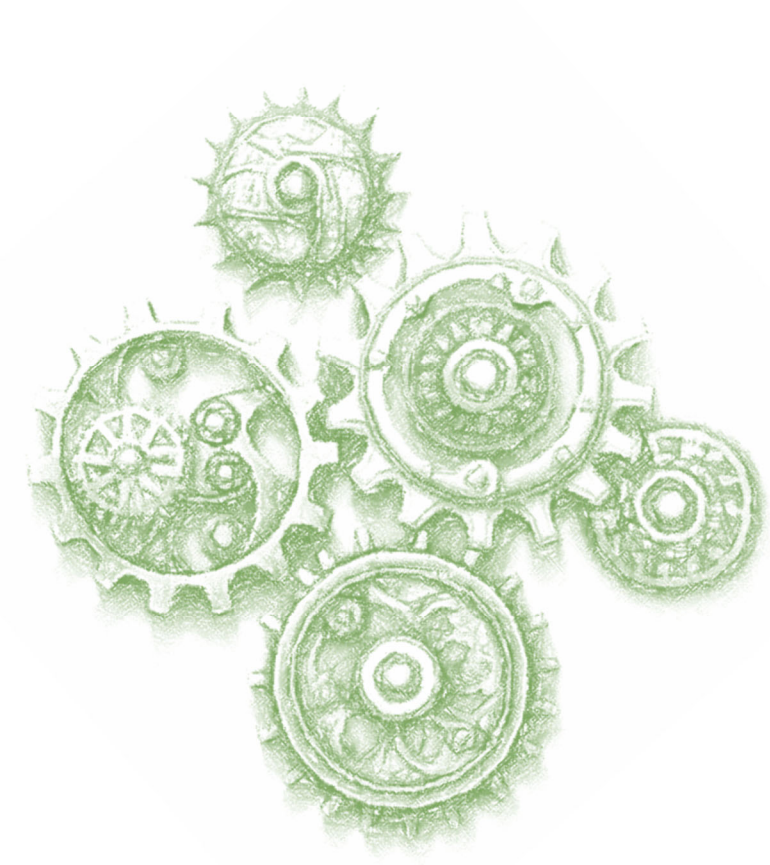
Control of replication associated DNA damage responses by Mismatch Repair

Ijsselsteiijn, R.



Chapter 5:

Induction of mismatch repair deficiency, compromised DNA damage signaling and compound hypermutagenesis by a dietary mutagen in a cell-based model for Lynch syndrome

Robbert Ijsselsteijn[1], Sandrine van Hees[1], Mark Drost[1], Jacob G Jansen[1], Niels de Wind[1]

[1]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands.

DOI: 10.1093/carcin/bgab108

Carcinogenesis. 2022 Mar 24;43(2):160-169

Published December 2021

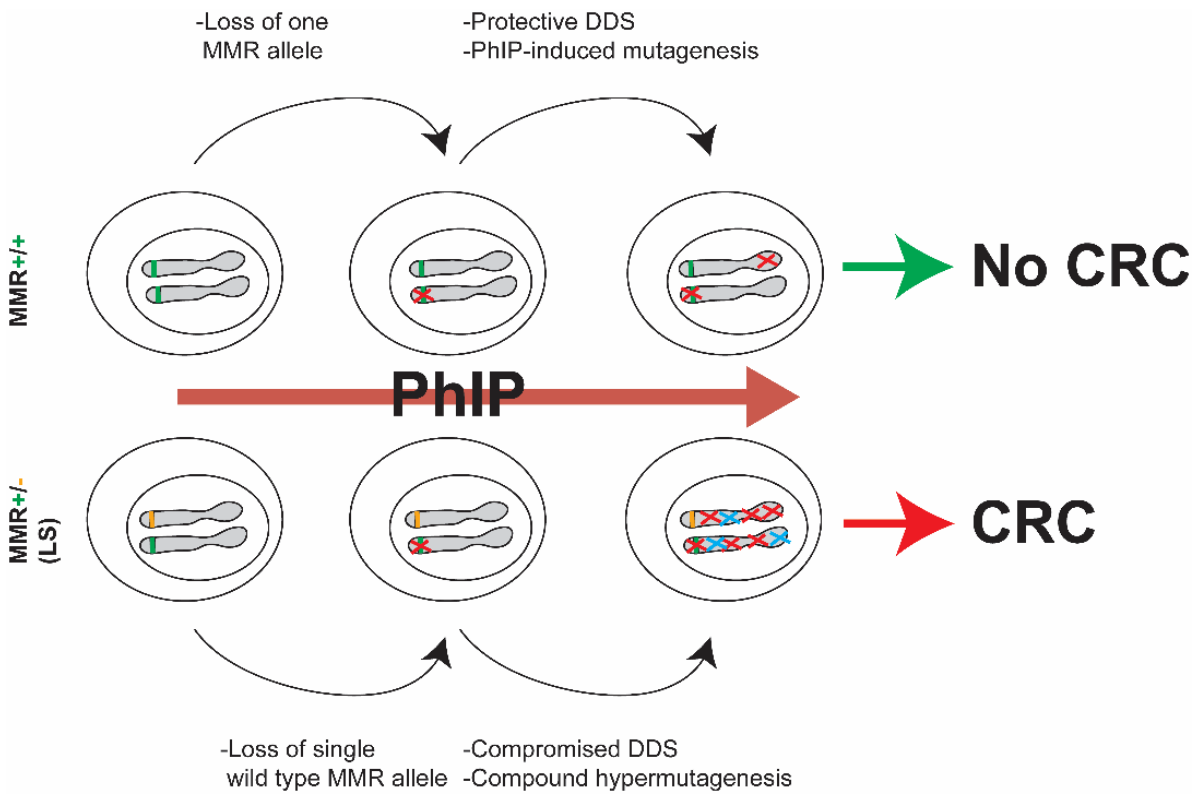
Abstract

The prevalent cancer predisposition Lynch syndrome (LS, OMIM #120435) is caused by an inherited heterozygous defect in any of the four core DNA mismatch repair (MMR) genes *MSH2*, *MSH6*, *MLH1* or *PMS2*. MMR repairs errors by the replicative DNA polymerases in all proliferating tissues. Its deficiency, following somatic loss of the wild type copy, results in a spontaneous mutator phenotype that underlies the rapid development of, predominantly, colorectal cancer (CRC) in LS. Here we have addressed the hypothesis that aberrant responses of intestinal stem cells to diet-derived mutagens may be causally involved in the restricted cancer tropism of LS. To test this we have generated a panel of isogenic mouse embryonic stem (mES) cells with heterozygous or homozygous disruption of multiple MMR genes and investigated their responses to the common dietary mutagen and carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP). Our data reveal that PhIP can inactivate the wild type allele of heterozygous mES cells via the induction of either loss of heterozygosity (LOH) or intragenic mutations. Moreover, while protective DNA damage signaling (DDS) is compromised, PhIP induces more mutations in *Msh2*, *Mlh1*, *Msh6* or *Pms2*-deficient mES cells than in wild type cells. Combined with their spontaneous mutator phenotypes, this results in a compound hypermutator phenotype. Together, these results indicate that dietary mutagens may promote CRC development in LS at multiple levels, providing a rationale for dietary modifications in the management of LS.

Summary

A dietary mutagen induces loss of the wild type allele in mismatch repair gene-heterozygous cells. This results in impaired DNA damage signaling and in compound hypermutagenesis. These findings provide insights into the etiology of colorectal carcinogenesis in Lynch syndrome.

Graphical abstract



Graphical abstract. Proposed roles of dietary mutagens in the etiology of colorectal cancer in Lynch syndrome.
While individuals normally carry two intact alleles of the four MMR genes (top panel, green rectangles), LS patients carry a germ-line defect in one allele of any of these genes (bottom panel, orange cross). The single wild type allele of that gene may inadvertently be lost in stem cells in a colonic crypt, either spontaneously or induced by exposure to an intestinal mutagen, such as PhIP (red cross). This results in compromised DNA damage signaling (DDS) and, possibly, loss of protective cell cycle responses to protracted PhIP exposure. MMR deficiency not only results in a spontaneous mutator phenotype (blue crosses), but also in increased mutability by PhIP (red crosses). The consequent compound hypermutator phenotype may strongly predispose to oncogenic derailment.

Introduction

Colorectal cancer (CRC) is the third most common cancer type world-wide, with the highest incidence in developed countries (1). Many lifestyle factors increase the risk for CRC, including obesity, smoking and consumption of red, processed and cooked meat (2, 3). Ten percent of all CRC has an underlying genetic predisposition. Of these, Lynch syndrome (LS) is the most prevalent (1 in 279 individuals), accounting for 3-5% of all CRCs (4). LS is caused by an inherited heterozygous defect in one of four genes involved in DNA mismatch repair (MMR), *MSH2*, *MSH6*, *MLH1* and *PMS2* Although LS is inherited in an autosomal dominant fashion, the wild type allele of the germ-line-defective gene must be somatically lost to procure cancer development (5).

Canonical MMR removes misincorporations by DNA polymerases during replication of undamaged DNA. Consequently, loss of MMR is associated with increased spontaneous mutagenesis. Since MMR operates in all proliferating tissues, the cause of the restricted cancer tropism of LS thus far is unclear. Interestingly, MMR-deficient cells also display aberrant mutagenic and cell cycle responses to agents that induce helix-distorting DNA lesions (6, 7). This appears paradoxical since helix-distorting DNA lesions do not induce base-base mismatches, the natural substrates for MMR. Data on responses of MMR-deficient cells to UV light have suggested the involvement of MMR proteins in the removal of misincorporations induced by DNA translesion synthesis polymerases opposite helix-distorting photolesions. This prevents their mutagenicity, while simultaneously inducing protective DNA damage signaling (DDS) (7).

Heterocyclic amines are an important class of dietary genotoxic compounds, present in meat cooked at high temperatures, that induce helix-distorting nucleotide lesions. The most abundant of these heterocyclic amines is 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP). Indeed, dietary intake of PhIP is positively correlated with CRC (3). Here we have investigated the hypothesis that PhIP (or similar diet-derived genotoxic compounds) may direct the development of CRC in LS at three levels (see the Graphical Abstract and Ref. 7): (i) by inducing loss of the wild type allele of the heterozygous MMR gene, (ii) by compromised PhIP-induced DDS in these MMR-deficient cells and (iii) by enhanced mutability of these cells by the mutagen. To address this tripartite hypothesis we have used mouse embryonic stem (mES) cells, heterozygous for *Msh2* or *Mlh1,* as models for colonic crypt stem cells in LS patients. We show that PhIP indeed induces loss of heterozygosity (LOH) or deleterious nucleotide substitutions, insertions, and deletions at the wild type allele. Consequently, in the resulting *Msh2*-deficient mES cells, PhIP-induced DDS is reduced. Moreover, in *Msh2*-, *Msh6*-, *Mlh1*-, or *Pms2*-deficient cells, the mutagenicity of PhIP is much higher than in wild type cells. This exacerbated mutability is additive to the spontaneous mutator phenotype of MMR-deficient cells, resulting in a compound hypermutator phenotype. These results provide support for the hypothesis that intestinal mutagens are involved at multiple levels in the development of CRC in LS and suggest the feasibility of dietary intervention as a preventive approach.

Materials and Methods

Cell culture and cell lines generated and used and their validation

mES cells were cultured on sub-lethally irradiated mouse embryonic fibroblast (MEF) feeder cells in complete medium consisting of KO DMEM supplemented with 10% fetal calf serum, 1% glutamax, 1% non-essential amino acids, 1mM pyruvate, 100U penicillin/100µg streptomycin, 0.1mM β-mercapto-ethanol and leukemia inhibitory factor. During experiments complete medium was mixed in a 1:1 ratio with Buffalo rat liver cell-conditioned medium (called 50/50) to allow for growth on gelatin-coated culture dishes.

Wild type mES cell line E14 (8) was used as a parental line to all cell lines generated and used in this study (Fig. 1A). The line was karyotyped before constructing derivative mutant mES cell lines. The *Msh2* and *Mlh1*-heterozygous mES cell lines used to study spontaneous and PhIP-induced loss of MMR (called Msh2-Bsd and Mlh1-Bsd) contain a *Blasticidin* selection cassette at the *Kcnk12* (3' of the wild type *Msh2* allele) or *Llrfip2* (3' of the wild type *Mlh1* allele) locus, respectively, introduced by classical gene targeting. The heterozygous *Msh2* and *Mlh1* deletions in these lines were generated using CRISPR/Cas9 (Supplemental Methods Table 1; manuscript in preparation). Briefly, two complementary oligonucleotides with BbsI overhangs were annealed and ligated into CRISPR-Cas9 vector PX330-Puro. mES cells were transfected with these plasmids using Lipofectamine 2000. After transfection, the introduction of a hemizygous deletion at *Msh2* or *Mlh1* was confirmed by allele-specific PCR (see below). The presence of the *Blasticidin* cassette linked to the wild type *Mlh1* or *Msh2* allele was validated by the appearance of clones, surviving 4 hours incubation with 20µM 6-thioguanine (6tG) (Sigma-Aldrich), a hallmark of MMR deficiency (9). All these spontaneous 6tG-tolerant clones had fortuitously lost the wild type allele by LOH and also the (linked) *Blasticidin* cassette, which resulted in re-acquired Blasticidin sensitivity (Supplemental Fig. S2 and the Results section).

The *Msh2*-1, *Msh6*-1 (10), *Mlh1*-1, *Mlh1*-s, *Pms2*-1 and *Pms2*-s knock-out cell lines (Fig. 1A) have been generated using CRISPR-Cas9, as described above (Supplemental Methods Table 1). After transfection, the cells were selected for tolerance of 6tG to more easily acquire MMR-deficient clones, with the exception of the *Pms2*-deficient lines that were identified by allele-specific PCR. Their phenotype was confirmed by Western blotting (Fig. 1B and Supplemental data only for review).

The *Msh2*-s, the *Msh6*-s and *Msh3* knock-out cell lines have been generated by conventional gene targeting (Fig. 1A) and have been published before by us (11, 12). All lines were validated prior to use by allele-specific PCR, as described (11, 12). In addition, the *Msh2*-1 and *Msh6*-1 lines were validated prior to use by testing for 6tG tolerance and by performing Western blotting (see below; Fig. 1B and Supplemental data only for review). Cell lines were validated regularly by Western blotting and/or allele-specific PCR. An overview of the origin or construction of all cell lines generated and used here and their validation/authentication are provided in Supplemental methods Table 2.

Western blotting

SDS-PAGE was performed using 4-12% Criterion XT Bis-Tris gels (Bio-rad). Proteins were transferred onto Protran 0,45µM nitrocellulose membranes (GE Healthcare). Membranes were blocked for one hour using blocking reagent (Rockland) diluted with PBS-0.1% Tween. Then, membranes were incubated overnight at 4°C with primary antibodies, diluted in Rockland-PBS-Tween. Membranes were washed using PBS-Tween (0.1%) and incubated with secondary anti-mouse and anti-rabbit HRP (ThermoFisher Scientific), depending on primary antibody isotype, diluted 1:50000 in Rockland-PBS-Tween for one hour at room temperature. Bands were visualized using Amersham ECL select western blotting detection reagent (GE Healthcare). Antibodies used: Anti-Msh2: mouse mAb (FE11, Calbiochem), anti-Msh6 (Abcam, clone 44), anti-Mlh1: rabbit polyclonal (C-20, Santa Cruz Biotechnology), anti-Pms2: mouse Ab (A16-4, BD Pharmingen), anti-Kap-1^{p} (Bethyl, polyclonal A300-767A), anti-Chk1^{p} (clone 133D3, Cell Signaling Technology) and anti-PCNA (clone PC10 clone, Santa Cruz).

Induction of MMR deficiency by PhIP

A day prior to treatment with PhIP (Apollo Scientific) or vehicle (dimethyl sulfoxide, DMSO), 5x10^{6} wild type cells or cells heterozygous for *Msh2* or *Mlh1* were seeded in 50/50 medium in gelatin-coated p90 dishes. Cells were washed twice with PBS and then incubated with 18µM PhIP or DMSO in complete medium in the presence of 10% S9 rat liver extract (Trinova), for three hours. Subsequently, the cells were washed twice with PBS and 50/50 medium was added with or without 5U/ml of Blasticidin (Invivogen). After two days, the cells were trypsinized and cultured for two passages to allow fixation of mutations. Afterwards, 2x10^{6} cells were seeded in a p90 dish and the next day the cells were incubated for four hours with 20µM 6tG in 50/50 medium to select for MMR-deficient clones. After one week, this selection was repeated and the cells were grown for another week in 50/50 medium to allow for the formation of clones. Clones were either stained with methylene blue and counted or picked and grown until confluency in 96-wells plates in the presence of 5mM Hypoxanthine, 20µM Aminopterin, 0.8µM Thymidine (HAT) (50x diluted, Thermo Fisher Scientific) to select against the inadvertent loss of *Hprt* that also yields 6tG resistance (albeit to a much higher concentration).

Amplification of *Msh2* and *Mlh1* for LOH analysis

Individual clones were lysed using 50µl DirectPCR lysis reagent (Viagen) with 8U/ml proteinase K (Invitrogen) for one hour at 37^{o}C. After heat inactivation of proteinase K (5 min at 85^{o}C), one µl of cell lysate was used in a multiplex PCR with three oligonucleotide primers to analyze LOH (Supplemental Figure 1, Supplemental Methods Table 3). LOH at Msh2 was analyzed using primers 1 (0.4µM) and 2 (0.08µM) to amplify the intact *Msh2* allele and primers 1 and 3 (0.32µM) to amplify the disrupted *Msh2* allele. For *Mlh1*, primers 4 (0.4µM) and 5 (0.16µM) amplify the intact *Mlh1* allele and 4 (0.4µM) and 6 (0.4µM) the disrupted *Mlh1* allele. PCR products were created using GoTaq polymerase (0.625U, Promega) in GoTaq buffer with dNTPs (0,5mM) and primers during 35 cycles, each consisting of 30 sec at 95°C, 30 sec at 57°C (*Msh2*)

or 61°C (*Mlh1*) and two minutes at 72°C. A final extension was performed for 10 minutes at 72°C. DNA products were examined using 3% agarose gel electrophoresis.

Production and sequence analysis of *Msh2* and *Mlh1* cDNA

RNA was isolated from individual clones lysed in 100µl TRIzol reagent (Invitrogen) according to the manufacturer's protocol (TRIzol, Invitrogen). The RNA pellet was dissolved in 15µl TE buffer and cDNA was generated using Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific). First, a final volume of 14.5µl containing one µl RNA, OligodT primer (6.9µM, Thermo Fisher Scientific) and dNTPs (0.69mM, Invitrogen) was incubated for five minutes at 65°C. Then, 1µl Maxima Polymerase (200U/µl), 4µl of 5x Maxima buffer and 20U of RNAsin (Promega) were added to a final volume of 20µl. cDNA synthesis was performed for one hour at 57°C. After heat inactivation at 85°C for five minutes, cDNA was PCR amplified using GoTaq polymerase as described above with an annealing temperature of 55°C and primers 7/8 and 9/10 for the 5' and 3' end of *Msh2* and primers 11/12 and 13/14 for both ends of *Mlh1*. PCR products were analyzed using 2% agarose gel electrophoresis. PCR products were purified using a QIAquick PCR purification kit (Qiagen) following manufacturer's protocol and eluted in a final volume of 20µl deionized H_2O. Five µl of purified PCR product was used for Sanger sequencing using 1,25µM of primer 8/10/12/15 for the 5' and 3' end of *Msh2* and *Mlh1*, respectively. Sequences were analyzed using ContigExpress from VectorNTI Suite 9. Primer sequences are listed in Supplemental Methods table 3.

Analysis of pathogenicity of PhIP-induced *Msh2* and *Mlh1* mutations

To analyze whether PhIP-induced mutations in MMR heterozygous mES cells reflect pathogenic mutations in humans, we used a combination of an *in silico* analysis (13), databases used for the classification of human MMR variants (ClinVar (14), InSiGHT (15) and COSMIC (16)), and two functional assays (CIMRA (17, 18) and reverse diagnosis catalogues ((19, 20), manuscript in preparation). For *in silico* analyses a probability of pathogenicity of less than 0.1 and more than 0.9 were used for classification as likely pathogenic or benign, respectively (13). Insertions/deletions that caused frameshifts were also classified as pathogenic, while in frame insertion/deletions are categorized as unknown.

Analysis of PhIP-induced toxicity and DDS

mES cells were seeded at a density of $5x10^4$ cells per well of a 6 wells plate, one day before treatment with a dose-range of 0 to 25µM PhIP, as described above. Cells were grown in 50/50 medium for 4 days post-treatment. The number of surviving cells was counted using a Beckman coulter counter.

The induction of DDS by PhIP was assessed as follows: one day prior to treatment with 18µM PhIP or vehicle for one hour, $1x10^6$ wild type and *Msh2*-deficient cells were seeded per well of a 6 wells plate. PhIP/vehicle treatment was performed as described above. After two washes of PBS, cells were grown for another two hours (PhIP) or five hours (PhIP or vehicle) in 50/50 medium before lysates were made in Laemmli sample buffer, gel electrophoresis and W blotting, as described above.

Determination of *Hprt*-mutant frequencies

mES cells were grown in HAT-supplemented 50/50 medium for six days to eliminate pre-existing *Hprt*-mutant cells. Then, $5x10^6$ cells were seeded in p90 dishes, one day prior to PhIP or mock treatment. PhIP (18 µM) or mock treatment was performed as described above. Next, $5x10^6$ cells were seeded in gelatin-coated p90 dishes and grown for six days in 50/50 medium to allow fixation of mutations. Next, $2x10^6$ cells were grown 50/50 medium supplemented with 6tG (30µM) to select *Hprt* deficient clones. In parallel, 3x250 cells were seeded in 50/50 medium in p60 dishes to determine cloning efficiencies. After 7-10 days, clones were stained with methylene blue and the *Hprt*-mutant frequency was calculated by determining the number of 6tG-resistant clones divided by the total number of clone-forming cells seeded in 6tG containing medium.

Next-generation sequencing of *Hprt*-mutant clones

Approximately 400 6tG-resistant clones from PhIP-treated wild type and *Msh2*-deficient mES cell cultures and 400 6tG-resistant clones from mock treated *Msh2* deficient ES cells were used for Next Generation Sequencing (NGS). Per experimental condition clones were pooled and cells were lysed with 1,6ml TRIzol reagent for total RNA isolation. A baseline wild type sequence was established from $8x10^6$ cells, corresponding to the total number of cells of 400 6tG-resistant clones. *Hprt* cDNA was generated using Maxima Reverse Transcriptase (ThermoFisher Scientific) and *Hprt*-specific oligonucleotide primer 21. PCR products for NGS were generated in a two-step procedure. First, three separate PCRs were performed using primer combinations 16/17, 18/19 and 20/21 to amplify three overlapping amplicons of the *Hprt* cDNA. The reaction mixture of 40 µl consisted of Phusion High-Fidelity DNA polymerase (0.4U) (Thermo Fisher Scientific), Phusion polymerase reaction buffer, dNTPs (0.4µM), Forward primer (0.5µM), Reverse primer (0.5µM) and two µl cDNA. Using a thermocycler, PCR products were generated by incubating the reaction mixture for two minutes at 95°C, followed by 15 seconds at 95°C, 30 seconds at 57°C and one minute 72°C for 25 cycles, and a final elongation step of 72°C for five minutes. PCR products were purified using AMPure XP beads (Beckman Coulter) following manufacturer's protocol and eluted in 20µl deionized H_2O. Subsequently, another Phusion PCR was

performed as previously described with the exception of cycling for 8 PCR cycles instead of 25. For this PCR, forward (22-25) and reverse (26-29) primers were used with unique barcodes for each condition (Supplemental Methods table 3). Following purification using AMPure beads, the size of each PCR products was assessed using a Qiaxcel Advanced System (Qiagen). Finally, 50ng pooled PCR products was sequenced using Illumina Paired-End sequencing (GenomeScan).

Analysis of spectra of PhIP-induced mutations in the *Hprt* coding region

To detect alterations in cDNA the paired-end data was first filtered to contain exclusively high-quality reads. Thus, only reads with a maximum error probability of 0.05 were kept. Paired-end reads were merged using Flash (21) and mapped to the *Hprt* reference sequence using in-house software (van Schendel *et al*. manuscript in preparation). Additional filtering was applied to ensure each read started and ended with the primer combinations used (see above). Finally, each mapped read was compared to the reference *Hprt* sequence and annotated into wild type (no difference compared to reference), single nucleotide substitution, multi nucleotide substitution, deletion or insertion. To identify unique mutations from background noise we considered mutations to be real if the allele frequency was > 0.001.

Results

An isogenic set of MMR gene-disrupted mES cells

mES cells are primary, diploid and display strong DNA damage responses, thus providing good models for intestinal stem cells in LS patients. We have previously described the *Msh2*-s, Msh6-1 and *Msh6*-s lines (10-12).

Since the mutator phenotype conferred by MMR deficiency predisposes to genetic drift, we decided to also use independently constructed isogenic mES cell lines, each carrying a targeted disruption of both alleles of one of the four genes MMR (Fig. 1A). These cell lines were generated using CRISPR-Cas9, as described in the Materials and Methods section, resulting in cell lines called *Msh2*-1, *Mlh1*-1, *Mlh1*-s, *Pms2*-1 and *Pms2*-s (Fig. 1B). As models for MMR gene-heterozygous colonic stem cells in LS patients, we employed recently generated mES cell lines, heterozygous for *Msh2* or for *Mlh1*. In these cells, the wild type *Msh2* or *Mlh1* allele was flanked by a *Blasticidin* resistance cassette (M. Drost et al, in preparation; Fig. 2A). This *Blasticidin* resistance cassette enables us to distinguish between clones that have lost the wild type allele by an intragenic mutation, since these retain the linked cassette, and clones that have undergone LOH, which results in concomitant loss of Blasticidin resistance. Thus, these cells allow to sensitively investigate the mechanistic basis of loss of the wild type allele in MMR gene-heterozygous cells.

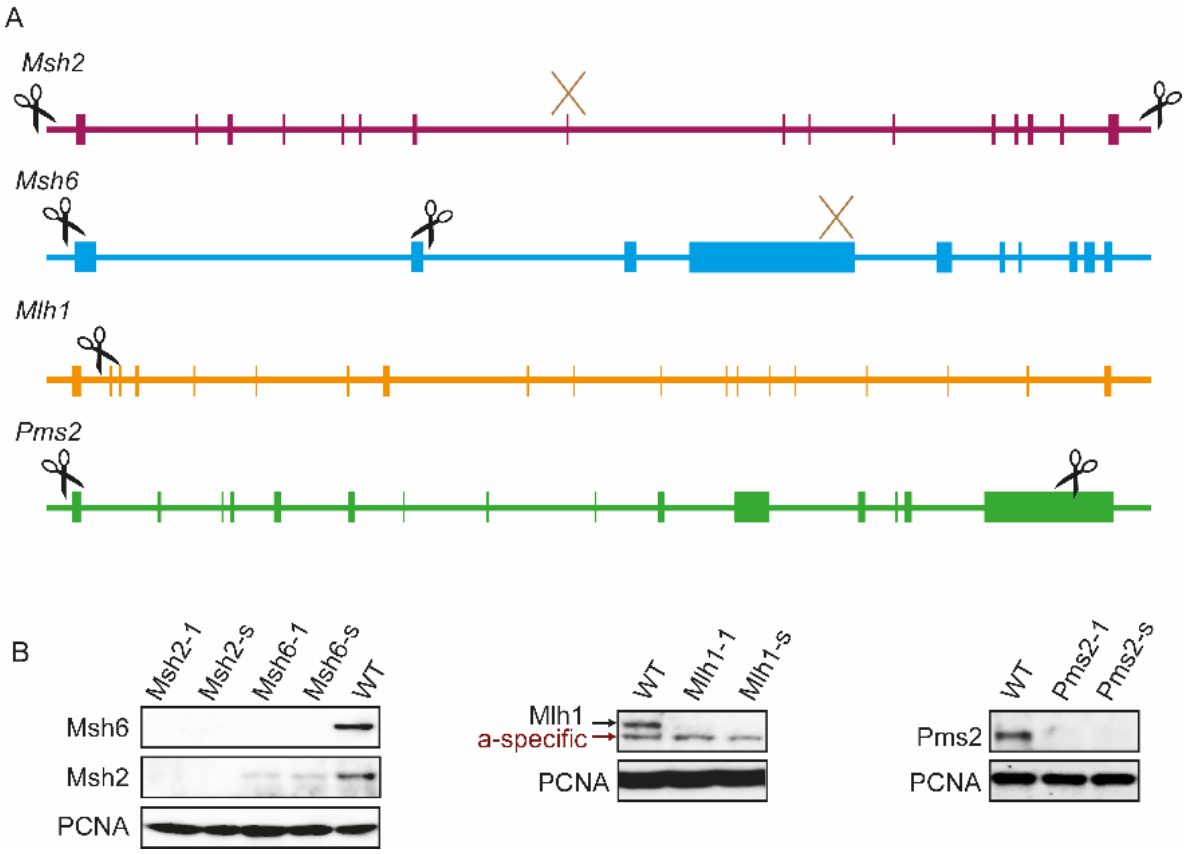


Figure 1. canonical MMR gene alleles used in this study.
(A) Graphical representation of the four core MMR genes. Vertical bars represent exons. Locations where guide RNAs directed a CRISPR-Cas9-induced break are symbolized by scissors. Brown crosses depict the sites where a hygromycin (*Msh2*) or puromycin (*Msh6*) resistance cassette was integrated to disrupt the gene to provide for completely independent knockouts. (B) Western blot validation of mES cell lines deficient for the four MMR genes. PCNA is used as a loading control. The Msh2 and Msh6 panels were derived from the same blot. The Msh2-s and Msh6-s lines were published before (9-11). Of note, some Msh2 protein is visible in the *Msh6* lines, resulting from dimerization of Msh2 with Msh3.

PhIP induces MMR deficiency in MMR gene-heterozygous cells

Loss of the single functional MMR allele resulting from intragenic mutations or LOH is a prerequisite for the development of CRC in LS individuals (22). We investigated whether, and how, PhIP can induce loss of the wild type allele of a heterozygous MMR gene by using the *Blasticidin* cassette-tagged mES cells, heterozygous for *Msh2* or *Mlh1*. We exploited the acquired tolerance of MMR-deficient cells to the nucleoside analog 6-thioguanine (6tG; Fig. 2A, Supplemental Fig. S1) to select individual clones that have lost the wild type allele in a MMR gene-heterozygous mES cell line (9). 6tG selection was followed by HAT counterselection to eliminate cells that inadvertently have acquired 6tG resistance by an inactivating mutation in the *Hprt* gene.

In the absence of PhIP treatment, selection of *Msh2* or *Mlh1*-heterozygous mES cell lines with 6tG yielded clones, suggesting a significant frequency of spontaneous loss of the wild type allele (Fig. 2B). These clones had all become Blasticidin-sensitive, indicating that the wild type allele occasionally is lost by LOH, in the absence of a

mutagen (Fig. 2C). Exposure to PhIP significantly increased the frequency of 6tG-tolerant clones in each heterozygous mES cell line (Fig. 2B-D, Supplemental table S1), suggesting that PhIP induces increased loss of the wild type allele. Of note, the frequency of both spontaneous and PhIP-induced loss of the wild type allele in *Msh2*-heterozygous cells was higher than that in *Mlh1*-heterozygous cells. This was not a consequence of higher 6tG tolerance, and therefore of more efficient selection, of *Msh2*-deficient than of *Mlh1*-deficient cells (Supplemental Fig. S1). Thus, we infer that both spontaneous and PhIP-induced allelic loss is more efficient at *Mhs2* than at *Mlh1*. In each cell line, approximately 75% of the PhIP-induced 6tG-tolerant clones were Blasticidin-sensitive, indicating that PhIP predominantly induces LOH at the wild type allele. Nevertheless, about 25% of the PhIP-induced 6tG-tolerant clones had retained Blasticidin resistance, suggesting that in these clones the wild type allele was lost by a PhIP-induced intragenic mutation. To provide more evidence for the latter, we performed an allele-specific multiplex PCR on isolated clones. Indeed, all Blasticidin-resistant clones had retained one allele of *Msh2* or *Mlh1*, whereas the majority of the Blasticidin-unselected clones had undergone complete loss of the wild type allele, confirming the induction of LOH by PhIP (Supplemental Fig. S2).

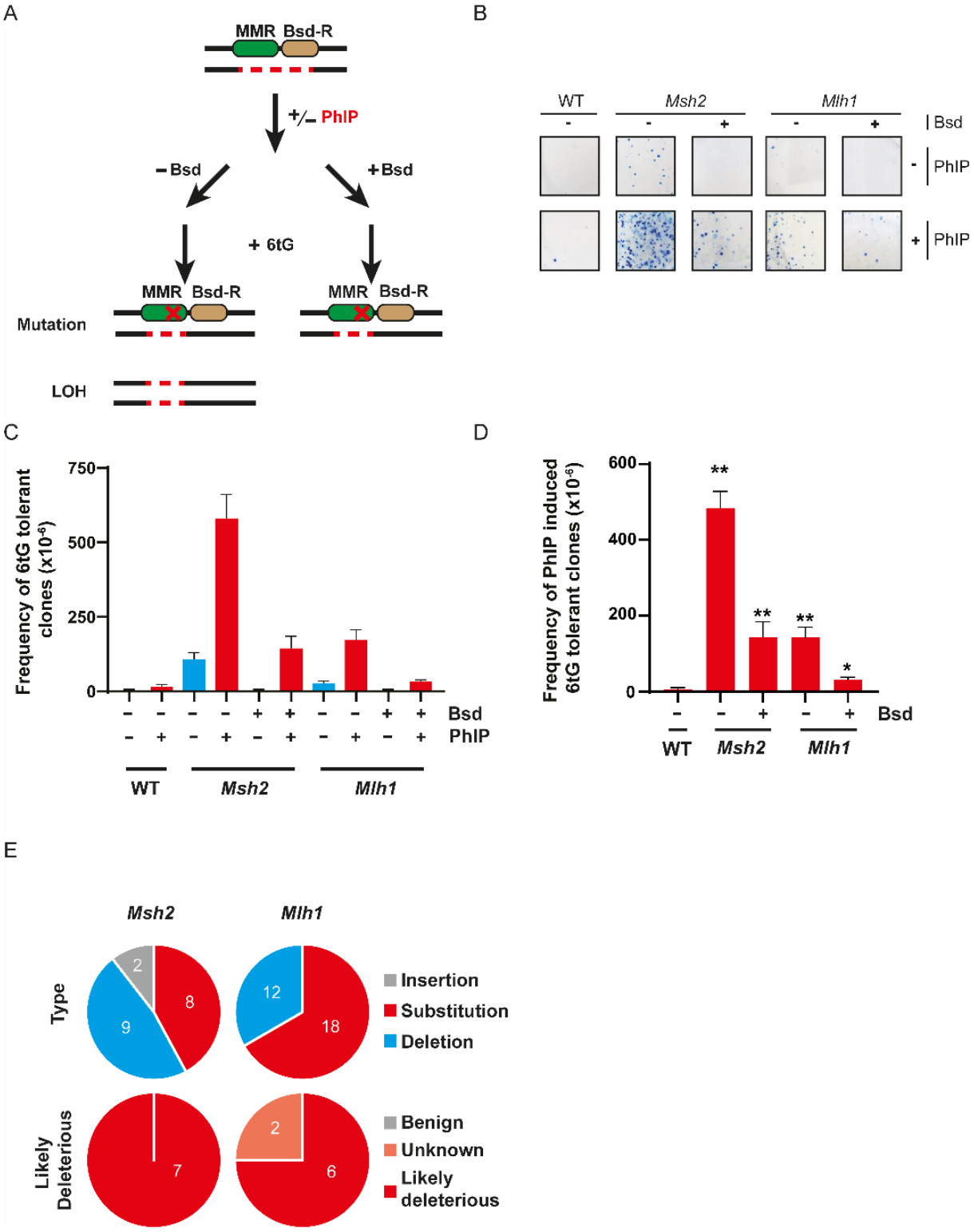

A
MMR Bsd-R
+/– PhIP
– Bsd
+ Bsd
+ 6tG
MMR Bsd-R
MMR Bsd-R
Mutation
LOH
B
WT
Msh2
Mlh1
Bsd
PhIP
C
Frequency of 6tG tolerant clones (x10⁻⁶)
750
500
250
0
Bsd
PhIP
WT
Msh2
Mlh1
D
Frequency of PhIP induced 6tG tolerant clones (x10⁻⁶)
600
400
200
0
Bsd
WT
Msh2
Mlh1
E
Msh2
Mlh1
Type
Likely Deleterious
Insertion
Substitution
Deletion
Benign
Unknown
Likely deleterious

Figure 2. PhIP-induces loss of MMR in MMR gene-heterozygous mES cells.
(A) Selection pipeline to investigate PhIP-induced loss of MMR in MMR gene-heterozygous cells. One allele of either *Msh2* or *Mlh1* was deleted using CRISPR-Cas9 in mES cells (red dashes) whilst the wild type allele was marked at the 3' side with a *Blasticidin* resistance gene (Bsd-R). These cell lines were treated with PhIP to investigate the induction of loss of the wild type allele by either an intragenic mutation (which leads to preservation of Bsd resistance) or by loss of heterozygosity (LOH, which results in concomitant loss of the Bsd-R gene). 6tG was used to select for MMR-deficient cells, with or without concomitant Bsd selection. (B) Wild type, *Msh2* and *Mlh1*-heterozygous cells were exposed to PhIP or vehicle, cultured for a week, and treated with 6tG. Surviving clones were stained with methylene blue. (C) Spontaneous or PhIP-induced MMR-deficient clones (n=3). Half of the plates were also treated with Blasticidin (Bsd) to select for intragenic events. Error bars, SEM. (D) PhIP-induced MMR-deficient clones (after subtraction of spontaneous MMR-deficient clones). Error bars, SEM.*, P ≤ 0,05, **, P ≤ 0,01; unpaired T-test of groups compared to WT. (E) Upper panel: PhIP-induced mutations in *Msh2* and *Mlh1* cDNA. Lower panel: predicted pathogenicity of PhIP-induced substitution mutations. Number, number of clones.

Inactivation of the wild type allele in MMR-heterozygous cells by PhIP-induced mutations

To confirm that in these 6tG-selected, Blasticidin-resistant, clones the wild type *Msh2* or *Mlh1* allele was inactivated by a PhIP-induced intragenic mutation, Sanger sequencing was performed on cDNA. The results showed that for both genes the spectrum of PhIP-induced mutations consisted mostly of single nucleotide substitutions (SNS) and intragenic deletions (Supplemental Tables S2, S3). All deletions started at a guanine and SNS were dominated by G.C > T.A transversions. These data are consistent with the adduct spectrum of PhIP (typically at G nucleotides) and its mutational fingerprint (23, 24). The mutations were distributed over *Msh2* and *Mlh1* (Supplemental Fig. S3). To confirm that the mutations resulted in inactivation of the wild type allele of these genes we used *in silico* analyses, data derived from functional assays, and variant databases. These databases list human variants and thus only mutations of residues conserved between mouse and human DNA could be studied (Supplemental Tables S2, S3). In particular, an SNS was deemed likely pathogenic in case (i) the *in silico* predicted probability of pathogenicity (13) exceeded 0.9, (ii) when MMR functionality was compromised in functional assays such as the biochemical MMR (CIMRA) assay (17, 18), (iii) when the mutation was identified by large-scale genetic screens for deleterious variants ((19, 20), manuscript in preparation), (iv) when identified in LS patients as likely pathogenic or pathogenic in the ClinVar (14) or InSiGHT (15) databases, (v) when listed in the COSMIC cancer somatic database (16). Seven out of 8 and 8 out of 12 SNS in *Msh2* and *Mlh1*, respectively, comprised conserved residues and 7/7 and 6/8, respectively, were predicted pathogenic using the aforementioned analyses (Fig. 2E, Supplemental Tables S2, S3). Only in *Mlh1* two substitutions could not be classified while none of the PhIP-induced substitutions in either gene were predicted to be benign. In addition, almost all intragenic deletions that we identified were either large in size and/or frameshifting and thus also deemed pathogenic. Combined, these data show that MMR gene-heterozygous stem cells lose their wild type MMR gene by spontaneous LOH and, at a high frequency, by PhIP-induced LOH or intragenic mutations.

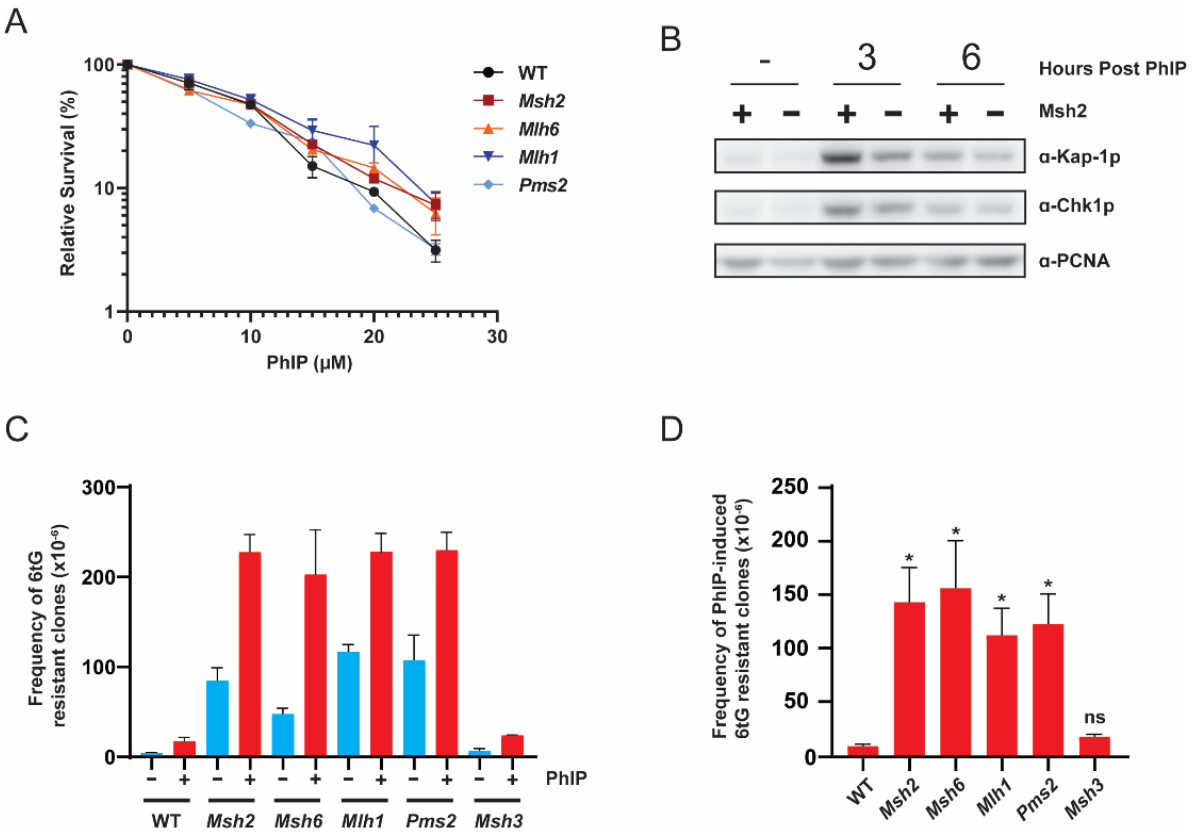


Figure 3. MMR deficient cell lines Msh2-1, Msh6-1, Mlh-1 and Pms2-1 are hypermutable and display reduced DDS following PhIP exposure.
(A) Cell survival after 5 days following exposure to a dose range of PhIP (n=3). Error bars, SEM. (B) Western blot probing for the DDS markers phosphorylated Kap-1 and phosphorylated Chk1. PCNA is used as loading control. Representative image of 3 independent experiments. (C) Frequency of *Hprt* mutants in isogenic MMR-proficient and -deficient mES cell lines following mock or PhIP treatment (n=3). Error bars, SEM. See Supplemental Fig. S4 for experiments using independent MMR gene mutants. (D) Frequency of *Hprt*-mutant clones induced by PhIP treatment (corrected for spontaneous mutants). Error bars, SEM. *, P ≤ 0,05, ns, not statistically significant; unpaired T-test comparing groups to wild type.

MMR deficiency leads to impaired DDS in response to PhIP treatment

To test whether the MMR status affects the toxicity of PhIP, survival of isogenic mES cells, defective for any of the four canonical MMR genes, was determined following exposure to a dose range of PhIP. No significant difference in survival between wild type and MMR-deficient lines *Msh2-1, Msh6-1, Mlh-1* and *Pms2-1* was found (Fig. 3A). Then, we assessed the induction of DDS by PhIP in wild type and *Msh2*-deficient mES cells. This was done by western blotting using antibodies against phosphorylated Chk1 and Kap-1, DDS markers for the formation of single-stranded and double-stranded DNA breaks, respectively. In wild type cells, both markers were increased at 3 and 6 hours post PhIP addition. Deficiency of *Msh2* resulted in a significant mitigation of Kap-1 phosphorylation and a, less pronounced, decrease in Chk1 phosphorylation (Fig. 3B). These data indicate that *Msh2* is involved in provoking DDS induced by PhIP,

which suggests that *Msh2*-deficient stem cells might partially have lost protective checkpoint responses to intestinal mutagens, a result consistent with the previously observed dependence on *Msh2* of DDS and checkpoint responses induced by UV-induced DNA damage (7).

MMR-deficient cells are hypermutable by PhIP

To investigate whether the mutagenicity of PhIP is affected by MMR deficiency we quantified the frequency of mutations induced by PhIP in wild type and the mES cell lines *Msh2-1, Msh6-1, Mlh1-1* and *Pms2-1*, using the *Hprt* gene as a reporter (Fig. 3C-D, and Supplemental Figs. S4A and S4B for independent cell lines). Similar to selection of MMR-deficient clones (see above), selection for mutational inactivation of *Hprt* employs 6tG although, rather than by a pulse, selection for deleterious *Hprt* mutations is continuous and uses a higher dose. Almost no *Hprt*-mutant clones were found in vehicle-treated wild type cells or cells deficient for the minor MMR gene *Msh3*, an alternative binding partner of Msh2 involved in repairing relatively large insertion/deletion loops (25) (Fig. 3C). PhIP treatment of these cell lines yielded a moderate increase of the mutant frequencies ranging between approximately 10 to 20 x 10^{-6} mutants (Fig. 3D). Whereas the MMR-deficient cell lines displayed spontaneous mutant frequencies varying between 45 and 110 x 10^{-6}, higher frequencies of *Hprt*-mutant clones were obtained following PhIP treatment of these cell lines (approximately 220 x 10^{-6}; Fig. 3C). Subtracting the *Hprt* mutant frequencies of the vehicle-treated cells from those of the PhIP-treated cells revealed that PhIP induced much higher frequencies of mutants in all MMR-deficient cells than in wild type or *Msh3*-deficient cells (Fig. 3D). Importantly, by using a completely independent set of isogenic mES cell lines *Msh2-s, Msh6-s, Mlh1-s and Pms2*-s) we obtained identical results (Supplemental Fig. S4), excluding that secondary mutations are responsible in the observed phenotypes. Thus, in addition to suppressing the mutagenicity of spontaneous replication errors, MMR genes suppress the induction of mutations by PhIP. We conclude that the combination of spontaneous and enhanced PhIP-induced mutagenesis in MMR-deficient cells procures a compound hypermutator phenotype.

PhIP-induced mutant spectra in wildtype and MMR-deficient cells

To investigate whether MMR genes, in addition to suppressing PhIP-induced mutations, also affect the PhIP-induced mutation spectrum we treated wild type and *Msh2-1* cells with PhIP or vehicle, selected for *Hprt* mutants, and sequenced pooled *Hprt* cDNAs using an amplicon based next-generation sequencing approach to analyze most of the *Hprt* cDNA. Using a novel sequencing analysis pipeline (R. van Schendel, manuscript in preparation), 58 unique mutations in PhIP treated wild type cells, 49 mutations in mock-treated *Msh2*-deficient cells and 58 mutations in PhIP treated *Msh2*-deficient cells were identified at a 0,1% cut-off for allele frequency (Supplemental tables S4-S6). Only one spontaneous mutation was found in mock-treated wild type samples, confirming minimal sequencing noise.

In both wild type and *Msh2*-deficient cells, 75% of both spontaneous and PhIP-induced mutations were SNS (Fig. 4A). To a lesser extent, *Hprt* was inactivated by multi nucleotide substitutions (MNS, e.g. CGAT > ATTA), insertions and deletions. MNS were exclusively found in cells treated with PhIP, whereas insertions were associated with loss of *Msh2*. By multiplying the fractions of the SNS spectra (Fig. 4A) with the absolute mutant frequencies that were obtained earlier for mock and PhIP-exposed conditions (Fig 3C, 3D) we then obtained the frequency of each type of spontaneous and PhIP-induced nucleotide substitution in both cell lines (Fig. 4B, 4C, Supplemental Table S7). In wild type cells most PhIP-induced SNSs were G.C > T.A transversions, at a frequency of 12.2 x 10^{-6}. In mock-treated *Msh2*-deficient cells, spontaneous SNS consisted mostly of G.C > A.T and A.T > G.C transitions (at a frequency of 25.7 x 10^{-6} each). In *Msh2*-deficient cells treated with PhIP, mutations were mainly comprised of 104.9 x 10^{-6} G.C > T.A transversions, 74.0 x 10^{-6} G.C > A.T transitions and 61.7 x 10^{-6} A.T > G.C transitions. Other substitutions were very rare. Spectra of PhIP-induced SNS were determined by correcting spectra obtained after PhIP treatment for the spontaneous spectra (Fig. 4C). This revealed that *Msh2* specifically suppresses PhIP-induced G.C >T.A transversions, the dominant SNS type also induced in wild type cells. Interestingly, also the frequencies of G.C > A.T and A.T > G.C transitions, that are also part of the spontaneous mutation spectrum in MMR-deficient cells (Fig. 4B), were increased following PhIP treatment. Finally, in both PhIP-treated wild type and *Msh2*-deficient cells there was a strong bias towards nucleotide substitutions derived from guanines in the non-transcribed strand (Supplemental Fig. S5), consistent with the removal of most PhIP-adducted nucleotides from the transcribed strand by transcription-coupled nucleotide excision repair, which precludes their mutagenicity.

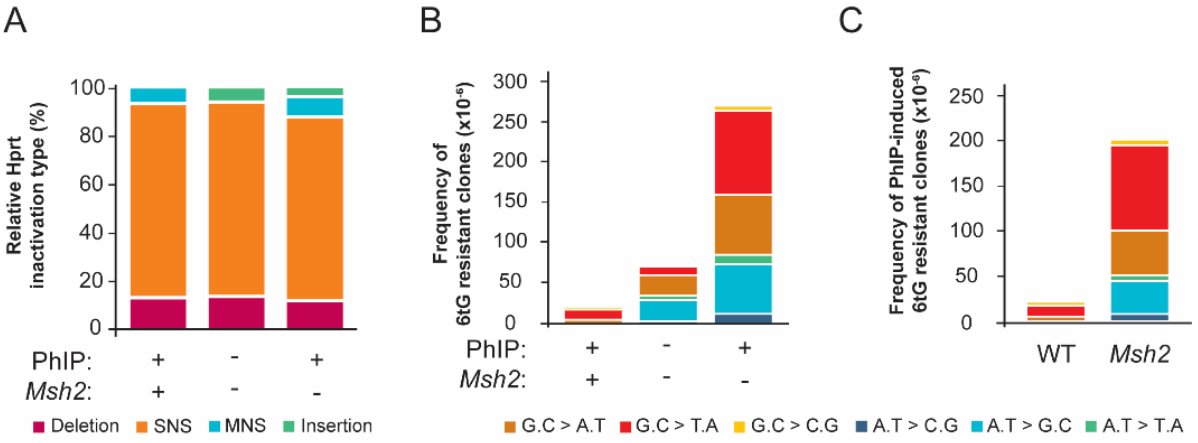


Figure 4. Analysis of PhIP-induced mutations at Hprt in MMR-proficient and -deficient backgrounds.
(A) Percentage of *Hprt* mutants by event type, relative to the total number of unique mutants per condition. SNS, single nucleotide substitution. MNS, multi nucleotide substitution. (B) Distribution of different single nucleotide substitutions. (C) Contribution of different nucleotide substitutions to PhIP-induced *Hprt*-mutant clones.

Discussion

Here we have generated and used isogenic panels of mES cells with targeted heterozygous deficiencies in the MMR genes *Msh2* or *Mlh1* or homozygous deficiencies in *Msh2*, *Msh6*, *Mlh1, Pms2* or *Msh3* as models for intestinal stem cells in LS patients and investigated responses of these cells to the prototypic heterocyclic amine PhIP.

While spontaneous LOH in *Msh2* or *Mlh1*-heterozygous cell lines was readily detectable, in both genotypes frequencies of loss of the wild type allele were strongly increased by a single exposure to PhIP (Fig. 2B-2D). Compared with *Msh2*-heterozygous cells, a lower number of MMR-deficient colonies was obtained following PhIP treatment of *Mlh1*-heterozygous cells, possibly reflecting locus-specific differences in allelic loss. The presence of a *Blasticidin* resistance cassette, linked to the wild type *Msh2* or *Mlh1* allele in these cell lines, allowed us to directly distinguish between either loss of the wild type allele by LOH (26), or by an intragenic deleterious mutation. Indeed, in approximately 25% of all MMR-deficient clones, PhIP treatment had induced inactivation of the wild type allele of *Msh2* or *Mlh1* by an intragenic mutation (Fig. 2D, Supplemental Fig. S2). The spectrum of these intragenic mutations is dominated by G.C > T.A transversions, in line with the spectrum induced by PhIP (23, 24, 27-30). Using a variety of analytical approaches we confirmed that the SNS induced by PhIP in the wild type *Msh2* or *Mlh1* alleles disrupt gene function (Fig. 2E). In support with the involvement of dietary mutagens in the induction of loss of MMR in the intestine of LS patients, our findings reflect observations in MMR-deficient tumors in LS patients where loss of the wild type allele frequently is caused by LOH (31, 32) while MMR inactivation by SNS also occurs albeit a lower frequency (33). Importantly, deleterious SNS might not result in loss of protein expression *per se,* which warrants caution as to the use of immunohistochemical staining for loss of MMR gene expression in CRC as a criterium to screen for MMR gene mutations in individuals suspected of LS (34).

Loss of DDS, that mediate protective checkpoint responses, senescence or apoptosis, is a prerequisite for early steps of carcinogenesis (35). Previous work has shown that loss of *Msh2* results in defective DDS in response to DNA damage induced by methylating agents and UV light (7, 36). We show here that also in response to PhIP, DDS of double-stranded DNA breaks (phospho-Kap-1) or (albeit to a lesser extent) persistent single stranded DNA (phospho-Chk1) partially depend on the MMR status (Fig. 3B). After exposure to a single dose of UV light, MMR-deficient cells displayed enhanced cell cycle progression compared with MMR-proficient cells (7). However, we did not observe a similar response after PhIP exposure (not shown), indicating that cell cycle responses to PhIP may only weakly be dependent on the MMR status. Nevertheless, the observed defective DDS might provide a long-term selective advantage of MMR-deficient over heterozygous intestinal stem cells when chronically exposed to PhIP or other diet-derived genotoxic agents.

Previously, we have found that *Msh2* and *Msh6*-deficient cells are hypermutable by UV light (7, 9), possibly by correcting misincorporations opposite photolesions by mutagenic translesion synthesis polymerases (37). Here we found that also the frequencies of PhIP-induced mutations are significantly higher in cells deficient for any of the four core MMR genes than in wild type cells (Fig. 3D and Supplemental Fig. S3). This exacerbated mutagenicity of helix-distorting nucleotide lesions acts in conjunction with the spontaneous mutagenesis in MMR-deficient cells, resulting in a compound hypermutagenesis phenotype. Importantly, these results were fully reproducible in an independent set of mutant cell lines (Supplemental Fig. S4). Mutational spectra analysis revealed that the frequency of the prevalent SNS induced by PhIP in wild type cells, G.C > T.A transversions, was increased most strongly in the absence of *Msh2* (Fig. 4C). This result suggests that MMR directly suppresses the mutagenicity of PhIP, possibly by removing misincorporations by translesion synthesis opposite the major lesion, dG-C8-PhIP-adducted guanines (30). This would be analogous to the proposed removal of misincorporations generated by translesion synthesis opposite UV light-induced photolesions (7).

G.C > A.T and A.T > G.C transitions dominate the spontaneous mutant spectrum of *Msh2*-deficient cells (Fig. 4B). Surprisingly, although these transitions do not represent typical PhIP-induced substitutions in wild type cells, their frequency was increased by PhIP treatment of *Msh2*-deficient cells (Fig. 4B, 4C). This might reflect the efficient activity of MMR at translesion synthesis-dependent misincorporations opposite PhIP-induced nucleotide adducts. Alternatively, the reduced DDS in the Msh2-deficient cells (Fig. 3B) might enable the survival of cells carrying excessive numbers of misincorporations by translesion synthesis, that otherwise would be eliminated by apoptosis. The wild-type survival of *Msh2*-deficient cells in response to PhIP (Fig. 3A), however, argues against this possibility. Finally, it cannot be excluded that in the absence of *Msh2*, mutagenic translesion synthesis is deregulated, as suggested previously (38), leading to more misincorporations in response to PhIP treatment. Irrespective of the underlying mechanism, these data show that PhIP is significantly more mutagenic in MMR-deficient cells than in MMR-proficient cells.

Taken together, our data support the possibility that dietary mutagens such as PhIP may direct CRC in LS patients at multiple levels (see the Graphical Abstract). First, PhIP exposure increases the frequency of somatic inactivation of the wild type allele in MMR-heterozygous cells by LOH or by a nucleotide substitution, resulting in MMR deficiency (Graphical Abstract). Second, the resulting MMR-deficient cells can no longer efficiently activate protective DDS in response to PhIP (Fig. 3B). Last, exposure of such MMR-deficient cells to PhIP results in compound hypermutagenesis resulting from both the spontaneous mutator phenotype at undamaged nucleotides associated with MMR deficiency (Fig, 3C, blue bars) as well as with the strongly increased mutability of these cells by PhIP as compared with wild type cells (Fig. 3D). Following validation of these observations in *in vivo* models, our findings may provide a rationale for dietary modification as a preventative strategy for LS-associated CRC.

Supplemental Materials

Supplemental Methods table 1: DNA sequences of guide RNAs.

CRISPR table	gRNA 1	gRNA 2	Target	Protein	6tG tolerant
Msh2-S	GATGAGCTACACTAGTGATG	ACACCAAAACCGCTGAGTTG	Entire gene	No protein	Yes
Msh6-1	GGAGCCTCCGCTTCCCGCGG	CCGCGGGAAGCGGAGGCTCC	Exon 1-2	No protein	Yes
Mlh1-1	GCATCAGAGTAGTTGCAA	-	Exon 2	No protein	Yes
Mlh1-S	GCATCAGAGTAGTTGCAA	-	Exon 2	No protein	Yes
Pms2-1	CGGCGCGCTAGACTGGACGAGGG	GTGAAGTCCAGGCGGCAGTTAGG	Entire gene	No protein	No
Pms2-S	TACCTGCACGTGGCCCGCGTCGG	GTGAAGTCCAGGCGGCAGTTAGG	Entire gene	No protein	No

Supplemental Methods table 2: Generation, analysis and validation of cell lines.

mES cell line	**Genotype**	**Origin**	**Year of publication/creation**	**Validation**
E14	Wild type	Ref. 8	1985	2016: Karyotyping
Msh2-Bsd	*Msh2*-heterozygous	Drost et al., in prep.	2016	2016: Fortuitous, 6tG-tolerant, Blasticidin-sensitive progeny clones; allele-specific PCR
Mlh1-Bsd	*Mlh1*-heterozygous	Drost et al., in prep.	2016	2016: Fortuitous, 6tG-tolerant, Blasticidin-sensitive progeny clones; allele-specific PCR
Msh2-1	*Msh2*-deficient	This work	2018	2018: PCR, 6tG tolerance; W blotting
Msh2-s	*Msh2*-deficient	Ref. 11	1995	2018: PCR, 6tG tolerance; W blotting
Msh6-1	*Msh6*-deficient	Ref. 10	2018	2018: 6tG tolerance; W blotting
Msh6-s	*Msh6*-deficient	Ref. 12	1998	2018: 6tG tolerance; W blotting
Mlh1-1	*Mlh1*-deficient	This work	2015	2018: 6tG tolerance; W blotting
Mlh1-s	*Mlh1*-deficient	This work	2015	2018: 6tG tolerance; W blotting
Pms2-1	*Pms2*-deficient	This work	2018	2018: PCR, W blotting
Pms2-s	*Pms2*-deficient	This work	2018	2018: PCR, W blotting
Msh3	*Msh3*-deficient	Ref. 11	1998	2019: PCR

Supplemental Methods table 3: DNA sequences of PCR primers.

Primer:	Sequence
1	CTCCTGAGAGCTGGGGTTTCAAG
2	GGGGAAAAAAGATAGTTCTTTGT
3	CCCTGGGACTCACTACGTTC
4	CCATGCACTGGTGTACGAAG
5	TCTACGAGCGGTTTGTAGGG
6	TTCCCGTCCTTGTTCTCAAC
7	ATGGCGGTGCAGCCTAAGGAG
8	GTTCTTGTTGTTGCGAAGC
9	GATCCTAACCTGAGTGAAC
10	CCCAGCTACAGACGGTAATGT
11	GACGGGCAACTCTGGCG
12	CTGGGCTGTGTCTGAAGTC
13	GTCTACGCTTACCAGATGGT
14	CCAATCCACAGGTGAAGTAC
15	GCACCAATACTGGGATACAG
16	GATGTGTATAAGAGACAGCAGTCCCAGCGTCGTGATTAG
17	CGTGTGCTCTTCCGATCTATCCAGCAGGTCAGCAAAGAA
18	GATGTGTATAAGAGACAGGCCATCACATTGTGGCCCTC
19	CGTGTGCTCTTCCGATCTAGTTTGCATTGTTTTACCAGTGTC
20	GATGTGTATAAGAGACAGTGACACTGGTAAAACAATGCAA
21	CGTGTGCTCTTCCGATCTTTTGCAGATTCAACTTGCGCT
22	AATGATACGGCGACCACCGAGATCTACACGAATTCGCTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
23	AATGATACGGCGACCACCGAGATCTACACGCAAGCATTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
24	AATGATACGGCGACCACCGAGATCTACACGGCATCGATCGTCGGCAGCGTCAGATGTGTATAAGAGACA G
25	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
26	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
27	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
28	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
29	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

Supplemental Figures

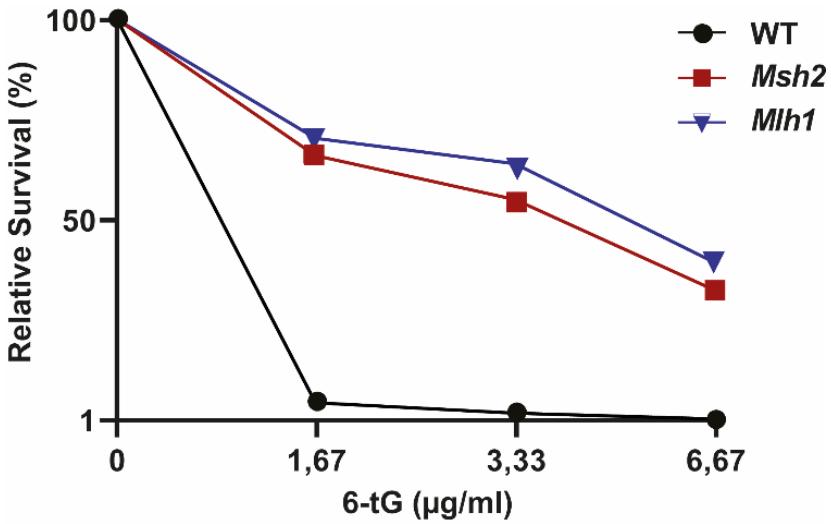


Figure S1: Tolerance of Msh2-1 and Msh6-1 mES cell lines to 6-tG
400 cells per well were seeded in triplicate in 6-well plates (day 1). The indicated concentration of 6-tG was added to the wells for four hours on days 2 and 7. Colonies were stained with methylene blue on day 10 and survival relative to untreated wells was calculated.

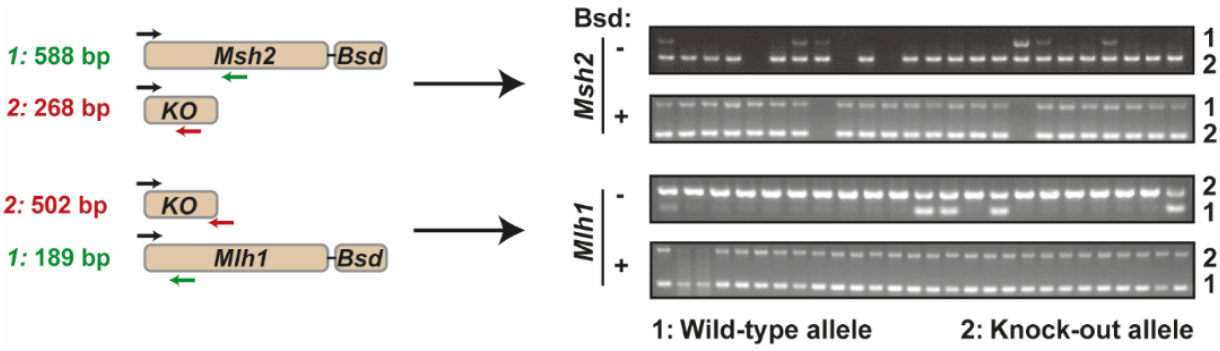


Figure S2: PCR analysis of genomic DNA from 6tG-tolerant clones that were selected with or without Blasticidin.
Primer pair 1, PCR of the wild type allele; primer pair 2, PCR of knockout allele. *Bsd*, Blasticidin. KO: knockout allele of *Msh2 or Mlh1*.

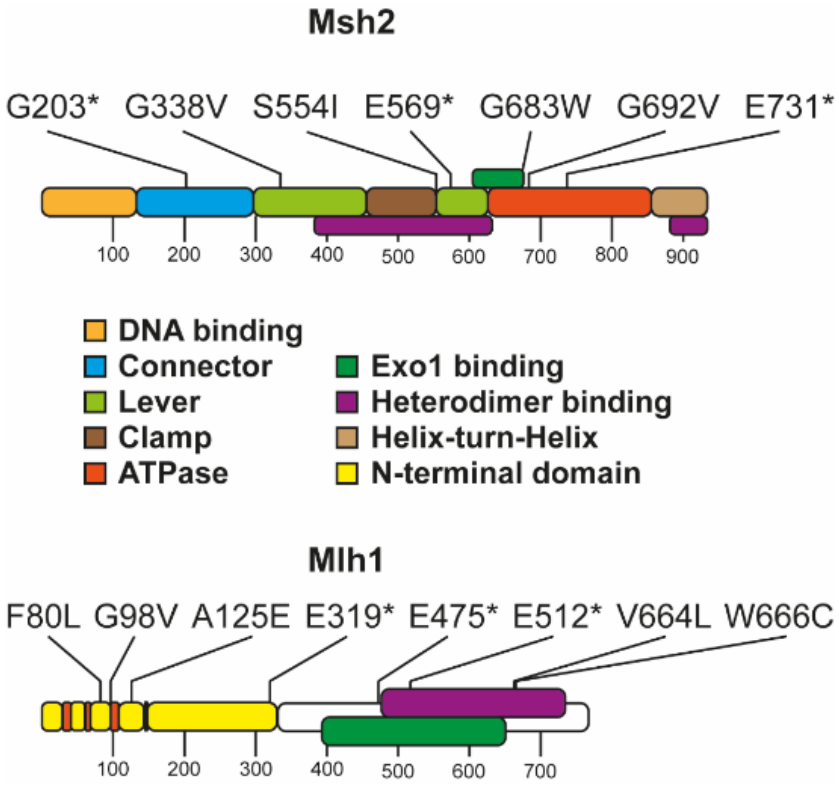


Figure S3: PhIP-induced amino acid changes in Msh2 or Mlh1 in 6tG-tolerant and Blasticidin-resistant clones.
Colors depict functional domains.

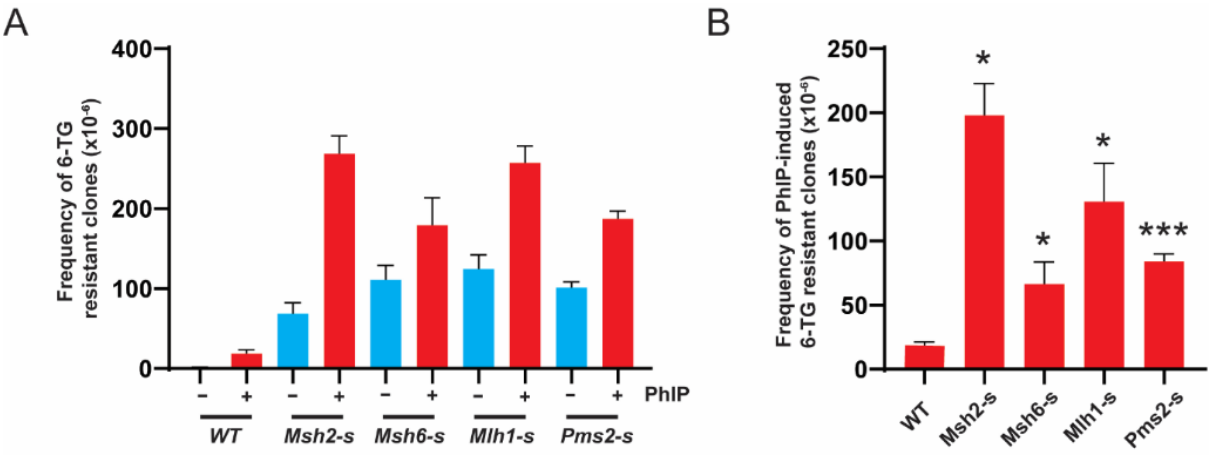


Figure S4: Spontaneous and PhIP-induced mutant frequencies in independent MMR-deficient mES cell lines
A: Cells were exposed to PhIP for 3 hours, cultured for a week in the presence of 6tG to select for clones that lost *Hprt*. Error bars, SEM. -s, independent line from those shown in Fig. 3 B: Quantification of 6tG-resistant clones induced by PhIP, corrected for spontaneous mutants. Error bars, SEM.*, P ≤ 0,05, ***, P ≤ 0,001; unpaired T-test of groups compared to WT.

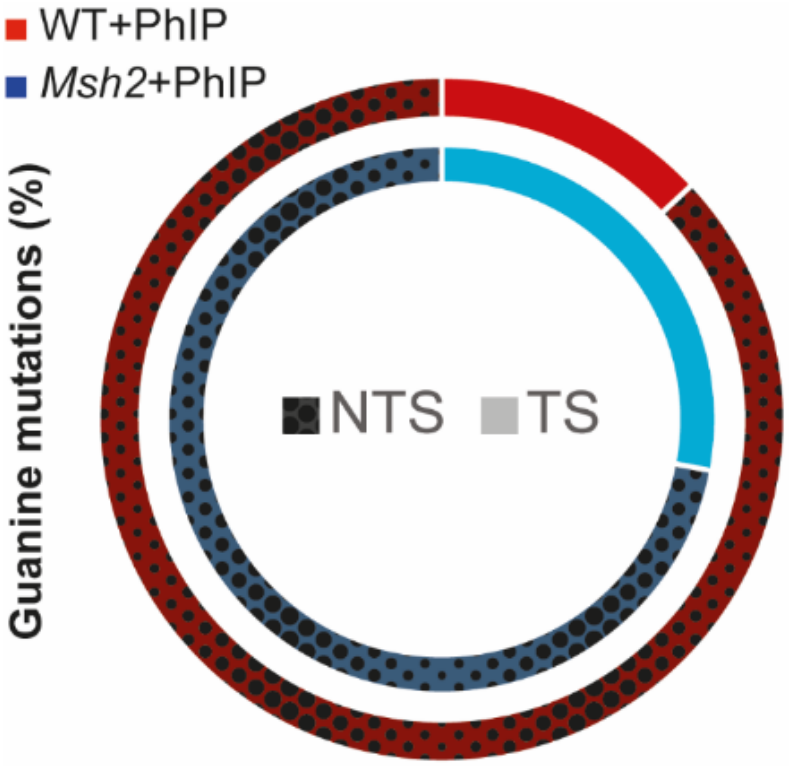


Figure S5: Strand bias of guanine mutations in PhIP-treated wild-type and Msh2-deficient cells.
Dark spotted colors, non-transcribed strand (NTS); bright colors, transcribed strand (TS). Blue, *Msh2*-deficient cells treated without PhIP; red, *Msh2*-deficient cells treated with PhIP.

Supplemental Tables

Table S1: Analysis of MMR inactivation type following mock and PhIP treatment
Cells were exposed to PhIP or vehicle for 3 hours, grown for a week and pulsed with 6tG to select for clones that lost MMR. Half of the plates were also treated with Blasticidin (Bsd) to select for intragenic events opposed to LOH. Clones were stained using methylene blue and quantified. The contribution of intragenic and LOH events was calculated by subtracting the Bsd population (intragenic population) from the total amount of counted colonies. Error = SEM.

	- PhIP			**+ PhIP**			**PhIP-induced**		
	Intragenic	LOH	Total	Intragenic	LOH	Total	Intragenic	LOH	Total
WT	n/a	n/a	**7.7** ± 7.2	n/a	n/	**26.7** ± 22.9	n/a	n/a	**19** ± 15.6
Msh2	**5** ± 1.6	**192.3** ± 77.9	**197.3** ± 78.2	**299.6** ± 119.5	**872.8** ± 243.6	**1172.3** ± 215.9	**294.6** ± 118.0	**680.4** ± 172.4	**975** ± 138.9
Mlh1	**4.2** ± 1.4	**59.2** ± 30.1	**63.3** ± 31.4	**75.2** ± 18.9	**282.8** ± 93.9	**358** ± 99.6	**71.1** ± 18.8	**223.6** ± 66.4	**294.7** ± 68.7

Table S2: Inactivating mutations in Msh2 cDNA following exposure of Msh2-heterozygous mES cells to PhIP

A SNS is deemed likely pathogenic (LP) when the *in silico* prior probability exceeded 0.9 when found as (likely) pathogenic in human MMR variant databases (Insight and ClinVar), or when the functionality of the protein is disturbed in biochemical assays (CIMRA) or *in vivo* assays (Reverse Diagnosis Catalogues). Insertions or deletions are classified as likely pathogenic when these cause frameshifts, otherwise they are unknown.

Mouse				Human		Analysis of Pathogenicity						
cDNA	Protein	Frameshift	Target sequence	cDNA	Protein	Prior probability	Insight	ClinVar	Cosmic	CIMRA	Reverse Diagnosis	Outcome
c. Δ140	p. 47	-1	GCACG[Δ]GAGAGG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. 607 G>T	p. G203*		TACCA[G]GAGGA	c. 607 G>T	p. G203*	N/A	Not found	Not found	Found	Not found	Found	LP
c. Δ213_Δ367	p. Δ71_Δ122	-2	GCAGG[Δ155]CTTCT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. Δ643_789	p. Δ215_263		TGAGG[Δ147]CAGGT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Unknown
c. Δ791_Δ940	p. Δ264_Δ313	-2	GAATC[Δ150]AGGGT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Unknown
c. 1013 G>T	p. G338V		TCAAG[G]ACAAA	c. 1013 G>T	p. G338V	0,94	Not found	Class 3	Found	Not found	Not found	LP
c. Δ1036_Δ1046	p. Δ346_Δ349		AGTGG[Δ11]GCTCA	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. Δ1539	p. 513	-1	AAATT[Δ G]GACTC	c. Δ1539	p. Δ513	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. 1661 G>T	p. S554L		CAACA[G]TGAAT	c. 1661 G>T	p. S554L	0,55	Not found	Not found	Not found	Not found	Found	LP
c.1661 G > 25ins	p. 554 G > ins(9)		AACAA[G]AACTT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. 1663 G>T	p. E555*		ACAGT[G]AATTG	Not conserved	Not conserved	N/A	N/A	N/A	N/A	N/A	N/A	N/A
c. 1705 G>T	p. E569*		AAGGC[G]AGTAT	c. 1705 G>T	p. E569*	N/A	Not found	Not found	Not found	Not found	Found	LP
c. Δ2008_Δ2460	p. Δ670_Δ820		CTGGT[Δ453]GTCTG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Unknown
c. 2047 G>T	p. G683W		AGACC[G]GGGTG	c. 2047 G>T	p. 683 G/W	0,89	Class 3	Not found	Not found	Not found	Found	LP
c. 2075 G>T	p. G692V		AATCG[G]TGTT	c. 2075 G>T	p. 692 G/V	0,938875343	Class 4	Class 4	Not found	5,90%	Not found	LP
c. 2191 G>T	p. E731*		TGCTG[G]AGACT	c. 2191 G>T	p. E731*	N/A	Class 5	Not found	Not found	Not found	Found	LP
c. Δ2583	p. 861	-1	TCGCT[Δ G]GGATG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. Δ2633_2636	p. 877	-2	AAGAG[ΔAAGCA]AGGTG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c.2634_2635ins ATCG	p. 878_879ins(2)		AGAGA[insATCG]CAAGGT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP

Likely deleterious (LD) Class 4/5
Unknown Class 3
Benign Class 1/2
N/A

Table S3: Inactivating mutations in Mlh1 cDNA following exposure of Mlh1-heterozygous mES cells to PhIP

A SNS is deemed likely pathogenic (LP) when the *in silico* prior probability exceeded 0.9 when found as (likely) pathogenic in human MMR variant databases (Insight and ClinVar), or when the functionality of the protein is disturbed in biochemical assays (CIMRA) or *in vivo* assays (Reverse Diagnosis Catalogues). Insertions or deletions are classified as likely pathogenic when these cause frameshifts, otherwise they are unknown.

Mouse				Human		Analysis of Pathogenicity						
cDNA	Protein	Frameshift	Target sequence	cDNA	Protein	Prior probability	Insight	ClinVar	Cosmic	CIMRA	Reverse Diagnosis	Outcome
c. 240 C>A	p. F80L		AGGTT[C]ACTAC	c. 240 C>A	p. F80L	0.70	Not found	Not found	Not found	Not found	Found	LP
c. 293 G>T	p. G98V		CTATG[G]CTTTC	c. 293 G>T	p. G98V	0.94	Not found	Not found	Not found	Not found	Not found	LP
c. 374 C>A	p. A125E		ATGTG[C]GTACA	c. 374 C>A	p. A125E	0.97	Not found	Class 3	Not found	13%	Found	LP
c. Δ453_Δ544	p. Δ151_Δ181	-2	ATCAC[92]GGTAT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. 520 G>T	p. G174*		AGTAC[G]GAAAA	Not Conserved	Not conserved	N/A	N/A	N/A	N/A	N/A	N/A	N/A
c. Δ531_Δ545	Δ177_Δ182		ATTTT[15]GTATT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Unknown
c. Δ589_Δ592	Δ197	-1	AAAAA[CAAG]GTGAG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. 955 G>T	p. E319*		TGCAC[G]AGGAG	c. 955 G>T	p. E319*	N/A	Class 5	Class 5	Not found	Not found	Found	LP
c. 1243 C>T	p. R415*		CTGTC[C]GAGGG	Not Conserved	Not conserved	N/A	N/A	N/A	N/A	N/A	N/A	N/A
c. 1281 G>T	p. T427T		GCCAC[G]CGGGA	Not Conserved	Not conserved	N/A	N/A	N/A	N/A	N/A	N/A	N/A
c. 1435 G>T	p. E479*		ATCGG[G]AGGAC	c. 1423 G>T	p. E475*	N/A	Not found	Not found	Not found	Not found	Not found	Unknown
c. 1546 G>T	p. E516*		AGGAA[G]AGATT	c. 1534 G>T	p. E512*	N/A	Class 5	Class 5	Found	Not found	Not found	LD
c. Δ1597_Δ 1598	p. 532	-2	ACCAT[TC]CTTTG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. Δ1613_Δ1614	p. 537_538	-2	CTGTG[TG]AATCC	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	LP
c. 2002 G>C	p. V668L		CTGAG[G]TGAAT	c. 1990 G>C	p. V664L	0.59	Class 3	Not found	Not found	Not found	Not found	Unknown
c. 2010 G>T	p. W670C		AATTG[G]GATGA	c. 1998 G>T	p. W666C	0.94	Not found	Not found	Not found	Not found	Not found	LP
c. Δ2040_Δ2153	p. Δ680_Δ717	-2	AGTCT[114]GACTG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Unknown
c. 2276 G>T	p. R759L		TGAGC[G]GTGTT	Not Conserved	Not conserved	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Likely deleterious (LD) Class 4/5
Unknown Class 3
Benign Class 1/2
N/A

Table S4: List of mutations in the Hprt gene in PhIP-treated wild type cells.
Flanking sequence; sequence surrounding the mutated bases, parentheses surround the mutated base, non-transcribed strand sequence is shown. Strand; in which strand the mutated guanine is found, the base which is known to mutable by PhIP. NTS, non-transcribed strand. TS, transcribed strand. Ins: inserted nucleotide(s), del: deleted nucleotide(s).

	Position	Exon	Base change	Flanking Sequence	Amino acid change	Strand
SNS	40	2	G > U	GAT(G)AAC	E > *	NTS
	47	2	G > A	CAG(G)TTA	G > D	NTS
	74	2	C > G	TAC(C)TAA	P > R	TS
	88	2	G > U	GCC(G)AGG	E > *	NTS
	97	2	G > U	TTG(G)AAA	E > *	NTS
	113	2	C > U	TTC(C)TCA	P > L	TS
	118	2	G > U	CAT(G)GAC	G > *	NTS
	119	2	G > C	ATG(G)ACT	G > A	NTS
	119	2	G > U	ATG(G)ACT	G > V	NTS
	134	2	G > U	ACA(G)GAC	R > M	NTS
	134	2	G > A	ACA(G)GAC	R > K	NTS
	135	3	G > U	CAG(G)ACT	R > S	NTS
	139	3	G > A	ACT(G)AAA	E > K	NTS
	145	3	C > U	AGA(C)TTG	L > F	TS
	148	3	G > A	CTT(G)CTC	A > T	NTS
	149	3	C > U	TTG(C)TCG	A > V	TS
	172	3	G > U	ATG(G)GAG	G > *	NTS
	208	3	G > U	AAG(G)GGG	G > W	NTS
	208	3	G > C	AAG(G)GGG	G > R	NTS
	209	3	G > A	AGG(G)GGG	G > E	NTS
	211	3	G > C	GGG(G)GCT	G > R	NTS
	212	3	G > A	GGG(G)CTA	G > D	NTS
	212	3	G > U	GGG(G)CTA	G > V	NTS
	222	3	C > A	GTT(C)TTT	F > L	TS
	229	3	G > U	GCT(G)ACC	D > Y	NTS

Table S4: Continued

	Position	Exon	Base change	Flanking Sequence	Amino acid change	Strand
SNS	238	3	G > U	CTG(G)ATT	D > Y	NTS
	463	4	C > A	AGC(C)CCA	P > T	TS
	464	4	C > A	GCC(C)CAA	P > H	TS
	482	4	C > A	TTG(C)AAG	A > E	TS
	500	7	G > U	AAA(G)GAC	R > M	NTS
	527	7	C > A	GGC(C)AGA	P > Q	TS
	539	8	G > U	TTG(G)ATT	G > V	NTS
	539	8	G > A	TTG(G)ATT	G > E	NTS
	551	8	C > A	TTC(C)AGA	P > Q	TS
	565	8	G > C	GTT(G)TTG	V > L	NTS
	568	8	G > U	GTT(G)GAT	G > *	NTS
	568	8	G > C	GTT(G)GAT	G > R	NTS
	569	8	G > U	TTG(G)ATA	G > V	NTS
	569	8	G > A	TTG(G)ATA	G > E	NTS
	580	8	G > U	CTT(G)ACT	D > Y	NTS
	582	8	C > A	TGA(C)TAT	D > E	TS
	599	8	G > A	TCA(G)GGA	R > K	NTS
	600	8	G > C	CAG(G)GAT	R > S	NTS
	601	8	G > U	AGG(G)ATT	D > Y	NTS
	610	9	C > A	AAT(C)ACG	H > N	TS
	612	9	C > G	TCA(C)GTT	H > Q	TS
MNS	152	3	GAGA > TG			
	236	3	TGG > CTGA			
	555	8	CA > G			
	597	8	CA > AT			
Deletion	208	3	del 111bp			
	319	4, 5	del 84bp			
	323	4, 5	del 66bp			
	337	4	del G			
	486	7	del 47bp			
	533	8	del 77bp			
	533	8	del 21bp			
	557	8	del A			

Table S5: List of mutations in the Hprt gene in mock-treated Msh2-deficient cells.
Flanking sequence; sequence surrounding the mutated bases, parentheses surround the mutated base, non-transcribed strand sequence is shown. Strand; in which strand the mutated guanine is found, the base which is known to mutable by PhIP. NTS, non-transcribed strand. TS, transcribed strand. Ins: inserted nucleotide(s), del: deleted nucleotide(s).

	Position	Exon	Base change	Flanking Sequence	Amino acid change
SNS	58	2	G > T	CTA(G)ATT	D > Y
	74	2	C > A	TAC(C)TAA	P > H
	103	2	G > A	AAA(G)TGT	V > M
	104	2	T > C	AAG(T)GTT	V > A
	119	2	G > A	ATG(G)ACT	G > E
	122	2	T > C	GAC(T)GAT	L > P
	130	2	G > T	ATG(G)ACA	D > Y
	131	2	A > G	TGG(A)CAG	D > G
	134	2	G > T	ACA(G)GAC	R > M
	134	2	G > A	ACA(G)GAC	R > K
	151	3	C > T	GCT(C)GAG	R > *
	158	3	T > C	ATG(T)CAT	V > A
	170	3	T > C	AGA(T)GGG	M > T
	172	3	G > T	ATG(G)GAG	G > *
	202	3	C > T	GTG(C)TCA	L > F
	208	3	G > A	AAG(G)GGG	G > R
	209	3	G > A	AGG(G)GGG	G > E
	212	3	G > A	GGG(G)CTA	G > D
	220	3	T > C	AAG(T)TCT	F > L
	233	3	T > C	ACC(T)GCT	L > P
	245	3	T > A	ACA(T)TAA	I > N
	248	3	A > G	TTA(A)AGC	K > R
	254	3	T > C	CAC(T)GAA	L > P
	295	3	T > C	GAT(T)TTA	F > L
	305	3	T > G	GAC(T)GAA	L > R

Table S5: Continued

	Position	Exon	Base change	Flanking Sequence	Amino acid change
SNS	358	4	G > T	GGA(G)ATG	D > Y
	446	4	T > C	CCC(T)GGT	L > P
	464	4	C > T	GCC(C)CAA	P > L
	527	7	C > T	GGC(C)AGA	P > L
	548	8	T > A	AAA(T)TCC	I > N
	550	8	C > T	ATT(C)CAG	P > S
	563	8	T > C	TTG(T)TGT	V > A
	572	8	A > C	GAT(A)TGC	Y > S
	574	8	G > A	TAT(G)CCC	A > T
	577	8	C > T	GCC(C)TTG	L > F
	595	8	T > A	TAC(T)TCA	F > I
	610	9	C > T	AAT(C)ACG	H > Y
	611	9	A > G	ATC(A)CGT	H > R
	614	9	T > C	ACG(T)TTG	V > A
Insertion	103	2	ins A		
	126	2	ins GAT		
	562	8	ins T		
Deletion	323	4, 5	del 66bp		
	345	4	del A		
	486	7	del 47bp		
	533	8	del 77bp		
	533	8	del 21bp		
	585	8	del TA		
	586	8	del AAT		

Table S6: List of mutations in the Hprt gene in PhIP-treated Msh2-deficient cells.
Flanking sequence; sequence surrounding the mutated bases, parentheses surround the mutated base, non-transcribed strand sequence is shown. Strand; in which strand the mutated guanine is found, the base which is known to mutable by PhIP. NTS, non-transcribed strand. TS, transcribed strand. Ins: inserted nucleotide(s), del: deleted nucleotide(s).

	Position	Exon	Base change	Flanking Sequence	Amino acid change	Strand
SNS	47	2	G > A	CAG(G)TTA	G > D	NTS
	47	2	G > T	CAG(G)TTA	G > V	NTS
	119	2	G > A	ATG(G)ACT	G > E	NTS
	122	2	T > C	GAC(T)GAT	L > P	
	130	2	G > C	ATG(G)ACA	D > H	NTS
	134	2	G > A	ACA(G)GAC	R > K	NTS
	135	3	G > T	CAG(G)ACT	R > S	NTS
	143	3	G > T	AAA(G)ACT	R > I	NTS
	149	3	C > T	TTG(C)TCG	A > V	TS
	166	3	G > T	AAG(G)AGA	E > *	NTS
	170	3	T > C	AGA(T)GGG	M > T	
	208	3	G > T	AAG(G)GGG	G > W	NTS
	208	3	G > A	AAG(G)GGG	G > R	NTS
	209	3	G > A	AGG(G)GGG	G > E	NTS
	212	3	G > A	GGG(G)CTA	G > D	NTS
	212	3	G > T	GGG(G)CTA	G > V	NTS
	220	3	T > C	AAG(T)TCT	F > L	
	223	3	T > A	TTC(T)TTG	F > I	
	229	3	G > T	GCT(G)ACC	D > Y	NTS
	233	3	T > C	ACC(T)GCT	L > P	
	238	3	G > T	CTG(G)ATT	D > Y	NTS
	296	3	T > C	ATT(T)TAT	F > S	
	358	4	G > T	GGA(G)ATG	D > Y	NTS
	365	4	T > G	ATC(T)CTC	L > R	

Table S6: Continued

	Position	Exon	Base change	Flanking Sequence	Amino acid change
SNS	446	6	T > C	CCC(T)GGT	L > P
	463	6	C > A	AGC(C)CCA	P > T
	464	6	C > T	GCC(C)CAA	P > L
	472	6	G > T	ATG(G)TTA	V > F
	478	6	G > T	AAG(G)TTG	V > F
	491	7	T > C	TGC(T)GGT	L > P
	495	7	G > A	GGT(G)AAA	V > V
	501	7	G > T	AAG(G)ACC	R > S
	536	8	T > G	TTG(T)TGG	V > G
	539	8	G > T	TTG(G)ATT	G > V
	553	8	G > T	CCA(G)ACA	D > Y
	554	8	A > G	CAG(A)CAA	D > G
	569	8	G > T	TTG(G)ATA	G > V
	574	8	G > A	TAT(G)CCC	A > T
	586	8	A > G	TAT(A)ATG	N > D
	599	8	G > A	TCA(G)GGA	R > K
	601	8	G > A	AGG(G)ATT	D > N
	612	9	C > A	TCA(C)GTT	H > Q
	614	9	T > C	ACG(T)TTG	V > A
	614	9	T > A	ACG(T)TTG	V > D
MNS	112	2	CCT > GCC		
	210	3	GG > AT		
	486	7	CT > AG		
	491	7	TGG > GGA		
	553	8	GAC > TAA		
Insertion	103	2	ins A		
	562	8	ins T		
Deletion	48	2, 3	del 88bp		
	319	4	del 9bp		
	323	4, 5	del 66bp		
	486	7	del 47bp		
	533	8	del 77bp		
	533	8	del 21bp		
	535	8	del G		

Table S7: Frequencies of PhIP-induced single base pair substitutions in wild type and Msh2-deficient backgrounds (corrected for spontaneous substitutions).

WT		**Mock**			**PhIP**			**PhIP-induced**	
		Absolute	Relative (%)	6tG-resistant clones ($x10^{-6}$)	Absolute	Relative	6tG-resistant clones ($x10^{-6}$)	Relative (%)	6tG-resistant clones ($x10^{-6}$)
Transitions	G.C > A.T	0	0	0	12	26,1	5,6	26,1	5,6
	A.T > G.C	0	0	0	0	0,0	0,0	0,0	0,0
Transversions	G.C > T.A	0	0	0	26	56,5	12,2	56,5	12,2
	G.C > C.G	0	0	0	8	17,4	3,8	17,4	3,8
	A.T > T.A	0	0	0	0	0,0	0,0	0,0	0,0
	A.T > C.G	0	0	0	0	0,0	0,0	0,0	0,0
Total		0	0	0	46	100,0	21,6	100,0	21,6

Msh2 $^{-/-}$		**Mock**			**PhIP**			**PhIP induced**	
		Absolute	Relative (%)	6tG-resistant clones ($x10^{-6}$)	Absolute	Relative	6tG-resistant clones ($x10^{-6}$)	Relative (%)	6tG-resistant clones ($x10^{-6}$)
Transitions	G.C > A.T	14	35,9	25,7	12	27,3	74,0	24,2	48,3
	A.T > G.C	14	35,9	25,7	10	22,7	61,7	18,0	36,0
Transversions	G.C > T.A	6	15,4	11,0	17	38,6	104,9	47,0	93,9
	G.C > C.G	0	0,0	0,0	1	2,3	6,2	3,1	6,2
	A.T > T.A	3	7,7	5,5	2	4,5	12,3	3,4	6,8
	A.T > C.G	2	5,1	3,7	2	4,5	12,3	4,3	8,7
Total		39	100,0	71,6	44	100,0	271,5	100,0	199,8

References

1. IARC. Globocan 2018: Cancer Fact Sheets - Colorectal Cancer http://gco.iarc.fr/today/data/factsheets/cancers/10_8_9-Colorectum-fact-sheet.pdf2018.
2. Derry MM, Raina K, Agarwal C, Agarwal R. Identifying Molecular Targets of Lifestyle Modifications in Colon Cancer Prevention. Frontiers in Oncology. 2013;3.
3. Chiavarini M, Bertarelli G, Minelli L, Fabiani R. Dietary Intake of Meat Cooking-Related Mutagens (HCAs) and Risk of Colorectal Adenoma and Cancer: A Systematic Review and Meta-Analysis. Nutrients. 2017;9(5):514.
4. Win AK, Jenkins MA, Dowty JG, Antoniou AC, Lee A, Giles GG, et al. Prevalence and Penetrance of Major Genes and Polygenes for Colorectal Cancer. Cancer Epidemiol Biomarkers Prev. 2017;26(3):404-12.
5. Lynch HT, Snyder CL, Shaw TG, Heinen CD and Hitchins MP Milestones of Lynch syndrome: 1895-2015. Nat. Rev. Cancer 2015;15(3):181-194.
6. Ijsselsteijn R, Jansen JG, De Wind N. DNA mismatch repair-dependent DNA damage responses and cancer. DNA Repair. 2020;93:102923.
7. Tsaalbi-Shtylik A, Ferras C, Pauw B, Hendriks G, Temviriyanukul P, Carlee L, et al. Excision of translesion synthesis errors orchestrates responses to helix-distorting DNA lesions. J Cell Biol. 2015;209(1):33-46.
8. Hooper M, Hardy K, Handyside A, Hunter S, Monk M. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature. 1987;326(6110):292-5.
9. Borgdorff V, van Hees-Stuivenberg S, Meijers CM, de Wind N. Spontaneous and mutagen-induced loss of DNA mismatch repair in Msh2-heterozygous mammalian cells. Mutat Res. 2005;574(1-2):50-7.
10. Van Gool IC, Rayner E, Osse EM, Nout RA, Creutzberg CL, Tomlinson IPM, et al. Adjuvant Treatment for POLE Proofreading Domain-Mutant Cancers: Sensitivity to Radiotherapy, Chemotherapy, and Nucleoside Analogues. Clin Cancer Res. 2018;24(13):3197-203.
11. de Wind N, Dekker M, Berns A, Radman M, te Riele H. Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. Cell. 1995;82(2):321-30.
12. de Wind N, Dekker M, Claij N, Jansen L, van Klink Y, Radman M, et al. HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. Nat Genet. 1999;23(3):359-62.
13. Thompson BA, Greenblatt MS, Vallee MP, Herkert JC, Tessereau C, Young EL, et al. Calibration of multiple in silico tools for predicting pathogenicity of mismatch repair gene missense substitutions. Hum Mutat. 2013;34(1):255-65.
14. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018;46(D1):D1062-D7.
15. Thompson BA, Spurdle AB, Plazzer JP, Greenblatt MS, Akagi K, Al-Mulla F, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. Nat Genet. 2014;46(2):107-15.
16. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 2019;47(D1):D941-D7.
17. Drost M, Zonneveld JBM, Van Hees S, Rasmussen LJ, Hofstra RMW, De Wind N. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Human Mutation. 2012;33(3):488-94.
18. Drost M, Zonneveld JÉBM, Van Dijk L, Morreau H, Tops CM, Vasen HFA, et al. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Human Mutation. 2010;31(3):247-53.
19. Drost M, Lutzen A, van Hees S, Ferreira D, Calleja F, Zonneveld JB, et al. Genetic screens to identify pathogenic gene variants in the common cancer predisposition Lynch syndrome. Proc Natl Acad Sci U S A. 2013;110(23):9403-8.
20. Drost M, Tiersma Y, Glubb D, Kathe S, van Hees S, Calleja F, et al. Two integrated and highly predictive functional analysis-based procedures for the classification of MSH6 variants in Lynch syndrome. Genet Med. 2020;22(5):847-56.
21. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011;27(21):2957-63.
22. Sekine S, Mori T, Ogawa R, Tanaka M, Yoshida H, Taniguchi H, et al. Mismatch repair deficiency commonly precedes adenoma formation in Lynch Syndrome-Associated colorectal tumorigenesis. Modern Pathology. 2017;30(8):1144-51.
23. Zhang S, Lloyd R, Bowden G, Glickman BW, De Boer JG. Msh2 DNA mismatch repair gene deficiency and the food-borne mutagen 2-amino-1-methy1-6-phenolimidazo [4,5-b] pyridine (PhIP) synergistically affect mutagenesis in mouse colon. Oncogene. 2001;20(42):6066-72.

24. Smith-Roe SL, Hegan DC, Glazer PM, Buermeyer AB. Mlh1-dependent suppression of specific mutations induced in vivo by the food-borne carcinogen 2-amino-1-methyl-6-phenylimidazo [4,5-b] pyridine (PhIP). Mutat Res. 2006;594(1-2):101-12.
25. Genschel J, Littman SJ, Drummond JT, Modrich P. Isolation of MutSβ from Human Cells and Comparison of the Mismatch Repair Specificities of MutSβ and MutSα. Journal of Biological Chemistry. 1998;273(31):19895-901.
26. Thiagalingam S, Laken S, Willson JK, Markowitz SD, Kinzler KW, Vogelstein B, et al. Mechanisms underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci U S A. 2001;98(5):2698-702.
27. Lynch AM, Gooderham NJ, Davies DS, Boobis AR. Genetic analysis of PHIP intestinal mutations in MutaTMMouse. Mutagenesis. 1998;13(6):601-5.
28. Stuart GR, Thorleifson E, Okochi E, De Boer JG, Ushijima T, Nagao M, et al. Interpretation of mutational spectra from different genes: analyses of PhIP-induced mutational specificity in the lacI and cII transgenes from colon of Big Blue® rats. 2000;452(1):101-21.
29. Yadollahi-Farsani M, Gooderham NJ, Davies DS, Boobis AR. ACCELERATED PAPER: Mutational spectra of the dietary carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5- b]pyridine (PhIP) at the Chinese hamster hprt locus. 1996;17(4):617-24.
30. Carothers AM, Yuan W, Hingerty BE, Broyde S, Grunberger D, Snyderwine EG. Mutation and repair induced by the carcinogen 2-(hydroxyamino)-1-methyl-6-phenylimidazo[4,5-b]pyridine (N-OH-PhIP) in the dihydrofolate reductase gene of Chinese hamster ovary cells and conformational modeling of the dG-C8-PhIP adduct in DNA. Chem Res Toxicol. 1994;7(2):209-18.
31. Ollikainen M, Hannelius U, Lindgren CM, Abdel-Rahman WM, Kere J, Peltomäki P. Mechanisms of inactivation of MLH1 in hereditary nonpolyposis colorectal carcinoma: a novel approach. Oncogene. 2007;26(31):4541-9.
32. Zhang J, Lindroos A, Ollila S, Russell A, Marra G, Mueller H, et al. Gene conversion is a frequent mechanism of inactivation of the wild-type allele in cancers from MLH1/MSH2 deletion carriers. Cancer Res. 2006;66(2):659-64.
33. Lu SL, Akiyama Y, Nagasaki H, Nomizu T, Ikeda E, Baba S, et al. Loss or somatic mutations of hMSH2 occur in hereditary nonpolyposis colorectal cancers with hMSH2 germline mutations. Jpn J Cancer Res. 1996;87(3):279-87.
34. Shia J. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. The Journal of molecular diagnostics : JMD. 2008;10(4):293-300.
35. Bartkova J, Hořejší Z, Koed K, Krämer A, Tort F, Zieger K, et al. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature. 2005;434(7035):864-70.
36. Jiricny J. The multifaceted mismatch-repair system. Nature reviews Molecular cell biology. 2006;7(5):335-46.
37. Pilzecker B, Buoninfante OA, Jacobs H. DNA damage tolerance in stem cells, ageing, mutagenesis, disease and cancer therapy. Nucleic Acids Research. 2019;47(14):7163-81.
38. Lv L, Wang F, Ma X, Yang Y, Wang Z, Liu H, et al. Mismatch repair protein MSH2 regulates translesion DNA synthesis following exposure of cells to UV radiation. Nucleic Acids Res. 2013;41(22):10312-22.