DNA mismatch repair haploinsufficiency leads to low-level microsatellite instability in normal mouse intestine

Kul Shrestha[1,2], Minna Tuominen[1], Liisa Kauppi*[1,2]

[1]Systems Oncology (ONCOSYS) Research Program, Research Programs Unit, Faculty of Medicine, University of Helsinki, Helsinki, Finland

[2]Department of Biochemistry and Developmental Biology, Faculty of Medicine, University of Helsinki, Helsinki, Finland

*Corresponding author: Liisa.Kauppi@helsinki.fi

Abstract

Tumors in Lynch syndrome (LS) patients display high levels of microsatellite instability (MSI), which results from complete loss of DNA mismatch repair (MMR). Why certain tissues, in particular the gastrointestinal tract, are especially prone to tumorigenesis in LS is not well understood. We hypothesized that MMR haploinsufficiency may compromise microsatellite stability in a tissue-specific manner before tumorigenesis. Using mouse genetics, we tested how levels of MLH1, a central MMR protein, contribute to microsatellite stability *in vivo* and whether elevated MSI is detectable prior to neoplastic growth. We assayed MSI by single-molecule PCR in normal jejunum and spleen of mice with different *Mlh1* gene dosage (*Mlh1*$^{+/+}$, *Mlh1*$^{+/-}$, *Mlh1*$^{-/-}$) at 4 and 12 months, and correlated our findings with Mlh1 mRNA and MLH1 protein expression levels.

While spleen MLH1 levels of *Mlh1*$^{+/-}$ mice were consistently reduced to, as expected, approximately 50% compared to *Mlh1*$^{+/+}$ mice, in jejunum, MLH1 levels varied between individual *Mlh1*$^{+/-}$ mice and decreased with age (2-54% and 1-30% at 4- and 12-months, respectively), which was due to sporadic and progressive *Mlh1* promoter methylation. The degree of MLH1 loss (prior to complete loss of the protein) correlated with MSI severity in *Mlh1*$^{+/-}$ jejunum, while spleens remained microsatellite-stable. Our findings reveal MSI as a common pre-tumorigenic event that precedes (complete) loss of MMR function in *Mlh1*$^{+/-}$ mice. A similar mechanism likely also operates in the GI epithelium of LS patients.

Introduction

Lynch syndrome (LS) is an autosomal-dominant cancer syndrome characterized by high risk of developing colorectal cancer (CRC), endometrial cancer and various other cancers. LS accounts for 1-5% of all CRC cases; individuals with LS have more than 80% lifetime risk of developing CRC, and are diagnosed with CRC at an average age of 44 years. LS patients carry germline heterozygous mutations in one of the DNA mismatch repair (MMR) genes, primarily *MLH1* and *MSH2* [1-3]. MMR is required for the repair of single base pair mismatches and small insertion-deletion (indel) mutations due to strand-slippage DNA replication errors [4, 5]. Microsatellites (short DNA tandem repeats), due to their repetitive nature, are highly prone to such replication errors [4, 6] leading to unrepaired indel mutations in MMR defective cells, a phenotype known as microsatellite instability (MSI) [4, 7]. High-level MSI is a molecular hallmark of LS patients' tumors and reflects defective MMR [8, 9]. The conventional understanding of MSI in LS-associated CRC is that it occurs after the only remaining functional MMR allele is lost (i.e. "second hit") by somatic mutations, epigenetic inactivation or loss of heterozygosity (LOH), leading to complete loss of MMR (MMR deficiency) which initiates the MSI mutator phenotype, and in turn, instigates tumorigenesis [10-12].

Certain observations in LS patients, however, suggests that MMR genes may be haploinsufficient, i.e. loss of function of just one allele may lead to loss of genome integrity. LS patients show low-level of MSI in peripheral blood leukocytes, and in normal colon mucosa [13-15]. Additionally, several *in vivo* and *in vitro* studies suggest MMR haploinsufficiency: *Mlh1*$^{+/-}$ mice show increased tumor incidence and mortality [16], cell-free extracts from *Mlh1*$^{+/-}$ mouse embryonic fibroblasts show decreased MMR activity [7], decreased MMR levels reduce MMR efficiency in non-neoplastic human cell lines [17, 18], and *Mlh1* hemizygosity increases indel mutations in cancer cell lines [19]. Given the clues of MMR haploinsufficiency, we aimed to understand the tissue-specificity of this phenomenon prior to tumorigenesis. We hypothesized that reduced levels of MMR protein may have more profound effects on microsatellite stability in certain tissues, in particular highly proliferating intestine and colon. To test this, we determined MLH1 levels and tissue-specific appearance of MSI in *Mlh1*$^{+/-}$ mice [7, 20].

The MSI mutator phenotype and MMR-deficiency associated tumor spectrum has been extensively studied and characterized in *Mlh1*$^{-/-}$ mice; *Mlh1*$^{-/-}$ mice have high incidence of MMR-deficient MSI-high GI tumors [7, 16, 20, 21]. Here, we used *Mlh1*$^{+/-}$ mice as model of LS to study the tissue-specific MMR haploinsufficiency. As in LS, *Mlh1*$^{+/-}$ mice have early-onset of GI tract tumors, and increased mortality [16]. As strand slippage during DNA replication is the primary source of endogenous MSI,

highly proliferating tissues should have higher chance of accumulating MSI. To obtain an estimate of possible tissue-specific MMR haploinsufficiency, we assayed MSI and MLH1 levels in tissues with different proliferation rates: highly proliferative small intestine [22], and less proliferating spleen [23]. Additionally, we tested loss of heterozygosity (LOH) and assayed *Mlh1* promoter methylation to get the mechanistic insights to MLH1 protein expression regulation.

Methods

Mice and genotyping

Mlh1 mice (B6.129- *Mlh1*^{tm1Rak}, strain 01XA2, National Institutes of Health, Mouse Repository, NCI-Frederick) [7] were bred and maintained according to national and institutional guidelines (Animal Experiment Board in Finland and Laboratory Animal Centre of the University of Helsinki). *Mlh1* genotyping was performed using ear pieces. DNA lysate was prepared by incubating the ear pieces overnight in 100 µl of lysis buffer supplemented with µg/ml proteinase K, next day boiling the lysate for 10 minutes and spinning down for 1 minute at 14,000 rpm at room temperature. 0.5 µl of DNA lysate was used for genotyping. Genotyping PCR was performed using Platinum green hot start PCR master mix (Invitrogen, Carlsbad, CA). Lysis buffer components, PCR program and primers [24] used for genotyping are listed in **Supplementary table 1.**

Tissue collection for DNA, RNA and protein analysis

The small intestine was flushed with cold phosphate-buffered saline (PBS), cut open longitudinally and visually inspected for visible tumors. A 3 cm piece of jejunum was cut approximately 15 cm from the pyloric sphincter. This tissue piece, henceforth referred to simply as “jejunum”, was approximately the center of the small intestine. Jejunum was inspected for any macroscopic tumor-like outgrowth under a stereoscope, and only normal tissue pieces were used for the experiments. The jejunum was snap-frozen and stored at - 80°C until further use. Other tissues (brain, kidney, liver and spleen) and *Mlh1*$^{-/-}$ intestinal tumors were also collected, snap-frozen and stored at - 80°C until further use. The frozen tissues were used for subsequent DNA, RNA and protein analysis.

DNA extraction and single-molecule MSI analysis by PCR

DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to manufacturer’s instruction. Approximately 5 mg of tissue was used for DNA extraction, the extracted DNA was quantified using Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA). The MSI

assay was performed using single-molecule PCR (SM-PCR) [25-27]. Three microsatellites, two mononucleotide repeats A27 and A33 [28], and a dinucleotide repeat D14Mit15 [7] were assayed for MSI. The number of amplifiable DNA molecules per reaction for each microsatellite was calculated using Poisson analysis of number of positive and negative reactions [29]. SM-PCR programs and primers [7, 28] are listed in **Supplementary table 1**. D14Mit15 and A33 were assayed in the same PCR reaction, and a separate PCR was run for A27. Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA) was used for SM-PCR. Fragment analysis was performed by combining 1µl of each PCR product, with an internal size standard (GeneScan™ 500 LIZ™ dye Size Standard, Applied Biosystems, Waltham, MA) on an ABI3730xl DNA Analyzer (Thermo Fisher Scientific). Between 116 and 205 amplifiable DNA molecules per tissue per mouse were assayed for MSI. Fragman R package was used for visualizing the electropherograms and for mutant scoring [30]. We used stringent mutant scoring criteria (adopted from [26, 31]) and thus, the mutation rates reported are likely a conservative estimate. Criteria for scoring microsatellite mutations were as follows:

I. A true microsatellite signal should have stutter peaks of lower intensities. Stutter peaks should have appropriate size difference (1 base for mononucleotide repeats, and 2 bases for dinucleotide repeats) from the dominant peak. Peaks without stutter were considered artifacts.
II. In order for an allele to be considered as mutant, both the highest peak and the stutter peaks should shift as a single unit. Shift of the highest peak alone was not considered as mutant.
III. If a wild-type and a mutant allele co-occurred in a single PCR, the reaction was scored as wild-type. The non wild-type peaks were presumed to be result of an artifact due to replication slippage during the early rounds of PCR.

For each microsatellite, MSI was calculated separately for insertions and deletions, from here onwards termed deletion% and insertion%:

deletion% = (total no. of DNA molecules with deletion mutation observed / total DNA molecules analyzed) x 100

insertion% = (total no. of DNA molecules with insertion mutation observed / total DNA molecules analyzed) x 100

RNA extraction and RNA expression analysis

RNA was extracted using AllPrep DNA/RNA/Protein Mini Kit (Qiagen) according to manufacturer's instructions. Approximately 5 mg of tissue was used for RNA extraction. Extracted RNA was measured using NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). 500

ng of extracted RNA was reverse transcribed using SuperScript VILO cDNA Synthesis Kit (Invitrogen). Quantitative PCR (qPCR) was performed using SsoAdvance Universal SYBR Green Supermix system (Bio-Rad, Hercules, CA). PCR programs and primers used in qPCR can be found in supplementary table 1. Gene expression data for Mlh1 was normalized to beta-actin. Data was analyzed using Bio-Rad CFX Maestro (version 1.1).

Western blot analysis

Frozen tissue pieces were thawed on ice, mechanically homogenized in RIPA buffer supplemented with protease inhibitor cocktail (Roche, Basel, Switzerland), incubated for 30 minutes on ice and centrifuged at 14000g for 10 minutes at 4°C. The supernatant was collected and frozen until further use. Total protein concentration was measured using Pierce™ BCA™ protein assay kit (Thermo Fisher Scientific), 30 µg of total protein extract was used for western blotting. The denatured protein was run in 4–20% Mini-Protean TGX gels (Bio-Rad) and transferred to 0.2 µm nitrocellulose membrane using Trans-Blot Transfer Pack (Bio-Rad). To confirm complete protein transfer, membranes were stained with Ponceau solution for 5 minutes at room temperature. The membrane was blocked with 5% milk for 1 hour at room temperature, and incubated overnight at 4°C with primary antibodies against MLH1 and Beta-actin, followed the next day by IRDye 800CW and IRDye 680RD secondary antibody incubation for 1 hour at room temperature. Details of antibodies and their dilutions are listed in **Supplementary table 1**. LI-cor Odyssey FC system (LI-COR, Nebraska, USA) was used to scan the membranes and LI-cor Image Studio lite (version 5.2) was used for image analysis. MLH1 protein signal intensity was normalized to Beta-actin signal intensity.

Immunohistochemistry (IHC) and histological image analysis

The small intestine was flushed with cold PBS, fixed overnight in 4% paraformaldehyde, cut open longitudinally and embedded into paraffin blocks as a “swiss roll” [32], and sectioned at 4µm. Antigen retrieval was performed with 10mM citrate buffer (pH 6). MLH1 was detected using rabbit monoclonal antibody ab92312 (clone EPR3894, Abcam, Cambridge, UK, diluted 1:1500 dilution) and visualized using BrightVision Poly HRP goat anti-rabbit IgG (cat. no. DPVR55HRP, ImmunoLogic, Duiven, The Netherlands). Sections were counterstained with hematoxylin-eosin. Slides were scanned using 3DHistech Panoramic 250 FLASH II digital slide scanner (3DHistech, Budapest, Hungary), and images were visualized using CaseViewer (version 2.2). For each mouse, MLH1 protein levels were quantified for five randomly selected sites in the jejunum, each site consisting of approximately five villus-crypt units. Quantification was performed using IHC Profiler plugin in ImageJ [33].

Mlh1 LOH and promoter methylation analysis by PCR

LOH and methylation analysis was performed on the same extracted genomic DNA used for MSI analysis. For LOH testing, 50 ng of DNA was used for PCR, and PCR reactions was performed using the genotyping PCR conditions and primers. Methylation analysis for *Mlh1* was performed using methylation-specific PCR (MSP) [34]. For MSP, 500 ng of DNA was bisulfite converted using EZ DNA Methylation-Direct Kit (Zymo Research, Irvine, CA) according to manufacturer's instruction. 1µl bisulfite converted DNA was used for PCR, which was performed using 2xZymo Taq premix. Details of the PCR program and primers [35] are in **supplementary table 1**. For both LOH testing and methylation analysis, the PCR products were run in 1.5% agarose gel and visualized using ethidium bromide.

Statistical analysis

Deletion% and insertion% were tested using an unpaired t-test, and linear regression fitting was tested using F-test. Two-tailed p-values < 0.05 were considered statistically significant. P-values are either stated as numerical values or indicated as follows: p-value ≥ 0.05 = not significant (ns), p-value < 0.05 = *, p-value <0.01 = **, and p-value < 0.005 = ***.

Results

Normal jejunum of Mlh1$^{+/-}$ mice show low-level MSI

To investigate tissue- and age-specific appearance of MSI in *Mlh1*$^{+/-}$ mice, we performed single-molecule MSI analysis by SM-PCR in jejunum and spleen of 4- and 12-month-old *Mlh1*$^{+/-}$ mice (**Fig. 1A**). We used three microsatellites for the MSI assay: A27, A33, and D14Mit15. A27 is an intergenic microsatellite approximately 2 kb downstream to *Epas1* gene, and A33 is an intronic microsatellite within *Epas1* gene. D14Mit15 is an intergenic microsatellite at 40 kb distance from *Ptpn20* gene. We separately scored MSI as deletion% and insertion%.

Compared to age-matched *Mlh1*$^{+/+}$ jejuna, in *Mlh1*$^{+/-}$ jejuna we observed an increase in deletions at mononucleotide repeats, while the dinucleotide repeat D14Mt15 was stable (**Fig. 2A, Supplementary fig. 1A and Supplementary fig. 1B**). Deletions observed at the mononucleotide repeats in *Mlh1*$^{+/-}$ jejuna were predominantly single repeat unit, i.e. 1 bp shifts (**Fig. 2A**), and the average deletion% was 6% and 7% at 4 and 12 months, respectively (**Fig. 2B**). Compared to *Mlh1*$^{+/+}$

jejuna, the deletion% in *Mlh1*$^{+/-}$ jejuna were 3.2- and 3.5-fold higher, and compared to *Mlh1*$^{-/-}$ jejuna 5.2- and 6.4- fold lower at 4 and 12 months, respectively. Unlike *Mlh1*$^{+/-}$ jejuna, *Mlh1*$^{+/-}$ spleens were microsatellite stable at both time points (**Fig. 2A and Fig. 2B**).

In contrast to *Mlh1*$^{+/-}$ tissues, in *Mlh1*$^{-/-}$ tissues, deletion alleles had bigger size-shifts and were more frequent at all the three microsatellites at both time points (**Fig. 2A, Supplementary fig. 1A)**. *Mlh1*$^{-/-}$ jejuna showed bigger size-shifts and had more frequency of deletion alleles than *Mlh1*$^{-/-}$ spleens, which further increased (in both deletion size-shifts and frequency) in *Mlh1*$^{-/-}$ intestinal tumor (**Fig. 2A, Supplementary fig. 1A**). In *Mlh1*$^{-/-}$ jejuna, the average deletion% at mononucleotide repeats were 32% and 46%, and in *Mlh1*$^{-/-}$ spleen 21% and 32% at 4- and 12-month time points, respectively (**Fig. 2B**). For comparison, an *Mlh1*$^{-/-}$ intestinal tumor showed 52% deletion (**Fig. 2B**). The deletion% at D14Mit15 was 9% and 22% in *Mlh1*$^{-/-}$ jejuna, and 3% and 6% *Mlh1*$^{-/-}$ spleen at 4- and 12- month time points, respectively, and 24% in the *Mlh1*$^{-/-}$ intestinal tumor. Both the *Mlh1*$^{+/+}$ tissues showed 2% deletion at mononucleotide repeats and no MSI at D14Mit15 at both the time points (**Fig. 2B**).

Unlike deletions, insertions at all three microsatellites were not influenced by *Mlh1* gene dosage, i.e. *Mlh1*$^{+/+}$, *Mlh1*$^{+/-}$, *Mlh1*$^{-/-}$ mice showed minimal or no difference in insertion% in both jejunum and spleen at both time points (except at A33, at the 12-month time point) (**Fig. 2, Supplementary fig. 1A, Supplementary fig. 1B, Supplementary fig. 2**). At 12 months, A33 showed a decreasing trend in insertion% with decreased *Mlh1* gene dosage in both tissues (**Supplementary fig. 2**). Irrespective of genotype, tissue-type or age, all observed insertions were almost exclusively single repeat unit in size (1 bp and 2 bp for mono- and dinucleotide repeats, respectively) (**Fig. 2A and Supplementary fig. 1A**).

One 4-month-old *Mlh1*$^{+/-}$ mouse showed high deletion% in jejunum (25% and 4% at mononucleotide markers and at D14Mit15, respectively) (indicated by an arrow in **Fig. 2B**). With Grubbs' test, this *Mlh1*$^{+/-}$ mouse was classified as an outlier ($p < 0.05$). We also detected a slight increase in deletion% at mononucleotide repeats (5%) in the spleen of this mouse, while D14Mit15 was stable**.** As in other *Mlh1*$^{+/-}$ mice, deletions observed in this mouse were predominantly single repeat unit in size (**Supplementary fig. 3)**. In addition to the extreme case of the outlier mouse, we also generally observed inter-individual variation in deletion% at mononucleotide repeats in jejunum between *Mlh1*$^{+/-}$ mice within both age groups (**Fig. 2B**).

Mlh1$^{+/-}$ mice show sporadic decrease in Mlh1 expression levels in jejunum

To assess whether the inter-individual variation in deletion% in *Mlh1$^{+/-}$* jejuna can be explained by *Mlh1* expression levels, we quantified Mlh1 mRNA and MLH1 protein levels in jejunum of 4- and 12-month time points *Mlh1$^{+/-}$* mice; we used *Mlh1$^{+/-}$* spleens, which were stable for deletions (**Fig. 2B**), as control tissue. To further test tissue-specificity of the *Mlh1* expression, we analyzed Mlh1 mRNA expression in brain, kidney and liver. We also performed immunohistochemistry (IHC) in jejunum of 4-month old mice as an alternative approach to assess variation in MLH1 expression.

In *Mlh1$^{+/-}$* jejuna, the average *Mlh1* expression (both at mRNA and protein levels) was less than the expected approximately 50% compared to *Mlh1$^{+/+}$* jejuna. Expression further decreased with age, indicating age-dependent mechanism of MLH1 depletion; additionally, expression levels varied between individual *Mlh1$^{+/-}$* mice; 4-month old mice showed more variation than 12-month mice. At 4 months, *Mlh1$^{+/-}$* jejuna showed in average 26% Mlh1 mRNA (range: 8%-46%) and 24% (range: 1%-54%) MLH1 protein expression compared to *Mlh1$^{+/+}$* jejuna, and at 12 months the Mlh1 mRNA and MLH1 protein expression decreased in average to 15% (range: 9%-23%) and 17% (range: 1%-29%), respectively **(Fig. 3A**). However, these age-specific decreases in Mlh1 mRNA and MLH1 protein were not statistically significant. Irrespective of age, we observed approximately expected 50% Mlh1 mRNA and MLH1 protein levels in all the *Mlh1$^{+/-}$* spleens (**Fig. 3A and Fig. 3C)**. Jejunum of the outlier *Mlh1$^{+/-}$* mouse showed only 8% Mlh1 mRNA expression (**Fig. 3A**) and no detectable MLH1 protein (**Fig. 3C**). Spleen of this mouse showed expected levels of Mlh1 mRNA and MLH1 protein (**Fig. 3A and Fig. 3C**).

We also assayed Mlh1 mRNA expression in brain, kidney and liver of two 4 months *Mlh1$^{+/-}$* mice expressing below age-average Mlh1 mRNA levels in jejunum; Mlh1 mRNA expression in these tissues was not lowered than expected (60%, 56% and 56%, respectively). (**Supplementary fig. 4**).

Further, using IHC, we also observed inter-individual variation in MLH1 protein expression in *Mlh1$^{+/-}$* jejuna. Based on MLH1 staining intensity, jejuna of half the *Mlh1$^{+/-}$* mice (n=6) were scored as positive and the other half was scored as low positive by IHC profiler (**Fig. 3D**). *Mlh1$^{+/-}$* jejuna scored as low positive had low MLH1 staining intensity though out the tissue with minimal variation (see SD in stacked bar in **Fig. 3D**) indicating tissue-wide lower expression of MLH1, additionally we also observed an increase in the number of MLH1-negative cells in jejunum of these mice (stack-bars

in **Fig. 3D**). Upon further analyzing crypts and villi separately, we observed decreased MLH1 staining intensity in both crypt and villi in jejuna scored as low positive, additionally, all MLH1-negative cells were located in villi **(Supplementary fig. 5)**.

Sporadic depletion of MLH1 in Mlh1$^{+/-}$ jejuna is primarily due to Mlh1 promoter methylation and not LOH

We observed variable MLH1 protein levels in jejunum of *Mlh1*$^{+/-}$ mice ranging from approximately expected 50% (compared to *Mlh1*$^{+/+}$ mice) to no expression (**Fig. 3C**). To investigate the underlying mechanism of this sporadic MLH1 depletion, we performed LOH assay (to test the genetic mechanism (loss of heterozygosity)) and MSP assay (to test the epigenetic mechanism (methylation of *Mlh1* promoter)).

Decrease in MLH1 protein levels in normal jejunum of *Mlh1*$^{+/-}$ mice was primarily due *Mlh1* promoter methylation. Summary of the *Mlh1* LOH and promoter methylation assay can be seen in **Table 1**, over all, the number of *Mlh1*$^{+/-}$ mice with *Mlh1* promoter methylation in jejunum increased with age (**Table 1**). Irrelevant of age, all the *Mlh1*$^{+/-}$ mice (n=4 and n=2 for 4 and 12 months respectively) with extremely low (less than 3%) MLH1 protein expression in jejunum had *Mlh1* promoter methylation (**Fig. 5B**). Additionally, 3 out of 4, 12 months old *Mlh1*$^{+/-}$ mice with higher level of MLH1 expression (range: 18%-29%) also had *Mlh1* promoter methylation (**Fig. 5B**), while none of the 4 month *Mlh1*$^{+/-}$ mice with higher MLH1 levels (range: 18%-52%) (n=4) had *Mlh1* promoter methylation. LOH was only observed in the outlier 4-month old *Mlh1*$^{+/-}$ mice which also had *Mlh1* promoter methylation (**Fig. 5B**), all the other 4- and 12-months *Mlh1*$^{+/-}$ mice assayed (n=22) still had the wild-type allele (i.e. no LOH had occurred) (**Table 1**).

MSI correlates with MLH1 protein level in jejunum of Mlh1$^{+/-}$ mice

In order to establish a tissue- and age specific relationship between deletion% at mononucleotide repeats and MLH1 protein expression we performed linear regression analysis.

In *Mlh1*$^{+/-}$ jejuna, deletion% highly correlated with MLH1 protein expression at both time points (F test: R^2 = 0.87 and 0.80, F test: p-value = 1.9×10^{-6} and 2.9×10^{-4} for 4 and 12 months, respectively)

(**Fig. 5A**). Further, based on the correlation, *Mlh1*$^{+/-}$ mice could be divided into two sub-groups. The first sub-group showed close to the expected 50% of wild type MLH1 protein levels at 4 months (41%), which decreased to 24% at 12 months. This sub-group showed 5% and 7% deletions at the 4- and 12-month time points, respectively. The second sub-group, at both time points, had very low MLH1 protein levels (2% and 1% of wild type expression at 4 and 12 months, respectively) and showed 10% deletions at both time points (**Fig. 5B**). Compared to *Mlh1*$^{+/+}$ mice, both *Mlh1*$^{+/-}$ sub-groups had higher deletion% at mononucleotide repeats in jejunum at both 4 and 12 months (**Fig. 5B**). Compared to *Mlh1*$^{+/+}$ jejuna, the deletion% increased by 2.5- and 5-fold in jejunum of *Mlh1*$^{+/-}$ sub-grp.1 and sub-grp.2 at 4-month time point, and by 3- and 5- fold at 12 months, respectively. *Mlh1*$^{+/-}$ spleens of both sub-groups showed expected approximately 50% MLH1 protein and no increase in deletion% compared to *Mlh1*$^{+/+}$ spleens at both time points (**Supplementary fig. 7**). The *Mlh1*$^{+/-}$ outlier mouse (indicated by arrow) had no detectable MLH1 protein and showed frequent (25%) deletions at mononucleotide repeats in jejunum, comparable to the deletion% in age-matched *Mlh1*$^{-/-}$ mice (32%) (**Fig. 5**).

Discussion

In mouse-specific MSI studies, there has always been discrepancies due to the use of different microsatellite panels. Our microsatellite panel consisted of three microsatellites, two highly unstable mononucleotide repeats A33 and A27 [28], and a conventional dinucleotide D14MIt15. In line with previous studies [28, 36, 37], we observed mononucleotide repeats to be more unstable than dinucleotide repeats. In addition to sensitive markers, we used a highly sensitive SM-PCR based MSI assay, which can detect MSI levels as low as 1%, whereas standard PCR based MSI has a detection limit of 20-25% [38]. The combination of sensitive microsatellite markers and sensitive detection method gave us an unprecedented advantage of detecting very low MSI. Our study underscores the importance of MSI assay sensitivity. We detected MSI as low as 1%; others using less-sensitive marker and conventional MSI assay have failed to detect MSI even in neoplastic tissues of *Mlh1*$^{+/-}$ mice [39].

We observed an *Mlh1* dosage-dependent increase in deletion% at microsatellite repeats, insertions were mostly stable. Other studies [40, 41] have also reported increase in deletions only in MMR deficient conditions, here we report that the deletion biasness also persists in *Mlh1*$^{+/-}$ condition. Surprisingly, in our study, microsatellite A33 had moderately high 1bp insertions in both jejunum

and spleen of *Mlh1*$^{+/+}$ mice which further increased with age (**Supplementary fig. 2**). Further, we observed a decreasing trend of these 1bp insertions with *Mlh1* gene dosage, which further decreased to almost none in intestinal tumor from 12-month-old *Mlh1*$^{-/-}$ mice (**Supplementary fig. 2**). A similar tendency of decrease in insertions was also reported by Sai et al. in *Mlh1*$^{-/-}$ yeast compared to wild type yeast [42], we first time show this phenomenon in higher eukaryotes, further also show an intermediate trend of this decrease in *Mlh1*$^{+/-}$ condition. Consistent to observation made by Sai et al., almost all insertions observed at A33 in our study were single repeat unit in size (i.e. 1bp) (**Fig. 2**).

We did not observe a substantial increase in deletions in jejunum of 12-month-old *Mlh1*$^{+/-}$ mice compared to 4-month-old *Mlh1*$^{+/-}$ mice; in *Mlh1*$^{-/-}$ mice this increase was seen prominently (**Supplementary fig. 8**). This observation i.e. no strong correlation between MSI and age in MMR heterozygous condition is in line with previous studies [13, 14]. This can be explained by intra-age variation in MSI in young *Mlh1*$^{+/-}$ mice, 42% of 4-month *Mlh1*$^{+/-}$ mice in our study already showed deletion frequency comparable to 12 months *Mlh1*$^{+/-}$ mice (**Supplementary fig. 8**). As with MSI levels, we did not observe a substantial decrease in Mlh1 mRNA and MLH1 protein levels between the two age groups. However, a significant decrease in Mlh1 mRNA levels was observed when young (1-month old) *Mlh1*$^{+/-}$ mice compared to 12-month old *Mlh1*$^{+/-}$ mice (unpaired t test: p-value= 0.0283) (**Supplementary fig. 9 and Figure 3A**). As seen in 4- and 12-month old mice, jejunum of 1-month old *Mlh1*$^{+/-}$ mice also showed lower than expected (30%) average Mlh1 mRNA level compared to age-matched *Mlh1*$^{+/+}$ jejuna (**Supplementary fig. 9**). Spleen of 1 month *Mlh1*$^{+/-}$ mice, showed the expected approximately 50% Mlh1 mRNA expression as the other age groups. Therefore, the moderate MSI detected in jejunum of 4-month-old *Mlh1*$^{+/-}$ mice may occur already at an earlier age.

LS associated colorectal adenomas and carcinomas are MMR deficient, which led to the conventional understanding of MMR genes function that they follow the Knudson's two-hit hypothesis [43]. However, more evidence of presence of MMR protein in LS associated adenomas are emerging suggesting that the "second hit" of MMR may arise after the tumorigenesis initiation [44], making the idea of MMR haploinsufficiency more plausible, which we showed as low-level MSI in normal jejunum of *Mlh1*$^{+/-}$ mice even with expected levels (i.e. 50% compared to wildtype mice) of MLH1. Further, we showed that the "second hit" occurs in a sporadic manner and is primarily due to *Mlh1* promoter methylation and not LOH, which further elevates MSI in normal jejunum. Additionally,

number of *Mlh1*$^{+/-}$ mice with *Mlh1* promoter methylation in jejunum increases with age which shows the progressive nature of the "second hit".

Lifespan was reduced in *Mlh1*$^{+/-}$ mice compared to *Mlh1*$^{+/+}$ mice (**Supplementary fig. 10**): 30% of total *Mlh1*$^{+/-}$ mice died by age of 8 months, similar observation was reported by Edelmann et al. [16]. However, unlike Edelmann et al., no easily identifiable tumors were observed in the small intestine of *Mlh1*$^{+/-}$ mice even at 12 months. Our tumor observations were only limited to mice that were scarified to harvest tissues; we did not dissect post-mortem the 30% *Mlh1*$^{+/-}$ mice which died in our care to see tumor incidence, though the reason for death of these mice is unknown, the most likely assumption is they died because of MMR associated abnormalities. Similar to our observation, Prolla et al. did not observe increase in tumor incidence in *Mlh1*$^{+/-}$ mice compared to *Mlh1*$^{+/+}$ mice [21]. Also, Pussila et al. did not observe any tumor in the colon of 12 months old *Mlh1*$^{+/-}$ mice [39] (colon and jejunum have similar tumor incidence in *Mlh1*$^{+/-}$ mice [16]).

MSI reports on global MMR-dependent genome instability, but MSI per se does not necessarily lead to tumorigenesis [21]. Tumorigenesis initiates when MSI occurs in MSI-target genes giving the cells selective advantage and making them malignant. These genes include tumor suppressors and oncogenes involved in different regulatory pathways, including cell proliferation, cell cycle, apoptosis, and DNA repair [45, 46]. As genes (particularly exons) have more efficient MMR activity [47], it is highly unlikely that MSI-target genes would have early mutation accumulation, which could potentially alter gene function. None of the three microsatellites we tested were exonic, and we did not observe any tumors in normal *Mlh1*$^{+/-}$ jejunum, which suggests that the MSI detected in normal jejunum of *Mlh1*$^{+/-}$ mice in our study represents early stage tissue-specific genomic instability prior to tumorigenesis. The 4-month-old outlier *Mlh1*$^{+/-}$ mouse with 25% deletions in jejunum (comparable to age-matched *Mlh1*$^{-/-}$ mice) would have had higher chance of accumulating mutations in MSI-target genes and thus, if aged more, may have potentially triggered tumorigenesis in GI tract.

We performed experiments in whole jejunum (i.e. mucosa, submucosa, and muscularis externa). A question that arises is what cell population in *Mlh1*$^{+/-}$ jejunum is the source of MSI we detected. Given that the endogenous source of MSI is strand slippage during DNA replication, MSI more likely originates in dividing cells. Such cells reside in the crypts of the intestinal epithelium of mucosa: namely, intestinal stem cells and the transit amplifying (TA) cells, as such, the probable origin of MSI detected in our study. Ionov et al. have also proposed the stem cells or TA cells to be the origin of

MSI [48]. By IHC we showed that all crypt cells of *Mlh1*$^{+/-}$ mice (n=6) were MLH1-positive (**Supplementary fig. 5**), therefore, it is likely that MSI originated in crypt cells in presence of MLH1.

To date, there is no study examining the relationship between MLH1 expression levels and MSI in normal tissues. Low-level MSI in non-neoplastic tissues has been reported, but studies have proposed MMR deficiency rather than MMR haploinsufficiency as possible mechanism [13-15]. We detected low-level MSI in jejunum of 4 months old *Mlh1*$^{+/-}$ mice expressing approximately expected 50% of MLH1 protein (4 months old *Mlh1*$^{+/-}$ sub. grp.1 in **Fig. 5B**) which indicates MMR haploinsufficiency. 12-month old *Mlh1*$^{+/-}$ mice with almost half the expected MLH1 levels in jejunum (12 months old *Mlh1*$^{+/-}$ sub. grp.1 in **Fig. 5B**) showed comparable MSI levels to the 4 months counterparts, suggesting that even lower (24%) MLH1 protein levels maintain MMR function. Further, we observed that MMR protein levels should be greatly reduced (0-2%) in order to have substantial increase in MSI (as seen in 4 and 12 months old *Mlh1*$^{+/-}$ sub. grp.2 in **Fig. 5B**). MMR function appears to be dependent on a critical threshold of MMR proteins, depletion of which is independent of age. Further, all these observations applied specifically to jejunum only and not in other tissues tested.

We demonstrate here tissue-specific reduction in MMR protein levels, accompanied by low-level MSI at mononucleotide repeats in MMR-proficient normal jejunum. This indicates tissue-specific MMR haploinsufficiency in the (non-neoplastic) GI tract, which provides an explanation for vulnerability to early onset of MMR-associated cancer.

Author contributions

K.S. designed and performed the experiments, analyzed the data, interpreted the results and wrote the manuscript. L.K. conceived and designed the study, interpreted the data, reviewed and edited the manuscript, acquired funding and supervised K.S. M.T. assisted in tissue collection, MSI assay, qPCR experiments, and genotyping.

Acknowledgments

We are grateful to Elli Aska, Erika Gucciardo, Manuela Tumiati and Saara Ollila for critical reading of the manuscript, and to members of the Kauppi lab for constructive comments on the figures. We would like to thank the staff of the following core facilities at University of Helsinki: Laboratory Animal Center team, Reeta Hautala and Helena Taskinen from Tissue preparation and histochemistry

unit for IHC services, Kaisa Laajanen and Laura Häkkinen at the Genome Biology Unit for digital slide scanning services, Suvi Kyttänen and Tiina Hannunen from FIMM sequencing unit for fragment analysis runs.

Disclosure of conflicts of interest

No conflicts of interest were disclosed.

Funding

This work was supported by the Academy of Finland grants: 263870, 292789, 256996, and 306026 (to L.K.), Sigrid Jusélius Foundation (to L.K.) and the Väre Foundation (to K.S.).

References

1. Burt, R.W., *Colon cancer screening.* Gastroenterology, 2000. **119**(3): p. 837-53.
2. Vasen, H.F., et al., *Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts.* Gut, 2013. **62**(6): p. 812-23.
3. Loukola, A., et al., *Microsatellite instability in adenomas as a marker for hereditary nonpolyposis colorectal cancer.* Am J Pathol, 1999. **155**(6): p. 1849-53.
4. Strand, M., et al., *Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair.* Nature, 1993. **365**(6443): p. 274-6.
5. Jiricny, J., *Postreplicative mismatch repair.* Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a012633.
6. Pray, L.A., *DNA Replication and Causes of Mutation.* Nature Education, 2008. **1**(1).
7. Edelmann, W., et al., *Meiotic pachytene arrest in MLH1-deficient mice.* Cell, 1996. **85**(7): p. 1125-34.
8. Yamamoto, H. and K. Imai, *Microsatellite instability: an update.* Arch Toxicol, 2015. **89**(6): p. 899-921.
9. Aaltonen, L.A., et al., *Clues to the pathogenesis of familial colorectal cancer.* Science, 1993. **260**(5109): p. 812-6.
10. Hemminki, A., et al., *Loss of the wild type MLH1 gene is a feature of hereditary nonpolyposis colorectal cancer.* Nat Genet, 1994. **8**(4): p. 405-10.
11. Mangold, E., et al., *Spectrum and frequencies of mutations in MSH2 and MLH1 identified in 1,721 German families suspected of hereditary nonpolyposis colorectal cancer.* Int J Cancer, 2005. **116**(5): p. 692-702.
12. Peltomaki, P., *Epigenetic mechanisms in the pathogenesis of Lynch syndrome.* Clin Genet, 2014. **85**(5): p. 403-12.
13. Alazzouzi, H., et al., *Low levels of microsatellite instability characterize MLH1 and MSH2 HNPCC carriers before tumor diagnosis.* Hum Mol Genet, 2005. **14**(2): p. 235-9.
14. Coolbaugh-Murphy, M.I., et al., *Microsatellite instability in the peripheral blood leukocytes of HNPCC patients.* Hum Mutat, 2010. **31**(3): p. 317-24.
15. Parsons, R., et al., *Mismatch repair deficiency in phenotypically normal human cells.* Science, 1995. **268**(5211): p. 738-40.
16. Edelmann, W., et al., *Tumorigenesis in Mlh1 and Mlh1/Apc1638N mutant mice.* Cancer Res, 1999. **59**(6): p. 1301-7.
17. Kansikas, M., et al., *Assessing how reduced expression levels of the mismatch repair genes MLH1, MSH2, and MSH6 affect repair efficiency.* Hum Mutat, 2014. **35**(9): p. 1123-7.
18. Kasela, M., M. Nystrom, and M. Kansikas, *PMS2 expression decrease causes severe problems in mismatch repair.* Hum Mutat, 2019.
19. Wang, L., et al., *Whole-exome sequencing of human pancreatic cancers and characterization of genomic instability caused by MLH1 haploinsufficiency and complete deficiency.* Genome Res, 2012. **22**(2): p. 208-19.
20. Baker, S.M., et al., *Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over.* Nat Genet, 1996. **13**(3): p. 336-42.
21. Prolla, T.A., et al., *Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair.* Nat Genet, 1998. **18**(3): p. 276-9.
22. Barker, N., M. van de Wetering, and H. Clevers, *The intestinal stem cell.* Genes Dev, 2008. **22**(14): p. 1856-64.
23. Crippen, T.L. and I.M. Jones, *Cell proliferation in the bone marrow, thymus and spleen of mice studied by continuous, in vivo bromodeoxycytidine labelling and flow cytometric analysis.* Cell Tissue Kinet, 1989. **22**(3): p. 203-12.
24. Pussila, M., et al., *Cancer-predicting gene expression changes in colonic mucosa of Western diet fed Mlh1+/- mice.* PLoS One, 2013. **8**(10): p. e76865.
25. Jeffreys, A.J., et al., *Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells.* Nucleic Acids Res, 1988. **16**(23): p. 10953-71.
26. Beal, M.A., et al., *Single-molecule PCR analysis of an unstable microsatellite for detecting mutations in sperm of mice exposed to chemical mutagens.* Mutat Res, 2015. **775**: p. 26-32.
27. Kraytsberg, Y. and K. Khrapko, *Single-molecule PCR: an artifact-free PCR approach for the analysis of somatic mutations.* Expert Rev Mol Diagn, 2005. **5**(5): p. 809-15.

28. Kabbarah, O., et al., *A panel of repeat markers for detection of microsatellite instability in murine tumors.* Mol Carcinog, 2003. **38**(4): p. 155-9.
29. Jeffreys, A.J., et al., *Complex gene conversion events in germline mutation at human minisatellites.* Nat Genet, 1994. **6**(2): p. 136-45.
30. Covarrubias-Pazaran, G., et al., *Fragman: an R package for fragment analysis.* BMC Genet, 2016. **17**: p. 62.
31. Berg, K.D., et al., *Detection of microsatellite instability by fluorescence multiplex polymerase chain reaction.* J Mol Diagn, 2000. **2**(1): p. 20-8.
32. Moolenbeek, C. and E.J. Ruitenberg, *The "Swiss roll": a simple technique for histological studies of the rodent intestine.* Lab Anim, 1981. **15**(1): p. 57-9.
33. Varghese, F., et al., *IHC Profiler: an open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples.* PLoS One, 2014. **9**(5): p. e96801.
34. Herman, J.G., et al., *Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9821-6.
35. Fraga, M.F., et al., *A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors.* Cancer Res, 2004. **64**(16): p. 5527-34.
36. Webster, M.T., N.G. Smith, and H. Ellegren, *Microsatellite evolution inferred from human-chimpanzee genomic sequence alignments.* Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8748-53.
37. Bacher, J.W., et al., *Use of mononucleotide repeat markers for detection of microsatellite instability in mouse tumors.* Mol Carcinog, 2005. **44**(4): p. 285-92.
38. Coolbaugh-Murphy, M., et al., *Estimating mutant microsatellite allele frequencies in somatic cells by small-pool PCR.* Genomics, 2004. **84**(2): p. 419-30.
39. Pussila, M., et al., *Mlh1 deficiency in normal mouse colon mucosa associates with chromosomally unstable colon cancer.* Carcinogenesis, 2018. **39**(6): p. 788-797.
40. Lu, Y., T.D. Soong, and O. Elemento, *A novel approach for characterizing microsatellite instability in cancer cells.* PLoS One, 2013. **8**(5): p. e63056.
41. Keysselt, K., et al., *Different in vivo and in vitro transformation of intestinal stem cells in mismatch repair deficiency.* Oncogene, 2017. **36**(19): p. 2750-2761.
42. Sia, E.A., et al., *Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes.* Mol Cell Biol, 1997. **17**(5): p. 2851-8.
43. Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma.* Proc Natl Acad Sci U S A, 1971. **68**(4): p. 820-3.
44. Valo, S., et al., *DNA hypermethylation appears early and shows increased frequency with dysplasia in Lynch syndrome-associated colorectal adenomas and carcinomas.* Clin Epigenetics, 2015. **7**: p. 71.
45. Duval, A. and R. Hamelin, *Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability.* Cancer Res, 2002. **62**(9): p. 2447-54.
46. Vilkki, S., et al., *Screening for microsatellite instability target genes in colorectal cancers.* J Med Genet, 2002. **39**(11): p. 785-9.
47. Frigola, J., et al., *Reduced mutation rate in exons due to differential mismatch repair.* Nat Genet, 2017.
48. Ionov, Y., et al., *Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis.* Nature, 1993. **363**(6429): p. 558-61.

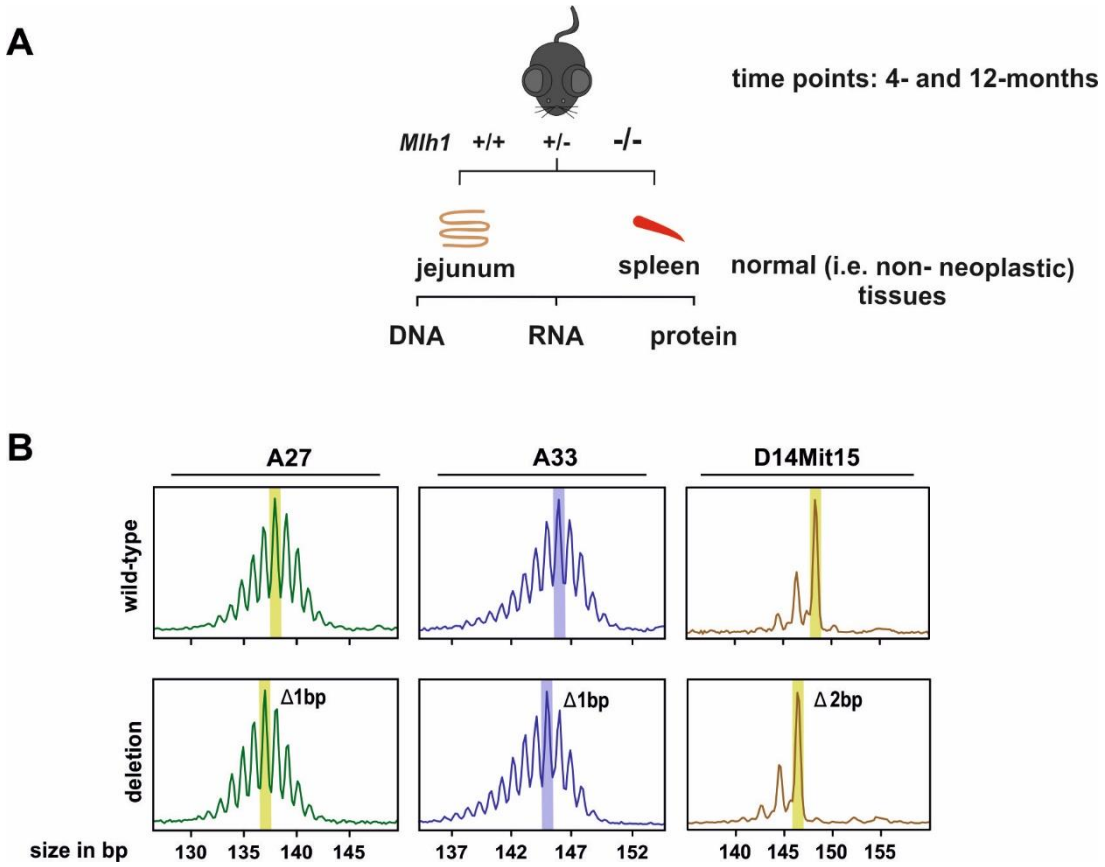


Fig. 1. Study approach. (**A**) Workflow of the study. We tested microsatellite instability (MSI) at single-DNA molecule resolution along with analyzing *Mlh1* expression both at mRNA and protein level. (**B**) Representative capillary electrophoresis electropherograms of single-molecule PCR (SM-PCR) based MSI assay. We tested MSI at three microsatellites: two mononucleotide repeats A27 and A33, and a dinucleotide repeat D14Mit15. The top and bottom rows show electropherograms scored as wild-type and single repeat unit deletion mutant allele, respectively. The highlighted peak represents the highest peak that was scored; smaller peaks flanking the highest peak are stutter peaks, which are typical feature of microsatellite markers.

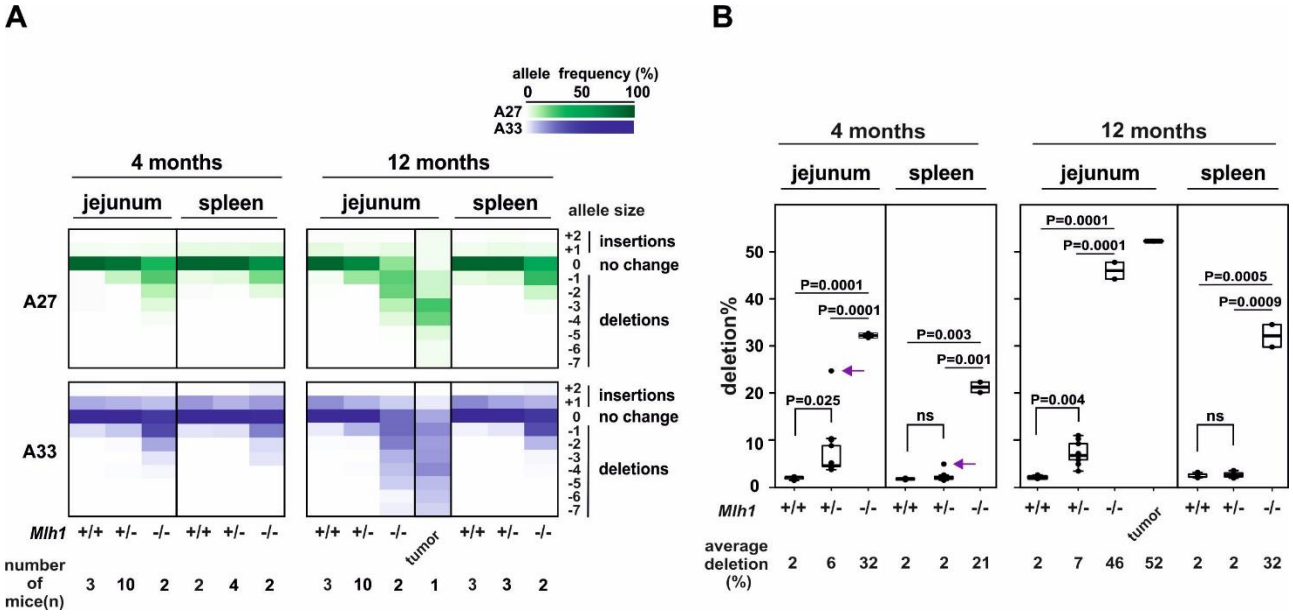


Fig. 2. Single-molecule MSI analysis at mononucleotide repeats in jejunum and spleen at 4- and 12-month time points. (**A**) Heat map of the frequency of different alleles of mononucleotide repeats A27 and A33. The outlier 4-month old *Mlh1*$^{+/-}$ mouse (indicated by an arrow in **Fig. 2B**) was excluded from the heat map, and is shown separately in **Supplementary fig. 5.** (**B**) Deletion% in jejunum and spleen of *Mlh1*$^{+/-}$ mice. Arrow indicates the outlier *Mlh1*$^{+/-}$ mouse.

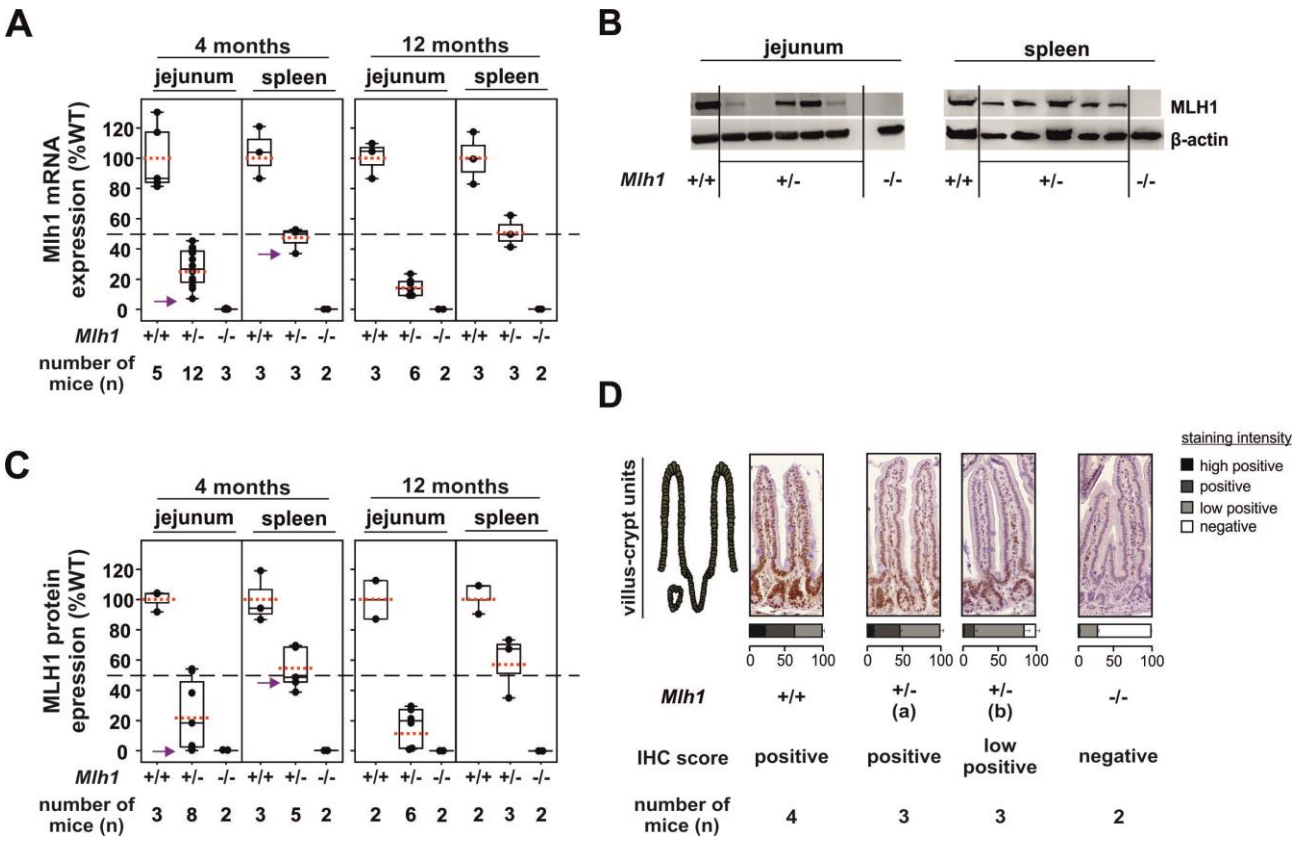


Fig. 3. *Mlh1* gene expression analysis in jejunum and spleen of *Mlh1*$^{+/-}$ mice at 4- and 12-month time point. (**A**) Mlh1 mRNA expression analysis. Arrow indicates the outlier *Mlh1*$^{+/-}$ mouse that showed high level of deletion% at mononucleotide repeats. The red-dotted line across each box-plot represents the average value. The dashed-horizontal line across the chart represents the 50% Mlh1 mRNA expression. (**B**), (**C**) and (**D**) MLH1 protein levels analysis. (**B**) Representative image of western blot. The corresponding lanes in the two adjacent western blots are tissues from the same mouse. Empty lanes in western blot do not necessarily mean absence of MLH1 but possibly MLH1 protein levels below detectable range of our western blot assay (**Supplementary fig. 6**). (**C**) MLH1 protein expression analysis. The red-dotted line across each box-plot represents the average value. The dashed-horizontal line across the chart represents the 50% MLH1 protein expression. (**D**) Representative IHC images of jejunum. Middle two IHC images shows a side-by-side comparison of villus-crypt units scored as (a) positive and (b) low positive by IHC profiler. The 100% stacked bars (mean + SD) below the representative IHC images show distribution of MLH1 staining intensities for all the villus-crypt units analyzed for each *Mlh1* genotype.

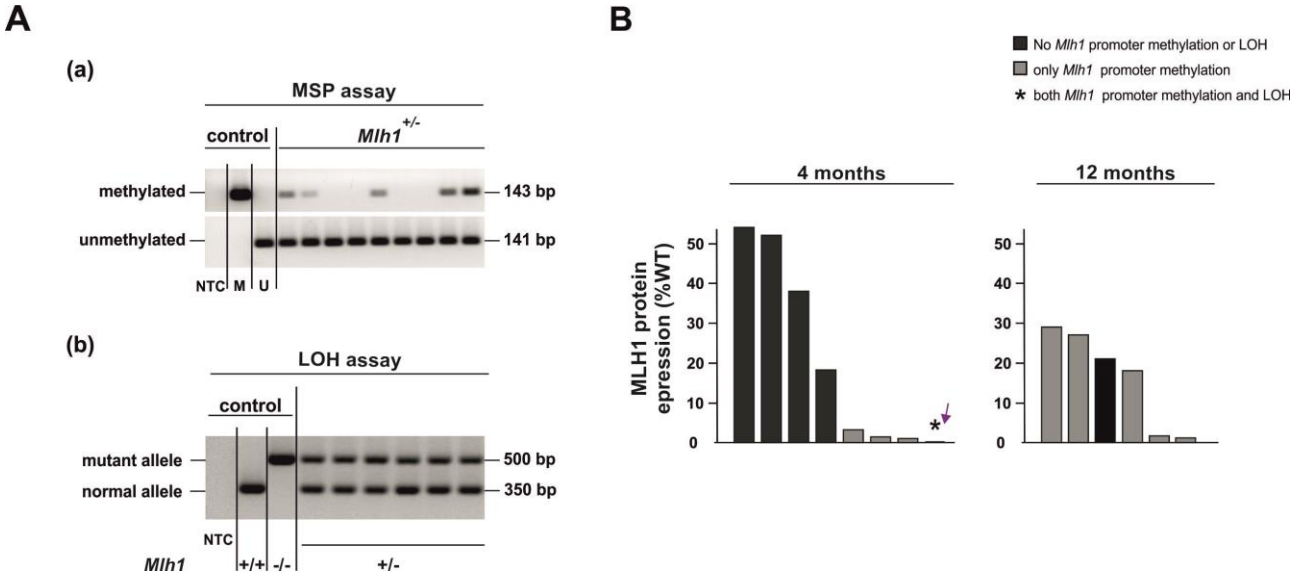


Fig. 4. Mechanism underlying MLH1 depletion in *Mlh1*$^{+/-}$ jejuna. (**A**) Representative image of (a) Methylation specific PCR (MSP). Universally methylated mouse DNA standard (cat. no. D5012, Zymo Research) (M) and genomic DNA extracted from normal spleen (U) was used as positive and negative controls, respectively. (b) Loss of heterozygosity (LOH) assay. Genomic DNA from *Mlh1*$^{+/+}$ and *Mlh1*$^{-/-}$ was used as mutant and normal allele controls, respectively. (**B**) MLH1 expression levels, and *Mlh1* promoter methylation and LOH status. Arrow indicates the outlier Mlh1$^{+/-}$ mouse.

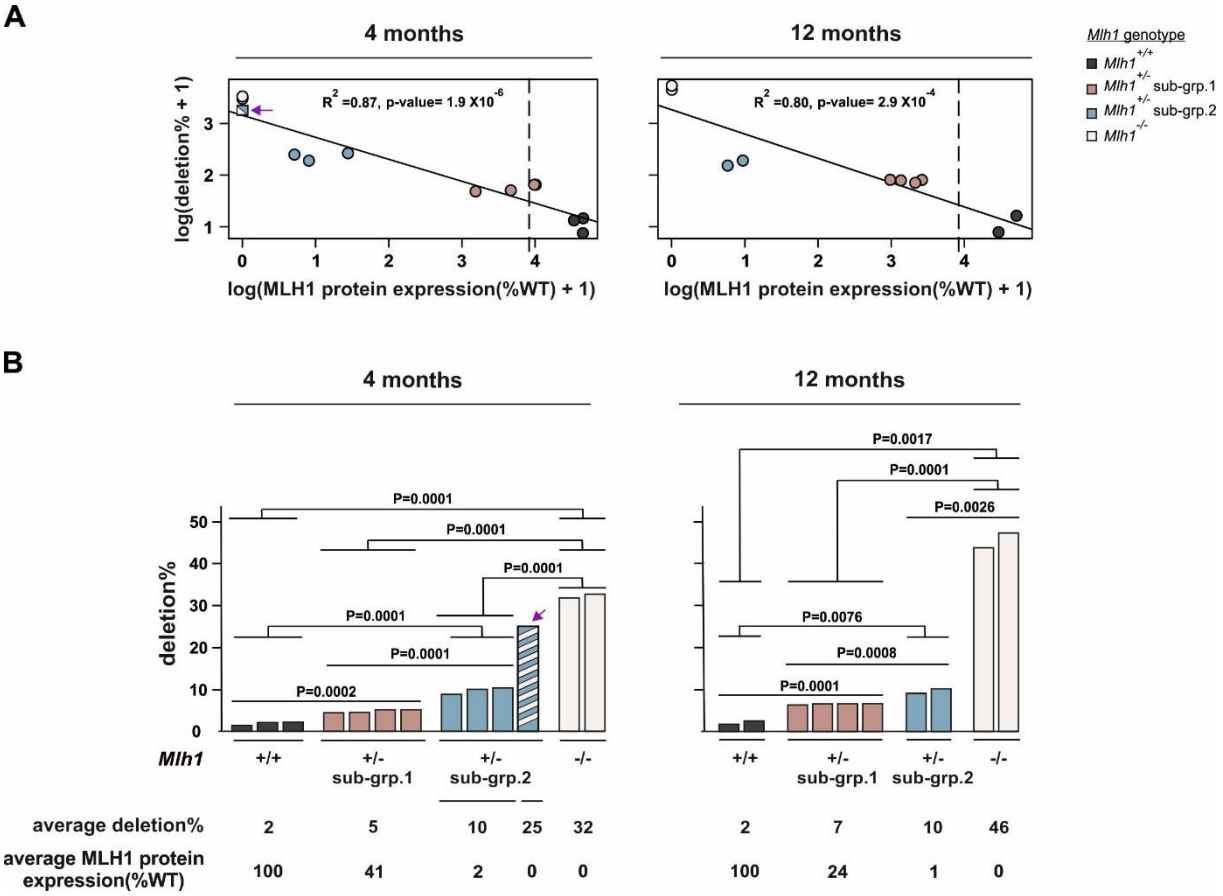
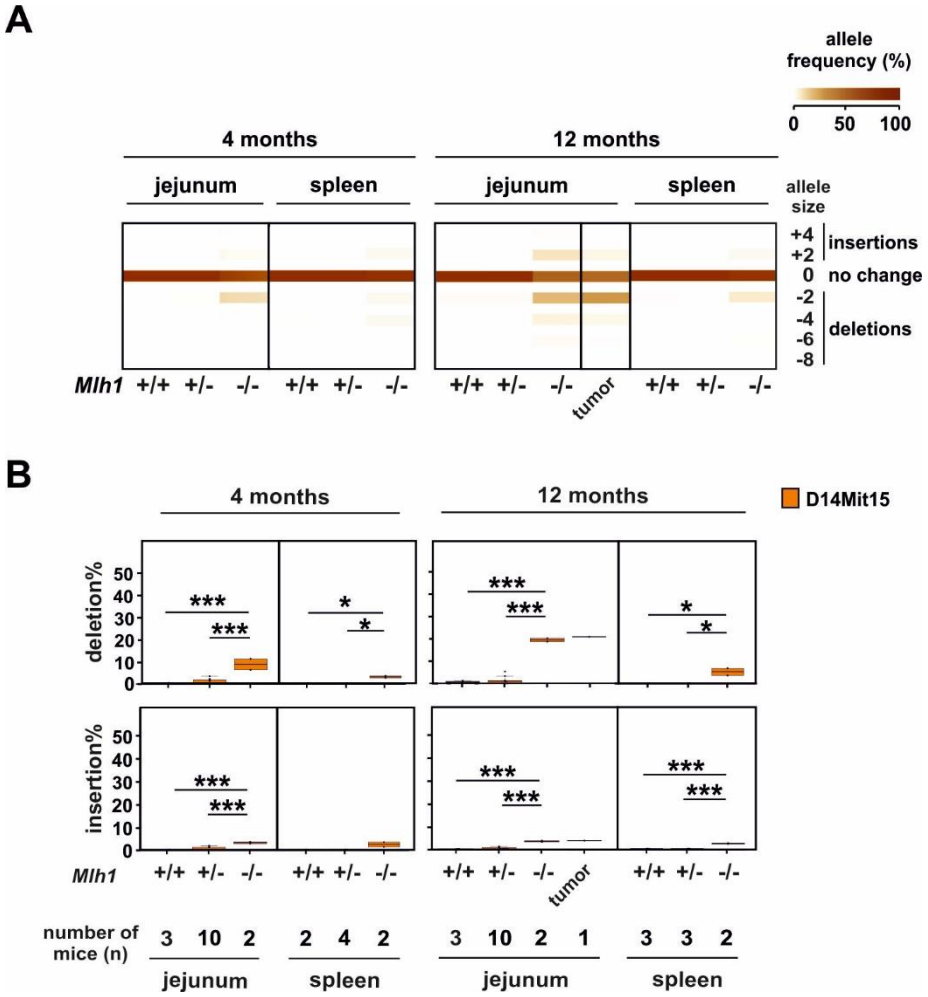


Fig. 5. Correlation between deletion% at mononucleotide repeats and MLH1 protein levels in jejunum of *Mlh1*$^{+/-}$ mice. (**A**) Linear regression fitting curve of log(deletions%+1) ~ log(MLH1 protein expression(%WT)+1)) at 4- and 12-month time point. (**B**) Comparison of deletion% between the two *Mlh1*$^{+/-}$ sub-groups. The arrow indicates the outlier *Mlh1*$^{+/-}$ mice.

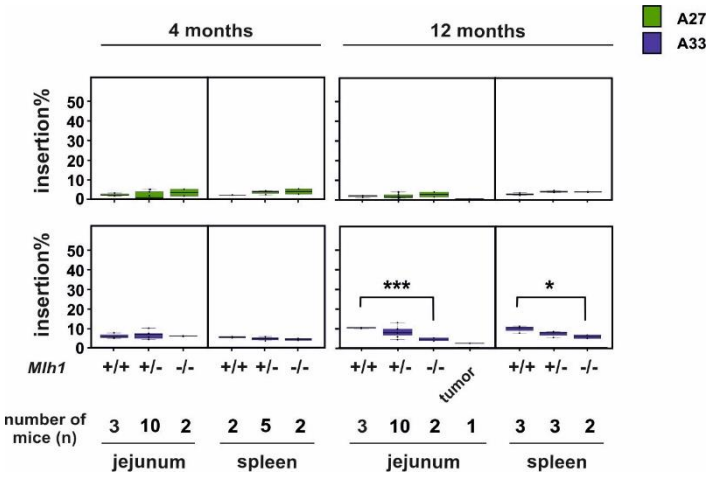
Table 1. Total number of *Mlh1*$^{+/-}$ jejuna assayed, and their *Mlh1* promoter methylation and LOH status.

	4 months	12 months
***Mlh1* promoter methylation**	7 out of 12 (58.3%)	9 out of 11 (81.8%)
***Mlh1* LOH**	1* out of 12 (8.3%)	0 out of 11 (0%)

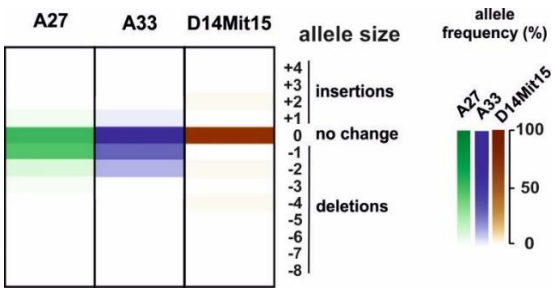
* Outlier 4-month *Mlh1*$^{+/-}$ mouse



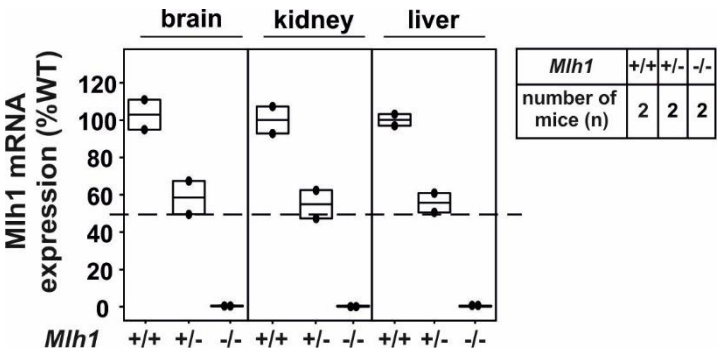
Supplementary fig. 1. Single-molecule MSI analysis at dinucleotide D14Mit15. (**A**) Heat map of the frequency of different alleles of D14Mit15 observed in the study. (**B**) MSI in normal jejunum and spleen. Indicated are significant p-values for the insertion% and deletion% comparisons between i) *Mlh1*$^{+/+}$ and *Mlh1*$^{+/-}$, ii) *Mlh1*$^{+/+}$ and *Mlh1*$^{-/-}$, and iii) *Mlh1*$^{+/-}$ and *Mlh1*$^{-/-}$ mice for each tissue.



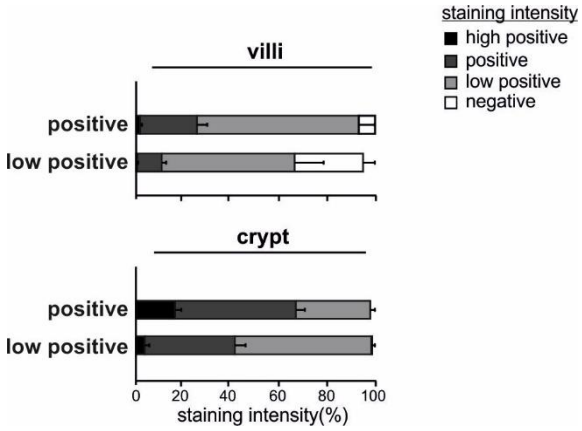
Supplementary fig. 2. Insertion% at mononucleotide repeats A27 and A33 in jejunum and spleen. Top- and bottom-panel shows insertion% at A27 and A33, respectively. Indicated are statistically significant p-values for insertion% comparison between i) *Mlh1*$^{+/+}$ and *Mlh1*$^{+/-}$, ii) *Mlh1*$^{+/+}$ and *Mlh1*$^{-/-}$, and iii) *Mlh1*$^{+/-}$ and *Mlh1*$^{-/-}$ mice for each tissue for both time points.



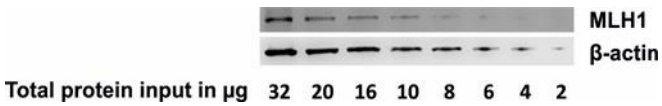
Supplementary fig. 3. MSI in the outlier *Mlh1*$^{+/-}$ mouse. Heat map of the frequency of different alleles of A27, A33 and D14Mit15 observed in jejunum of outlier *Mlh1*$^{+/-}$ mouse showing high-level of deletion%.



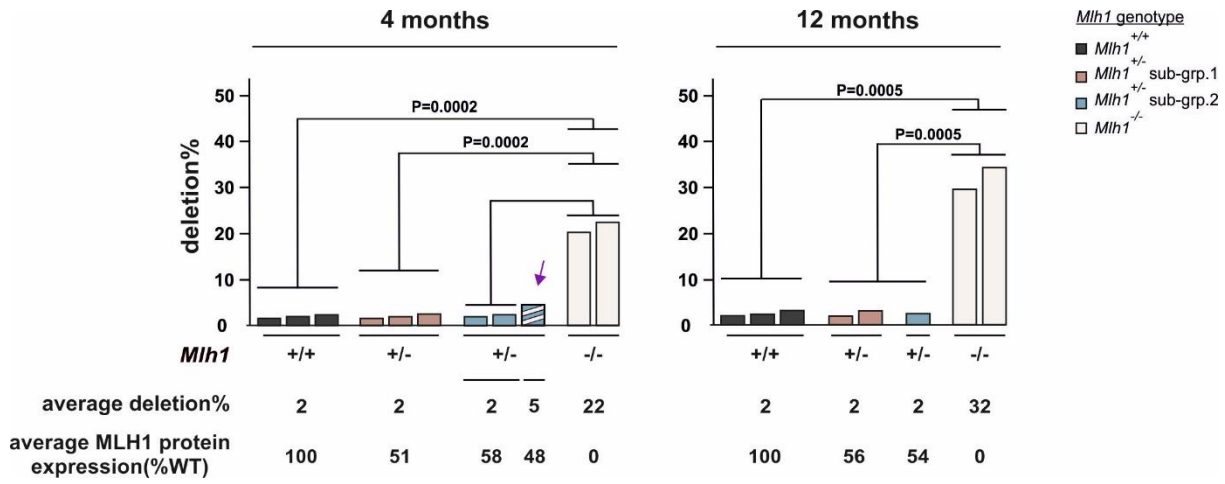
Supplementary fig. 4. Mlh1 mRNA expression analysis in brain, kidney and liver of Mlh1$^{+/-}$ mice expressing low Mlh1 mRNA levels in jejunum. The dashed-horizontal line represents the 50% Mlh1 mRNA expression.



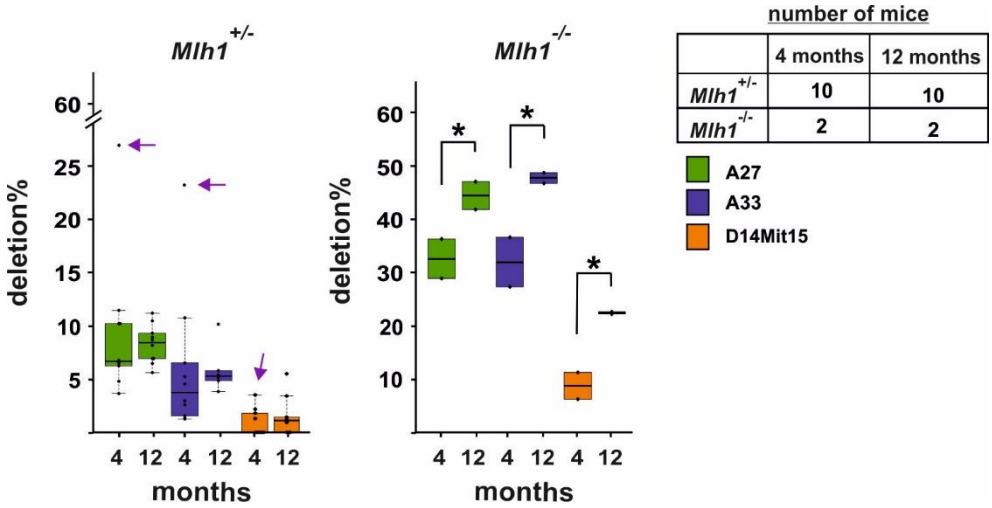
Supplementary fig. 5. Comparison of MLH1 staining intensity in crypts and villi of MLH1-positive and MLH1-low positive *Mlh1*$^{+/-}$ jejuna.



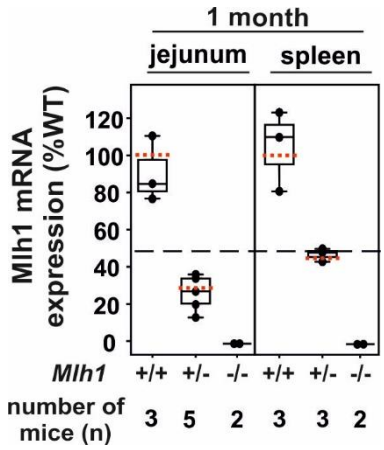
Supplementary fig. 6. Sensitivity of the western blot assay. A dilution series of protein was run to determine sensitivity of the western blot assay; Infra-red signal was not detectable below 4 µg of protein for MLH1.



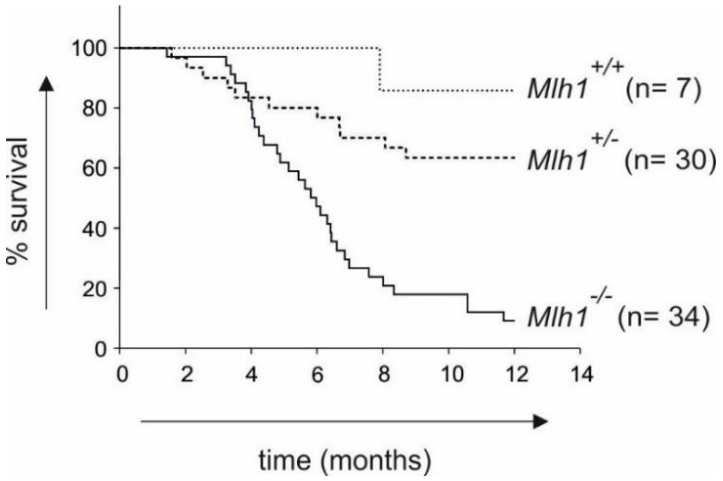
Supplementary fig. 7. Relationship between MSI and MLH1 protein levels in spleen of *Mlh1*$^{+/-}$ mice. The arrow indicates the outlier *Mlh1*$^{+/-}$ mice.



Supplementary fig. 8. Deletion% and age. The arrow indicates the outlier *Mlh1*$^{+/-}$ mice showing high deletion%.



Supplementary fig. 9. Mlh1 mRNA expression analysis at 1-month time point. The dashed-horizontal line represents the 50% Mlh1 mRNA expression.



Supplementary fig. 10. Kaplan-Meier survival curves for Mlh1 mutant mice.

Supplementary table 1.

Western blotting:

Antibodies	Company and catalog number	Dilution used
MLH1	ab92312 (Abcam)	1:1000
Beta-actin	A5441 (Sigma)	1:5000
IRDye 800CW	926-32211 (Li-cor)	1:5000
IRDye 680RD	926-68070 (Li-cor)	1:5000

SM-PCR primers and program:

SM-PCR primers

Microsatellite loci	Paper reference	Labeled forward primer (5´- 3´)	Reverse primer (5´- 3´)	Amplicon size (bp)	Flurescent label	concetration of F and R primers / PCR reaction
A27	Kabbarah et al. [28]	TCCCTGTATAACCCTGGCTGACT	GCAACCAGTTGTCCTGGCGTGGA	138	6-FAM	0.5 µM
A33	Kabbarah et al. [28]	TACAGAGGATTGTCCTCTTGGAG	GCTGCTTCACTTGGACATTGGCT	147	VIC	0.2 µM
D14Mit15	Edelmann et al. [7]	TTGGCTGCTCACTTGCAG	TTACCCTCCCCATAACTCCC	148	NED	0.1 µM

PCR programs:

PCR A: multiplexed A33 and D14Mit15

98°C for 30 seconds
98°C for 10 seconds
66°C for 30 seconds — 35 cycles
72°C for 5 seconds
72°C for 2 minutes

PCR B: A27

98°C for 30 seconds
98°C for 10 seconds
70°C for 30 seconds — 35 cycles
72°C for 5 seconds
72°C for 2 minutes

qPCR primers and program:

qPCR primers

Gene	Forward primer (5´- 3´)	Reverse primer (5´- 3´)	Amplicon size (bp)
Mlh1	GGGAGGACTCTGATGTGGAA	AGAGCTTGGTCTGGTGCTGT	216
beta-actin	AGACTTCGAGCAGGAGATGG	AGGTCTTTACGGATGTCAACG	310

qPCR program:

95°C for 30 seconds
95°C for 10 seconds
60°C for 30 seconds — 40 cycles
72°C for 5 seconds

MSP primers and PCR program for *Mlh1* :

MSP primers

Paper reference	Methylated/unmethylated	Forward primer (5´- 3´)	Reverse primer (5´- 3´)	Amplicon size (bp)
Fraga et al. [35]	Methylated	GAATTTGAGCGTGAGGAGTTC	TAACCGACCGCTAAATAACTTCC	143
Fraga et al. [35]	Unmethylated	AGAATTTGAGTGTGAGGAGTTT	CCAACCACTAAATAACTTCCC	141

MSP PCR program:

95°C for 10 seconds
95°C for 30 seconds
62°C for 30 seconds — 35 cycles
72°C for 60 seconds
72°C for 7 minutes

LOH assay and genotyping for *Mlh1:*

Lysis buffer components:

50 mM KCl
10 mM Tris-HCl pH 8.3
2.5 mM MgCl2
0.1 mg/mL gelatin
0.45% NP-40
0.45% Tween20

genotyping primers

Paper reference	Forward primer	Mutant reverse primer (5´- 3´)	WT reverse primer (5´- 3´)	Amplicon size (bp)
Pussila et al. [24]	TGTCAATAGGCTGCCCTAGG	TGGAAGGATTGGAGCTACGG	TTTTCAGTGCAGCCTATGCTC	WT:350 /mutant:500

genotyping PCR program:

94°C for 2 minutes

94°C for 30 seconds	
55°C for 30 seconds	38 cycles
72°C for 45 seconds	