

Review Article

Spectrum of DNA mismatch repair failures viewed through the lens of cancer genomics and implications for therapy

 David Mas-Ponte¹,  Marcel McCullough¹ and  Fran Supek^{1,2}

¹Genome Data Science, Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute for Science and Technology, Baldori Reixac 10, Barcelona 08028, Spain;

²Catalan Institution for Research and Advanced Studies (ICREA), Pg Lluís Companys, 23, Barcelona 08010, Spain

Correspondence: Fran Supek (fran.supek@irbbarcelona.org)



Genome sequencing can be used to detect DNA repair failures in tumors and learn about underlying mechanisms. Here, we synthesize findings from genomic studies that examined deficiencies of the DNA mismatch repair (MMR) pathway. The impairment of MMR results in genome-wide hypermutation and in the ‘microsatellite instability’ (MSI) phenotype—occurrence of indel mutations at short tandem repeat (microsatellite) loci. The MSI status of tumors was traditionally assessed by molecular testing of a selected set of MS loci or by measuring MMR protein expression levels. Today, genomic data can provide a more complete picture of the consequences on genomic instability. Multiple computational studies examined somatic mutation distributions that result from failed DNA repair pathways in tumors. These include analyzing the commonly studied trinucleotide mutational spectra of single-nucleotide variants (SNVs), as well as of other features such as indels, structural variants, mutation clusters and regional mutation rate redistribution. The identified mutation patterns can be used to rigorously measure prevalence of MMR failures across cancer types, and potentially to subcategorize the MMR deficiencies. Diverse data sources, genomic and pre-genomic, from human and from experimental models, suggest there are different ways in which MMR can fail, and/or that the cell-type or genetic background may result in different types of MMR mutational patterns. The spectrum of MMR failures may direct cancer evolution, generating particular sets of driver mutations. Moreover, MMR affects outcomes of therapy by DNA damaging drugs, antimetabolites, nonsense-mediated mRNA decay (NMD) inhibitors, and immunotherapy by promoting either resistance or sensitivity, depending on the type of therapy.

Cancer sequencing studies have yielded a repository of genomic data of an extraordinary scale, which is growing by the day. Major catalogs of tumor genomes include The Cancer Genome Atlas (TCGA), containing approximately 10000 tumor whole-exome sequences (WESs) [1,2], the Pan-cancer Analysis of Whole Genomes (PCAWG) consortium [3], with approximately 2700 whole-genome sequences (WGSs), and the Hartwig Medical Foundation (HMF) dataset of approximately 2500 WGS metastatic tumors as published [4], and growing. A major motivation behind sequencing a large number of cancer exomes or genomes was to discover new cancer driver genes. This was largely a success: approximately 300–800 cancer genes, depending on methodology [1,5], were identified as recurrently mutated, indicating positive selection in tumors [6,7] and suggesting a role in driving cancer evolution. In addition to driver gene discovery, the tumor genomic data were used to study mutational processes by analyzing the passenger (non-selected) mutations, which are the majority of mutations observed in a typical cancer genome. These analyses helped elucidate mechanisms of mutagenesis and DNA repair that operate in tumor cells

Received: 01 December 2021
Revised: 02 February 2022
Accepted: 28 February 2022

Version of Record published:
11 March 2022

[8–12]. The mutation spectra are often used to derive ‘mutational signatures’, an output of a computer algorithm that captures the DNA sequence propensity towards distinct mutational processes that are active across many tumors; see below for more detail. These recent advances in genomic data generation and in the methodologies for the extraction of mutational signatures have revolutionized the study of DNA repair deficiencies in tumors. These are a common occurrence and an important mechanism that promotes cancer evolution.

Components of the mismatch repair pathway and their causative role in cancer risk

DNA mismatch repair (MMR) is a highly conserved system with a major role in maintaining genomic stability. A proficient MMR mechanism greatly increases DNA replication fidelity by detecting and initiating the replacement of DNA base mismatches and small loops (reviewed in [13–15]). MMR can also terminate erroneous recombination attempts (reviewed by [16]) and initiate a DNA damage response signal [17,18]. This article will not be dissecting the MMR pathway and will instead focus on the effects of a deficient MMR machinery on the genomic landscape, and the relevance thereof to cancer therapy. Nonetheless, we will describe in brief the most relevant MMR factors, further referring the reader to other excellent reviews on the molecular biology of the MMR system [15,19–22].

MMR has two crucial components, MutS, which detects the mismatches, and MutL, necessary for mismatch processing and excision. In human cells, mismatches and short loops of one to two nucleotides are detected by MutS α (a heterodimer of MSH2 and MSH6) and loops of varying sizes are detected by MutS β (MSH2 associated with MSH3) [23–25]; unrepaired mismatches result in single-nucleotide substitution mutations, while unrepaired loops in DNA result in indel mutations. MutL is also a heterodimer with three different types in human cells: MutL α (MLH1 and PMS2), MutL β (MLH1 and PMS1) and MutL γ (MLH1 and MLH3) [26–28]. Currently, only MutL α is considered to have a notable role in the general process of post-replicative MMR. The EXO1 nuclease is another important protein that participates in the MMR process during mismatch excision [29].

The most severe phenotypes, both in human cancers [30] and in mouse studies, have been observed when the cells lack MSH2 or MLH1, with either knockout behaving similarly [31]. This indicates that MMR cannot resolve the mutations without either MutS or MutL complexes. Mutations in other MutS and MutL genes, such as *MSH6* and *PMS2*, have also been associated with cancer development at lower rates [15,21,32]. In mouse knockout studies, *Msh6*-deficient mice have a similar phenotype to *Msh2*-null mice, but can repair larger insertion-deletions and tumors with longer latencies [33], mirroring human tumors lacking MSH6 [20]. *Msh3* (MSH3 in human) deficiencies have much milder phenotypes [34,35]. As for the MutL heterodimer, *Pms2*^{-/-} mice develop cancers less frequently than *Mlh1*^{-/-} and *Msh2*^{-/-}, suggesting that *Pms1* and/or *Mlh3* might be partially substituting for *Pms2* [36,37].

Detecting microsatellite instability resulting from MMR failures in tumors

The MMR pathway sometimes malfunctions in cancer. Dysfunctional MMR is often diagnosed in the clinic by measuring the levels of MLH1, MSH2, MSH6 and PMS2 proteins by immunohistochemistry. However, some cancers with defective MMR have normal expression levels of these proteins and so the activity of the MMR pathway can be more directly observed via a ‘microsatellite instability’ (MSI) mutational phenotype. In MSI, the lengths of microsatellite (MS) loci in DNA become unstable (these loci are also called ‘short tandem repeats’ or ‘simple sequence repeats’, here referring to those consisting of 1-mer to 4-mer repeat units, which are used to infer MMR activity; of note, a common definition of MS includes repeats with up to 6-mer units). Effectively, such DNA repeats contract and expand via indel mutations at very high rates, such that a population of cells may contain a mixture of alleles of various lengths at a single locus. This property allows MMR failures to be measured via a PCR and capillary electrophoresis-based assay, which provides a facile readout of MS lengths at chosen loci. Typically, five loci are examined that are a part of the Bethesda panel [38]: two mononucleotide repeats (Big Adenine Tract [BAT]-25 and BAT-26) and three dinucleotide repeats (D2S123, D5S346 and D17S250). Instability at ≥ 2 Bethesda loci means the tumor is MSI-H (unstable, high), at 0 loci is MSS (stable), while the intermediate case with only 1 unstable locus is labeled MSI-L (low). By this commonly applied criterion, >10% of uterus, colorectal, and stomach cancers are MSI-H, as well as 1–10% in some other cancer types (esophagus, stomach, ovary, prostate, leukemia, brain); additionally, there exist rare examples of MSI-H in other cancers such as breast or skin cancers.

Alternative systems also exist, e.g. a pentaplex panel was derived by screening 266 repeats to focus on five loci that were highly discriminative for MSI samples [39]; they are all mononucleotide repeats and are almost monomorphic (BAT-25, BAT-26, NR-21, NR-24 and NR-27). These loci may provide a more accurate readout than the Bethesda

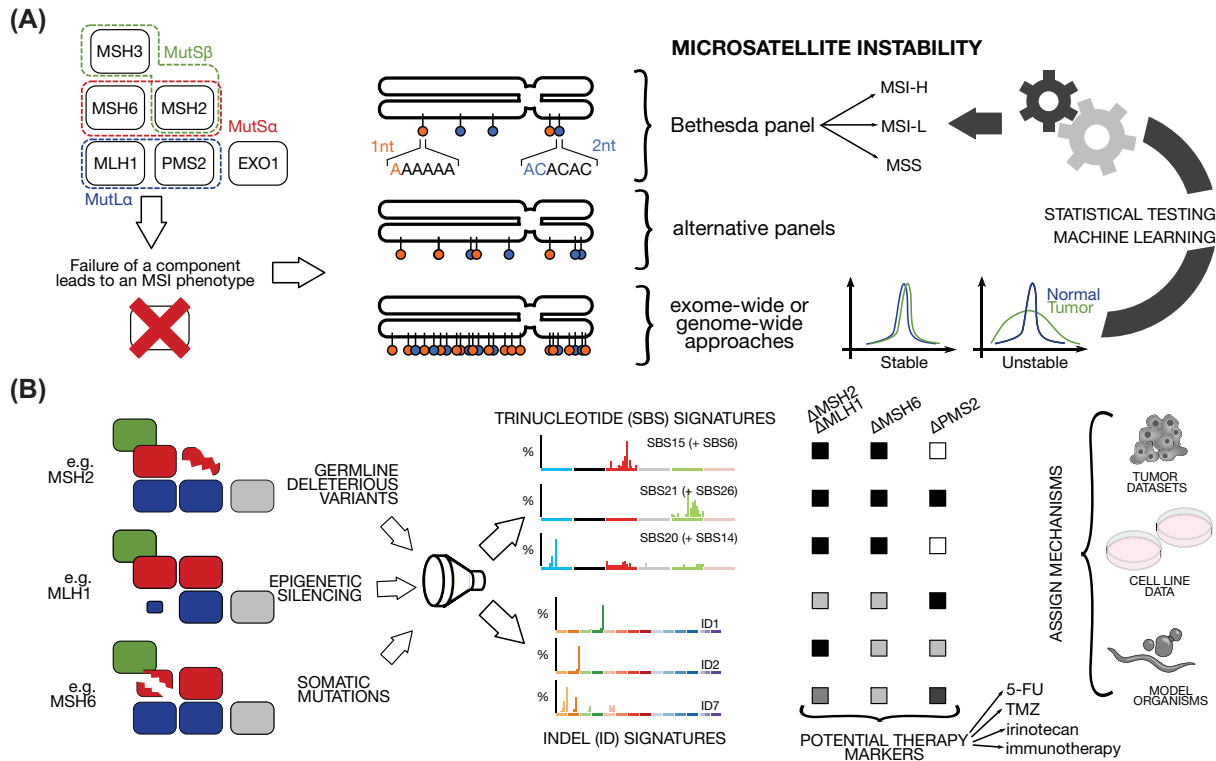


Figure 1. Genomic signatures of MMR failures, and possible subtypes thereof

(A) Methods for detecting MMR deficiencies (left) are based on assaying the instability of MS loci across the genome (middle). This can be done either by a PCR+electrophoresis experimental assay of various MS panels such as the commonly employed Bethesda panel (middle, above), or more recently by a statistical analysis of WGS/WES data across many MS loci simultaneously by various bioinformatics tools (middle, below). In both cases, the output is the binary-labeled MSS (stable) or MSI (unstable) (right). The Bethesda panel can distinguish MSI-H (high) versus MSI-L (low), however the significance of this distinction is unclear. (B) Different molecular mechanisms are known to cause MMR failures, either through pathogenic germline variation, somatic mutations and/or copy-number alterations, or epigenetic silencing of different MMR genes or MMR-associated genes (left). All converge on a set of broadly similar genomic patterns involving a high burden of SNV and indel mutations. However recent data suggest that these mechanisms may be separable based on classifying the SNVs and indels by type, as defined in mutational signatures (middle) observed in genomes of cancers and of experimental models where MMR deficiencies were induced (right). For example, *MSH6* LOF may generate similar burden of SNVs as *MLH1* or *MSH2* loss, however with fewer indels generated. Another example is that *PMS2* loss might generate a different SNV trinucleotide spectrum than the *MSH2* or *MLH1* loss. Abbreviation: SNV, single-nucleotide variant.

panel [40,41], for which it was noted that MMR-deficient cancers resulting from loss-of-function (LOF) in the *MSH6* gene are challenging to classify correctly, since *MSH6* loss does not cause strong instability at dinucleotide repeats. Simpler systems based on experimentally assaying only a single repeat locus (e.g. CAT-25 or HT-17) were also proposed, as well as more complex panels consisting of more than ten MS loci (see [42] for a recent review, and the references therein).

The advent of genomics has brought advances to the methods for determining MSI status. Instead of experimentally assaying only selected loci, a WES or a WGS can be analyzed, searching for indels at many MSs at once. Next, statistical analysis integrates the instability signal across various MS loci, and is used to predict global MSI or MSS status for each tumor sample (Figure 1A). These predictions often have high concordance with the standard PCR-based MSI assays, suggesting genomics can largely supplant the standard PCR+capillary electrophoresis methodology. In particular, in a head-to-head comparison of three genomic software tools against the PCR MSI calls as a gold standard in colorectal, uterine and stomach cancers, the concordance ranged from 83 to 97% for different tools [43]. It should be noted that the genomic algorithms may also reproduce the errors of these PCR-based MSI assays, whose output they were explicitly designed to mimic. Another limitation in using genomics to determine MSI status is the cost of

the sequencing, and also that some bioinformatics expertise is required to analyze the data. However, the tumor WES or WGS that have been generated for other diagnostic purposes (e.g. identifying actionable driver mutations) can also be used in the detection of MSI, thus reducing the global cost and complexity of the analysis.

Many software tools can predict MSI status from WES or WGS data, including MSISensor [44], mSINGs [45], Mantis [46] or MSISeq [47]. While the implementations differ, the general principle uniting these tools is that they analyze the distributions of MS lengths at many loci in short-read sequencing data. However, repeat-containing loci are particularly challenging for WGS/WES analyses, because of difficulties in aligning the short-reads to the reference genome and other types of artefacts (e.g. during library preparation, if using PCR-based protocols). These concerns are usually addressed by comparing the lengths of each locus between a tumor and a normal sample, which would, presumably, both be similarly affected by various biases, and a significant difference would reflect a genuine biological process. This is followed by a step where data are integrated across all loci to classify the sample as MSI, usually requiring that a certain percentage of the loci genome-wide are unstable (see a comparison in [43]). However, instead of simply counting stable versus unstable loci, more complex machine-learning approaches can also be used to classify MSI versus MSS, performing implicit selection of loci that maximize ability to discriminate known MSI and MSS examples on a training dataset [47,48].

Mutation signature analysis from cancer genomes can identify DNA repair phenotypes

Mutational processes that were operating in a tumor genome are often quantified by estimating risk of each possible trinucleotide to accrue different types of single-nucleotide variants (SNVs) [8]. The process to obtain this ‘mutational spectrum’ consists in the classification of the somatic mutations according to the sequence context features thought to be relevant for mutation rates. The mutation spectra of many tumor samples are then jointly input to a Non-negative Matrix Factorization (NMF) algorithm, which aims to generate independent patterns (‘signatures’) corresponding to mutational mechanisms, as well as the weights of each pattern in each tumor sample (often referred to as ‘exposures’); of note, in addition to NMF other factorization algorithms were also applied, such as topic models [49]. Commonly, the process is performed using the 5′ and 3′ adjacent nucleotides for the six possible mutation types (considered strand-symmetrically, e.g. a T > C mutation is tallied together with an A > G mutation), for a total of 96 categories. This concept was also extended to pentanucleotides [10,50,51], indels [8,51,52], double-nucleotide substitutions [51] and copy number alterations (CNAs) [53–56]. Due to the independence of each signature, it is often assumed that each pattern represents a single molecular mechanism, thus describing how mutations in that tumor were generated. For some mutational signatures, underlying mechanisms were identified via associations with carcinogenic exposures, for instance C > A mutations correlated with tobacco smoking exposure, YC > YT substitutions (where Y = C or T) correlated with UV exposure, A > T changes correlated with exposure to the aristolochic acid carcinogen, and TC > TG and TC > TT substitutions correlated with expression levels of the *APOBEC3A* gene encoding a mutagenic enzyme. Experimental work using controlled exposures of human cell lines to mutagenic agents usually recapitulated the signatures suspected to be generated by that agent in tumors [57–59], suggesting that mutational signatures are a useful tool to ascertain the mutagenic exposures of a tumor cell (or its non-cancerous progenitor). This is not limited to chemicals or radiation: also the DNA repair failures engineered into human cultured cells or into model organisms [52,60,61] result in mutation patterns resembling those found in tumor genomes. This suggests that genomic data analysis approaches can distinguish different DNA repair failure mechanisms. It also bears mentioning that results of mutational signature inference do vary depending on the computational method employed and on the parameters used in running the algorithm, and might also be affected by tissue-specificity to some extent [62,63]; additional variables that affect results are the method used to call the mutations, and the quality of sequencing data. Because of this variation, it is not straightforward to establish the mechanism underlying an observed mutational pattern by comparing to mutation signatures generated from model systems. There may not always be a one-to-one mapping—at least using current methods—between a specific DNA repair mechanism and an individual mutation signature. Extracted mutational signatures are often difficult to resolve, and some signatures may represent mixtures of mechanisms in various proportions. Additionally, currently most of the SNV trinucleotide signatures observed in tumors have not been linked to a specific mechanism at all; elucidating their mechanistic underpinnings is a matter of much interest. Some of the mutational signatures observed are thought to capture artefacts, either from sequencing or from idiosyncrasies of the computational analysis. Statistical methods for investigating mutational signatures are still actively developed. Recent and upcoming work [49,64–67] may address many of the standing issues with the methodologies.

In addition to trinucleotide signatures and derivatives thereof such as clustered mutation signatures [68,69], other types of mutation patterns may provide deeper insight into mutagenic mechanisms. For example, signatures can be derived from spectra of short indels [51], or from CNAs or other types of large-scale structural variation (also referred to as rearrangements) [54,70,71]. Even though mechanisms remain unclear for most of the proposed indel and rearrangement signatures, these methods did demonstrate practical utility. Applications to cancers deficient in homologous recombination repair (HRR) [55,72] has revealed distinct rearrangement patterns between subtypes of HRR deficiencies. The two genomic HRR deficiency signature subtypes associate with pathogenic variants in the *BRCA1* gene, versus *BRCA2* or *PALB2* genes; it is conceivable that this dichotomy might have implications for disease prognosis and treatment response. More generally, mutational signatures derived from SNVs and from CNA spectra were suggested to be predictive of survival in various tumors, for example ovarian and esophageal cancers [56,73].

In addition to SNVs, indels and rearrangements, other types of patterns in cancer genomes can inform mutational signature analyses. This includes, e.g. strand-biases of mutations, which can quantitate transcription-coupled and DNA replication-coupled mechanisms [74,75]. Another example is the local density of mutations, which correlates with activity of MMR and/or nucleotide excision repair pathways in that tumor [76–78]. Clustered mutagenesis describes mutations due to activity of the TLS enzymes (error-prone DNA polymerases) or to APOBEC3 mutagenic cytosine deaminases [68,69]. These various types of mutation patterns can be incorporated into statistical frameworks to jointly inform mutational signature analysis and help subclassify processes that generate or prevent mutations, including DNA repair pathways [67,79]. In particular, faults in the DNA MMR pathway are reflected in several different types of ‘mutational scars’ observed in cancer genomes. Joint analysis of different mutation pattern types has the potential to pinpoint the mechanism of how this important DNA repair pathway failed in each particular tumor.

Mutational patterns associated with MMR failures observed in tumor genomes

Genome-wide statistical analyses of MS indels across many different cancer types [48,80–82] have identified cancer driver mutations emerging at MS loci. Additionally, and more pertinent to the topic of this text, these studies have also characterized the genome-wide landscape of somatic indel mutations at all MS loci accessible to short-read WGS or WES (i.e. the majority of the human genome, except the highly repetitive DNA such as pericentromeric regions). One suggestive finding of these global analyses of thousands of tumors is that MSI may be present in additional cancer types than was previously appreciated. For instance, adrenocortical carcinoma, cervical cancer and mesothelioma present instances of genomically MSI-like tumors, which (as supporting evidence for MMR failures) also have higher overall mutation burdens [46] suggesting they are *bona fide* cases of failed MMR. Interestingly, these analyses hint at a tissue-specificity of MSI phenotypes—different MS loci appear unstable in different cancer types [81]. This is consistent with regional patterns observed in somatic SNVs, which also vary across cancer types, likely due to heterogeneity in the activity of the MMR pathway coupled to replication timing changes across tissues [76,83].

In addition to the MS-based genomics studies described above, the analysis of the SNV mutational signatures [8,51] has shown that several trinucleotide signatures associate with the MSI-H phenotype (COSMIC Signature 6, 14, 15, 20, 21, 26 and 44) in tumors from different organs. Why there exist so many different MMR-associated SNV signatures is enigmatic. Some of them—in particular Signature 14 and Signature 20—associate with concurrent mutations in replicative DNA polymerases (ϵ and δ , respectively) and so might plausibly reflect different combinations of mechanisms, however for the remainder of the MMR signatures differences in etiology are less clear. One possible explanation is that there are tissue-specific factors which can modulate a mutational process (e.g. differences in free nucleotide pools), thus resulting in somewhat different signatures resulting from that same process in each tissue [63].

The SNV trinucleotide mutational signatures can, similarly as a MS length analysis, be used to construct an MSI classification model based on genomics data (see e.g. MMRDetect, [60]). The more recently introduced indel mutational signatures also associate with MSI-H status in tumor genomes [51] (COSMIC signatures ID1, ID2 and ID7, containing an abundance of indels at MSs) and thus would be predictive, however this type of indel signatures would be largely redundant with the MS length analysis. In addition to genome sequences from human tumors, MMR failures have been induced in human cell lines, followed by WGS. This yielded increased mutational burdens and distinct trinucleotide SNV mutational patterns in cell lines, which were broadly comparable with those observed in cancer genomes for MMR genes commonly inactivated in cancer, *MLH1*, *MSH6*, *MSH2* and *PMS2* [60,84,85]. However it should be noted there were also differences in mutation patterns between these individual genotypes, as detailed below.

In summary, recent genomic comparative analyses of the MSI versus the MSS tumors [48,80–82] have revealed indel-associated cancer driver genes, diverse indel mutation patterns at MS loci and also diverse trinucleotide SNV mutation signatures resulting from failed MMR (Figure 1B). This genomics work hints that the MMR failures might occur in multiple forms and that the classical MSI-H versus MSS dichotomy might be an oversimplification. Moreover, some data presented in these studies suggest that MMR failures might be more common across tumors than anticipated by the classical, PCR and electrophoresis-based MSI tests based on the Bethesda panel (see below).

Evidence for different subtypes of MMR failures observed in human cancer

As detailed above, genomics analyses of diverse types of mutation patterns can comprehensively identify MMR failures, providing an estimate of their prevalence. We posit that future statistical analyses of genome-wide mutational patterns may discover subtypes of MMR failures, as well as the relevance of such subtypes for cancer evolution and for the clinical course of the disease, including response to chemotherapy and immunotherapy. This reasoning is supported by a considerable amount of pre-genomic studies suggesting that there may exist different subtypes of MMR failures and/or different ways in which MMR failures may become manifest, perhaps depending on the cell type or on the genetic background of each tumor.

The EMAST phenotype

In addition to the commonly recognized MSI-H phenotype, there exists a distinct repeat instability phenomenon termed EMAST, for ‘elevated MS alterations at selected tetranucleotide repeats’. This generates instability mostly at the AAAG or ATAG tetranucleotide repeats; in contrast, MSI-H is detected at mononucleotide and/or dinucleotide repeats. There is evidence to suggest that the two phenotypes affect cancer types differentially. While MSI-H is usually seen in uterus, colorectal, and gastric cancers, EMAST is also common in other cancer types: one-third or more of lung cancers, skin cancers and bladder cancers were reported to show EMAST [86,87]. This suggests that EMAST is a mechanistically distinct phenomenon from MSI. In support of this, ~40% of the MSS colon cancers show some degree of EMAST associated with the loss of the MSH3 expression [88], while MSI tends to be associated with loss of activity (via mutation, damaging germline variant or epigenetic silencing) of other MMR genes, in particular *MLH1*, *MSH6* and *MSH2*. This dichotomy is consistent with the known molecular function of the MSH3/MSH2 dimer (MutS β), which preferentially recognizes DNA loops of 2–13 nucleotides [89], while the competing MSH6/MSH2 dimer (MutS α) recognizes single base mispairs and two-nucleotide bulges. Thus, it is expected that MSH6 efficiently prevents 1–2 nt long but not longer indels, even in absence of MSH3. This is consistent with the absence of classical MSI (largely observed as indels—often 1-nt deletions—at mononucleotide repeats) in many EMAST tumors (which exhibit indels at tetranucleotide repeats, possibly due to loss of MSH3 activity). In a WGS study of genetically engineered human cell lines, the *MSH6*^{-/-} cells indeed generated a large excess of SNVs and 1-nt indels, while the increase in 2+ nt indels was somewhat milder relative to e.g. *MSH2*^{-/-} and *MLH1*^{-/-} cells (the latter are expected to indirectly abrogate also the MSH3 activity) [60]; note that *MSH3* LOF was not explicitly assayed therein, nor the instability of these mutants at particular loci.

The MSI-L phenotype

Further, the MSI-L (low) phenotype readout is defined as 1 (of 5) Bethesda MS locus affected by instability. This was originally intended to signify an intermediate degree of MMR failure between MSS (0 loci) and MSI-H (≥ 2 loci), or alternatively it may represent those MSS cases that randomly accrued an indel in one of the five marker loci. However, it was also proposed that MSI-L may be a distinct type of MMR failure, rather than simply a lower-intensity MSI-H phenotype [90]. In particular MSI-L seems to be more commonly associated with instability at dinucleotide rather than mononucleotide repeat loci, suggesting a different underlying mechanism. MSI-L is observed in melanoma and pancreatic tumors; these tumor types are infrequently MSI-H. MSI-L is associated with gene expression patterns that are distinct from MSI-H, supporting that it is generated by a distinct biological process [91]. Evidence from mouse models implicates loss of activity of the *PMS2* gene homolog in MSI-L [92]. In the Mlh1/Pms2 MMR complex in mouse, the loss of Pms2 functionality can be partially replaced by the Mlh3 subunit [37]. Thus, some activity of the MMR pathway is retained upon Pms2 loss, which appears compatible with the observation of a reduced-intensity MSI-like phenotype. Consistently, a human *PMS2*^{-/-} cell line genome sequence exhibits a different spectrum of indel mutations at MS repeats, with an enrichment of insertions compared with *MLH1*^{-/-} and *MSH2*^{-/-} cell lines (which were similar to each other) [60], although it is not clear whether specific MS loci were differentially affected. Apart from MMR, there is evidence to implicate base excision repair pathway and/or translesion DNA polymerases

as contributing to MSI-L [90], and moreover an overlap was reported between the MSI-L designation of colorectal cancers and the MSH3-deficient EMAST phenotype [88]. Additionally, it was suggested that deficiencies in the POLE protein (replicative DNA polymerase ϵ) can also result in MSI-L [80]. These examples further illustrates how different DNA repair pathway disturbances might converge on to apparently similar mutational phenotypes. Genome-wide studies using statistical methods designed to distinguish MSI-H from MSS tumors usually group MSI-L together with MSS, and not with MSI-H [93], however such analyses do not necessarily focus on stratifying mutagenic processes by repeat unit type and other characteristics. Future genomics analyses may shed light on the question whether MSI-L tumors result from distinct mechanisms of DNA repair failures, rather than simply representing a milder version of the MSI-H phenotype.

Multiple SNV signatures associated with MMR failures

Analysis of trinucleotide mutational signatures suggest that up to seven different trinucleotide mutational signatures are associated with the MSI-H label across cancer types (as mentioned earlier) [51], suggesting multiple underlying mechanisms of MMR failures. Two of seven signatures are due to concurrent mutations in replicative DNA polymerases [94,95], but the mechanisms underlying the other five are unaccounted for. Clearly one possibility is that these result from deficiencies in different core MMR factors: for example the *PMS2*^{-/-} cultured cells exhibited a different mutational spectrum, considerably more rich in A>G/T>C mutations, compared with generally C>T/G>A-rich spectra of *MSH2* or *MSH6* mutant cells [60,96]. Furthermore, alterations in trans-acting factors can impact on MMR efficiency, for example the overexpression of the microRNA, miR-155, in colon cancers [97], deletions in genes *MTOR* (*FRAP1*), *HERC1*, *PRKCZ* and *PIK3C2B* in leukemia [98] and loss of the *SETD2* chromatin modifier [99], which occurs in kidney, lung and colon cancers; it is well conceivable that some of these might result in particular mutation signatures. LOF in the *ARID1A* chromatin remodeler was suggested to result in an MMR-like mutational phenotype [100,101]. A somewhat related explanation underlying this diversity of MMR signatures is that there may exist tissue-specific variation in the ‘expression’ of the same mutational process such as MMR failing [63]. This could happen conceivably due to different activity of accessory or backup DNA repair proteins or relevant metabolites in different tissues, which change the signature of MMR failure sufficiently such that it is detected as discrete mutational signatures. While this tissue-specificity may or may not be pathologic in origin (i.e. it may not represent a failure of DNA repair *per se*), arguably the overall mutagenic mechanism is different and may result in varied implications for cancer evolution and for therapy response. Additionally, some of the MMR signatures might result from interactions of failed MMR with exposures to external DNA-damaging agents [52,102,103], in which case they would not, strictly speaking, be considered subtypes of MMR failures. Finally, it cannot be excluded that the multiplicity of MMR signatures observed could, in part, be due to limitations of the signature inference algorithms that have overestimated their numbers.

The data on tissue specificity, possible underlying genes/mechanisms, and genomic instability signatures of known and anticipated subtypes of MMR failures are summarized in Supplementary Table S1.

Evidence for a higher prevalence for MMR failures across cancer types

Furthermore, recent analyses of genomic data suggest MMR failures are not only of diverse types, but that they might also be more prevalent than currently appreciated. There are various circumstantial evidence anticipating that some degree of MMR dysfunction might be more common in some tumor types than reported by the MSI-H phenotype of the Bethesda MS panel, although currently this remains speculative and should be followed up in future work. In particular:

- Previously reported—not based on genome-wide searches but on assaying various MSI loci individually—frequencies of MSI in ovarian cancer were approximately 10–12% [104,105]. However, a classifier that reproduces the Bethesda MSI-H label (using genomics) classed only 3.2% ovarian tumors as MSI-H [48]. This suggests that the commonly employed Bethesda panel based readout might underestimate MSI prevalence in some cancer types such as ovarian cancer, head-and-neck and cervical cancer [48]. However, MMR failures would be expected to also generate a high burden of SNV mutations in those tumors, which was not always reported.
- Studies that used WES/WGS to examine instability of MS loci showed that some tumors from tissues such as breast [80] or kidney and liver [81] or adrenocortical and cervical cancer and mesothelioma [46], which are rarely classed as MSI-H, do sometimes have substantial numbers of indel events at MSs. This suggests that a milder MSI-like

phenotype might be widespread (or, alternatively, that there exists other factors that result in unstable MSs but are not related with MMR). For adenocortical and cervical cancer, the genomics-based observation of MSI was also supported in the increased mutation burden of SNVs [46].

- Because the distribution of the number of MS indels in tumor genomes is not strictly bimodal, it was suggested that MSI is a continuous phenotype (which also associates with survival outcomes, meaning this continuum is physiologically relevant) [48,80,81], rather than a discrete subdivision of MSI-H *versus* MSS groups. This suggests that some tumors may have sufficiently functional MMR to get an MSS readout on the Bethesda panel (perhaps because a small number [$n=5$] of loci examined) but might still deviate from a baseline MSS state, suggesting a possibility of a subtle MMR dysfunction.
- Trinucleotide SNV signature analysis suggests that the signature resulting from a *PMS2* gene knockout in cell lines resembles the SNV Signature 12 (also abbreviated as SBS12, for ‘single base substitution’) [60]. SBS12 is common in liver hepatocellular tumors and is sometimes seen in other cancer types [51] and, importantly, it was not reported to be associated with MSI-H. Thus, some cancers—hypothetically, resulting from a *PMS2* gene silencing or mutation—may have MMR failures that are not seen as MSI by the Bethesda panel. Another study has reported a mutational signature called RefSig MMR2 [63] with the trinucleotide spectrum resembling SBS12 (and also SBS26, a known MSI-associated signature) and commonly occurring across multiple tissues, particularly ovary, liver, lymphoid, bone and soft tissue, central nervous system and breast.
- A related observation is that of the indel mutational signature ‘ID4’ [51], which consists of short deletions only, however it does not appear associated with the canonical MSI-H phenotypes. This was observed in approximately 10–30% kidney, lymphoma/CLL, bladder and breast tumors (i.e. not the classically MSI-prone tissues). ID4 also occurs in the sometimes MSI-prone tissues of esophagus and ovary and uterus, but is not associated with MSI status thereof. An important caveat about interpreting SNV and indel mutational signatures is that the statistical methodologies to infer the signatures are still maturing, and the signature catalogs might change substantially with newer methods and arrival of more data. Thus results from mutational signature analyses should, currently, not be taken as definitive evidence of prevalence of MMR failures.
- It was proposed that differential activity of MMR is a cause of tumor mutation rate heterogeneity at the megabase scale; thus MSI-H cancers lose such heterogeneity—their genome-wide mutation landscapes become ‘flat’ [76]. However some cancers that are not MSI-H also have ‘flat’ mutation rate landscapes, seen in e.g. kidney or breast or pancreas [unpublished data]. A part of these cases can be explained by a presence of mutational signatures which do not obey the normal mutation rate variability patterns such as APOBEC mutagenesis [69,106], but many of the remaining cases of the ‘redistribution’ in mutation rates [107] remain unexplained. For example this occurs in kidney cancer, where APOBEC mutation patterns are rare, however the *MLH1* gene is sometimes deleted (since it is located near *VHL*, a major tumor suppressor gene). This suggests MMR activity loss as a possible explanation for the unusual distribution of mutation rates across these kidney genomes.
- Overall, genomics has generated various reports hinting at that MMR failures might have diverse causes, which are not fully captured by the unidimensional Bethesda panel MSI-H *versus* MSS dichotomy. These studies moreover suggest that MMR might fail in more tumors than anticipated, however at more moderate levels than the extreme MSI-H, although further supporting evidence is needed to ascertain whether these genomic patterns correspond to *bona fide* MMR deficiencies.

Impact of MMR failures on cancer evolution

MMR status as a determinant of mutation rates across model organisms

The MSI-H phenotype has strong potential to affect gene function as the increased accumulation of indels and SNVs can lead to disruption of gene coding sequences and regulatory sequences. Studies in model organisms have measured these increased mutation rates by comparing strains deficient in core MMR genes to background mutation rates in the *wild-type* controls [108]. Whole-genome mutation rate estimates were measured in yeast strains deficient in *MSH2* [109,110]. These studies yielded a mutation rate of tens ($\sim 32\text{--}71$) of mutations Gd^{-1} (per Gbp and per division), a $\sim 150\times$ increased mutation rate compared with their MMR-proficient counterparts.

Earlier reports [111,112] in an animal model (the worm *Caenorhabditis elegans*), estimated a similar increase in mutation rates (~ 100 -times) in *MutS α* (worm genes *msh-2* and *msh-6*, whose human orthologs are *MSH2* and *MSH6*, respectively)-deficient strains. However, these reports used mutational gene reporter assays, which can give

biased estimates. More recently, WGS of worm strains defective in MutL α (*mlh-1* and *pms-2*) [52,113] have yielded similar ($\sim 71.5 \text{ Gd}^{-1}$) mutation rate estimates [52].

In humans, cultured cells were sequenced after knocking out 43 DNA repair genes and letting the cells undergo tens of divisions. Of those, nine genetic backgrounds reported a significant increase in mutation rates, six of which were MMR genes [60]. This result highlights the importance of MMR in controlling the stability of genomes of unexposed dividing cells. The average mutation rate for the clones deficient in the core MMR genes (*MLH1*, *MSH2*, *MSH6*, *PMS2* and *EXO1*) yielded a value of ~ 93 SNV per division (31 Gd^{-1}) and ~ 69 indel mutations per division (23 Gd^{-1}) [60], which is roughly an order of magnitude increase over the baseline rate in the same cell line with a control gene k.o. Other human cell lines/organoids were examined for k.o. of the *MLH1* [85] and *MSH6* [84] genes, similarly reporting mutation rate increases, and additionally confirming results from genomic analysis suggesting altered trinucleotide spectra [8,114] and genome-wide distributions of mutations [76,115] in MMR-deficient tumors.

Increased cancer incidence due to inheritable MMR deficiencies

Overall, the major increase in mutation rate due to MMR deficiency in the different experimental models is remarkably consistent, highlighting an important role of MMR in maintaining genome integrity, across the tree of life.

Thus, it is not surprising that the germline inactivation of MMR genes causes a higher predisposition and earlier onset for several types of cancer, a consequence of increased mutation rates. Lynch syndrome (also called the hereditary non-polyposis colorectal cancer syndrome, HNPCC) is the most studied case of an MMR deficiency condition, responsible for many hereditary colorectal cancer cases [116,117], as well as for multiple other cancer types, including but not limited to uterus (endometrial), stomach, ovarian and bladder tumors. Although most of the causal variants affect the core MMR genes: *MSH2*, *MLH1*, *MSH6* and *PMS2*, the pathogenic variants in these genes can be detected in only a fraction of familial colon cancer cases [117,118]. A potential explanation is that another condition apart from Lynch syndrome underlies the familial colorectal cancer risk; this includes genes not involved in MMR such as *APC*, or the base excision repair genes *MUTYH* and *NTHL1* or the replicative DNA polymerases *POLE* or *POLD1* [119]. Remarkably, in approximately half of the MMR-deficient colorectal cancers a genetic cause cannot be determined. This ‘Lynch-like syndrome’ also runs in families and has a broadly similar tissue-type spectrum as the standard Lynch syndrome [120–122]. This might perhaps be ascribed to variants can affect accessory genes that modulate MMR activity (this was demonstrated for somatic deletions in several genes in leukemia [98], as well as somatic mutations in some chromatin modifiers; see above). There is a clear need for more studies to identify causes of yet-unexplained part of familial tumor risk, which is likely to be in part due to MMR failures [118].

Although Lynch syndrome has a dominant pattern of inheritance, the single mutated allele is not sufficient to cause the disease. A second-hit somatic alteration is needed to develop cancer, which can occur via loss of heterozygosity (LOH) [123], somatic mutations [124] or by methylation of the promoter for *MLH1* [117,125]. Interestingly, the tissue spectrum of Lynch patients may differ depending on the gene affected, where for instance *MSH6* and *PMS2* carriers were reported to have increased breast cancer risk, unlike *MLH1* and *MSH2* carriers [126,127]. A condition closely related with Lynch syndrome is the Muir–Torre syndrome, where sebaceous skin tumors co-occur with gastrointestinal tumors, and this results from germline variants in *MSH2* or *MLH1* genes [128].

A rare type of a germline MMR syndrome is the constitutional MMR deficiency (CMMRD), which drastically increases the rates of early-onset cancers in the brain and blood [11]. In these cases, the germline inactivating variants are homozygous, thus generating an increased mutation rate much earlier in life. As in Lynch syndrome, the core MMR genes are affected [11,129], however the most commonly associated genes are *PMS2* and *MSH6*, while in the standard Lynch syndrome the *MSH2* and *MLH1* genes predominate [130]. Somatic mutation patterns in CMMRD tend to be different from sporadic MMR-deficient cancers and from Lynch cancers [131]. Some patients develop a characteristic ultramutator phenotype, which results from the combination of the MMR loss and mutations in either *POLE* or *POLD1* replicative DNA polymerases [11]. A closely related condition is Turcot syndrome, where colon polyps occur together with tumors of the central nervous system. A part of the Turcot syndrome cases result from germline variants in *MLH1*, *MSH2* or *PMS2* genes which are homozygous or compound heterozygous [132].

Consequences of MMR failures to the integrity of gene sequences

The most characteristic mutation type in MSI tumor samples is the lengthening or shrinkage of MS repeats. This is caused by the slippage of DNA strands, resulting in a loop, which is, during replication, converted into an indel at the MS locus—a type of mutation that is effectively prevented by active MMR. The usual functional consequence of indels in the protein-coding regions is a frameshift mutation, which commonly results in downstream termination

codons and thus causes the silencing of the gene via nonsense-mediated mRNA decay (NMD) surveillance pathway [133].

Although the majority of MS repeats are located outside coding sequences, there are nonetheless ~380000 candidate repeats in exons in human [48,80]. By analogy to estimating positive selection on driver genes from recurrence of SNVs in exons [6,7], methods exist that estimate the positive selection via recurrence of MS events in gene exons [80]. The genes that appear selected due to the accumulation of MS indel events are in general tumor-suppressor genes, consistent with the tendency of MS mutations to generate LOF events. The most recurrently mutated gene is *ACVR2A* which was previously identified [81,134] as a cancer gene in colorectal and gastric cancer. Other examples of tumor suppressor genes with recurrent MS events are *MSH3* (curiously, an MMR gene itself), *RNF43* (also previously identified in [135]), *JAK1* and *PRDM2* [48,80,81]; see also the cases of *MRE11* and *RAD50* described below.

The MS mutational events can also target non-coding regions, which may be important for gene regulation. One example is the accumulation of MS mutations in UTRs of certain genes such as *FOXP1* and *STC1* [48,136]. A significant accumulation of MS mutations in the 3' UTR was also systematically detected [48] suggesting that a potential mechanism underlying this could be the selection for altered regulation of the mRNA stability of oncogenes [48,137]. Data from MMR-deficient cell lines [136] seem however to suggest a depletion of expression. Thus, the causal role of somatic MS events within the non-coding portions of the genome such as UTRs are still unclear, as is more generally the case for other kinds of somatic mutations in UTRs, although it seems likely that individual examples of cancer genes will be affected [138–140].

One of the important considerations of MSI that became evident only from WGS is that, in addition to overall increased mutation rates, there is a re-distribution of mutation density genome-wide [107]. Normally, somatic mutations accumulate preferentially in the late-replicating, heterochromatic parts of the genome [141,142], which contain fewer genes and fewer other functional elements. In MMR-deficient samples however, mutation densities lose their normal correlation with replication time, wherein the protection for domains with a high density of active genes is removed [76,85,115,143]. One consequence of this is that the genes located in very early-replicating regions will suffer a disproportionately large increase in mutation rates in MSI tumors, while genes in later-replicating regions are expected to have a smaller increase in mutation rates in MSI. This redistribution is expected to affect the repertoire of cancer driver genes mutated in MMR-proficient *versus* MMR-deficient cancers, although this to our knowledge still remains to be studied.

Mutagenic DNA mismatch repair in cancer

Recent genomic studies have raised interesting possibilities that MMR activity (rather than lack thereof) may actually promote accumulation of certain mutation types under certain circumstances. One example of this scenario is the non-pathogenic process of somatic hypermutation (SHM), which diversifies antibody genes during maturation of B lymphocytes. The MMR pathway is recruited to the SHM target regions in its non-canonical form, which recruits the error-prone DNA polymerase η (POLH protein) that increases mutation rates during repair [144]. Genomic evidence suggests that POLH mutational signatures appear abundant in cancer, are associated with activity of normal MMR, and also with exposures to certain carcinogenic agents like UV, oxidative stress, tobacco smoking or alcohol consumption [68,145]. Hypothetically, under these exposures, the association of MMR activity with the early-replicating, gene-dense part of the genome is detrimental, as errors from the error-prone synthesis of POLH accumulate therein [68].

Another example of how MMR changes the distribution of mutations can be found in the activity of the APOBEC enzymes, important mutators in many types of human cancers [146–148]. APOBEC activity is confined to single-stranded DNA, which can commonly occur as an intermediate during repair of double-stranded breaks [147–149]; APOBEC activity thereon results in a characteristic pattern of ‘mutation showers’ (*kataegis*). Moreover, experimental evidence suggested that a potential source of this substrate could be originated by MMR, in particular the excision by the EXO1 nuclease (possibly, by hijacking a repair intermediate during BER) [150,151]. In our recent genomics study [69], we analyzed mutation clustering patterns in human tumors, detecting a prevalent ‘mutation fog’ pattern (*omikli*), which can be associated with MMR activity. This mechanism may explain around two-thirds of all the APOBEC mutations in cancer, with increased accumulation in early replicating, gene-rich regions thus increasing the functional impact of APOBEC mutagenesis [69].

MMR genes as pro-apoptotic tumor suppressors

MMR genes are tumor suppressor genes not only because they control mutation rates, but also because of the role of MMR as a pro-apoptotic pathway under certain DNA damaging conditions. A clear example where this role is relevant is treatment by the DNA alkylating agent and anti-cancer drug TMZ. TMZ is metabolized and reacts with DNA to form O6- and N7-methylguanine (MeG) or N3-methyladenine (MeA) [152,153]. The O6-MeG can be reversed by MGMT, a direct DNA repair enzyme [154], while N7-MeG and N7-MeA are repaired by base-excision repair (reviewed in [153,155]). Importantly, O6-MeG lesion does not block the replicative DNA polymerase during copying, but results in mismatches (the methylated G base-pairs with a T) that can be recognized and processed by the MMR machinery [156]. TMZ is mostly employed in brain tumors with a silenced *MGMT* gene [157,158] but these patients can develop resistance to the treatment. These relapsing brain tumors often carry a hypermutator phenotype and are associated with inactivating mutations in the *MSH6* gene [102,103,159,160]. The mutational signature SBS11 can be extracted from genomes of these hypermutators [8,113] suggesting its association with the TMZ treatment. Further reports in *agt-1* (an *MGMT* homolog) deficient *C. elegans* worms [113] provided an experimental validation for this association.

Overall, the above suggests a genetic interaction of both a first *MGMT* depletion event, combined with a posterior MMR deficiency that promotes the survival of relapsing clones. This phenotype is possibly due to the avoidance of the ‘futile cycling’ of MMR [157], which occurs when MMR repetitively tries to but fails to correct a mismatch (here, involving a methylated nucleotide), signaling apoptosis.

MMR failures generate genetic events that sensitize to therapeutics

The mutations resulting from MMR failures also can provide markers of resistance or sensitivity to chemotherapeutic drugs (Figure 2). This can occur in two ways: either the process itself can be targeted, or its downstream consequences can be targeted. Firstly, the failed MMR pathway itself may present a vulnerability as it increases reliance on auxiliary pathways for DNA replication or repair of certain lesions. Secondly, the mutations that accumulate at increased rates after failed MMR can disable specific genes—most notably those genes bearing MS repeats, which are hotspots for mutagenesis—therefore rendering the tumor cells vulnerable to interventions.

Several studies have found an association between the MSI background of a tumor and sensitivity towards therapies using irinotecan [161–164]. The proposed mechanism is that the failure of MMR is not in itself sensitizing the cancer cells to irinotecan, but that MSI generates mutations in other DNA repair genes that would otherwise be able to fend off the damage caused by the drug. Irinotecan binds to topoisomerase I, inhibiting the religation of DNA; this can result in dsDNA breaks upon conflicts with the replication machinery. Once the breaks occur, *MRE11* together with *RAD50* and *NBS1* can repair them, avoiding apoptosis. Both *MRE11* and *RAD50* genes are known to be commonly mutated in MSI cancers [164–166] due to the MSs within their coding sequence (for *RAD50*) [165] or intronic sequence (for *MRE11*; despite being intronic the indels were reported to be deleterious to gene function) [166,167]. Further studies using cancer cell lines with mutated *MRE11* and *RAD50* have also reported increased sensitivity to irinotecan [164]. Considered together with experiments using isogenic cell lines differing only in *MRE11* status, these indicate that MMR activity itself is not directly causal to the sensitivity to irinotecan [163,168]. Notwithstanding, with approximately three-quarters of MSI cancers bearing mutations in *MRE11* [166,169], MSI tumor status could serve as a proxy biomarker for the effectiveness of an irinotecan treatment.

In a similar fashion, PARP-1 inhibitors have been tested against MSI cell lines and the success of the treatment was negatively correlated to the expression levels of *MRE11* or the activity of the interacting protein *NBS1* [170–172], further demonstrating roles in resolving double-strand damage and suggesting MSI as a possible marker for PARPi usage. However, the PARPi olaparib as a single-agent treatment did not benefit colorectal cancer patients, neither to the MSI nor to MSS group [173]. This outcome seems consistent with an application of the PARPi niraparib to a panel of colorectal cancer cell lines, where only a modest, non-significant MSI-associated trend was observed [174].

In addition to *MRE11* and *RAD50* genes, also the *MBD4* gene of the base excision repair pathway has been proposed to be affected by coding region indels in some MSI cancers, resulting in a truncated form of *MBD4* protein that may sensitize to 5-fluorouracil (5-FU) and to trifluridine [175,176].

Failed MMR was additionally proposed to result in (potentially exploitable) vulnerabilities to cancer cells not only by LOF mutations, as described above, but also by gain-of-function mutations. In particular, indel mutations can generate frameshifted and/or truncated, potentially toxic polypeptides. One known example of this is the *HSP110* chaperone protein. Mutations in an intronic microsatellite repeat of the *HSP110* gene, commonly occurring in MSI cells, cause skipping of exon 9 and generate a dominant-negative acting truncated protein [177]. Colorectal cancers

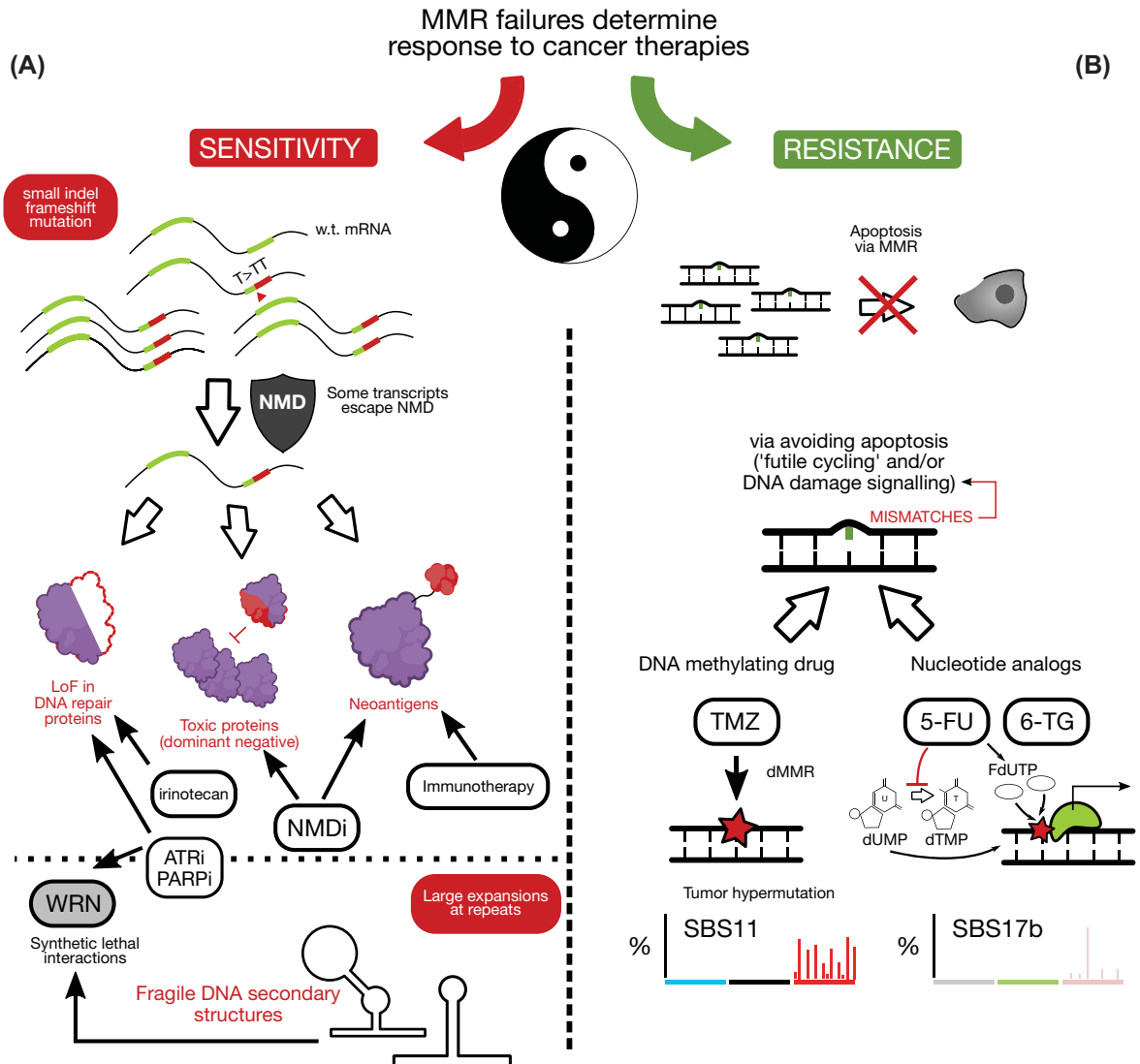


Figure 2. MMR failures in cancers can cause resistance or sensitization to certain therapies

(A) MMR dysfunction may sensitize tumors to various therapies via different mechanisms; some of these are well established and applied in the clinic (immunotherapy), while others are proposed based on known molecular mechanisms and/or studies on experimental models, however awaiting clinical validation (in panel A). A common consequence of MMR failures are indel mutations at MS repeats, which can be in coding regions and so generate frameshifted, and usually truncated, mutant proteins (note that some indels at repeats also in introns can induce missplicing, consequently again resulting in frameshifting and truncation of coding regions). Expression of such truncated proteins is often silenced by the NMD surveillance pathway that degrades mutant mRNA. However some frameshifted mRNAs can escape NMD detection and generate neoantigens that sensitize to immunotherapy, or that generate toxic proteins (e.g. HSP110 mutant), suggesting that beneficial effects of therapy might be potentiated by NMD inhibition. Frameshifting indels also can cause LOF in DNA repair proteins e.g. MRE11 and RAD50, sensitizing to irinotecan and potentially to DNA damage-signaling drugs e.g. ATR inhibitors. Finally, expansions at repeats due to MMR failure increase reliance on the WRN DNA helicase. (B) MMR inactivation causes resistance to some commonly used cancer drugs, which can trigger MMR activity—either by methylation of nucleobases [TMZ, dacarbazine], or by incorporation of chemically modified nucleotides [6-TG, 5-FU] that can register as mismatches—potentially resulting in apoptosis via DNA damage signalling. A dysfunctional MMR results in reduced clinical utility of such drugs, as well as in increased mutation rates in the MMR-deficient tumor upon treatment, which can generate further mutations in cancer driver genes or in drug resistance genes. Abbreviations: 5-FU, 5-fluorouracil; 6-TG, 6-thioguanine.

bearing the mutations may have better survival and better response to chemotherapy [178]. Pharmacological inhibition of the NMD pathway, which can silence nonsense-mutation bearing mRNAs, further promoted expression of these HSP110 mutants and possibly other similarly affected proteins, selectively poisoning MSI cells and tumors [179].

Reduced MMR function confers resistance to some chemotherapies

In contrast with the above mechanism, where mutations occurring after MMR failure inactivate various genes and thereby sensitize to interventions, MMR activity itself can result in sensitivity or resistance to therapeutics (Figure 2). The TMZ example was already covered above; other drugs can also have reduced benefits when given to patients suffering from an MSI cancer. One example is 5-FU, an antimetabolite commonly administered to breast and colon cancer patients, among others [180]. The introduction of this drug into regimens to treat MSS colorectal cancers drastically improves the outcome, but not so for MMR-deficient tumors, where 5-FU does not provide a benefit [181–184]. 5-FU depletes thymidine triphosphate by inhibiting the thymidylate synthase enzyme; due to the nucleotide pool imbalance, 5-FU (in the form of FdUTP) and uracil may be incorporated during DNA synthesis instead of thymine [180,185]. This nucleotide misincorporation, upon repair by MMR, can lead to DNA breaks accumulated during S phase that can activate the apoptotic program [185]. Additionally, the MMR machinery can recognize and bind to the incorporated 5-FU and then trigger cell death [186,187]. Due to this evidence, administration of 5-FU is currently not advised in MSI colorectal cancers [188].

A similar mechanism as for 5-FU also underlies treatments using 6-thioguanine (6-TG), a therapeutic for various types of leukemia, which is also used as an immunosuppressant. MMR recognizes the mismatch when the DNA antimetabolite is incorporated into the DNA strands [189,190] and fails to repair the damage, potentially leading to apoptosis [189]. Thus, cells that are MMR-deficient are more resistant to the cytotoxic effects of 6-TG [191,192]. Additionally, the tumor likely obtains an evolutionary benefit from the increased mutation rates caused by the unrepaired 6-TG mismatches, which can produce further cancer driver mutations [192]. One example is a *TP53* gene disruption that makes the cancer resistant to many treatments; other mutagenic hotspots related with 6-TG treatment were reported in drug resistance genes [192]. In addition to 6-TG, cytotoxic activity of other thiopurine drugs (6-mercaptopurine, azathioprine) is likely at least in part mediated via MMR acting on mismatched bases resulting from the treatment [193].

Implications of MMR failures in tumors for immunotherapy

Finally, occurrence of a high burden of somatic mutations (that can stem from MMR failures) predicts response to immunotherapy. Thus, ascertaining the extent and characteristics of MMR failures in tumors has important implications for stratifying patients for immunotherapy.

The introduction of immune checkpoint inhibitors (also called checkpoint blockade or simply referred to by the more general term immunotherapy) has had a dramatic impact on overall patient survival in many types of cancers such as melanoma or lung adenocarcinoma. However, the response to these therapies tends to be quite variable across patients, which prompted a search for predictive biomarkers of efficacy. Currently, the tumor mutational burden (TMB) is considered to be a useful predictor of efficacy for these immune checkpoint inhibitors [194–197]. This is consistent with known mechanisms by which the immune system can recognize mutated proteins and, in response, activate the aberrant cell's programmed death [198–201]; thus, cancerous cells holding a greater mutational load should be more readily recognized and cleared. Consistently, cancer types with commonly higher TMBs—such as melanomas [202,203] and smokers' lung cancers [204], where most of the mutations result from the exposure to external mutagens—respond better to immune checkpoint inhibitors. Another important cause of the cancer's somatic mutational burden can be MMR failures, which can be estimated by exome or genome sequencing, or indirectly via an MSI assay. An MSI cancer will typically acquire many times more mutations than their MSS counterparts [205,206], and will thus likely express more neoantigens [206] recognizable by the immune cells. The predictive power of MSI-status for immune checkpoint inhibitor patient-response has made MSI the first FDA-approved genetic biomarker valid across different cancer types [207].

It is likely that MMR failures sensitize to immunotherapy mainly through generating a much increased mutation burden. Consistently, a statistical analysis of multiple determinants of immunotherapy response did not find an independent association with MSI status, after controlling for mutation burden [208]. Similarly, other causes of high TMB unrelated with MMR similarly predict a favourable immunotherapy response, including UV-exposed skin cancer, and tobacco-smoke exposed lung cancer cells. Additionally, kidney cancers often respond even though they do not commonly exhibit MSI (at least not in the classical sense, see above), which was ascribed to their relatively high burden of

indel mutations [209]. Recent studies by us and others [133,210] have highlighted specifically the frameshifting indel mutations—which are abundantly occurring upon MMR dysfunction—rather than single-nucleotide substitutions, as more robust predictors of immunotherapy response (Figure 2). These genomic analyses further highlighted the role of the NMD pathway in silencing mRNAs containing frameshifting indel mutations in human cancers [133,210], in line with prior reports showing that inhibition of NMD boosts antitumor immunity in mouse models [211].

MMR failures result in synthetic lethalties with other DNA repair mechanisms

As noted above, MMR-deficient tumours can respond to drugs differently than MMR-proficient ones. They resist specific chemotherapeutics (such as TMZ, 5-FU and 6-TG), but can also be sensitized to other therapeutics via mutating genes participating in various DNA repair pathways, or via giving rise to neoantigens targetable by immunotherapy. Additionally, MMR genes may, similarly as many other tumor suppressor genes, participate in synthetic lethal interactions allowing another window of opportunity to target MMR-deficient cells. A recently discovered synthetic lethal interaction with the Werner syndrome helicase (WRN) [212–215] has revealed that a mutational process caused by the MMR deficiency, over time, can also lead to vulnerabilities that can be exploited [216]. In this case, MSI cells develop a dependency on WRN due to large expansions at A:T repeats and thus require the activity of WRN for normal DNA replication; without it the cells suffer large-scale chromosome breakage [216]. Currently, there is no approved WRN inhibitor drug, but the proposed mechanism suggests that ATR inhibitors might show some selectivity towards MSI cells, presumably by preventing to recruit WRN by ATR [216].

Concluding remarks and outlook

DNA MMR is a key pathway for ensuring genome stability in all eukaryotic cells. Loss of MMR activity causes vastly increased mutation rates and, consequently, an increase in cancer risk. MMR-deficient tumors are commonly detected by immunohistochemistry to measure MMR protein levels, or by PCR and electrophoresis-based assays that detect instability at microsatellite repeats (MSI), which become hypermutable when MMR function is compromised. Importantly, loss of MMR activity is clinically relevant as it contributes to resistance or sensitivity to various cancer therapies such as certain nucleobase analogs, DNA alkylating agents, topoisomerase inhibitors and immune checkpoint inhibitors. Other experimental therapeutic strategies such as NMD inhibition to express toxic or immunogenic mutated proteins might be used to target MMR-deficient cancer cells. Recent developments in genome sequencing and bioinformatics tools facilitate routine detection of MSI status, however genomes can also provide information in addition to the binary MSI-H *versus* MSS classification. Interestingly, bioinformatics analyses of mutational signatures suggest that MMR deficiencies may be multifaceted, consistent with LOF of individual MMR genes resulting in partial failures of the pathway. Genomics also provides evidence that a certain level of MMR malfunction may be more widespread among cancer types than anticipated previously. We envision that studies of genomes of MMR-deficient tumors, as well as experimental models (e.g. engineered cell lines) will further elucidate subtypes of MMR failures and their prevalence in tumors by analysis of various mutational patterns. We anticipate that such mutational patterns will provide refined predictive markers for applying chemo-, radio- and immunotherapies to MMR-deficient tumors.

Data Availability

The statement is not applicable: there are no primary data in this review article.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work in the F.S.'s lab was supported by the ERC Starting Grant 'HYPER-INSIGHT' [grant number 757700]; the Horizon2020 RIA project 'DECIDER' [grant number 965193]; the Spanish Ministry of Science, Education and Universities project 'REPAIRSCAPE' [grant number PID2020-118795GB-I00]; the State Agency for Research of the Ministry of Science and Innovation - Severo Ochoa Centre of Excellence Award [grant number CEX2019-000913-S]; and the SO-FPI Fellowship [grant number BES-2017-079820 (to D.M.-P.)].

Acknowledgements

We apologize to colleagues whose work was not referenced due to article length constraints.

Abbreviations

CMMRD, constitutional MMR deficiency; CNA, copy number alteration; EMAST, elevated microsatellite alterations at selected tetranucleotide repeats; HRR, homologous recombination repair; LOF, loss-of-function; MeA, N3-methyladenine; MeG, O6- and N7-methylguanine; MMR, mismatch repair; MS, microsatellite; MSI, microsatellite instability; NMD, nonsense-mediated mRNA decay; SHM, somatic hypermutation; SNV, single-nucleotide variant; TMB, tumor mutational burden; UTR, untranslated region; WES, whole-exome sequence; WGS, whole-genome sequence; 5-FU, 5-fluorouracil; 6-TG, 6-thioguanine.

References

- 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.e18–385.e18, <https://doi.org/10.1016/j.cell.2018.02.060>
- 2 Ellrott, K., Bailey, M.H., Saksena, G., Covington, K.R., Kandoth, C., Stewart, C. et al. (2018) Scalable open science approach for mutation calling of tumor exomes using multiple genomic pipelines. *Cell Syst.* **6**, 271.e7–281.e7, <https://doi.org/10.1016/j.cels.2018.03.002>
- 3 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, <https://doi.org/10.1038/s41586-020-1969-6>
- 4 Priestley, P., Baber, J., Lolkema, M.P., Steeghs, N., de Bruijn, E., Shale, C. et al. (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **575**, 210–216, <https://doi.org/10.1038/s41586-019-1689-y>
- 5 Dietlein, F., Weghorn, D., Taylor-Weiner, A., Richters, A., Reardon, B., Liu, D. et al. (2020) Identification of cancer driver genes based on nucleotide context. *Nat. Genet.* **52**, 208–218, <https://doi.org/10.1038/s41588-019-0572-y>
- 6 Martincorena, I., Raine, K.M., Gerstung, M., Dawson, K.J., Haase, K., Loo, P.V. et al. (2017) Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041, e21., <https://doi.org/10.1016/j.cell.2017.09.042>
- 7 Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A. et al. (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218, <https://doi.org/10.1038/nature12213>
- 8 Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V. et al. (2013) Signatures of mutational processes in human cancer. *Nature* **500**, 415–421, <https://doi.org/10.1038/nature12477>
- 9 de Bruin, E.C., McGranahan, N., Mitter, R., Salm, M., Wedge, D.C., Yates, L. et al. (2014) Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* **346**, 251–256, <https://doi.org/10.1126/science.1253462>
- 10 Chan, K., Roberts, S.A., Klimczak, L.J., Sterling, J.F., Saini, N., Malc, E.P. et al. (2015) An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. *Nat. Genet.* **47**, ng.3378, <https://doi.org/10.1038/ng.3378>
- 11 Shlien, A., Campbell, B.B., de Borja, R., Alexandrov, L.B., Merico, D., Wedge, D. et al. (2015) Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nat. Genet.* **47**, 257–262, <https://doi.org/10.1038/ng.3202>
- 12 Johnson, B.E., Mazar, T., Hong, C., Barnes, M., Aihara, K., McLean, C.Y. et al. (2014) Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* **343**, 189–193, <https://doi.org/10.1126/science.1239947>
- 13 Hsieh, P. and Yamane, K. (2008) DNA mismatch repair: molecular mechanism, cancer, and ageing. *Mech. Ageing Dev.* **129**, 391–407, <https://doi.org/10.1016/j.mad.2008.02.012>
- 14 Iyer, R.R., Pluciennik, A., Burdett, V. and Modrich, P.L. (2006) DNA mismatch repair: functions and mechanisms. *Chem. Rev.* **106**, 302–323, <https://doi.org/10.1021/cr0404794>
- 15 Kunkel, T.A. and Erie, D.A. (2005) DNA mismatch repair. *Annu. Rev. Biochem.* **74**, 681–710, <https://doi.org/10.1146/annurev.biochem.74.082803.133243>
- 16 Chen, W. and Jinks-Robertson, S. (1998) Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. *Mol. Cell Biol.* **18**, 6525–6537, <https://doi.org/10.1128/MCB.18.11.6525>
- 17 Yoshioka, K., Yoshioka, Y. and Hsieh, P. (2006) ATR kinase activation mediated by MutS α and MutL α in response to cytotoxic O6-methylguanine adducts. *Mol. Cell* **22**, 501–510
- 18 Kim, W.-J., Rajasekaran, B. and Brown, K.D. (2007) MLH1- and ATM-dependent MAPK signaling is activated through c-Abl in response to the alkylator N-methyl-N'-nitro-N'-nitrosoguanidine*. *J. Biol. Chem.* **282**, 32021–32031, <https://doi.org/10.1074/jbc.M701451200>
- 19 Li, G.-M. (2008) Mechanisms and functions of DNA mismatch repair. *Cell Res.* **18**, 85–98, <https://doi.org/10.1038/cr.2007.115>
- 20 Kolodner, R.D., Tytell, J.D., Schmeits, J.L., Kane, M.F., Gupta, R.D., Weger, J. et al. (1999) Germ-line msh6 mutations in colorectal cancer families. *Cancer Res.* **59**, 5068–5074
- 21 Modrich, P. and Lahue, R. (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **65**, 101–133, <https://doi.org/10.1146/annurev.bi.65.070196.000533>
- 22 Jiricny, J. (2006) The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.* **7**, 335–346, <https://doi.org/10.1038/nrm1907>
- 23 Marsischky, G.T., Filosi, N., Kane, M.F. and Kolodner, R. (1996) Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes Dev.* **10**, 407–420, <https://doi.org/10.1101/gad.10.4.407>
- 24 Genschel, J., Littman, S.J., Drummond, J.T. and Modrich, P. (1998) Isolation of MutS β from human cells and comparison of the mismatch repair specificities of MutS β and MutS α *. *J. Biol. Chem.* **273**, 19895–19901, <https://doi.org/10.1074/jbc.273.31.19895>
- 25 Umar, A., Risinger, J.I., Glaab, W.E., Tindall, K.R., Barrett, J.C. and Kunkel, T.A. (1998) Functional overlap in mismatch repair by human MSH3 and MSH6. *Genetics* **148**, 1637–1646, <https://doi.org/10.1093/genetics/148.4.1637>
- 26 Prolla, T.A., Pang, Q., Alani, E., Kolodner, R.D. and Liskay, R.M. (1994) MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science* **256**, 1091–1093, <https://doi.org/10.1126/science.8066446>

- 27 Li, G.M. and Modrich, P. (1995) Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1950–1954, <https://doi.org/10.1073/pnas.92.6.1950>
- 28 Wang, T.-F., Kleckner, N. and Hunter, N. (1999) Functional specificity of MutL homologs in yeast: Evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13914–13919, <https://doi.org/10.1073/pnas.96.24.13914>
- 29 Wei, K., Clark, A.B., Wong, E., Kane, M.F., Mazur, D.J., Parris, T. et al. (2003) Inactivation of Exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. *Genes Dev.* **17**, 603–614, <https://doi.org/10.1101/gad.1060603>
- 30 Kastrinos, F., Stoffel, E.M., Balmaña, J., Steyerberg, E.W., Mercado, R. and Syngal, S. (2008) Phenotype comparison of MLH1 and MSH2 mutation carriers in a cohort of 1,914 individuals undergoing clinical genetic testing in the United States. *Cancer Epidemiol. Prev. Biomark.* **17**, 2044–2051, <https://doi.org/10.1158/1055-9965.EPI-08-0301>
- 31 Hegan, D.C., Narayanan, L., Jirik, F.R., Edelmann, W., Liskay, R.M. and Glazer, P.M. (2006) Differing patterns of genetic instability in mice deficient in the mismatch repair genes Pms2, Mlh1, Msh2, Msh3 and Msh6. *Carcinogenesis* **27**, 2402–2408, <https://doi.org/10.1093/carcin/bgl079>
- 32 Lynch, H.T. and de la Chapelle, A. (2003) Hereditary colorectal cancer. *N. Engl. J. Med.* **348**, 919–932, <https://doi.org/10.1056/NEJMra012242>
- 33 Edelmann, W., Yang, K., Umar, A., Heyer, J., Lau, K., Fan, K. et al. (1997) Mutation in the mismatch repair gene msh6 causes cancer susceptibility. *Cell* **91**, 467–477, [https://doi.org/10.1016/S0092-8674\(00\)80433-X](https://doi.org/10.1016/S0092-8674(00)80433-X)
- 34 Miao, H.-K., Chen, L.-P., Cai, D.-P., Kong, W.-J., Xiao, L. and Lin, J. (2015) MSH3 rs26279 polymorphism increases cancer risk: a meta-analysis. *Int. J. Clin. Exp. Pathol.* **8**, 11060–11067
- 35 Adam, R., Spier, I., Zhao, B., Kloth, M., Marquez, J., Hinrichsen, I. et al. (2016) Exome sequencing identifies biallelic MSH3 germline mutations as a recessive subtype of colorectal adenomatous polyposis. *Am. J. Hum. Genet.* **99**, 337–351, <https://doi.org/10.1016/j.ajhg.2016.06.015>
- 36 Prolla, T.A., Baker, S.M., Harris, A.C., Tsao, J.-L., Yao, X., Bronner, C.E. et al. (1998) Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nat. Genet.* **18**, 276–279, <https://doi.org/10.1038/ng0398-276>
- 37 Chen, P.-C., Dudley, S., Hagen, W., Dizon, D., Paxton, L., Reichow, D. et al. (2005) Contributions by MutL homologues Mlh3 and Pms2 to DNA mismatch repair and tumor suppression in the mouse. *Cancer Res.* **65**, 8662–8670, <https://doi.org/10.1158/0008-5472.CAN-05-0742>
- 38 Boland, C.R., Thibodeau, S.N., Hamilton, S.R., Sidransky, D., Eshleman, J.R., Burt, R.W. et al. (1998) A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* **58**, 5248–5257
- 39 Murphy, K.M., Zhang, S., Geiger, T., Hafez, M.J., Bacher, J., Berg, K.D. et al. (2006) Comparison of the Microsatellite Instability Analysis System and the Bethesda Panel for the determination of microsatellite instability in colorectal cancers. *J. Mol. Diagn.* **8**, 305–311, <https://doi.org/10.2353/jmoldx.2006.050092>
- 40 Goel, A., Nagasaka, T., Hamelin, R. and Boland, C.R. (2010) An optimized pentaplex PCR for detecting DNA mismatch repair-deficient colorectal cancers. *PLoS ONE* **5**, e9393, <https://doi.org/10.1371/journal.pone.0009393>
- 41 You, J.-F., Buhard, O., Ligtenberg, M.J.L., Kets, C.M., Niessen, R.C., Hofstra, R.M.W. et al. (2010) Tumours with loss of MSH6 expression are MSI-H when screened with a pentaplex of five mononucleotide repeats. *Br. J. Cancer* **103**, 1840–1845, <https://doi.org/10.1038/sj.bjc.6605988>
- 42 Baudrin, L.G., Deleuze, J.-F. and How-Kit, A. (2018) Molecular and computational methods for the detection of microsatellite instability in cancer. *Front. Oncol.* **8**, 621, <https://doi.org/10.3389/fonc.2018.00621>
- 43 Kautto, E.A., Bonneville, R., Miya, J., Yu, L., Krook, M.A., Reeser, J.W. et al. (2016) Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS. *Oncotarget* **8**, 7452–7463, <https://doi.org/10.18632/oncotarget.13918>
- 44 Niu, B., Ye, K., Zhang, Q., Lu, C., Xie, M., McLellan, M.D. et al. (2014) MSIsensor: microsatellite instability detection using paired tumor-normal sequence data. *Bioinformatics* **30**, 1015–1016, <https://doi.org/10.1093/bioinformatics/btt755>
- 45 Salipante, S.J., Scroggins, S.M., Hampel, H.L., Turner, E.H. and Pritchard, C.C. (2014) Microsatellite instability detection by next generation sequencing. *Clin. Chem.* **60**, 1192–1199, <https://doi.org/10.1373/clinchem.2014.223677>
- 46 Bonneville, R., Krook, M.A., Kautto, E.A., Miya, J., Wing, M.R., Chen, H.-Z. et al. (2017) Landscape of microsatellite instability across 39 cancer types. *JCO Precis. Oncol.* **1**, 1–15, <https://doi.org/10.1200/PO.17.00073>
- 47 Huang, M.N., McPherson, J.R., Cutcutache, I., Teh, B.T., Tan, P. and Rozen, S.G. (2015) MsiSeq: software for assessing microsatellite instability from catalogs of somatic mutations. *Sci. Rep.* **5**, 1–10, <https://doi.org/10.1038/srep13321>
- 48 Cortes-Ciriano, I., Lee, S., Park, W.-Y., Kim, T.-M. and Park, P.J. (2017) A molecular portrait of microsatellite instability across multiple cancers. *Nat. Commun.* **8**, 15180, <https://doi.org/10.1038/ncomms15180>
- 49 Funnell, T., Zhang, A.W., Grewal, D., McKinney, S., Bashashati, A., Wang, Y.K. et al. (2019) Integrated structural variation and point mutation signatures in cancer genomes using correlated topic models. *PLoS Comput. Biol.* **15**, e1006799, <https://doi.org/10.1371/journal.pcbi.1006799>
- 50 Shiraishi, Y., Tremmel, G., Miyano, S. and Stephens, M. (2015) A simple model-based approach to inferring and visualizing cancer mutation signatures. *PLoS Genet.* **11**, e1005657, <https://doi.org/10.1371/journal.pgen.1005657>
- 51 Alexandrov, L.B., Kim, J., Haradhvala, N.J., Huang, M.N., Tian Ng, A.W. et al. (2020) The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101, <https://doi.org/10.1038/s41586-020-1943-3>
- 52 Meier, B., Volkova, N.V., Hong, Y., Schofield, P., Campbell, P.J., Gerstung, M. et al. (2018) Mutational signatures of DNA mismatch repair deficiency in *C. elegans* and human cancers. *Genome Res.* **28**, 666–675, <https://doi.org/10.1101/gr.226845.117>
- 53 Nik-Zainal, S., Davies, H., Staaf, J., Ramakrishna, M., Glodzik, D., Zou, X. et al. (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54, <https://doi.org/10.1038/nature17676>
- 54 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* **578**, 112–121, <https://doi.org/10.1038/s41586-019-1913-9>

- 55 Nguyen, L., Martens, J.W.M., Van Hoeck, A. and Cuppen, E. (2020) Pan-cancer landscape of homologous recombination deficiency. *Nat. Commun.* **11**, 5584, <https://doi.org/10.1038/s41467-020-19406-4>
- 56 Macintyre, G., Goranova, T.E., De Silva, D., Ennis, D., Piskorz, A.M., Eldridge, M. et al. (2018) Copy number signatures and mutational processes in ovarian carcinoma. *Nat. Genet.* **50**, 1262–1270, <https://doi.org/10.1038/s41588-018-0179-8>
- 57 Kucab, J.E., Zou, X., Morganello, S., Joel, M., Nanda, A.S., Nagy, E. et al. (2019) A compendium of mutational signatures of environmental agents. *Cell* **177**, 821.16–836.16, <https://doi.org/10.1016/j.cell.2019.03.001>
- 58 Szikriszt, B., Póti, Á., Pipek, O., Krzystanek, M., Kanu, N., Molnár, J. et al. (2016) A comprehensive survey of the mutagenic impact of common cancer cytotoxics. *Genome Biol.* **17**, 99, <https://doi.org/10.1186/s13059-016-0963-7>
- 59 Olivier, M., Weninger, A., Ardin, M., Huskova, H., Castells, X., Vallée, M.P. et al. (2014) Modelling mutational landscapes of human cancers in vitro. *Sci. Rep.* **4**, 4482, <https://doi.org/10.1038/srep04482>
- 60 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, <https://doi.org/10.1038/s43018-021-00200-0>
- 61 Zámboreszky, J., Szikriszt, B., Gervai, J.Z., Pipek, O., Póti, Á., Krzystanek, M. et al. (2017) Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions. *Oncogene* **36**, 746–755, <https://doi.org/10.1038/ncr.2016.243>
- 62 Islam, S.M.A., Wu, Y., Díaz-Gay, M., Bergstrom, E.N., He, Y., Barnes, M. et al. (2021) Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor. *bioRxiv* 2020.12.13.422570, <https://doi.org/10.1101/2020.12.13.422570>
- 63 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, <https://doi.org/10.1038/s43018-020-0027-5>
- 64 Sason, I., Chen, Y., Leiserson, M.D.M. and Sharan, R. (2021) A mixture model for signature discovery from sparse mutation data. *Genome Med.* **13**, 173, <https://doi.org/10.1186/s13073-021-00988-7>
- 65 Lal, A., Liu, K., Tibshirani, R., Sidow, A. and Ramazzotti, D. (2021) De novo mutational signature discovery in tumor genomes using SparseSignatures. *PLoS Comput. Biol.* **17**, e1009119, <https://doi.org/10.1371/journal.pcbi.1009119>
- 66 Gilad, G., Leiserson, M.D.M. and Sharan, R. (2021) A data-driven approach for constructing mutation categories for mutational signature analysis. *PLoS Comput. Biol.* **17**, e1009542, <https://doi.org/10.1371/journal.pcbi.1009542>
- 67 Wojtowicz, D., Hoinka, J., Amgalan, B., Kim, Y.-A. and Przytycka, T.M. (2021) RepairSig: deconvolution of DNA damage and repair contributions to the mutational landscape of cancer. *Cell Syst.* **12**, 994.e4–1003.e4, <https://doi.org/10.1016/j.cels.2021.07.004>
- 68 Supek, F. and Lehner, B. (2017) Clustered mutation signatures reveal that error-prone DNA repair targets mutations to active genes. *Cell* **170**, 534.e23–547.e23, <https://doi.org/10.1016/j.cell.2017.07.003>
- 69 Mas-Ponte, D. and Supek, F. (2020) DNA mismatch repair promotes APOBEC3-mediated diffuse hypermutation in human cancers. *Nat. Genet.* **52**, 958–968, <https://doi.org/10.1038/s41588-020-0674-6>
- 70 Glodzik, D., Morganello, S., Davies, H., Simpson, P.T., Li, Y., Zou, X. et al. (2017) A somatic-mutational process recurrently duplicates germline susceptibility loci and tissue-specific super-enhancers in breast cancers. *Nat. Genet.* **49**, 341–348, <https://doi.org/10.1038/ng.3771>
- 71 Wang, Y.K., Bashashati, A., Anglesio, M.S., Cochrane, D.R., Grewal, D.S., Ha, G. et al. (2017) Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat. Genet.* **49**, 856–865, <https://doi.org/10.1038/ng.3849>
- 72 Davies, H., Glodzik, D., Morganello, 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, <https://doi.org/10.1038/nm.4292>
- 73 Secier, M., Li, X., de Silva, N., Eldridge, M.D., Contino, G., Borschein, J. et al. (2016) Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. *Nat. Genet.* **48**, 1131–1141, <https://doi.org/10.1038/ng.3659>
- 74 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, <https://doi.org/10.1016/j.cell.2015.12.050>
- 75 Tomkova, M., Tomek, J., Kriaiucionis, S. and Schuster-Böckler, B. (2018) Mutational signature distribution varies with DNA replication timing and strand asymmetry. *Genome Biol.* **19**, 129, <https://doi.org/10.1186/s13059-018-1509-y>
- 76 Supek, F. and Lehner, B. (2015) Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature* **521**, 81–84, <https://doi.org/10.1038/nature14173>
- 77 Zheng, C.L., Wang, N.J., Chung, J., Moslehi, H., Sanborn, J.Z., Hur, J.S. et al. (2014) Transcription restores DNA repair to heterochromatin, determining regional mutation rates in cancer genomes. *Cell Rep.* **9**, 1228–1234, <https://doi.org/10.1016/j.celrep.2014.10.031>
- 78 Polak, P., Lawrence, M.S., Haugen, E., Stoletzki, N., Stojanov, P., Thurman, R.E. et al. (2014) Reduced local mutation density in regulatory DNA of cancer genomes is linked to DNA repair. *Nat. Biotechnol.* **32**, 71–75, <https://doi.org/10.1038/nbt.2778>
- 79 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, <https://doi.org/10.1038/s41467-021-23551-9>
- 80 Maruvka, Y.E., Mouw, K.W., Karlic, R., Parasuraman, P., Kamburov, A., Polak, P. et al. (2017) Analysis of somatic microsatellite indels identifies driver events in human tumors. *Nat. Biotechnol.* **35**, 951–959, <https://doi.org/10.1038/nbt.3966>
- 81 Hause, R.J., Pritchard, C.C., Shendure, J. and Salipante, S.J. (2016) Classification and characterization of microsatellite instability across 18 cancer types. *Nat. Med.* **22**, 1342–1350, <https://doi.org/10.1038/nm.4191>
- 82 Fujimoto, A., Fujita, M., Hasegawa, T., Wong, J.H., Maejima, K., Oku-Sasaki, A. et al. (2020) Comprehensive analysis of indels in whole-genome microsatellite regions and microsatellite instability across 21 cancer types. *Genome Res.* **30**, 334–346, <https://doi.org/10.1101/gr.255026.119>
- 83 Polak, P., Karlič, R., Koren, A., Thurman, R., Sandstrom, R., Lawrence, M.S. et al. (2015) Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **518**, 360–364, <https://doi.org/10.1038/nature14221>

- 84 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, <https://doi.org/10.1038/s41467-018-04052-8>
- 85 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, <https://doi.org/10.1126/science.aao3130>
- 86 Ahrendt, S.A., Decker, P.A., Doffek, K., Wang, B., Xu, L., Demeure, M.J. et al. (2000) Microsatellite instability at selected tetranucleotide repeats is associated with p53 mutations in non-small cell lung cancer. *Cancer Res.* **60**, 2488–2491
- 87 Danaee, H., Nelson, H.H., Karagas, M.R., Schned, A.R., Ashok, T.D.S., Hirao, T. et al. (2002) Microsatellite instability at tetranucleotide repeats in skin and bladder cancer. *Oncogene* **21**, 4894–4899, <https://doi.org/10.1038/sj.onc.1205619>
- 88 Haugen, A.C., Goel, A., Yamada, K., Marra, G., Nguyen, T.-P., Nagasaka, T. et al. (2008) Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. *Cancer Res.* **68**, 8465–8472, <https://doi.org/10.1158/0008-5472.CAN-08-0002>
- 89 Lang, W.H., Coats, J.E., Majka, J., Hura, G.L., Lin, Y., Rasnik, I. et al. (2011) Conformational trapping of Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E837–E844, <https://doi.org/10.1073/pnas.1105461108>
- 90 Hile, S.E., Shabashev, S. and Eckert, K.A. (2013) Tumor-specific microsatellite instability: do distinct mechanisms underlie the MSI-L and EMAST phenotypes? *Mutat. Res. Mol. Mech. Mutagen.* **743–744**, 67–77, <https://doi.org/10.1016/j.mrfmmm.2012.11.003>
- 91 Mori, Y., Selaru, F.M., Sato, F., Yin, J., Simms, L.A., Xu, Y. et al. (2003) The impact of microsatellite instability on the molecular phenotype of colorectal tumors. *Cancer Res.* **63**, 4577–4582
- 92 Yao, X., Buermeier, A.B., Narayanan, L., Tran, D., Baker, S.M., Prolla, T.A. et al. (1999) Different mutator phenotypes in Mlh1- versus Pms2-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6850–6855, <https://doi.org/10.1073/pnas.96.12.6850>
- 93 Kim, T.-M., Laird, P.W. and Park, P.J. (2013) The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell* **155**, 858–868, <https://doi.org/10.1016/j.cell.2013.10.015>
- 94 Haradhvala, N.J., Kim, J., Maruvka, Y.E., Polak, P., Rosebrock, D., Livitz, D. et al. (2018) Distinct mutational signatures characterize concurrent loss of polymerase proofreading and mismatch repair. *Nat. Commun.* **9**, 1746, <https://doi.org/10.1038/s41467-018-04002-4>
- 95 Shinbrot, E., Henninger, E.E., Weinhold, N., Covington, K.R., Göksemen, A.Y., Schultz, N. et al. (2014) Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Res.* **24**, 1740–1750, <https://doi.org/10.1101/gr.174789.114>
- 96 Németh, E., Lovrics, A., Gervai, J.Z., Seki, M., Rospo, G., Bardelli, A. et al. (2020) Two main mutational processes operate in the absence of DNA mismatch repair. *DNA Rep. (Amst.)* **89**, 102827, <https://doi.org/10.1016/j.dnarep.2020.102827>
- 97 Valeri, N., Gasparini, P., Fabbri, M., Braconi, C., Veronese, A., Lovat, F. et al. (2010) Modulation of mismatch repair and genomic stability by miR-155. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6982–6987, <https://doi.org/10.1073/pnas.1002472107>
- 98 Diouf, B., Cheng, Q., Krynetskaia, N.F., Yang, W., Cheok, M., Pei, D. et al. (2011) Somatic deletions of genes regulating MSH2 protein stability cause DNA mismatch repair deficiency and drug resistance in human leukemia cells. *Nat. Med.* **17**, 1298–1303, <https://doi.org/10.1038/nm.2430>
- 99 Huang, Y., Gu, L. and Li, G.-M. (2018) H3K36me3-mediated mismatch repair preferentially protects actively transcribed genes from mutation. *J. Biol. Chem.* **293**, 7811–7823, <https://doi.org/10.1074/jbc.RA118.002839>
- 100 Shen, J., Ju, Z., Zhao, W., Wang, L., Peng, Y., Ge, Z. et al. (2018) ARID1A deficiency promotes mutability and potentiates therapeutic antitumor immunity unleashed by immune checkpoint blockade. *Nat. Med.* **24**, 556–562, <https://doi.org/10.1038/s41591-018-0012-z>
- 101 Allo, G., Bernardini, M.Q., Wu, R.-C., Shih, I.-M., Kaloger, S., Pollett, A. et al. (2014) ARID1A loss correlates with mismatch repair deficiency and intact p53 expression in high-grade endometrial carcinomas. *Mod. Pathol.* **27**, 255–261, <https://doi.org/10.1038/modpathol.2013.144>
- 102 Touat, M., Li, Y.Y., Boynton, A.N., Spurr, L.F., Iorgulescu, J.B., Bohrsen, C.L. et al. (2020) Mechanisms and therapeutic implications of hypermutation in gliomas. *Nature* **580**, 517–523, <https://doi.org/10.1038/s41586-020-2209-9>
- 103 Hunter, C., Smith, R., Cahill, D.P., Stephens, P., Stevens, C., Teague, J. et al. (2006) A hypermutation phenotype and somatic MSH6 mutations in recurrent human malignant gliomas after alkylator chemotherapy. *Cancer Res.* **66**, 3987–3991, <https://doi.org/10.1158/0008-5472.CAN-06-0127>
- 104 Pal, T., Permuth-Wey, J., Kumar, A. and Sellers, T.A. (2008) Systematic review and meta-analysis of ovarian cancers: estimation of microsatellite-high frequency and characterization of mismatch repair deficient tumor histology. *Clin. Cancer Res.* **14**, 6847–6854, <https://doi.org/10.1158/1078-0432.CCR-08-1387>
- 105 Murphy, M.A. and Wentzensen, N. (2011) Frequency of mismatch repair deficiency in ovarian cancer: a systematic review This article is a US Government work and, as such, is in the public domain of the United States of America. *Int. J. Cancer* **129**, 1914–1922, <https://doi.org/10.1002/ijc.25835>
- 106 Seplyarskiy, V.B., Andrianova, M.A. and Bazykin, G.A. (2017) APOBEC3A/B-induced mutagenesis is responsible for 20% of heritable mutations in the TpCpW context. *Genome Res.* **27**, 175–184, <https://doi.org/10.1101/gr.210336.116>
- 107 Supek, F. and Lehner, B. (2019) Scales and mechanisms of somatic mutation rate variation across the human genome. *DNA Rep. (Amst.)* **81**, 102647, <https://doi.org/10.1016/j.dnarep.2019.102647>
- 108 Lujan, S.A. and Kunkel, T.A. (2021) Stability across the whole nuclear genome in the presence and absence of DNA mismatch repair. *Cells* **10**, 1224, <https://doi.org/10.3390/cells10051224>
- 109 Lang, G.I., Parsons, L. and Gammie, A.E. (2013) Mutation rates, spectra, and genome-wide distribution of spontaneous mutations in mismatch repair deficient yeast. *G3* **3**, 1453–1465, <https://doi.org/10.1534/g3.113.006429>
- 110 Lujan, S.A., Clausen, A.R., Clark, A.B., MacAlpine, H.K., MacAlpine, D.M., Malc, E.P. et al. (2014) Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res.* **24**, 1751–1764, <https://doi.org/10.1101/gr.178335.114>
- 111 Denver, D.R., Feinberg, S., Estes, S., Thomas, W.K. and Lynch, M. (2005) Mutation rates, spectra and hotspots in mismatch repair-deficient *Caenorhabditis elegans*. *Genetics* **170**, 107–113, <https://doi.org/10.1534/genetics.104.038521>

- 112 Tijsterman, M., Pothof, J. and Plasterk, R.H.A. (2002) Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient *Caenorhabditis elegans*. *Genetics* **161**, 651–660, <https://doi.org/10.1093/genetics/161.2.651>
- 113 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**, 2169, <https://doi.org/10.1038/s41467-020-15912-7>
- 114 Póti, Á., Gyergyák, H., Németh, E., Rusz, O., Tóth, S., Kovácsházi, C. et al. (2019) Correlation of homologous recombination deficiency induced mutational signatures with sensitivity to PARP inhibitors and cytotoxic agents. *Genome Biol.* **20**, 240, <https://doi.org/10.1186/s13059-019-1867-0>
- 115 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**, 11383, <https://doi.org/10.1038/ncomms11383>
- 116 Lynch, H.T., Boland, C.R., Gong, G., Shaw, T.G., Lynch, P.M., Fodde, R. et al. (2006) Phenotypic and genotypic heterogeneity in the Lynch syndrome: diagnostic, surveillance and management implications. *Eur. J. Hum. Genet.* **14**, 390–402, <https://doi.org/10.1038/sj.ejhg.5201584>
- 117 de la Chapelle, A. (2004) Genetic predisposition to colorectal cancer. *Nat. Rev. Cancer* **4**, 769–780, <https://doi.org/10.1038/nrc1453>
- 118 Schubert, S.A., Morreau, H., de Miranda, N.F.C.C. and van Wezel, T. (2020) The missing heritability of familial colorectal cancer. *Mutagenesis* **35**, 221–231, <https://doi.org/10.1093/mutage/gez027>
- 119 Bellido, F., Pineda, M., Aiza, G., Valdés-Mas, R., Navarro, M., Puente, D.A. et al. (2016) POLE and POLD1 mutations in 529 kindred with familial colorectal cancer and/or polyposis: review of reported cases and recommendations for genetic testing and surveillance. *Genet. Med.* **18**, 325–332, <https://doi.org/10.1038/gim.2015.75>
- 120 Buchanan, D.D., Rosty, C., Clendenning, M., Spurdle, A.B. and Win, A.K. (2014) Clinical problems of colorectal cancer and endometrial cancer cases with unknown cause of tumor mismatch repair deficiency (suspected Lynch syndrome). *Appl. Clin. Genet.* **7**, 183–193
- 121 Bucksch, K., Zachariae, S., Aretz, S., Büttner, R., Holinski-Feder, E., Holzapfel, S. et al. (2020) Cancer risks in Lynch syndrome, Lynch-like syndrome, and familial colorectal cancer type X: a prospective cohort study. *BMC Cancer* **20**, 460, <https://doi.org/10.1186/s12885-020-06926-x>
- 122 Rodríguez-Soler, M., Pérez-Carbonell, L., Guarinos, C., Zapater, P., Castillejo, A., Barberá, V.M. et al. (2013) Risk of cancer in cases of suspected lynch syndrome without germline mutation. *Gastroenterology* **144**, 926.e1–932.e1, <https://doi.org/10.1053/j.gastro.2013.01.044>
- 123 Hemminki, A., Peltomäki, P., Mecklin, J.-P., Järvinen, H., Salovaara, R., Nyström-Lahti, M. et al. (1994) Loss of the wild type MLH1 gene is a feature of hereditary nonpolyposis colorectal cancer. *Nat. Genet.* **8**, 405–410, <https://doi.org/10.1038/ng1294-405>
- 124 Liu, B., Nicolaides, N.C., Markowitz, S., Willson, J.K.V., Parsons, R.E., Jen, J. et al. (1995) Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat. Genet.* **9**, 48–55, <https://doi.org/10.1038/ng0195-48>
- 125 Kane, M.F., Loda, M., Gaida, G.M., Lipman, J., Mishra, R., Goldman, H. et al. (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res.* **57**, 808–811
- 126 Roberts, M.E., Jackson, S.A., Susswein, L.R., Zeinomar, N., Ma, X., Marshall, M.L. et al. (2018) MSH6 and PMS2 germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. *Genet. Med.* **20**, 1167–1174, <https://doi.org/10.1038/gim.2017.254>
- 127 Sheehan, M., Heald, B., Yanda, C., Kelly, E.D., Grobmyer, S., Eng, C. et al. (2020) Investigating the link between lynch syndrome and breast cancer. *Eur. J. Breast Health* **16**, 106–109, <https://doi.org/10.5152/ejbh.2020.5198>
- 128 *Online Mendelian Inheritance in Man, OMIM*[®], Johns Hopkins University, Baltimore, MD, MIM Number: #158320
- 129 Sanders, M.A., Vöhringer, H., Forster, V.J., Moore, L., Campbell, B.B., Hooks, Y. et al. (2021) Life without mismatch repair. *bioRxiv* 2021.04.14.437578, <https://doi.org/10.1101/2021.04.14.437578>
- 130 Aronson, M., Colas, C., Shuen, A., Hampel, H., Foulkes, W.D., Feldman, H.B. et al. (2021) Diagnostic criteria for constitutional mismatch repair deficiency (CMMRD): recommendations from the international consensus working group. *J. Med. Genet.*, <https://doi.org/10.1136/jmedgenet-2020-107627>
- 131 Chung, J., Maruvka, Y.E., Sudhaman, S., Kelly, J., Haradhvala, N.J., Bianchi, V. et al. (2021) DNA polymerase and mismatch repair exert distinct microsatellite instability signatures in normal and malignant human cells. *Cancer Discov.* **11**, 1176–1191, <https://doi.org/10.1158/2159-8290.CD-20-0790>
- 132 *Online Mendelian Inheritance in Man, OMIM*[®], Johns Hopkins University, Baltimore, MD, MIM Number: #276300
- 133 Lindeboom, R.G.H., Supek, F. and Lehner, B. (2016) The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat. Genet.* **48**, 1112–1118, <https://doi.org/10.1038/ng.3664>
- 134 Jung, B., Doctolero, R.T., Tajima, A., Nguyen, A.K., Keku, T., Sandler, R.S. et al. (2004) Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers. *Gastroenterology* **126**, 64–659, <https://doi.org/10.1053/j.gastro.2004.01.008>
- 135 Giannakis, M., Hodis, E., Jasmine Mu, X., Yamauchi, M., Rosenbluh, J., Cibulskis, K. et al. (2014) RNF43 is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* **46**, 1264–1266, <https://doi.org/10.1038/ng.3127>
- 136 Yoon, K., Lee, S., Han, T.-S., Moon, S.Y., Yun, S.M., Kong, S.-H. et al. (2013) Comprehensive genome- and transcriptome-wide analyses of mutations associated with microsatellite instability in Korean gastric cancers. *Genome Res.* **23**, 1109–1117, <https://doi.org/10.1101/gr.145706.112>
- 137 Mayr, C. and Bartel, D.P. (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684, <https://doi.org/10.1016/j.cell.2009.06.016>
- 138 Supek, F., Miñana, B., Valcárcel, J., Gabaldón, T. and Lehner, B. (2014) Synonymous mutations frequently act as driver mutations in human cancers. *Cell* **156**, 1324–1335, <https://doi.org/10.1016/j.cell.2014.01.051>
- 139 Juul, M., Bertl, J., Guo, Q., Nielsen, M.M., Świtnicki, M., Hornshøj, H. et al. (2017) Non-coding cancer driver candidates identified with a sample- and position-specific model of the somatic mutation rate. *eLife* **6**, e21778, <https://doi.org/10.7554/eLife.21778>
- 140 Rheinbay, E., Nielsen, M.M., Abascal, F., Wala, J.A., Shapira, O., Tiao, G. et al. (2020) Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* **578**, 102–111, <https://doi.org/10.1038/s41586-020-1965-x>
- 141 Schuster-Böckler, B. and Lehner, B. (2012) Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* **488**, 504–507, <https://doi.org/10.1038/nature11273>

- 142 Hodgkinson, A., Chen, Y. and Eyre-Walker, A. (2012) The large-scale distribution of somatic mutations in cancer genomes. *Hum. Mutat.* **33**, 136–143, <https://doi.org/10.1002/humu.21616>
- 143 Li, F., Mao, G., Tong, D., Huang, J., Gu, L., Yang, W. et al. (2013) The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutS α . *Cell* **153**, 590–600, <https://doi.org/10.1016/j.cell.2013.03.025>
- 144 Peña-Díaz, J., Bregenhorn, S., Ghodgaonkar, M., Follonier, C., Artola-Borán, M., Castor, D. et al. (2012) Noncanonical mismatch repair as a source of genomic instability in human cells. *Mol. Cell* **47**, 669–680, <https://doi.org/10.1016/j.molcel.2012.07.006>
- 145 Rogozin, I.B., Goncarenco, A., Lada, A.G., De, S., Nudelman, G., Panchenko, A.R. et al. (2018) DNA polymerase η mutational signatures are found in a variety of different types of cancer. *Cell Cycle* **17**, 1–31, <https://doi.org/10.1080/15384101.2017.1404208>
- 146 Nik-Zainal, S., Alexandrov, L.B., Wedge, D.C., Van Loo, P., Greenman, C.D., Raine, K. et al. (2012) Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993, <https://doi.org/10.1016/j.cell.2012.04.024>
- 147 Roberts, S.A., Sterling, J., Thompson, C., Harris, S., Mav, D., Shah, R. 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, <https://doi.org/10.1016/j.molcel.2012.03.030>
- 148 Taylor, B.J., Nik-Zainal, S., Wu, Y.L., Stebbings, L.A., Raine, K., Campbell, P.J. et al. (2013) DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *eLife* **2**, e00534, <https://doi.org/10.7554/eLife.00534>
- 149 Sakofsky, C.J., Roberts, S.A., Malc, E., Mieczkowski, P.A., Resnick, M.A., Gordenin, D.A. et al. (2014) Break-induced replication is a source of mutation clusters underlying kataegis. *Cell Rep.* **7**, 1640–1648, <https://doi.org/10.1016/j.celrep.2014.04.053>
- 150 Chen, J., Miller, B.F. and Furano, A.V. (2014) Repair of naturally occurring mismatches can induce mutations in flanking DNA. *eLife* **3**, e02001, <https://doi.org/10.7554/eLife.02001>
- 151 Chen, J. and Furano, A.V. (2015) Breaking bad: the mutagenic effect of DNA repair. *DNA Rep. (Amst.)* **32**, 43–51, <https://doi.org/10.1016/j.dnarep.2015.04.012>
- 152 Bull, V.L. (1988) *Studies on the Mode of Cytotoxicity of Imidazotetrazinones*, University of Aston, Birmingham
- 153 Chien, C.-H., Hsueh, W.-T., Chuang, J.-Y. and Chang, K.-Y. (2021) Dissecting the mechanism of temozolomide resistance and its association with the regulatory roles of intracellular reactive oxygen species in glioblastoma. *J. Biomed. Sci.* **28**, 18, <https://doi.org/10.1186/s12929-021-00717-7>
- 154 Pegg, A.E. (1990) Mammalian O6-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogenic and therapeutic agents. *Cancer Res.* **50**, 6119–6129
- 155 Newlands, E.S., Stevens, M.F., Wedge, S.R., Wheelhouse, R.T. and Brock, C. (1997) Temozolomide: a review of its discovery, chemical properties, pre-clinical development and clinical trials. *Cancer Treat. Rev.* **23**, 35–61, [https://doi.org/10.1016/S0305-7372\(97\)90019-0](https://doi.org/10.1016/S0305-7372(97)90019-0)
- 156 Karran, P., Macpherson, P., Ceccotti, S., Dogliotti, E., Griffin, S. and Bignami, M. (1993) O6-methylguanine residues elicit DNA repair synthesis by human cell extracts. *J. Biol. Chem.* **268**, 15878–15886, [https://doi.org/10.1016/S0021-9258\(18\)82335-3](https://doi.org/10.1016/S0021-9258(18)82335-3)
- 157 Kaina, B., Christmann, M., Naumann, S. and Roos, W.P. (2007) MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. *DNA Rep. (Amst.)* **6**, 1079–1099, <https://doi.org/10.1016/j.dnarep.2007.03.008>
- 158 Yu, W., Zhang, L., Wei, Q. and Shao, A. (2020) O6-methylguanine-DNA methyltransferase (MGMT): challenges and new opportunities in glioma chemotherapy. *Front. Oncol.* **9**, 1547, <https://doi.org/10.3389/fonc.2019.01547>
- 159 Yip, S., Miao, J., Cahill, D.P., Iafrate, A.J., Aldape, K., Nutt, C.L. et al. (2009) MSH6 mutations arise in glioblastomas during temozolomide therapy and mediate temozolomide resistance. *Clin. Cancer Res.* **15**, 4622–4629, <https://doi.org/10.1158/1078-0432.CCR-08-3012>
- 160 Wang, J., Cazzato, E., Ladewig, E., Frattini, V., Rosenbloom, D.I.S., Zairis, S. et al. (2016) Clonal evolution of glioblastoma under therapy. *Nat. Genet.* **48**, 768–776, <https://doi.org/10.1038/ng.3590>
- 161 Jacob, S., Aguado, M., Fallik, D. and Praz, F. (2001) The role of the DNA mismatch repair system in the cytotoxicity of the topoisomerase inhibitors camptothecin and etoposide to human colorectal cancer cells. *Cancer Res.* **61**, 6555–6562
- 162 Fallik, D., Borriani, F., Boige, V., Viguier, J., Jacob, S., Miquel, C. et al. (2003) Microsatellite instability is a predictive factor of the tumor response to irinotecan in patients with advanced colorectal cancer. *Cancer Res.* **63**, 5738–5744
- 163 Rodriguez, R., Hansen, L.T., Phear, G., Scorah, J., Spang-Thomsen, M., Cox, A. et al. (2008) Thymidine selectively enhances growth suppressive effects of camptothecin/irinotecan in MSI+ cells and tumors containing a mutation of MRE11. *Clin. Cancer Res.* **14**, 5476–5483, <https://doi.org/10.1158/1078-0432.CCR-08-0274>
- 164 Vilar, E., Scaltriti, M., Balmaña, J., Saura, C., Guzman, M., Arribas, J. et al. (2008) Microsatellite instability due to hMLH1 deficiency is associated with increased cytotoxicity to irinotecan in human colorectal cancer cell lines. *Br. J. Cancer* **99**, 1607–1612, <https://doi.org/10.1038/sj.bjc.6604691>
- 165 Kim, N.-G., Choi, Y.R., Baek, M.J., Kim, Y.H., Kang, H., Kim, N.K. et al. (2001) Frameshift mutations at coding mononucleotide repeats of the hRAD50 gene in gastrointestinal carcinomas with microsatellite instability. *Cancer Res.* **61**, 36–38
- 166 Giannini, G., Rinaldi, C., Ristori, E., Ambrosini, M.I., Cerignoli, F., Viel, A. et al. (2004) Mutations of an intronic repeat induce impaired MRE11 expression in primary human cancer with microsatellite instability. *Oncogene* **23**, 2640–2647, <https://doi.org/10.1038/sj.onc.1207409>
- 167 Giannini, G., Ristori, E., Cerignoli, F., Rinaldi, C., Zani, M., Viel, A. et al. (2002) Human MRE11 is inactivated in mismatch repair-deficient cancers. *EMBO Rep.* **3**, 248–254, <https://doi.org/10.1093/embo-reports/kvf044>
- 168 Wen, Q., Scorah, J., Phear, G., Rodgers, G., Rodgers, S. and Meuth, M. (2008) A mutant allele of MRE11 found in mismatch repair-deficient tumor cells suppresses the cellular response to DNA replication fork stress in a dominant negative manner. *Mol. Biol. Cell* **19**, 1693–1705, <https://doi.org/10.1091/mbc.e07-09-0975>
- 169 Miquel, C., Jacob, S., Grandjouan, S., Aimé, A., Viguier, J., Sabourin, J.-C. et al. (2007) Frequent alteration of DNA damage signalling and repair pathways in human colorectal cancers with microsatellite instability. *Oncogene* **26**, 5919–5926, <https://doi.org/10.1038/sj.onc.1210419>
- 170 Koppensteiner, R., Samartzis, E.P., Noske, A., von Teichman, A., Dedes, I., Gwerder, M. et al. (2014) Effect of MRE11 loss on PARP-inhibitor sensitivity in endometrial cancer in vitro. *PLoS ONE* **9**, e100041, <https://doi.org/10.1371/journal.pone.0100041>

- 171 Vilar, E., Bartnik, C.M., Stenzel, S.L., Raskin, L., Ahn, J., Moreno, V. et al. (2011) MRE11 deficiency increases sensitivity to poly(ADP-ribose) polymerase inhibition in microsatellite instable colorectal cancers. *Cancer Res.* **71**, 2632–2642, <https://doi.org/10.1158/0008-5472.CAN-10-1120>
- 172 McCabe, N., Turner, N.C., Lord, C.J., Kluzek, K., Bialkowska, A., Swift, S. et al. (2006) Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. *Cancer Res.* **66**, 8109–8115, <https://doi.org/10.1158/0008-5472.CAN-06-0140>
- 173 Leichman, L., Groshen, S., O'Neil, B.H., Messersmith, W., Berlin, J., Chan, E. et al. (2016) Phase II Study of Olaparib (AZD-2281) after standard systemic therapies for disseminated colorectal cancer. *Oncologist* **21**, 172–177, <https://doi.org/10.1634/theoncologist.2015-0319>
- 174 Genthner Williams, S.M., Kuznicki, A.M., Andrade, P., Dolinski, B.M., Elbi, C., O'Hagan, R.C. et al. (2015) Treatment with the PARP inhibitor, niraparib, sensitizes colorectal cancer cell lines to irinotecan regardless of MSI/MSS status. *Cancer Cell Int.* **15**, 14, <https://doi.org/10.1186/s12935-015-0162-8>
- 175 Suzuki, S., Iwaizumi, M., Tseng-Rogenski, S., Hamaya, Y., Miyajima, H., Kanaoka, S. et al. (2016) Production of truncated MBD4 protein by frameshift mutation in DNA mismatch repair-deficient cells enhances 5-fluorouracil sensitivity that is independent of hMLH1 status. *Cancer Biol. Ther.* **17**, 760–768, <https://doi.org/10.1080/15384047.2016.1178430>
- 176 Suzuki, S., Iwaizumi, M., Yamada, H., Sugiyama, T., Hamaya, Y., Furuta, T. et al. (2017) MBD4 frameshift mutation caused by DNA mismatch repair deficiency enhances cytotoxicity by trifluridine, an active antitumor agent of TAS-102, in colorectal cancer cells. *Oncotarget* **9**, 11477–11488
- 177 Dorard, C., de Thonel, A., Collura, A., Marisa, L., Svrcek, M., Lagrange, A. et al. (2011) Expression of a mutant HSP110 sensitizes colorectal cancer cells to chemotherapy and improves disease prognosis. *Nat. Med.* **17**, 1283–1289, <https://doi.org/10.1038/nm.2457>
- 178 Collura, A., Lagrange, A., Svrcek, M., Marisa, L., Buhard, O., Guilloux, A. et al. (2014) Patients with colorectal tumors with microsatellite instability and large deletions in HSP110 T17 have improved response to 5-fluorouracil-based chemotherapy. *Gastroenterology* **146**, 401.e1–411.e1, <https://doi.org/10.1053/j.gastro.2013.10.054>
- 179 Bokhari, A., Jonchere, V., Lagrange, A., Bertrand, R., Svrcek, M., Marisa, L. et al. (2018) Targeting nonsense-mediated mRNA decay in colorectal cancers with microsatellite instability. *Oncogenesis* **7**, 1–9, <https://doi.org/10.1038/s41389-018-0079-x>
- 180 Longley, D.B., Harkin, D.P. and Johnston, P.G. (2003) 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**, 330–338, <https://doi.org/10.1038/nrc1074>
- 181 Bertagnoli, M.M., Niedzwiecki, D., Compton, C.C., Hahn, H.P., Hall, M., Damas, B. et al. (2009) Microsatellite instability predicts improved response to adjuvant therapy with irinotecan, fluorouracil, and leucovorin in Stage III colon cancer: Cancer and Leukemia Group B Protocol 89803. *J. Clin. Oncol.* **27**, 1814–1821, <https://doi.org/10.1200/JCO.2008.18.2071>
- 182 Jover, R., Zapater, P., Castells, A., Llor, X., Andreu, M., Cubiella, J. et al. (2006) Mismatch repair status in the prediction of benefit from adjuvant fluorouracil chemotherapy in colorectal cancer. *Gut* **55**, 848–855, <https://doi.org/10.1136/gut.2005.073015>
- 183 Carethers, J.M., Smith, E.J., Behling, C.A., Nguyen, L., Tajima, A., Doctolero, R.T. et al. (2004) Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer. *Gastroenterology* **126**, 394–401, <https://doi.org/10.1053/j.gastro.2003.12.023>
- 184 Ribic, C.M., Sargent, D.J., Moore, M.J., Thibodeau, S.N., French, A.J., Goldberg, R.M. et al. (2003) Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N. Engl. J. Med.* **349**, 247–257, <https://doi.org/10.1056/NEJMoa022289>
- 185 De Angelis, P.M., Svendsrud, D.H., Kravik, K.L. and Stokke, T. (2006) Cellular response to 5-fluorouracil (5-FU) in 5-FU-resistant colon cancer cell lines during treatment and recovery. *Mol. Cancer* **5**, 20, <https://doi.org/10.1186/1476-4598-5-20>
- 186 Iwaizumi, M., Tseng-Rogenski, S. and Carethers, J.M. (2011) DNA mismatch repair proficiency executing 5-fluorouracil cytotoxicity in colorectal cancer cells. *Cancer Biol. Ther.* **12**, 756–764, <https://doi.org/10.4161/cbt.12.8.17169>
- 187 Li, L.S., Morales, J.C., Veigl, M., Sedwick, D., Greer, S., Meyers, M. et al. (2009) DNA mismatch repair (MMR)-dependent 5-fluorouracil cytotoxicity and the potential for new therapeutic targets. *Br. J. Pharmacol.* **158**, 679–692, <https://doi.org/10.1111/j.1476-5381.2009.00423.x>
- 188 Jin, Z. and Sinicrope, F.A. (2021) Prognostic and predictive values of mismatch repair deficiency in non-metastatic colorectal cancer. *Cancers* **13**, 300, <https://doi.org/10.3390/cancers13020300>
- 189 Evensen, N.A. and Carroll, W.L. (2021) dMMR and thiopurines strand TP53 for ALL relapse. *Nat. Cancer* **2**, 777–779, <https://doi.org/10.1038/s43018-021-00248-y>
- 190 Brem, R. and Karran, P. (2012) Oxidation-mediated DNA cross-linking contributes to the toxicity of 6-thioguanine in human cells. *Cancer Res.* **72**, 4787–4795, <https://doi.org/10.1158/0008-5472.CAN-12-1278>
- 191 Evensen, N.A., Madhusoodhan, P.P., Meyer, J., Saliba, J., Chowdhury, A., Araten, D.J. et al. (2018) MSH6 haploinsufficiency at relapse contributes to the development of thiopurine resistance in pediatric B-lymphoblastic leukemia. *Haematologica* **103**, 830–839, <https://doi.org/10.3324/haematol.2017.176362>
- 192 Yang, F., Brady, S.W., Tang, C., Sun, H., Du, L., Barz, M.J. et al. (2021) Chemotherapy and mismatch repair deficiency cooperate to fuel TP53 mutagenesis and ALL relapse. *Nat. Cancer* **2**, 819–834, <https://doi.org/10.1038/s43018-021-00230-8>
- 193 Karran, P. (2006) Thiopurines, DNA damage, DNA repair and therapy-related cancer. *Br. Med. Bull.* **79–80**, 153–170, <https://doi.org/10.1093/bmb/ldl020>
- 194 Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J.M., Desrichard, A. et al. (2014) Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199, <https://doi.org/10.1056/NEJMoa1406498>
- 195 Riaz, N., Blecua, P., Lim, R.S., Shen, R., Higginson, D.S., Weinhold, N. et al. (2017) Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. *Nat. Commun.* **8**, 857, <https://doi.org/10.1038/s41467-017-00921-w>
- 196 Yarmoan, M., Hopkins, A. and Jaffee, E.M. (2017) Tumor mutational burden and response rate to PD-1 inhibition. *N. Engl. J. Med.* **377**, 2500–2501, <https://doi.org/10.1056/NEJMc1713444>
- 197 Litchfield, K., Reading, J.L., Puttick, C., Thakkar, K., Abbosh, C., Bentham, R. et al. (2021) Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* **184**, 596.e14–614.e14, <https://doi.org/10.1016/j.cell.2021.01.002>

- 198 Segal, N.H., Parsons, D.W., Peggs, K.S., Velculescu, V., Kinzler, K.W., Vogelstein, B. et al. (2008) Epitope landscape in breast and colorectal cancer. *Cancer Res.* **68**, 889–892, <https://doi.org/10.1158/0008-5472.CAN-07-3095>
- 199 Schwitalle, Y., Kloor, M., Eiermann, S., Linnebacher, M., Kienle, P., Knaebel, H.P. et al. (2008) Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. *Gastroenterology* **134**, 988–997, <https://doi.org/10.1053/j.gastro.2008.01.015>
- 200 Saeterdal, I., Bjorheim, J., Lislud, K., Gjertsen, M.K., Bukholm, I.K., Olsen, O.C. et al. (2001) Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13255–13260, <https://doi.org/10.1073/pnas.231326898>
- 201 Linnebacher, M., Gebert, J., Rudy, W., Woerner, S., Yuan, Y.P., Bork, P. et al. (2001) Frameshift peptide-derived T-cell epitopes: a source of novel tumor-specific antigens. *Int. J. Cancer* **93**, 6–11, <https://doi.org/10.1002/ijc.1298>
- 202 Hodi, F.S., O’Day, S.J., McDermott, D.F., Weber, R.W., Sosman, J.A., Haanen, J.B. et al. (2010) Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* **363**, 711–723, <https://doi.org/10.1056/NEJMoa1003466>
- 203 Larkin, J., Minor, D., D’Angelo, S., Neyns, B., Smylie, M., Miller, W.H. et al. (2018) Overall Survival in Patients With Advanced Melanoma Who Received Nivolumab Versus Investigator’s Choice Chemotherapy in CheckMate 037: a randomized, controlled, open-label Phase III trial. *J. Clin. Oncol.* **36**, 383–390, <https://doi.org/10.1200/JCO.2016.71.8023>
- 204 Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D.R., Steins, M., Ready, N.E. et al. (2015) Nivolumab versus docetaxel in advanced non-squamous non-small cell lung cancer. *N. Engl. J. Med.* **373**, 1627–1639, <https://doi.org/10.1056/NEJMoa1507643>
- 205 Timmermann, B., Kerick, M., Roehr, C., Fischer, A., Isau, M., Boerno, S.T. et al. (2010) Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis. *PLoS ONE* **5**, e15661, <https://doi.org/10.1371/journal.pone.0015661>
- 206 Le, D.T., Uram, J.N., Wang, H., Bartlett, B.R., Kemberling, H., Eyring, A.D. et al. (2015) PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520, <https://doi.org/10.1056/NEJMoa1500596>
- 207 Duffy, M.J. and Crown, J. (2019) Biomarkers for predicting response to immunotherapy with immune checkpoint inhibitors in cancer patients. *Clin. Chem.* **65**, 1228–1238, <https://doi.org/10.1373/clinchem.2019.303644>
- 208 Chowell, D., Yoo, S.-K., Valero, C., Pastore, A., Krishna, C., Lee, M. et al. (2021) Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat. Biotechnol.* 1–8, <https://doi.org/10.1038/s41587-021-01070-8>
- 209 Turajlic, S., Litchfield, K., Xu, H., Rosenthal, R., McGranahan, N., Reading, J.L. et al. (2017) Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol.* **18**, 1009–1021, [https://doi.org/10.1016/S1470-2045\(17\)30516-8](https://doi.org/10.1016/S1470-2045(17)30516-8)
- 210 Litchfield, K., Reading, J.L., Lim, E.L., Xu, H., Liu, P., Al-Bakir, M. et al. (2020) Escape from nonsense-mediated decay associates with anti-tumor immunogenicity. *Nat. Commun.* **11**, 3800, <https://doi.org/10.1038/s41467-020-17526-5>
- 211 Pastor, F., Kolonias, D., Giangrande, P.H. and Gilboa, E. (2010) Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. *Nature* **465**, 227–230, <https://doi.org/10.1038/nature08999>
- 212 Behan, F.M., Iorio, F., Picco, G., Gonçalves, E., Beaver, C.M., Migliardi, G. et al. (2019) Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* **568**, 511–516, <https://doi.org/10.1038/s41586-019-1103-9>
- 213 Chan, E.M., Shibue, T., McFarland, J.M., Gaeta, B., Ghandi, M., Dumont, N. et al. (2019) WRN helicase is a synthetic lethal target in microsatellite unstable cancers. *Nature* **568**, 551–556, <https://doi.org/10.1038/s41586-019-1102-x>
- 214 Kategaya, L., Perumal, S.K., Hager, J.H. and Belmont, L.D. (2019) Werner syndrome helicase is required for the survival of cancer cells with microsatellite instability. *iScience* **13**, 488–497, <https://doi.org/10.1016/j.isci.2019.02.006>
- 215 Lieb, S., Blaha-Ostermann, S., Kamper, E., Rippka, J., Schwarz, C., Ehrenhöfer-Wölfer, K. et al. (2019) Werner syndrome helicase is a selective vulnerability of microsatellite instability-high tumor cells. *eLife* **8**, e43333, <https://doi.org/10.7554/eLife.43333>
- 216 van Wietmarschen, N., Sridharan, S., Nathan, W.J., Tubbs, A., Chan, E.M., Callen, E. et al. (2020) Repeat expansions confer WRN dependence in microsatellite-unstable cancers. *Nature* **586**, 292–298, <https://doi.org/10.1038/s41586-020-2769-8>

Supplementary Table S1. Overview of subtypes of MMR deficiencies observed in cancers, their tissue-specificity, mutational footprints and underlying mechanisms.

MMR failure subtype	Type of genomic instability	Cancer Types involved	Genes and mechanisms
MSI-H, estimated by Bethesda panel	Indels at microsatellite loci (five mononucleotide or dinucleotide repeats are examined), usually assayed using PCR and capillary electrophoresis to measure length distribution of the locus.	<p>Commonly: colorectal, uterine, stomach.</p> <p>Sometimes: ovarian, esophageal, prostate, leukemia, brain.</p> <p>Also seen infrequently in many other cancer types.</p>	<p>Commonly: silencing of <i>MLH1</i> by promoter hypermethylation, or germline variants in <i>MSH2</i>, <i>MLH1</i>, <i>PMS2</i> or <i>MSH6</i> genes with a somatic inactivation of the second allele (Lynch syndrome).</p> <p>Sometimes: somatic mutations in <i>MSH2</i>, <i>MLH1</i>, <i>PMS2</i> or <i>MSH6</i>.</p> <p>Note: Bethesda panel classifiers were reported to sometimes miss <i>MSH6</i>-deficient cancers.</p>
MSI-H, estimated by genomics classifiers	High number of indels at many microsatellite loci considered jointly, estimated exome-wide or genome-wide from sequencing data.	As above, with additional tissues implicated from genomic analysis such as kidney, liver, and breast cancers (it is unclear to what extent this instability at many MS loci actually reflects MMR failures, or is due to other causes).	Likely same as above; additional factors might be involved. For example, loss-of-function of the chromatin modifiers SETD2 and ARID1A has been associated with some MMR mutational phenotypes.
MSI-L	<p>Instability at 1 out of 5 Bethesda loci is observed.</p> <p>Associated with instability at dinucleotide rather than mononucleotide repeat loci.</p> <p>May in some cases result from a random occurrence of an indel at 1 of the 5 Bethesda loci in an otherwise MSS</p>	Same as MSI-H types, and also additional cancer types e.g. melanoma and pancreatic cancer.	<p>Possibly, loss of activity of the <i>PMS2</i> gene.</p> <p>Some cases of <i>MSH6</i> deficient tumors are observed as MSI-L.</p> <p>Some evidence for involvement of <i>MSH3</i> and <i>MGMT</i> proteins, as well as proteins in base excision repair, translesion synthesis, and replicative DNA</p>

	tumor		polymerases.
EMAST ("elevated microsatellite alterations at selected tetranucleotide repeats")	Tetranucleotide repeats, primarily assayed at AAAG or ATAG repeats.	Lung, skin, bladder, colon cancer.	<i>MSH3</i> gene alterations are identified as one of the causes, however given the reported prevalence of EMAST additional causes seem likely.
Mutational signatures known to be associated with MSI-H status	High 'exposure' (mutation burden from WGS/WES) of signatures Cosmic SBS6/14/15/20/21/26/44 or Cosmic ID1/2/7 of RefSig-MMR1	Same as MSI-H cancer types.	Likely same as MSI-H genes, with possible additional causes (alterations in trans- acting genes that regulate MMR).
Additional mutational signatures not clearly associated with the Bethesda MSI-H status	High 'exposure' (i.e. mutation burden measured by WGS of WES) of the mutational signatures Cosmic SBS12 or ID4, or the RefSig MMR-2.	Various cancer types, e.g. liver, ovarian, lymphoid, breast, kidney.	Unclear. <i>PMS2</i> gene alterations are one possible cause. Loss-of-function of the chromatin modifiers SETD2 and ARID1A has been associated with some MMR mutational phenotypes.