The Mutagenic Impact of Environmental Exposures in Human Cells and Cancer: Imprints Through Time

Axel Rosendahl Huber[1,2], Arne Van Hoeck[2,3] and Ruben Van Boxtel[1,2]*

[1]Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands, [2]Oncode Institute, Utrecht, Netherlands, [3]Center for Molecular Medicine, University Medical Centre Utrecht, Utrecht, Netherlands

During life, the DNA of our cells is continuously exposed to external damaging processes. Despite the activity of various repair mechanisms, DNA damage eventually results in the accumulation of mutations in the genomes of our cells. Oncogenic mutations are at the root of carcinogenesis, and carcinogenic agents are often highly mutagenic. Over the past decade, whole genome sequencing data of healthy and tumor tissues have revealed how cells in our body gradually accumulate mutations because of exposure to various mutagenic processes. Dissection of mutation profiles based on the type and context specificities of the altered bases has revealed a variety of signatures that reflect past exposure to environmental mutagens, ranging from chemotherapeutic drugs to genotoxic gut bacteria. In this review, we discuss the latest knowledge on somatic mutation accumulation in human cells, and how environmental mutagenic factors further shape the mutation landscapes of tissues. In addition, not all carcinogenic agents induce mutations, which may point to alternative tumor-promoting mechanisms, such as altered clonal selection dynamics. In short, we provide an overview of how environmental factors induce mutations in the DNA of our healthy cells and how this contributes to carcinogenesis. A better understanding of how environmental mutagens shape the genomes of our cells can help to identify potential preventable causes of cancer.

Keywords: mutagens, environmental carcinogens, mutational signatures, cancer, genomics

INTRODUCTION

Somatic mutations accumulate gradually in the DNA of cells as we age (Blokzijl et al., 2016; Cagan et al., 2021). These mutations are incorporated during replication as a result of replicating damaged DNA, error-prone DNA repair or stochastic errors by DNA polymerases (Kunkel and Bebenek, 2000). DNA can be damaged through spontaneous chemical processes, such as hydrolysis causing deamination of nucleotides. In addition, cell intrinsic processes can be mutagenic, such as oxidative stress-induced DNA damage as a consequence of cellular metabolism (Lindahl and Barnes, 2000), stochastic DNA replication errors and expression of APOBEC enzymes that actively deaminate cytosine residues (Roberts et al., 2013), as reviewed in (Tubbs and Nussenzweig, 2017). Besides these endogenous mutagenic processes, exposure to environmental genotoxins can also cause mutagenic DNA damage. The mutagenic action of the different types of DNA damage can be counteracted by the activity of a wide repertoire of DNA repair pathways. Although these repair mechanisms are

highly effective, some DNA lesions escape repair, are incorrectly repaired or are fixed as DNA mutations following mispairings generated during replication, resulting in an annual accumulation of 15–40 mutations in healthy human stem cells (Blokzijl et al., 2016). Indeed, loss of DNA repair activity results in a tremendously increased rate of mutation accumulation, depending on the affected pathway and presence of DNA damage (Drost et al., 2017; Zou et al., 2018; Zou et al., 2021; Sanders et al., 2021). Therefore, the mutational landscape of a cell is shaped by a balance between DNA damage induction and the efficiency of the repair thereof (Volkova et al., 2020).

Most of the mutations that accumulate in healthy tissues during normal ageing are induced by endogenous mutagenic processes (Blokzijl et al., 2016; Franco et al., 2018; Lee-Six et al., 2018; Lodato et al., 2018; Osorio et al., 2018; Franco et al., 2019; Lee-Six et al., 2019; Machado et al., 2021). This notion potentially explains why aging is the main risk factor for developing cancer (Edwards et al., 2002; Tomasetti and Vogelstein, 2015). However, in addition to aging, epidemiological data indicates that exposure to carcinogenic environmental exposures greatly increases the risk for developing cancer. For example, 80–90% of lung cancers are thought to be induced by smoking (Peto et al., 2000) and 86% of melanoma cases by UV-radiation (Parkin et al., 2011). Cancer is a global health problem and the leading or second largest cause of premature death in 112 out of 183 countries (Sung et al., 2021). Thus, one of the most effective strategies to prevent cancer is to reduce the exposure of individuals to environmental carcinogens (Emmons and Colditz, 2017). While various carcinogenic environmental agents have been identified, for many cancers the underlying etiology remains unclear. Identification of environmental genotoxins inducing cancer may aid in the design of effective preventive measures (Green et al., 2011; Spira et al., 2017).

Large-scale DNA sequencing of cancer, normal, and cultured cells have revolutionized our understanding of the mutagenic and DNA repair processes that can shape the mutational landscapes in the genomes of human cells. In this review, we will provide an overview of how these DNA sequencing studies have contributed to our understanding on how environmental exposures induce mutations. In addition, we address how the topography of these mutations can provide mechanistic insight into the mutagenicity of environmental genotoxins. Finally, we will discuss how these environmental genotoxins could contribute to the development of cancer, which may be key in the design of strategies to prevent cancer in the future.
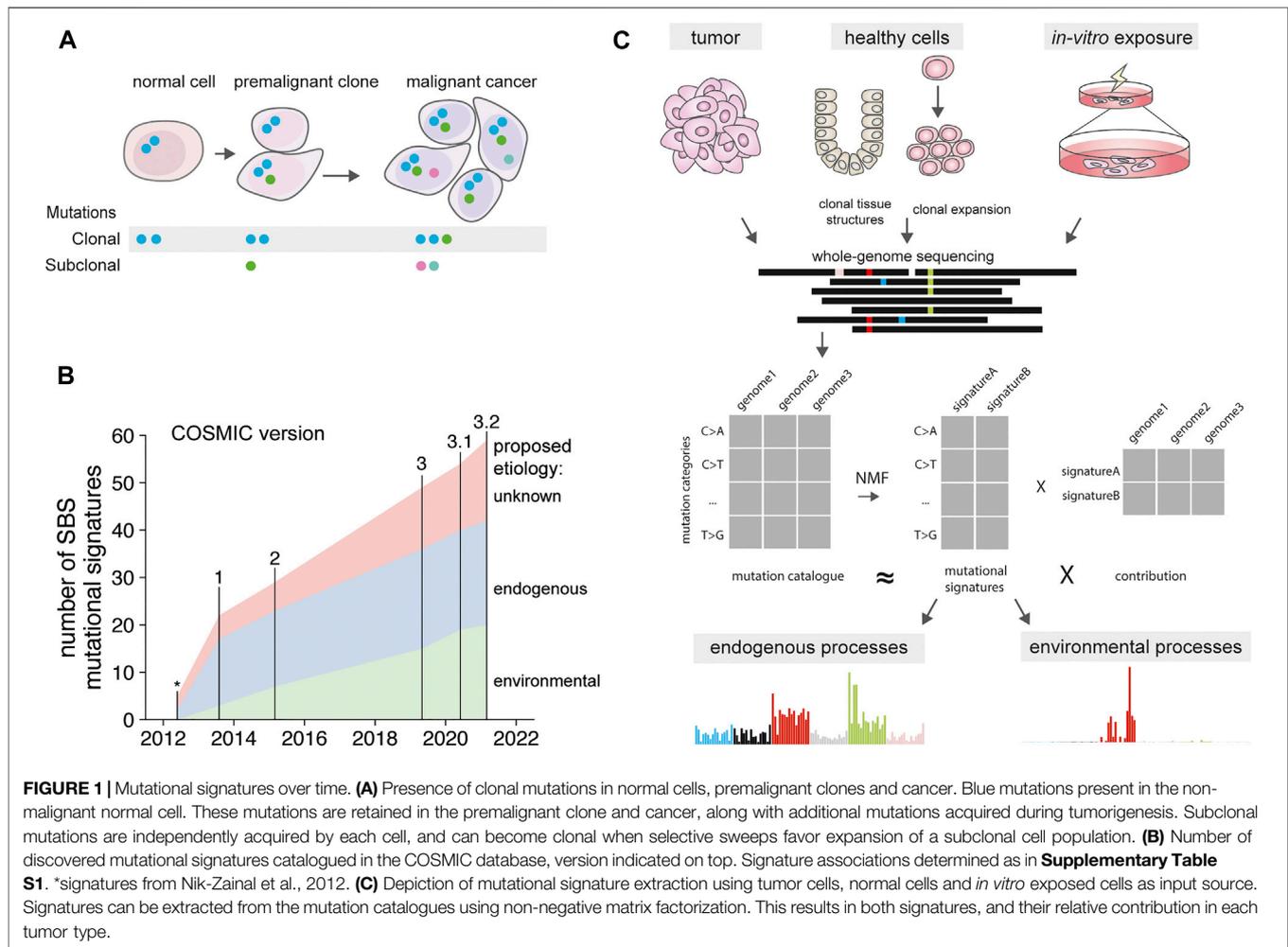
DETECTION OF MUTATIONS INDUCED BY ENVIRONMENTAL GENOTOXINS

Our initial understanding of the mutational consequences of environmental genotoxins relied on several biological assays to assess to mutagenic potential of chemical compounds. Of these, the most well-known is the Ames' test, which was developed in 1973 (Ames et al., 1975) and is still used for assessing mutagenicity of environmental and medical compounds (Zeiger, 2019). While positivity in the Ames test is a good predictor for mutagenicity as well as carcinogenic potential in rodents (Mortelmans and Zeiger, 2000), bacterial systems do not completely recapitulate the DNA structure and maintenance of mammalian cells (Johnson, 2012). To study mutagenicity in a mammalian context, the use of reporter genes, such as LacZ in mouse models has been employed (Gossen et al., 1989). In this assay, transgenic mice with chromosomally integrated LacZ reporter genes are exposed to mutagens. After the exposure, LacZ fragments are cloned into *Escherichia coli* and inactivating mutations can be selected and quantified as a measure for mutagenicity. However, such an experimental approach is impossible in humans. To overcome this, endogenous reporter genes present in the human genome, such as *HPRT* (Furth et al., 1981) or *TP53* (Pfeifer et al., 2002) have been used. In the *HPRT*-assay, inactivation of the *HPRT* gene by mutations is used as a selection marker. This gene encodes for hypoxanthine-guanine phosphoribosyltransferase (HPRT), which plays a central role in the generation of purine nucleotides through the purine salvage pathway (Sculley et al., 1992). However, cells with HPRT activity also process 6-thioguanine (6-TG) into a toxic guanine analogue, which ultimately leads to cell death (Hayes et al., 2020). The mutagenicity of a compound can be determined by counting the number of cells in an originally HPRT-proficient population that survive 6-TG selection, because they accumulated inactivating *HPRT* mutations (Furth et al., 1981). Further insight into the underlying DNA damaging processes can be obtained by analyzing the spectrum of mutations identified in the reporter genes. In the case of UV-light, induction of specific CC > TT double-base substitutions have been observed in the *HPRT* assay (Hutchinson, 1994). Indeed, sequencing cancer reporter genes, such as *TP53*, revealed the presence of CC > TT mutations in melanoma, suggesting that these mutations were induced by UV-light (Ziegler et al., 1993). In lung cancer, C > A mutations were overrepresented in *TP53* (Hollstein et al., 1991). This difference in mutational spectra between cancer types indicates that different environmental genotoxins caused distinct mutation characteristics. Despite these initial insights, sequencing of reporter genes limits mutation detection to small DNA fragments, which are biased in their sequence makeup and genomic location, making it difficult to extrapolate these findings to the entire genome.

Sequencing of Cancer Genomes

With the advent of next-generation sequencing technologies, mutation detection in whole genomes has become possible. Currently, thousands of cancer exomes and genomes have been sequenced in large consortium-based efforts (Bailey et al., 2018; Priestley et al., 2019b; Campbell et al., 2020). The data of these large-scale genome projects have been made available to the biomedical research community and efforts to increase the number of included patients are ongoing. These collections of somatic mutations are providing an unprecedented amount of information about the activity of mutagenic processes before and after carcinogenesis. As cancer is the result of a clonal expansion originating from a single founder cell, all the mutations present in that ancestral cell will be shared by all cells in the tumor

FIGURE 1 | Mutational signatures over time. **(A)** Presence of clonal mutations in normal cells, premalignant clones and cancer. Blue mutations present in the non-malignant normal cell. These mutations are retained in the premalignant clone and cancer, along with additional mutations acquired during tumorigenesis. Subclonal mutations are independently acquired by each cell, and can become clonal when selective sweeps favor expansion of a subclonal cell population. **(B)** Number of discovered mutational signatures catalogued in the COSMIC database, version indicated on top. Signature associations determined as in **Supplementary Table S1**. *signatures from Nik-Zainal et al., 2012. **(C)** Depiction of mutational signature extraction using tumor cells, normal cells and *in vitro* exposed cells as input source. Signatures can be extracted from the mutation catalogues using non-negative matrix factorization. This results in both signatures, and their relative contribution in each tumor type.

(Nik-Zainal et al., 2012b). Depending on the depth of sequencing and the clonal makeup of the tumor, subclonal mutations from the most predominant subclones can be detected as well (**Figure 1A**) (Miller et al., 2014; Roth et al., 2014). Thus, mutagenic exposure during the lifetime of the ancestral cell or during early carcinogenesis will be captured in the tumor genome (Nik-Zainal et al., 2012b). Mutagenic processes that are active later after tumor initiation and stochastically present in a single or a small subset of tumor cells are not detectable with traditional cancer genome sequencing methods. Sequencing of single cancer cells is required to accurately detect each subclonal mutation (Wang et al., 2014; Roerink et al., 2018). This approach gives a detailed insight in the mutagenic processes active during tumor progression at individual cell resolution. Clonal expansions of cells also take place *in vivo* during tumor development, such as in the case of tumor relapses, metastases or intratumoral selection pressures (e.g., treatment) favoring the outgrowth of a specific subclone. During these steps, low frequency subclonal mutations in the original tumor can become clonal and detectable by sequencing (**Figure 1A**) (Priestley et al., 2019a; Li et al., 2019). Both exome and whole genome sequencing can be employed to study cancer genomes. However, whole genome sequencing

provides a higher resolution, as it enables the detection of a large number of passenger mutations (Helleday et al., 2014). In addition, whole genome sequencing enables detailed characterization of broad (karyotype), local ploidy changes (loss-off-heterozygosity), and accurate estimates of cancer purity, which is ideally suited for studying mutational spectra and clonal compositions.

Mutational Signatures

Different mutagenic exposures during the lifetime of a cell can be disentangled by identifying recurrent mutational patterns, or "mutational signatures" across cancer genomes. These mutational signatures are defined by individual mutation classes, such as single base substitutions (SBS), double base substitutions (DBS), or short insertions and deletions (indels). For the most common mutation class, SBS, different mutagenic processes can induce specific base changes, such as UV resulting in C > T mutations and smoking in C > A mutations. These base changes are by convention always reported from the perspective of the pyrimidine base, such as C > A, or T > G, because the strand containing the mutagenic damage cannot directly be deduced. In addition to the type of substitution, the causative

processes can display a preference for the direct 5' and 3' bases flanking the mutated base. Therefore, single base substitutions are usually depicted as trinucleotide changes, in which the middle base is mutated, resulting in 96 different possibilities (Alexandrov L. B. et al., 2013). When various mutagenic processes are differentially active in a set of tumor genomes, they can be extracted from mutation data using dimension reduction algorithms (Alexandrov L. B. et al., 2013). These algorithms, such as non-negative matrix factorization, reduce the mutation spectra of a multitude of individual cancer genomes into a limited set of recurrent 96-trinucleotide signatures, and the contribution of these signatures in each of the cancer genomes (**Figure 1C**). The higher the diversity in mutation spectra in the set of assessed genomes, for example, by analyzing across various cancer types with different exposures, and the larger the cohort of samples, the more distinct signatures can be extracted. Indeed, by analyzing increasing numbers of cancer genomes, more and more signatures have been defined in the last decade (Nik-Zainal et al., 2012a; Alexandrov L. L. B. et al., 2013; Alexandrov et al., 2020) (**Figure 1B**). Each of these signatures is regarded to reflect a specific mutational process (Helleday et al., 2014) and new signatures are still being discovered (Pleguezuelos-Manzano et al., 2020; de Kanter et al., 2021). For many signatures the biological cause remains unknown, for other signatures an underlying molecular association has been proposed and for a subset the underlying causative process has been experimentally confirmed. These molecular causes range from endogenous processes that are active in all cells of the body, to exposure to specific chemotherapeutic agents (Pich et al., 2019; Alexandrov et al., 2020). While the first mutational signatures were composed of single-base substitution patterns (SBS), signatures have now also been defined for double base substitutions (DBS), and indels (ID), and are catalogued as part of the COSMIC database (Alexandrov et al., 2020). Future developments in signature extraction are focusing on the integration of additional genomic characteristics, such as specific genomic regions (Vöhringer et al., 2021), tissue-specific signatures (Degasperi et al., 2020) as well as signatures from structural variants (Li et al., 2020). While most of known signatures have been discovered in cancer genomes, the activity of the underlying mutagenic processes is often not limited to tumor cells alone and can be operative in normal cells.

Determining Mutation Accumulation in Normal Cells

As a tumor genome can serve as an historical archive, it will contain mutations that were acquired before the onset of tumorigenesis in a precancerous normal cell (Helleday et al., 2014). Most cancers, however, are characterized by a mutator phenotype (Loeb et al., 2003), which may be caused by excessive proliferation, loss of specific DNA-repair components, treatment or distorted cellular metabolism. Therefore, it is difficult to exactly determine which mutations were acquired before the malignant transformation (Stratton et al., 2009). Sequencing the genomes of normal cells can be used to identify which mutagenic processes are already active in normal non-malignant cells. However, detection of somatic

mutations in bulk normal tissue is difficult due to the stochastic nature of mutation accumulation and the polyclonal architecture of most tissues. In addition, the amount of DNA of a single cell is not sufficient for standard sequencing technologies and needs to be amplified. To achieve this, three methods have been frequently employed. First, *in vitro* expansion of single stem/progenitor cells into clonal cultures has been used to obtain sufficient DNA of the parental cell (Jager et al., 2018; Rosendahl Huber et al., 2019). This method has been used to determine the mutation accumulation in hematopoietic, intestinal, colorectal, liver, skin, muscle, kidney, and lung cells (Welch et al., 2012; Behjati et al., 2014; Blokzijl et al., 2016; Franco et al., 2018; Lee-Six et al., 2018; Osorio et al., 2018; Franco et al., 2019; Yoshida et al., 2020). While *in vitro* expansion is a highly accurate approach to amplify genomic DNA, as cells are well equipped to copy their own DNA faithfully, and circumvents the need for specialized library preparation methods, it is limited to cells with sufficient replicative potential, such as stem cells. A second approach is based on the microdissection of naturally occurring clonal structures that exist within normal tissues, such as skin, esophagus, colonic crypts, bladder, and endometrium (Martincorena et al., 2015; Martincorena et al., 2018; Brunner et al., 2019; Lee-Six et al., 2019; Yokoyama et al., 2019; Lawson et al., 2020; Moore et al., 2020; Olafsson et al., 2020; Yoshida et al., 2020; Sanders et al., 2021). The downside of this approach is that it relies on the presence of clonal structures or expansions within the tissue of choice, which are not always present, such as in the brain. In addition, only mutations that are present in all cells of the clonal population are captured, whereas more recent mutagenic events are missed. To determine the somatic mutations present in these cells, whole genome amplification (WGA) using strand displacement polymerases has been employed (Lodato et al., 2015; Gawad et al., 2016; Vijg et al., 2017; Lodato et al., 2018). However, WGA-based methods are notorious for introducing amplification biases, resulting in overrepresentation of random loci, and allelic dropouts (Hou et al., 2012). These biases make WGA-based methods less suited for assessing mutations in samples with a relative low mutation burden, such as normal cells. Nonetheless, new promising methods to directly study somatic mutations in single cells have been developed that address amplification-induced artifacts, such as primary template-directed amplification (PTA) (Gonzalez-Pena et al., 2021). PTA relies on the introduction of exonuclease-resistant amplification terminators, resulting in a more uniform amplification of the genome. In addition, duplex-sequencing, such as Nanorate sequencing (Nanoseq), enables the sequencing of two complimentary DNA strands. By comparing the sequences, errors that arise during sequencing can be corrected for, as these are not shared by the two strands. This approach has been employed to detect somatic mutations in cells with no replicative potential, such as differentiated cells (Abascal et al., 2021).

Mutagenic Processes Active in the Genomes of Normal Cells

Using the methods described above, it has been determined that the genomes of healthy cells gradually accumulate mutations during life (Blokzijl et al., 2016; Lee-Six et al., 2018; Osorio et al., 2018; Machado et al., 2021). This rate varies between ~40

mutations per year in liver, small intestine, and colon, while hematopoietic stem cells and muscle stem cells are on the lower end with about 13–15 novel mutations each year (Franco et al., 2018; Lee-Six et al., 2018; Osorio et al., 2018). A more pronounced variation has been observed for the mutation spectra and signatures across different organs, suggesting a tissue-specific activity of mutagenic processes. Of these, some processes are active in a clock-like manner in most tissues causing mutation accumulation at a rate that is constant within a tissue (Alexandrov et al., 2015). One of these processes is reflected by signature SBS1, which is induced by spontaneous deamination of methylated cytosines, present in a CpG dinucleotide context, into thymine residues. SBS1 seems more predominant in genomes of fast-dividing cells, such as colon and intestine as well as tumors originating from these tissues (Alexandrov et al., 2015; Blokzijl et al., 2016; Lee-Six et al., 2019). Another clock-like signature, SBS5, is observed in practically all tissues (Alexandrov et al., 2015). Finally, hematopoietic stem and progenitor cells display a unique aging-related signature, termed HSPC signature (Maura et al., 2019; Brandsma et al., 2021). The cause for these latter two signatures remains unknown, but their continuous action in normal cells suggests a role for endogenous mutagenic sources, which are constantly present. In addition to these clock-like processes, the mutational consequences of a variety of environmental mutagenic processes can sometimes be observed in normal cells. For example, in colonic crypts, approximately half of all studied individuals displayed a specific mutational signature, which is characterized by T > N single base substitutions in an ANN<u>T</u> context (of which the underlined base is mutated) and deletions of a single thymidine in short T-homopolymers preceded by adenine (Lee-Six et al., 2019). This signature is caused by exposure to colibactin, a mutagenic toxin produced by *pks + E. coli* (Pleguezuelos-Manzano et al., 2020). These colibactin-induced mutations are shared in phylogenetically related crypts, indicating that these mutations have been induced early during life (Lee-Six et al., 2019). The mutational consequences of *pks + E. coli* can also in rare cases be observed in the genomes of bladder, neuroendocrine and head, and neck cancers. The presence of this mutational signature in these cancers is likely indicative for a history of colibactin exposure, which potentially increased the risk for developing cancer as a direct result of enhanced mutation accumulation (Boot et al., 2020; Pleguezuelos-Manzano et al., 2020). Another example of variable damage induced by environmental exposure is in the skin, where UV-induced mutations can be observed in both melanoma and healthy melanocytes (Tang et al., 2020). In skin not exposed to sunlight, melanocytes harbor a low number of UV-induced mutations. Interestingly, compared to melanocytes present in skin continuously exposed to UV-light, melanocytes from intermittently UV-exposed skin contain a higher number of UV-induced mutations (Tang et al., 2020). This observation suggests that the number of genotoxin-induced mutations does not necessarily has to correlate with the dose of exposure. In addition to melanocytes, the mutational signatures attributed to UV-damage have also been observed in skin-residing lymphocytes (Machado et al., 2021) and sporadically in T-cell lymphoma (Jones et al., 2021), pointing to a past exposure to UV-radiation.

Exposure to genotoxins in cigarette smoke can induce mutations in a patchy manner, affecting some cells but not others, as was recently demonstrated by the sequencing of bronchial epithelial cells in smokers, non-smokers, and ex-smokers (Yoshida et al., 2020). As expected, the non-smokers accumulated aging-related mutations at a constant rate with little variance amongst each assessed cell within the same donor. In contrast, this variance was increased in smokers as well as ex-smokers. Most bronchial epithelial cells displayed a several-fold increase in mutation load, which could be attributed to the tobacco smoke-associated signature SBS4. However, some cells in the lungs of smokers and ex-smokers had no additional mutation load, implying that some bronchial epithelial cells were not or less exposed to the mutagen. Interestingly, these cells with a near-normal mutation load were present in higher numbers in ex-smokers, which may explain why after years of quitting smoking the risk of developing lung cancer is reduced (Yoshida et al., 2020).

In Vitro Assays Mutagenicity

The mutagenic properties of environmental components can also be experimentally determined using cell culture assays. In these experiments, primary cells, cell lines or organoids are exposed *in vitro* to mutagens, followed by clonal expansion and sequencing analysis (Jager et al., 2018). During culture, cells experience both background mutagenesis, which seems to be driven by oxidative stress (Kuijk et al., 2020), and mutation accumulation as a result of the genotoxic exposure. Such *in vitro* experiments can confirm associations between a mutational signature and hypothesized mutagenic exposure (Kucab et al., 2019). Recently, a landmark study has examined the mutational impact of 79 different environmental genotoxins, providing a resource to causally link mutational signatures to specific mutagenic exposures (Kucab et al., 2019). One such carcinogen is aflatoxin B1, which is produced by strains of the fungus *Aspergillus* that grows on contaminated food. This carcinogen is a known mutagen, which induces SBS24 mutations as determined by signature extraction using cancer genomes (Alexandrov L. L. B. et al., 2013) and confirmed by *in vitro* mutagenicity assays (Huang et al., 2017; Kucab et al., 2019). Another carcinogenic compound, aristocholic acid, causes signature SBS22, and has been implicated with bladder and liver cancer (Poon et al., 2015; Ng et al., 2017). Also this association between signature and environmental genotoxin has indeed been confirmed *in vitro* (Nik-Zainal et al., 2015). While aflatoxin B1 and aristocholic acid directly alkylate the DNA, several indirect mutagenic processes have been identified. External exposures can result in unsuccessful apoptosis (Ichim et al., 2015; Hawkins and Miles, 2021) or altered expression of DNA replication and repair enzymes, elevating mutational loads (Russo et al., 2019; Cipponi et al., 2020). As the mutagenicity of these processes has been determined using colony outgrowth assays or reporter gene assays, mutational signatures from these processes are lacking, and could be established in the future using WGS-based methods.

Treatment-Induced Signatures

Most chemotherapeutic drugs act by fatally damaging the DNA or blocking the replication thereof (Helleday et al., 2008). However, noncancerous cells can also be damaged by treatment (Quispe-Tintaya et al., 2018). This can result in the accumulation of DNA mutations in normal tissues with potentially adverse effects later in life, such as an increased risk for developing a secondary malignancy (Morton et al., 2019). Indeed, diverse chemotherapeutic drugs have been shown to cause specific mutational signatures. Platinum-based compounds, which cause inter- and intrastrand crosslinks between at guanine-guanine residues, induce mutational signatures SBS31, and SBS35 *in vivo* (Alexandrov et al., 2020) as well as *in vitro* (Boot et al., 2018; Kucab et al., 2019). Other chemotherapies known to induce specific mutational signatures in cancer and *in vitro* are 6-mercaptopurine (6-MP) (Li et al., 2019), which causes SBS87, and 5-fluorouracil (5-FU), which is known to be one of the underlying causes of SBS17 (Christensen et al., 2019; Pich et al., 2019). More recently, it was shown that ganciclovir, a synthetic guanosine analogue used as an antiviral drug to treat reactivation of cytomegalovirus in immunocompromised patients, can induce a highly specific C > A signature at CpA sites in hematopoietic stem and progenitor cells of patients as well as *in vitro* (de Kanter et al., 2021).

The loss of specific DNA repair activity can alter the mutational profile caused by an environmental mutagen. For example, temozolomide exposure has been associated with two different mutational signatures. Patients treated with this alkylating agent can display signature SBS11 in their tumors genomes, which mainly consists of C > T changes in an CpC or CpT context (Alexandrov L. L. B. et al., 2013). However, *in vitro* exposure of induced pluripotent stem cells (iPSCs) yielded a very different signature, which comprised of T > N changes (Kucab et al., 2019). This apparent difference in temozolomide-induced mutational signatures was explained by a defective DNA mismatch repair in the tumor cells of the assessed patients, resulting in SBS11 mutations (Touat et al., 2020).

In contrast, a very similar mutational signature can also be induced by multiple factors, suggesting that some agents induce mutations in an indirect manner. For example, as indicated above, exposure to the chemotherapeutic drug 5-FU/capecitabine can cause SBS17a/b mutations, which are characterized by T > G changes in a CpTpT context (Christensen et al., 2019; Pich et al., 2019). However, this signature is also observed in the genomes of treatment-naïve esophageal and stomach tumors, which was hypothesized to be caused by gastric acid exposure (Dulak et al., 2013; Secrier et al., 2016). In addition, organoids derived from the mouse intestine also accumulate SBS17a/b-like mutations during culturing (Behjati et al., 2014). While the exact underlying mechanism remains unknown, it has been proposed that SBS17a/b mutations might be caused by the incorporation of oxidized guanine residues opposite adenine in the DNA during replication (Tomkova et al., 2018). The similarity between the mutation profiles in (5-FU untreated) tumors exhibiting SBS17a/b mutations and 5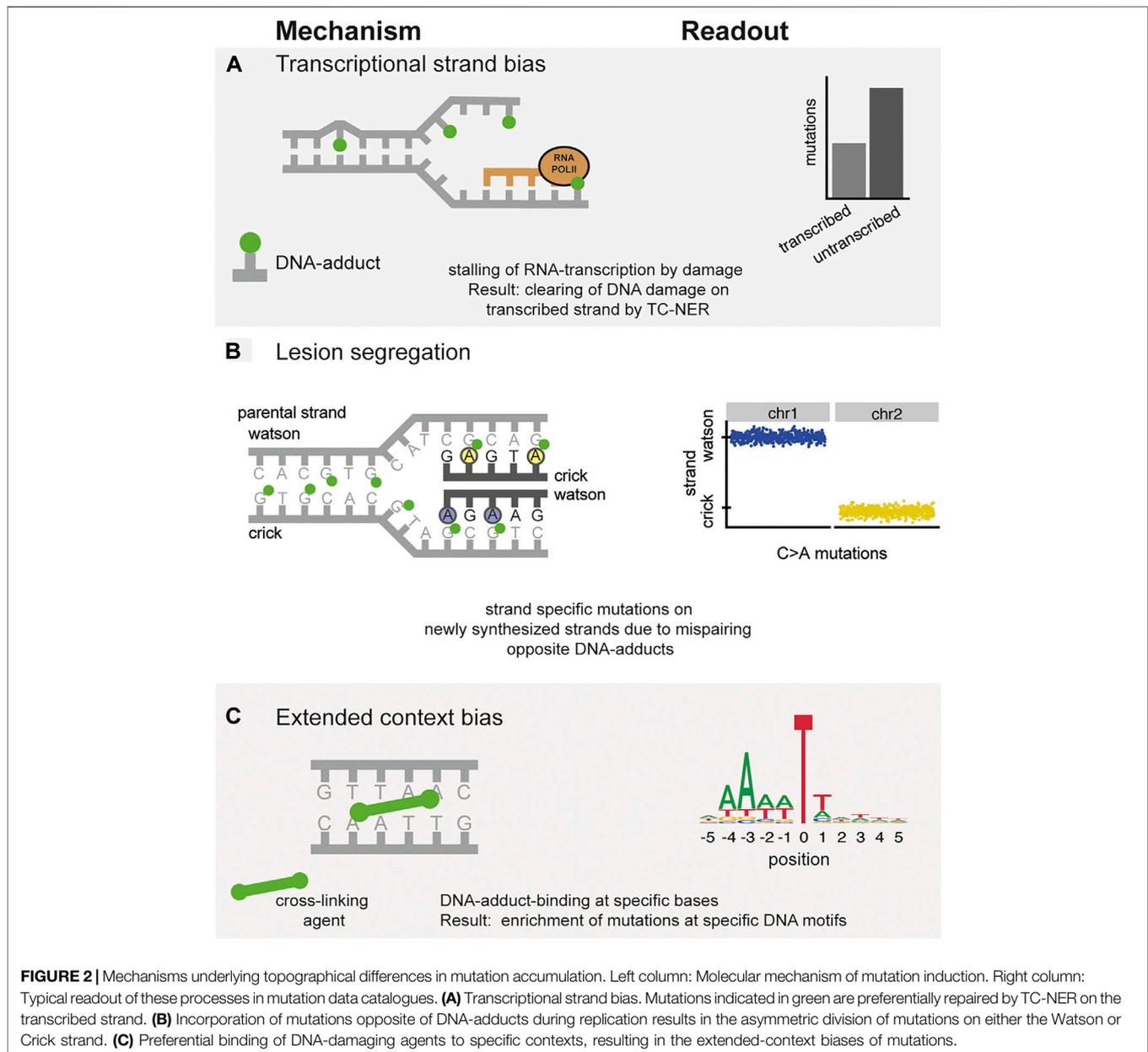-FU exposed organoids suggest that distinct mutation-inducing processes may converge in the same outcome, resulting in similar signatures.

TOPOGRAPHIES OF ENVIRONMENTALLY INDUCED MUTATIONS

In vitro experiments can prove causality between exposure to a certain environmental mutagen and a mutational signature observed in the genomes of healthy and/or tumor cells. However, the mechanism by which a mutagen induces mutations cannot always be directly inferred from the signature it causes. Nonetheless, mutations can harbor additional information, which can help in revealing the causative mechanism, such as genomic distribution and strand asymmetries (**Figure 2**) (Makova and Hardison, 2015; Haradhvala et al., 2016; Aitken et al., 2020). Together, these characteristics form a "mutational topography", which can yield further insights into the etiology of mutational signatures (Morganella et al., 2016). If the causative factor is known, the topography of a mutational profile can help to understand the molecular mechanism by which the factor induced mutations. In the following paragraphs, we will highlight some of the most informative features important for the study of environmentally induced mutations.

Transcriptional Strand Asymmetry

One of the major conformation-changing processes that occurs in DNA is its transcription into RNA. This process requires a temporary separation of the two DNA strands to synthesize a complementary RNA molecule. When RNA polymerase II (RNA Pol II) encounters a blocking DNA lesion, it cannot proceed and will stall (Vermeulen and Fousteri, 2013). To continue transcription, RNA Pol II recruits transcription-coupled nucleotide excision repair (TC-NER), which initiates the repair of the blocking lesion. As this process is dependent on transcription, TC-NER can only take place in actively transcribed regions of the genome. Importantly, only the transcribed (template) strand will repaired be by TC-NER. This preferential repair of the transcribed strand results in a characteristic mutational strand asymmetry in expressed regions (Green et al., 2003; Haradhvala et al., 2016) (**Figure 2A**). The bias of the mutational strand asymmetry (i.e., which base is on the transcribed strand and thus protected) can indicate which nucleotide of the mutated base pair was originally damaged. For instance, the mutational signature induced by smoking, SBS4, displays a depletion of C > A mutations when the guanine of that mutated base pair was on the transcribed strand. This bias indicates that damaged guanine residues underlie the signature as these are preferentially repaired on the transcribed strand. Indeed, mutagenic agents in tobacco smoke, such as benzo(a)pyrene, are known to react with guanine and thereby damaging the DNA. Thus, the presence of a transcriptional strand bias in a specific signature can give clues into the DNA damage that cause specific signatures.

FIGURE 2 | Mechanisms underlying topographical differences in mutation accumulation. Left column: Molecular mechanism of mutation induction. Right column: Typical readout of these processes in mutation data catalogues. **(A)** Transcriptional strand bias. Mutations indicated in green are preferentially repaired by TC-NER on the transcribed strand. **(B)** Incorporation of mutations opposite of DNA-adducts during replication results in the asymmetric division of mutations on either the Watson or Crick strand. **(C)** Preferential binding of DNA-damaging agents to specific contexts, resulting in the extended-context biases of mutations.

Influence of DNA-Replication State

Recent studies implicate that the replicative state of a cell is an important determinant for incorporating chemotherapy-induced DNA damage (Pich et al., 2021). According to this model, the turnover of damaged nucleotides is so quickly, that only cells in a replicative state during chemotherapy exposure will incorporate damaged nucleotides during DNA-replication. After DNA replication, damaged nucleotides are converted in mutations. This mechanism might be important in more mutagenic exposures where the damaged nucleotides are concerned, such as in thiopurine treatment, which also can induce a mutational signature (Li et al., 2019). In the future it might be feasible to take this mechanism into account during chemotherapy treatment to spare healthy replicating cells.

Strand Lesion Segregation

When DNA-lesions encounter a replication fork, these lesions can cause mispairing (Boiteux and Jinks-Robertson, 2013). The mutagenicity of such lesions can depend on the activity of different polymerases on the leading and lagging strands (Seplyarskiy et al., 2018). DNA replication can result in asymmetric distribution of mutations between Watson and Crick strands after short mutagenic exposures. During DNA replication, mispairing opposite damaged bases on the parental DNA strand can result in mutations in the newly synthesized strand of DNA in a strand-specific manner (**Figure 2B**). This strand lesion segregation was first observed in mice receiving the highly mutagenic agent diethylnitrosamine (DEN), which induces liver cancer (Aitken et al., 2020). As the exposure to

DEN was a single short burst, a clear separation of mutations in a Watson-Crick strand asymmetry could be observed in the genomes of the liver tumors. It is supposed that such a mechanism is involved during every DNA replication when damaged bases result in mispairing during DNA synthesis. However, a clear strand-specific separation is generally not visible in the genomes of tumor and normal cells, as exposure to a mutagenic source across multiple cell divisions will result in mutation accumulation on alternating Watson and Crick strands (Aitken et al., 2020). This prolonged exposure will result in DNA damage on both Watson and Crick strands, and therefore in most cancer genomes this mutational asymmetry is not observed. However, strand lesion segregation is observed in renal, liver and biliary tract-tumors containing aristocholic acid or aflatoxin exposure (Aitken et al., 2020). This segregation therefore indicates that the exposure to these environmental genotoxins was limited to a single replication round.

Double Strand Breaks Induced by Radiation

Ionizing radiation particles have the potential to induce DNA double strand breaks (DSBs) in the genomes of affected cells. These DSBs, when repaired incorrectly, can result in short insertions or deletions, and larger chromosomal aberrations. Indeed, in radiation exposed cells, mice and tumors indel mutations and larger structural variants displaying the hallmarks of erroneous DSB repair can become visible (Behjati et al., 2016; Rose Li et al., 2020; Kocakavuk et al., 2021). While there is no single hallmark of radiation-induced cancers, and mutation induction may vary on the specific type of radiation (Rose Li et al., 2020), detection of radiation-induced mutations post-treatment could be used to determine sensitivity to radiation in the future (Kocakavuk et al., 2021).

Variation in Regional Mutation Burden

Different chromatin states, gene expression, timing of DNA replication during the S-phase and differential DNA repair across the genome can result in regional differences in the density and types of mutations (Makova and Hardison, 2015; Supek and Lehner, 2015). These regional mutation densities are so specific that the distribution of mutations across cancer genomes can be used to predict the tissue type of origin of tumors (Jiao et al., 2020). In addition, the three-dimensional conformation of the genome influences regional mutation rates at boundaries between different genomic topologically associating domains (Akdemir et al., 2020). In addition, steric hindrance of DNA-binding molecules can impair the repair of DNA damage at specific DNA loci. In melanomas, it has become apparent that transcription factor-binding sites are enriched for mutations, as nucleotide excision repair is not able to access these sites resulting in mutations enriched at transcription-factor binding motifs (Perera et al., 2016; Sabarinathan et al., 2016).

Extended Mutation Context

Some mutagenic processes display a clear preference for a specific context beyond the direct 5' and 3' flanking bases. Assessing an additional base on each side of the mutation results in a 1,536 different pentanucleotide categories. While it is technically possible to delineate signatures using pentanucleotide changes, spreading mutations across a much larger number of categories results in much sparser data, complicating analyses (Alexandrov L. B. et al., 2013). Moreover, nucleotides even further away from the mutated base may be specifically enriched at mutated bases, suggesting a mechanistic cause. Such an extended context has been demonstrated for mutations induced by the cytidine deaminase enzyme APOBEC. These mutations are enriched at DNA hairpin sites, indicating that these secondary structures provide an optimal substrate for APOBEC enzymes (Buisson et al., 2019; Langenbucher et al., 2021). In addition, the mutations induced by colibactin-producing pks^+ E. coli, which are mainly T > N SBS mutations display a striking enrichment for adenines at the -3 position. As colibactin is known to preferentially bind to adenine (Wernke et al., 2020), the broader sequence context suggests that colibactin causes a cross-link between the -3 positioned adenine and the adenine opposite the mutated base (Boot et al., 2020; Pleguezuelos-Manzano et al., 2020) (**Figure 2C**).

CANCER DRIVER INDUCTION BY ENVIRONMENTAL GENOTOXINS

Induction of Driver Gene Mutations

Environmental genotoxins have the potential to induce many somatic mutations, often recognizable by the mutational signatures they leave behind in the genomes of cells (Kucab et al., 2019). However, the presence of such a mutational signature does not necessarily mean that the exposure contributed to carcinogenesis. As mutations in specific driver genes are required for carcinogenesis, evidence of driver gene induction by environmental mutagens could be used to obtain additional evidence to link environmental genotoxins to the initiation of cancer. The specific type and context characteristics of driver mutations can be used to attribute cancer driver mutations to specific mutational signatures (Poulos et al., 2018; Temko et al., 2018). In addition, the topography of mutations can provide further evidence for the involvement of specific genotoxins in causing the oncogenic mutations driving cancer. In the colorectal cancer-driving APC gene, 5.3% of the mutations display the extended motif characteristic for colibactin-induced mutations (SBS88/ID18), implying a causative role for pks^+ E. coli in the induction of colorectal cancer (Pleguezuelos-Manzano et al., 2020; Terlouw et al., 2020). However, the analysis of all APC mutations in a large colorectal cancer cohort indicates that most mutations are C > T substitutions at CpG sites, pointing to a major role for the endogenous clock-like deamination of methylated cytosines (SBS1) in the induction of cancer driving mutations (Blokzijl et al., 2016). However, even if a relatively large fraction of driver mutations is induced by endogenous processes, a tumor harbors typically between 2 and 10 cancer driver mutations (Martincorena et al., 2017). Additional driver mutations can be induced by environmental genotoxins on top of the driver mutations caused by endogenous processes, which may be sufficient to induce full malignancy. In this scenario, environmental carcinogens may be responsible for the induction of a larger fraction of cancers than suggested by the fraction of driver mutations linked to these carcinogens (Tomasetti et al., 2017; Volkova et al., 2020).

Driver Mutations in Non-malignant Tissue

Despite the well-established role of driver mutations in cancer, recent sequencing studies have shown that normal tissues can harbor many cancer driver mutations without being malignant. In 83% of skin naevi (common moles), oncogenic *BRAF* V600E mutations have been found (Pollock et al., 2003). While melanoma can originate from a naevus, most naevi never progress to melanoma. Moreover, in the skin of eyelids of elderly individuals between 55 and 73 years of age, 18–32% of small clonal populations of skin cells can contain mutations in classical cancer driver genes, such as *NOTCH1* and *TP53*. These mutations might underlie the clonal expansion of these driver-containing cells, but the tissue is phenotypically normal. Deep-sequencing of esophageal epithelium revealed a similar presence of clonal expansions, which contained cancer driver mutations in the same genes as found in normal skin (Martincorena et al., 2015; Martincorena et al., 2018; Yokoyama et al., 2019). While some of these expansions can be detected in young individuals, their frequency and size are increased in older individuals. Despite the presence of these driver mutations in up to 30% of the cells, the tissue still functions normally and appears nonmalignant. A similar clonal expansion of cells has been observed in the hematopoietic system, named clonal hematopoiesis (CH). The incidence of CH is higher in elderly individuals and was initially detected by somatic mutations in *DNMT3A*, *ASXL1,* and *TET2* (Genovese et al., 2014; Jaiswal et al., 2014). These genes are frequently mutated in leukemia and are considered to have a leukemic driving potential (Martínez-Jiménez et al., 2020). Later studies discovered CH without driver mutations is highly prevalent (Zink et al., 2017), raising the question whether CH is a pre-cancerous state. These observations raise the question whether these clonal expansions represent pre-cancer states, or if the enhanced clonality of a tissue is a characteristic of normal aging (Brash, 2015; Colom et al., 2021).
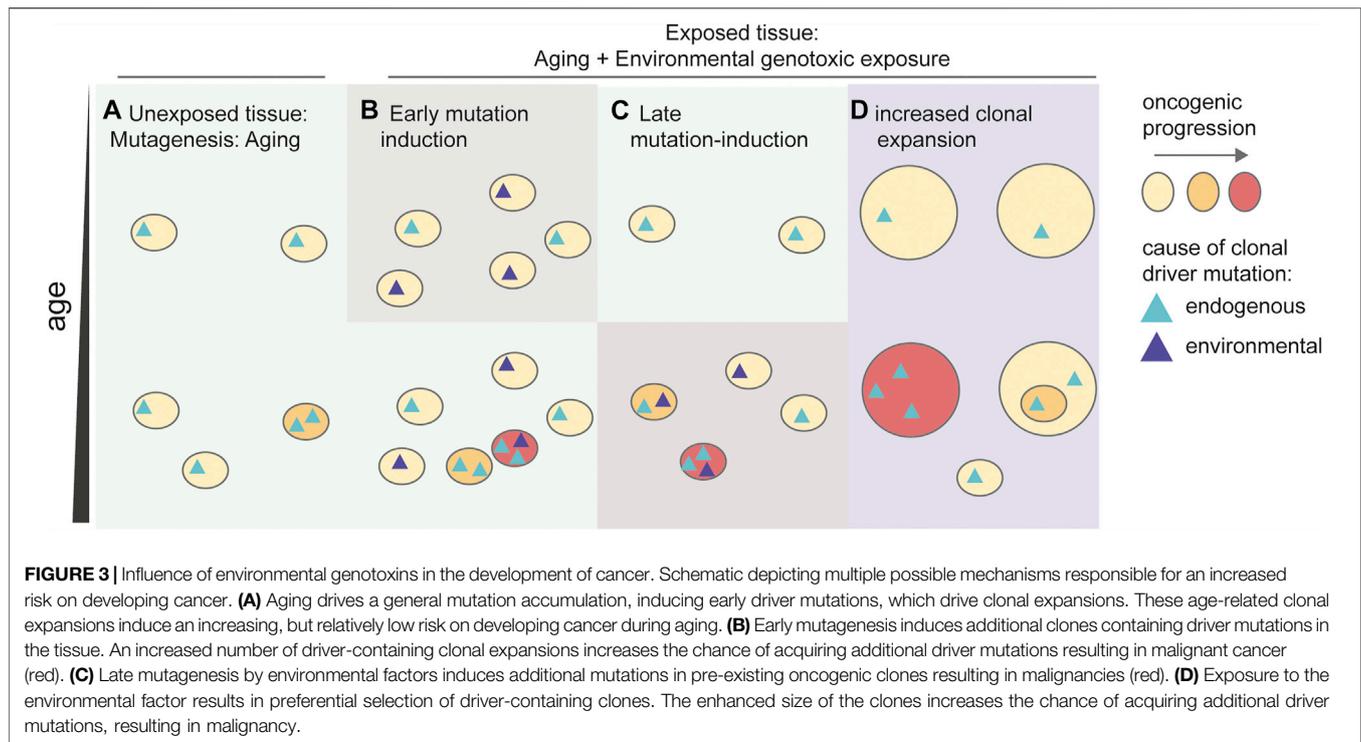
DISCUSSION

Sequencing the DNA of tumor, healthy and *in vitro* exposed tissues has provided a wealth of insights into the mutagenicity of environmental genotoxins, and the mechanisms by which they might contribute to carcinogenesis. These insights can be particularly useful to identify factors that increase the risk for developing cancer and help to design preventive measures. During the last decade, tremendous gains in knowledge have been achieved by whole genome analysis of somatic mutations. Signatures can be used to identify past mutagenic processes using somatic mutation data alone. Hypothesized causes of these signatures can be experimentally tested using experimental setups (Kucab et al., 2019). In comparison to many endogenous signatures, which can provide information on the presence of genetic predisposition (Drost et al., 2017), and acquired targetable vulnerabilities (Davies et al., 2017; Nguyen et al., 2020), environmental signatures are not (yet) used in clinical decision-making (Van Hoeck et al., 2019). However, presence of certain environmental signatures can be associated with clinical outcomes, as in the case of radiotherapy-induced deletions in metastatic cancer, which are associated with a poor prognosis (Kocakavuk et al., 2021).

As the genomics revolution has greatly increased our knowledge on mutation accumulation in our cells, it might be reasoned that now most major mutational processes are known. However, new signatures are still being discovered by analyzing genomes of healthy cells as well as cancer (Pleguezuelos-Manzano et al., 2020; de Kanter et al., 2021; Gurjao et al., 2021). An example of using somatic mutation data to assess cancer risk in exposed humans, is the recent study on Chernobyl survivors (Morton et al., 2021; Yeager et al., 2021). In these studies, the mutagenic effects of ionizing radiation could be observed in papillary thyroid carcinoma, but no transgenerational germline *de novo* mutations could be detected. Monitoring of individuals who have been exposed to high doses of environmental genotoxins, such as cancer survivors, may help to assess risk of developing cancer. In addition to targeted assessment of potential risk groups, large scale cancer genomic datasets should focus on obtaining the most diverse dataset possible, as such a dataset will capture a wider variety of cancer-causing exposures (Balmain, 2020; Ginsburg et al., 2021).

As discussed above, clonal expansions with and without driver mutations are highly prevalent in multiple tissues, and the incidence of these clonal populations increases as we age (**Figure 3A**). While some clonal expansions harbor driver mutations, these expansions are not malignant, as they lack the hallmark characteristics of cancer (Hanahan and Weinberg, 2011). Here, we propose three mechanisms how environmental genotoxins may accelerate the onset of cancer (**Figure 3**). First, exposure to environmental mutagens early in life can result in expansion of a subset of mutated cells. Later during life these cells may expand further and develop into cancer (**Figure 3B**). Enhanced mutagenesis can also play a role in later stages of cancer development, as environmental genotoxins can induce additional driver mutations in cells, or cell populations already containing a single or low number of driver mutations, resulting in cancer (**Figure 3C**). Finally, other rate-limiting steps beyond mutation induction can be involved. As stated above, clonal expansions are highly prevalent in the population, and may precede cancer onset. Thus, exposure to environmental genotoxins can alter cellular selection and promote the expansion of clones that contain specific driver mutations (**Figure 3D**). This mechanism does not require a direct induction of driver mutations by the genotoxin. Indeed, it was recently shown that a large number of carcinogens does not cause an elevated mutation load or specific mutational signatures (Riva et al., 2020), raising the question how cancer is induced, if not *via* mutation induction.

A potential mechanistic explanation for non-mutagenic cancer induction has been proposed in two studies on therapy-associated AML (t-AML), which pointed to a causal role for chemotherapy *via* selection of cells that are resistant to cytotoxic DNA damage. Two genes involved in the induction of apoptosis after DNA damage, *TP53,* and *PPM1D* were more frequently mutated in t-AML compared to *de novo*, treatment-naïve AML (Wong et al., 2015; Hsu et al., 2018). The enrichment of *TP53* and *PPM1D* mutations in t-AML can be explained by preferential selection of pre-existing HSCs, which already harbor *TP53* and *PPM1D* mutations prior to chemotherapy exposure. As both genes are involved in inducing apoptosis because of elevated DNA damage levels, cells that lack the function of these genes may have an increased chance to survive the genotoxic stress induced by cancer

FIGURE 3 | Influence of environmental genotoxins in the development of cancer. Schematic depicting multiple possible mechanisms responsible for an increased risk on developing cancer. **(A)** Aging drives a general mutation accumulation, inducing early driver mutations, which drive clonal expansions. These age-related clonal expansions induce an increasing, but relatively low risk on developing cancer during aging. **(B)** Early mutagenesis induces additional clones containing driver mutations in the tissue. An increased number of driver-containing clonal expansions increases the chance of acquiring additional driver mutations resulting in malignant cancer (red). **(C)** Late mutagenesis by environmental factors induces additional mutations in pre-existing oncogenic clones resulting in malignancies (red). **(D)** Exposure to the environmental factor results in preferential selection of driver-containing clones. The enhanced size of the clones increases the chance of acquiring additional driver mutations, resulting in malignancy.

treatment. Indeed, pre-existing clones containing *TP53* mutations can be detected in patients before treatment (Wong et al., 2015) and genotoxic exposure, such as radio- and chemotherapy, can alter selection dynamics favoring survival and expansion of these mutated cells (Bolton et al., 2020). This non-mutagenic promotion of clonal populations containing driver mutations may be an explanation why some carcinogens may induce cancer without inducing any additional mutations (Riva et al., 2020).

In humans, the incidence of cancer is low during reproductive age, presumably due to evolutionary pressure (de Magalhães, 2013). From an evolutionary perspective, however, there is no added benefit for reducing pre-malignant states cells beyond the threshold of malignant cancer induction. Therefore, there is no requirement to inhibit clonal expansions containing driver mutations if the tissue still functions normally (Martincorena and Campbell, 2015; Martincorena et al., 2018). However, aging, or exposure to environmental mutagens can both lead to elevated mutation levels and/or alter clonal dynamics. These additional processes may be enough to accumulate the additional characteristics and reach the "tipping point", where cells become malignant (Lee-Six, 2018). To reach this tipping point, a relatively modest increase in mutation load, or promotion of pre-cancerous expansions by environmental factors may already be enough to increase the incidence of cancer (Tomasetti et al., 2017; Volkova et al., 2020). Now is the time to systematically assess which environmental genotoxins can induce cancer, and in which manner. Determining how these agents induce mutational signatures or alter clonal tissue dynamics could lead to further insights, and aid in the design of future strategies in the prevention of cancer.

AUTHOR CONTRIBUTIONS

ARH and RvB wrote the manuscript with input from AvH. ARH prepared the figures. All authors approved the manuscript for publication.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.760039/full#supplementary-material

Supplementary Table S1 | Mutational signatures and their suggested etiologies over time.

REFERENCES

Abascal, F., Harvey, L. M. R., Mitchell, E., Lawson, A. R. J., Lensing, S. V., Ellis, P., et al. (2021). Somatic Mutation Landscapes at Single-Molecule Resolution. *Nature* 593, 405. doi:10.1038/s41586-021-03477-4

Aitken, S. J., Anderson, C. J., Connor, F., Pich, O., Sundaram, V., Feig, C., et al. (2020). Pervasive Lesion Segregation Shapes Cancer Genome Evolution. *Nature* 583, 265–270. doi:10.1038/s41586-020-2435-1

Akdemir, K. C., Le, V. T., Kim, J. M., Killcoyne, S., King, D. A., Lin, Y. P., et al. (2020). Somatic Mutation Distributions in Cancer Genomes Vary with Three-Dimensional Chromatin Structure. *Nat. Genet.* 52, 1178–1188. doi:10.1038/s41588-020-0708-0

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J., and Stratton, M. R. (2013a). Cell Reports Resource Deciphering Signatures of Mutational Processes Operative in Human Cancer. *Cell Rep.* 3, 246–259. doi:10.1016/j.celrep.2012.12.008

Alexandrov, L. L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R. J. R., Behjati, S., Biankin, A. V., et al. (2013b). Signatures of Mutational Processes in Human Cancer. *Nature* 500, 415–421. doi:10.1038/nature12477

Alexandrov, L. B., Jones, P. H., Wedge, D. C., Sale, J. E., Campbell, P. J., Nik-Zainal, S., et al. (2015). Clock-like Mutational Processes in Human Somatic Cells. *Nat. Genet.* 47, 1402–1407. doi:10.1038/ng.3441

Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., Wu, Y., et al. (2020). The Repertoire of Mutational Signatures in Human Cancer. *Nature* 578, 94–101. doi:10.1038/s41586-020-1943-3

Ames, B. N., McCann, J., and Yamasaki, E. (1975). Methods for Detecting Carcinogens and Mutagens with the Salmonella/mammalian-microsome Mutagenicity Test. *Mutat. Res. Environ. Mutagen. Relat. Subj.* 31, 347–363. (Netherlands). doi:10.1016/0165-1161(75)90046-1

Bailey, M. H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., et al. (2018). Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* 173, 371–385.e18. doi:10.1016/j.cell.2018.02.06010.1016/j.cell.2018.07.034

Balmain, A. (2020). The Critical Roles of Somatic Mutations and Environmental Tumor-Promoting Agents in Cancer Risk. *Nat. Genet.* 52, 1139–1143. doi:10.1038/s41588-020-00727-5

Behjati, S., Huch, M., Van Boxtel, R., Karthaus, W., Wedge, D. C., Tamuri, A. U., et al. (2014). Genome Sequencing of normal Cells Reveals Developmental Lineages and Mutational Processes. *Nature* 513, 422–425. doi:10.1038/nature13448

Behjati, S., Gundem, G., Wedge, D. C., Roberts, N. D., Tarpey, P. S., Cooke, S. L., et al. (2016). Mutational Signatures of Ionizing Radiation in Second Malignancies. *Nat. Commun.* 7, 1–8. doi:10.1038/ncomms12605

Blokzijl, F., De Ligt, J., Jager, M., Sasselli, V., Roerink, S., Sasaki, N., et al. (2016). Tissue-specific Mutation Accumulation in Human Adult Stem Cells during Life. *Nature* 538, 260–264. doi:10.1038/nature19768

Boiteux, S., and Jinks-Robertson, S. (2013). DNA Repair Mechanisms and the Bypass of DNA Damage in *Saccharomyces cerevisiae*. *Genetics* 193, 1025–1064. doi:10.1534/genetics.112.145219

Bolton, K. L., Ptashkin, R. N., Gao, T., Braunstein, L., Devlin, S. M., Kelly, D., et al. (2020). Cancer Therapy Shapes the Fitness Landscape of Clonal Hematopoiesis. *Nat. Genet.* 52, 1219–1226. doi:10.1038/s41588-020-00710-0

Boot, A., Huang, M. N., Ng, A. W. T., Ho, S. C., Lim, J. Q., Kawakami, Y., et al. (2018). In-depth Characterization of the Cisplatin Mutational Signature in Human Cell Lines and in Esophageal and Liver Tumors. *Genome Res.* 28, 654–665. doi:10.1101/gr.230219.117

Boot, A., Ng, A. W. T., Chong, F. T., Ho, S. C., Yu, W., Tan, D. S. W., et al. (2020). Characterization of Colibactin-Associated Mutational Signature in an Asian Oral Squamous Cell Carcinoma and in Other Mucosal Tumor Types. *Genome Res.* 30, 803–813. doi:10.1101/gr.255620.119

Brandsma, A. M., Bertrums, E. J. M., van Roosmalen, M. J., Hofman, D. A., Oka, R., Verheul, M., et al. (2021). Mutation Signatures of Pediatric Acute Myeloid Leukemia and normal Blood Progenitors Associated with Differential Patient Outcomes. *Blood Cancer Discov.* 2, 484–499. doi:10.1158/2643-3230.BCD-21-0010

Brash, D. E. (2015). Preprocancer. *Science* 348, 867–868. doi:10.1126/science.aac4435

Brunner, S. F., Roberts, N. D., Wylie, L. A., Moore, L., Aitken, S. J., Davies, S. E., et al. (2019). Somatic Mutations and Clonal Dynamics in Healthy and Cirrhotic Human Liver. *Nature* 574, 538–542. doi:10.1038/s41586-019-1670-9

Buisson, R., Langenbucher, A., Bowen, D., Kwan, E. E., Benes, C. H., Zou, L., et al. (2019). Passenger Hotspot Mutations in Cancer Driven by APOBEC3A and Mesoscale Genomic Features. *Science* 80, 364. doi:10.1126/science.aaw2872

Cagan, A., Baez-Ortega, A., Brzozowska, N., Abascal, F., Coorens, T. H. H., Sanders, M. A., et al. (2021). Somatic Mutation Rates Scale with Lifespan Across Mammals. *bioRxiv*. doi:10.1101/2021.08.19.456982

Campbell, P. J., Getz, G., Korbel, J. O., Stuart, J. M., Jennings, J. L., Stein, L. D., et al. (2020). Pan-cancer Analysis of Whole Genomes. *Nature* 578, 82–93. doi:10.1038/s41586-020-1969-6

Christensen, S., Van der Roest, B., Besselink, N., Janssen, R., Boymans, S., Martens, J. W. M., et al. (2019). 5-Fluorouracil Treatment Induces Characteristic T>G Mutations in Human Cancer. *Nat. Commun.* 10, 4571. doi:10.1038/s41467-019-12594-8

Cipponi, A., Goode, D. L., Bedo, J., McCabe, M. J., Pajic, M., Croucher, D. R., et al. (2020). MTOR Signaling Orchestrates Stress-Induced Mutagenesis, Facilitating Adaptive Evolution in Cancer. *Science* 368, 1127–1131. doi:10.1126/science.aau8768

Colom, B., Herms, A., Hall, M. W. J., Dentro, S. C., King, C., Sood, R. K., et al. (2021). Precancer: Mutant Clones in normal Epithelium Outcompete and Eliminate Esophageal Micro-tumors. *bioRxiv*. doi:10.1101/2021.06.25.449880

Davies, H., Glodzik, D., Morganella, S., Yates, L. R., Staaf, J., Zou, X., et al. (2017). HRDetect Is a Predictor of BRCA1 and BRCA2 Deficiency Based on Mutational Signatures. *Nat. Med.* 23, 517–525. doi:10.1038/nm.4292

de Kanter, J. K., Peci, F., Bertrums, E., Rosendahl Huber, A., van Leeuwen, A., van Roosmalen, M. J., et al. (2021). Antiviral Treatment Causes a Unique Mutational Signature in Cancers of Transplantation Recipients. *Cell Stem Cell* 28 (10), 1726–1739.e6. doi:10.1016/j.stem.2021.07.012

de Magalhães, J. P. (2013). How Ageing Processes Influence Cancer. *Nat. Rev. Cancer* 13, 357–365. doi:10.1038/nrc3497

Degasperi, A., Amarante, T. D., Czarnecki, J., Shooter, S., Zou, X., Glodzik, D., et al. (2020). A Practical Framework and Online Tool for Mutational Signature Analyses Show Intertissue Variation and Driver Dependencies. *Nat. Cancer* 1, 249–263. doi:10.1038/s43018-020-0027-5

Drost, J., Van Boxtel, R., Blokzijl, F., Mizutani, T., Sasaki, N., Sasselli, V., et al. (2017). Use of CRISPR-Modified Human Stem Cell Organoids to Study the Origin of Mutational Signatures in Cancer. *Science* 358, 234–238. doi:10.1126/science.aao3130

Dulak, A. M., Stojanov, P., Peng, S., Lawrence, M. S., Fox, C., Stewart, C., et al. (2013). Exome and Whole-Genome Sequencing of Esophageal Adenocarcinoma Identifies Recurrent Driver Events and Mutational Complexity. *Nat. Genet.* 45, 478–486. doi:10.1038/NG.2591

Edwards, B. K., Howe, H. L., Ries, L. A. G., Thun, M. J., Rosenberg, H. M., Yancik, R., et al. (2002). Annual Report to the Nation on the Status of Cancer, 1973-1999, Featuring Implications of Age and Aging on U.S. Cancer burden. *Cancer* 94, 2766–2792. doi:10.1002/cncr.10593

Emmons, K. M., and Colditz, G. A. (2017). Realizing the Potential of Cancer Prevention — the Role of Implementation Science. *N. Engl. J. Med.* 376, 986–990. doi:10.1056/NEJMsb1609101

Franco, I., Johansson, A., Olsson, K., Vrtačnik, P., Lundin, P., Helgadottir, H. T., et al. (2018). Somatic Mutagenesis in Satellite Cells Associates with Human Skeletal Muscle Aging. *Nat. Commun.* 9, 800. doi:10.1038/s41467-018-03244-6

Franco, I., Helgadottir, H. T., Moggio, A., Larsson, M., Vrtačnik, P., Johansson, A., et al. (2019). Whole Genome DNA Sequencing Provides an Atlas of Somatic Mutagenesis in Healthy Human Cells and Identifies a Tumor-Prone Cell Type. *Genome Biol.* 20, 1–22. doi:10.1186/s13059-019-1892-z

Furth, E. E., Thilly, W. G., Penman, B. W., Liber, H. L., and Rand, W. M. (1981). Quantitative Assay for Mutation in Diploid Human Lymphoblasts Using Microtiter Plates. *Anal. Biochem.* 110, 1–8. doi:10.1016/0003-2697(81)90103-2

Gawad, C., Koh, W., and Quake, S. R. (2016). Single-cell Genome Sequencing: Current State of the Science. *Nat. Rev. Genet.* 17, 175–188. doi:10.1038/nrg.2015.16

Genovese, G., Kähler, A. K., Handsaker, R. E., Lindberg, J., Rose, S. A., Bakhoum, S. F., et al. (2014). Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N. Engl. J. Med.* 371, 2477–2487. doi:10.1056/NEJMoa1409405

Ginsburg, O., Ashton-Prolla, P., Cantor, A., Mariosa, D., and Brennan, P. (2021). The Role of Genomics in Global Cancer Prevention. *Nat. Rev. Clin. Oncol.* 18, 116–128. doi:10.1038/s41571-020-0428-5

Gonzalez-Pena, V., Natarajan, S., Xia, Y., Klein, D., Carter, R., Pang, Y., et al. (2021). Accurate Genomic Variant Detection in Single Cells with Primary Template-Directed Amplification. *Proc. Natl. Acad. Sci.* 118, e2024176118. doi:10.1073/PNAS.2024176118

Gossen, J. A., De Leeuw, W. J. F., Tan, C. H. T., Zwarthoff, E. C., Berends, F., Lohman, P. H. M., et al. (1989). Efficient rescue of Integrated Shuttle Vectors from Transgenic Mice: A Model for Studying Mutations *In Vivo. Proc. Natl. Acad. Sci. U. S. A.* 86, 7971–7975. doi:10.1073/pnas.86.20.7971

Green, P., Ewing, B., Miller, W., Thomas, P. J., Thomas, J., Touchman, J., et al. (2003). Transcription-associated Mutational Asymmetry in Mammalian Evolution. *Nat. Genet.* 33, 514–517. doi:10.1038/ng1103

Green, A. C., Williams, G. M., Logan, V., and Strutton, G. M. (2011). Reduced Melanoma after Regular Sunscreen Use: Randomized Trial Follow-Up. *J. Clin. Oncol.* 29, 257–263. doi:10.1200/JCO.2010.28.7078

Gurjao, C., Zhong, R., Haruki, K., Li, Y. Y., Spurr, L. F., Lee-Six, H., et al. (2021). Discovery and Features of an Alkylating Signature in Colorectal Cancer. *Cancer Discov.* 11, candisc.1656.2020. doi:10.1158/2159-8290.CD-20-1656

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674. doi:10.1016/j.cell.2011.02.013

Haradhvala, N. J., Polak, P., Stojanov, P., Covington, K. R., Shinbrot, E., Hess, J. M., 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

Hawkins, C. J., and Miles, M. A. (2021). Mutagenic Consequences of Sublethal Cell Death Signaling. *Int. J. Mol. Sci.* 22, 6144. doi:10.3390/IJMS22116144

Hayes, A. W., Wang, T., and Dixon, D. (2020). Toxicologic Testing Methods. *Loomis's Essentials Toxicol.*, 189–222. doi:10.1016/B978-0-12-815921-7.00013-2

Helleday, T., Petermann, E., Lundin, C., Hodgson, B., and Sharma, R. A. (2008). DNA Repair Pathways as Targets for Cancer Therapy. *Nat. Rev. Cancer* 8, 193–204. doi:10.1038/nrc2342

Helleday, T., Eshtad, S., and Nik-Zainal, S. (2014). Mechanisms Underlying Mutational Signatures in Human Cancers. *Nat. Rev. Genet.* 15, 585–598. doi:10.1038/nrg3729

Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. (1991). p53 Mutations in Human Cancers. *Science* 253, 49–53. doi:10.1126/science.1905840

Hou, Y., Song, L., Zhu, P., Zhang, B., Tao, Y., Xu, X., et al. (2012). Single-cell Exome Sequencing and Monoclonal Evolution of a JAK2-Negative Myeloproliferative Neoplasm. *Cell* 148, 873–885. doi:10.1016/j.cell.2012.02.028

Hsu, J. I., Dayaram, T., Tovy, A., De Braekeleer, E., Jeong, M., Wang, F., et al. (2018). PPM1D Mutations Drive Clonal Hematopoiesis in Response to Cytotoxic Chemotherapy. *Cell Stem Cell* 23, 700–713.e6. doi:10.1016/j.stem.2018.10.004

Huang, M. N., Yu, W., Teoh, W. W., Ardin, M., Jusakul, A., Ng, A. W. T., et al. (2017). Genome-scale Mutational Signatures of Aflatoxin in Cells, Mice, and Human Tumors. *Genome Res.* 27, 1475–1486. doi:10.1101/gr.220038.116

Hutchinson, F. (1994). Induction of Tandem-Base Change Mutations. *Mutat. Res. Regul. Pap.* 309, 11–15. doi:10.1016/0027-5107(94)90037-X

Ichim, G., Lopez, J., Ahmed, S. U., Muthalagu, N., Giampazolias, E., Delgado, M. E., et al. (2015). Limited Mitochondrial Permeabilization Causes DNA Damage and Genomic Instability in the Absence of Cell Death. *Mol. Cel* 57, 860–872. doi:10.1016/j.molcel.2015.01.018

Jager, M., Blokzijl, F., Sasselli, V., Boymans, S., Janssen, R., Besselink, N., et al. (2018). Measuring Mutation Accumulation in Single Human Adult Stem Cells by Whole-Genome Sequencing of Organoid Cultures. *Nat. Protoc.* 13, 59–78. doi:10.1038/nprot.2017.111

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., et al. (2014). Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N. Engl. J. Med.* 371, 2488–2498. doi:10.1056/NEJMoa1408617

Jiao, W., Atwal, G., Polak, P., Karlic, R., Cuppen, E., Danyi, A., et al. (2020). A Deep Learning System Accurately Classifies Primary and Metastatic Cancers Using Passenger Mutation Patterns. *Nat. Commun.* 11, 1–12. doi:10.1038/s41467-019-13825-8

Johnson, G. E. (2012). Mammalian Cell HPRT Gene Mutation Assay: Test Methods. *Methods Mol. Biol.* 817, 55–67. doi:10.1007/978-1-61779-421-6_4

Jones, C. L., Degasperi, A., Grandi, V., Amarante, T. D., Ambrose, J. C., Arumugam, P., et al. (2021). Spectrum of Mutational Signatures in T-Cell Lymphoma Reveals a Key Role for UV Radiation in Cutaneous T-Cell Lymphoma. *Sci. Rep.* 11, 3962. doi:10.1038/s41598-021-83352-4

Kocakavuk, E., Anderson, K. J., Varn, F. S., Johnson, K. C., Amin, S. B., Sulman, E. P., et al. (2021). Radiotherapy Is Associated with a Deletion Signature that Contributes to Poor Outcomes in Patients with Cancer. *Nat. Genet.* 53, 1088–1096. doi:10.1038/s41588-021-00874-3

Kucab, J. E., Zou, X., Morganella, S., Joel, M., Nanda, A. S., Nagy, E., et al. (2019). A Compendium of Mutational Signatures of Environmental Agents. *Cell* 177, 821–836.e16. doi:10.1016/j.cell.2019.03.001

Kuijk, E., Jager, M., van der Roest, B., Locati, M. D., Van Hoeck, A., Korzelius, J., et al. (2020). The Mutational Impact of Culturing Human Pluripotent and Adult Stem Cells. *Nat. Commun.* 11, 2493. doi:10.1038/s41467-020-16323-4

Kunkel, T. A., and Bebenek, K. (2000). DNA Replication Fidelity. *Annu. Rev. Biochem.* 69, 497–529. doi:10.1146/annurev.biochem.69.1.497

Langenbucher, A., Bowen, D., Sakhtemani, R., Bournique, E., Wise, J. F., Zou, L., et al. (2021). An Extended APOBEC3A Mutation Signature in Cancer. *Nat. Commun.* 12, 1602. doi:10.1038/s41467-021-21891-0

Lawson, A. A. R. J., Abascal, F., Coorens, T. H. H., Hooks, Y., O'Neill, L., Latimer, C., et al. (2020). Extensive Heterogeneity in Somatic Mutation and Selection in the Human Bladder. *Science* 370, 75–82. doi:10.1126/science.aba8347

Lee-Six, H., Øbro, N. F., Shepherd, M. S., Grossmann, S., Dawson, K., Belmonte, M., et al. (2018). Population Dynamics of normal Human Blood Inferred from Somatic Mutations. *Nature* 561, 473–478. doi:10.1038/s41586-018-0497-0

Lee-Six, H., Olafsson, S., Ellis, P., Osborne, R. J., Sanders, M. A., Moore, L., et al. (2019). The Landscape of Somatic Mutation in normal Colorectal Epithelial Cells. *Nature* 574, 532–537. doi:10.1038/s41586-019-1672-7

Lee-Six, H. (2018). *Somatic Evolution in Human Blood and colon.* PhD Thesis Cambridge. doi:10.17863/CAM.37059

Li, B., Brady, S. W., Ma, X., Shen, S., Zhang, Y., Li, Y., et al. (2019). Therapy-Induced Mutations Drive the Genomic Landscape of Relapsed Acute Lymphoblastic Leukemia. *Blood* 135, 41–55. doi:10.1182/blood.2019002220

Li, Y., Roberts, N. D., Wala, J. A., Shapira, O., Schumacher, S. E., Kumar, K., et al. (2020). Patterns of Somatic Structural Variation in Human Cancer Genomes. *Nature Commun.* 578, 112–121. doi:10.1038/s41586-019-1913-9

Lindahl, T., and Barnes, D. E. (2000). Repair of Endogenous DNA Damage. *Cold Spring Harb. Symp. Quant. Biol.* 65, 127–133. doi:10.1101/sqb.2000.65.127

Lodato, M. M. A., Woodworth, M. B., Lee, S. S., Evrony, G. D., Mehta, B. K., Karger, A., et al. (2015). Somatic Mutation in Single Human Neurons Tracks Developmental and Transcriptional History. *Science* 350, 94–98. doi:10.1126/science.aab1785

Lodato, M. A., Rodin, R. E., Bohrson, C. L., Coulter, M. E., Barton, A. R., Kwon, M., et al. (2018). Aging and Neurodegeneration Are Associated with Increased Mutations in Single Human Neurons. *Science* 359, 555–559. doi:10.1126/science.aao4426

Loeb, L. A., Loeb, K. R., and Anderson, J. P. (2003). Multiple Mutations and Cancer. *Proc. Natl. Acad. Sci.* 100, 776–781. doi:10.1073/PNAS.0334858100

Machado, H. E., Mitchell, E., Øbro, N. F., Kübler, K., Davies, M., Maura, F., et al. (2021). Genome-wide Mutational Signatures of Immunological Diversification in normal Lymphocytes. *bioRxiv.* doi:10.1101/2021.04.29.441939

Makova, K. D., and Hardison, R. C. (2015). The Effects of Chromatin Organization on Variation in Mutation Rates in the Genome. *Nat. Rev. Genet.* 164, 213–223. doi:10.1038/nrg3890

Martínez-Jiménez, F., Muiños, F., Sentís, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C., et al. (2020). A Compendium of Mutational Cancer Driver Genes. *Nat. Rev. Cancer* 20, 555–572. doi:10.1038/s41568-020-0290-x

Martincorena, I., and Campbell, P. J. (2015). Somatic Mutation in Cancer and normal Cells. *Science* 349, 1483–1489. doi:10.1126/science.aab4082

Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., McLaren, S., et al. (2015). High burden and Pervasive Positive Selection of Somatic Mutations in normal Human Skin. *Science* 348 (6237), 880–886. doi:10.1126/science.aaa6806

Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Van Loo, P., et al. (2017). Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* 171, 1029–1041.e21. doi:10.1016/j.cell.2017.09.042

Martincorena, I., Fowler, J. C., Wabik, A., Lawson, A. R. J., Abascal, F., Hall, M. W. J., et al. (2018). Somatic Mutant Clones Colonize the Human Esophagus with Age. *Science* 362, 911–917. doi:10.1126/science.aau3879

Maura, F., Degasperi, A., Nadeu, F., Leongamornlert, D., Davies, H., Moore, L., et al. (2019). A Practical Guide for Mutational Signature Analysis in Hematological Malignancies. *Nat. Commun.* 10, 2969 doi:10.1038/s41467-019-11037-8

Miller, C. A., White, B. S., Dees, N. D., Griffith, M., Welch, J. S., Griffith, O. L., et al. (2014). SciClone: Inferring Clonal Architecture and Tracking the Spatial and Temporal Patterns of Tumor Evolution. *PLOS Comput. Biol.* 10, e1003665. doi:10.1371/JOURNAL.PCBI.1003665

Moore, L., Leongamornlert, D., Coorens, T. H. H., Sanders, M. A., Ellis, P., Dentro, S. C., et al. (2020). The Mutational Landscape of normal Human Endometrial Epithelium. *Nature* 580, 640–646. doi:10.1038/s41586-020-2214-z

Morganella, S., Alexandrov, L. B., Glodzik, D., Zou, X., Davies, H., Staaf, J., et al. (2016). The Topography of Mutational Processes in Breast Cancer Genomes. *Nat. Commun.* 7 (1), 11383. doi:10.1038/ncomms11383

Mortelmans, K., and Zeiger, E. (2000). The Ames Salmonella/microsome Mutagenicity Assay. *Mutat. Res.* 455, 29–60. doi:10.1016/s0027-5107(00)00064-6

Morton, L. M., Dores, G. M., Schonfeld, S. J., Linet, M. S., Sigel, B. S., Lam, C. J. K., et al. (2019). Association of Chemotherapy for Solid Tumors with Development of Therapy-Related Myelodysplastic Syndrome or Acute Myeloid Leukemia in the Modern Era. *JAMA Oncol.* 5, 318–325. doi:10.1001/jamaoncol.2018.5625

Morton, L. M., Karyadi, D. M., Stewart, C., Bogdanova, T. I., Dawson, E. T., Steinberg, M. K., et al. (2021). Radiation-related Genomic Profile of Papillary Thyroid Cancer after the Chernobyl Accident. *Science* 2538, 1–21. doi:10.1126/science.abg2538

Ng, A. W. T., Poon, S. L., Huang, M. N., Lim, J. Q., Boot, A., Yu, W., et al. (2017). Aristolochic Acids and Their Derivatives Are Widely Implicated in Liver Cancers in Taiwan and throughout Asia. *Sci. Transl. Med.* 9, eaan6446. doi:10.1126/scitranslmed.aan6446

Nguyen, L., Martens, W. M. J., Van Hoeck, A., and Cuppen, E. (2020). Pan-Cancer Landscape of Homologous Recombination Deficiency. *Nat. Commun.* 11, 1–12. doi:10.1038/s41467-020-19406-4

Nik-Zainal, S., Alexandrov, L. B., Wedge, D. C., Van Loo, P., Greenman, C. D., Raine, K., et al. (2012a). Mutational Processes Molding the Genomes of 21 Breast Cancers. *Cell* 149, 979–993. doi:10.1016/j.cell.2012.04.024

Nik-Zainal, S., Van Loo, P., Wedge, D. C., Alexandrov, L. B., Greenman, C. D., Lau, K. W., et al. (2012b). The Life History of 21 Breast Cancers. *Cell* 149, 994–1007. doi:10.1016/j.cell.2012.04.023

Nik-Zainal, S., Kucab, J. E., Morganella, S., Glodzik, D., Alexandrov, L. B., Arlt, V. M., et al. (2015). The Genome as a Record of Environmental Exposure. *Mutagenesis* 30, 763–770. doi:10.1093/mutage/gev073

Olafsson, S., McIntyre, R. E., Coorens, T., Butler, T., Jung, H., Robinson, P. S., et al. (2020). Somatic Evolution in Non-neoplastic IBD-Affected Colon. *Cell* 182, 672–684.e11. doi:10.1016/j.cell.2020.06.036

Osorio, F. G., Rosendahl Huber, A., Oka, R., Verheul, M., Patel, S. H., Hasaart, K., et al. (2018). Somatic Mutations Reveal Lineage Relationships and Age-Related Mutagenesis in Human Hematopoiesis. *Cel Rep.* 25, 2308–2316.e4. doi:10.1016/j.celrep.2018.11.014

Parkin, D. M., Mesher, D., and Sasieni, P. (2011). Cancers Attributable to Solar (Ultraviolet) Radiation Exposure in the UK in 2010. *Br. J. Cancer* 105, S66–S69. doi:10.1038/bjc.2011.486

Perera, D., Poulos, R. C., Shah, A., Beck, D., Pimanda, J. E., and Wong, J. W. H. (2016). Differential DNA Repair Underlies Mutation Hotspots at Active Promoters in Cancer Genomes. *Nature* 532, 259–263. Available at: https://www.nature.com/articles/nature17437 (Accessed September 29, 2021). doi:10.1038/nature17437

Peto, R., Darby, S., Deo, H., Silcocks, P., Whitley, E., and Doll, R. (2000). Smoking, Smoking Cessation, and Lung Cancer in the UK since 1950: Combination of National Statistics with Two Case-Control Studies. *BMJ* 321, 323–329. doi:10.1136/bmj.321.7257.323

Pfeifer, G. P., Denissenko, M. F., Olivier, M., Tretyakova, N., Hecht, S. S., and Hainaut, P. (2002). Tobacco Smoke Carcinogens, DNA Damage and P53 Mutations in Smoking-Associated Cancers. *Oncogene* 21 (48), 7435–7451. doi:10.1038/sj.onc.1205803

Pich, O., Muiños, F., Lolkema, M. P., Steeghs, N., Gonzalez-Perez, A., and Lopez-Bigas, N. (2019). The Mutational Footprints of Cancer Therapies. *Nat. Genet.* 51, 1732–1740. doi:10.1038/s41588-019-0525-5

Pich, O., Cortes-Bullich, A., Muiños, F., Pratcorona, M., Gonzalez-Perez, A., and Lopez-Bigas, N. (2021). The Evolution of Hematopoietic Cells under Cancer Therapy. *Nat. Commun.* 12, 1–11. doi:10.1038/s41467-021-24858-3

Pleguezuelos-Manzano, C., Puschhof, J., Rosendahl Huber, A., van Hoeck, A., Wood, H. M., Nomburg, J., et al. (2020). Mutational Signature in Colorectal Cancer Caused by Genotoxic Pks + E. coli. *Nature* 580, 269–273. doi:10.1038/s41586-020-2080-8

Pollock, P. M., Harper, U. L., Hansen, K. S., Yudt, L. M., Stark, M., Robbins, C. M., et al. (2003). High Frequency of BRAF Mutations in Nevi. *Nat. Genet.* 33, 19–20. doi:10.1038/ng1054

Poon, L. S., Huang, N. M., Choo, Y., McPherson, J. R., Yu, W., Heng, L. H., et al. (2015). Mutation Signatures Implicate Aristolochic Acid in Bladder Cancer Development. *Genome Med.* 7, 38. doi:10.1186/s13073-015-0161-3

Poulos, R. C., Wong, Y. T., Ryan, R., Pang, H., and Wong, J. W. H. (2018). Analysis of 7,815 Cancer Exomes Reveals Associations between Mutational Processes and Somatic Driver Mutations. *Plos Genet.* 14, 1–20. doi:10.1371/journal.pgen.1007779

Priestley, P., Baber, J., Lolkema, M., Steeghs, N., de Bruijn, E., Shale, C., et al. (2019a). Pan-cancer Whole Genome Analyses of Metastatic Solid Tumors. *bioRxiv*, 415133. doi:10.1101/415133

Priestley, P., Baber, J., Lolkema, M. P., Steeghs, N., de Bruijn, E., Shale, C., et al. (2019b). Pan-cancer Whole-Genome Analyses of Metastatic Solid Tumours. *Nature* 575, 210–216. doi:10.1038/s41586-019-1689-y

Quispe-Tintaya, W., Lee, M., Dong, X., Weiser, D. A., Vijg, J., and Maslov, A. Y. (2018). Bleomycin-induced Genome Structural Variations in normal, Non-tumor Cells. *Sci. Rep.* 8, 16523. doi:10.1038/s41598-018-34580-8

Riva, L., Pandiri, A. R., Li, Y. R., Droop, A., Hewinson, J., Quail, M. A., et al. (2020). The Mutational Signature Profile of Known and Suspected Human Carcinogens in Mice. *Nat. Genet.* 52, 1189–1197. doi:10.1038/s41588-020-0692-4

Roberts, S. A., Lawrence, M. S., Klimczak, L. J., Grimm, S. A., Fargo, D., Stojanov, P., et al. (2013). An APOBEC Cytidine Deaminase Mutagenesis Pattern Is Widespread in Human Cancers. *Nat. Genet.* 45, 970–976. doi:10.1038/ng.2702

Roerink, S. F., Sasaki, N., Lee-Six, H., Young, M. D., Alexandrov, L. B., Behjati, S., et al. (2018). Intra-tumour Diversification in Colorectal Cancer at the Single-Cell Level. *Nature* 556, 437–462. doi:10.1038/s41586-018-0024-3

Rose Li, Y., Halliwill, K. D., Adams, C. J., Iyer, V., Riva, L., Mamunur, R., et al. (2020). Mutational Signatures in Tumours Induced by High and Low Energy Radiation in Trp53 Deficient Mice. *Nat. Commun.* 11. doi:10.1038/s41467-019-14261-4

Rosendahl Huber, A., Manders, F., Oka, R., and van Boxtel, R. (2019). Characterizing Mutational Load and Clonal Composition of Human Blood. *J. Vis. Exp* 149, e59846. doi:10.3791/59846

Roth, A., Khattra, J., Yap, D., Wan, A., Laks, E., Biele, J., et al. (2014). PyClone: Statistical Inference of Clonal Population Structure in Cancer. *Nat. Methods* 11, 396–398. doi:10.1038/nmeth.2883

Russo, M., Crisafulli, G., Sogari, A., Reilly, N. M., Arena, S., Lamba, S., et al. (2019). Adaptive Mutability of Colorectal Cancers in Response to Targeted Therapies. *Science* 366, 1473–1480. doi:10.1126/science.aav4474

Sabarinathan, R., Mularoni, L., Deu-Pons, J., Gonzalez-Perez, A., and Lopez-Bigas, N. (2016). Nucleotide Excision Repair Is Impaired by Binding of Transcription Factors to DNA. *Nature* 532. doi:10.1038/nature17661

Sanders, A. M. A., Vöhringer, H., Forster, V. J., Moore, L., and Brittany, B. (2021). Life without Mismatch Repair. *bioRxiv*. doi:10.1101/2021.04.14.437578

Sculley, D. G., Dawson, P. A., Emmerson, B. T., and Gordon, R. B. (1992). A Review of the Molecular Basis of Hypoxanthine-Guanine Phosphoribosyltransferase (HPRT) Deficiency. *Hum. Genet.* 90, 195–207. doi:10.1007/BF00220062

Secrier, M., Li, X., De Silva, N., Eldridge, M. D., Contino, G., Bornschein, J., et al. (2016). Mutational Signatures in Esophageal Adenocarcinoma Define Etiologically Distinct Subgroups with Therapeutic Relevance. *Nat. Genet.* 48, 1131–1141. doi:10.1038/ng.3659

Seplyarskiy, V. B., Akkuratov, E. E., Akkuratova, N., Andrianova, M. A., Nikolaev, S. I., Bazykin, G. A., et al. (2018). Error-prone Bypass of DNA Lesions during Lagging-Strand Replication Is a Common Source of Germline and Cancer Mutations. *Nat. Genet.* 51, 36–41. doi:10.1038/S41588-018-0285-7

Spira, A., Yurgelun, M. B., Alexandrov, L., Rao, A., Bejar, R., Polyak, K., et al. (2017). Precancer Atlas to Drive Precision Prevention Trials. *Cancer Res.* 77, 1510–1541. doi:10.1158/0008-5472.CAN-16-2346

Stratton, M. R. M. R., Campbell, P. J. P. J., and Futreal, P. A. A. (2009). The Cancer Genome. *Nature* 458, 719–724. doi:10.1038/nature07943

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Supek, F., and Lehner, B. (2015). Differential DNA Mismatch Repair Underlies Mutation Rate Variation across the Human Genome. *Nature* 521, 81–84. doi:10.1038/nature14173

Tang, J., Fewings, E., Chang, D., Zeng, H., Liu, S., Jorapur, A., et al. (2020). The Genomic Landscapes of Individual Melanocytes from Human Skin. *Nature* 586, 600–605. doi:10.1038/s41586-020-2785-8

Temko, D., Tomlinson, I. P. M., Severini, S., Schuster-Böckler, B., and Graham, T. A. (2018). The Effects of Mutational Processes and Selection on Driver Mutations across Cancer Types. *Nat. Commun.* 9 (1), 1857. doi:10.1038/s41467-018-04208-6

Terlouw, D., Suerink, M., Boot, A., van Wezel, T., Nielsen, M., and Morreau, H. (2020). Recurrent APC Splice Variant c.835-8A>G in Patients with Unexplained Colorectal Polyposis Fulfilling the Colibactin Mutational Signature. *Gastroenterology* 159, 1612–1614.e5. doi:10.1053/j.gastro.2020.06.055

Tomasetti, C., and Vogelstein, B. (2015). Variation in Cancer Risk Among Tissues Can Be Explained by the Number of Stem Cell Divisions. *Science* 347, 78–81. doi:10.1126/science.1260825

Tomasetti, C., Li, L., and Vogelstein, B. (2017). Stem Cell Divisions, Somatic Mutations, Cancer Etiology, and Cancer Prevention. *Science* 355, 1330–1334. doi:10.1126/science.aaf9011

Tomkova, M., Tomek, J., Kriaucionis, S., and Schuster-Böckler, B. (2018). Mutational Signature Distribution Varies with DNA Replication Timing and Strand Asymmetry. *Genome Biol.* 19, 1–12. doi:10.1186/s13059-018-1509-y

Touat, M., Li, Y. Y., Boynton, A. N., Spurr, L. F., Iorgulescu, J. B., Bohrson, C. L., et al. (2020). Mechanisms and Therapeutic Implications of Hypermutation in Gliomas. *Nature* 580, 517–523. doi:10.1038/s41586-020-2209-9

Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell* 168, 644–656. doi:10.1016/j.cell.2017.01.002

Van Hoeck, A., Tjoonk, N. H., Van Boxtel, R., and Cuppen, E. (2019). Portrait of a Cancer: Mutational Signature Analyses for Cancer Diagnostics. *BMC Cancer* 19, 1–14. doi:10.1186/s12885-019-5677-2

Vermeulen, W., and Fousteri, M. (2013). Mammalian Transcription-Coupled Excision Repair. *Cold Spring Harb. Perspect. Biol.* 5, a012625. doi:10.1101/cshperspect.a012625

Vijg, J., Dong, X., and Zhang, L. (2017). A High-Fidelity Method for Genomic Sequencing of Single Somatic Cells Reveals a Very High Mutational burden. *Exp. Biol. Med.* 242, 1318–1324. doi:10.1177/1535370217717696

Vöhringer, H., Hoeck, A. V., Cuppen, E., and Gerstung, M. (2021). Learning Mutational Signatures and Their Multidimensional Genomic Properties with TensorSignatures. *Nat. Commun.* 12, 3628. doi:10.1038/s41467-021-23551-9

Volkova, N. V., Meier, B., González-Huici, V., Bertolini, S., Gonzalez, S., Vöhringer, H., et al. (2020). Mutational Signatures Are Jointly Shaped by DNA Damage and Repair. *Nat. Commun.* 11, 1–15. doi:10.1038/s41467-020-15912-7

Wang, Y., Waters, J., Leung, M. L., Unruh, A., Roh, W., Shi, X., et al. (2014). Clonal Evolution in Breast Cancer Revealed by Single Nucleus Genome Sequencing. *Nature* 512, 155–160. doi:10.1038/NATURE13600

Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., et al. (2012). The Origin and Evolution of Mutations in Acute Myeloid Leukemia. *Cell* 150 (2), 264–278. doi:10.1016/j.cell.2012.06.023

Wernke, K. M., Xue, M., Tirla, A., Kim, C. S., Crawford, J. M., and Herzon, S. B. (2020). Structure and Bioactivity of Colibactin. *Bioorg. Med. Chem. Lett.* 30, 127280. doi:10.1016/j.bmcl.2020.127280

Wong, T. N., Ramsingh, G., Young, A. L., Miller, C. A., Touma, W., Welch, J. S., et al. (2015). Role of TP53 Mutations in the Origin and Evolution of Therapy-Related Acute Myeloid Leukaemia. *Nature* 518, 552–555. doi:10.1038/nature13968

Yeager, M., Machiela, M. J., Kothiyal, P., Dean, M., Bodelon, C., Suman, S., et al. (2021). Lack of Transgenerational Effects of Ionizing Radiation Exposure from the Chernobyl Accident. *Science* 372, 725–729. doi:10.1126/SCIENCE.ABG2365

Yokoyama, A., Kakiuchi, N., Yoshizato, T., Nannya, Y., Suzuki, H., Takeuchi, Y., et al. (2019). Age-related Remodelling of Oesophageal Epithelia by Mutated Cancer Drivers. *Nature* 565, 312–317. doi:10.1038/s41586-018-0811-x

Yoshida, K., Gowers, K. H. C., Lee-Six, H., Chandrasekharan, D. P., Coorens, T., Maughan, E. F., et al. (2020). Tobacco Smoking and Somatic Mutations in Human Bronchial Epithelium. *Nature* 578, 266–272. doi:10.1038/s41586-020-1961-1

Zeiger, E. (2019). The Test that Changed the World: The Ames Test and the Regulation of Chemicals. *Mutat. Res. Toxicol. Environ. Mutagen.* 841, 43–48. doi:10.1016/j.mrgentox.2019.05.007

Ziegler, A., Leffell, D. J., Kunala, S., Sharma, H. W., Gailani, M., Simon, J. A., et al. (1993). Mutation Hotspots Due to Sunlight in the P53 Gene of Nonmelanoma Skin Cancers. *Proc. Natl. Acad. Sci. U. S. A.* 90, 4216–4220. doi:10.1073/pnas.90.9.4216

Zink, F., Stacey, S. N., Norddahl, G. L., Frigge, M. L., Magnusson, O. T., Jonsdottir, I., et al. (2017). Clonal Hematopoiesis, with and without Candidate Driver Mutations, Is Common in the Elderly. *Blood* 130, 742–752. doi:10.1182/BLOOD-2017-02-769869

Zou, X., Owusu, M., Harris, R., Jackson, S. P., Loizou, J. I., and Nik-Zainal, S. (2018). Validating the Concept of Mutational Signatures with Isogenic Cell Models. *Nat. Commun.* 9, 1–16. doi:10.1038/s41467-018-04052-8

Zou, X., Koh, G. C. C., Nanda, A. S., Degasperi, A., Urgo, K., Roumeliotis, T. I., et al. (2021). A Systematic CRISPR Screen Defines Mutational Mechanisms Underpinning Signatures Caused by Replication Errors and Endogenous DNA Damage. *Nat. Cancer* 2, 643–657. doi:10.1038/s43018-021-00200-0