Chronic inflammation induces a novel epigenetic program that is conserved in intestinal adenomas and in colorectal cancer

Monther Abu-Remaileh, Sebastian Bender, Günter Raddatz, Ihab Ansari, Daphne Cohen, Julian Gutekunst, Tanja Musch, Heinz Linhart, Achim Breiling, Eli Pikarsky, Yehudit Bergman, and Frank Lyko

1Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel-Canada, Hebrew University Medical School, Jerusalem, Israel; 2Division of Epigenetics, DKFZ-ZMBH Alliance, German Cancer Research Center, Heidelberg, Germany; 3Salem Medical Center, University of Heidelberg, Heidelberg, Germany; 4Department of Hematology/Oncology, University of Heidelberg Medical Center, Heidelberg, Germany; 5The Lautenberg Center for Immunology, Institute for Medical Research Israel-Canada, Hebrew University Medical School, Jerusalem, Israel; 6Co-first author; 7Co-senior author

Corresponding authors: Yehudit Bergman
Institute of Medical Research Israel-Canada
Hebrew University Medical School, Jerusalem, Israel
phone: ++972-2-6758362, fax: ++972-2-6757482
email: yehuditb@ekmd.huji.ac.il

Frank Lyko
Deutsches Krebsforschungszentrum (DKFZ)
Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
phone: ++49-6221-423800, fax: ++49-6221-423802
email: f.lyko@dkfz.de

Disclosure of Potential Conflicts of Interest: No potential conflicts of interest were disclosed.

Running title: DNA methylation links inflammation and cancer
Summary

Chronic inflammation represents a major risk factor for tumor formation, but the underlying mechanisms have remained largely unknown. Epigenetic mechanisms can record the effects of environmental challenges on the genome level, and could therefore play an important role in the pathogenesis of inflammation-associated tumors. Using single-base methylation maps and transcriptome analyses of a colitis-induced mouse colon cancer model, we identified a novel epigenetic program that is characterized by hypermethylation of DNA methylation valleys that are characterized by low CpG density and active chromatin marks. This program is conserved and functional in mouse intestinal adenomas and results in silencing of active intestinal genes that are involved in gastrointestinal homeostasis and injury response. Further analyses reveal that the program represents a prominent feature of human colorectal cancer and can be used to correctly classify colorectal cancer samples with high accuracy. Together, our results show that inflammatory signals establish a novel epigenetic program that silences a specific set of genes which contribute to inflammation-induced cellular transformation.
Introduction

Inflammation has been linked to the pathogenesis of tumors in a substantial fraction of human cancers (1,2). Inflammatory processes play a causal role in cancer development through cellular pathways that involve genotoxicity, aberrant tissue repair, proliferative responses, invasion and metastasis (2). The inflammatory microenvironment is an essential component of almost all tumors, including some for which a causal relationship with inflammation is not yet proven (3). One of the most widely studied and most prevalent inflammation-induced cancers is Inflammatory Bowel Disease (IBD)-associated colorectal cancer. Patients suffering from long-standing IBD have a significantly higher risk of developing colorectal cancer (4).

Altered DNA methylation patterns have long been associated with tumor formation and represent one of the earliest and most consistent molecular markers of human cancers (5-7). Traditionally, most studies of DNA methylation have focused mainly on CpG islands which are unmethylated and generally associated with promoters. It is well known that subsets of CpG islands can become aberrantly hypermethylated in cancer (6-10). More recently, genome-wide methylation pattern analyses have begun to systematically uncover key features of tumor methylomes (8,9,11). These include not only hypermethylated CpG islands, but also large (several 100 kb to several Mb) partially methylated domains (PMDs) that are gene-poor and co-localize with lamina-associated domains (8,9). Another novel feature of methylomes that shows altered DNA methylation patterns in tumors has been termed DNA methylation valleys (DMVs). These regions extend over several kilobases of DNA, are strongly hypomethylated in most tissues and are enriched for transcription factors and developmental genes (12). DMVs have been shown to become hypermethylated in colorectal cancer and may thus contribute to the aberrant epigenetic programming of tumor cells (12). Regions similar to DMVs, termed methylation canyons, were recently identified in hematopoietic stem cells and were shown to be distinct from CpG islands and CpG island-associated features (13).
While a few initial studies have suggested aberrant methylation patterns in mouse models of IBD (14,15), both the effect of inflammation on DNA methylation and the functional outcome have not been investigated in detail yet. We have now used whole-genome bisulfite sequencing (16) to characterize the methylomes of the AOM/DSS mouse model at single-base resolution. In this model, mice are treated with dextran sodium sulfate (DSS) to induce colitis (17). When this treatment is preceded by injections of the weak carcinogen azoxymethane (AOM) the mice develop intestinal tumors (18). Our results identify hypermethylated DMVs as a prominent feature of the colitis methylome that is conserved in intestinal adenocarcinomas and in human colon cancer. Our findings provide strong support for the hypothesis that inflammatory signals induce a higher risk for cancer development by establishing a novel epigenetic program in enterocytes.

Materials and Methods

Preparation of samples
Male mice (C57BL/6) were obtained from Harlan Laboratory (Israel) at 6-7 weeks of age. The mice were left for one week for adaptation and then treated to induce inflammation and/or cancer. In the adenoma analysis both male and female mice at 4-6 months of age were used. For inflammation induction, mice were given 2% of 36-50 kDa DSS (MP Biomedicals) for one week in their drinking water, and then left for two weeks on regular drinking water. This cycle was repeated 3 times. In order to induce inflammation-associated cancer the same protocol was used but the DSS cycles were preceded by two i.p doses of 12.5 mg/kg AOM (Sigma) given with one week difference (Supplementary Fig. S1). Mice were housed and cared for under SPF conditions and all the animal procedures were approved by the Animal Care and Use Committee of the Hebrew University of Jerusalem.

For epithelial cell preparation from colon, we followed a previously published protocol (19). Briefly, the colon was flushed with ice-cold PBS, then cut open longitudinally. Distal and
proximal colon regions were processed separately. Each part was minced into small pieces in cold PBS. Enterocytes were mechanically isolated by shaking in PBS with 30 mM EDTA, at 37°C. Crypts were recovered and then stored at -80°C for DNA and RNA extraction. Only samples from the distal colon were used in our study. For histological analysis of colon and tumor samples, representative pieces of distal and proximal colon were used. Cancer cells were extracted from distal colon crypts of samples showing a high percentage of tumor tissue using histological assessments, as described above. Preparation and maintenance of organoid cultures were performed as previously reported (19,20). 4-methylpyrazol treatment was initiated directly after organoid derivation using 1 mM final concentration. Organoids were monitored daily and size was measured using NIS software. Gene expression analysis was performed on organoids grown for 8 days in culture.

Analysis of DMVs

DMVs were identified as described previously (12) and further analyzed by visual inspection of methylation tracks in the UCSC Genome Browser. The refSeq gene annotation database was used for the identification of DMV-associated genes. A DMV that overlapped the gene body or promotor region of a gene was assigned to that gene. Analyses of the chromatin status of DMVs and DMV-surrounding regions were performed using the cistrome online platform (21). Intestinal ChIP-seq data from the mouse ENCODE project were used in our analysis (GSE31039).

454 DNA bisulfite sequencing

Deep DNA bisulfite sequencing was performed as described previously (22), see Supplementary Materials and Methods for details.

Support vector machine based classification of tumor and control samples

For the 30 genes that were found to be associated with hypermethylated DMVs and downregulated in the AOM/DSS model, we identified 28 human homologues. Of these, 18
were associated with a DMV or a DMV-like structure in a normal human colon methylome (23). We then used a support vector machine with a linear kernel (24) (Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm) to classify the RNA-seq data of the TCGA colon cancer set, based on 10-fold cross validation. To identify a minimum gene set for sample classification, the gene with the lowest F-score was removed, as long as the classification accuracy did not drop below 100%.

Analysis of TCGA datasets

We obtained expression and methylation data for colon adenocarcinoma from the TCGA database (as of February/March 2014). Expression data comprised normalized gene expression values, from Illumina HighSeq2000 RNASeqV2 experiments, for 300 samples (259 tumor, 41 normal). Methylation analysis was performed on genome-mapped beta values (third data level) derived from the TCGA Infinium HumanMethylation450 data set consisting of 329 samples (291 tumor, 38 normal). For each gene (NR5A2, ECH1 and FOXP2) three adjacent Infiniums probes were selected with regard to their positioning in DMVs. Boxplots were computed using the statistical computing language R (version 3.0.2). The additional package rgl provided tools for visualizing the combined gene expression. Welch two sample t-tests were performed to assess statistical significance.

Data access

Sequencing data have been deposited in the GEO database under the accession number GSE57569 and can be accessed using the following reviewer link:


Results
Whole-genome DNA methylation analysis of a mouse model for inflammation-induced intestinal carcinogenesis

To gain insight into the potential contribution of DNA methylation to intestinal inflammation and tumorigenesis, we used the AOM/DSS mouse model of inflammation-induced colon cancer (Supplementary Fig. S1). In this model, inflammatory bowel disease (IBD) was induced by adding 2% dextran sodium sulfate (DSS) to the drinking water during weeks 2, 5 and 8. In parallel experiments, intestinal tumorigenesis was induced by additional injections of azoxymethane (AOM). We determined base-resolution DNA methylomes of normal as well as inflamed (DSS-treated) and cancer (AOM/DSS-treated) intestinal epithelial cells. To obtain sufficient amounts of DNA for whole-genome bisulfite sequencing (WGBS), equimolar sample pools were prepared from intestinal epithelial cells or tumors from 3 mice for each group, resulting in 3 libraries (control, IBD and tumor) for sequencing. Paired-end sequencing was performed on an Illumina HiSeq 2000 platform (see Supplementary Materials and Methods for details), with average genome coverage of 19.4x (control), 13.8x (IBD) and 11.4x (tumor). We also determined the bisulfite conversion rate by analyzing mitochondrial sequences that were co-purified during the sample preparation and that are known to be unmethylated. These sequences showed a bisulfite conversion rate of >99.5% (see Supplementary Materials and Methods for details), suggesting highly effective bisulfite treatment.

Initial data analysis revealed many basal features of mammalian methylomes. This included the specificity for CpG dinucleotides and a characteristic bimodal enrichment for completely unmethylated and completely methylated CpG residues (data not shown). The average methylation ratios of the three datasets were very similar (Supplementary Fig. S2A), and a direct comparison of the methylation landscapes showed a high degree of similarity between the three datasets (Supplementary Fig. S2B). Partially methylated domains, which represent large regions spanning several 100 kb with reduced methylation levels, and which have been observed in a variety of human and mouse tumor methylomes (8,9,11,25), were not observed in any of the datasets (Supplementary Fig. S2C).
Inflammation-induced hypermethylation of DNA methylation valleys (DMVs)

Our results indicate that AOM/DSS treatments induce localized aberrations rather than the formation of large-scale DNA methylation changes. We therefore segmented our datasets according to gene substructure and observed a progressive (from control to IBD and then to tumor) and highly significant ($P<0.01$) increase in the methylation levels of promoters, 5'-UTRs and exons (Fig. 1A). This effect was not observed for 3'-UTRs and introns, which had highly similar methylation levels in all three datasets (Fig. 1A). Furthermore, a minor but significant ($P<0.01$) decrease in methylation levels was observed for intergenic regions in the IBD and tumor datasets (Fig. 1A). These findings indicate that hypermethylation in AOM/DSS inflammation-induced cancer occurs at 5' gene regions.

The two epigenetic features that are commonly associated with 5' regions of genes are CpG islands and DMVs. We therefore compared the methylation levels of all annotated CpG islands and DMVs in our datasets. This analysis failed to reveal any significant methylation changes in CpG islands, but interestingly showed a progressive and highly significant ($P<0.01$, two-tailed t-test) hypermethylation of DMVs in the IBD and tumor datasets (Fig. 1B). Of the 1753 DMVs that were identified in the normal intestinal methylome, 1048 were shared between all three samples, while several hundred were unique for specific samples (Fig. 1C), suggesting that a subset of DMVs undergoes dynamic methylation changes in our model. In agreement with previous findings (12), the DMVs that we identified in the control intestinal methylome showed robust sequence conservation (Fig. 1D), and a moderate enrichment for the H3K4me1 and H3K27me3 histone modifications (Fig. 1E). However, a stronger and more defined enrichment was observed for H3K4me3 and H3K27ac (Fig. 1E), two modifications that are usually associated with active gene expression. This finding is in agreement with the notion that DMVs are often associated with active genes (12).

Having established that DNA methylation is indeed dynamic at specific DMVs during inflammation, we performed a more detailed analysis of the methylome data. Interestingly,
this survey revealed that DMVs found in normal intestinal methylome were globally hypermethylated in the IBD and tumor samples (Supplementary Fig. S3). We also identified a set of 99 DMVs that showed robust hypermethylation (methylation difference >0.1) in the IBD or the tumor sample (see Fig. 2A, for an example). Remarkably, the majority of these DMVs became hypermethylated in epithelial cells from IBD mice (Fig. 2B). A closer analysis of the hypermethylated DMVs revealed additional novel features that expand previous analyses (12). For example, average DMV methylation profiles showed that hypermethylation was confined to valley floors and did not spread beyond the borders of the DMVs (Fig. 2C). Furthermore and of note, hypermethylated DMVs were strongly depleted for CpG islands and also showed reduced GC densities, when compared to the complete set of DMVs (Fig. 2D, E). Finally, DMVs that became hypermethylated were also enriched for the active chromatin marks H3K27ac and H3K4me3 in normal mouse intestine (Fig. 2F). Together, these data suggest that inflammation-related methylation changes are targeted to a specific subset of DMVs that are associated with active genes and thus define an epigenetic program that is clearly distinct from the previously described H3K27me3-associated de novo methylation of CpG island-associated cancer genes (26-28).

**Inflammation-induced gene expression changes**

To better understand the role of this unique set of hypermethylated DMVs, we performed transcriptome sequencing of control, IBD and tumor samples. mRNA was prepared from 3-4 individual mice, then barcoded and sequenced on an Illumina HiSeq 2000 platform (see Supplemental Experimental Procedures for details). Data analysis showed high expression levels of epithelial marker genes, such as *Epcam* and *Cdh1* (Supplementary Table S1). Immune genes were only expressed at low levels, with no significant increase in the IBD or tumor samples (Supplementary Table S1), reflecting the high purity of our epithelial cell preparations. Furthermore, expression levels for genes encoding DNA methylation enzymes, such as *Dnmt1-3* and *Tet1-3* and their known cofactors *Dnmt3l* and *Uhrf1* also appeared...
very similar in all samples (Supplementary Table S1), consistent with the absence of large-scale methylation changes in our mouse model.

To identify genes that were differentially expressed in the IBD and tumor samples, we used DESeq (29). This analysis revealed 791 and 1597 differentially expressed (Q<0.05) genes in the IBD and tumor samples, respectively (Supplementary Fig. S4A). Of these, a large fraction showed relatively minor quantitative changes (Supplementary Fig. S4A). We therefore applied an additional cutoff of ≥2-fold expression change, which revealed that the majority of genes that were deregulated in the IBD samples also showed concordant deregulation in the tumor samples (Supplementary Fig. S4B), thus illustrating the close relationship between intestinal inflammation and tumorigenesis. Pathway analysis of the 299 commonly upregulated genes revealed a highly significant enrichment of genes involved in immunological, gastrointestinal, and inflammatory diseases categories (Supplementary Fig. S4C), consistent with the phenotypic changes observed in the mice. Interestingly, a closer analysis of the downregulated genes revealed several genes that were associated with a differentially methylated DMV. To systematically investigate the relationship between differential expression and differential DNA methylation of nearby DMVs in the intestinal epithelium of the inflamed and tumorigenic tissue, we integrated our methylome and transcriptome datasets. This identified 102 genes that are associated with DMVs that disappeared (due to hypermethylation) and became significantly (Q<0.05) downregulated in the IBD or tumor samples. Pathway analysis of these genes revealed a significant enrichment of genes involved in biological processes directly relevant to tumor cell phenotypes (Fig. 2G). These results provide additional insight into the characteristic features of the inflammation-induced epigenetic program.

**Hypermethylation-induced silencing of DMV-associated genes**

Further data analysis identified 58 genes that are associated with strongly hypermethylated DMVs (methylation difference >0.1) in the IBD or tumor samples. Among these, 30 genes showed relatively high expression levels in the control sample and reduced expression levels
in the IBD or tumor samples, indicating methylation-dependent silencing in the inflamed and cancer tissues (Supplementary Table S2). To validate and further analyze the methylation changes at DMVs chosen from the above-described subgroup, we used targeted deep bisulfite sequencing. DNA was prepared from 3 individual mice per group and bisulfite converted. 8 regions with predicted differential methylation (as observed in our WGBS datasets) were then PCR amplified and sequenced by 454 technology, which routinely generates sequencing coverages exceeding 100x. The resulting methylation profiles provided strong confirmation for the WGBS-based results and also confirmed DMV hypermethylation in tumors from AOM mice (Supplementary Table S3). A prominent example is \textit{Nr5a2}, which encodes a nuclear receptor that has been shown to have a protective function against inflammatory bowel disease (30). Our 454 bisulfite sequencing data show that the \textit{Nr5a2} DMV was largely unmethylated in intestinal epithelial cells from 3 independent control mice (Fig. 3A). In contrast, epithelial cells from 3 independent DSS-treated mice and intestinal tumors from 3 independent DSS/AOM-treated mice showed robust \textit{de novo} methylation in the \textit{Nr5a2} DMV (Fig. 3A). These findings are consistent with the notion that hypermethylation of specific DMVs is an early event in inflammation-associated intestinal tumorigenesis. Similar effects were observed for 5 additional loci from our validation set (\textit{Adh1, B3gnt7, Kcne3, Sorbs2, Vdr}), with little variation between individual mice (Fig. 3B). Only 2 DMVs (\textit{Cx1} and \textit{Prelp}) showed a less pronounced hypermethylation and appeared to be more strongly affected by individual variation (Fig. 3B).

To specifically investigate the effect of the experimentally confirmed DMV methylation changes on gene expression we re-analyzed our RNA-seq datasets. Expression data for the set of 8 genes included in our 454 bisulfite sequencing analysis were available from 10 mice and showed a noticeably consistent reduction in gene expression levels for DSS-treated and AOM/DSS-treated mice, relative to controls (Fig. 3C). This effect was observed for all 8 DMV-associated genes that were included in our analysis and was highly significant (Q<0.05) in the majority of cases (Fig. 3C). Together, these findings indicate coordinated hypermethylation-induced gene silencing of specific loci during the inflammatory process.
Functional significance of DMV-associated gene silencing in intestinal cancer

To test whether hypermethylation of our unique set of DMVs is a general process occurring in other intestinal mouse malignancies, we analyzed their DNA methylation pattern in adenoma samples derived from \textit{APC (Min)} mice (31), using 454 technology. Interestingly, we found increased methylation levels in three tested DMVs relative to their wt counterparts (Fig. 4A, B and Supplementary Fig. S5). Consistent with this methylation increase, a significant down-regulation was observed in their associated genes (Fig. 4C). This suggests that the inflammation-associated epigenetic changes are conserved in mouse premalignant adenomas.

We also sought to investigate the functional significance of genes that were regulated by our specific set of hypermethylated DMVs. To this end, we used organoid cultures, derived from adenomas developing in \textit{APC (Min)} mice (adenoma organoids) or wt mice (wt organoids) (20). These cultures provide a pure system to investigate whether the effect of these epigenetic lesions are maintained independently of stroma and infiltrating immune cells, as well as a mean to study the role of these lesions in cellular transformation. For the first time, we have adopted this system to decipher the role of DNA methylation in inflammation-induced cancer. Adenoma organoids showed an increased activity of Wnt signaling (Fig. 5A), a hallmark of intestinal adenomas, exemplifying their independence from external Wnt ligands, as previously reported (20). Interestingly, in these organoids 4 out of 5 DMV-associated genes were significantly hypermethylated (Fig. 5B) and their expression level was strongly reduced (Fig. 5C). This clearly shows that the organoid cultures mimic the \textit{in vivo} methylation and expression patterns of our DMV-associated genes, and thus are suitable for studying their role in inflammation-induced intestinal cancer.

The availability of organoid cultures also provided an excellent opportunity to functionally characterize a novel candidate gene that is subjected to DMV-associated hypermethylation and gene silencing. Treatment of wt organoids with 4-methylpyrazole (4-MP), a competitive inhibitor of alcohol dehydrogenases, including Adh1, led to enhanced
growth of freshly derived organoids (Fig. 5D, E). This loss-of-function phenotype was also observed in long-term cultured organoids (Supplementary Fig. S6). Remarkably, inhibition of Adh1 activity was associated with increased Wnt signaling in wt organoids (Fig. 5F), as well as with a moderately but significantly ($p<0.05$) increased expression of the proliferation associated genes $Ki67$ and $Top2A$ (Fig. 5G). Altogether, these results further confirm the functional relevance of our novel set of DMV-associated genes, that are targeted for efficient hypermethylation and silencing, to facilitate inflammation-induced cellular transformation.

Hypermethylation-induced silencing of DMV-associated genes is conserved in human colon cancer samples

To further explore the conservation of our set of hypermethylated DMVs, we analyzed the expression of DMV-associated genes in human colorectal cancer datasets. Of the 30 genes that we found to be associated with hypermethylated DMVs and downregulated in the AOM/DSS model, we identified 28 human homologues. Of these, 18 were associated with a DMV or a DMV-like structure (Supplementary Fig. S7) in a human healthy colon methylome (23). We then analyzed the expression patterns of these 18 genes in a published RNA-seq dataset from The Cancer Genome Atlas (TCGA) (32) that was available for 259 colon cancer samples and 41 control mucosa samples. This revealed a distinct clustering of control samples that was defined by relatively high expression levels for the majority of DMV-associated genes (data not shown). Furthermore, we used a support vector machine approach to define a minimum subset of genes for sample classification. After starting with all 18 DMV-associated genes, we iteratively removed the gene with the lowest F-score as long as the classification accuracy did not drop below 100% (Supplementary Fig. S8). This identified an optimized set of 3 genes, $NR5A2$, $FOXP2$ and $ECH1$, with a remarkable classification accuracy (10-fold cross-validation) of 100%. Interestingly, the expression levels of all three genes were significantly reduced in tumor samples when compared to control tissues (Fig. 6A). Moreover, low expression levels were coordinately observed for all three
genes in the majority of tumor samples (Fig. 6B), which is in agreement with the notion that they are regulated by a common epigenetic silencing mechanism.

To investigate the methylation patterns of these DMVs in human samples, we analyzed array-based methylation datasets from TCGA (32) that were available for 291 colon cancer samples and 38 control mucosa samples. Data analysis showed that several methylation probes from the array were located in the DMVs associated with NR5A2, ECH1 and FOXP2. The NR5A2 DMV probes showed a robust and significant methylation increase in the tumor samples (Fig. 6C). For the ECH1 DMV, significant hypermethylation was also detected in tumors (Fig. 6C), but the effect size appeared smaller, which may indicate that hypermethylation is restricted to a subset of tumor cells. The FOXP2 probes again showed markedly increased methylation levels in tumors (Fig. 6C). Together, our data clearly demonstrate that hypermethylation and silencing of our novel inflammation-associated set of DMVs is a general process conserved in mouse and human colon cancers.

Discussion

The incidence of inflammatory bowel disease (IBD) is increasing among adults and children in the developed and developing world. Current estimates attribute roughly 10% of the IBD heritability to identified genetic variants (33), while environmental factors interacting with genetic predisposition are discussed as major determinants for disease manifestation (34). Importantly, IBD significantly increases the risk of colorectal cancer (4). Since epigenetic patterns are highly tissue-specific and can be influenced by environmental factors, they might explain key features of IBD and IBD-associated cancers (35,36). Indeed, two previous studies have identified candidate risk loci for IBD, using low-coverage DNA methylation arrays (37,38).

We have now used whole-genome bisulfite sequencing of an established mouse model for IBD and IBD-associated intestinal cancer to investigate disease-associated DNA
methylation changes at single-base resolution. Our results identify the hypermethylation of DMVs as a dynamic and prominent feature, not only of the cancer methylome (12), but also of the inflamed intestinal methylome. Moreover, our results clearly show that DMV hypermethylation is an early event in inflammation-induced intestinal tumorigenesis. Over the past years, many studies have associated the mechanisms of epigenetic deregulation with cancer (39). Cancer-specific DNA methylation patterns are often explained through H3K27me3-associated de novo methylation of CpG islands in the vicinity of genes that may already be repressed in the corresponding normal cell types (26-28). This mechanism has also been implicated in the etiology of mouse intestinal adenomas (40). However, the molecular mechanism identified in our study appears to be different since the hypermethylated DMVs showed a robust enrichment for the activating chromatin marks H3K27ac and H3K4me3, were depleted of CpG islands and characterized by intermediate to low GC densities (Fig. 7). This indicates that during the inflammatory process aberrant DMV methylation occurs in a programmed and coordinated manner, and, as a consequence, silences the expression of nearby genes, which are active in the normal cells. Furthermore, our data clearly demonstrate that inflammation-associated hypermethylation and silencing of a specific set of DMVs is a general process, conserved in other intestinal mouse and human colon malignancies.

Recent epigenomic analyses have indicated the existence of two distinct gene silencing mechanisms during cellular differentiation (12,41): while the majority of lineage specific genes of differentiating stem cells were found to be GC-rich and silenced by PcG-mediated repression in non-expressing lineages, gene silencing in later stages of development was mostly characterized by DNA methylation of GC-poor genes. Our results are consistent with the latter mechanism and suggest that PcG-independent gene silencing is relevant for epigenetic deregulation in inflammation and creates a signature that is maintained through malignant tumor stages. Our data also raise the possibility of additional, inflammation-induced methylation changes, including the establishment of de-novo DMVs in
AOM/DSS-treated mice (Fig. 1B). These findings will have to be validated and further analyzed in future studies.

Remarkably, inflammation-induced DMV hypermethylation affected a number of genes that have previously been reported to be methylated or and silenced in human cancers, such as *B3GNT7*, *CDX1*, *NR5A2* and *VDR* (42-45). This confirms previous findings of DNA hypermethylation in DMVs that are associated with cancer genes (12), and further establishes that hypermethylated DMVs define a novel epigenetic program that links intestinal inflammation to colon cancer.

It should be noted that our novel set of inflammation-induced DMV-silenced genes is also hypermethylated and silenced in adenoma organoid cultures, which recapitulate the complete intestinal stem cell differentiation hierarchy. This *in vitro* system allows confirmation of the functional relevance of these genes. Indeed, we found that pharmacological inhibition of Adh1, an enzyme that is encoded by a DMV-associated gene, significantly enhanced intestinal organoid proliferation, possibly via the activation of Wnt-dependent signaling. Moreover, the functional relevance of this program is clearly illustrated by the 3 DMV-associated genes that correctly classified human colon cancer: *NR5A2* encodes a nuclear receptor with protective functions against inflammatory bowel disease in mice (30). *ECH1* encodes enoyl coenzyme A hydratase, an essential enzyme for fatty acid metabolism. Fatty acid metabolism has been shown to play an important role in IBD and in colon cancer (46). *FOXP2* encodes a transcription factor that is commonly associated with brain function but that has also been shown to be involved in gut development (47). Thus, using an integrated approach, we were able to provide evidence that inflammation-associated epigenetic lesions could have causal effects in cancer development.

The diagnostic potential of aberrant DNA methylation patterns in colorectal cancer is well established (48,49), and our data suggest that DMV hypermethylation might be a promising biomarker for early cancer detection. Comprehensive methylation profiling of IBD samples and polyps will be required to systematically identify inflammation-related epigenetic biomarkers for the risk stratification of patients. Moreover, human diseases known to be
associated with increased cancer incidence, including obesity, are known to involve a systemic low grade inflammation (2). However, biomarkers that identify patients at risk for cancer development are lacking. We suggest that the methylation status of a minimal gene subset can be used in large retrospective studies to evaluate their potential as risk stratification biomarkers.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ contributions

Conception and design: Y. Bergman, Frank Lyko

Acquisition of data: M. Abu-Remaileh, I. Ansari, D. Cohen, T. Musch, H. Linhart, E. Pikarsky

Analysis and interpretation of data: M. Abu-Remaileh, S. Bender, G. Raddatz, J. Gutekunst, A. Breiling

Writing, review, and/or revision of the manuscript: Y. Bergman, Frank Lyko

Study supervision: Y. Bergman, Frank Lyko

Acknowledgements

We thank Haya Hamza, Yinon Ben-Neriah and Guy Ludwig, The Hebrew University Medical School, for providing organoids and adenoma samples, respectively. We further thank the
DKFZ Genomics and Proteomics Core Facility for sequencing services. This work was supported in part by the Cooperation Program in Cancer Research of the Deutsches Krebsforschungszentrum (DKFZ) and Israel's Ministry of Science, Technology and Space (MOST) (F.L., Y.B.). This work was further supported by a research grant from the Israel Cancer Research Foundation (Y.B.). M.A.R. was supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities.

References


**Figure legends**

**Figure 1.** Characterization of DNA Methylation Valleys (DMVs) in the AOM/DSS model. A, Average DNA methylation ratios of various intragenic subsegments are shown for the control (blue), IBD (green) and tumor sample (red), respectively. Visible methylation differences between individual samples were highly significant (P<0.01), as determined by a paired t-test. B, Boxplots showing the distribution of methylation ratios for CpG islands (left) and DMVs (right) in the control (blue), IBD (green) and tumor sample (red), respectively. Only DMVs with >20 CpGs were included in this analysis to ensure comparability and analytical robustness. Differences between samples were not significant (P>0.05, two-tailed t-test) for CpG islands, but highly significant for DMVs (as indicated). C, Venn diagram illustrating the numbers and conservation of DMVs between the samples. C, Average sequence conservation level around DMVs. D, Chromatin profile of DMVs. The highest enrichment is observed for H3K4me3 (green), followed by H3K27ac (purple), H3K4me1 (red) and H3K27me3 (black).

**Figure 2.** Inflammation-induced hypermethylation of DNA methylation valleys (DMVs). A, *Adh1* as a representative example for a DMV-associated gene. DNA methylation and histone modifications are shown as UCSC Genome Browser tracks. The DMV (shaded in orange) overlaps the promoter, TSS and 2 exons. B, Heatmap showing methylation ratios of DMVs that became hypermethylated (methylation difference >0.1) in the IBD or the tumor sample. C, Average methylation profile of hypermethylated DMVs. Hypermethylated DMVs were normalized for their different lengths. D, Hypermethylated DMVs are depleted for CpG islands. Bars show the fraction of DMVs that contain CpG islands. Results are shown for all DMVs (all) and for hypermethylated DMVs (hyper). E, Hypermethylated DMVs show reduced GC density. Bars show the GC density for all DMVs (all) and for hypermethylated DMVs (hyper). The red line indicates the average GC density of the mouse genome. F, Chromatin profile of hypermethylated DMVs. Robust enrichment is observed for H3K4me3 (green) and
Abu-Remaileh et al.

H3K27ac (purple). No enrichment could be observed for H3K4me1 (red) and H3K27me3 (black). G, Ingenuity pathway analysis of the 102 genes that are associated with DMVs which were lost (due to hypermethylation) and became downregulated in the IBD or tumor samples. In addition to "Cancer", the five most highly enriched functions are shown.

**Figure 3.** DMV hypermethylation is associated with gene silencing. A, 454 bisulfite sequencing results for the *Nr5a2* DMV in 3 different groups of mice (control, DSS, AOM/DSS). PCR amplification was performed on DNA samples from individual mice and sequencing results are shown as heatmaps. Each row represents one sequence read, blue boxes represent methylated CpG dinucleotides, yellow boxes represent unmethylated CpG dinucleotides. Sequencing coverage ranged from 312 to 495 reads, as indicated. B, 454 bisulfite sequencing results for the complete validation set. The heatmap shows average methylation ratios of 8 DMV amplicons from 9 individual mice. C, Relative expression levels of genes associated with the differentially methylated DMVs shown in (B). Gene expression data was extracted from RNAseq datasets, gene expression levels are indicated relative to the maximum gene-specific expression level across all individual samples. Q-values (as determined by DESeq) are indicated relative to controls.

**Figure 4.** DNA methylation-induced silencing of inflammation-associated DMVs in adenomas from *APC(Min)* mice. A, Representative 454 bisulfite sequencing results for the Vdr DMV in wild-type (4 mice) and adenoma-derived (4 mice) epithelium. PCR amplification was performed on DNA samples from individual mice and sequencing results are shown as heatmaps. Each row represents one sequence read, blue boxes represent methylated CpG dinucleotides, yellow boxes represent unmethylated CpG dinucleotides. Sequencing coverage ranged from 646 to 896 reads, as indicated. B, Quantification of 454 bisulfite sequencing results for 3 DMV amplicons from 9 individual mice. Normal intestinal epithelial cells were derived from wild-type mice (n=4) and adenoma epithelial cells from Min mice (n=5). Values are means and error bars indicate s.e.m. P-values were calculated using a t-
test (* ≤ 0.05, ** ≤ 0.001). C, Normalized expression levels of genes associated with the
differentially methylated DMVs in wild-type- (n=6) and adenoma (n=4) -derived epithelium.
Values are means and error bars indicate s.e.m. P-values were calculated using a t-test (* ≤ 0.01, ** ≤ 0.001).

Figure 5. Functional silencing of DMV-associated genes in adenoma-derived organoids. A,
Normalized expression levels of Wnt target genes in wild-type (wt) and adenoma-derived
intestinal organoids. B, 454 bisulfite sequencing results for 4 DMVs in freshly isolated (8 days
in culture) wt- and adenoma-derived organoids. C, Expression levels of DMV-associated
genes in wt- and adenoma-derived intestinal organoids. Values are means and error bars
indicate s.d. The results are representative of two independent derivations of organoids in
each group. D, Representative images of wt intestinal organoids treated or not treated with 1
mM 4-methylpyrazole (4-MP) for 4 days (see Materials and Methods). Scale bar is 50 µm. E,
Quantification of organoid size using area size in µm². ** represents a P-value ≤ 0.05, as
calculated by a t-test. Normalized expression levels of (F) Wnt signaling target genes and (G)
proliferation genes in wt- and adenoma-derived intestinal organoids. Values are means and
error bars indicate s.d. The results are representative of two independent derivations of
organoids in each group.

Figure 6. DNA methylation-induced silencing of inflammation-associated DMVs in human
colon cancer. A, Boxplots showing mRNA expression levels of the minimum gene set for
tumor classification in normal mucosa (N, n=41) and tumor (T, n=259) samples from the
TCGA colon cancer set. B, Combined expression levels of NR5A2, ECH1 and FOXP2 in
colon adenocarcinomas. Each dot represents a tumor sample from the TCGA colon cancer
set. C, Boxplots showing methylation levels of 3 adjacent probes each from the NR5A2,
ECH1 and FOXP2 DMVs in control (n=38) and tumor (n=291) samples from the TCGA colon
cancer set. cg numbers indicate probe numbers (see Methods for details).
Figure 7. Model illustrating the differences between the epigenetic program induced by intestinal inflammation (right panel) and the known epigenetic program associated with gene silencing in cancer (left panel). While the known mechanism of epigenetic gene silencing in cancer is associated with poorly expressed genes that are characterized by high CpG density and Polycomb-dependent repressive chromatin modifications, inflammation-induced epigenetic gene silencing is targeted to highly expressed genes with low CpG density and active chromatin modifications.
Figure 2
Figure 3

A

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mouse</th>
<th>CpGs</th>
<th>Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>10</td>
<td>9</td>
<td>312</td>
</tr>
<tr>
<td>Control</td>
<td>17</td>
<td>9</td>
<td>449</td>
</tr>
<tr>
<td>Control</td>
<td>18</td>
<td>9</td>
<td>390</td>
</tr>
<tr>
<td>DSS</td>
<td>5</td>
<td>9</td>
<td>343</td>
</tr>
<tr>
<td>DSS</td>
<td>13</td>
<td>9</td>
<td>377</td>
</tr>
<tr>
<td>DSS</td>
<td>15</td>
<td>9</td>
<td>384</td>
</tr>
<tr>
<td>AOM</td>
<td>1</td>
<td>9</td>
<td>348</td>
</tr>
<tr>
<td>AOM</td>
<td>2</td>
<td>9</td>
<td>470</td>
</tr>
<tr>
<td>AOM</td>
<td>22</td>
<td>9</td>
<td>395</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>Gene</th>
<th>control</th>
<th>DSS</th>
<th>AOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adh1</td>
<td>0.12</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>B3gnt7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Cdx1</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Kcne3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Nr5a2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prelp</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Sorbs2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Vdr</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

C

<table>
<thead>
<tr>
<th>Gene</th>
<th>control</th>
<th>DSS</th>
<th>AOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adh1</td>
<td>0.12</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>B3gnt7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Cdx1</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Kcne3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Nr5a2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prelp</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Sorbs2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Vdr</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Figure 5

A

Bar graph showing relative expression levels of Axin2, c-Myc, CD44, and CyclinD1 in wt organoids and adenoma organoids.

B

Heatmaps displaying the expression of Vdr, Nr5a2, Cdx1, and Prelp in wildtype, adenoma with 4 CpGs, 9 CpGs, 29 CpGs, and 5 CpGs.

C

Bar graph showing expression levels of Vdr, Nr5a2, Cdx1, Prelp, and Adh1 in wt and adenoma organoids.

D

Micrographs of untreated and 1 mM 4-MP treated organoids.

E

Scatter plot showing size (μm x 10⁷) of organoids in untreated and 1 mM 4-MP treated conditions.

F

Bar graph showing relative expression levels of Axin2, c-Myc, CD44, and CyclinD1 in untreated and 1 mM 4-MP treated conditions.

G

Bar graph showing relative expression levels of K67 and Top2A in untreated and 4-MP treated conditions.
Figure 6

A

NR5A2

ECH1

FOXP2

normalized counts (TCGA)

N                  T

P = 5.7x10^{-22}

P = 3.8x10^{-24}

P = 2.2x10^{-4}

B

NR5A2

FOXP2

ECH1

C

NR5A2

cg17759535  cg03257172  cg18203366

beta value

N                  T

1.4x10^{-47}  1.3x10^{-152}  1.7x10^{-144}

ECH1

cg22017573  cg18615328  cg16018184

N                  T

1.8x10^{-6}  3.3x10^{-6}  7.9x10^{-6}

FOXP2

cg13486406  cg05994320  cg25941299

N                  T

3.2x10^{-30}  6.9x10^{-41}  4.4x10^{-34}
Figure 7

Polycomb-hypermethylated genes

Normal

H3K4me3
H3K27me3
H3K27ac

unmethylated CpG
methylated CpG

Cancer/Inflammation

Inflammation-hypermethylated genes


Chronic inflammation induces a novel epigenetic program that is conserved in intestinal adenomas and in colorectal cancer

Monther Abu-Remaileh, Sebastian Bender, Günter Raddatz, et al.

Cancer Res  Published OnlineFirst March 25, 2015.

Updated version  Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-14-3295
Supplementary Material  Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2015/03/27/0008-5472.CAN-14-3295.DC1
Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions  To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/early/2015/03/25/0008-5472.CAN-14-3295. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.