure 4) [134,175,176]. FHIT is frequently lost very early during tumour development, causing replication stress due to deoxythymidine triphosphate depletion [177], thus the mutagenic potential of A3B may be unleashed in the absence of FHIT [176]. When genomic sequences from lung adenocarcinomas were stratified by A3B and FHIT expression, those with high A3B and FHIT loss showed significantly higher A3 signature mutation loads than high A3B expressers with normal FHIT levels [176].

Elevated A3B expression was reported to correlate with poor prognosis in lung cancers and other types of cancers, pointing to A3B as the key mutation driver in human cancers [34]. Without including A3-induced mutations, an analysis of chromosomal instability (CIN) quartiles in NSCLC has revealed that intermediate thresholds of CIN appear to exist and correlate with the poorest clinical outcomes, in contrast with excessive or minimal CIN in cases where the outcomes are better [178,179]. Also, evidence of preferential benefit from therapeutic approaches has emerged in patients with tumours with the highest mutational load [180]. This suggests that the evolutionary trade-off for increased fitness brought about by an increased mutation rate is the risk of tumour neo-antigenic presentation and immune control (figure 4) [181,182]. A3 enzymes are an important part of this balance, the effects of which appear to be specific to the tumour [34,180].

A3B-related mutational processes fuelling cancer heterogeneity and treatment resistance remains a challenge for NSCLC treatment. Although A3B upregulation is associated with poor NSCLC prognosis [34], targeting A3B in this cancer is still a big challenge. It was proposed that cancer progression would be suppressed through A3B inhibition, but there are no available drugs that can inhibit A3B expression or function. NSCLC is commonly treated by radiation, surgery and chemical therapy, but the approval of the antibodies targeting immune checkpoints PD-1 and ligand (PD-L1) had a big impact for the immunotherapy of this and other cancers. It

was demonstrated in advanced NSCLC, in patients treated with an antibody targeting PD-1, response rates of 17–21% with some responses being remarkably durable [183]. Although clinical studies have shown promise for targeting PD-1, PD-L1 signalling in NSCLC, the factors that predict which patients will be responsive to checkpoint blockade are not fully understood. Based on existing publications, the predictive markers for immune checkpoint inhibitor therapy include: PD-L1 expression [184,185], tumour mutational load [181,182], DNA mismatch repair deficiency [186] and CD8+ T-cell activity [187,188]. A recent study was focused on the correlation of A3B expression with immune gene expression and immunotherapy response in NSCLC [133]. Although A3H Hap I has been implicated in lung cancer [5,173,174], this study stated that similar to A3B, overexpression of A3A, but not A3H, predicts poor NSCLC prognosis (figure 4) [133]. Thus, A3 family members can play distinct and overlapping functions in NSCLC. Through combined cancer genomic mutation analysis and gene expression analysis, A3B upregulation is significantly associated with immune gene expression, and A3B expression positively correlates with known immunotherapy response biomarkers, including PD-L1 expression and T-cell infiltration in NSCLC (figure 4) [133]. The APOBEC mutational signature is specifically enriched in patients with durable clinical benefit after immunotherapy and APOBEC mutation count can be better than total mutation count in predicting immunotherapy responses. This study implicates A3B and APOBEC mutational signatures as novel predictive biomarkers for checkpoint blockade immunotherapy response in NSCLC and suggests immunotherapy as a novel treatment option for A3B overexpressing NSCLC [133].

## 5.3. Bladder cancer

With almost 550 000 new cases in 2018, BLCA is considered the sixth most commonly occurring cancer in men and the 17th most commonly occurring cancer in women [167]. The risk factors for BLCA include smoking, physical inactivity, unhealthy nutrition, schistosomiasis, occupational exposure to aromatic amines, hair dye and contaminants in drinking water [189]. Urothelial carcinoma is the most common type of BLCA while squamous cell carcinoma, adenocarcinoma, small cell carcinoma and sarcoma are the less common types. BLCA is also described as non-muscle-invasive or muscle-invasive (MIBC), depending on whether it has grown into or through the muscle of the bladder wall. About 75% of patients have non-muscle-invasive BLCA and 25% have MIBC or metastatic disease. Approximately 50% of non-muscle-invasive BLCA are low grade, whereas most MIBC or metastatic tumours are high grade [190].

APOBEC mutagenesis is the predominant mutational pattern in BLCA [11–13]. About 80% of bladder tumours in the TCGA have an APOBEC mutation signature that is also frequently found in BRCA, lung, head and neck, and cervical cancers [3,11–13,134,191]. Although several studies have been focused on the linkage of A3B expression with mutagenesis [3,13,101], its expression alone does not fully explain the APOBEC mutational signature, and A3A can play a significant role as recently demonstrated for BRCA [4,24] and also for BLCA [36,134,135]. Middlebrooks et al. in 2016 [134] demonstrated that expression of both A3A and A3B can be induced in BLCA cell lines (HT-1376, HTB-9 and RT-4) that represent some of the major clinical subtypes of bladder tumours by
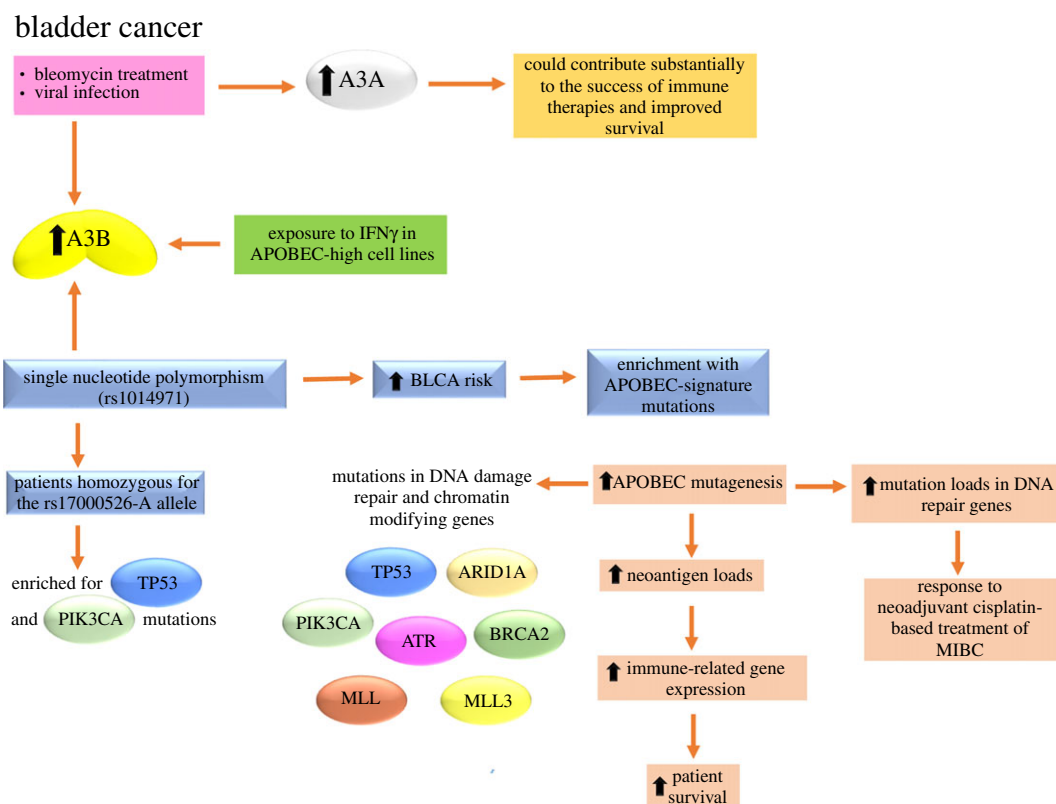
**Figure 5.** Causes and effects of A3s upregulation in BLCA. The expression of both A3A and A3B can be induced by bleomycin, a DNA-damaging agent and by an RNA virus that induces an IFN response. Also, the direct exposure to IFNγ in APOBEC-high cell lines increases the A3B expression. Hypermutation in BLCA (mainly by A3A) could contribute substantially to the success of immune therapies and improved survival. The single nucleotide polymorphism, rs1014971, was associated with BLCA risk, increased A3B expression and enrichment with APOBEC-signature mutations in bladder tumours. The tumours from patients homozygous for the rs17000526-A allele were enriched for TP53 and PIK3CA mutations. Increased APOBEC mutagenesis in BLCA patients have a positive association with neoantigen loads, the relative abundances of immune-related genes and the improved survival. Also, increased mutation loads, especially in DNA repair genes, were associated with a response to neoadjuvant cisplatin-based treatment of MIBC. Tumours enriched for APOBEC mutagenesis had better survival and were more likely to have mutations in both DNA damage repair and chromatin-modifying genes such as TP53, PIK3CA, ATR, BRCA2, MLL, MLL3 and ARID1A. For more details, see §5.3.

bleomycin, a DNA-damaging agent and by an RNA virus that induces an interferon (IFN) response [134] (figure 5). Both A3A and A3B were induced in all cell lines by bleomycin, but the effect was more robust for A3B. By contrast, A3A expression was uniformly induced four- to 167-fold by viral infection with in BLCA cell lines. This range of induction suggests additional cell-type-specific factors that may affect sensitivity to different environmental exposures [134]. The analysis of TCGA BLCA patient datasets revealed that a single nucleotide polymorphism, rs1014971, but not the germline A3A/A3B deletion [36], was associated with BLCA risk, increased A3B expression and enrichment with APOBEC-signature mutations in bladder tumours (figure 5). Also, this group demonstrated that TCGA BLCA patients with increased APOBEC mutagenesis had significantly improved survival, and that the tumours from patients homozygous for the rs17000526-A allele were enriched for TP53 and PIK3CA mutations [134] (figure 5). The more efficient immune surveillance due to neoantigens and synthetic lethality of tumour cells could contribute to improved survival for patients with higher APOBEC mutagenesis. In line with this observation, other groups reported a significant positive association between APOBEC mutational signature and neoantigen loads in BLCA [36,192] that also has an association with the relative abundances of immune-related gene expression [36,135] (figure 5). When evaluating the effects of all A3 isoforms on survival, the effect of A3B expression was comparable to that of

rs17000526 and similar in treated and untreated patients, while the effect of A3A expression was much stronger in treated compared to untreated patients suggesting that mutagenesis caused by A3B may represent a genetically regulated mechanism contributing to cancer initiation, while mutagenesis caused by A3A may represent events occurring in tumours and influenced by the tumour-specific environment, including treatment [134]. Increased mutation loads, especially in DNA repair genes, were also associated with a response to neoadjuvant cisplatin-based treatment of MIBC (figure 5). This was attributed to the inability of cancer cells to recover after treatment-induced DNA damage [134].

Robertson *et al.* in 2017 [192] highlighted essential findings from the complete cohort of 412 MIBC samples characterized by multiple TCGA analytical platforms. They confirmed that MIBCs show high overall mutation rates similar to those of melanoma and NSCLC, and these high rates are principally associated with mutation signatures from APOBEC enzymes [134]. Most BLCA mutations are clonal, suggesting that APOBEC's mutagenic activity occurs early in BLCA development [134,172,192]. For instance, mutations in specific cancer genes as TP53 and ARID1A show a tendency to be clonal, but focusing on subclonal mutations in known cancer driver genes, in APOBEC-associated BLCA more than 45% of subclonal mutations in driver genes occurred in an APOBEC context [172]. Curiously, in BLCA, it appears that one of the two A3 mutation signatures defined by Alexandrov and co-workers

(signature 13) is enriched early, while the other (signature 2) becomes enriched in subclones [11,172]. A better understanding of the origin and regulation of APOBEC expression and activity in normal bladder could lead to preventive strategies that target APOBEC as a key mutagenic source in BLCA. Robertson *et al.* in 2017 [192] also corroborated a positive correlation between the improved survival of subjects with higher mutational APOBEC burden and higher neoantigen load as described by Middlebrooks *et al.* [134] and proposed that this is due to a natural host immune reaction to the high mutation burden, curbing further tumour growth and metastasis [192]. Also, chromatin modifier gene mutations are common in BLCA and open potential therapeutic opportunities through rebalancing acetylation and deacetylation, and through other chromatin modifications. Integrating RNA subtype classification, pathway information, epithelial–mesenchymal transition and low carcinoma-*in situ* signatures, and immune infiltrate analyses led to proposal of a model of mRNA-based expression subtypes that may be associated with unique responses to therapies that can be prospectively tested in clinical trials [192]. Neoadjuvant cisplatin-based chemotherapy is the current standard of care in cisplatin-eligible patients without risk stratification. However, as not all patients derive benefit from chemotherapy, subtype-specific personalized therapies could help to optimize global patient outcome, while preventing unnecessary toxicity to non-responders. Also, the results obtained by this group suggested that mRNA subtype classification may be possible with a reduced gene set, enabling validation in independent cohorts and informing clinical trial designs that test new personalized therapies [192].

Recently, several groups investigated the APOBEC mutational signature in the TCGA, Beijing Genomics Institute and Cancer Cell Line Encyclopedia BLCA datasets and its relationship with specific mutations, molecular subtype, gene expression and survival. The results obtained by Glaser *et al.* in 2018 [135] were in good agreement with the results obtained by Middlebrooks *et al.* in 2016 [134] and also with those obtained by Robertson *et al.* in 2017 [192]. They hypothesized that tumours with high levels of APOBEC-mediated mutagenesis would be enriched for mutations in DNA damage response genes and express genes related to activation of the immune system at higher levels, while tumours with low levels of APOBEC-mediated mutagenesis may have enrichments for oncogenes. They found that the expression of A3A and A3B were the only APOBEC enzymes that directly correlated with the total mutation burden in every BLCA subtype. A3A was expressed at a significantly higher level in the basal subtype than in luminal, p53-like or claudin-low subtypes, while A3B was evenly expressed across subtypes. Tumours enriched for APOBEC mutagenesis had better survival and were more likely to have mutations in both DNA damage repair and chromatin-modifying genes such as TP53, PIK3CA (primarily at E542 K and E545 K), ATR, BRCA2, MLL, MLL3 and ARID1A (figure 5). The APOBEC mutagenesis signature was associated with immune signatures and with increased expression of immune-related genes [135]. Bladder tumours not enriched for APOBEC mutagenesis were more likely to have mutations in FGFR3 and the RAS family of oncogenes (KRAS/HRAS/NRAS), which are mutually exclusive, and these patients had poor overall survival. The mutational pattern described above was also confirmed by analysis of 20 BLCA cell lines [135]. Finally, to further evaluate the

association of APOBEC mutational pattern, A3B enzyme expression and the immune environment, they analysed A3B expression in two APOBEC-low cell lines (RT4 and KU-19–19) and two APOBEC-high cell lines (HT-1376 and UM-UC-3) after exposure to IFN$\gamma$. The expression of A3B increased after exposure to IFN$\gamma$ in APOBEC-high cell lines (figure 5), but not in APOBEC-low cell lines suggesting that urothelial cancers with high APOBEC activity may have a feed-forward mechanism resulting in increased APOBEC expression upon immune activation [135].

As mentioned for lung cancer, a preferential benefit from therapeutic approaches has emerged in patients with tumours with a high prevalence of APOBEC mutagenesis. It has been suggested that hypermutation could enhance the effectiveness of immune stimulation therapy to treat cancer, by means of the generation of tumour-specific neoantigens that might trigger targeted destruction by the immune system. BLCA often has high levels of enrichment for APOBEC mutagenesis and A3A-like signatures [4]. The clinical observations in BLCA patients treated with available immune therapies [193,194] raise the intriguing possibility that hypermutation in BLCA (mainly by A3A) could contribute substantially to the success of immune therapies [4].

## 5.4. HPV-associated cancers

HPVs are small, non-enveloped double-stranded DNA viruses consisting of an 8 kb circular genome encased in a viral capsid [195]. There are more than 300 different genotypes, as well as thousands of variants, many of which may have been generated because of A3-induced mutagenesis [196]. These viruses have tropisms for the cells in the basal layer of either cutaneous or mucosal epithelia, in which the viral life cycle is tightly linked to and dependent upon keratinocyte differentiation [197]. At least 14 HPV types are carcinogenic, and these 'high-risk' (HR) types, among which HPV16 and HPV18 are the most studied, cause human cancers in the mucosal epithelia of several sites, including the cervix, vulva, vagina, penis, anus, and head and neck, especially those from the oropharynx that includes the tonsils and tongue base [198,199]. HPV16 infection is a major risk factor in cervical, head and neck squamous cell carcinoma (HNSCC) and oropharyngeal cancers. Particularly, cervical cancer has received the most attention because it ranks as the fourth leading cause of female cancer in the world with about 570 000 new cervical cancer cases diagnosed annually and 311 000 deaths [200].

A3s enzymes are all expressed, albeit at vastly different levels, in epithelial cells, which are the natural hosts of HPV infection [201]. While their cytidine deaminase activity causes C-to-T mutations during viral genome synthesis, A3s also restrict viral replication through cytidine deaminase-independent mechanisms [42]. Surprisingly, HR-HPVs have not evolved strategies to counteract restriction by A3 enzymes. The strong enrichment of the APOBEC signature in cervical cancer exomes [11–13] and the previous evidence for A3 editing of HPV genomes in plantar warts and pre-cancerous cervical lesions [202] suggest that the presence of HPV in cells might somehow induce or potentiate A3 activity, damaging the host genome and resulting in the observed enrichment of these mutational signatures in HPV-associated cancers [203].

Proteins E6 and E7 from HR-HPV types are oncogenes that are important for carcinogenesis and have some key
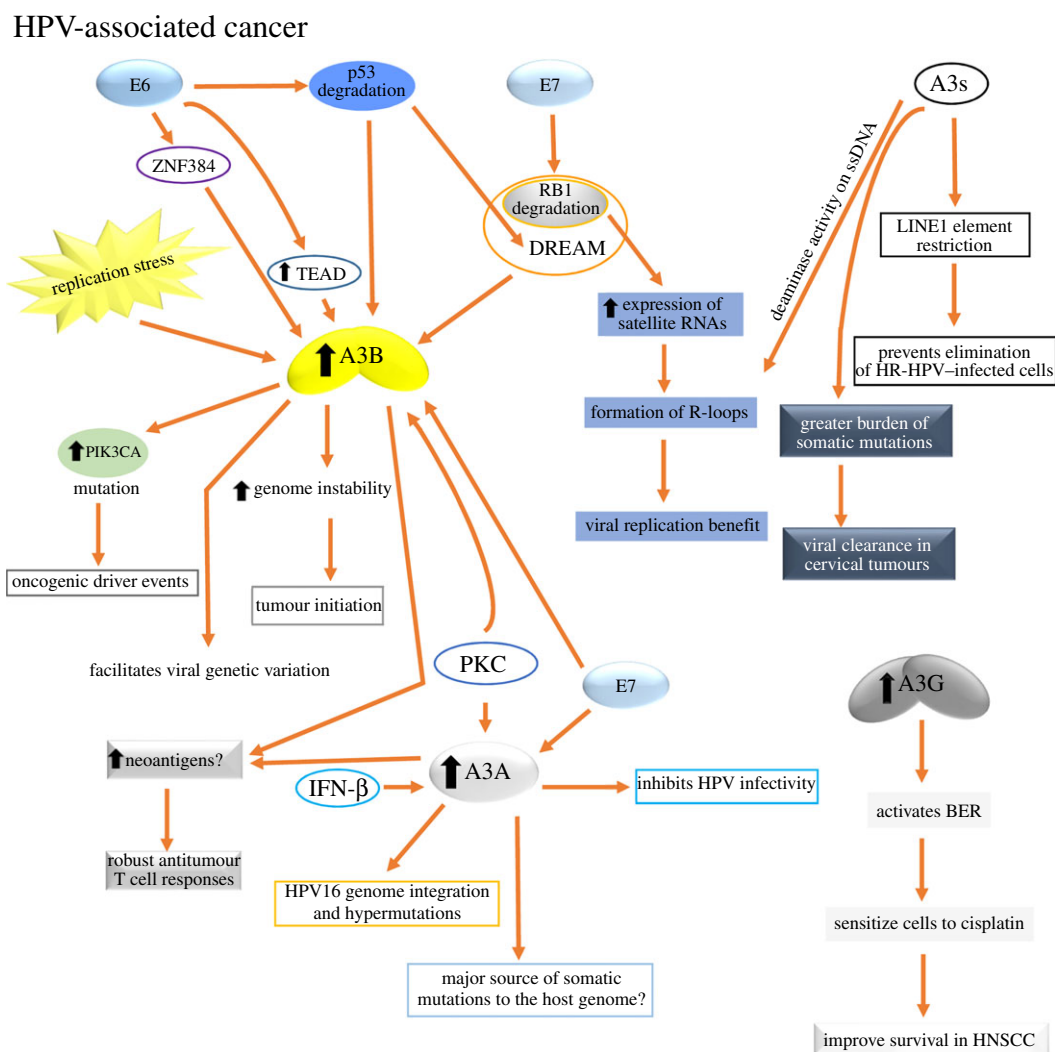
HPV-associated cancer



**Figure 6.** Causes and effects of A3s upregulation in HPV-associated cancer. In these types of cancers, the downregulation of pRB and p53 tumour suppressors exerted by E7 and E6 oncoproteins, respectively, seems to be the main mechanisms by which the mutagenic activity of A3B and A3A is triggered. Other causes for A3B upregulation are replication stress and PKC signalling activation, with the latter being also a cause for A3A upregulation. The upregulation of A3B may benefit the virus by facilitating viral genetic variation or by helping to transform the cells. It was proposed that A3A, rather than A3B, may be the major source of somatic mutations to the host genome in HPV-associated cancer. Abundant neoantigens in these types of cancers may be associated with the deaminase activity of up-regulated A3A and A3B expression that may be beneficial for tumour immunotherapy. A3A is also involved in the inhibition of the HPV infectivity, genome integration and hypermutation. The activity of A3s against retroelements could ameliorate the loss of LINE1 silencing caused by E7 inhibition of RB1. Also, HR-HPV-mediated RB1 degradation causes high-level expression of satellite RNAs that can lead to the formation of R-loops. A3s target the ssDNA in R-loops and can thereby also activate the DNA damage response, which benefits the virus replication. Cervical infections with a greater burden of somatic HPV16 A3-induced mutations are more likely to be benign or subsequently cleared. Lastly, higher A3G expression in HNSCC activates BER, sensitizing the tumour cells to cisplatin treatment and, consequently, improving the patient survival. For more details, see §5.4.

activities such as the induction of replication stress, host DNA repair responses and downregulation of the pRB and p53 tumour suppressors [204] that may serve to trigger the mutagenic activity of A3 proteins seen in HPV-associated cancers (figure 6). Several studies indicate that NF-κB pathway activation, p53 inactivation by HPV oncoprotein E6 activation, or loss-of-function mutations in the TP53 gene and replication stress activation are responsible for transcriptional activation of APOBEC, in particular, A3B (figure 6) [95,149,205,206]. Both the E6 and E7 proteins from HPV16 can act independently to increase A3B expression in immortalized keratinocytes through this pathway; E6 via p53 degradation, with E7 likely acting through its effects on the p107 and p130 pRb family pocket proteins in the DREAM (DP1, RB-like, E2F4 and MuvB) complex [150], thus also offering a mechanistic basis for the E7-mediated A3B upregulation previously described (figure 6) [207]. The loss of p53

activity through mutations (e.g. in BRCA [3,142,146]) or HPV-16 E6/E7-mediated downregulation, causes A3B upregulation. Thus, inactivation of p53 by viral protein E6 activation or loss of function of p53 mutations can activate A3B function, increase genome instability and promote tumour initiation. Removing p53 allows HPV not only to activate A3B transcription, but possibly also allows the A3B protein to accumulate to levels that would not otherwise be tolerated in normal cells.

In addition to the role of HPV E6/E7 described above in upregulating A3B, it was recently proposed that the activity of A3s against retroelements could ameliorate the loss of LINE1 silencing caused by E7 inhibition of RB1, thus providing a potential explanation for why HPV causes A3 upregulation (figure 6) [208]. A3s are well known to restrict expression of repetitive elements, including LINEs, via predominantly deaminase-independent activity [68]. RB1

protein plays a key role in the epigenetic silencing of repetitive elements that may provide an alternative explanation as to why it may be beneficial for HR-HPVs not to counteract A3 restriction (figure 6). RB1 silences repetitive elements by associating with a unique E2F1 transcription factor complex and the degradation of RB1 by HR-HPV E7 proteins is, therefore, predicted to cause transcription of repetitive elements [209]. Translation of LINE1 results in neoantigen expression, so transcription of repetitive elements would put HR-HPV-infected cells at risk of extinction through adaptive immune responses. A3 restriction of repetitive elements may protect HR-HPV-infected cells from undergoing excessive, lethal DNA damage and genomic instability and will prevent the elimination of HR-HPV-infected cells by adaptive immune responses to neoantigen expression due to expression of repetitive elements. Therefore, HR-HPV-infected cells would gain significant advantages from this A3-dependent restriction of LINE elements. Also, HR-HPV-mediated RB1 degradation causes high-level expression of satellite RNAs that can lead to the formation of R-loops, which causes replication forks to stall [209]. A3s target the ssDNA in R-loops and can thereby also activate the DNA damage response [210], which benefits virus replication (figure 6).

Other mechanisms of A3 dysregulation have been described in the context of viral carcinogenesis. HPV16 or HPV18 induces A3B expression in cultured cells of BRCA and HNSCC, and the virus-encoded protein E6 directly binds the A3B promoter and triggers transcription. Two functional regions responsive to E6 have been identified in the promoter: a distal region (from −200 to −51), required for basal promoter activity, and a proximal region (from +1 to +45), which exerts an inhibitory effect on gene expression. Through the regulatory functions of the cellular zinc finger protein ZNF384, E6 relieves this inhibition [211] and interacts with transcriptional-enhanced associate domain (TEAD) transcription factors at the distal region [212]. E6-mediated p53 degradation, therefore, not only de-represses A3B transcription via the DREAM complex, but also results in increased levels of TEAD expression, further activating the A3B promoter (figure 6). Infection with several polyomaviruses specifically upregulated A3B expression and activity and the viral T antigen was shown to be sufficient to mediate this response [43], suggesting that A3B upregulation appears to be a conserved response to small DNA tumour viruses.

Upon inspection of the genes most frequently A3-mutated in HPV-associated cancers, the PIK3CA proto-oncogene is almost exclusively mutated at two helical domain hotpots [143]. The distribution of PIK3CA-activating mutations is different in head and neck cancers, with exclusively helical domain C-to-T transitions observed in HPV-positive tumours and a combination of helical domain and kinase domain mutations in HPV-negative tumours [143]. The HPV-positive tumours have 5′-TGA-to-TAA transitions (complementary strand 5′-TCA-to-TTA) that convert both helical domain Glu542 and Glu545 to Lys, whereas HPV-negative tumours often have a 5′-CAT-to-CGT transition mutation resulting in a kinase domain His1047 to Arg substitution. Similar helical domain biases have also been reported for PIK3CA mutations in other APOBEC-signature tumour types, implying as the common denominator the APOBEC mutagenesis and not viral infection [143,172]. This mutation distribution in the PIK3CA oncogene is significant because it implies that A3B is the predominant source of the helical domain mutations

[213], although the other A3s have not been investigated. The distinctive pattern of APOBEC-signature mutations in exon 9 of the PIK3CA proto-oncogene in HPV-positive HNSCC and in other cancer types displaying the APOBEC mutational signature implicates APOBEC activity in the generation of oncogenic driver events [143,214], findings that were confirmed by analysis of TCGA HPV-positive HNSCC cohorts [215,216]. The strong link between A3B upregulation and viral infection, specifically HPV in HNSCC, cervical and perhaps some bladder tumours, suggests that other cancer types may have an unknown association with a virus or DNA-based genomic parasite. The upregulation of A3B may benefit the virus by facilitating viral genetic variation or by helping to transform the cell and thereby increase the virus's chances of spreading (figure 6). There are also likely to be non-viral mechanisms for A3B upregulation, and understanding what activates APOBECs is particularly important, as it could help in the development of methods to prevent activation. Activation of PKC signalling during differentiation of HPV-infected keratinocytes is a likely means by which at least A3A and possibly A3B could become upregulated (figure 6) during productive HPV infections, potentially triggering viral genome editing alongside viral amplification [28].

As recently described for BRCA [24], A3B mRNA levels are not correlated with A3 signature mutation burden in cancers associated with the HPV infection [143,217], but A3B expression is consistently elevated in HPV-associated cancers in comparison to both normal tissue and to HPV-independent cancers arising at equivalent anatomical sites [214]. It could be that when the mutations are occurring during development of these tumours, they are correlated with the expression of the A3 responsible, but that this relationship is lost following subsequent downregulation, possibly because of the role of A3s as transient hypermutators [129]. The preponderance of A3-induced mutations in HPV-driven cervical cancer, together with the observation that A3-induced mutations are enriched in the HPV-associated subset of HNSCC, suggest a possible off-target response to the virus [12,13,143,217]. Consistent with this, HPV16 infection upregulates by means of E7 the expression of A3A and A3B mRNA in keratinocytes, and both enzymes are upregulated in pre-invasive cervical lesions [207,213]. E7 from HR-HPV types can stabilize A3A protein by blocking its polyubiquitination by cullin-RING-based E3 ubiquitin ligase complexes; thus, HPVs appear to modulate A3 expression at multiple levels (figure 6) [218]. Also, A3A expression is associated with HPV16 genome integration and hypermutations in oropharyngeal cancers [219]. These observations and the previous studies demonstrating A3B upregulation by HR-HPV types [207,213] suggest an important role for A3B not only in HPV-associated cancer but also possibly in the viral life cycle. Periyasamy et al. [162] have shown that A3B associates with the ER in BRCA cell lines and co-activates ER target genes, involving deamination of promoter sites by A3B, leading to recruitment of DNA repair proteins and local chromatin remodelling. The cervical epithelium is an oestrogen-responsive tissue; indeed, HPV E6/E7-driven cervical cancer development in transgenic mice can be promoted by oestradiol infusion over several months [220]. Considering these factors, it is a possibility that A3B could also fuel cervical carcinogenesis via this non-mutagenic but nonetheless deaminase-dependent transcriptional activity.

The regulation of A3B by HPV has been the focus of attention; however, it is important to consider the roles that other A3 genes may play, both in the response to HPV infection and in HPV-associated cancer. Vartanian *et al.* [202] first reported the evidence for APOBEC editing of HPV in human cells and noted that HPV1a DNA co-transfected with A3A, A3C and A3H but not A3B displayed evidence of cytidine deamination, and while low-risk HPV genomes isolated from warts display evidence of A3 editing, several tested low-risk E6 variants did not upregulate A3B in cultured keratinocytes [213]. Taken together these findings [202,213] and the study from Warren *et al.* [207] suggesting that A3A but not A3B inhibits HPV infectivity, there were two hypotheses considered. The first hypothesis considered the possibility that the A3 response to HPV infection (mediated by A3A and/or A3C, A3H) is entirely separate from any role in host mutagenesis (mediated by A3B) during cancer development. The other hypothesis was that although A3B is induced by HPV, it is not responsible for the mutations seen in either viral or host genomes. Consistent with this hypothesis is the analysis of tumour exome data, in which there was much greater enrichment of the 5′YTCA (A3A) signature across multiple tumour types including cervical cancer [4]. That apparent preference of A3A is also supported by *in vitro* studies using purified enzyme [221,222]. Although further investigations will be necessary, these observations suggest A3A, rather than A3B, may be the major source of somatic mutations to the host genome in HPV-associated cancer. Also, A3A could contribute to the loss of HPV genomes in persistently infected cells, given that A3A is an IFN-inducible protein in keratinocytes [223], and that A3A can eliminate foreign DNA [52].

The treatment with IFN-β significantly restricts HPV infection in keratinocytes as well as represses HPV DNA replication in infected keratinocytes [207,224–227]. A3s are IFN-inducible proteins that target retroviruses and DNA viruses, so Wang *et al.* studied whether A3s are also involved in the IFN-β-mediated response against HPV infections. This study showed that IFN-β treatment upregulated A3A expression in cervical keratinocytes, and that knockdown of A3A expression reduced IFN-β-induced hypermutation of the viral E2 gene (figure 6) [223]. Using a high-yield HPV production system, it has been shown that virions packaged in cells overexpressing A3A are dramatically less infectious in keratinocytes [228]. By contrast, the expression of other A3s localized to the nucleus (A3B and A3C) had no effect on restricting viral infection [207]. HPV restriction by A3A is deaminase-dependent, as a catalytically inactive mutant A3A was unable to restrict HPV infection [207]. Further analysis of whole viral genome or RNA sequences are needed to identify critical A3A mutation targets that disrupt HPV infectivity.

Several clinical trials have shown that HPV-positive HNSCC patients have a better response to chemotherapy and radiation therapy than HPV-negative cases [229,230]. The reasons for this different behaviour can be found in the opposite genetic features which characterize these two types of tumours. Cisplatin is used for the HNSCC treatment to induce DNA adducts including interstrand cross-links (ICLs) and previous reports have shown that HPV-positive HNSCC patients respond better to cisplatin therapy. The loss of BER and mismatch repair (MMR) results in cisplatin resistance and UNG is required to initiate the BER response to cisplatin treatment and maintain drug sensitivity. Specific cytidine deaminases could play an important role in the cisplatin response by activating the BER pathway to mediate drug sensitivity. A recent study used TCGA HNSCC data to assess the association between the expression of the seven proteins in the A3 cytidine deaminase family, HPV-status and survival outcomes [231]. Higher A3G expression in HPV-positive tumours corresponds with better overall survival (figure 6). The results obtained suggest that A3s activate BER in HNSCC, mediate repair of cisplatin ICLs and thereby, sensitize cells to cisplatin which likely contributes to the improved patient responses observed in HPV-infected patients [231].

Contrary to the idea that A3-mediated somatic mutations may drive HPV-positive cancer progression, recent cancer immunology studies have shown that high levels of somatic mutations favour anti-tumour immune responses that also coincide with better prognosis after immunotherapies [232–234]. Tumour neoantigens are recognized as emerging targets for personalized cancer immunotherapies, implying that cancers with a high level of A3 mutation signatures may be beneficial for immunotherapies that induce robust anti-tumour T-cell responses specific to neoantigens generated by A3-mediated mutations. One study has revealed that tumour-infiltrating lymphocytes in cervical cancer are more reactive to neoantigens than to HPV viral epitopes [235]. This finding suggests that abundant neoantigens in HPV-positive cancers may be associated with the deaminase activity of upregulated A3A and A3B expression (figure 6). In this regard, A3-mediated mutations could be used beneficially to identify T-cell epitopes and treat HPV-positive cancer patients. Also, Boichard *et al.* [159] found that 'kataegis' and A3 overexpression participate in regulation of PD-L1 expression. Thus, it would be interesting to investigate if A3 mutation loads in patients correlate to better outcome following current immunotherapies targeting immune checkpoint blockades, e.g. PD-L1 that is not expressed in normal cervical tissue but is expressed in 95% of cervical intraepithelial neoplasia (grades 1 and 2) [236]. Lastly, one study analysed the viral genomes of 5328 HPV16-positive case–control samples to investigate mutational signatures and the role of human A3-induced mutations in viral clearance and cervical carcinogenesis [132]. This analysis revealed that cervical infections with a greater burden of somatic HPV16 A3-induced mutations are more likely to be benign or subsequently clear (figure 6), suggesting they may reduce persistence, and thus progression, within the host [132].

# 6. Concluding remarks

Despite the intense research completed to delineate the specific contribution of each A3 enzyme involved in cancer, this is still an area that requires further investigation. In general, the studies have been performed in yeast, cell lines and tumours using bioinformatics resources. The comparison among them has been difficult because those studies were done under different conditions or using different cell lines. In some cases, though not conclusively, we are gaining some knowledge about the relationship between A3 enzymes and cancer. However, the mechanistic studies informing us whether A3s are associated with the cell transformation or which A3s are involved in tumour progression are lacking in the field. Also, the evaluation in a complex system using animal models will be necessary because the cell culture cannot mimic the interactions that take place *in vivo*,

**Table 1.** Beneficial and detrimental outcomes of the A3s in cancer.

| beneficial outcomes | detrimental outcomes |
|---|---|
| — A3-mediated RNA editing is positively associated with elevated immune activity and improved survival in BRCA [22] | — A3B overexpression is associated with adverse outcomes in ER+ BRCA [140,141] |
| — germline A3B deletion influence APOBEC mutational signature, neoantigen loads and relative immune cell compositions in BRCA that is beneficial for cancer immunotherapy [36] | — A3B deletion increases the risk of BRCA [155,156] and increases tumour mutational burden [152] |
| — A3B and APOBEC mutational signature is specifically enriched in patients with durable clinical benefit after immunotherapy in NSCLC [133] | — APOBEC is a mutagenic source, fuelling cancer heterogeneity and subclonal diversification in lung adenocarcinoma, squamous cell carcinoma, BRCA, HNSCC and BLCA [172] |
| — increased APOBEC mutagenesis in BLCA patients have a significant positive association with neoantigen loads, the relative abundances of immune-related gene expression and the improved survival [36,134,135,192] | — A3B upregulation is associated with poor NSCLC prognosis [34] |
| — increased mutation loads, especially in DNA repair genes, were associated with responsiveness to neoadjuvant cisplatin-based treatment of MIBC [134] | — A3H Hap I may cause early clonal mutations in lung adenocarcinoma [5] |
| — hypermutation in BLCA (mainly by A3A) could contribute substantially to the success of immune therapies [4] | — the loss of A3H Hap I activity through the K121E variant may benefit the lung cancer and be detrimental to the host [174] |
| — higher A3G expression sensitizes cells to cisplatin which likely contributes to the improved patient responses observed in HNSCC [231] | — the single nucleotide polymorphism, rs1014971, was associated with BLCA risk [134] |
| — tumour-infiltrating lymphocytes in cervical cancer are more reactive to neoantigens than to HPV viral epitopes [235], and the abundance of those neoantigens could be associated with the deaminase activity of upregulated A3A and A3B expression | — A3A expression associates with HPV16 genome integration and hypermutations in oropharyngeal cancers [219] |
| — cervical infections with a greater burden of somatic HPV16 A3-induced mutations are more likely to be benign or subsequently clear [132] | |

**Table 2.** A3 candidates involved in mutagenesis, prognosis, carcinogenesis or cancer progression in BRCA, lung cancer, BLCA and HPV-associated cancer.

| | A3s candidates for | | | |
|---|---|---|---|---|
| | BRCA | lung cancer | BLCA | HPV-associated cancer |
| mutagenesis | A3A and A3B [3,4,9,24,134,139], A3H Hap I [5] | A3A and A3B [4,134,175,176], A3H Hap I [5] | A3A and A3B [4,134,135] | A3A and A3B [4,28,207,213] |
| prognosis | A3B [140,141] | A3A and A3B [34,133] | A3A [134] | A3-induced mutations [132,134] |
| carcinogenesis | A3B [3,144] | | | A3B [220] |
| cancer progression | A3B [147] | | | |

particularly when studying enhanced immune recognition due to A3-induced neoantigens.

The leading candidates for inducing mutations in cancer are A3B, A3A and A3H Hap I. Notably, according to the current knowledge, the role of each A3 needs to be evaluated in each cancer separately because it is not obvious if one finding translates to additional cancer types. A3B, for example, could provide a mutation rate that is in some cases beneficial favouring immune responses following immunotherapy with a positive outcome for the cancer patients [133], but in other cases, it is related with poor clinical outcomes [34,140,141]. Understanding the role of A3A in cancer evolution, drug resistance or response to therapies requires further investigation due to the latest research suggesting its prominent role in BRCA [24], which was previously attributed to A3B [3,11–13] and the effect of A3A expression in BLCA which was much stronger than A3B in treated versus untreated patients and was associated with improved survival [134]. The role of the cytidine deaminase A3H Hap I in cancer has been less studied, but it was associated with BRCA and lung cancer [5,173,174]. Table 1 summarizes the beneficial and detrimental outcomes of the A3s overexpression. This table and the pathways of A3 activation (figures 3–6) that have emerged from different cancers exemplify the complexity of discerning the role of each A3 in each specific cancer. Additionally, table 2 summarizes where A3s candidates have been related to mutagenesis, prognosis, carcinogenesis or cancer progression for the cancers included in this review.

Undoubtedly, knowledge of A3 mutagenesis in cancer may yield significant diagnostic and prognostic value and could open the doors towards new therapeutic opportunities. Understanding what activates APOBECs is also of critical importance, as it could allow the development of methods to prevent their activation. Given evidence for the 'just-right' threshold of genome instability in cancers, increasing genomic instability in APOBEC-high tumours may also be beneficial. We are only at the beginning of research on the effects of APOBEC-mediated mutagenesis in cancer and many questions remain to be answered. In this sense, we want to leave open some questions for further consideration: (i) is overexpression of A3 enzymes the cause for cancer initiation or is it a consequence during cancer development? (ii) How and when can we avoid the off-target A3-induced mutations that results in poor clinical outcomes, and (iii) why does A3 overexpression sometimes have a clinical benefit for cancer patients? All the knowledge about the relationship between A3 enzymes and cancer will be decisive for combined therapy in cancer and will hopefully lead to refined personalized therapy in future years.

# References

1. Hanahan D, Weinberg RA. 2011 Hallmarks of cancer: the next generation. *Cell* **144**, 646–674. (doi:10.1016/j.cell.2011.02.013)

2. Luch A. 2005 Nature and nurture—lessons from chemical carcinogenesis. *Nat. Rev. Cancer* **5**, 113–125. (doi:10.1038/nrc1546)

3. Burns MB et al. 2013 APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* **494**, 366–370. (doi:10.1038/nature11881)

4. Chan K et al. 2015 An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. *Nat. Genet.* **47**, 1067–1072. (doi:10.1038/ng.3378)

5. Starrett GJ et al. 2016 The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *Nat. Commun.* **7**, 12918. (doi:10.1038/ncomms12918)

6. Berger MF et al. 2012 Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature* **485**, 502–506. (doi:10.1038/nature11071)

7. Wei X et al. 2011 Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nat. Genet.* **43**, 442–446. (doi:10.1038/ng.810)

8. Barnes DE, Lindahl T. 2004 Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* **38**, 445–476. (doi:10.1146/annurev.genet.38.072902.092448)

9. Nik-Zainal S et al. 2012 Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993. (doi:10.1016/j.cell.2012.04.024)

10. Roberts SA et al. 2012 Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. *Mol. Cell.* **46**, 424–435. (doi:10.1016/j.molcel.2012.03.030)

11. Alexandrov LB et al. 2013 Signatures of mutational processes in human cancer. *Nature* **500**, 415–421. (doi:10.1038/nature12477)

12. Roberts SA et al. 2013 An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.* **45**, 970–976. (doi:10.1038/ng.2702)

13. Burns MB, Temiz NA, Harris RS. 2013 Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat. Genet.* **45**, 977–983. (doi:10.1038/ng.2701)

14. Albin JS, Harris RS. 2010 Interactions of host APOBEC3 restriction factors with HIV-1 *in vivo*: implications for therapeutics. *Expert Rev. Mol. Med.* **12**, e4. (doi:10.1017/S1462399409001343)

15. Conticello SG. 2008 The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* **9**, 229. (doi:10.1186/gb-2008-9-6-229)

16. Chiu YL, Greene WC. 2008 The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu. Rev. Immunol.* **26**, 317–353. (doi:10.1146/annurev.immunol.26.021607.090350)

17. de Oliveira GR, Vieira VC, Barral MF, Dowich V, Soares MA, Concalves CV, de Martinez AM. 2013 Risk factors and prevalence of HPV infection in patients from basic health units of an university hospital in Southern Brazil. *Rev. Bras. Ginecol. Obstet.* **35**, 226–232. (doi:10.1590/s0100-72032013000500007)

18. Milewska A, Kindler E, Vkovski P, Zeglen S, Ochman M, Thiel V, Rajfur Z, Pyrc K. 2018 APOBEC3-mediated restriction of RNA virus replication. *Sci. Rep.* **8**, 5960. (doi:10.1038/s41598-018-24448-2)

19. Sharma S, Patnaik SK, Taggart RT, Kannisto ED, Enriquez SM, Gollnick P, Baysal BE. 2015 APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. *Nat. Commun.* **6**, 6881. (doi:10.1038/ncomms7881)

20. Sharma S, Patnaik SK, Taggart RT, Baysal BE. 2016 The double-domain cytidine deaminase APOBEC3G is a cellular site-specific RNA editing enzyme. *Sci. Rep.* **6**, 39100. (doi:10.1038/srep39100)

21. Sharma S, Patnaik SK, Kemer Z, Baysal BE. 2017 Transient overexpression of exogenous APOBEC3A causes C-to-U RNA editing of thousands of genes. *RNA Biol.* **14**, 603–610. (doi:10.1080/15476286.2016.1184387)

22. Asaoka M, Ishikawa T, Takabe K, Patnaik SK. 2019 APOBEC3-mediated RNA editing in breast cancer is associated with heightened immune activity and improved survival. *Int. J. Mol. Sci.* **20**, 5621. (doi:10.3390/ijms22025621)

23. Alexandrov LB et al. 2020 The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101. (doi:10.1038/s41586-020-1943-3)

24. Cortez LM, Brown AL, Dennis MA, Collins CD, Brown AJ, Mitchell D, Mertz TM, Roberts SA. 2019 APOBEC3A is a prominent cytidine deaminase in breast cancer. *PLoS Genet.* **15**, e1008545. (doi:10.1371/journal.pgen.1008545)

25. Swanton C, McGranahan N, Starrett GJ, Harris RS. 2015 APOBEC enzymes: mutagenic fuel for cancer evolution and heterogeneity. *Cancer Discov.* **5**, 704–712. (doi:10.1158/2159-8290.CD-15-0344)

26. Adolph MB, Love RP, Chelico L. 2018 Biochemical basis of APOBEC3 deoxycytidine deaminase activity on diverse DNA substrates. *ACS Infect. Dis.* **4**, 224–238. (doi:10.1021/acsinfecdis.7b00221)

27. Petljak M, Maciejowski J. 2020 Molecular origins of APOBEC-associated mutations in cancer. *DNA Repair (Amst)* **94**, 102905. (doi:10.1016/j.dnarep.2020.102905)

28. Smith NJ, Fenton TR. 2019 The APOBEC3 genes and their role in cancer: insights from human papillomavirus. *J. Mol. Endocrinol.* **62**, R269–R287. (doi:10.1530/JME-19-0011)

29. Green AM, Weitzman MD. 2019 The spectrum of APOBEC3 activity: from anti-viral agents to anti-cancer opportunities. *DNA Repair (Amst)* **83**, 102700. (doi:10.1016/j.dnarep.2019.102700)

30. Silvas TV, Schiffer CA. 2019 APOBEC3s: DNA-editing human cytidine deaminases. *Protein Sci.* **28**, 1552–1566. (doi:10.1002/pro.3670)

31. Henderson S, Fenton T. 2015 APOBEC3 genes: retroviral restriction factors to cancer drivers. *Trends Mol. Med.* **21**, 274–284. (doi:10.1016/j.molmed.2015.02.007)

32. Walker BA et al. 2015 APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. *Nat. Commun.* **6**, 6997. (doi:10.1038/ncomms7997)

33. Xu L, Chang Y, An H, Zhu Y, Yang Y, Xu J. 2015 High APOBEC3B expression is a predictor of recurrence in

patients with low-risk clear cell renal cell carcinoma. *Urol. Oncol.* **33**, e341–e348. (doi:10.1016/j.urolonc.2015.05.009)

34. Yan S, He F, Gao B, Wu H, Li M, Huang L, Liang J, Wu Q, Li Y. 2016 Increased APOBEC3B predicts worse outcomes in lung cancer: a comprehensive retrospective study. *J. Cancer* **7**, 618–625. (doi:10.7150/jca.14030)

35. Zhang J, Wei W, Jin HC, Ying RC, Zhu AK, Zhang FJ. 2015 The roles of APOBEC3B in gastric cancer. *Int. J. Clin. Exp. Pathol.* **8**, 5089–5096.

36. Chen Z, Wen W, Bao J, Kuhs KL, Cai Q, Long J, Shu XO, Zheng W, Guo X. 2019 Integrative genomic analyses of APOBEC-mutational signature, expression and germline deletion of APOBEC3 genes, and immunogenicity in multiple cancer types. *BMC Med. Genomics* **12**, 131. (doi:10.1186/s12920-019-0579-3)

37. Cannataro VL et al. 2019 APOBEC-induced mutations and their cancer effect size in head and neck squamous cell carcinoma. *Oncogene* **38**, 3475–3487. (doi:10.1038/s41388-018-0657-6)

38. Mariaggi AA et al. 2018 Presence of human papillomavirus (HPV) apolipoprotein B messenger RNA editing, catalytic polypeptide-like 3 (APOBEC)-related minority variants in HPV-16 genomes from anal and cervical samples but not in HPV-52 and HPV-58. *J. Infect. Dis.* **218**, 1027–1036. (doi:10.1093/infdis/jiy287)

39. Macduff DA, Harris RS. 2006 Directed DNA deamination by AID/APOBEC3 in immunity. *Curr. Biol.* **16**, R186–R189. (doi:10.1016/j.cub.2006.02.035)

40. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. 2010 Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res.* **38**, 4274–4284. (doi:10.1093/nar/gkq174)

41. Koning FA, Newman EN, Kim EY, Kunstman KJ, Wolinsky SM, Malim MH. 2009 Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J. Virol.* **83**, 9474–9485. (doi:10.1128/JVI.01089-09)

42. Warren CJ, Westrich JA, Doorslaer KV, Pyeon D. 2017 Roles of APOBEC3A and APOBEC3B in human papillomavirus infection and disease progression. *Viruses* **9**, 233. (doi:10.3390/v9080233)

43. Verhalen B, Starrett GJ, Harris RS, Jiang M. 2016 Functional upregulation of the DNA cytosine deaminase APOBEC3B by polyomaviruses. *J. Virol.* **90**, 6379–6386. (doi:10.1128/JVI.00771-16)

44. Siriwardena SU, Chen K, Bhagwat AS. 2016 Functions and malfunctions of mammalian DNA-cytosine deaminases. *Chem. Rev.* **116**, 12 688–12 710. (doi:10.1021/acs.chemrev.6b00296)

45. Friedberg EC. 2006 *DNA repair and mutagenesis*, 2nd edn. Washington, DC: ASM Press.

46. Krokan HE, Saetrom P, Aas PA, Pettersen HS, Kavli B, Slupphaug G. 2014 Error-free versus mutagenic processing of genomic uracil—relevance to cancer. *DNA Repair (Amst)* **19**, 38–47. (doi:10.1016/j.dnarep.2014.03.028)

47. Nelson JR, Lawrence CW, Hinkle DC. 1996 Deoxycytidyl transferase activity of yeast REV1 protein. *Nature* **382**, 729–731. (doi:10.1038/382729a0)

48. Auerbach P, Bennett RA, Bailey EA, Krokan HE, Demple B. 2005 Mutagenic specificity of endogenously generated abasic sites in *Saccharomyces cerevisiae* chromosomal DNA. *Proc. Natl Acad. Sci. USA* **102**, 17 711–17 716. (doi:10.1073/pnas.0504643102)

49. Chan K, Resnick MA, Gordenin DA. 2013 The choice of nucleotide inserted opposite abasic sites formed within chromosomal DNA reveals the polymerase activities participating in translesion DNA synthesis. *DNA Repair (Amst)* **12**, 878–889. (doi:10.1016/j.dnarep.2013.07.008)

50. Vieira VC, Soares MA. 2013 The role of cytidine deaminases on innate immune responses against human viral infections. *Biomed. Res. Int.* **2013**, 683095. (doi:10.1155/2013/683095)

51. Refsland EW, Hultquist JF, Harris RS. 2012 Endogenous origins of HIV-1 G-to-A hypermutation and restriction in the nonpermissive T cell line CEM2n. *PLoS Pathog.* **8**, e1002800. (doi:10.1371/journal.ppat.1002800)

52. Stenglein MD, Burns MB, Li M, Lengyel J, Harris RS. 2010 APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nat. Struct. Mol. Biol.* **17**, 222–229. (doi:10.1038/nsmb.1744)

53. Arias JF, Koyama T, Kinomoto M, Tokunaga K. 2012 Retroelements versus APOBEC3 family members: no great escape from the magnificent seven. *Front. Microbiol.* **3**, 275. (doi:10.3389/fmicb.2012.00275)

54. Sheehy AM, Gaddis NC, Choi JD, Malim MH. 2002 Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650. (doi:10.1038/nature00939)

55. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. 2003 Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**, 99–103. (doi:10.1038/nature01709)

56. Iwatani Y et al. 2007 Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res.* **35**, 7096–7108. (doi:10.1093/nar/gkm750)

57. Berger G, Durand S, Fargier G, Nguyen XN, Cordeil S, Bouaziz S, Muriaux D, Darlix JL, Cimarelli A. 2011 APOBEC3A is a specific inhibitor of the early phases of HIV-1 infection in myeloid cells. *PLoS Pathog.* **7**, e1002221. (doi:10.1371/journal.ppat.1002221)

58. Conticello SG, Harris RS, Neuberger MS. 2003 The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr. Biol.* **13**, 2009–2013. (doi:10.1016/j.cub.2003.10.034)

59. Marin M, Golem S, Rose KM, Kozak SL, Kabat D. 2008 Human immunodeficiency virus type 1 Vif functionally interacts with diverse APOBEC3 cytidine deaminases and moves with them between cytoplasmic sites of mRNA metabolism. *J. Virol.* **82**, 987–998. (doi:10.1128/JVI.01078-07)

60. Cheng AZ et al. 2019 Epstein–Barr virus BORF2 inhibits cellular APOBEC3B to preserve viral genome integrity. *Nat. Microbiol.* **4**, 78–88. (doi:10.1038/s41564-018-0284-6)

61. Janahi EM, McGarvey MJ. 2013 The inhibition of hepatitis B virus by APOBEC cytidine deaminases. *J. Viral Hepat.* **20**, 821–828. (doi:10.1111/jvh.12192)

62. Norman JM, Mashiba M, McNamara LA, Onafuwa-Nuga A, Chiari-Fort E, Shen W, Collins KL. 2011 The antiviral factor APOBEC3G enhances the recognition of HIV-infected primary T cells by natural killer cells. *Nat. Immunol.* **12**, 975–983. (doi:10.1038/ni.2087)

63. Casartelli N, Guivel-Benhassine F, Bouziat R, Brandler S, Schwartz O, Moris A. 2010 The antiviral factor APOBEC3G improves CTL recognition of cultured HIV-infected T cells. *J. Exp. Med.* **207**, 39–49. (doi:10.1084/jem.20091933)

64. Bogerd HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS, Moran JV, Cullen BR. 2006 Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl Acad. Sci. USA* **103**, 8780–8785. (doi:10.1073/pnas.0603313103)

65. Chiu YL, Witkowska HE, Hall SC, Santiago M, Soros VB, Esnault C, Heidmann T, Greene WC. 2006 High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proc. Natl Acad. Sci. USA* **103**, 15 588–15 593. (doi:10.1073/pnas.0604524103)

66. Carmi S, Church GM, Levanon EY. 2011 Large-scale DNA editing of retrotransposons accelerates mammalian genome evolution. *Nat. Commun.* **2**, 519. (doi:10.1038/ncomms1525)

67. Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. 2014 APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. *eLife* **3**, e02008. (doi:10.7554/eLife.02008)

68. Koito A, Ikeda T. 2013 Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front. Microbiol.* **4**, 28. (doi:10.3389/fmicb.2013.00028)

69. Kazazian Jr HH. 2004 Mobile elements: drivers of genome evolution. *Science* **303**, 1626–1632. (doi:10.1126/science.1089670)

70. Kouno T et al. 2017 Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. *Nat. Commun.* **8**, 15024. (doi:10.1038/ncomms15024)

71. Yang H et al. 2020 Understanding the structural basis of HIV-1 restriction by the full length double-domain APOBEC3G. *Nat. Commun.* **11**, 632. (doi:10.1038/s41467-020-14377-y)

72. Bohn JA, Thummar K, York A, Raymond A, Brown WC, Bieniasz PD, Hatziioannou T, Smith JL. 2017 APOBEC3H structure reveals an unusual mechanism of interaction with duplex RNA. *Nat. Commun.* **8**, 1021. (doi:10.1038/s41467-017-01309-6)

73. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, Navaratnam N. 2002 An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79**, 285–296. (doi:10.1006/geno.2002.6718)

74. Harris RS, Liddament MT. 2004 Retroviral restriction by APOBEC proteins. *Nat. Rev. Immunol.* **4**, 868–877. (doi:10.1038/nri1489)

75. Navarro F, Bollman B, Chen H, Konig R, Yu Q, Chiles K, Landau NR. 2005 Complementary function of the two catalytic domains of APOBEC3G. *Virology* **333**, 374–386. (doi:10.1016/j.virol.2005.01.011)

76. Hache G, Liddament MT, Harris RS. 2005 The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J. Biol. Chem* **280**, 10 920–10 924. (doi:10.1074/jbc.M500382200)

77. Siriwardena SU, Guruge TA, Bhagwat AS. 2015 Characterization of the catalytic domain of human APOBEC3B and the critical structural role for a conserved methionine. *J. Mol. Biol.* **427**, 3042–3055. (doi:10.1016/j.jmb.2015.08.006)

78. Dang Y, Wang X, Esselman WJ, Zheng YH. 2006 Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *J. Virol.* **80**, 10 522–10 533. (doi:10.1128/JVI.01123-06)

79. Carpenter MA et al. 2012 Methylcytosine and normal cytosine deamination by the foreign DNA restriction enzyme APOBEC3A. *J. Biol. Chem.* **287**, 34 801–34 808. (doi:10.1074/jbc.M112.385161)

80. Refsland EW, Harris RS. 2013 The APOBEC3 family of retroelement restriction factors. *Curr. Top. Microbiol. Immunol.* **371**, 1–27. (doi:10.1007/978-3-642-37765-5_1)

81. Chelico L, Pham P, Calabrese P, Goodman MF. 2006 APOBEC3G DNA deaminase acts processively 3′ –> 5′ on single-stranded DNA. *Nat. Struct. Mol. Biol.* **13**, 392–399. (doi:10.1038/nsmb1086)

82. Gallois-Montbrun S, Holmes RK, Swanson CM, Fernandez-Ocana M, Byers HL, Ward MA, Malim MH. 2008 Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins. *J. Virol.* **82**, 5636–5642. (doi:10.1128/JVI.00287-08)

83. Li J, Chen Y, Li M, Carpenter MA, McDougle RM, Luengas EM, Macdonald PJ, Harris RS, Mueller JD. 2014 APOBEC3 multimerization correlates with HIV-1 packaging and restriction activity in living cells. *J. Mol. Biol.* **426**, 1296–1307. (doi:10.1016/j.jmb.2013.12.014)

84. Shaban NM et al. 2018 The antiviral and cancer genomic DNA deaminase APOBEC3H is regulated by an RNA-mediated dimerization mechanism. *Mol. Cell* **69**, 75–86. (doi:10.1016/j.molcel.2017.12.010)

85. Ito F, Yang H, Xiao X, Li SX, Wolfe A, Zirkle B, Arutiunian V, Chen XS. 2018 Understanding the structure, multimerization, subcellular localization and mC selectivity of a genomic mutator and anti-HIV factor APOBEC3H. *Sci. Rep.* **8**, 3763. (doi:10.1038/s41598-018-21955-0)

86. Feng Y, Wong L, Morse M, Rouzina I, Williams MC, Chelico L. 2018 RNA-mediated dimerization of the human deoxycytidine deaminase APOBEC3H influences enzyme activity and interaction with nucleic acids. *J. Mol. Biol.* **430**, 4891–4907. (doi:10.1016/j.jmb.2018.11.006)

87. Li MM, Emerman M. 2011 Polymorphism in human APOBEC3H affects a phenotype dominant for

88. Basu U et al. 2011 The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* **144**, 353–363. (doi:10.1016/j.cell.2011.01.001)

89. Pefanis E, Basu U. 2015 RNA exosome regulates AID DNA mutator activity in the B cell genome. *Adv. Immunol.* **127**, 257–308. (doi:10.1016/bs.ai.2015.04.002)

90. Mertz TM, Harcy V, Roberts SA. 2017 Risks at the DNA replication fork: effects upon carcinogenesis and tumor heterogeneity. *Genes (Basel)* **8**, 46. (doi:10.3390/genes8010046)

91. Lackey L, Law EK, Brown WL, Harris RS. 2013 Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell Cycle* **12**, 762–772. (doi:10.4161/cc.23713)

92. Nowarski R et al. 2012 APOBEC3G enhances lymphoma cell radioresistance by promoting cytidine deaminase-dependent DNA repair. *Blood* **120**, 366–375. (doi:10.1182/blood-2012-01-402123)

93. Seplyarskiy VB, Soldatov RA, Popadin KY, Antonarakis SE, Bazykin GA, Nikolaev SI. 2016 APOBEC-induced mutations in human cancers are strongly enriched on the lagging DNA strand during replication. *Genome Res.* **26**, 174–182. (doi:10.1101/gr.197046.115)

94. Kazanov MD, Roberts SA, Polak P, Stamatoyannopoulos J, Klimczak LJ, Gordenin DA, Sunyaev SR. 2015 APOBEC-induced cancer mutations are uniquely enriched in early-replicating, gene-dense, and active chromatin regions. *Cell Rep.* **13**, 1103–1109. (doi:10.1016/j.celrep.2015.09.077)

95. Kanu N et al. 2016 DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. *Genome Biol.* **17**, 185. (doi:10.1186/s13059-016-1042-9)

96. Hoopes JI, Cortez LM, Mertz TM, Malc EP, Mieczkowski PA, Roberts SA. 2016 APOBEC3A and APOBEC3B preferentially deaminate the lagging strand template during DNA replication. *Cell Rep.* **14**, 1273–1282. (doi:10.1016/j.celrep.2016.01.021)

97. Haradhvala NJ et al. 2016 Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell* **164**, 538–549. (doi:10.1016/j.cell.2015.12.050)

98. Saini N, Gordenin DA. 2020 Hypermutation in single-stranded DNA. *DNA Repair (Amst)* **91–92**, 102868. (doi:10.1016/j.dnarep.2020.102868)

99. Landry S, Narvaiza I, Linfesty DC, Weitzman MD. 2011 APOBEC3A can activate the DNA damage response and cause cell-cycle arrest. *EMBO Rep.* **12**, 444–450. (doi:10.1038/embor.2011.46)

100. Green AM, Landry S, Budagyan K, Avgousti DC, Shalhout S, Bhagwat AS, Weitzman MD. 2016 APOBEC3A damages the cellular genome during DNA replication. *Cell Cycle* **15**, 998–1008. (doi:10.1080/15384101.2016.1152426)

101. Nikkila J et al. 2017 Elevated APOBEC3B expression drives a kataegic-like mutation signature and replication stress-related therapeutic vulnerabilities

in p53-defective cells. *Br. J. Cancer* **117**, 113–123. (doi:10.1038/bjc.2017.133)

102. Sakofsky CJ, Roberts SA, Malc E, Mieczkowski PA, Resnick MA, Gordenin DA, Malkova A. 2014 Break-induced replication is a source of mutation clusters underlying kataegis. *Cell Rep.* **7**, 1640–1648. (doi:10.1016/j.celrep.2014.04.053)

103. Sakofsky CJ, Saini N, Klimczak LJ, Chan K, Malc EP, Mieczkowski PA, Burkholder AB, Fargo D, Gordenin DA. 2019 Repair of multiple simultaneous double-strand breaks causes bursts of genome-wide clustered hypermutation. *PLoS Biol.* **17**, e3000464. (doi:10.1371/journal.pbio.3000464)

104. Kohli RM, Abrams SR, Gajula KS, Maul RW, Gearhart PJ, Stivers JT. 2009 A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J. Biol. Chem.* **284**, 22 898–22 904. (doi:10.1074/jbc.M109.025536)

105. Logue EC, Bloch N, Dhuey E, Zhang R, Cao P, Herate C, Chauveau L, Hubbard SR, Landau NR. 2014 A DNA sequence recognition loop on APOBEC3A controls substrate specificity. *PLoS ONE* **9**, e97062. (doi:10.1371/journal.pone.0097062)

106. Feng Y, Baig TT, Love RP, Chelico L. 2014 Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. *Front. Microbiol.* **5**, 450. (doi:10.3389/fmicb.2014.00450)

107. Feng Y, Love RP, Ara A, Baig TT, Adolph MB, Chelico L. 2015 Natural polymorphisms and oligomerization of human APOBEC3H contribute to single-stranded DNA scanning ability. *J. Biol. Chem.* **290**, 27 188–27 203. (doi:10.1074/jbc.M115.666065)

108. Mitra M et al. 2015 Sequence and structural determinants of human APOBEC3H deaminase and anti-HIV-1 activities. *Retrovirology* **12**, 3. (doi:10.1186/s12977-014-0130-8)

109. Holden LG, Prochnow C, Chang YP, Bransteitter R, Chelico L, Sen U, Stevens RC, Goodman MF, Chen XS. 2008 Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121–124. (doi:10.1038/nature07357)

110. Maiti A, Myint W, Kanai T, Delviks-Frankenberry K, Sierra Rodriguez C, Pathak VK, Schiffer CA, Matsuo H. 2018 Crystal structure of the catalytic domain of HIV-1 restriction factor APOBEC3G in complex with ssDNA. *Nat. Commun.* **9**, 2460. (doi:10.1038/s41467-018-04872-8)

111. Carpenter MA, Rajagurubandara E, Wijesinghe P, Bhagwat AS. 2010 Determinants of sequence-specificity within human AID and APOBEC3G. *DNA Repair (Amst)* **9**, 579–587. (doi:10.1016/j.dnarep.2010.02.010)

112. Bransteitter R, Pham P, Calabrese P, Goodman MF. 2004 Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase. *J. Biol. Chem.* **279**, 51 612–51 621. (doi:10.1074/jbc.M408135200)

113. Berg OG, Winter RB, von Hippel PH. 1981 Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929–6948. (doi:10.1021/bi00527a028)

114. von Hippel PH, Berg OG. 1989 Facilitated target location in biological systems. *J. Biol. Chem.* **264**, 675–678.

115. Redding S, Greene EC. 2013 How do proteins locate specific targets in DNA? *Chem. Phys. Lett.* **570**, 35. (doi:10.1016/j.cplett.2013.03.035)

116. Halford SE, Marko JF. 2004 How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* **32**, 3040–3052. (doi:10.1093/nar/gkh624)

117. Casellas R, Basu U, Yewdell WT, Chaudhuri J, Robbiani DF, Di Noia JM. 2016 Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. *Nat. Rev. Immunol.* **16**, 164–176. (doi:10.1038/nri.2016.2)

118. Taylor BJ, Nik-Zainal S, Wu YL, Stebbings LA, Raine K, Campbell PJ, Rada C, Stratton MR, Neuberger MS. 2013 DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *eLife* **2**, e00534. (doi:10.7554/eLife.00534)

119. Adolph MB, Love RP, Feng Y, Chelico L. 2017 Enzyme cycling contributes to efficient induction of genome mutagenesis by the cytidine deaminase APOBEC3B. *Nucleic Acids Res.* **45**, 11 925–11 940. (doi:10.1093/nar/gkx832)

120. Gibb B, Ye LF, Gergoudis SC, Kwon Y, Niu H, Sung P, Greene EC. 2014 Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. *PLoS ONE* **9**, e87922. (doi:10.1371/journal.pone.0087922)

121. Ma CJ, Gibb B, Kwon Y, Sung P, Greene EC. 2017 Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. *Nucleic Acids Res.* **45**, 749–761. (doi:10.1093/nar/gkw1125)

122. Saini N, Roberts SA, Sterling JF, Malc EP, Mieczkowski PA, Gordenin DA. 2017 APOBEC3B cytidine deaminase targets the non-transcribed strand of tRNA genes in yeast. *DNA Repair (Amst)* **53**, 4–14. (doi:10.1016/j.dnarep.2017.03.003)

123. Loeb LA. 2011 Human cancers express mutator phenotypes: origin, consequences and targeting. *Nat. Rev. Cancer* **11**, 450–457. (doi:10.1038/nrc3063)

124. Nik-Zainal S *et al.* 2012 The life history of 21 breast cancers. *Cell* **149**, 994–1007. (doi:10.1016/j.cell.2012.04.023)

125. Nowarski R, Kotler M. 2013 APOBEC3 cytidine deaminases in double-strand DNA break repair and cancer promotion. *Cancer Res.* **73**, 3494–3498. (doi:10.1158/0008-5472.CAN-13-0728)

126. Chan K, Gordenin DA. 2015 Clusters of multiple mutations: incidence and molecular mechanisms. *Annu. Rev. Genet.* **49**, 243–267. (doi:10.1146/annurev-genet-112414-054714)

127. Law EK *et al.* 2020 APOBEC3A catalyzes mutation and drives carcinogenesis *in vivo*. *J. Exp. Med.* **217**, e20200261. (doi:10.1084/jem.20200261)

128. Stephens PJ *et al.* 2012 The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**, 400–404. (doi:10.1038/nature11017)

129. Roberts SA, Gordenin DA. 2014 Clustered and genome-wide transient mutagenesis in human cancers: hypermutation without permanent mutators or loss of fitness. *Bioessays* **36**, 382–393. (doi:10.1002/bies.201300140)

130. Petljak M *et al.* 2019 Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. *Cell* **176**, 1282–1294. (doi:10.1016/j.cell.2019.02.012)

131. Ng JC. F., Quist J, Grigoriadis A, Malim MH, Fraternali F. 2019 Pan-cancer transcriptomic analysis dissects immune and proliferative functions of APOBEC3 cytidine deaminases. *Nucleic Acids Res.* **47**, 1178–1194. (doi:10.1093/nar/gky1316)

132. Zhu B *et al.* 2020 Mutations in the HPV16 genome induced by APOBEC3 are associated with viral clearance. *Nat. Commun.* **11**, 886. (doi:10.1038/s41467-020-14730-1)

133. Wang S, Jia M, He Z, Liu XS. 2018 APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. *Oncogene* **37**, 3924–3936. (doi:10.1038/s41388-018-0245-9)

134. Middlebrooks CD *et al.* 2016 Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. *Nat. Genet.* **48**, 1330–1338. (doi:10.1038/ng.3670)

135. Glaser AP, Fantini D, Wang Y, Yu Y, Rimar KJ, Podojil JR, Miller SD, Meeks JJ. 2018 APOBEC-mediated mutagenesis in urothelial carcinoma is associated with improved survival, mutations in DNA damage response genes, and immune response. *Oncotarget* **9**, 4537–4548. (doi:10.18632/oncotarget.23344)

136. Global Burden of Disease Cancer C, Fitzmaurice C *et al.* 2018 Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: a systematic analysis for the global burden of disease study. *JAMA Oncol.* **4**, 1553–1568. (doi:10.1001/jamaoncol.2018.2706)

137. Cancer Genome Atlas N. 2012 Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70. (doi:10.1038/nature11412)

138. Nik-Zainal S, Morganella S. 2017 Mutational signatures in breast cancer: the problem at the DNA level. *Clin. Cancer Res.* **23**, 2617–2629. (doi:10.1158/1078-0432.CCR-16-2810)

139. Harris RS. 2015 Molecular mechanism and clinical impact of APOBEC3B-catalyzed mutagenesis in breast cancer. *Breast Cancer Res.* **17**, 8. (doi:10.1186/s13058-014-0498-3)

140. Law EK *et al.* 2016 The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. *Sci. Adv.* **2**, e1601737. (doi:10.1126/sciadv.1601737)

141. Sieuwerts AM *et al.* 2014 Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. *Horm. Cancer* **5**, 405–413. (doi:10.1007/s12672-014-0196-8)

142. Silwal-Pandit L *et al.* 2014 TP53 mutation spectrum in breast cancer is subtype specific and has distinct prognostic relevance. *Clin. Cancer Res.* **20**, 3569–3580. (doi:10.1158/1078-0432.CCR-13-2943)

143. Henderson S, Chakravarthy A, Su X, Boshoff C, Fenton TR. 2014 APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. *Cell Rep.* **7**, 1833–1841. (doi:10.1016/j.celrep.2014.05.012)

144. Harris RS. 2013 Cancer mutation signatures, DNA damage mechanisms, and potential clinical implications. *Genome Med.* **5**, 87. (doi:10.1186/gm490)

145. Gyorffy B, Surowiak P, Kiesslich O, Denkert C, Schafer R, Dietel M, Lage H. 2006 Gene expression profiling of 30 cancer cell lines predicts resistance towards 11 anticancer drugs at clinically achieved concentrations. *Int. J. Cancer* **118**, 1699–1712. (doi:10.1002/ijc.21570)

146. Cescon DW, Haibe-Kains B, Mak TW. 2015 APOBEC3B expression in breast cancer reflects cellular proliferation, while a deletion polymorphism is associated with immune activation. *Proc. Natl Acad. Sci. USA* **112**, 2841–2846. (doi:10.1073/pnas.1424869112)

147. Udquim KI, Zettelmeyer C, Banday AR, Lin SH, Prokunina-Olsson L. 2020 APOBEC3B expression in breast cancer cell lines and tumors depends on the estrogen receptor status. *Carcinogenesis* **41**, 1030–1037. (doi:10.1093/carcin/bgaa002)

148. Yager JD, Davidson NE. 2006 Estrogen carcinogenesis in breast cancer. *N. Engl. J. Med.* **354**, 270–282. (doi:10.1056/NEJMra050776)

149. Menendez D, Nguyen TA, Snipe J, Resnick MA. 2017 The cytidine deaminase APOBEC3 family is subject to transcriptional regulation by p53. *Mol. Cancer Res.* **15**, 735–743. (doi:10.1158/1541-7786.MCR-17-0019)

150. Periyasamy M *et al.* 2017 p53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. *Nucleic Acids Res.* **45**, 11 056–11 069. (doi:10.1093/nar/gkx721)

151. Curtis C *et al.* 2012 The genomic and transcriptomic architecture of 2000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352. (doi:10.1038/nature10983)

152. Nik-Zainal S *et al.* 2014 Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat. Genet.* **46**, 487–491. (doi:10.1038/ng.2955)

153. Kidd JM, Newman TL, Tuzun E, Kaul R, Eichler EE. 2007 Population stratification of a common APOBEC gene deletion polymorphism. *PLoS Genet.* **3**, e63. (doi:10.1371/journal.pgen.0030063)

154. Wellcome Trust Case Control C, Craddock N *et al.* 2010 Genome-wide association study of CNVs in 16 000 cases of eight common diseases and 3000 shared controls. *Nature* **464**, 713–720. (doi:10.1038/nature08979)

155. Long J *et al.* 2013 A common deletion in the APOBEC3 genes and breast cancer risk. *J. Natl Cancer Inst.* **105**, 573–579. (doi:10.1093/jnci/djt018)

156. Xuan D, Li G, Cai Q, Deming-Halverson S, Shrubsole MJ, Shu XO, Kelley MC, Zheng W, Long J. 2013 APOBEC3 deletion polymorphism is associated with

breast cancer risk among women of European ancestry. *Carcinogenesis* **34**, 2240–2243. (doi:10.1093/carcin/bgt185)

157. Klonowska K *et al.* 2017 The 30 kb deletion in the APOBEC3 cluster decreases APOBEC3A and APOBEC3B expression and creates a transcriptionally active hybrid gene but does not associate with breast cancer in the European population. *Oncotarget* **8**, 76 357–76 374. (doi:10.18632/oncotarget.19400)

158. Wen WX *et al.* 2016 Germline APOBEC3B deletion is associated with breast cancer risk in an Asian multi-ethnic cohort and with immune cell presentation. *Breast Cancer Res.* **18**, 56. (doi:10.1186/s13058-016-0717-1)

159. Boichard A, Tsigelny IF, Kurzrock R. 2017 High expression of PD-1 ligands is associated with kataegis mutational signature and APOBEC3 alterations. *Oncoimmunology* **6**, e1284719. (doi:10.1080/2162402X.2017.1284719)

160. Caval V, Suspene R, Shapira M, Vartanian JP, Wain-Hobson S. 2014 A prevalent cancer susceptibility APOBEC3A hybrid allele bearing APOBEC3B 3′UTR enhances chromosomal DNA damage. *Nat. Commun.* **5**, 5129. (doi:10.1038/ncomms6129)

161. Mussil B, Suspene R, Aynaud MM, Gauvrit A, Vartanian JP, Wain-Hobson S. 2013 Human APOBEC3A isoforms translocate to the nucleus and induce DNA double strand breaks leading to cell stress and death. *PLoS ONE* **8**, e73641. (doi:10.1371/journal.pone.0073641)

162. Periyasamy M *et al.* 2015 APOBEC3B-mediated cytidine deamination is required for estrogen receptor action in breast cancer. *Cell Rep.* **13**, 108–121. (doi:10.1016/j.celrep.2015.08.066)

163. Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, Lawrence MS. 2019 Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* **364**, aaw282. (doi:10.1126/science.aaw2872)

164. Ito F, Fu Y, Kao SA, Yang H, Chen XS. 2017 Family-wide comparative analysis of cytidine and methylcytidine deamination by eleven human APOBEC proteins. *J. Mol. Biol.* **429**, 1787–1799. (doi:10.1016/j.jmb.2017.04.021)

165. Niavarani A, Currie E, Reyal Y, Anjos-Afonso F, Horswell S, Griessinger E, Luis Sardina J, Bonnet D. 2015 APOBEC3A is implicated in a novel class of G-to-A mRNA editing in WT1 transcripts. *PLoS ONE* **10**, e0120089. (doi:10.1371/journal.pone.0120089)

166. Jalili P *et al.* 2020 Quantification of ongoing APOBEC3A activity in tumor cells by monitoring RNA editing at hotspots. *Nat. Commun.* **11**, 2971. (doi:10.1038/s41467-020-16802-8)

167. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018 Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424. (doi:10.3322/caac.21492)

168. Travis WD. 2002 Pathology of lung cancer. *Clin. Chest Med.* **23**, 65–81. (doi:10.1016/s0272-5231(03)00061-3)

169. Govindan R *et al.* 2012 Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**, 1121–1134. (doi:10.1016/j.cell.2012.08.024)

170. de Bruin EC *et al.* 2014 Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* **346**, 251–256. (doi:10.1126/science.1253462)

171. Zhang J *et al.* 2014 Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* **346**, 256–259. (doi:10.1126/science.1256930)

172. McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C. 2015 Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci. Transl. Med.* **7**, 283ra254. (doi:10.1126/scitranslmed.aaa1408)

173. Silvestrov P, Maier SJ, Fang M, Cisneros GA. 2018 DNArCdb: a database of cancer biomarkers in DNA repair genes that includes variants related to multiple cancer phenotypes. *DNA Repair (Amst)* **70**, 10–17. (doi:10.1016/j.dnarep.2018.07.010)

174. Hix MA, Wong L, Flath B, Chelico L, Cisneros GA. 2020 Single-nucleotide polymorphism of the DNA cytosine deaminase APOBEC3H haplotype I leads to enzyme destabilization and correlates with lung cancer. *NAR Cancer* **2**, zcaa023. (doi:10.1093/narcan/zcaa023)

175. Rebhandl S, Huemer M, Greil R, Geisberger R. 2015 AID/APOBEC deaminases and cancer. *Oncoscience* **2**, 320–333. (doi:10.18632/oncoscience.155)

176. Waters CE, Saldivar JC, Amin ZA, Schrock MS, Huebner K. 2015 FHIT loss-induced DNA damage creates optimal APOBEC substrates: insights into APOBEC-mediated mutagenesis. *Oncotarget* **6**, 3409–3419. (doi:10.18632/oncotarget.2636)

177. Saldivar JC, Miuma S, Bene J, Hosseini SA, Shibata H, Sun J, Wheeler LJ, Mathews CK, Huebner K. 2012 Initiation of genome instability and preneoplastic processes through loss of Fhit expression. *PLoS Genet.* **8**, e1003077. (doi:10.1371/journal.pgen.1003077)

178. Roylance R *et al.* 2011 Relationship of extreme chromosomal instability with long-term survival in a retrospective analysis of primary breast cancer. *Cancer Epidemiol. Biomarkers Prev.* **20**, 2183–2194. (doi:10.1158/1055-9965.EPI-11-0343)

179. Birkbak NJ *et al.* 2011 Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res.* **71**, 3447–3452. (doi:10.1158/0008-5472.CAN-10-3667)

180. Serebrenik AA *et al.* 2020 The DNA cytosine deaminase APOBEC3B is a molecular determinant of platinum responsiveness in clear cell ovarian cancer. *Clin. Cancer Res.* **26**, 3397–3407. (doi:10.1158/1078-0432.CCR-19-2786)

181. Rizvi NA *et al.* 2015 Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128. (doi:10.1126/science.aaa1348)

182. Snyder A, Wolchok JD, Chan TA. 2015 Genetic basis for clinical response to CTLA-4 blockade. *N. Engl. J. Med.* **372**, 783. (doi:10.1056/NEJMc1415938)

183. Topalian SL *et al.* 2012 Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* **366**, 2443–2454. (doi:10.1056/NEJMoa1200690)

184. Herbst RS *et al.* 2014 Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567. (doi:10.1038/nature14011)

185. Shukuya T, Carbone DP. 2016 Predictive markers for the efficacy of anti-PD-1/PD-L1 antibodies in lung cancer. *J. Thorac. Oncol.* **11**, 976–988. (doi:10.1016/j.jtho.2016.02.015)

186. Le DT *et al.* 2015 PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520. (doi:10.1056/NEJMoa1500596)

187. Tang H, Wang Y, Chlewicki LK, Zhang Y, Guo J, Liang W, Wang J, Wang X, Fu YX. 2016 Facilitating T cell infiltration in tumor microenvironment overcomes resistance to PD-L1 blockade. *Cancer Cell* **30**, 500. (doi:10.1016/j.ccell.2016.08.011)

188. Dong ZY, Wu SP, Liao RQ, Huang SM, Wu YL. 2016 Potential biomarker for checkpoint blockade immunotherapy and treatment strategy. *Tumour Biol.* **37**, 4251–4261. (doi:10.1007/s13277-016-4812-9)

189. Burger M *et al.* 2013 Epidemiology and risk factors of urothelial bladder cancer. *Eur. Urol.* **63**, 234–241. (doi:10.1016/j.eururo.2012.07.033)

190. Kamat AM, Hahn NM, Efstathiou JA, Lerner SP, Malmstrom PU, Choi W, Guo CC, Lotan Y, Kassouf W. 2016 Bladder cancer. *Lancet* **388**, 2796–2810. (doi:10.1016/S0140-6736(16)30512-8)

191. Lawrence MS *et al.* 2013 Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218. (doi:10.1038/nature12213)

192. Robertson AG *et al.* 2017 Comprehensive molecular characterization of muscle-invasive bladder cancer. *Cell* **171**, 540–556. (doi:10.1016/j.cell.2017.09.007)

193. Casey RG, Catto JW, Cheng L, Cookson MS, Herr H, Shariat S, Witjes JA, Black PC. 2015 Diagnosis and management of urothelial carcinoma *in situ* of the lower urinary tract: a systematic review. *Eur. Urol.* **67**, 876–888. (doi:10.1016/j.eururo.2014.10.040)

194. Powles T *et al.* 2014 MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* **515**, 558–562. (doi:10.1038/nature13904)

195. Scheurer ME, Tortolero-Luna G, Adler-Storthz K. 2005 Human papillomavirus infection: biology, epidemiology, and prevention. *Int. J. Gynecol. Cancer* **15**, 727–746. (doi:10.1111/j.1525-1438.2005.00246.x)

196. Mirabello L *et al.* 2017 HPV16 E7 genetic conservation is critical to carcinogenesis. *Cell* **170**, 1164–1174. (doi:10.1016/j.cell.2017.08.001)

197. Doorbar J. 2005 The papillomavirus life cycle. *J. Clin. Virol.* **32** (Suppl. 1), S7–S15. (doi:10.1016/j.jcv.2004.12.006)

198. Arbyn M *et al.* 2012 EUROGIN 2011 roadmap on prevention and treatment of HPV-related disease. *Int. J. Cancer* **131**, 1969–1982. (doi:10.1002/ijc.27650)

199. Bussu F *et al.* 2013 HPV infection in squamous cell carcinomas arising from different mucosal sites of the head and neck region. Is p16 immunohistochemistry a reliable surrogate marker? *Br. J. Cancer* **108**, 1157–1162. (doi:10.1038/bjc. 2013.55)

200. Arbyn M, Weiderpass E, Bruni L, de Sanjose S, Saraiya M, Ferlay J, Bray F. 2020 Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob. Health* **8**, e191–e203. (doi:10.1016/S2214-109X(19)30482-6)

201. Harden ME, Prasad N, Griffiths A, Munger K. 2017 Modulation of microRNA-mRNA target pairs by human papillomavirus 16 oncoproteins. *mBio* **8**, e02170-16. (doi:10.1128/mBio.02170-16)

202. Vartanian JP, Guetard D, Henry M, Wain-Hobson S. 2008 Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. *Science* **320**, 230–233. (doi:10.1126/science. 1153201)

203. Kuong KJ, Loeb LA. 2013 APOBEC3B mutagenesis in cancer. *Nat. Genet.* **45**, 964–965. (doi:10.1038/ ng.2736)

204. Munger K, Jones DL. 2015 Human papillomavirus carcinogenesis: an identity crisis in the retinoblastoma tumor suppressor pathway. *J. Virol.* **89**, 4708–4711. (doi:10.1128/JVI.03486-14)

205. Maruyama W *et al.* 2016 Classical NF-kappaB pathway is responsible for APOBEC3B expression in cancer cells. *Biochem. Biophys. Res. Commun.* **478**, 1466–1471. (doi:10.1016/j.bbrc.2016.08.148)

206. Gao J, Choudhry H, Cao W. 2018 Apolipoprotein B mRNA editing enzyme catalytic polypeptide-like family genes activation and regulation during tumorigenesis. *Cancer Sci.* **109**, 2375–2382. (doi:10. 1111/cas.13658)

207. Warren CJ, Xu T, Guo K, Griffin LM, Westrich JA, Lee D, Lambert PF, Santiago ML, Pyeon D. 2015 APOBEC3A functions as a restriction factor of human papillomavirus. *J. Virol.* **89**, 688–702. (doi:10.1128/ JVI.02383-14)

208. Wallace NA, Munger K. 2018 The curious case of APOBEC3 activation by cancer-associated human papillomaviruses. *PLoS Pathog.* **14**, e1006717. (doi:10.1371/journal.ppat.1006717)

209. Ishak CA *et al.* 2016 An RB-EZH2 complex mediates silencing of repetitive DNA sequences. *Mol. Cell* **64**, 1074–1087. (doi:10.1016/j.molcel.2016.10.021)

210. Sollier J, Cimprich KA. 2015 Breaking bad: R-loops and genome integrity. *Trends Cell Biol.* **25**, 514–522. (doi:10.1016/j.tcb.2015.05.003)

211. Mori S, Takeuchi T, Ishii Y, Kukimoto I. 2015 Identification of APOBEC3B promoter elements responsible for activation by human papillomavirus type 16 E6. *Biochem. Biophys. Res. Commun.* **460**, 555–560. (doi:10.1016/j.bbrc.2015.03.068)

212. Mori S, Takeuchi T, Ishii Y, Yugawa T, Kiyono T, Nishina H, Kukimoto I. 2017 Human papillomavirus 16 E6 upregulates APOBEC3B via the TEAD transcription factor. *J. Virol.* **91**, e02413-16. (doi:10. 1128/JVI.02413-16)

213. Vieira VC *et al.* 2014 Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *mBio* **5**, e02234-14. (doi:10.1128/mBio.02234-14)

214. Chakravarthy A, Henderson S, Thirdborough SM, Ottensmeier CH, Su X, Lechner M, Feber A, Thomas GJ, Fenton TR. 2016 Human papillomavirus drives tumor development throughout the head and neck: improved prognosis is associated with an immune response largely restricted to the oropharynx. *J. Clin. Oncol.* **34**, 4132–4141. (doi:10.1200/JCO.2016. 68.2955)

215. Gillison ML *et al.* 2019 Human papillomavirus and the landscape of secondary genetic alterations in oral cancers. *Genome Res.* **29**, 1–17. (doi:10.1101/ gr.241141.118)

216. Qin T *et al.* 2018 Expressed HNSCC variants by HPV-status in a well-characterized Michigan cohort. *Sci. Rep.* **8**, 11458. (doi:10.1038/s41598-018-29599-w)

217. Ojesina AI *et al.* 2014 Landscape of genomic alterations in cervical carcinomas. *Nature* **506**, 371–375. (doi:10.1038/nature12881)

218. Westrich JA, Warren CJ, Klausner MJ, Guo K, Liu CW, Santiago ML, Pyeon D. 2018 Human papillomavirus 16 E7 stabilizes APOBEC3A protein by inhibiting Cullin 2-dependent protein degradation. *J. Virol.* **92**, e01318-17. (doi:10.1128/JVI.01318-17)

219. Kondo S *et al.* 2017 APOBEC3A associates with human papillomavirus genome integration in oropharyngeal cancers. *Oncogene* **36**, 1687–1697. (doi:10.1038/onc.2016.335)

220. Brake T, Lambert PF. 2005 Estrogen contributes to the onset, persistence, and malignant progression of cervical cancer in a human papillomavirus-transgenic mouse model. *Proc. Natl Acad. Sci. USA* **102**, 2490–2495. (doi:10.1073/pnas.0409883102)

221. Shi K *et al.* 2017 Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat. Struct. Mol. Biol.* **24**, 131–139. (doi:10.1038/nsmb.3344)

222. Silvas TV, Hou S, Myint W, Nalivaika E, Somasundaran M, Kelch BA, Matsuo H, Kurt Yilmaz N, Schiffer CA. 2018 Substrate sequence selectivity of APOBEC3A implicates intra-DNA interactions. *Sci. Rep.* **8**, 7511. (doi:10.1038/s41598-018-25881-z)

223. Wang Z, Wakae K, Kitamura K, Aoyama S, Liu G, Koura M, Monjurul AM, Kukimoto I, Muramatsu M. 2014 APOBEC3 deaminases induce hypermutation in human papillomavirus 16 DNA upon beta interferon stimulation. *J. Virol.* **88**, 1308–1317. (doi:10.1128/ JVI.03091-13)

224. Herdman MT, Pett MR, Roberts I, Alazawi WO, Teschendorff AE, Zhang XY, Stanley MA, Coleman N. 2006 Interferon-beta treatment of cervical keratinocytes naturally infected with human papillomavirus 16 episomes promotes rapid reduction in episome numbers and emergence of latent integrants. *Carcinogenesis* **27**, 2341–2353. (doi:10.1093/carcin/bgl172)

225. Chang YE, Pena L, Sen GC, Park JK, Laimins LA. 2002 Long-term effect of interferon on keratinocytes that maintain human papillomavirus type 31. *J. Virol.* **76**, 8864–8874. (doi:10.1128/jvi.76.17. 8864-8874.2002)

226. Terenzi F, Saikia P, Sen GC. 2008 Interferon-inducible protein, P56, inhibits HPV DNA replication by binding to the viral protein E1. *EMBO J.* **27**, 3311–3321. (doi:10.1038/emboj.2008.241)

227. Warren CJ, Griffin LM, Little AS, Huang IC, Farzan M, Pyeon D. 2014 The antiviral restriction factors IFITM1, 2 and 3 do not inhibit infection of human papillomavirus, cytomegalovirus and adenovirus. *PLoS ONE* **9**, e96579. (doi:10.1371/journal.pone.0096579)

228. Pyeon D, Lambert PF, Ahlquist P. 2005 Production of infectious human papillomavirus independently of viral replication and epithelial cell differentiation. *Proc. Natl Acad. Sci. USA* **102**, 9311–9316. (doi:10. 1073/pnas.0504020102)

229. D'Souza G, Agrawal Y, Halpern J, Bodison S, Gillison ML. 2009 Oral sexual behaviors associated with prevalent oral human papillomavirus infection. *J. Infect. Dis.* **199**, 1263–1269. (doi:10.1086/597755)

230. Kjaer SK *et al.* 2001 High-risk human papillomavirus is sexually transmitted: evidence from a follow-up study of virgins starting sexual activity (intercourse). *Cancer Epidemiol. Biomarkers Prev.* **10**, 101–106.

231. Conner KL, Shaik AN, Ekinci E, Kim S, Ruterbusch JJ, Cote ML, Patrick SM. 2020 HPV induction of APOBEC3 enzymes mediate overall survival and response to cisplatin in head and neck cancer. *DNA Repair (Amst)* **87**, 102802. (doi:10.1016/j.dnarep. 2020.102802)

232. Schumacher TN, Schreiber RD. 2015 Neoantigens in cancer immunotherapy. *Science* **348**, 69–74. (doi:10.1126/science.aaa4971)

233. McGranahan N *et al.* 2016 Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469. (doi:10.1126/science.aaf1490)

234. Anagnostou V *et al.* 2017 Evolution of neoantigen landscape during immune checkpoint blockade in non-small cell lung cancer. *Cancer Discov.* **7**, 264–276. (doi:10.1158/2159-8290.CD-16-0828)

235. Stevanovic S *et al.* 2017 Landscape of immunogenic tumor antigens in successful immunotherapy of virally induced epithelial cancer. *Science* **356**, 200–205. (doi:10.1126/science.aak9510)

236. Mezache L, Paniccia B, Nyinawabera A, Nuovo GJ. 2015 Enhanced expression of PD L1 in cervical intraepithelial neoplasia and cervical cancers. *Mod. Pathol.* **28**, 1594–1602. (doi:10.1038/modpathol. 2015.108)