


RESEARCH

Open Access



Loss of *Msh2* and a single-radiation hit induce common, genome-wide, and persistent epigenetic changes in the intestine

Maria Herberg^{1†}, Susann Siebert^{2,3†}, Marianne Quaas^{1,4}, Torsten Thalheim¹, Karen Rother^{4,5}, Michelle Hussong^{2,6}, Janine Altmüller⁷, Christiane Kerner⁴, Joerg Galle^{1†}, Michal R. Schweiger^{2,6†} and Gabriela Aust^{4*†} 

Abstract

Background: Mismatch repair (MMR)-deficiency increases the risk of colorectal tumorigenesis. To determine whether the tumors develop on a normal or disturbed epigenetic background and how radiation affects this, we quantified genome-wide histone H3 methylation profiles in macroscopic normal intestinal tissue of young radiated and untreated MMR-deficient *VCMsh2^{LoxP/LoxP} (Msh2^{-/-})* mice months before tumor onset.

Results: Histone H3 methylation increases in *Msh2^{-/-}* compared to control *Msh2^{+/+}* mice. Activating H3K4me3 and H3K36me3 histone marks frequently accumulate at genes that are H3K27me3 or H3K4me3 modified in *Msh2^{+/+}* mice, respectively. The genes recruiting H3K36me3 enrich in gene sets associated with DNA repair, RNA processing, and ribosome biogenesis that become transcriptionally upregulated in the developing tumors. A similar epigenetic effect is present in *Msh2^{+/+}* mice 4 weeks after a single-radiation hit, whereas radiation of *Msh2^{-/-}* mice left their histone methylation profiles almost unchanged.

Conclusions: MMR deficiency results in genome-wide changes in histone H3 methylation profiles preceding tumor development. Similar changes constitute a persistent epigenetic signature of radiation-induced DNA damage.

Keywords: Intestine, Mismatch repair deficiency, Radiation, Msh2, Histone H3 methylation

Introduction

Mismatch repair (MMR) deficiency originates from loss of function of MMR proteins. Germline alterations of MMR genes have been related to Lynch syndrome [1] which is associated with a predisposition to a variety of malignant tumors including colorectal cancer. About 10% to 15% of all colorectal cancers are MMR-deficient. *MSH2* and *MLH1* are the genes that are most commonly altered [2, 3], and a loss of function results in an immediate increase of mismatch mutations [4] which accumulate particularly in microsatellites giving rise to the phenomenon of microsatellite instability (MSI) [2, 5].

The *VCMsh2^{LoxP/LoxP} (Msh2^{-/-})* mouse is a suitable model to quantify the severity and time course of molecular alterations caused by MMR deficiency [6, 7]. In organoids derived from intestinal stem cells (ISC) of these mice, MSI increases continuously from birth and affects microsatellites in all organoids already at a mouse age of 2 months. However, intestinal carcinomas become detectable only at a mouse age of 12 months and their occurrence is paralleled by tissue-wide changes in gene expression [7]. This obvious time delay between the early MSI in organoids and the late tumor onset as well as tissue-wide transcriptional changes suggest additional, as yet hidden, (epi)genetic changes.

Among the epigenetic changes frequently associated with tumor formation are histone methylations [8]. Depending on the lysine which is methylated, they can induce transcriptional activation or repression [9].

* Correspondence: gabriela.aust@medizin.uni-leipzig.de

[†]Maria Herberg, Susann Siebert, Joerg Galle, Michal R. Schweiger, and Gabriela Aust contributed equally to this work.

⁴Department of Surgery, Research Laboratories, Leipzig University, Liebigstr. 19, D-04103 Leipzig, Germany

Full list of author information is available at the end of the article



Tri-methylation of histone H3 at lysine 4 (H3K4me3) facilitates initiation of gene transcription, while H3K36me3 stabilizes elongation. H3K27me3 is frequently associated with suppression of gene transcription. To clarify whether histone methylation changes are already present in young *Msh2*^{-/-} mice, and whether they are capable of inducing the tissue-wide transcriptional changes present in aged *Msh2*^{-/-} mice [7], we performed genome-wide chromatin immunoprecipitation with high-throughput sequencing (ChIP-seq) comparing 4-month-old *Msh2*^{-/-} and control *Msh2*^{+/+} mice.

Based on our results, the question arose whether the observed changes are associated with loci-specific DNA repair or are part of a response on genomic instability at the cell level. Therefore, we analyzed whether a similar epigenetic response can be induced in the intestine by a single-radiation hit. ChIP-sequencing (ChIP-seq) analysis was performed on purpose 4 weeks after radiation when short term transcriptional and epigenetic responses are abated [10, 11]. At this time point, most intestinal crypts had already undergone monoclonal conversion, that is, they had become populated by the progeny of a single ISC, while all other ISC and their clones had vanished due to competition for niche space [12]. Potential selection of stress-tolerant ISC is mostly completed by clonal competition [13]. Moreover, these measurements enable us to answer the question of whether a single-radiation hit induces additional epigenetic changes in MMR-deficient tissue, which is controversially discussed at the gene level [14].

By comparing the epigenetic changes induced by loss of *Msh2* and a single-radiation hit, we identified a common, genome-wide, and persistent epigenetic response.

Materials and methods

Mouse and human tissue

The mouse strains B6.SJL-Tg(Villin-cre)997Gum/J and B6.Cg-*Msh2*^{tm2.1Rak}/J were obtained from The Jackson Laboratory (Bar Harbor, USA). By crossing both, the conditional *Msh2* allele has been placed under the control of the Villin-Cre transgene. Mice were genotyped as described [6]. *VC*^{+/?}*Msh2*^{LoxP/LoxP} (*Msh2*^{-/-}) and *VC*^{-/-}*Msh2*^{LoxP/LoxP} (*Msh2*^{+/+}) mice were bred under specific pathogen-free conditions. A mouse doublet consisted of one *Msh2*^{-/-} and one *Msh2*^{+/+} mouse from one litter. In addition, we analyzed normal human colon samples of patient 7 (74 years old, female, *MSH6* mutation carrier) of the cohort described in [3].

Radiation of mice and collection of tissue samples

Three-month-old mice were placed in the radiation unit (X-ray generator, Gulmay D3225, round tube 170 mm, radiation level at 2.5 cm: 1.068 Gy/min). One pair of mice of each genotype was irradiated with a total dose of 0.5 Gy, while the other pair was left non-irradiated. All

mice were killed after 28 days. Samples of the proximal jejunum were dissected, shock-frozen without delay, and stored in liquid nitrogen until further use.

ChIP

Chromatin was prepared from up to 50 mg jejunum using SDS Shearing Buffer and the truChIP Tissue Chromatin Shearing Kit (Covaris, Brighton, UK). Chromatin immunoprecipitations (ChIPs) were run on the IP-Star compact system using the Auto iDeal ChIP-seq kit for histones (Diagenode, Seraing, Belgium). For ChIP-seq validation, shredded intestinal tissue was cross-linked in 1% formaldehyde for 8 min at room temperature and homogenized in ZR 2.0 mm Bashing-Bead™ Lysis Tubes (Zymo Research Corp., Tustin, USA) with a Precellys24 homogenization device (VWR International GmbH, Darmstadt, Germany). ChIP assays were performed with minor modifications as described [15]. DNA from 15 mg tissue was used in each IP with H3K4me3, H3K27me3, H3K36me3, and H3pan premium rabbit polyclonal antibodies (Diagenode). A non-targeting IgG rabbit antibody (Cell Signaling, Leiden, Netherlands) was used to control non-specific signals. One microliter of purified chromatin was applied for semi-quantitative real-time PCR using the absolute qPCR SYBR Green Mix (Thermo) performed with the ABI 7500 Real-Time PCR System (Applied Biosystems, Foster City, USA). Primers are given in the Additional file 1. All PCR results were normalized to input controls and to an unmodified region (background).

Quantitative PCR and Illumina sequencing

As quality control for ChIP precipitates, a quantitative real-time PCR analysis was performed in a 384-well format in 10 µl volume using the GoTaq qPCR Master Mix (Promega, Mannheim, Germany) and run on a LightCycler 480 (Roche Diagnostics, Mannheim, Germany). Primers are given in the Additional file 1. The relative enrichment was calculated using the % input method. Library preparations were performed according to the TrueSeq LT PCR free or TruSeq Nano DNA Kit instructions (Illumina, San Diego, USA). Sequencing was performed on the Illumina HiSeq 4000 with 76 bp paired-end sequencings.

Data pre-processing and peak calling

The sequenced reads underwent a quality trimming using Cutadapt [16] to ensure a minimum read length of 40 bases and a Phred quality score above 20. Quality reports were generated and analyzed using Fastqc [17]. The high-quality reads were mapped to the mouse reference genome NCBI37/mm9 (GRCh37/hg19) using the software tool segemehl 0.2.0 [18]. Histone modification peaks, i.e., genomic regions, in which modifications are

enriched, were identified by applying MACS 1.4.2 [19, 20] with the H3pan data as controls and the following, modification independent parameter settings: `--nomodel --space=30 --bw 300 --pvalue 1e-5`. The parameter `--shiftsize` has been set to 200 for sharp (H3K4me3) and to 600 for broad (H3K27me3, H3K36me3) peaks.

Published H3K4me3, H3K27me3, H3K36me3, and DNA methylation sequencing data have been reanalyzed using read data provided [21] and calling peaks as described or using the peak data provided [22] directly. In the case of RNA-seq data [3, 22], the raw counts provided were normalized to reads per kilobase million (RPKM) values and transformed into log-scale for further analysis [23]. Accession numbers are provided at the end of this manuscript.

ChIP-seq data analysis

We generated summarized quality peak lists for each genotype and treatment. Each list contains the peaks that are either consistently detected in both replicates or have a high reliability (MACS fold enrichment > 5.0). For peaks detected in both replicates the maximum peak breadth and tag density was taken. Subsequently, we identified peaks within the promoter region of genes (defined as transcriptional start site, \pm 1000 bases) and the gene bodies (defined as the region between the transcriptional start site and the last base of the gene). The RefSeq gene reference lists for mouse and human were taken from UCSC Table Browser. To avoid gender-specific artifacts, we excluded genes and peaks of the X- and Y-chromosome from the analysis. If a peak had a minimum overlap of 5% with a promoter region and/or a gene body, the respective gene-associated histones were considered to carry the modification in a binary present or non-present manner. Mouse and human data were compared after filtering for conserved genes using BioMarts [Ensembl release 95] [24]. Gene set enrichment analysis was performed using the gene ontology enrichment analysis tool PANTHER Version 12.0 [25, 26]. Data analysis and images were conducted using the statistic software R [3.4.4] (<https://www.r-project.org/>) [27]. In addition, the data were analyzed by the self-organizing map (SOM)-technique using the R package oposSOM [2.0.0] [28].

Microarray analysis

mRNA from intestine samples of a 4-month-old radiated and untreated *Msh2*^{+/+} and *Msh2*^{-/-} mice was analyzed with the Bead Chip Array MouseRef-8 v2 (Illumina, San Diego, USA) [7]. Data were selected that had a *p* value < 5% from the background corrected raw data and subject to quantile normalization. Further analysis was performed together with similar data on 12-month-old

untreated *Msh2*^{+/+} and *Msh2*^{-/-} mice and tumors harvested from the *Msh2*^{-/-} mice [7] using the SOM-technique (see above).

Results

To characterize the histone modification profiles of the intestine of 4-month-old mice, we measured the histone marks H3K4me3, H3K27me3, H3K36me3, and control H3pan [29] of four doublets of mice. The study design is summarized in Fig. 1a. The number of mapped high-quality reads is shown in Fig. 1b. Based on these data, we constructed summary lists of the genomic regions in which modified histones were detected for each modification and mouse pair. The modified regions are hereinafter referred to as peaks. The number of peaks measured in the individual mice is shown in Fig. 1c. For further analysis, quality peaks have been selected (see “Materials and Methods” trump). Details about the peak breadths and read densities are provided in the Additional file 2.

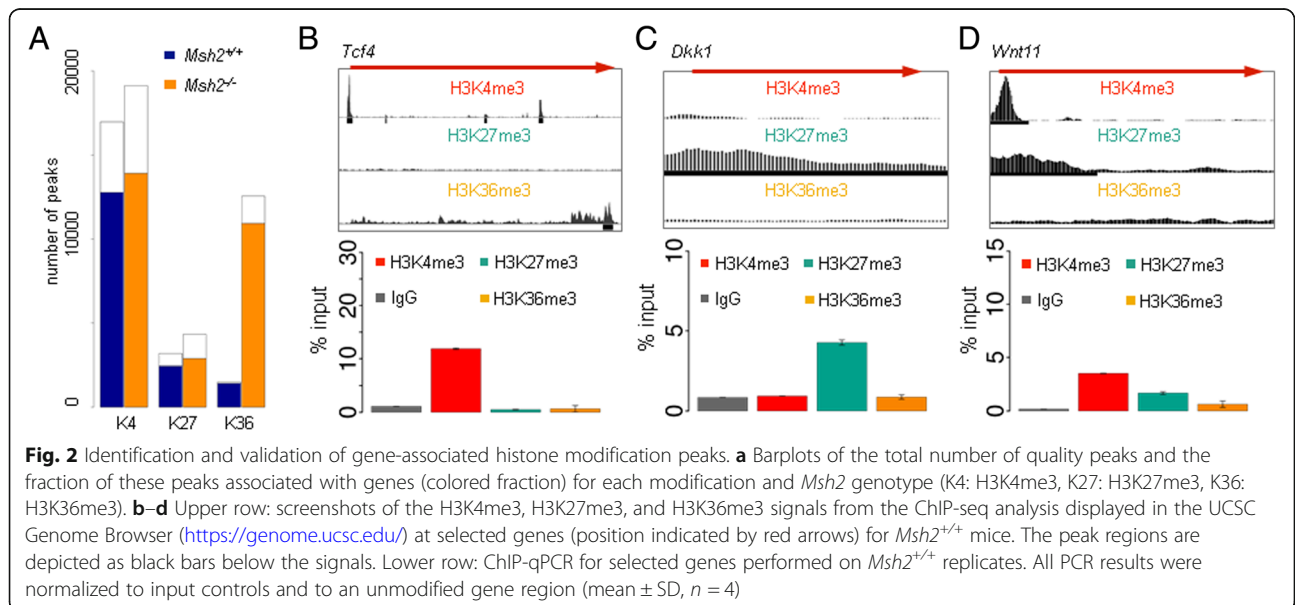
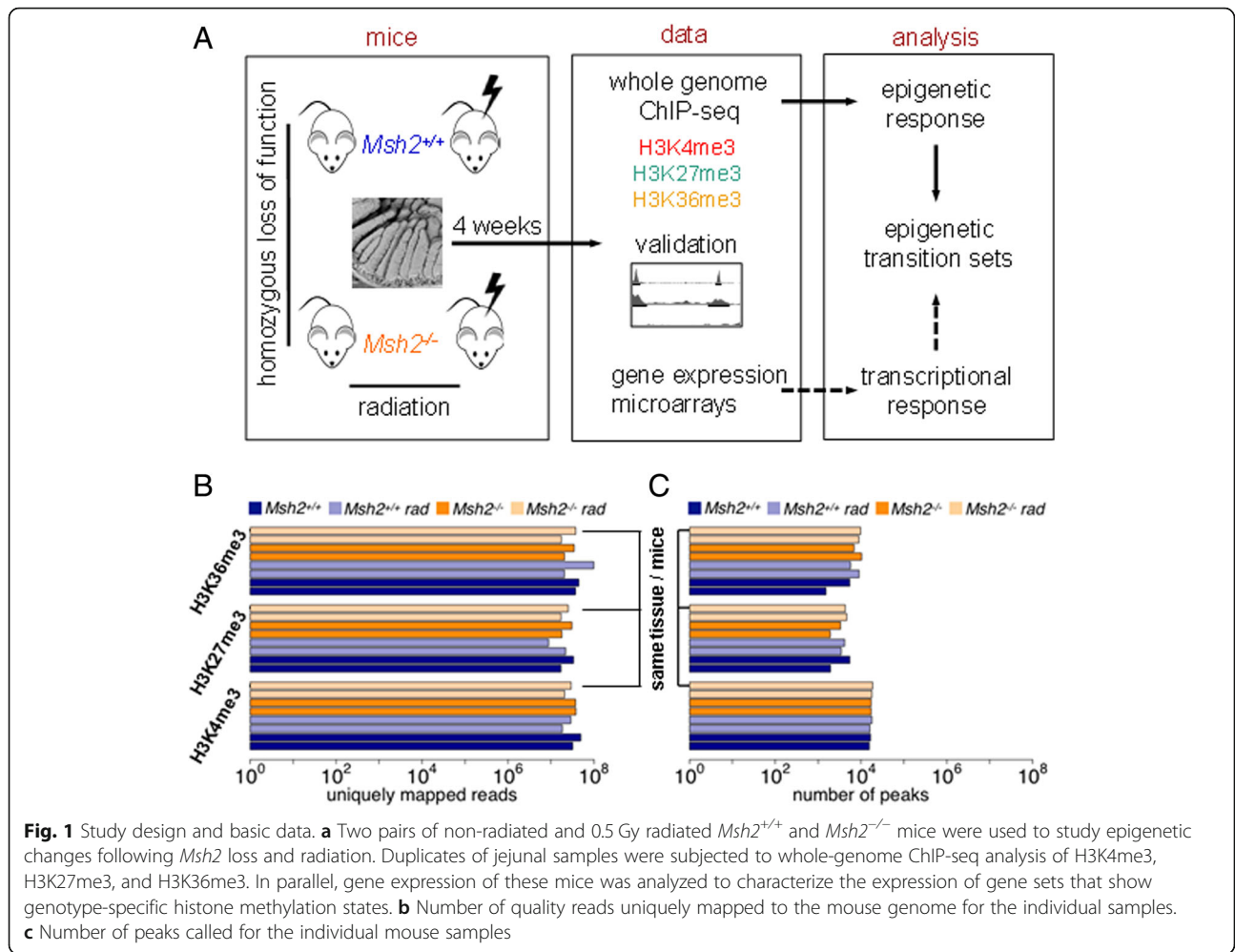
Histone methylation is enhanced in *Msh2*^{-/-} mice

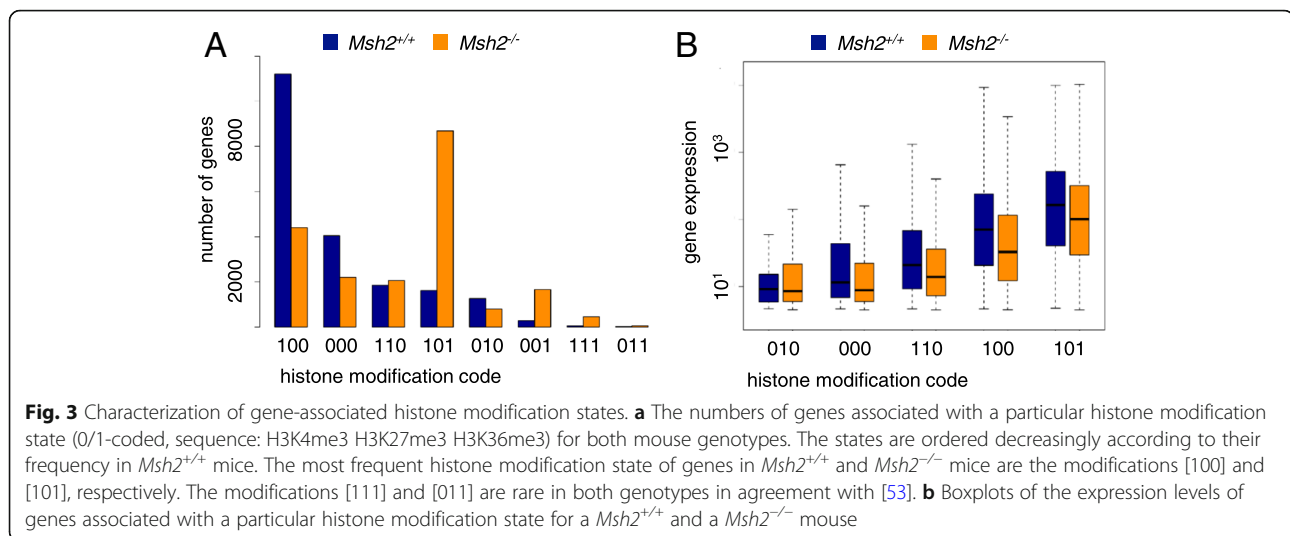
In total, *Msh2*^{-/-} mice show a higher number of peaks, i.e., methylated histones, compared to control mice (peak numbers: 35982 *Msh2*^{-/-}, 21650 *Msh2*^{+/+}). Independent of the mouse genotype most of the analyzed histone marks are associated with genes (Fig. 2a), i.e., their peaks overlap with promoter and/or gene body regions (91.7% *Msh2*^{-/-}, 72.9% *Msh2*^{+/+}). The total numbers of H3K4me3 and H3K27me3 gene-associated peaks are similar, while the number of H3K36me3 modified genes is higher in *Msh2*^{-/-} mice compared to *Msh2*^{+/+} mice.

To validate our results, we examined the histone modification state of selected genes using ChIP-qPCR. Examples are shown in Fig. 2b–d (*Msh2*^{+/+} mice) and Additional file 3 (*Msh2*^{-/-} mice). We confirmed exclusive enrichment for H3K4me3 in the promoter region of *Tcf4*, a transcriptional effector of Wnt signaling (Fig. 2b), and for H3K27me3 in the promoter region of *Dkk1*, a Wnt antagonist (Fig. 2c). *Wnt11* was found to be bivalent H3K4me3-H3K27me3 modified (Fig. 2d).

Genes exhibit distinctive histone modification states in *Msh2*^{-/-} and *Msh2*^{+/+} mice

Because multiple histone methylations can cause combinatorial effects [30], we determined the histone modification state of all genes. Therefore, each gene obtained a 0/1-triplet code depending on whether the associated histones carry a particular modification (1) or not (0). The code sequence is [H3K4me3 H3K27me3 H3K36me3]. *Msh2*^{-/-} clearly differed from *Msh2*^{+/+} mice in their histone modification profiles. The numbers of genes associated with a particular histone modification state are shown in Fig. 3a. Genes with a signature of





actively and stably transcribed genes, i.e., H3K4me3-H3K36me3 modified genes (signature [101]) are increased in *Msh2*^{-/-} compared with *Msh2*^{+/+} mice (42.71% vs. 7.95%).

Correlating the histone modification state with the expression of genes, we confirmed well-known relationships (Fig. 3b). The highest mean expression is observed for H3K4me3-H3K36me3 genes [101]. Genes with promoters carrying an H3K4me3 modification [100] show a higher expression than genes with an H3K27me3 modified promoter [010]. Unmodified [000] and bivalent H3K4me3-H3K27me3 genes [110] have an intermediate expression [31].

Large sets of genes become epigenetically activated in *Msh2*^{-/-} mice

Among the genes that are differentially modified between both genotypes (10,767 genes), in *Msh2*^{+/+} mice genes with no modification [000] or H3K4me3 only [100] are the most frequent, whereas in *Msh2*^{-/-} mice H3K4me3-H3K36me3 [101] modified genes are most common (Fig. 4a). To provide information on the number and direction of gene-specific epigenetic changes between *Msh2*^{+/+} and *Msh2*^{-/-} mice, we calculated a transition matrix (Fig. 4b). We arranged the five most frequent histone modification states in this matrix ascending from epigenetic repression [010] to stable activation [101]. Each square in the matrix comprises genes which show the same histone modification change between *Msh2*^{+/+} and *Msh2*^{-/-} mice. In the following, we refer to these genes as “transition sets” denoted by XYZ → X'Y'Z'.

We detected two large transition sets by comparing *Msh2*^{+/+} and *Msh2*^{-/-} mice. H3K36me3 recruitment is frequently observed in genes that are associated with histones carrying only an H3K4me3 mark in *Msh2*^{+/+} mice (100 → 101, 6931 genes; transition gene set 1 for

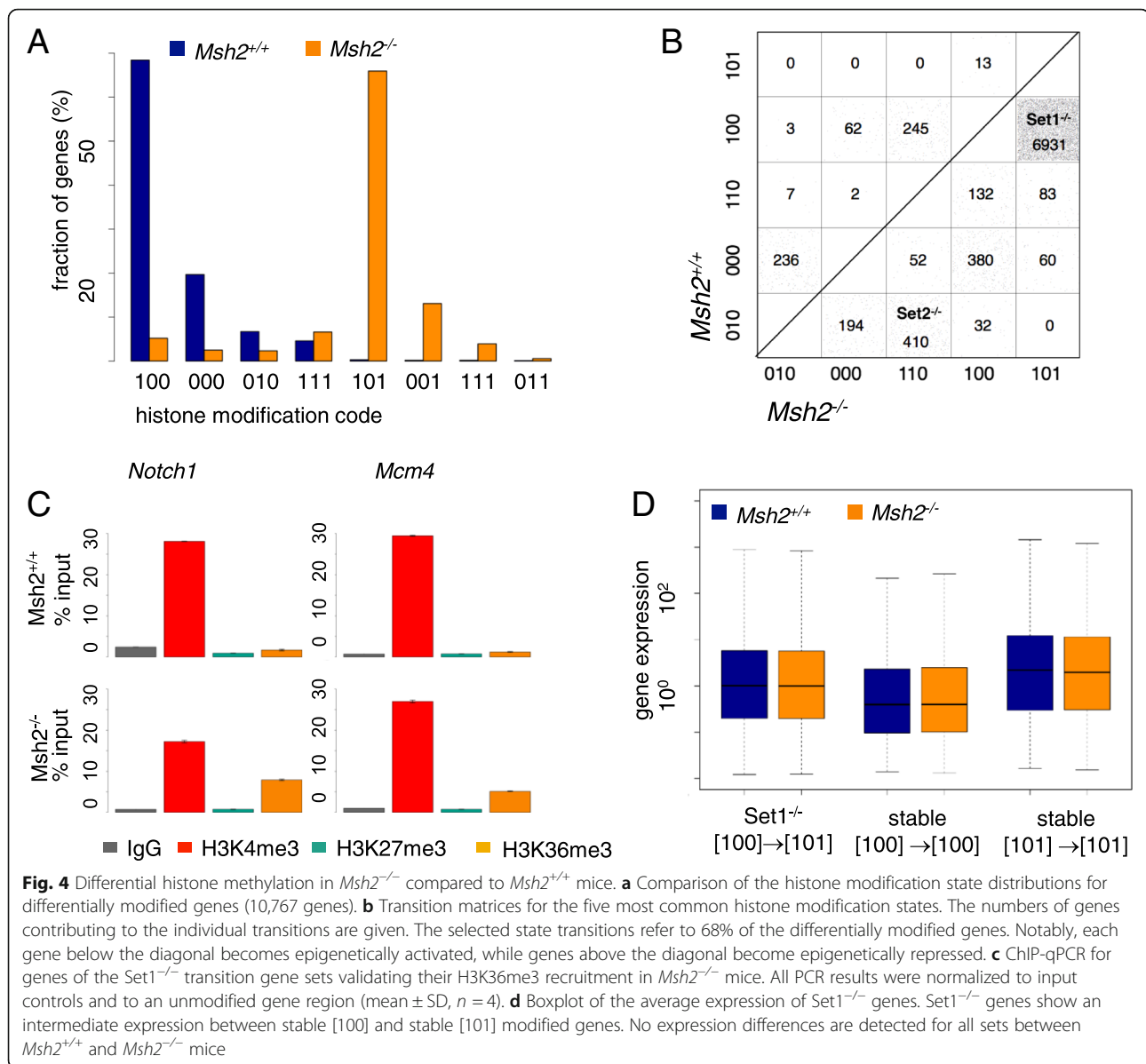
transition *Msh2*^{+/+} to *Msh2*^{-/-}, in short Set1^{-/-}), while recruitment of H3K4me3 is mainly detected for genes associated with histones carrying only a H3K27me3 modification in *Msh2*^{+/+} mice (010 → 110, 410 genes; Set2^{-/-}).

For selected genes the histone modification differences were exemplarily validated using ChIP-qPCR (Fig. 4c). As in ChIP-seq, *Mcm4* and *Notch1* show an increased level of H3K36me3 in *Msh2*^{-/-} compared to *Msh2*^{+/+} mice. These genes are essential for the initiation of replication and lineage specification of the progeny of ISC, respectively.

Because recruitment of the H3K36me3 histone methyltransferase Setd2 is RNA polymerase II-dependent [32], we tested whether Set1^{-/-} genes are differentially expressed between *Msh2*^{+/+} and *Msh2*^{-/-} mice. Surprisingly, the expression of these genes is similar suggesting another way of H3K36me3 enrichment (Fig. 4d). Set1^{-/-} genes are more highly expressed in the *Msh2*^{+/+} intestine than those which are [100] modified in both mice. Due to the low expression of H3K27me3 genes, we could not verify Set2^{-/-} genes for their expression.

A single-radiation hit causes an epigenetic activation in *Msh2*^{+/+} mice similar to *Msh2*-loss

To study the long-term epigenetic response of the intestine to radiation, we determined the histone modification profiles 4 weeks after a single-radiation hit of 0.5 Gy. The radiation hit induced phosphorylation of p53 at Ser15 (Additional file 4), a common marker of double-strand break response [33]. We found 10,389 differentially modified genes in radiated *Msh2*^{+/+} mice (Fig. 5a). Most of these genes either recruit H3K36me3 to H3K4me3 target genes (100 → 101, 6306 genes; Set1^{rad}) or H3K4me3 to H3K27me3 target genes (010 → 110, 508 genes, Set2^{rad}). In addition, unmodified genes recruit H3K4me3 (000 → 100, 814 genes) (Fig. 5b).



Surprisingly, radiation of *Msh2*^{-/-} mice induces fewer changes (3495 genes, Fig. 5c) without a prominent state transition (Fig. 5d). We confirmed the histone modification states for selected genes of untreated and radiated mice using ChIP-qPCR (Additional file 3).

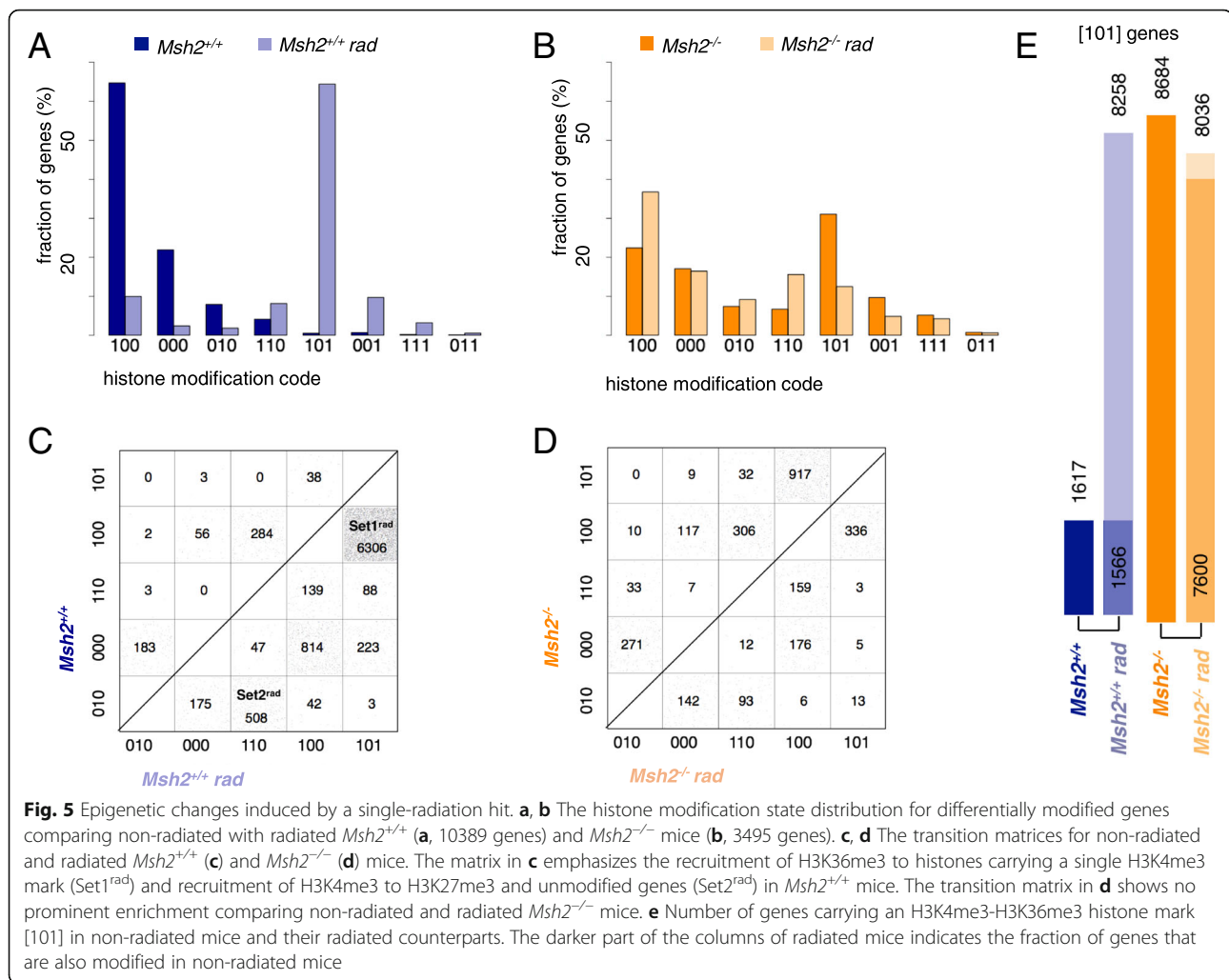
The average expression of Set1^{rad} genes was not increased in radiated *Msh2*^{+/+} compared to untreated *Msh2*^{+/+} mice. Similar to Set1^{-/-} genes, a higher expression of Set1^{rad} genes was detected compared to [100] modified genes that remain [100] modified in radiated mice (Additional file 4).

A common epigenetic response

Counting the number of genes with H3K4me3-H3K36me3 histone marks [101] in all mice (Fig. 5e), we found similar

numbers of [101] modified genes in radiated *Msh2*^{+/+} (8258) and *Msh2*^{-/-} (8036) mice. Moreover, 7600 out of the 8036 [101] modified genes observed in radiated *Msh2*^{-/-} mice are already [101] modified in untreated *Msh2*^{-/-} mice. Together, these results suggest that *Msh2* loss and a single-radiation hit cause a common epigenetic response.

Focusing on genes which belong to Set1, which acquire H3K36me3 marks at promoters already modified by H3K4me3, we found that after *Msh2* loss or radiation a large gene set is affected under both conditions (Fig. 6a). In fact, 5984 genes are part of both Set1 (100 → 101) transition sets. Similarly, we found 357 genes to be part of both Set2 (010 → 110) transition sets. Notably, for a small set of genes, radiation-specific



epigenetic activation is seen. Details are provided in the Additional file 5.

Functional annotation of the common epigenetic response

To test whether Set1 and Set2 genes are associated with functional gene sets, we performed a gene ontology (GO) enrichment analysis [26, 34]. In Set1, we found enrichment in gene sets linked, among others, to ribosomal biogenesis, translation, RNA metabolic processes, mitotic cell cycle, DNA damage checkpoints, and repair mechanisms (Additional file 6). These are enrichments typically found for H3K4me3 target genes, i.e. highly expressed genes. The genes of Set2 enrich in gene sets linked to developmental processes, cell differentiation, and fate commitment, as expected for H3K27me3 target genes (Additional file 6) [35]. Among these genes are many *Wnt*, *Fox*, *Sox*, *Pax*, and *Fgf* family members and components of receptor and channel complexes. A heatmap using a Gene Set Z-scoring function (GSZ score) of modification [28] demonstrates that both Set1 and Set2 genes dominate

entire functional gene sets (Fig. 6b). Details on selected gene sets are provided in the Additional file 5.

Subsequently, we analyzed, whether the gene sets dominated by Set1 and Set2 genes are transcriptionally regulated in the macroscopic normal intestine during aging and tumorigenesis. For this purpose a SOM-analysis was performed based on gene expression data of 4- and 12-month-old mice. The data of the older mice including tumor data have been published recently [7]. Comparing the SOM portraits, a differential activation of genes in 4- and 12-month-old mice and tumors is observed (Fig. 7). Actually, several of the gene sets that are dominated by Set1 and Set2 genes are among those becoming upregulated in tumors. This is exemplified for the genes of the biological process (BP) set “DNA repair” and the cellular component (CC) set “ribosome biogenesis”.

Set1 and Set2 genes can be predicted based on their molecular characteristics

The validity of the identified histone modification changes depends in particular on the epigenetic states determined

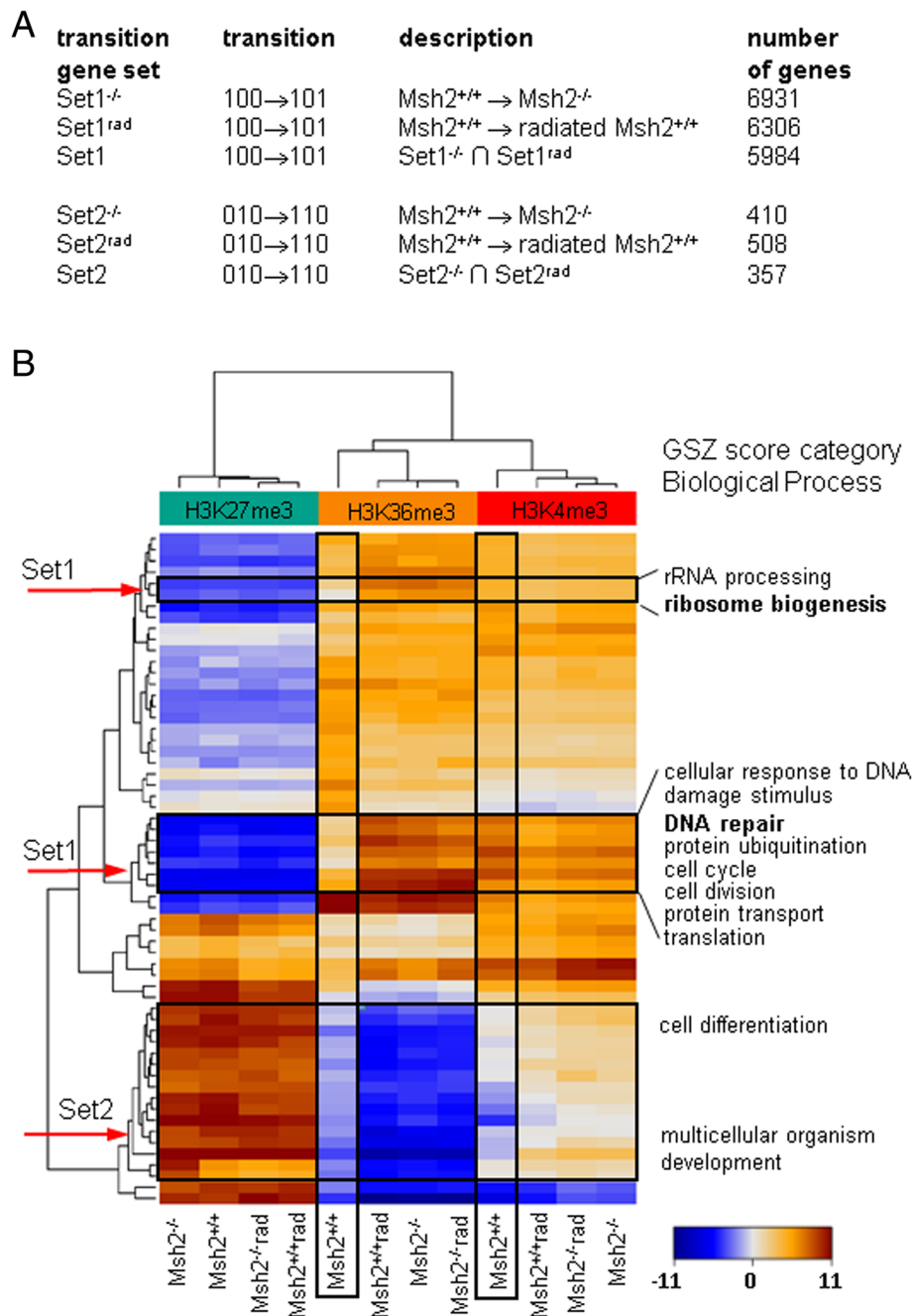


Fig. 6 Features of the common epigenetic response. **a** Epigenetic transition gene sets. The individual sets are given together with their intersection (*n*) and the numbers of genes belonging to them. **b** Heatmap of the Gene Set Z-scoring function (GSZ score) of modification for selected gene sets. These sets are enriched by genes contributing to the overexpression spots of the SOM calculated using the histone modification data (see Additional file 5). Entire sets are dominated by Set1 or Set2 genes (red arrows)

in untreated control *Msh2*^{+/+} mice. Thus, we validated histone modifications of Set1 and Set2 genes in further controls by comparing our results with tissue data from the mouse ENCODE consortium [21]. As these experiments measured the histone modifications with different sensitivity, we compared equal numbers of best quality, gene-associated peaks. We found 4893 (81.9%) out of the

Set1 genes to be [100] modified and 267 (74.8%) out of the Set2 genes to be [010] modified, supporting our results. In addition, we compared our data with H3K4me3 and H3K27me3 data on isolated mouse intestinal epithelial cells provided by Kazakevych et al. [22] including data on cells of embryonic days E12.5 and E14.5, ISC and adult enterocytes (AE). Modification profiles seen in ISC and

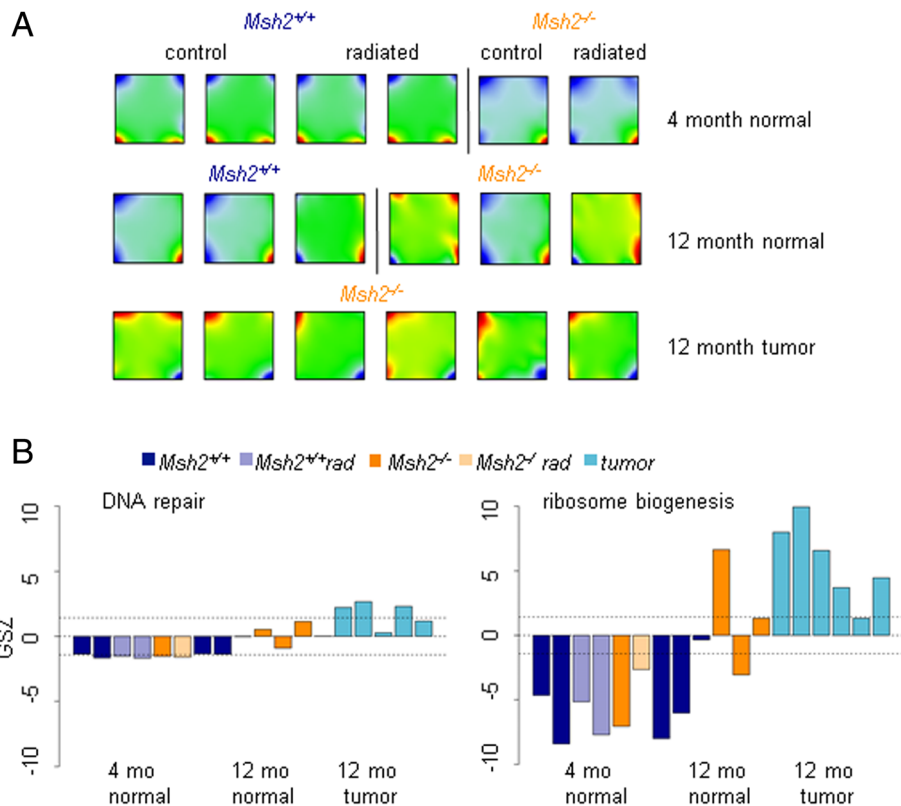


Fig. 7 SOM-analysis of gene expression data. **a** SOM-portraits of gene expression in normal intestinal samples of untreated or radiated 4- and 12-month-old mice and tumors of 12-month-old *Msh2*^{-/-} mice. Color codes for gene expression which was normalized to the maximum (red) and minimum (blue) expression within the individual sample. **b** GSZ score of expression for the “DNA repair” and “ribosome” gene sets. Following the loss of *Msh2* or a single-radiation hit, the epigenetic profile of these sets is dominated by Set1 genes (Fig. 6b, Additional file 5). An increased variability of the expression is seen in 12- compared to 4-month-old mice. Gene expression is strongly upregulated in tumors

AE are very similar to our tissue profiles, while those from embryonic cells deviate (Additional file 7). We conclude that the profiles identified for Set1 and Set2 genes are characteristic for absorptive cells of the adult mouse intestine.

Subsequently, we assigned Set1 and Set2 genes further molecular characteristics, reanalyzing methyl-CpG-binding domain (MBD) and RNA-seq data provided by Kazakevych et al. [22]. We found that, in isolated ISC and AE, Set1 (Set2) genes show relatively low promoter DNA methylation compared to genes [100] ([010]) modified in both untreated *Msh2*^{+/+} and *Msh2*^{-/-} mice (Additional file 8). Moreover, the genes of both sets are characterized by a specific gene body DNA methylation. As in our tissue samples, Set1 genes are higher expressed compared to genes [100] modified in both untreated *Msh2*^{+/+} and *Msh2*^{-/-} mice (Additional file 8). Among all [100] genes, the 1000 genes with the highest (lowest) expression in ISC cover 693 (119) Set1 genes which is a 1.23 (0.21)-fold enrichment (depletion). Together, these

results demonstrate that Set1 and Set2 genes carry specific properties that allow their prediction based on data from control intestinal tissue.

This led us to analyze the conservation of these properties in human. We measured histone modifications in normal human colon and combined these data with published matched RNA-seq data [3]. From the 5984 Set1 mouse genes, we found 5065 (85%) to be conserved in human. Among them, 49.4% show a [100] profile as in mice but 32.4% a [101] profile (Additional file 9). Assuming a comparable sensitivity of our measurements in mouse and human, this suggests an already present epigenetic activation in human. From the 357 Set2 mouse genes, we found 253 (71%) to be conserved in human. Among them, only 28.3% show a [010] profile as in mice and 34.6% a [110] profile. Thus, an epigenetic activation of these genes might be present as well.

Similar to our observations in *Msh2*^{+/+} mice, expression of Set1 genes in the human colon is above the expression of those [100] genes not contributing to Set1.

Discussion

Here, we show that macroscopic normal intestinal tissue of 4-month-old *Msh2*^{-/-} mice exhibit genome-wide epigenetic changes compared to age-matched control mice although intestinal tumors do not develop in *Msh2*^{-/-} mice until the age of 12 months [7]. Thus, tumorigenesis under MMR deficiency occurs on a disturbed epigenetic background. Comparing non-radiated with radiated *Msh2*^{+/+} mice, differences in the histone H3 methylation profiles are very similar to those induced by *Msh2* loss, suggesting a common epigenetic response which is most likely part of a response to genomic stress at the cell level. It is unlikely that the same large gene sets are damaged in both experimental settings, MMR deficiency, and a single-radiation hit.

Questions arise on the mechanisms controlling this common epigenetic response. Spontaneous mutation rates in MMR-deficient tissue are much higher than in control tissue. In the human colon, mismatch mutations increase immediately after the loss of *MLH1* and *MSH2* [4]. In non-conditional *Msh2*-knockout mice mutations rates at endogenous expanded simple tandem repeat (ESTR) DNA loci rise in the germline [36]. Comparable to our study, a single-radiation hit of these mice did not further increase the ESTR mutation rate, whereas a significant induction of mutations was observed in radiated controls. Burr et al. hypothesized that high cell killing in radiated MMR-deficient tissue explains the phenomenon [36]. A similar explanation might also hold true for our observed epigenetic changes. We suggest that the common epigenetic response results from a selection of ISC tolerant to genomic stress. Less tolerant ISC either undergo apoptosis or are, in agreement with experimental findings on radio-resistance of ISC, outcompeted in each individual crypt [37, 38]. This assumption consistently explains the epigenetic response in *Msh2*^{-/-} mice on radiation, where the selection of ISC might have already taken place. In agreement with a response to genomic stress at the cell level, a short-term epigenetic response following double-strand repair does not comprise changes in H3K36me3 [39, 40], which are obviously present 4 weeks after radiation or after *Msh2* loss.

A competitive advantage of ISC subsequent to the epigenetic activation in Set1 and Set2 genes may be explained as follows:

(1) H3K36me3 accumulates most frequently at [100] genes, i.e., genes that are H3K4me3 modified in control mice. These genes are important for ribosomal biogenesis, translation, and RNA processing. An impairment of any of these cellular processes can severely retard cell growth and perturb mammalian development [41]. H3K36me3 is required for proper recruitment of the MMR machinery [42] and homologous recombination (HR) [43], and it regulates the mode of double-strand

break repair in favor of non-homologous end joining (NHEJ) [44], the preferred repair mode in the intestine [45]. There is a strong inverse correlation between normalized H3K36me3 level and mutational frequency [46]. Thus, higher H3K36me3 at Set1 genes might ensure higher fidelity of ISC. In the absence of Setd2 and thus of H3K36me3, the capacity for clonogenic survival following DNA damage is significantly reduced [39].

(2) Accumulation of H3K4me3 was predominantly observed for [010] genes, i.e., genes that are repressed by H3K27me3 in control mice. Among these genes are many that control developmental processes and differentiation [25]. Upregulation of such genes might be important for responding to regenerative demands [47]. Here, the recruitment of H3K4me3 in Set2 genes might represent a mechanism to protect their promoters from DNA methylation and thus to ensure potential gene activation according to the following mechanism: DNA damage triggers promoter DNA methylation of H3K27me3 target genes [48]. Recruitment of H3K4me3 to H3K27me3-labeled gene loci can prevent DNA methylation in *Msh2*^{-/-} mice [49]. Consistently, loss of *Msh2* prevented promotor CpG island hypermethylation observed in inflammation-induced tumors [50].

Based on our data, we cannot exclude that the epigenetic response is an intrinsic regulation within all cells. Moreover, it remains unclear, how the enrichment of activating histone methylation marks is achieved and stabilized. We did not find an increased gene expression paralleling the broad H3K36me3 enrichment. Thus, changes in the activity of histone methyltransferases and/or demethylases of H3K36 are likely.

Until now, the epigenetic response following DNA damage has been predominantly studied within a short-term window (reviewed in [51, 52]). Here, we provide evidence that loss of *Msh2* and a single-radiation hit induce common epigenetic changes that persist for long time scales. We expect that the advantages of an epigenetic activation of the affected gene sets are accessed also by tumor cells. *Msh2*^{-/-} tumors show a transcriptional activation of many of the affected genes [7]. This suggests that the observed epigenetic changes protect DNA loci encoding essential cellular functions also in tumor cells. Indeed, tumorigenesis is often controlled by regulating ribosome biogenesis and global protein synthesis [41]. For about half of the Set1 genes, histone profiles and relative transcription levels seen in control mice are conserved in human. Thus, we expect similar response scenarios also in the human intestine. In fact, mismatch repair-deficient Lynch colorectal carcinoma with high mutational load show overexpression of gene sets associated with the GO gene sets “translation,” “ribosome,” and “cell cycle” [3] which we found

to be associated with Set1 genes. It would be interesting to see whether targeting H3K36me3 by SetD2 inhibitors selectively affects tumor cell survival in Lynch colorectal carcinoma.

In summary, MMR deficiency and a single radiation hit result in common genome-wide changes of histone H3 methylation profiles.

Additional files

Additional file 1: Primer sequences. List of all primers used in the study. (XLSX 10 kb)

Additional file 2: Peak characteristics of histone methylations of mouse intestinal samples [54]. (DOCX 543 kb)

Additional file 3: ChIP-qPCR for selected genes. (DOCX 114 kb)

Additional file 4: Details on the radiation response [33]. (DOCX 196 kb)

Additional file 5: Results of the SOM analysis of the epigenetic profiles. SOM portraits and additional information about affected gene sets. (DOCX 389 kb)

Additional file 6: Results of the GO-analysis for Set1 and Set2 genes. (XLSX 15 kb)

Additional file 7: Comparison of epigenetic states in intestinal tissue and isolated intestinal cells [22]. (DOCX 515 kb)

Additional file 8: Molecular characterization of Set1 and Set2 genes in isolated intestinal cells. Published MBD- and RNA-seq data on Set1 and Set2 genes [22]. (DOCX 73 kb)

Additional file 9: Set1 and Set2 genes in mouse and human intestinal tissue. Comparison of Set1 and Set2 genes in mouse and human tissue [3]. (DOCX 888 kb)

Additional file 10: Modification profiles of human genes. The list includes those genes that show a modification in at least one of the lysines K4, K27 or K36 of histone 3 only. (XLSX 3040 kb)

Abbreviations

AE: Adult enterocytes; ChIP: Chromatin immunoprecipitation; ISC: Intestinal stem cell; MMR: Mismatch repair; MSI: Microsatellite instability; SOM: Self-organizing map

Funding

This work was supported by the BMBF grant INDRA [grant number: 031A312; JG, MH, MQ, TT], by the Volkswagenstiftung [Lichtenberg program; MRS, SS], the Center for Molecular Medicine [MHU], and the German research foundation [DFG SCHW1605/1-1].

Availability of data and materials

The mouse jejunum ChIP-seq data is deposited in the Gene Expression Omnibus (GEO) repository (<https://www.ncbi.nlm.nih.gov/geo/>) under the series accession number GSE128842. The human rectum raw ChIP-seq and gene expression data are available on request. The human modification summary is available in the Supplement (Additional file 10).

Public available datasets

The sequencing data supporting the conclusions of this article were obtained from the Gene Expression Omnibus (GEO) repository (<https://www.ncbi.nlm.nih.gov/geo/>) and the Encyclopedia of DNA Elements (ENCODE) [21] repository (<https://www.encodeproject.org/>). The data of Kazakevych et al. [22] are deposited in GEO under the series accession number GSE89684. From ENCODE, we obtained mm9 alignment data associated with the accession numbers: ENCSR000CCS, ENCSR000CCT, ENCSR000CFR, ENCSR000CFS.

Authors' contributions

MH contributed to the data processing, analysis, interpretation, visualization, and manuscript writing. SS contributed to the data acquisition and analysis.

MQ and MRS contributed to the data acquisition, analysis, and interpretation. TT contributed to the data analysis and model simulation. KR, MHU, and JA contributed to the data acquisition. CK was responsible for the comprehensive technical support. JG helped in the study conception and design, data analysis and interpretation, and manuscript writing. GA helped in the data analysis and interpretation, manuscript writing, and resources. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki. We obtained ethics approval from the Landesdirektion Leipzig (TVV 53/14) for radiation of mice. Studies using human intestinal sample were approved by the ethics committees of the Universities of Duesseldorf, Bochum, and Bonn [3].

Consent for publication

All authors have approved the manuscript for submission.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Interdisciplinary Center for Bioinformatics (IZBI), Leipzig University, Leipzig, Germany. ²Laboratory for Translational Epigenetics and Tumor Genetics, University Hospital Cologne, Cologne, Germany. ³Graduate School for Biological Sciences (GSFB), University of Cologne, Cologne, Germany. ⁴Department of Surgery, Research Laboratories, Leipzig University, Liebigstr. 19, D-04103 Leipzig, Germany. ⁵Laboratory for Clinical and Experimental Hepatology (LCEHep) Section of Hepatology, Clinic for Gastroenterology and Rheumatology, University Hospital Leipzig, Leipzig, Germany. ⁶Center for Molecular Medicine Cologne, University of Cologne, Cologne, Germany. ⁷Cologne Center for Genomics, University of Cologne, Cologne, Germany.

Received: 21 August 2018 Accepted: 25 February 2019

Published online: 27 April 2019

References

- Guillotin D, Martin SA. Exploiting DNA mismatch repair deficiency as a therapeutic strategy. *Exp Cell Res*. 2014;329:110–5.
- Poulogiannis G, Frayling IM, Arends MJ. DNA mismatch repair deficiency in sporadic colorectal cancer and Lynch syndrome. *Histopathology*. 2010;56:167–79.
- Binder H, Hopp L, Schweiger MR, et al. Genomic and transcriptomic heterogeneity of colorectal tumours arising in Lynch syndrome. *J Pathol*. 2017;243:242–54.
- Staffa L, Echterdiek F, Nelius N, et al. Mismatch repair-deficient crypt foci in lynch syndrome—molecular alterations and association with clinical parameters. *PLoS One*. 2015;10. <https://doi.org/10.1371/journal.pone.0121980> eCollection 2015.
- de la Chapelle A, Hampel H. Clinical relevance of microsatellite instability in colorectal cancer. *J Clin Oncol*. 2010;28:3380–7.
- Kucherlapati MH, Lee K, Nguyen AA, et al. An Msh2 conditional knockout mouse for studying intestinal cancer and testing anticancer agents. *Gastroenterology*. 2010;138:993–1002.
- Keysselt K, Kreutzmann T, Rother K, et al. Different in vivo and in vitro transformation of intestinal stem cells in mismatch repair deficiency. *Oncogene*. 2017;36:2750–61.
- Toh TB, Lim JJ, Chow EK. Epigenetics in cancer stem cells. *Mol Cancer*. 2017;16. <https://doi.org/10.1186/s12943-017-0596-9>.
- Kimura H. Histone modifications for human epigenome analysis. *J Hum Genet*. 2013;58:439–45.
- Michna A, Schotz U, Selmsberger M, et al. Transcriptomic analyses of the radiation response in head and neck squamous cell carcinoma subclones with different radiation sensitivity: time-course gene expression profiles and gene association networks. *Radiat Oncol*. 2016;11. <https://doi.org/10.1186/s13014-016-0672-0>.

11. Fnu S, Williamson EA, De Haro LP, et al. Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. *Proc Natl Acad Sci U S A*. 2011;108:540–5.
12. Snippet HJ, van der Flier LG, Sato T, et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell*. 2010;143:134–44.
13. Otsuka K, Iwasaki T. Effects of dose rates on radiation-induced replenishment of intestinal stem cells determined by Lgr5 lineage tracing. *J Radiat Res*. 2015;56:615–22.
14. Martin LM, Marples B, Coffey M, et al. DNA mismatch repair and the DNA damage response to ionizing radiation: making sense of apparently conflicting data. *Cancer Treat Rev*. 2010;36:518–27.
15. Mannefeld M, Klassen E, Gaubatz S. B-MYB is required for recovery from the DNA damage-induced G2 checkpoint in p53 mutant cells. *Cancer Res*. 2009;69:4073–80.
16. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011. *EMBnetjournal*. <https://doi.org/10.14806/ej.17.1.200>.
17. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2011. Available online at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
18. Hoffmann S, Otto C, Kurtz S, et al. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol*. 2009;5. <https://doi.org/10.1371/journal.pcbi.1000502>.
19. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc*. 2012;7:1728–40.
20. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9:R137.
21. Yue F, Cheng Y, Breschi A, et al. A comparative encyclopedia of DNA elements in the mouse genome. *Nature*. 2014;515:355–64.
22. Kazakevych J, Sayols S, Messner B, Krienke C, Soshnikova N. Dynamic changes in chromatin states during specification and differentiation of adult intestinal stem cells. *Nucleic Acids Res*. 2017;45:5770–84.
23. Thalheim T, Hopp L, Binder H, Aust G, Galle J. On the cooperation between epigenetics and transcription factor networks in the specification of tissue stem cells. *Epigenomes*. 2018;2. <https://doi.org/10.3390/epigenomes2040020>.
24. Kinsella RJ, Kahari A, Haider S, et al. Ensembl BioMart: a hub for data retrieval across taxonomic space. *Database (Oxford)*. 2011;2011:bar030.
25. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev*. 2006;20:1123–36.
26. Mi H, Huang X, Muruganujan A, et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. 2017;45:D183–9.
27. R Core Team. A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2014.
28. Loffler-Wirth H, Kalcher M, Binder H. oposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor. *Bioinformatics*. 2015;31:3225–7.
29. Flensburg C, Kinkel SA, Keniry A, Blewitt ME, Oshlack A. A comparison of control samples for ChIP-seq of histone modifications. *Front Genet*. 2014;5. <https://doi.org/10.3389/fgene.2014.00329>.
30. Zhang T, Cooper S, Brockdorff N. The interplay of histone modifications - writers that read. *EMBO Rep*. 2015;16:1467–81.
31. Mikkelsen TS, Ku M, Jaffe DB, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*. 2007;448:553–60.
32. Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol*. 2005;25:3305–16.
33. Wittlinger M, Grabenbauer GG, Sprung CN, Sauer R, Distel LV. Time and dose-dependent activation of p53 serine 15 phosphorylation among cell lines with different radiation sensitivity. *Int J Radiat Biol*. 2007;83:245–57.
34. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25:25–9.
35. Boyer LA, Plath K, Zeitlinger J, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*. 2006;441:349–53.
36. Burr KL, van Duyn-Goedhart A, Hickenbotham P, Monger K, van Buul PP, Dubrova YE. The effects of MSH2 deficiency on spontaneous and radiation-induced mutation rates in the mouse germline. *Mutat Res*. 2007;617:147–51.
37. Roche KC, Gracz AD, Liu XF, Newton V, Akiyama H, Magness ST. SOX9 maintains reserve stem cells and preserves radioresistance in mouse small intestine. *Gastroenterology*. 2015;149:1553–63.
38. Otsuka K, Suzuki K, Fujimichi Y, Tomita M, Iwasaki T. Cellular responses and gene expression profiles of colonic Lgr5+ stem cells after low-dose/low-dose-rate radiation exposure. *J Radiat Res*. 2018;59. <https://doi.org/10.1093/jrr/rx078>.
39. Carvalho S, Vitor AC, Sridhara SC, et al. SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. *Elife*. 2014;3. <https://doi.org/10.7554/eLife.02482>.
40. Cao LL, Wei F, Du Y, et al. ATM-mediated KDM2A phosphorylation is required for the DNA damage repair. *Oncogene*. 2016;35:301–13.
41. Zhou X, Liao WJ, Liao JM, Liao P, Lu H. Ribosomal proteins: functions beyond the ribosome. *J Mol Cell Biol*. 2015;7:92–104.
42. Li F, Mao G, Tong D, et al. The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha. *Cell*. 2013;153:590–600.
43. Pfister SX, Ahrabi S, Zalmas LP, et al. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep*. 2014;7:2006–18.
44. Pai CC, Deegan RS, Subramanian L, et al. A histone H3K36 chromatin switch coordinates DNA double-strand break repair pathway choice. *Nat Commun*. 2014;5. <https://doi.org/10.1038/ncomms5091>.
45. Haines JW, Coster M, Bouffler SD. Impairment of the non-homologous end joining and homologous recombination pathways of DNA double strand break repair: impact on spontaneous and radiation-induced mammary and intestinal tumour risk in Apc min/+ mice. *DNA Repair*. 2015;35:19–26.
46. Mar BG, Chu SH, Kahn JD, et al. SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia. *Blood*. 2017;130:2631–41.
47. Tetteh PW, Basak O, Farin HF, et al. Replacement of lost Lgr5-positive stem cells through plasticity of their enterocyte-lineage daughters. *Cell Stem Cell*. 2016;18:203–13.
48. Schlesinger Y, Straussman R, Keshet I, et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat Genet*. 2007;39:232–6.
49. Thalheim T, Herberg M, Galle J. Linking DNA damage and age-related promoter DNA hyper-methylation in the intestine. *Genes*. 2018;9. <https://doi.org/10.3390/genes9010017>.
50. Maiuri AR, Peng M, Sriramkumar S, et al. Mismatch repair proteins initiate epigenetic alterations during inflammation-driven tumorigenesis. *Cancer Res*. 2017;77:3467–78.
51. Dantuma NP, van Attikum H. Spatiotemporal regulation of posttranslational modifications in the DNA damage response. *EMBO J*. 2016;35:6–23.
52. Chen Y, Zhu WG. Biological function and regulation of histone and non-histone lysine methylation in response to DNA damage. *Acta Biochim Biophys Sin*. 2016;48:603–16.
53. Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem*. 2011;286:7983–9.
54. Wong SH, Goode DL, Iwasaki M, et al. The H3K4-methyl epigenome regulates leukemia stem cell oncogenic potential. *Cancer Cell*. 2015;28:198–209.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

