Loss of MMR and TGFBR2 Increases the Susceptibility to Microbiota-Dependent Inflammation-Associated Colon Cancer

Elena Tosti,1 Ana S. Almeida,2 Tam T. Tran,3 Mariel Barbachan e Silva,4 Pilib Ó. Broin,4 Robert Dubin,1 Ken Chen,1 Amanda P. Beck,5 Andrew S. Mclellan,6 Eduardo Vilar-Sanchez,7 Aaron Golden,4 Paul W. O’Toole,2 and Winfried Edelmann1

SUMMARY

We present a murine model that demonstrates the synergistic effects of DNA mismatch repair deficiency and the early loss of TGFBR2 in inflammation-associated colon tumorigenesis (inflammatory bowel disease—associated colorectal cancer [IBD-CRC]). Importantly, we found that the mutational and transcriptional alterations in these IBD-CRCs were highly similar to those observed in human IBD-CRCs and highlight the importance of the initial microbiota composition in the development of IBD-CRCs.

BACKGROUND AND AIMS: Mutations in DNA mismatch repair (MMR) genes are causative in Lynch syndrome and a significant proportion of sporadic colorectal cancers (CRCs). MMR-deficient (dMMR) CRCs display increased mutation rates, with mutations frequently accumulating at short repetitive DNA sequences throughout the genome (microsatellite instability). The TGFBR2 gene is one of the most frequently mutated genes in dMMR CRCs. Therefore, we generated an animal model to study how the loss of both TGFBR2 signaling impacts dMMR-driven intestinal tumorigenesis in vivo and explore the impact of the gut microbiota recombinase in the intestinal epithelium. VCMsh2/Tgfbr2 mice were analyzed for their rate of intestinal cancer development and for the mutational spectra and gene expression profiles of tumors. In addition, we assessed the impact of chemically induced chronic inflammation and gut microbiota composition on colorectal tumorigenesis.

RESULTS: VCMsh2/Tgfbr2 mice developed small intestinal adenocarcinomas and CRCs with histopathological features highly similar to CRCs in Lynch syndrome patients. The CRCs in VCMsh2/Tgfbr2 mice were associated with the presence of colitis and displayed genetic and histological features that resembled inflammation-associated CRCs in human patients. The development of CRCs in VCMsh2/Tgfbr2 mice was strongly modulated by the gut microbiota composition, which in turn was impacted by the TGFBR2 status of the tumors.

CONCLUSIONS: Our results demonstrate a synergistic interaction between MMR and TGFBR2 inactivation in inflammation-associated colon tumorigenesis and highlight the crucial impact of the gut microbiota on modulating the incidence of inflammation-associated CRCs. (Cell Mol Gastroenterol Hepatol 2022; ■ ■ ■ ■; https://doi.org/10.1016/j.jcmgh.2022.05.010)

Keywords: DNA Mismatch Repair; Colon Cancer; Gut Microbiota; Inflammation.
Colorectal cancer (CRC) represents the third-most-common cancer type worldwide and is the second leading cause for cancer-related mortality in the United States. Multiple risk factors contribute to its etiology including genetic predisposition, genotoxic and environmental factors, prolonged exposure to chronic inflammation, and unbalanced diet. Importantly, risk factors for CRC are known to interact with gut microbes.\(^1\)\(^-\)\(^3\) The study of the gut microbiota involvement in CRC has generated multiple theories for how single or groups of taxa contribute to CRC development: altered microbial metabolites cause loss of intestinal epithelial barrier function, enhanced access of inflammatory bacteria to epithelial cells, combination of bacterial-driven inflammation, and genotoxic activity lead to mutation and hyperplastic transformation followed by adenocarcinoma.\(^4\)

DNA mismatch repair (MMR) plays a crucial role in maintaining the integrity of the genome by removing misincorporated nucleotides that result from errors in DNA replication and by mediating a DNA damage response after exposure to genotoxic agents.\(^5\) Deficient MMR (dMMR) leads to a 50- to 1000-fold increase in mutation rates and is causative for Lynch syndrome (LS) and 15%-20% of sporadic CRCs.\(^6\) The MMR genes most frequently affected in LS and sporadic CRCs are MSH2 and MLH1 and encode key components of MMR complexes involved in coordinating mismatch recognition and excision.

dMMR CRCs are characterized by insertion or deletion mutations at microsatellite DNA sequences, termed microsatellite instability (MSI). In dMMR CRCs, MSI frequently occurs in coding repeat sequences, and these mutations are thought to contribute to dMMR intestinal tumorigenesis.\(^7\) A remarkable example of coding MSI is the disruption of the TGFBR2 gene at a poly(A) repeat occurring in up to 90% of LS and 70%-80% of sporadic dMMR CRCs.\(^8\) TGFBR2 signaling regulates the differentiation of enterocytes in the crypts or villi, acts in re-establishing intestinal crypt homeostasis upon injury,\(^9\) and plays critical roles in controlling cell proliferation and differentiation.\(^8\) In tumorigenesis, TGFβ signaling is complex and mediates pro- and antitumoral activities depending on the concurrent gene mutations in cancer cells and characteristics of the cells infiltrating the tumor microenvironment.\(^9\)

The high prevalence of TGFBR2 mutations in MSI-positive CRCs suggests that TGFβ2 inactivation is an essential event in dMMR-driven tumorigenesis. Indeed, studies in Tgfbr2 knockout mice showed that TGFβ2 functions as a tumor suppressor when tumorigenesis is initiated genetically by Apc or Pten deletions, by oncogenic Kras activation, or chemically by genotoxic or colitogenic insult.\(^10\)-\(^13\)

TGFβ2 heterozygous mutations cause Loesys-Dietz syndrome, a rare autosomal dominant disease with an increased risk to inflammatory bowel disease (IBD).\(^14\)

Studies of inflammation-associated CRCs (IBD-CRCs) showed MSI in 20%-50% of advanced ulcerative colitis-associated CRCs.\(^15\) Interestingly, MSI-positive IBD-CRCs\(^15\) display a high incidence of TGFβ2 mutations (50%-76% of cases), which is also observed in early dysplastic lesions.\(^15\)-\(^17\) Interestingly, loss of TGF-beta signaling has been linked to gut microbiota alterations in mice.\(^18\)

Here, we demonstrate that the combined loss of MMR and TGFβR2 tumor suppressor genes in intestinal epithelial cells caused the development of mucinous CRCs with histologic and molecular features highly similar to human IBD-CRCs. The analysis of intestinal microbiota in VCMsh2/Tgfbr2 mice revealed alterations in their taxonomic composition during IBD-CRC tumorigenesis, delineating interactions between intestinal inflammation and microbial dysbiosis in the development of IBD-CRCs. Overall, our studies indicate a role for both MMR and TGFβR2 signaling in protecting from IBD-CRC consistent with the presence of MMR and TGFβR2 mutations in human IBD-CRCs.\(^19\)

**Results**

*Inactivation of Msh2 and Tgfbr2 in the Intestinal Epithelium Induces Colon Tumorigenesis*

To study how the loss of TGFβR2 affects dMMR intestinal tumorigenesis, we generated Villin-Cre;Msh2\(^{2loxP/lop}\) (VCMsh2) and Villin-Cre;Msh2\(^{2loxP/lop}\);Tgfbr2\(^{-}\)\(^{loxP/lop}\) mice (VCMsh2/Tgfbr2). VCMsh2/Tgfbr2 mice had a significantly reduced lifespan (median survival 8 months) compared with VCMsh2 mice (median survival 12 months) (Figure 1A).

Both mouse lines developed intestinal tumors at very high incidence (Table 1 in Figure 1). A large proportion of the intestinal tumors in VCMsh2/Tgfbr2 mice were localized to the colon (74.16% \(n = 66 \text{ of } 89\)) and in 20.2% (\(n = 18 \text{ of } 89\)) of mice they were co-occurring with tumors in the small intestine (SI) (Table 1 in Figure 1). A subset of VCMsh2/Tgfbr2 mice (23.6% \(n = 21\) of 89) developed only SI tumors. VCMsh2 mice developed tumors only in the SI (89.3% \(n = 25\) of 28). The analysis of age and tumor location revealed that VCMsh2/Tgfbr2 mice with CRCs died faster (average of 7 months), while VCMsh2/Tgfbr2 mice developing only SI tumors survived longer (9 months on average). VCMsh2/Tgfbr2 mice developed SI tumors significantly faster than VCMsh2 mice, indicating that the loss of TGFβ2 also accelerated dMMR SI tumorigenesis (Figure 1B).

The CRC-bearing VCMsh2/Tgfbr2 mice had a thickened mucosa by inflammation with abundant mucous and feces within the colon lumen. VCMsh2/Tgfbr2 CRCs did not form intraluminal masses characteristic of SI tumors in VCMsh2/Tgfbr2 mice.

**Abbreviations used in this paper:** 8-oxoG, 8-oxoguanine; CRC, colorectal cancer; dMMR, deficient mismatch repair; DSS, dextran sulfate sodium; ELISA, enzyme-linked immunosorbent assay; GSEA, gene set enrichment analysis; IBD-CRC, inflammatory bowel disease—associated colorectal cancer; IHC, immunohistochemistry; IL, interleukin; IPA, Ingenuity Pathway Analysis; LS, Lynch syndrome; MMR, mismatch repair; MSI, microsatellite instability; MSS, microsatellite stable; OUT, Operational Taxonomic Unit; qPCR, quantitative polymerase chain reaction; rRNA, ribosomal RNA; SI, small intestine; SPF, specific pathogen—free; TNF-α, tumor necrosis factor α.

© 2022 The Authors. Published by Elsevier Inc. on behalf of the AGA Institute. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.ccmgh.2022.05.010

FLA 5.6.0 DTD  ■ JCMGH1025 proof ■ 23 June 2022 ■ 1:49 pm ■ ce OB
Figure 1. Inactivation of Msh2 and Tgfbr2 in the intestinal epithelium induces colon tumorigenesis. (A) Kaplan-Meier analysis. VCMsh2/Tgfbr2 (n = 89) vs VCMsh2 (n = 28) P < .0001. (B) Comparison of tumor occurrence. VCMsh2/Tgfbr2 CRCs (n = 42; 7 months) vs SI tumors (n = 18; 9.5 months), **** P < .0001; VCMsh2/Tgfbr2 SI tumors (n = 18; 9.5 months) vs VCMsh2 SI tumors (n = 28, 12 months) ** P = .0019. (C) I, Macroscopic image of VCMsh2/Tgfbr2 CRC. Representative images: II, hematoxylin and eosin staining of VCMsh2/Tgfbr2 mucinous CRC (scale bar = 200 μm); III, dysplastic VCMsh2/Tgfbr2 colon (scale bar = 50 μm); IV, hypertrophy or hyperplasia in goblet cells of VCMsh2/Tgfbr2 colon mucosa (scale bar = 20 μm); V, colitis in VCMsh2/Tgfbr2 colon (scale bar = 20 μm); VI, VCTgfbr2 mucinous CRC (scale bars = 50 μm, 200 μm). (D) Kaplan-Meier analysis. FCMsh2/Tgfbr2 (n = 38) vs FCMsh2 (n = 31), *** P < .0001. (E) FCMsh2/Tgfbr2 mucinous CRC (scale bars = 50 μm, 200 μm).

mice but were highly invasive, extending through the bowel wall to form large masses on the serosal side (Figure 1C-I). Histologically, the CRCs were adenocarcinomas containing large mucin lakes that invaded from the base of the crypts through the lamina propria, submucosa and the muscular layer (Figure 1C-II). In VCMsh2/Tgfbr2 colonic crypts exhibited dysplasia (Figure 1C-III) and hypertrophy or hyperplasia of mucosal goblet cells (Figure 1C-IV); lymphocytes, plasma cells, and neutrophils expanded the lamina propria and extended multifocally into the underlying submucosa (Figure 1C-V). Single mutant VCTgfbr2 mice were generated during the generation of VCMsh2/Tgfbr2 mice as well. A small number of VCTgfbr2 mice developed mucinous CRCs (22.7% [n = 9 of 26]) at an average age of 13 months (Figure 1B-VI).

In addition, we generated Fabp1-Cre; Msh2loxP/loxP; Tgfbr2loxP/loxP (FCMsh2/Tgfbr2) mice that also displayed reduced survival and developed CRCs, while Fabp1-Cre; Msh2loxP/loxP (FCMsh2) mice developed exclusively SI tumors (Table 2 in Figure 1D). Histologically, FCMsh2/Tgfbr2 CRCs displayed the same mucinous phenotype as the CRCs in VCMsh2/Tgfbr2 mice (Figure 1E). Overall, the increased CRC incidence using 2 different intestine-specific Cre recombinase transgens indicates a crucial role for TGFBR2 in suppressing dMMR-driven colon tumorigenesis.

Histopathological and Molecular Features of VCMsh2/Tgfbr2 CRCs

The intestinal tumors in VCMsh2/Tgfbr2 mice displayed unique characteristics. The SI tumors differed between genotypes in mucinous content: VCMsh2/Tgfbr2 mice developed SI adenocarcinomas with extensive mucinous content similar to CRCs (Table 3 in Figure 2A-I/II), whereas the SI...
adenocarcinomas in VCMsh2 mice displayed mainly villous or tubulovillous growth patterns with minimal to no mucin accumulation (Table 3 in Figure 2A-III). The inflammation score in the colonic mucosa and submucosa of VCMsh2/Tgfbr2 mice was significantly higher compared with their SI tracts and with the colonic mucosa in VCMsh2 mice (Figure 2B). VCMsh2/Tgfbr2 CRCs displayed high frequency of MSI, as expected (Table 4 in Figure 2). Because MMR facilitates the repair of oxidative DNA damage and inflammation is associated with increased reactive oxygen species production, the loss of MMR and increased inflammation together could result in increased unrepaired oxidative DNA damage. Indeed, enzyme-linked immunosorbent assay (ELISA) and immunohistochemistry (IHC) analyses showed that the genomic DNA in VCMsh2/Tgfbr2 CRCs contained higher levels of unrepaired 8-oxoguanine (8-oxoG) lesions compared with SI tumors in both VCMsh2/Tgfbr2 and VCMsh2 mice (Figure 2C and D). A canonical feature of colorectal carcinogenesis is dysregulation of WNT signaling, mainly caused by mutation in the APC tumor suppressor gene, resulting in the nuclear accumulation of beta-catenin. VCMsh2/Tgfbr2 CRCs did not show any beta-catenin nuclear localization (Figure 2E), whereas SI tumors in both VCMsh2 and VCMsh2/Tgfbr2 mice frequently showed widespread beta-catenin nuclear localization (60% and 50%, respectively) (Table 5 in Figure 2E).

**Mutational Analysis of VCMsh2/Tgfbr2 Intestinal Tumors**

The histological and molecular differences between SI and colon tumors in VCMsh2/Tgfbr2 mice suggested that different genetic alterations underlie tumorigenesis in different regions of the intestine. We therefore analyzed the
Exome mutational spectra associated with the pathogenesis of SI tumors and CRCs in VCMsh2/Tgfbr2 mice (n = 5) within the same mouse. In SI tumors we detected an average of 622.8 mutations/tumor (ranging from 514 to 796), CRCs showed an average of 516.6 mutations/tumor (ranging from 201 to 1306). In both SI tumors and CRCs, missense mutations were the most frequent type of alterations, followed by frameshift deletions (Figure 3A). The distribution of base changes showed a preponderance of C>T and T>C transitions (Figure 3B). The third-most common type of mutations were C>A transversions, likely resulting from unrepaird oxidative DNA lesions (Figure 3B). Overall, the mutational signature resembled that of MSI-positive human tumors consistent with a mutational process caused by dMMR. The variants were annotated and functionally predicted: 374 high- and 503 moderate-impact mutations were present in SI tumors, while 298 high- and 474 moderate-impact variants were found in CRCs.

While these results indicated similarity in the overall variant classification between VCMsh2/Tgfbr2 SI tumors and CRCs, the analysis of the affected genes revealed significant differences. Consistent with the IHC analyses (Figure 2E, Table 5 in Figure 2), the Apc gene was mutated in 80% of SI tumors but not in any of the CRCs (Figure 3C and D). When comparing the genes with high-impact mutations in...
VCMsh2/Tgfbr2 CRCs and CRCs, only 27 were in common, all reported to be involved in intestinal tumorigenesis (Figure 3 E and F).

We next analyzed the mutational targets in VCMsh2/Tgfbr2 CRCs for similarities with human MSI-positive CRCs extracted from the cBioPortal database. High-impact mutations occurred recurrently in several genes (Xirp1, Ubr5, Tsnax1p, Trmt10a, Phactr4, Htr1f, Arid1a, Acvr2a) that are also mutated at higher frequencies in human MSI-positive CRCs (91 samples) compared with microsatellite stable (MSS) CRCs (1501 samples) (Figure 3G).

Core Ingenuity Pathway Analysis (IPA) was employed to investigate the biological functions of genes with high impact mutations in VCMsh2/Tgfbr2 CRCs. The major disease and functions identified were cancer and gastrointestinal disease (Table 6 in Figure 4). When investigating the top canonical pathways category, VCMsh2/Tgfbr2 CRC high-impact mutations were enriched in genes belonging to the epithelial adherens junction signaling, PPARα/RXRα activation, and GP6 signaling pathway categories (Table 7 in Figure 4). In addition, IPA Bio-profile analysis identified a group of mutated genes that were previously found associated with colitis or IBD-CRC either in patients or in mouse models (Figure 4A).

Finally, we investigated similarities with the available exome data from human IBD-CRCs using a dataset of
high- or moderate-impact mutations from 7 MSI-positive IBD-CRCs and 24 MSS IBD-CRCs. VCMsh2/Tgfbr2 CRCs shared a higher number of mutations with MSI-positive IBD-CRCs compared with MSS IBD-CRCs (Figure 4B). The genes in VCMsh2/Tgfbr2 CRCs that IPA Bio-profile analysis reported to be involved in colitis or IBD-CRC were also mutational targets predominantly in the human MSI IBD-CRCs (Figure 4C). We also observed a marked enrichment in mutations in genes belonging to the 3 top canonical pathways in VCMsh2/Tgfbr2 CRCs in human MSI IBD-CRCs but not in MSS IBD-CRCs (Figure 4D).

The Tumor Microenvironment in VCMsh2/Tgfbr2 CRCs Is Characterized by an Inflammation-Specific Gene Expression Signature

To study the impact of the mutational changes on gene expression during VCMsh2/Tgfbr2 colorectal tumorigenesis, we performed gene expression profiling comparing CRCs with matched tumor-free colon mucosa (n = 7). A multi-dimensional scaling plot showed a clear separation of tumor and colon mucosa samples (Figure 5A). Gene set enrichment analysis (GSEA) revealed significant similarities with chronic inflammation signatures in the human colon.
mucosa (Figure 5B-I, II) as well as in human MSI-positive inflammatory-CRCs (Figure 5B-III). IPA yielded a list of genes within the disease and disorders category that were significantly overrepresented in VCMsh2/Tgfbr2 CRCs, including genes in pathways involved in inflammatory response and gastrointestinal and metabolic disease (Table 8 in Figure 5). IPA disease and function annotations within the inflammatory response disease category predicted increased recruitment and activity of distinct immune cell types within the VCMsh2/Tgfbr2 CRC microenvironment (Table 9 in Figure 5). Consistent with this, CD45+ cells were significantly enriched within the tumor microenvironment compared with matched colonic mucosa (Figure 5C). We observed significant macrophage infiltration throughout CRC tissues and increased focal T cell infiltration, whereas B cell presence was limited to the lymphoid tissue associated with the colon tumors (Figure 5D). Upstream regulator IPA further predicted increased activation of proinflammatory cytokines that act as tumor growth and survival factors in human IBD-CRCs including interleukin (IL)-6, tumor necrosis factor α (TNF-α), and IL-17A (Table 10 in Figure 6). Their overexpression was confirmed by quantitative polymerase chain reaction (qPCR) (Figure 6A). TNF-α expression was elevated in tumor-free mucosa and significantly increased in tumors. TNF-α staining was particularly widespread in the colonic mucosa compared with SI mucosa in VCMsh2/Tgfbr2 CRCs (Figure 6B-I, III). A similarly increased TNF-α staining was

Table 10. IPA analysis of TNF-alpha, IL6 and IL17

<table>
<thead>
<tr>
<th>Upstream regulator</th>
<th>Predicted activation state</th>
<th>Activation z-score</th>
<th>p-value of overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF-alpha</td>
<td>activated</td>
<td>7.117</td>
<td>7.29E-70</td>
</tr>
<tr>
<td>IL6</td>
<td>activated</td>
<td>5.692</td>
<td>3.07E-29</td>
</tr>
<tr>
<td>IL17A</td>
<td>activated</td>
<td>3.89</td>
<td>1.02E-14</td>
</tr>
</tbody>
</table>

Figure 6. Characterization of inflammatory factors in the VCMsh2/Tgfbr2 CRCs tumor microenvironment. (A) qPCR. TNF-α: wild-type (WT) vs VCMsh2/Tgfbr2 CRCs colon (n = 7), *P = .01. IL-6: WT (n = 3) vs VCMsh2/Tgfbr2 CRCs (n = 5), *P = .036; VCMsh2/Tgfbr2 CRCs vs VCMsh2/Tgfbr2 CRCs colon (n = 5), **P = .0079. IL-17A WT (n = 5) vs VCMsh2/Tgfbr2 CRCs (n = 4), *P = .015; VCMsh2/Tgfbr2 CRCs vs VCMsh2/Tgfbr2 CRCs colon (n = 4), *P = .028. (B) Representative images of TNF-α staining: I, VCMsh2/Tgfbr2 colon tumor-free mucosa; II, VCMsh2/Tgfbr2 CRC (scale bar = 20 μm); III, VCMsh2/Tgfbr2 CRC (scale bar = 50 μm). (C) Representative images of staining for IL-6 and IL-17 comparing VCMsh2/Tgfbr2 colon tumor-free mucosa and tumor (scale bars = 200 μm, 50 μm). (D) GSEA of VCMsh2/Tgfbr2 CRCs vs matched colon mucosa with a signature indicative of TGFβ1-treated fibroblast. (E) qPCR. Tgfb1, CRCs fold change relative to tumor-free colon (n = 5). **P = .0079. (F) iNOS staining comparing VCMsh2/Tgfbr2 colon mucosa and CRC in VCMsh2/Tgfbr2 CRCs (scale bar = 200 μm).
evident when comparing CRCs with SI tumors (Figure 6B-II, VI). The increased expression of IL-6 and IL-17 in VCMsh2/Tgfbr2 CRCs was also confirmed by IHC analyses (Figure 6C). A list of gene expression changes predicted to be associated with immune cell infiltration in VCMsh2/Tgfbr2 CRCs is shown in Supplementary Excel File S6.

IPA upstream regulator analysis in VCMsh2/Tgfbr2 CRCs also predicted an activated state for some genes involved in inflammation or inhibition of genes lost in human CRCs (Table 11 in Figure 6). The VCMsh2/Tgfbr2 CRC microenvironment also shared significant similarities with the gene expression signature of human fibroblasts exposed to TGFB1 and its increased expression was verified by qPCR (Figure 6D and E). iNOS/NOS2 was found widely expressed in VCMsh2/Tgfbr2 CRC tissue, as predicted by the IPA activation analysis, consistent with an inflammatory CRC environment (Figure 6F).

To investigate specifically epithelial CRC cells, we performed additional transcriptomic analyses of sorted epithelial cells to exclude extensive infiltrating immune cells. GSEA highlighted similarities between the gene expression profiles in VCMsh2/Tgfbr2 CRC epithelial cells and transcriptional signatures derived from human cancer cell lines with PTEN knockdown or oncogenic KRAS overexpression (Figure 7A). IPA further showed a significant enrichment in molecules within the disease and disorders category belonging to cancer and gastrointestinal diseases (Table 12 in Figure 7). Several genes with altered expression profiles have also been reported in human CRC and IBD or mouse models of colitis (Table 13 in Figure 7).

Figure 7. Epithelial tumor cells expression profile analysis of VCMsh2/Tgfbr2 CRCs. (A) GSEA of sorted epithelial cells from VCMsh2/Tgfbr2 CRCs and matched tumor-free colon mucosa. (B) qPCR. Nlrp6: WT (n = 3) vs VCMsh2/Tgfbr2 CRCs (n = 5), *P = .035; VCMsh2/Tgfbr2 colon (n = 5) vs VCMsh2/Tgfbr2 CRCs (n = 5), **P = .0079.
Nlrp6 was of particular interest due to its roles in IBD-CRC and colonic microbial regulation. Indeed, we found Nlrp6 mutated in a VCMsh2/Tgfbr2 CRC (Figure 4C) and also generally downregulated in VCMsh2/Tgfbr2 CRCs, which was confirmed by qPCR (Figure 7B).

**Microbiota Composition Modulates the Susceptibility of VCMsh2/Tgfbr2 Mice to IBD-CRC**

To study the impact of the microbiota on the susceptibility to IBD-CRCs, VCMsh2/Tgfbr2 mice were rederived into specific pathogen–free (SPF) recipients and aged in an SPF environment. VCMsh2/Tgfbr2-SPF mice had a significantly longer median survival (12 months) compared with VCMsh2/Tgfbr2 mice in the conventional barrier (8 months) (Figure 8A). The SPF environment significantly affected tumor distribution: while 38% (n = 19 of 50) of SPF mice developed CRCs (compared with 74.15% of conventional mice) (Table 14 in Figure 8), a larger proportion developed SI adenocarcinomas (60% [n = 30 of 50]) compared with conventionally aged mice (23.6% [n = 21 of 89]). The colonic mucosa in CRC bearing VCMsh2/Tgfbr2-SPF mice displayed lower inflammation (Figure 8B) and expressed significantly less TNF-α, indicating a reduced ongoing proinflammatory state (Figure 8C). VCMsh2/Tgfbr2-SPF mice that developed only SI tumors displayed the lowest inflammation. These results suggested that the intestinal

![Table 14. Comparison of tumor incidence](image)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>N</th>
<th>Mice with colon tumor</th>
<th>Mice with SI tumor</th>
<th>Mice with colon + SI tumor</th>
<th>Mice with no tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCMsh2/Tgfbr2</td>
<td>89</td>
<td>48 (53.93%)</td>
<td>21 (23.59%)</td>
<td>18 (20.22%)</td>
<td>2 (2.24%)</td>
</tr>
<tr>
<td>VCMsh2/Tgfbr2-SPF</td>
<td>50</td>
<td>11 (22%)</td>
<td>30 (80%)</td>
<td>8 (16%)</td>
<td>1 (2%)</td>
</tr>
</tbody>
</table>

![Figure 8. Microbiota composition modulates the susceptibility of VCMsh2/Tgfbr2 mice to IBD-CRC.](image)

(A) Kaplan-Meier analysis. VCMsh2/Tgfbr2 mice (n = 89) vs VCMsh2/Tgfbr2-SPF mice aged in the SPF barrier (n = 50), P < .0001. (B) Inflammation score. VCMsh2/Tgfbr2 mice with CRC (n = 24), VCMsh2/Tgfbr2-SPF mice with CRC (n = 15), VCMsh2/Tgfbr2-SPF mice with SI tumors (n = 7); no significant differences. (C) qPCR. TNF-α: VCMsh2/Tgfbr2 (n = 7) vs VCMsh2/Tgfbr2-SPF colon (n = 6), **P = .008. (D) Relatedness (i̲-diversity) of fecal microbiota composition between the 2 different cohorts of young VCMsh2/Tgfbr2 mice (n = 3 conventional mice and n = 6 SPF mice). Principal coordinate (PC) analysis plot on Bray-Curtis distance matrix shows a clear separation between the 2 barriers (analyses of similarities test, R = -1, P = .014).
microbiota in VCMsh2/Tgfbr2 mice modulate the severity of colonic inflammation and susceptibility to CRC.

Therefore, we compared the microbiota from conventional barrier and SPF mice at 3 months of age (young and young-SPF, respectively) to determine the initial microbiota composition associated with either the increased or reduced IBD-CRC development by sequencing the 16S ribosomal RNA (rRNA) amplicon in fecal DNA. Differences in global microbiota relatedness were investigated by β-diversity analysis using multiple dissimilarity metrics (Table 15 in Figure 8), which showed clear separation of the VCMsh2/Tgfbr2 microbiota in the 2 barriers (Figure 8D). In young mice, major taxonomic differences specific to each barrier microbiota showed unique distributions between the 2 major phyla: while microbiota in conventional barrier mice were highly enriched in Bacteroides and reduced in Firmicutes, in SPF mice, the ratio of Bacteroides over Firmicutes was significantly reduced (Figure 9A and 9B). Specifically, at the genus level, conventional barrier microbiota contained higher relative proportions of Prevotella, Helicobacter, and...
Desulfovibrio, while the SPF microbiota had barely detectable levels of these pathogenic genera (Figure 9C). In contrast, no measurable levels of Operational Taxonomic Units (OTUs) for other bacteria such as Peptococcus and Clostridium XVIII were found in conventional barrier microbiota compared with SPF microbiota (Figure 9D).

We then analyzed the microbiota associated with spontaneous IBD-CRC development in conventional barrier mice and observed a marked increase in Verrucomicrobia. Three genera were significantly more abundant in conventional barrier VCMsh2/Tgfbr2 mice carrying inflammation-associated CRCs: Akkermansia, Coprobacillus, and Acetatifactor (Figure 9E). These genera were previously found associated with human and mouse CRCs.31-37 Interestingly, Akkermansia did not show any significant expansion in VCMsh2/Tgfbr2 mice that developed only SI tumors, but it was significantly more abundant in VCMsh2/Tgfbr2 mice with IBD-CRCs (Figure 9F and G).

TGFBR2 Status and Microbiota Composition
Affect Dextran Sulfate Sodium–Induced Colonic Inflammation and dMMR Colorectal Tumorigenesis

To assess the effects that either the TGFBR2 status or microbiota composition has on dMMR IBD-CRC development in a procolitogenic environment, VCMsh2, VCMsh2/Tgfbr2, and VCMsh2/Tgfbr2-SPF mice were exposed to DSS. The dextran sulfate sodium (DSS) treatment induced a 100% (n = 20 of 20) penetrant CRC phenotype in VCMsh2/Tgfbr2 mice with a median survival of 6 months (Figure 10A, Table 16 in Figure 10). In contrast, both the presence of TGFBR2 or the SPF environment had a protective effect against DSS-induced colorectal tumorigenesis, with VCMsh2 and VCMsh2/Tgfbr2-SPF mice displaying median survival of 9.3 and 13 months, respectively (Figure 10A). The increased survival was associated with reduced CRC incidence: 56.25% (n = 9 of 16) of VCMsh2 mice and 53.3% (n = 8 of 15) of VCMsh2/Tgfbr2-SPF mice developed DSS-induced CRCs (Table 16 in Figure 10). While TGFBR2 expression in VCMsh2 mice caused only a moderate reduction in DSS-induced inflammation compared with VCMsh2/Tgfbr2 mice, a significant reduction was seen in VCMsh2/Tgfbr2-SPF mice (Figure 10B). Furthermore, VCMsh2 mice rapidly gained weight upon cessation of DSS treatment, while VCMsh2/Tgfbr2 mice recovered at a significantly slower rate (Figure 10C). Tumor incidence analysis performed at early time points after cessation of DSS treatment showed that VCMsh2/Tgfbr2 mice had already developed CRCs during treatment (4 CRcs in 5 mice at 12 days; 6 CRcs in 5 mice at 30 days). In contrast, in VCMsh2 mice, only 1 adenoma was found at 12 days (1 in 3 mice) and none at 30 days post-DSS treatment (0 in 4 mice) (Table 17 in Figure 11).

The DSS-induced CRcs in VCMsh2 and VCMsh2/Tgfbr2 mice presented distinct macroscopic and histological features. While VCMsh2/Tgfbr2 mice developed mucinous adenocarcinomas similar to the CRcs during spontaneous tumor development (Figure 11A-I, II), VCMsh2 mice developed mainly adenomas with a polyplike appearance and low mucinous content (Figure 11B-I, II). All DSS-induced CRcs in VCMsh2 mice showed nuclear beta-catenin accumulation (n = 5) (Figure 11B-III), while none of the DSS-induced VCMsh2/Tgfbr2 CRcs showed nuclear beta-catenin (n = 5) (Figure 11A-III). These results demonstrate that both TGFBR2 signaling and a stable SPF environment protect from DSS-induced dMMR colorectal tumorigenesis.

TGFBR2 Loss and the Microbial Ecosystem
Cause Distinct Dysbiotic Shifts During DSS-Induced CRC Development

Because the DSS-induced CRcs in VCMsh2 and VCMsh2/Tgfbr2 mice display distinct phenotypic and molecular characteristics, we investigated whether they were associated with specific dynamic changes in the fecal microbiota. At 8 weeks of age, before DSS treatment, β-diversity analysis revealed significant separation of VCMsh2 and VCMsh2/Tgfbr2 microbiota with a low dissimilarity value (Figure 12A, Table 18 in Figure 12). After DSS-induced CRC development, the microbiota in the 2 mouse lines separated by a considerably greater dissimilarity value (Figure 12A, Table 18 in Figure 12) indicating that the TGFBR2 status has a significant impact on microbiota development over time. Specifically, DSS-induced CRC tumorigenesis was associated with a significant reduction in α-diversity in VCMsh2/Tgfbr2 mice but not in VCMsh2 mice (Figure 12B). β-Diversity measured by Bray-Curtis principal components metrics revealed significant differences in the microbiota in DSS-treated VCMsh2/Tgfbr2 mice compared with their pre-treatment composition (Figure 12C, Table 18 in Figure 12). Although DSS also caused a significant separation of microbiota in VCMsh2 mice, the observed dissimilarity between untreated and treated mice was much lower (Figure 12D, Table 18 in Figure 12).

The taxonomic composition in response to DSS-induced CRC within both mouse lines showed distinct genera expansion–contraction profiles (Figure 13A and B). The microbiota in DSS-treated VCMsh2/Tgfbr2 mice with mucinous CRcs were characterized by significant increases in the proportional abundances of Bacteroides, Parabacteroides, Akkermansia, and Desulfovibrio (Figure 13A). In contrast, VCMsh2 mice with nonmucinous CRcs showed pronounced abundance in Alistipes, Escherichia/Shigella, and Turicibacter genera but only a minimal increase in Akkermansia (Figure 13B). We also observed a reduction of distinct genera in both mouse lines: VCMsh2/Tgfbr2 microbiota were characterized by reductions in Lactobacillus and Rikenella abundance, whereas VCMsh2 mice mainly showed reduction in Barnesiella and Enterobadus abundance (Figure 13A and B).

Finally, we tested how DSS treatment affected the microbiota composition specific to CRC development in the restricted SPF environment. No reduction of α-diversity in VCMsh2/Tgfbr2-SPF microbiota was evident (Figure 13C), with only a discrete but significant difference in β-diversity (Figure 13D). The genera significantly affected by DSS...
treatment of VCMsh2/Tgfbr2-SPF mice included a significant increase in *Escherichia/Shigella* together with a significant decrease in *Mucispirillum* and *Roseburia* (Figure 13E). When the microbiota in DSS-treated VCMsh2/Tgfbr2 mice was compared with that of DSS-treated VCMsh2 or VCMsh2/Tgfbr2-SPF mice, an increase in *Bacteroides* and *Parabacteroides* but not in *Escherichia/Shigella* appeared to be a distinctive feature of IBD-CRC development in VCMsh2/Tgfbr2 mice (Figure 13F and G).

### Discussion

In the gastrointestinal tract, TGFβ signaling plays a central role in the control of cell growth and differentiation, the maintenance of intestinal homeostasis and, importantly, can either promote or suppress inflammation and cancer formation. Indeed, disruptive mutations in several TGFβ signaling pathway factors are found in human CRCs and colitis-associated CRCs. The aim of the current study was to determine how TGFBR2 mutations that occur at high
frequency in MSI-CRCs affect histological and molecular features of dMMR-driven colon tumorigenesis in an animal model. The loss of TGFBR2 signaling in VCMsh2/Tgfr2 mice not only induced a mucinous phenotype in intestinal dMMR tumors, but also increased intestinal inflammation and the susceptibility to CRCs with histopathological features characteristic of human IBD-CRCs. The mutational spectra and gene expression profiles in VCMsh2/Tgfr2 CRCs indicated that the combined effects of dMMR and loss of TGFBR2 signaling led to the disruption of intestinal homeostasis during the development of IBD-CRCs. The gene expression changes in VCMsh2/Tgfr2 CRCs indicate that several genes and pathways could be candidates for further investigation in preclinical studies. Importantly, an SPF environment had a protective effect lowering the incidence of VCMsh2/Tgfr2 IBD-CRCs, suggesting that elements of the gut microbiota play a crucial role in the development of these tumors.

The microenvironment of VCMsh2/Tgfr2 IBD-CRCs displayed increased immune cell infiltration and proinflammatory cytokine release including IL-6, TNF-α, and IL-17A, similar to human IBD-CRCs. Inflammatory processes induce oxidative DNA damage, which in part is repaired by MMR. Consistent with this, the mutational spectra in VCMsh2/Tgfr2 IBD-CRCs were representative of not only unrepaird replication errors but also unrepaird oxidative DNA lesions. The IBD-CRCs in VCMsh2/Tgfr2 mice showed no Apc mutations either during spontaneous or DSS-induced colorectal tumorigenesis in contrast to the DSS-induced CRCs in VCMsh2 mice or SI tumors in both mouse lines. This is reminiscent of most IBD-CRCs in human patients that do not harbor APC mutations or seem to acquire them at later stages during tumor development and indicates that TGFBR2 loss at the early stages of tumorigenesis is crucial for their etiology or progression. However, other genes with known roles in intestinal homeostasis or colorectal tumorigenesis carried mutations in VCMsh2/Tgfr2 CRCs. These include mutations in the PPARα/RXRα pathway, which plays a protective role against inflammation and colitis-associated colorectal tumorigenesis. Mutations in Ctnnd1, Acvr2a/1b, and Hnf1a were also found, suggesting that the loss of intestinal epithelial adherens junction signaling integrity is a feature in dMMR IBD-CRCs. Overall, the mutational landscape in VCMsh2/Tgfr2 IBD-CRCs shared significant similarities with human MSI-positive CRCs and included many genetic alterations that were reported in human IBD-CRCs or mouse models of colitis, such as mutations in Nrlp6, Arid1a, Kmt2c, and Sirt2.

IPA bio-function analysis revealed several gene expression changes specific to VCMsh2/Tgfr2 CRCs. For example, genes involved in liver fibrosis were upregulated, suggesting a role for fibrosis in remodeling the tumor microenvironment during prolonged inflammation in VCMsh2/Tgfr2 IBD-CRCs. Widespread expression of iNOS/NOS2 in VCMsh2/Tgfr2 IBD-CRCs (Table 11 in Figure 6, Figure 6F) was observed reminiscent of elevated iNOS expression in human IBD-affected intestines. The IBD-CRC microenvironment also shared significant similarities with the gene expression signature of human fibroblasts exposed to TGFβ1 (Figure 6D and E), which is also found in CRC patients at higher risk of relapse after therapy. These results are consistent with the notion that while TGF signaling suppresses the initiation of tumorigenesis in intestinal epithelial cells, at later stages, the increased release of

*Table 17. Comparison of tumor incidence*

<table>
<thead>
<tr>
<th>Genotype</th>
<th>N</th>
<th>Number of colon tumors</th>
<th>12 days post DSS</th>
<th>N</th>
<th>Number of colon tumors</th>
<th>30 days post DSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCMsh2/Tgfr2 +DSS</td>
<td>5</td>
<td>4 adenocarcinoma</td>
<td></td>
<td>5</td>
<td>6 adenocarcinoma</td>
<td></td>
</tr>
<tr>
<td>VCMsh2 -DSS</td>
<td>3</td>
<td>1 adenoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 11. The effect of Tgfr2 loss of function in DSS-induced dMMR CRCs.** (A) CRCs in VCMsh2/Tgfr2 mice treated with DSS (objectives magnification): I, macroscopic mucinous CRCs. II, Alcian blue staining showing blue mucin lakes (scale bar = 200 μm). III, Beta-catenin membrane-bound staining (scale bar = 50 μm). (B) CRCs in VCMsh2 mice treated with DSS. I, CRC with a polyp-like appearance. II, Alcian blue staining showing prominent villous or tubulovillous histology and absence of mucin lakes (scale bar = 200 μm). III, Beta-catenin nuclear accumulation (scale bar = 50 μm).
TGFB by malignant and noncancer cells within the tumor microenvironment can promote cancer progression/invasion and therapeutic resistance.\(^4^5\) Several in vitro studies tested the concept of restoring TGFBR2 signaling in microsatellite unstable CRC cell lines to suppress tumorigenesis but observed different outcomes.\(^4^6\)–\(^4^8\) While some studies found it reduced tumorigenicity, another study observed an increase in metastatic potential. A limitation of these studies is that they were mainly performed in vitro, and therefore the effect of Tgfbr2 restoration in CRC cells within the native tumor microenvironment has not been tested. It is also likely that restoration of TGFBR2 signaling at late stages in tumorigenesis will have protumorigenic effects or not be able to overcome the effects of existing mutations in tumor driver genes or the impact of severe inflammation and microbial dysbiosis that are associated with MSI-high IBD-CRCs development.

Figure 12. Tgfbr2 loss and the microbial ecosystem cause distinct dysbiotic shifts during DSS-induced CRC development. (A) Principal coordinate (PC) analysis plots of \(\beta\)-diversity analysis based on Bray-Curtis metrics comparing VCMsh2/Tgfbr2 and VCMsh2 mice before and after DSS treatment. Significances are shown in Table 18. (B) \(\alpha\)-Diversity measured before and after DSS treatment: VCMsh2/Tgfbr2 mice: Chao1 index, \(P = .002\); Shannon index, \(P = .014\); VCMsh2 mice \(P > .05\) (ns). (C, D) \(\beta\)-diversity by PC analysis of Bray-Curtis dissimilarity distances showing a significant separation before and after DSS treatment in VCMsh2/Tgfbr2 and in VCMsh2. Analyses of similarities test statistics in Table 18.
Figure 13. Dysbiotic shifts during DSS-induced VCMsh2/Tgfbr2 tumorigenesis. (A, B) Box plots showing the relative abundance of genera significantly different after DSS treatment in VCMsh2/Tgfbr2 mice and VCMsh2 mice. (C) No significant differences in α-diversity in VCMsh2/Tgfbr2-SPF mice before and after DSS treatment. (D) Principal coordinate (PC) analysis plots of β-diversity analysis based on Bray-Curtis metrics in VCMsh2/Tgfbr2-SPF before and after DSS treatment, analyses of similarities R = 0.26, P = .033. (E) Bar plots showing the relative abundance of genera significantly different after DSS treatment in VCMsh2/Tgfbr2-SPF mice. (F, G) Box plots of relative abundance comparing DSS-treated VCMsh2/Tgfbr2 mice with either DSS-treated VCMsh2 or VCMsh2/Tgfbr2-SPF mice, indicating common genera shifts signature (Supplementary Excel File S12).
The transcriptional signatures also indicated that the loss of MMR and TGFBR2 signaling had a disruptive impact on the overall functionality of epithelial CRC cells. This included changes in genes regulating intestinal stem cell fate: for example, Cdx2, a transcriptional regulator of cell fate specification and differentiation, and Spdef, which controls the differentiation and maturation of goblet cells. Changes in the expression of genes involved in intestinal epithelium permeability (Gfrt, Sle4a4, Sult1a1, Aqp1, Cldn15) indicate damage to the epithelial barrier as seen in human IBD epithelia. Nrp6 was mutated or downregulated in VCMsh2/Tgfbr2 CRCs (Figures 4C and 7B). Nrp6 is of particular interest since Nrp6 knockout mice show increased susceptibility to chemically induced colitis and inflammation-associated intestinal tumorigenesis. Interestingly, Nrp6 knockout mice display enrichment of A. muciniphila, similar to VCMsh2/Tgfbr2 IBD-CRCs.

A. muciniphila is the only known species of the Akkermansia genus in mice and humans with a mucin degrading activity, suggesting that this bacterium finds a favorable environment in the mucinous IBD-CRCs of VCMsh2/Tgfbr2 mice, which supports the hypothesis that specific genetic alterations induce a unique CRC microenvironment and thus modulate microbiota composition.

The colonic mucosa in VCMsh2/Tgfbr2 mice was more inflamed compared with their SI mucosa, which could be linked to the more dense and complex microbial communities in the colonic tract and the presence of pathobiont bacteria, gut species that exert pathogenic effects in particular genetic or environmental contexts. A critical role for pathogenic genera such as Helicobacter and Desulfovibrio in the development of VCMsh2/Tgfbr2 IBD-CRCs was observed by housing mice in an SPF environment devoid of these bacterial genera, which significantly reduced CRC incidence and increased survival. A recent study analyzing the effect of nonsteroidal anti-inflammatory drugs on intestinal tumorigenesis in another VCMsh2/Tgfbr2 mouse line did not report increased intestinal inflammation or development of IBD-CRCs. While no molecular analyses of intestinal tumors or microbiota were reported, the absence of IBD-CRCs in this mouse line could be caused by differences in microbiota composition in different animal facilities. For example, different Smad3 knockout mouse lines displayed different CRC phenotypes, which were proposed to be caused by the presence or absence of specific pathogenic microbiota members such as H. hepaticus. Recent studies have suggested more complex interactions with entire communities of pathogenic bacteria being responsible for the observed CRC phenotypes. Interestingly, while the inactivation of Tgfbr2 was reported to not cause the development of CRC in other studies, we found that a small number of VCTgfr2 mice that we generated during our experiments developed mucinous CRCs (Figure 1C-VI), further indicating the impact of pathogenic bacteria in the conventional microbiota on CRC development. However, while some VCMsh2/Tgfbr2 mice in the conventional barrier did not develop IBD-CRCs and the IBD-CRC incidence varied over time in our experiments, VCMsh2/Tgfbr2-SPF mice that developed IBD-CRCs displayed higher inflammation scores compared with mice that did not develop these tumors. This suggests that besides the microbiota, other factors, such as diet or genetic background, that can affect inflammatory processes in the colon are likely also involved in IBD-CRC development.

The analysis of microbiota alterations under DSS-induced colitogenic conditions using global metrics such as α- and β-diversity, and taxonomic analyses, provided insights into how comparable levels of intestinal epithelial damage can affect microbial dysbiosis depending on the Tgfbr2 genotype or the baseline microbial composition.

Neither α-diversity nor β-diversity was greatly affected during DSS-induced colorectal carcinogenesis in TGFBR2-expressing VCMsh2 mice or when TGFBR2-deficient VCMsh2/Tgfbr2 mice were housed under SPF condition. This analysis also identified specific genera that characterize the microbiota signatures in the different mouse lines during CRC development: while DSS-induced CRCs in VCMsh2/Tgfbr2 mice were associated with a marked increase in the abundance of Akkermansia, Desulfovibrio, Bacteroides, and Parabacteroides, DSS-induced CRCs in both VCMsh2 and VCMsh2/Tgfbr2-SPF mice were associated with pronounced enrichment of Escherichia/Shigella. This suggests that the DSS-induced epithelial damage that ultimately enhances exposure to luminal microbes and induces intestinal inflammation is modulated by both the synergistic effects of Msh2 and Tgfbr2 inactivation and the initial microbiota composition. Similar to these findings in IBD-CRC bearing VCMsh2/Tgfbr2 mice, both fecal and mucosal biopsies in LS patients with colorectal adenomas contained increased abundance in Desulfovibrio. Interestingly, a role for this microbe in inducing proinflammatory responses and DNA damage has been proposed. This study also found similar association between MSH2 mutations and enrichment in Coprobacillus.
will not only affect their histopathological features and the associated microbiota, but also may ultimately have important implications for diagnosis and treatment of these tumors.

**Materials and Methods**

**Animal Models**

Msh2\(^{loxP}\) mice were described previously.\(^{66}\) Villin-Cre transgenic mice were obtained from the Jackson Laboratory (Bar Harbor, ME) (B6.D2-Tg Villin-Cre). Tgfbr2\(^{loxP}\) and Fabp1-Cre mice were obtained from the National Cancer Institute Frederick Mouse Repository (strain code: 01XN5; B6.129S6 and 01XD8; FVB/N, respectively).\(^{67,68}\) All mice were genotyped by PCR as previously described\(^{66-68}\) and kept in conventional housing conditions under animal use protocol 00001194 approved by the Institutional Animal Care and Use Committee at the Albert Einstein College of Medicine.

To conditionally inactivate the Msh2\(^{loxP}\) and Tgfbr2\(^{loxP}\) genes in the intestinal epithelium, mice of each line were individually mated with Villin-Cre mice. Villin-Cre;Msh2\(^{loxP/loxP}\) (VCMsh2) or Villin-Cre;Tgfbr2\(^{loxP/loxP}\) (VCTgfbr2) mice were made to create VCMsh2\(^{loxP/loxP}\)/Tgfbr2\(^{loxP/loxP}\) (VCMsh2/Tgfbr2) mice. Fabp1-Cre;Msh2\(^{loxP/loxP}\)/Tgfbr2\(^{loxP/loxP}\) (FCMsh2/Tgfbr2) mice were generated in a similar manner.

8-oxoG ELISA

OxiSelect Oxidative DNA damage ELISA kit (Cell Biolabs, San Diego, CA) was used to detect and quantify 8-oxoG lesions in genomic DNA. The quantity of 8-oxoG in samples was determined by comparing absorbance to an 8-oxoG predetermined standard curve. Briefly, the 8-oxoG test samples or the 8-oxoG standards were first added to an 8-oxoG conjugate-coated plate. After a brief incubation, an anti-8-oxoG monoclonal antibody was added, followed by a horseradish conjugated peroxidase secondary antibody. The absorbance was measured at 450 nm using a plate reader. All steps were performed according to the manufacturer instructions. Values were normalized to adjacent tumor-free mucosa.

**MSI Analysis**

MSI was evaluated by standard procedures using a panel of 3 microsatellite markers previously published\(^{70}\) in which each primer set contained 1 FITC-labeled primer. The amplification products were diluted 1:20 in water, and 2 \(\mu L\) of the diluted product was added to 7.5 \(\mu L\) of HiDi Formamide mixed with 0.5 \(\mu L\) of Genescan Liz 600 size standard (Life Technologies, Carlsbad, CA) in a 96-well PCR plate. This reaction mixture was denatured at 95°C for 3 minutes and rapidly chilled to 4°C. The plate was loaded onto a 3730 DNA Analyzer (Life Technologies, Carlsbad, CA) for separation via capillary electrophoresis and data collection. The raw data (.fsa files) were analyzed with GeneMarker Software (SoftGenetics, State College, PA). MSI was defined by comparing tumors and matched normal mucosa and scoring for differences in peaks.

**Statistical Analysis**

Values are expressed as mean ± SEM. The 2-tailed, unpaired, nonparametric Mann-Whitney test was used to
evaluate significance between samples. Survival distributions were statistically compared by log-rank (Mantel-Cox) test.

**Exome Sequencing**

DNA was extracted from tumor and matched liver tissue using the DNA tissue extraction Kit (Qiagen) and evaluated for quality using the Pico green kit following the manufacturer instructions. Library preparation, exome sequencing and analysis were performed using standardized methods.

Genomic libraries from tumors and individually matched germline (liver) DNA were generated, enriched for exonic sequences using Agilent’s SureSelect Mouse All Exon V1 target enrichment kit (S0276129; Agilent, Santa Clara, CA) and sequenced using Illumina NovaSeq6000 S4 (2 × 150bp; Illumina, San Diego, CA). Design files describing enriched regions were converted from mm9 to mm10 coordinates using the UCSC online tool hgLiftOver. Raw fastq files were downloaded from the Psamagen server and file integrity confirmed by md5sum value.

Flanking adapters were removed (Trim Galore, v0.3.7; https://github.com/FelixKrueger/TrimGalore) and sequence quality assessed (FastQC v0.11.4; Fastq Screen v0.4.4 and summarized with MultiQC v1.7; https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/).

Reads were aligned to the mouse mm10 genome (bwa mem v0.7.15; http://bio-bwa.sourceforge.net/bwa.shtml); initial bam files were sorted and duplicates marked (Picard modules SortSam and MarkDuplicates v2.17.1; https://broadinstitute.github.io/picard/) and then filtered to retain properly paired reads with a MAPQ value ≥10 (samtools view v1.9).

Average coverage in the filtered bam files, over enriched regions described in the design file, was between 95x and 125x (GATK module DepthOfCoverage v4.1.7.0).

BQSR (base quality score recalibration) bam files were generated from the filtered bam files (GATK modules BaseRecalibrator and ApplyBQSR) and used as input for somatic variant calling. For each tumor–liver matched pair, somatic variants were initially identified using 3 variant callers, Mutect2 (GATK v4.1.7.0), Strelka2 (v2.9.10; assisted by Manta v1.6.0), and Lancet (v1.1.0). For each tumor–liver pair, pass-filter somatic variants identified by at least 2 of the variant callers (determined with bcftools module isec v1.9; http://www.htslib.org/doc/bcftools.html) were retained for further analysis and annotated using Snpeff and further manipulated using SnpSift (v4.3T). Data visualization was done using the MAFtools package. The cbioPortal platform was employed to analyze human mutation frequencies in MSI or MSS CRCs, combining all the available datasets from the bowel group. A full list of mutations for each sample is available in Supplementary Excel File S1. A full list of selected mutations from Din et al is available in Supplementary Excel File S2. IPA Bio-profile analysis list of genes previously found associated with colitis and/or IBD-CRC is shown in Supplementary Excel File S3.

**Quantitative PCR**

RNA extracted from tumor and tumor-free mucosa was used to synthesize complementary DNA using the SuperScript enzyme protocol (Invitrogen, Carlsbad, CA). VeriQuest SYBR Green PCR mix (Affymetrix, Santa Clara, CA) was used. Ribosomal U6 internal control was used to normalize the relative expression of each messenger RNA. Primers were used from previously published studies or designed to span 2 adjacent exons. Each sample was measured at least twice using a Viia7 machine (Thermo Fisher Scientific, Waltham, MA). Data are expressed as fold change relative to wild-type (Supplementary Excel File S12).

**Gene Expression Analysis**

RNA was isolated from tumor and adjacent tumor-free colon mucosa using the RNasey Mini-Kit (Qiagen). RNA extraction from sorted epithelial intestinal cells was performed using the RNasey Plus Micro Kit (Qiagen). RNA quantity was determined using a 2100 Bioanalyzer (Agilent).

Gene expression analysis of RNA isolated from bulk tissue or sorted epithelial cells was performed using the Mouse Gene ST 1.0 Array System (Affymetrix). Microarray data were preprocessed and analyzed using R-Bioconductor software (R Foundation for Statistical Computing, Vienna, Austria). Raw data were background corrected and RNA normalized, and expression values calculated (oligo package) were followed by annotation with the pd.mogene.1.0.st.v1 package (Carvalho B (2015). Pd.mogene.1.0.st.v1: Platform Design Info for Affymetrix MoGene-1.0-st-v1. R package version 3.14.1). Hierarchical clustering was performed using heatmap.2 to examine the pairwise correlations among all datasets (Pearson’s R). Statistical comparisons were made between the tumor and normal mucosal groups by linear modelling using the limma package in Bioconductor. The output of this analysis was saved to a GCT file (using exprs2gct) for further exploration. Genes evaluated as exhibiting significant differences in expression (as determined by the standard Volcano plot method) were subjected to further analysis for known and predicted regulatory relationships using IPA. In addition, GSEA was employed to investigate similarities with defined signatures from the GSEA Molecular Signatures database or from signatures that were extrapolated from previously published studies. Enrichment scores, normalized enrichment score, and P values were reported using GSEA analysis metrics.

A full list of genes is available in Supplementary Excel File S4 (bulk tissue) and S5 (sorted epithelial cells). A comprehensive list of growth factors and cytokines found to be changed in expression or predicted to be activated/inactivated by IPA upstream analysis is shown in Supplementary Excel File S6. A list of genes changed in expression involved in fibrosis is shown in Supplementary Excel File S6.
16S rRNA Amplicon Sequencing

Fecal DNA was extracted from mouse stool using the QIAamp DNA Stool Mini Kit (Qiagen). The 16S V3-V4 region was amplified using universal primers 341F-CTTACGGGNGGCWGCCAG and 805R-GACTACHVGGGTATCTAATCC86; libraries were prepared using the Nextera XT index kit (Illumina). Sequencing was performed using the Illumina MiSeq platform (2 x 300 bp reads/mouse sample).

Bioinformatics and Statistical Analysis of 16S rRNA Sequencing Data

Microbiota composition analysis of 16S rRNA amplicon sequencing data was conducted following standard computational approaches, as previously described.85 Paired-end sequencing reads (2 x 300 bp) were joined using FLASH version 1.2.10 primers were removed using cutadapt v1.8.386 (Martin, M. (2011)). Cutadapt removes adapter sequences from high-throughput sequencing reads.86

EMBNet J 17, 10–12). Demultiplexing and quality filtering were performed using the QIIME package version 1.9.1.87 Duplicated sequences and sequences out of a limited range of 373-473 nt were discarded using the USEARCH v8.1.86.88 After singleton removal, the remaining sequences were clustered into OTUs at a 97% identity level and chimera filtering was carried out withUCHIME89 against the GOLD reference database. Taxonomic classification of representative out sequences was carried out using mothur v1.36.190 to genus level against the 16S rRNA reference of RDP trainset 16, and using SPINGO version 1.391 to species levels against the RDP v11.2 database. ToutOTU table was generated with USEARCH by mapping the quality-filtered sequences against the representative OTUs sequenout. The OTU read counts were rarefied to the lowest read count in the dataset of 8244 reads for diversity calculations. A phylogenetic tree for UniFrac calculations was created using QIIME. α-Diversity indices (Shannon and Chao1) and β-diversity indices (Bray-Curtis, Weighted UniFrac, and Unweighted UniFrac) were generated using QIIME.

Statistical analysis and data visualization were carried out in R v4.0.4. Statistical differences for α-diversity were tested using Mann-Whitney U tests for 2 independent groups and Kruskal-Wallis tests for more than 2 independent groups. Principal coordinate analysis plots were created to explore β-diversity based on Bray-Curtis, Weighted UniFrac, and Unweighted UniFrac distance metrics using the ade4 package v1.7-15 with dudi.pco function. Two-dimensional principal coordinate analysis plots were created using the ggplot2 package v2.2.1. To test for statistical differences in β-diversity, analyses of similarities were carried out using the anosim function from the vegan package v2.5-6. The barplots showing different taxonomic level classification were created using the ggplot2 package. Taxa below 1% sample abundance and the unclassified taxa were grouped into the “other” category. Differentially abundant taxa were tested using DESeq2 package v1.30.1.92 P values from multiple pairwise comparisons were adjusted using the Benjamini-Hochberg method. Significance was assumed for adjusted P values <.05 if not stated otherwise.

Data availability: NGS run details for 16S sequencing data: PRJN759725. See Supplementary Excel File S7 for statistical analysis relative to Figure 9A and C–E. See Supplementary Excel File S8 for genus counts relative to Figure 9C and D. See Supplementary Excel File S9 for statistical analysis relative to Figure 9F. See Supplementary Excel File S10 for statistical analysis relative to Figure 13A, B, and D. See Supplementary Excel File S11 for statistical analysis relative to Figure 13F and G.

qPCR for A. muciniphila Analysis

The abundance of A. muciniphila was evaluated by real-time PCR using 16S species specific primers93,94 and normalized to total bacteria 16S rRNA using universal EUB primers55,56. Each sample was measured at least twice. For statistical analysis, analysis of variance was applied. Supplementary Excel File S12 lists all the primers97,98 and antibodies used in this study.

References


FLA 5.6.0 DTD ■ JCMGH1025 proof ■ 23 June 2022 ■ 1:49 pm ■ cc OB


22 Tosti et al

Cellular and Molecular Gastroenterology and Hepatology Vol. No.


Correspondence
Elena Tosti, PhD, Department of Cell Biology, Albert Einstein College of Medicine, 1301 Morris Park Avenue, Bronx, NY 10461. e-mail: elena.tosti@einstein.edu; fax: 718-430-8574.

Acknowledgments
The authors thank Bo Jin for assistance in mouse husbandry. The microarray data can be found with the accession number GSE183749. The BioProject accession number for the exome sequencing data is PRJNA759725. The BioProject accession number for the 16S sequencing data is PRJNA760488. The BioProject accession number for the 16S sequencing data is PRJNA670488.
Methodology: Lead; Project administration: Equal; Visualization: Lead; Writing – original draft: Lead; Writing – review and editing: Lead
Ana S. Almeida, PhD (Data curation: Supporting; Investigation: Supporting; Methodology: Supporting; Writing – review and editing: Supporting)
Tam T.T. Tran, PhD (Data curation: Supporting; Methodology: Equal; Software: Equal; Visualization: Equal; Writing – review and editing: Supporting)
Mariel Barbachan e Silva, MS (Data curation: Supporting; Methodology: Equal; Software: Equal; Visualization: Equal)
Pilib Ó Broin, PhD (Data curation: Supporting; Methodology: Equal; Software: Equal; Visualization: Equal; Writing – review and editing: Supporting)
Robert Dubin, PhD (Data curation: Supporting; Methodology: Equal; Software: Equal; Visualization: Equal; Writing – review and editing: Supporting)
Andrew S Mclellan, PhD (Data curation: Supporting; Methodology: Supporting; Software: Equal; Visualization: Supporting)
Aaron Golden, PhD (Data curation: Supporting; Methodology: Supporting; Resources: Supporting; Writing – review and editing: Supporting)
Paul W O’Toole, PhD (Conceptualization: Supporting; Supervision: Supporting; Writing – review and editing: Supporting)
Amanda P Beck, DVM (Formal analysis: Equal; Methodology: Equal; Visualization: Equal; Writing – review and editing: Supporting)
Ken Chen, MD (Methodology: Supporting)
Eduardo Vilar-Sanchez, MD (Resources: Supporting; Writing – review and editing: Supporting)
Ken Chen, MD (Methodology: Supporting)
Eduardo Vilar-Sanchez, MD (Resources: Supporting; Writing – review and editing: Supporting)

Conflicts of Interest
The authors disclose no conflicts.

Funding
This work was supported by National Institutes of Health grants CA248536 and CA222358 (to Winfried Edelmann), a Feinberg Family Donation (to Winfried Edelmann and Eduardo Vilar-Sanchez), Cancer Center grant CA13330 (to Albert Einstein College of Medicine), and Department of Defense fellowship CA15080 (to Elena Tosti).