AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems®

RESEARCH ARTICLE
Host-Microbe Biology

Human Colon Mucosal Biofilms and Murine Host Communicate via Altered mRNA and microRNA Expression during Cancer

Sarah Tomkovich,[a] Raad Z. Gharaibeh,[a] Christine M. Dejea,[c,d] Jillian L. Pope,[a] Jinmai Jiang,[f] Kathryn Winglee,[e] Josee Gauthier,[a] Rachel C. Newsome,[a] Ye Yang,[a] Anthony A. Fodor,[e] Thomas D. Schmittgen,[f] Cynthia L. Sears,[c,d] Christian Jobin[a,b]

[a]Department of Medicine, University of Florida, Gainesville, Florida, USA
[b]Department of Infectious Diseases and Immunology, University of Florida, Gainesville, Florida, USA
[c]Bloomberg-Kimmel Institute of Immunotherapy, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins Medical Institutions, Baltimore, Maryland, USA
[d]Department of Oncology and Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
[e]Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, North Carolina, USA
[f]Department of Pharmaceutics, College of Pharmacy, University of Florida, Gainesville, Florida, USA

Sarah Tomkovich and Raad Z. Gharaibeh are co-first authors, and Cynthia L. Sears and Christian Jobin are co-senior authors. The order of the first two authors was determined by drawing straws.

ABSTRACT Disrupted interactions between host and intestinal bacteria are implicated in colorectal cancer (CRC) development. However, activities derived from these bacteria and their interplay with the host are unclear. Here, we examine this interplay by performing mouse and microbiota RNA sequencing on colon tissues and 16S and small RNA sequencing on stools from germfree (GF) and gnotobiotic *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice associated with microbes from biofilm-positive human CRC tumor (BF+T) and biofilm-negative healthy (BF-bx) tissues. The bacteria in BF+T mice differentially expressed (DE) >2,900 genes, including genes related to bacterial secretion, virulence, and biofilms but affected only 62 host genes. Small RNA sequencing of stools from these cohorts revealed eight significant DE host microRNAs (miRNAs) based on biofilm status and several miRNAs that correlated with bacterial taxon abundances. Additionally, computational predictions suggest that some miRNAs preferentially target bacterial genes while others primarily target mouse genes. 16S rRNA sequencing of mice that were reassociated with mucosa-associated communities from the initial association revealed a set of 13 bacterial genera associated with cancer that were maintained regardless of whether the reassociation inoculums were initially obtained from murine proximal or distal colon tissues. Our findings suggest that complex interactions within bacterial communities affect host-derived miRNA, bacterial composition, and CRC development.

IMPORTANCE Bacteria and bacterial biofilms have been implicated in colorectal cancer (CRC), but it is still unclear what genes these microbial communities express and how they influence the host. MicroRNAs regulate host gene expression and have been explored as potential biomarkers for CRC. An emerging area of research is the ability of microRNAs to impact growth and gene expression of members of the intestinal microbiota. This study examined the bacteria and bacterial transcriptome associated with microbes derived from biofilm-positive human cancers that promoted tumorigenesis in a murine model of CRC. The murine response to different microbial communities (derived from CRC patients or healthy people) was evaluated through RNA and microRNA sequencing. We identified a complex interplay between biofilm-associated bacteria and the host during CRC in mice. These findings may lead to the development of new biomarkers and therapeutics for identifying and treating biofilm-associated CRCs.

Citation Tomkovich S, Gharaibeh RZ, Dejea CM, Pope JL, Jiang J, Winglee K, Gauthier J, Newsome RC, Yang Y, Fodor AA, Schmittgen TD, Sears CL, Jobin C. 2020. Human colon mucosal biofilms and murine host communicate via altered mRNA and microRNA expression during cancer. mSystems 5:e00451-19. https://doi.org/10.1128/mSystems.00451-19.

Editor David W. Cleary, University of Southampton

Address correspondence to Cynthia L. Sears, csears@jhmi.edu, or Christian Jobin, Christian.Jobin@medicine.ufl.edu.

Received 26 July 2019
Accepted 21 December 2019
Published 14 January 2020

KEYWORDS colorectal cancer, transcriptomics, microRNAs, microbiota, biofilm, germfree

Numerous 16S rRNA and shotgun metagenomic studies have demonstrated that colorectal cancer (CRC) patients have an altered intestinal microbiota compared to healthy controls (1, 2). Colibactin-producing *Escherichia coli*, enterotoxigenic *Bacteroides fragilis*, and *Fusobacterium nucleatum* among others, are implicated in CRC pathogenesis due to their abilities to produce genotoxins and adhesins which promote proliferation and modulate immune responses in preclinical models (3, 4). How these bacteria interact with the rest of the complex microbiota to influence CRC initiation and/or progression is still unclear. In fact, recent studies with *F. nucleatum* suggest these bacteria may be associated with later stages of disease and have less of an influence on CRC initiation (5–8). Testing the functional role of human CRC-associated bacterial communities in chemically induced mouse models of CRC have led to mixed results (9, 10). One group demonstrated stool communities from either individual CRC or healthy patients promoted polyp formation depending on the composition of the microbiota that established in mice (9). Another recent report revealed an increased tumorigenic phenotype in mice that received stools pooled from multiple CRC patients compared to stools from controls (10).

The lack of a consensus carcinogenic CRC-associated microbiota from patient stools suggests that other factors, including how the bacteria are organized/located or the genes they express, may be just as important to CRC pathogenesis. Polymicrobial bacterial biofilms, spanning $>$200 μm of epithelial surface were recently identified in $\sim$52% of human CRC patients (11) and also found in $\sim$13% of the healthy patients who were screened (12). We previously showed that human biofilm-forming bacterial communities from either CRC or healthy patients play a functional role in CRC development in multiple preclinical mouse models, emphasizing the contribution of bacterial organization to CRC (13).

CRC is an evolving disease, characterized by a series of molecular and microbial changes (14–16), suggesting a dynamic interplay between the host and intestinal microbiota as the disease progresses. MicroRNAs (miRNAs) have emerged as potential mediators of these host-microbe interactions with their ability to modulate both host (17) and bacterial genes, which can result in shifts in microbiota composition (18, 19). In turn, the microbiota is able to modulate host miRNA expression (18, 20, 21), with *F. nucleatum* targeting several miRNAs related to CRC pathogenesis (5, 22). However, it is uncertain how human CRC-associated microbial communities as a whole impact fecal miRNA expression and whether host miRNAs affect bacterial composition/gene expression during CRC.

To examine the bacterial activities associated with biofilm-positive microbes from CRC patients, we examined mouse and bacterial gene expression from colon tissues and mouse small RNA sequencing from stools collected from biofilm-positive associated $Apc^{Min\Delta 850/+}$;$Il10^{-/-}$ mice. We found that a number of bacterial virulence genes were increased in biofilm-positive communities and identified a conserved core group of transmissible biofilm-positive associated bacteria. Additionally, we demonstrate that biofilm status and CRC development alter miRNA expression and specific miRNAs correlate with biofilm-positive associated taxa.

RESULTS

Bacterial activities associated with biofilm status. In order to elucidate microbial activities associated with biofilm-forming bacteria derived from human CRC patients that promote tumorigenesis in $Apc^{Min\Delta 850/+}$;$Il10^{-/-}$ mice (13), we characterized mouse and microbial gene expression from colon tissue snips using RNA sequencing (see Fig. 1AI and II for experimental design). Principal-component analysis (PCA) of microbial community gene expression detected by both our *de novo* assembly and aligning the microbial transcriptome sequencing (RNA-seq) reads to the human gut microbiome

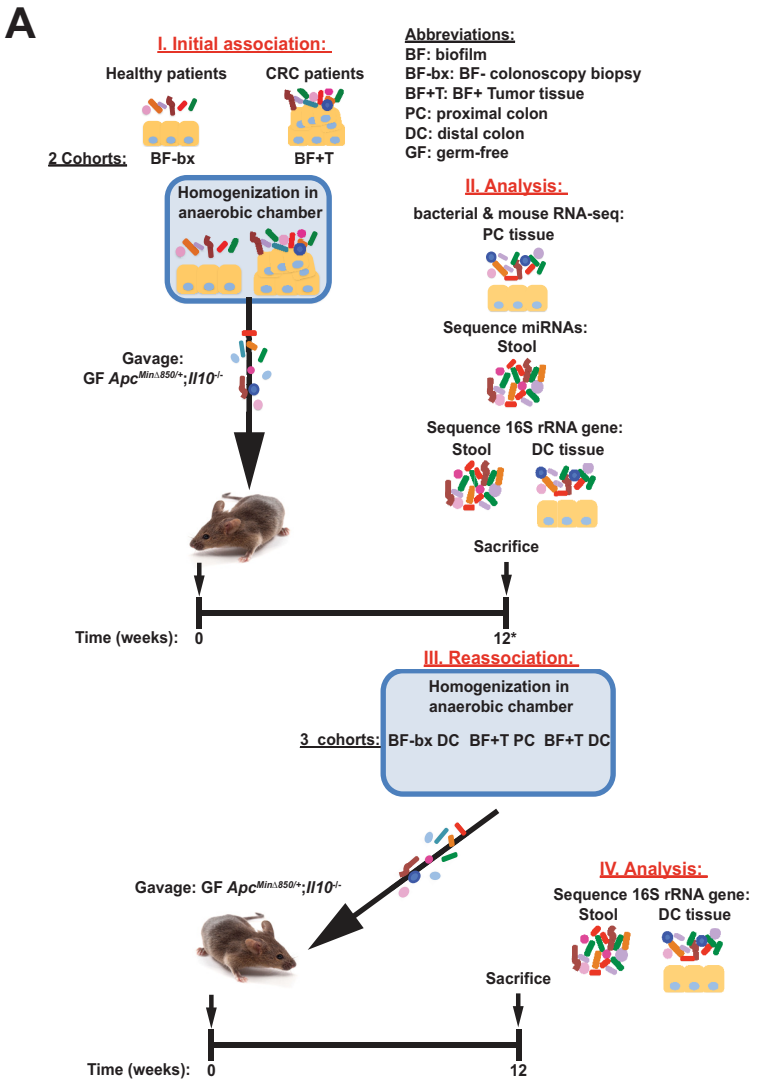


FIG 1 Experimental design and microbial community gene expression associated with biofilm status (related to Materials and Methods). (A) Schematic showing the setup of the gnotobiotic association (I)

(Continued on next page)

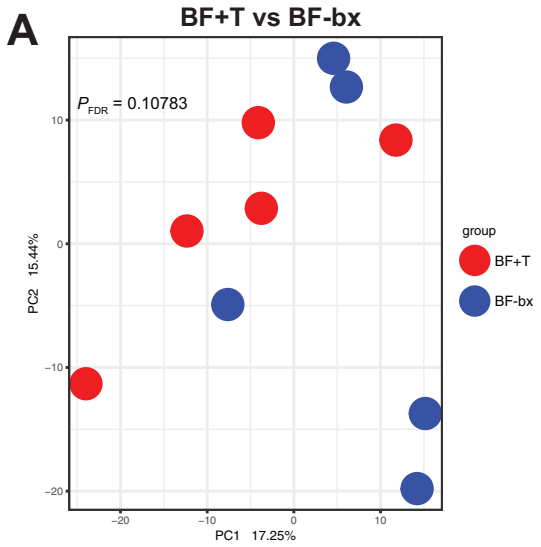
TABLE 1 Pathways enriched in BF+T samples using KEGG pathways from Trinity assembly or HUMAnN analysis of RNA-seq reads

Method and pathway	*P* value	*P*FDR
Trinity assembly		
Carbon fixation pathways in prokaryotes	1.06e−05	0.0006
Protein export	0.0002	0.0051
Bacterial secretion system	0.0002	0.0051
Valine, leucine, and isoleucine biosynthesis	0.0004	0.007
2-Oxocarboxylic acid metabolism	0.001	0.0131
HUMAnN analysis		
2-Oxocarboxylic acid metabolism	1.41e−05	0.0007
Inositol phosphate metabolism	1.42e−05	0.0007
Biosynthesis of amino acids	0.0004	0.0163
Flagellar assembly	0.001	0.0204
Protein export	0.0021	0.035

integrated gene catalog (IGC) showed separate clustering of biofilm-positive CRC tumor tissue (BF+T) associated *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice from biofilm-negative healthy patient tissue (BF-bx) associated mice (Fig. 1B and Fig. S2A in the supplemental material, respectively). Bacterial metatranscriptomic analysis found 2,918 significant differentially expressed (DE) genes (false-discovery rate-adjusted *P* value [P_{FDR}] < 0.05), the majority of which were increased in BF+T mice (2,739 increased genes and 179 decreased genes [see Table S1 at https://figshare.com/s/bd06b560c635de3ac830]) compared to BF-bx mice. Pathways related to protein export, bacterial secretion systems, carbon fixation, flagellar assembly, and biosynthesis of amino acids were increased in BF+T mice compared to BF-bx mice (Table 1 and Fig. S2B and C). Additional genes related to virulence and biofilm formation, including stress response, toxins, iron acquisition, mucin cleavage/transport, outer membrane polysaccharide importers, and adhesins were also significantly increased in BF+T mice (see Table S1 at https://figshare.com/s/bd06b560c635de3ac830). Increased toxin genes included *Clostridium difficile* toxins A and B, *Clostridium perfringens* Mu toxin, and *E. coli* colibactin (*clbG* and *clbI*). Weighted gene coexpression network analysis (23) identified 34 hub genes (see Table S1 at https://figshare.com/s/bd06b560c635de3ac830) from modules detected in BF+T mice and included outer membrane proteins involved in protein export and heat shock proteins involved in the stress response.

Despite the high number of microbial genes with increased expression in BF+T microbial communities, no separation by PCA analysis was found at the host gene expression level, and only 62 significant DE genes (Fig. 2A; see also Table S2 at https://figshare.com/s/4b593b780f756a4acf69) between BF+T and BF-bx mice were detected. Instead, the host was more responsive to the microbiota in general as opposed to the type of microbiota, since the host transcriptomes of either BF+T- or BF-bx-associated mice clustered separately from germfree (GF) mice (Fig. S1A and B). There were >3,000 significant DE host genes (~2,000 upregulated, ~1,300 downregulated) in the BF+T and BF-bx groups compared to GF mice (see Tables S3 at https://figshare.com/s/163676e591c87b0d6c35, S4 at https://figshare.com/s/652055bdb15e48b866ef, and S5 at https://figshare.com/s/c9adfcd56af278666c55) and pathway analysis revealed that the majority of upregulated genes in colonized mice belonged to immune-related pathways (Fig. S1C and D). Only the peroxisome proliferator-activated receptor (PPAR) signaling pathway was significantly upregulated ($P = 1.88e-05$, $P_{FDR} = 0.004$) in BF+T

FIG 1 Legend (Continued)
and reassociation (III) experiments, along with the analyses done on the stool and tissue samples (II and IV) at the end of the 12-week experiments. Twelve-week stool and/or DC tissue samples were used for RNA, miRNA, and 16S rRNA sequencing analyses (II). Tissue was collected from 12-week-associated BF-bx mice and 16- to 20-week-associated BF+T mice to make the reassociation inoculums (III). (B) PCA of bacterial transcriptomes from BF+T- and BF-bx-associated *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice generated from Trinity *de novo* assembly ($N = 5$ for BF+T and BF-bx).



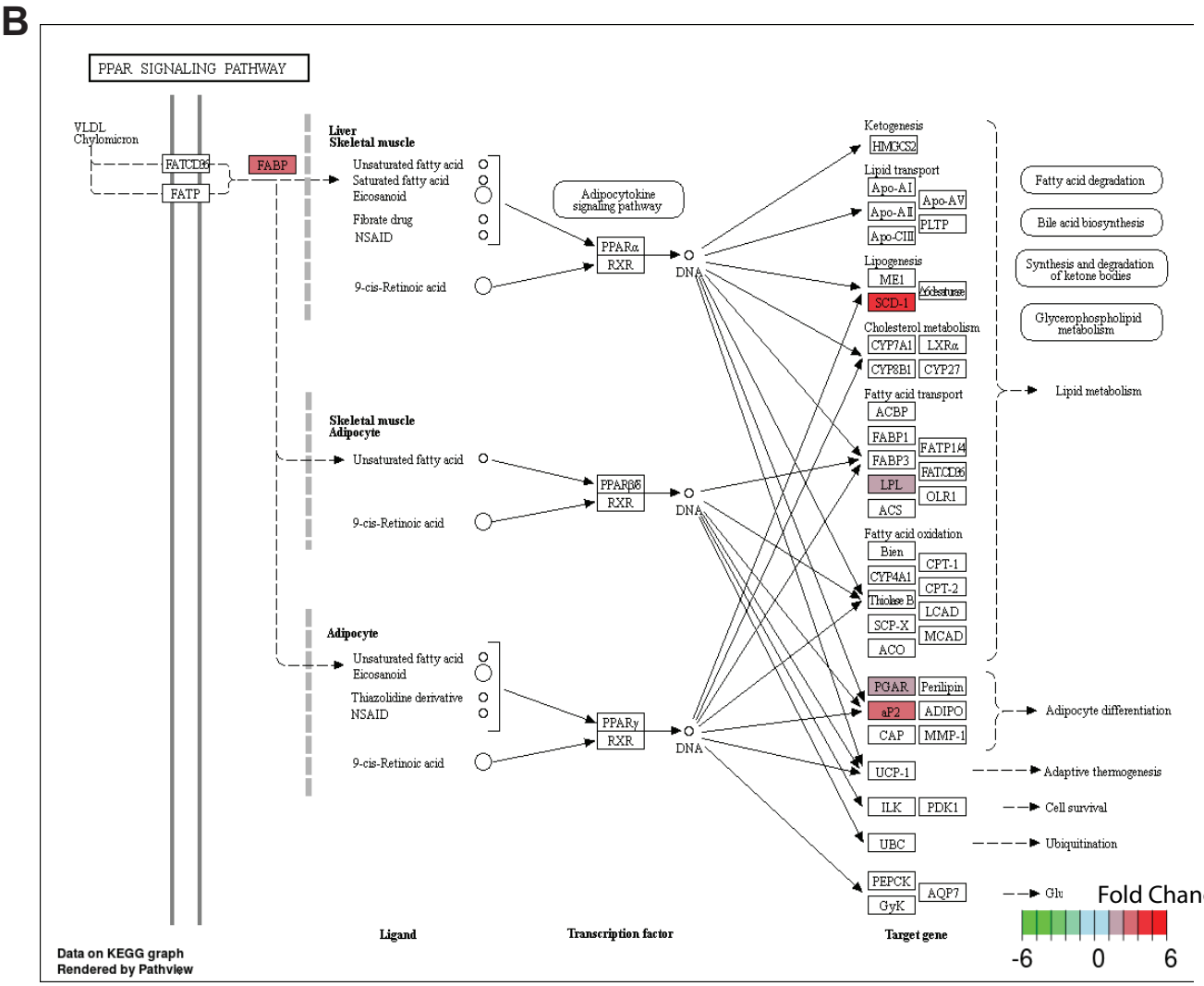


FIG 2 Mouse gene expression affected by biofilm status. (A) PCA of mouse transcriptomes from BF+T and BF-bx mice ($N = 5$ per group). (B) PPAR signaling pathway depicting genes significantly increased in BF+T associated mice in red. See Table S2 at https://figshare.com/s/4b593b780f756a4acf69 for a list of all significant DE genes in BF+T versus BF-bx associated mice. Boxes without color denote no significant change.

mice compared to BF-bx mice. The majority of the significant DE genes were upregulated (56 versus 6 downregulated) in BF+T mice and included genes related to lipid metabolism, iron scavenging, and the extracellular matrix (see Table S2 at https://figshare.com/s/4b593b780f756a4acf69). Cross-referencing the list of upregulated genes with The Cancer Genome Atlas (TCGA) colorectal microarray data using Oncomine (24) revealed 10 of the upregulated genes were also significantly overexpressed in colon adenocarcinomas compared to control tissues with SCD (stearoyl-coenzyme A [CoA] desaturase, PPAR pathway member), MMP10, and SLC22A3 increased more than 2-fold (see Table S6 at https://figshare.com/s/7248a5727e5972879086).

Biofilm status alters the fecal miRNA profile. Given the differential bacterial and host gene expression observed in BF+T mice that developed colon tumors and the potential of miRNAs to modulate interkingdom interactions (18), we next profiled stool miRNA expression. To examine the interplay between miRNAs and bacterial communities derived from human CRC or healthy patients, we sequenced the fecal small RNAs from GF, BF-bx, and BF+T *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice. The PCA plot shows clear separation between GF mice and BF-bx or BF+T-associated mice, demonstrating that microbial colonization alters the fecal miRNA profile in *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice (Fig. 3A; see also Table S7 at https://figshare.com/s/b7d7ff960f787eb31954). Biofilm/cancer status of the initial human-derived microbes also modulates miRNA expression since PCA analysis demonstrates separation of BF+T and BF-bx miRNAs (Fig. 3A; see also Table S7 at https://figshare.com/s/b7d7ff960f787eb31954). Pairwise comparisons between the three groups of mice revealed that 25 unique miRNAs were significantly DE (out of 142 total detected) (Fig. 3B; see also Tables S7 at https://figshare.com/s/b7d7ff960f787eb31954, S8 at https://figshare.com/s/2eaad72eea8a5890f7f2, and S9 at https://figshare.com/s/20a8d9c108e990242bc7). Next, we compared the significantly different miRNAs in BF+T or BF-bx versus GF mice and found that nine significant DE miRNAs overlapped (mmu-miR-6538, -146b-5p, -215-5p, -194-5p, -192-5p, -2137, and -5126 and mmu-let-7b-5p and mmu-let-7i-5p), suggesting host miRNAs targeting the microbiota (Fig. 3B; see also Tables S8 at https://figshare.com/s/2eaad72eea8a5890f7f2 and S9 at https://figshare.com/s/20a8d9c108e990242bc7). We were also able to identify eight miRNAs (mmu-miR-709, -690, -21a-5p, -142a-5p, -6240, -6239, and -148a-3p and mmu-let-7a-5p) that were significantly DE according to biofilm status (BF+T versus BF-bx [Fig. 3B; see also Table S7 at https://figshare.com/s/b7d7ff960f787eb31954]). These findings suggest that the microbiota modulates host miRNA expression. Together, both microbiota composition and disease status drive miRNA expression (BF+T versus BF-bx DE miRNAs).

Fecal miRNAs correlate with specific bacterial taxa and target bacterial and mouse genes. We next examined whether the fecal miRNAs of human microbiota-associated mice correlated with bacterial taxa previously identified in the mice (13). We identified 11 miRNAs that were significantly DE between GF, BF-bx, and BF+T mice and correlated with the relative abundances of bacterial taxa in the stool, distal colon tissue, or both compartments (Fig. 3B [colored green, orange, or underlined]; see also Tables S10 at https://figshare.com/s/55e156ec52d9d3e53394 and S11 at https://figshare.com/s/1c47b438d806062915c9). Additionally, we identified miRNAs that although not significantly DE between groups, correlated with five and eight bacterial genera in the stool and distal colon tissues, respectively (Fig. 4A and B; see also Tables S10 at https://figshare.com/s/55e156ec52d9d3e53394 and S11 at https://figshare.com/s/1c47b438d806062915c9). Five of these genera (*Bacteroides*, *Lachnospiracea incertae sedis*, *Anaerostipes*, *Clostridium* XVIII, and *Roseburia*) were significantly increased in BF+T mice (Fig. 4A and B; see also Tables S10 at https://figshare.com/s/55e156ec52d9d3e53394 and S11 at https://figshare.com/s/1c47b438d806062915c9). Furthermore, mmu-miR-140-3p correlated with *Lachnospiraceae incertae sedis* abundance and tumor number, both of which were increased in BF+T mice (Fig. 4A and Fig. S3A). Thus, CRC-associated microbial communities elicit a specific host miRNA profile and maintain a subset of miRNAs that correlate with specific microbial taxa.

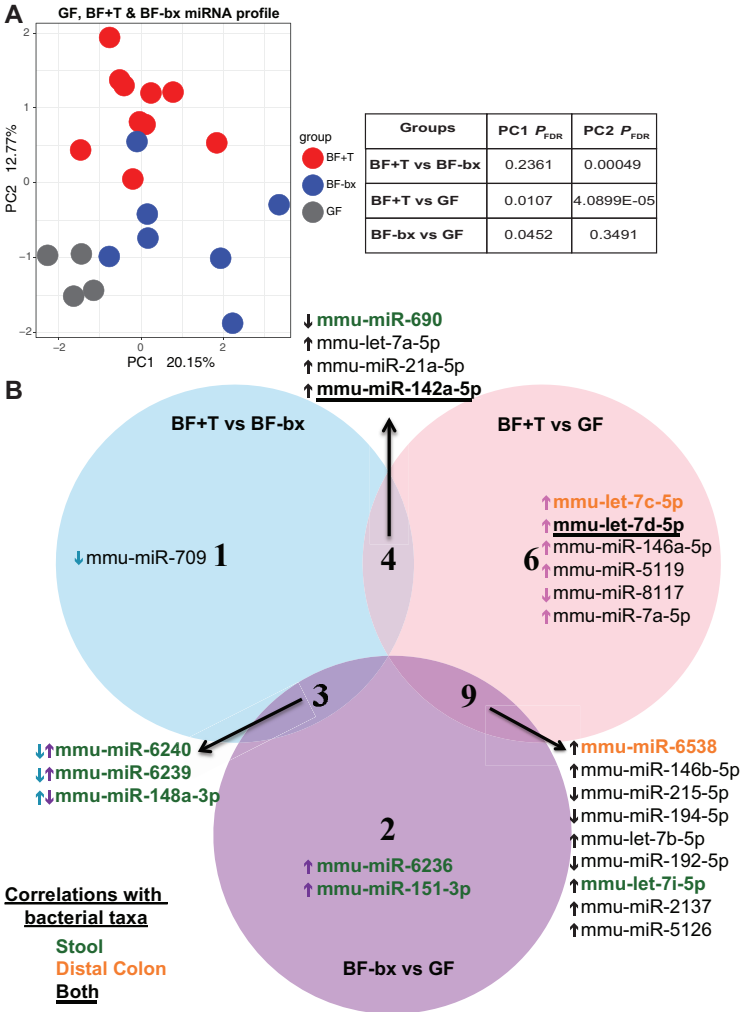


FIG 3 Fecal miRNA profiles and significant DE miRNAs according to biofilm status. (A) PCA plot of log-normalized mature miRNA counts in GF, BF-bx, and BF+T *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ stool ($N = 4$, 7, and 10 for GF, BF-bx, and BF+T, respectively). Symbols represent miRNAs from individual mice for BF-bx and BF+T groups. For the GF group, symbols represent miRNAs pooled from 2 to 5 mice (13 mice total, see Materials and Methods). (B) Significant ($P_{FDR} < 0.05$) DE miRNA between the three experimental groups. BF+T microbiota elicited more host miRNAs than the GF and BF-bx microbiota with six miRNAs uniquely DE in response to it. The arrows next to miRNA names indicate whether the miRNA is increased (↑) or decreased (↓) in the first group listed relative to the second for each comparison. Blue arrows indicate direction of miRNA expression in BF+T mice compared to BF-bx mice. Pink arrows indicate direction of miRNA expression in BF+T mice compared to GF mice. Purple arrows indicate direction of miRNA expression in BF-bx mice compared to GF mice. Black arrows indicate direction of miRNAs that are shared between comparison groups and go in the same direction. The miRNAs that are significantly correlated with taxa identified from 16S rRNA sequencing of the stool are highlighted in green, those that significantly correlated with taxa identified from the DC tissue are highlighted in orange, and those correlating with both are in black font and underlined (total of 11 miRNAs).

Interestingly, computationally predicting the bacterial genes targeted by the significant miRNAs revealed several miRNAs (mmu-miR-2137, mmu-miR-5126, and mmu-miR-6538) that primarily target bacterial genes (see Table S12 at https://figshare.com/s/33fb65da16763a5e3347). In contrast, other miRNAs (mmu-let-7s, mmu-miR-21a-5p, mmu-miR-142a-5p, mmu-miR-148a-3p, mmu-miR-194-5p, mmu-miR-690, and mmu-miR-709) are predicted to primarily target mouse genes (see Table S13 at https://figshare.com/s/be3f38607b2e9875e7f6), and there is a significant negative correlation between the number of predicted bacterial versus mouse gene targets for the significant miRNAs we identified (Fig. 4C and D and Fig. S3B and C). Taken together, these predictions suggest that specific miRNAs have differential roles in mediating host-

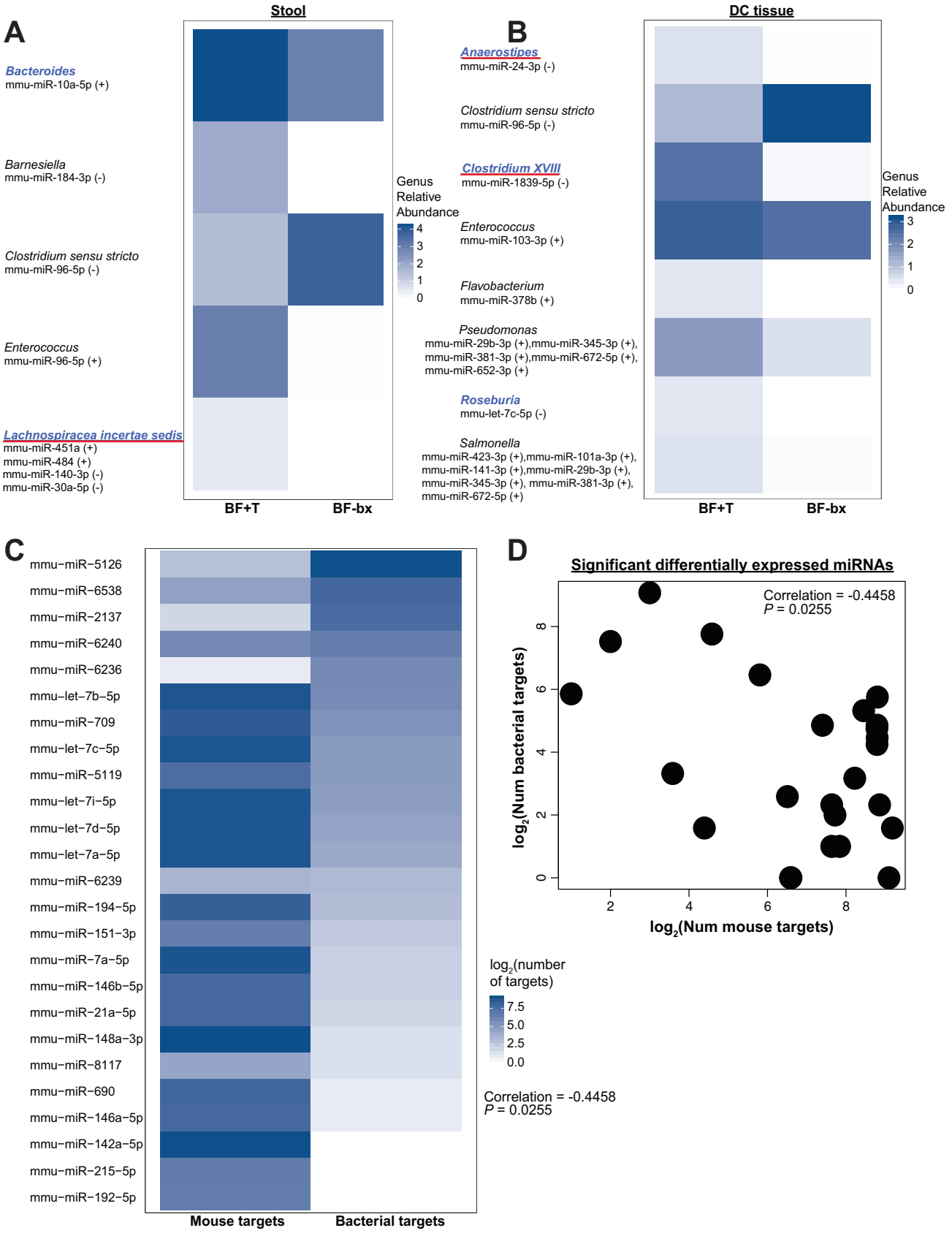


FIG 4 Microbial relative abundance at the genus level correlates with specific miRNAs. (A and B) Heatmap depicting the mean $\log_{10}$-normalized relative abundance of genera within the stool (A) or distal colon tissue (B) compartments that have significant ($P_{FDR} < 0.05$) correlation with miRNA expression.

(Continued on next page)

microbe interactions, with some mainly modulating host gene expression and others primarily impacting bacterial gene expression and/or abundance.

A core BF+T microbiota is transmissible. We previously showed that carcinogenic properties are retained over time as microbial inoculums from homogenized mouse colon tissues collected from the initial BF+T but not BF-bx tissue-associated microbes promoted tumors in new cohorts of GF *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice (13) (Fig. 1A.III).

We next examined microbial compositional differences between BF+T and BF-bx reassociated mice. 16S rRNA sequencing revealed separation of stool and distal colon (DC) tissue microbiota from reassociated mice according to the biofilm status of the initial association (Fig. 5A and B). Thirteen genera were significantly different (12 genera enriched and one genus depleted) in the stool and/or DC tissue of both the BF+T-associated and reassociated *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice (Fig. 5C), three of which also correlated with specific miRNAs (Fig. 4A and B). Six of these genera (*Clostridium* XVIII, *Erysipelotrichaceae incertae sedis*, *Escherichia/Shigella*, *Eubacterium*, *Parabacteroides*, and *Robinsoniella*) were increased in both the stool and DC tissue of BF+T-associated and reassociated mice compared to BF-bx mice. To determine how much the microbiota composition shifted after reassociation, we compared the microbiotas from the reassociated mice to the initial associated mice, whose tissues were used to generate the inoculums (Fig. S4 and S5). The BF-bx communities were highly transmissible, with no significant differences based on principal-coordinate analysis (PCoA) of the stool and DC tissue communities (Fig. S4). The BF+T communities shifted more after reassociation, with distinct clustering seen in the PCoA (Fig. S5A and B). Depending on the colon region the BF+T inoculum was derived from, there were 5 and 13 significantly different genera within the DC tissue or stool compartment; however, only 1 and 5 of these genera were significantly different based on biofilm status in the initial associations (Fig. S5C and D). We also examined how the location of the colon tissues used to make the BF+T reassociation inoculums impacts community composition by comparing the stool and DC tissue communities from mice reassociated with PC or DC tissue inoculums (Fig. S6A). We found that only 2 and 5 genera were significantly different (Fig. S6B) and only 2 of these genera (*Coprobacillus* and *Holdemania*) differed according to biofilm status in the initial associations. Out of the 12 genera that were increased in BF+T-associated and reassociated mice, only 1 (*Coprobacillus*) was not maintained in both groups of BF+T reassociated mice. Thus, regardless of the murine colon region of origin (proximal or distal), the majority of BF+T microbes (11 genera total; *Anaerostipes*, *Clostridium* XI, *Clostridium* XIVa, *Clostridium* XVIII, *Erysipelotrichaceae incertae sedis*, *Escherichia/Shigella*, *Eubacterium*, *Flavonifractor*, *Lachnospiraceae incertae sedis*, *Parabacteroides*, and *Robinsoniella*) are able to reestablish and promote cancer when transmitted to a new cohort of GF mice. Taken together, the data suggest there is a core set of bacteria and bacterial gene expression associated with biofilm-positive cancers.

DISCUSSION

In contrast to previous studies that tested the carcinogenicity of CRC-associated microbiotas by gavaging mice with stools from either CRC or healthy control patients (9, 10), mucosa-associated BF+T bacteria retain their carcinogenicity when transplanted into a new set of GF mice, regardless of whether they were extracted from the proximal or distal colon (13). The number of taxa that overlap between the initial

FIG 4 Legend (Continued)
The name of each miRNA is shown below the genus it correlates with. Genera in blue font are significantly different between BF-bx and BF+T *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice. The red underlined genera were significantly different based on biofilm status in both the initial association and reassociation experiments. The direction of correlation is shown within parentheses. Relative abundance data are from the subset of BF-bx and BF+T mice that were used for miRNA sequencing ($n = 7$ for BF-bx; $n = 10$ for BF+T). See Tables S10 at https://figshare.com/s/55e156ec52d9d3e53394 and S11 at https://figshare.com/s/1c47b438d806062915c9 for the full list of miRNAs that correlate with bacterial taxa and the corresponding correlation coefficients. (C) Heatmap comparing the $\log_2$-transformed number of predicted bacterial versus mouse gene targets for the miRNAs that were significantly DE between the GF, BF-bx, or BF+T group. There is a significant negative correlation (Pearson) between the number of predicted bacterial versus mouse gene targets for the set of significant DE miRNAs. (D) Scatter plot demonstrating the significant negative Pearson correlation between the $\log_2$-transformed number of mouse versus bacterial gene targets where each circle represents a unique mouse miRNA that was significantly DE between GF, BF-bx, or BF+T group.

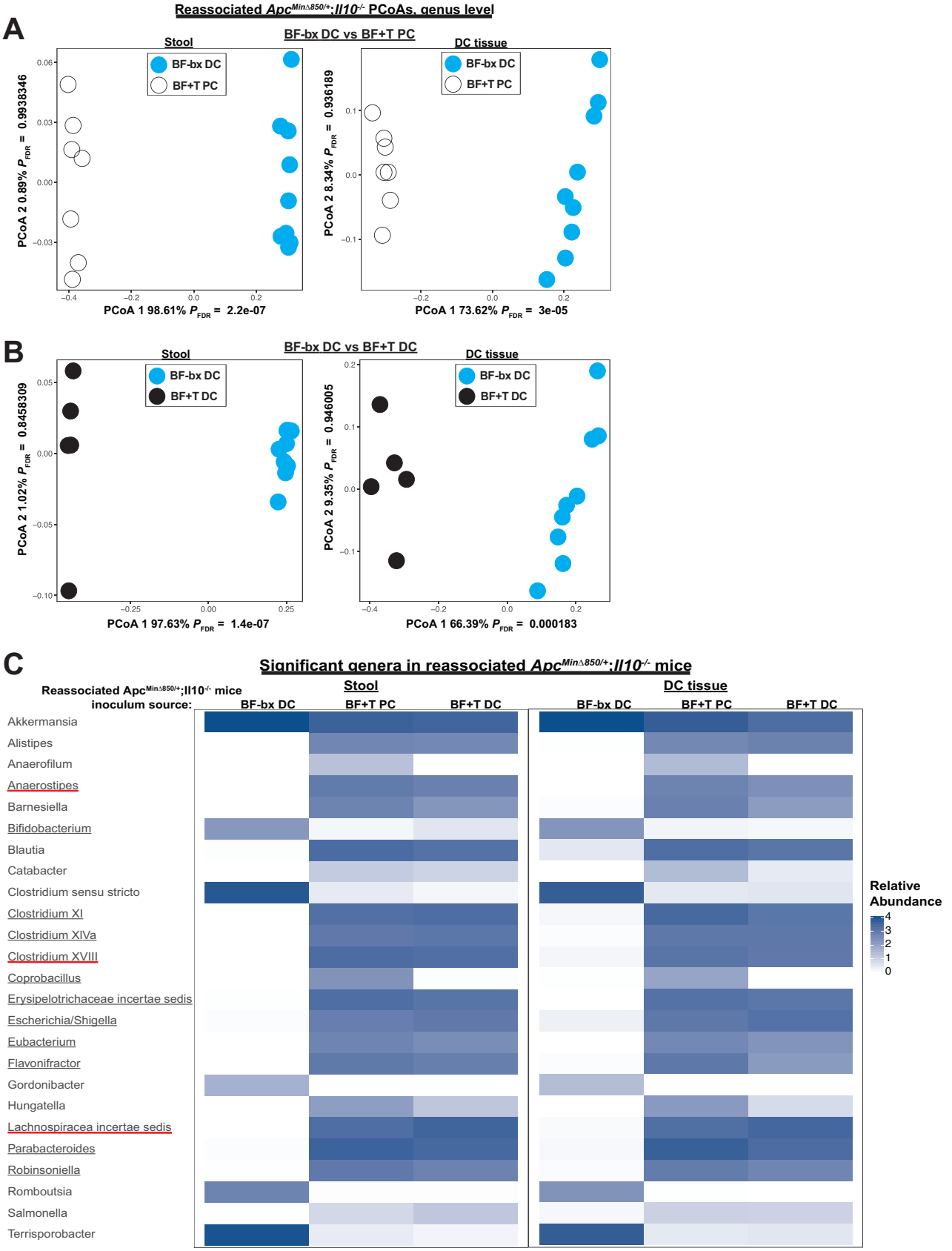


FIG 5 Biofilm status is associated with microbiota changes in reassociated *Apc*$^{Min\Delta850/+}$;*Il10*$^{-/-}$ mice. (A and B) PCoAs of BF-bx and BF+T reassociated mice, with the composition of the stool and DC tissue compartment on the left and right, respectively (N = 9, 7, and 5 for BF-bx DC, BF+T PC, and BF+T

(Continued on next page)

association and the reassociation experiments, shown herein (Fig. 5), suggest that there is a core, transmissible, cancer-promoting microbiota associated with biofilm-positive cancers.

Four of the BF+T core transmissible cancer-promoting genera overlap with human stool-derived taxa that established and were associated with cancer in mice. *Parabacteroides* correlated with high tumor numbers in a gnotobiotic azoxymethane (AOM)/dextran sulfate sodium (DSS) C57BL/6 model (9), and *Erysipelotrichaceae*, *Lachnospiraceae*, and a *Clostridium* XIVa derived from CRC patients were associated with polyp formation in an antibiotic-treated AOM C57BL/6 model (10). Additionally, increased *Erysipelotrichaceae*, *E. coli*, *Lachnospiraceae*, and *Parabacteroides* and decreased *Bifidobacterium* have previously been associated with human CRC patient samples by 16S rRNA gene sequencing studies (1, 25).

Metagenomic predictions generated from 16S rRNA data identified bacterial secretion systems and motility genes associated with human CRC stool communities (26) and host glycan utilization genes correlated with tumor numbers in human stool-associated AOM/DSS mice (9). Recent meta-analysis studies of metagenomes from CRC patient fecal samples identified gluconeogenesis, mucin degradation, and colibactin as associated with the CRC microbiome (27, 28). These bacterial pathways and genes were also increased in our BF+T metatranscriptome. PICRUSt analysis of biofilm-positive versus biofilm-negative human CRC tissues also demonstrated an increased sporulation capacity associated with biofilm-positive CRCs contributed by several taxa, among them the *Lachnospiraceae* family (11). Similarly, herein, we saw genera from the *Lachnospiraceae* family (*Anaerostipes*, *Clostridium* XIVa, *Lachnospiracea incertae sedis*, and *Robinsoniella*) and upregulation of multiple sporulation genes in mice transplanted with the BF+T community. There is also overlap between BF+T community gene expression and human periodontitis polymicrobial metatranscriptomes, particularly for genes related to iron acquisition, flagellar synthesis, and the stress response (29). These findings are interesting since biofilms have also been associated with periodontal disease and *F. nucleatum* and *Porphyromonas* spp. are associated with both oral biofilms and colon cancer (4, 11, 25, 30).

Multiple genes related to nutrient, envelope, DNA damage, and environmental stress responses were increased in the BF+T community that could be indicative of host immune pressures, but could also be associated with a competitive polymicrobial environment, a feature of biofilms (31). Host iron metabolism changes during inflammation and cancer can promote competition for iron within the intestinal microbiota (32). Multiple iron acquisition genes, including siderophores and transport receptors, were increased in BF+T mice (32). The expression of these metabolic and iron acquisition genes could be indicative of a low-nutrient environment, fostering interbacterial competition.

Bacterial adhesion genes are a critical colonization determinant and may also contribute to biofilm formation, in which attachment to host cells represents a key initiating step (33). There are a number of adhesins expressed in the BF+T community, including type I and IV pili, capsule genes, and proteins that bind to host extracellular matrix (ECM) components such as fibronectin and laminin. On the host side, BF+T mice exhibit upregulation of a laminin subunit and the ECM-degrading matrix metalloproteinase MMP10 (34), suggesting alterations to the host ECM. BF+T communities also expressed numerous moonlighting adhesins (such as flagellin, GroEL, DnaK, and elongation factor Tu), putative multifunctional proteins which have been demonstrated in some bacterial strains to bind host cells, mucin, or ECM components (35). Bacterial

FIG 5 Legend (Continued)
DC, respectively). (A) BF+T PC reassociation compared to the BF-bx DC reassociation. (B) BF+T DC reassociation compared to the BF-bx DC reassociation. (C) Heatmap depicting the mean $\log_{10}$-normalized relative abundances of genera that were significantly different in the stool and/or DC tissues of reassociated mice inoculated with murine colon tissue homogenates derived from human BF+T or BF-bx tissue-associated mice. The underlined genera were significantly different based on biofilm status in both the initial association and the reassociation, and the red underlines represent the three genera that correlated with specific miRNAs (Fig. 4A and B).

adherence has also been identified as an important feature of CRC-associated bacteria, including *F. nucleatum* (36) and *Streptococcus gallolyticus* subsp. *gallolyticus* (37), the latter of which is capable of forming biofilms on collagen IV, an ECM component (38). Furthermore, bacteria expressing adhesins that bind to ECM and host glycoproteins may have an additional colonization advantage, as host glycosylation is disrupted during inflammation and cancer with increases in sialylation and fucosylation that can result in decreased host cell adhesion to ECM components (39). Additionally, some of the effects of CRC-associated bacteria may be contact dependent; for example, colibactin-induced DNA damage requires direct contact between the bacteria and epithelial cells (40).

Coupled with its role in facilitating colonization and attachment, mucin also represents a source of nutrition for intestinal bacteria (41). Antibiotic treatment was shown to increase sialic acid levels in the lumen and promoted the expansion of pathogenic bacteria such as *C. difficile* and *Salmonella enterica* serovar Typhimurium (42). Interestingly, sialic acid and other mucin sugar cleavage and transport expression were increased in BF+T mice along with an increased abundance of *Clostridium* XI and *Salmonella*. The SusC and SusD outer membrane proteins, involved in oligosaccharide binding and transport (43), were also increased in the BF+T community (from *Bacteroides* and *Parabacteroides* spp.). The upregulation of stress responses, mucin, and other nutrient acquisition genes indicate environmental conditions that could in turn promote virulence expression. Nutrient- and iron-responsive global transcriptional factors such as cyclic AMP receptor protein and ferric uptake regulator (Fur), which were increased in BF+T mice, have also been shown to regulate bacterial virulence expression (44).

Shotgun metagenomic sequencing of patient stools has revealed that host glycan utilization and virulence factor genes are associated with the CRC microbiome (16) and that genes in these categories were also overexpressed in the BF+T microbial community. Host inflammation, bacterial iron (Fur), and stress response (Hsp90 chaperone) genes have all been implicated in colibactin regulation, and all of these genes were increased in BF+T mice (45–47). Additionally, iron acquisition genes have previously been associated with *E. coli* mucus colonization (48), and mucins have the capacity to induce *E. coli* virulence gene expression (41). Mucin and its components may also serve as a cue for virulence regulation of other BF+T community members, as they have also been linked to virulence regulation in *S. enterica* and *Campylobacter jejuni* (41). Although the expression of multiple *B. fragilis* genes were increased in the BF+T community, *B. fragilis* toxin (*bft*) was not detected. One possible explanation is that expression of the RprXY two-component system, recently implicated in *bft* suppression (49), was significantly increased in BF+T mice. However, even intermittent enterotoxigenic *B. fragilis* colonization as short as 2 weeks appears to be sufficient to induce tumor formation (50), so it is possible that *bft* expression occurred at an earlier time point in the $Apc^{Min\Delta 850/+}$; $Il10^{-/-}$ model. Taken together, the metatranscriptomic data suggest that the BF+T community expresses more pathogen-related virulence factors and metabolism genes that provide competitive advantages over commensals but may have detrimental side effects to the host.

Members of the *Erysipelotrichaceae* family, part of the core transmissible bacteria in BF+T mice, were also increased in the microbiota of Western or high-fat diet-fed mice (*Clostridium innocuum*, *Eubacterium dolichum*, and *Clostridium ramosum*) and were associated with increased fat deposition (51, 52). Diets high in fat and obesity are established risk factors for CRC (53). Conceivably, the activation of the PPAR signaling pathway within BF+T mice could be a response to colonization with these obesity-associated taxa. PPARs are nuclear hormone receptors that regulate key aspects of lipid and carbohydrate metabolism, including fatty acid synthesis, uptake, and storage (54), and have both suppressive and promotional effects in CRC (55).

Branched-chain amino acid (BCAA) biosynthesis was increased in the BF+T community, and serum BCAA have also been associated with metabolic disorders and

correlated with intestinal microbiota members such as *Bacteroides vulgatus* (56). BCAAs, which include valine, leucine, and isoleucine, may also contribute to fatty acid synthesis, which is also a feature of cancer metabolism (57). Valine and/or leucine secretion were previously associated with *E. coli* (58) and polymicrobial environmental biofilms (59). Polyamines were previously shown through metabolomics to be increased in CRC patients, with a rare polyamine, N1, N12-diacetylspermine detected in biofilm-positive CRCs (60), and we found that multiple microbial polyamine-related genes were increased in BF+T mice (61). Polyamines increase in proliferating cells and can promote tumor growth and invasion (62) and are also important to bacterial biofilm formation (61). Several transporter genes (*Slc22a3*, *Abcb1a*, and *Abcb1b*) that have been previously implicated in polyamine uptake (63) were upregulated in BF+T mice. Thus, biofilm-associated communities and their associated metabolism pathways have the potential to modulate host metabolism, which may promote cancer.

In addition to bacterial components and metabolites, miRNAs represent another mode of host-microbe interplay during cancer. Profiling the fecal miRNAs of $Apc^{Min\Delta 850/+}$;$Il10^{-/-}$ mice under different microbial conditions allowed us to identify specific miRNAs that were associated with biofilm/CRC status. A few of the CRC-associated miRNAs we identified have conserved sequences with human miRNAs (hsa-miR-21-5p, hsa-miR-142-5p, and hsa-miR-146a-5p) that are increased in CRC patients (17, 22, 64, 65). Mmu-miR-21a-5p was significantly increased in the BF+T mice, and *F. nucleatum* has previously been shown to increase miR-21 (22), suggesting that miR-21 may be targeted by multiple CRC-associated bacterial genera.

Although the depleted miRNAs in BF+T mice, miR-690 and miR-709, are not found in humans, they do share several CRC-related gene targets (such as *Ctnnb1*, *Il6ra*, *Stat3*, *Src*, and *Zeb1*) with other miRNAs depleted in CRC (17). Though none of these genes were significantly DE according to biofilm status in the mouse colon tissue RNA-seq data set, it is possible the luminal miRNAs might target other regions of the colon or specific intestinal epithelial cell types. Additionally, even though miRNAs have been shown to primarily control gene expression through mRNA degradation, translational repression is also possible (66). Alternatively, another mechanism of miRNA regulation could relate to targeting the RNA-induced silencing complex genes like *Argo1*, *Argo2*, *Argo3*, *Argo4*, *Cnot6*, and *Dcp2* (66), which are predicted targets of multiple significant DE miRNAs, the majority of which are increased in BF+T mice.

miRNAs also have the capacity to target bacterial genes and impact microbial composition (18), and our computational predictions indicate that newly discovered miRNAs (those with higher numbers in their names) preferentially target bacterial genes. These miRNAs include miR-2137, miR-5126, miR-6239, miR-6240, and miR-6538, which were also increased in DSS-treated mice (67), where microbiota composition also contributes to disease susceptibility (68). Many of these miRNAs were predicted to have redundant bacterial targets (including genes regulating motility, secretion, outer membrane proteins, stress response, iron acquisition, and carbohydrate utilization/transport) that overlap with genes that were increased in the BF+T microbial community. miR-6239 and -6240 were decreased in BF+T mice, but miR-2137, -5126, and -6538 were increased in mice, regardless of biofilm status.

Conclusions. The mechanisms by which miRNAs enter and regulate bacterial gene expression and how bacterial organization affects bacterial community gene expression warrant further investigation. A recent publication showed that plant-derived exosome-like nanoparticles contain miRNAs that altered intestinal microbial composition and gene function (69). Whether intestinal epithelial cell-derived exosomes carrying miRNAs could similarly target intestinal microbiota is under investigation. Our findings suggest a complex interplay between BF+T-associated bacteria, their gene expression, the host transcriptome, and miRNAs that may contribute to CRC pathogenesis (Fig. 6). Deciphering this complex interplay will likely identify new regulatory pathways and molecules with potential therapeutic implications.

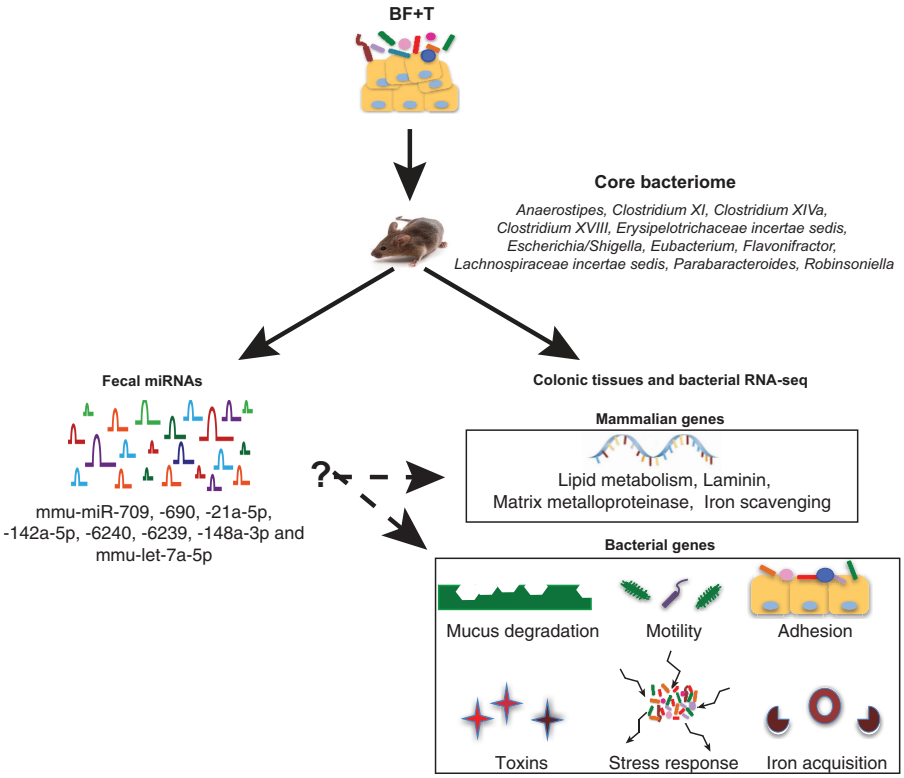


FIG 6 Schematic depiction of the major findings. The core, transmissible bacteria found in biofilm-positive tumor (BF+T) associated and reassociated mice are listed under core bacteriome. Some of the bacterial and mouse genes that were differentially expressed in the BF+T associated mice compared to biofilm-negative (BF-bx) associated mice are listed. Fecal miRNAs were also differentially expressed according to biofilm status, correlated with the relative abundances of some bacterial taxa, and were predicted to target mouse and bacterial genes.

MATERIALS AND METHODS

Animals. Germfree (GF) 129/SvEv *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice were transferred to separate gnotobiotic experimental isolators based on inoculum type for the duration of the association.

Initial associations with human tissue-associated microbes. GF 129/SvEv *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice were inoculated with pooled tissue-derived microbes from biofilm-negative tissues collected from healthy patients via colonoscopy biopsy (BF-bx) or biofilm-positive tumor tissues collected from CRC patients (BF+T) via surgical resection. Patient tissues were collected and screened for biofilm status via fluorescence *in situ* hybridization (FISH) (biofilm-positive criteria, polymicrobial, within the mucus layer, spanning 200 μm, and $>10^9$ bacteria/ml) as described previously (12). Tissues were analyzed with the universal bacterial probe (EUB338), and a nonsense probe (non338) was used as a negative control (12, 13). Additional FISH analysis was conducted on biofilm-positive tumor tissues with probes to detect *Bacteroidetes*, *Lachnospiraceae*, *Fusobacteria*, and *Proteobacteria* (12, 13). The probe sequences are listed in Table S4 in reference 12. Each inoculum was prepared anaerobically by homogenizing tissue (tissue pooled from five different patients) in phosphate-buffered saline (PBS), and FISH images for these tissues can be found in Fig. S1 in reference 13. Each mouse received 100 to 200 μl of inoculum, and associations were carried out for 12 to 20 weeks (Fig. 1AI). Tissues and/or stools from mice collected 12 weeks after association were used for transcriptome sequencing (RNA-seq), microRNA sequencing (miRNA-seq), and 16S rRNA gene sequencing analyses (Fig. 1AII).

Mouse reassociation inoculums. Mouse reassociation inoculums (Fig. 1AIII) were made from colon tissues from 12-week BF-bx-associated (cohort 2) and 16- to 20-week BF+T-associated (cohort 3) *Apc*$^{Min\Delta 850/+}$;*Il10*$^{-/-}$ mice (13). After the colon was flushed 1× with PBS, tissue snips were taken from both the distal colon (DC) and proximal colon (PC) and stored at −80°C until time of inoculum preparation. Each inoculum was prepared from colon tissue snips pooled from four mice. All inoculums were prepared anaerobically by mincing and homogenizing tissue snips in prereduced PBS. The BF-bx reassociation inoculum was prepared from inflamed (average inflammation score of 2.5/4) distal colon tissues (BF-bx DC). The two BF+T reassociation inoculums were prepared from mostly normal (average PC inflammation score of 0.9/4) proximal colon tissues (BF+T PC) or distal colon tissues (BF+T DC) from the same four mice with colitis and tumors (average DC inflammation score of 3.6/4; average number of tumors = 5.5, range = 3 to 10 tumors).

Reassociation with mouse tissue-associated microbes. Six- to 13-week GF 129/SvEv *Apc*$^{Min\Delta 850/+}$; *Il10*$^{-/-}$ (males and females) were transferred to gnotobiotic isolators (separate isolator for each exper-

imental group) and gavaged with 100 to 200 μl of inoculum (Fig. 1AIII). Mice were euthanized after 12 weeks, and the colon was flushed once with PBS and then cut and splayed longitudinally for macroscopic tumor counts. About 2 × 0.5 cm tissue snips from the proximal and distal colon were collected by flash freezing in liquid nitrogen and stored at −80°C until analysis.

Stool and distal colon tissue DNA extraction. Stool and DC tissue DNA (Fig. 1AII and IV) was extracted via phenol-chloroform separation by lysing the cells with phenol-chloroform-isoamyl alcohol (25:24:1) and 0.1-mm zirconium glass beads on a bead beater (Precellys), followed by phase separation with chloroform-isoamyl alcohol (24:1), DNA precipitation with ethanol, and subsequent purification with the DNeasy Blood & Tissue kit (Qiagen catalog no. 69506).

16S rRNA sequencing. The 16S rRNA V1-V3 hypervariable region was amplified using barcoded primer pairs 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 534R (5′-ATTACCGCGGCTGCTGG-3′) with universal Illumina paired-end adapter sequences (see Tables S14 at https://figshare.com/s/ffd8c3066929eb0c69fc and S15 at https://figshare.com/s/5e95b9a855710812950b). PCR products were purified, quantified, and pooled as described previously (7) and sequenced with an Illumina MiSeq (as two separate runs). The first 16S rRNA run (Run01, Fig. 1AII and Table S16 at https://figshare.com/s/97c3b0a48482b6fb0ecf) included stool and DC tissue samples from BF-bx cohort 1, BF+T cohorts 1 to 3, and two different biofilm-positive groups (13). Differentially abundant taxa were identified by comparing stool (1- and 12-week time points) and DC tissue (12-week time point) communities collected from BF-bx #1, BF+T#1-2, and two additional biofilm-positive groups reported in Tomkovich et al. (13). The second 16S rRNA run (Run02, Fig. 1AII and IV and Table S16 at https://figshare.com/s/97c3b0a48482b6fb0ecf) included stool and DC tissue samples from BF-bx cohort 2 and the three reassociation groups (BF-bx DC, BF+T PC, and BF+T DC). Comparisons between initial associations and reassociations were assessed by comparing the microbiota from mice whose tissues were used for the inoculums (BF-bx #2 and BF+T#3) to the reassociation microbiotas (BF-bx DC, BF+T PC, and BF+T DC).

16S rRNA sequencing analysis. Reads were preprocessed using Quantitative Insights into Microbial Ecology (QIIME) (70) version 1.9.1 including trimming and filtering at Q20. The final set of reads was fed to the RDP (Ribosomal Database Project) classifier (71) version 2.12 with the confidence set at 80%. Reads were grouped by genera, and samples with less than 1,000 total reads and genera with less than 5 reads were removed. The resulting counts were normalized and $\log_{10}$ transformed (72) using the following formula:

$$\log_{10}\left(\frac{RC}{n} \times \frac{\sum x}{N} + 1\right)$$

where *RC* is the read count for a particular taxon in a particular sample, *n* is the total number of reads in that sample, the sum of *x* is the total number of reads in all samples, and *N* is the total number of samples. The principal-coordinate analysis (PCoA) was generated from the Bray-Curtis distance of the normalized and $\log_{10}$-transformed counts using the phyloseq (73) R package (74).

Genera significant for biofilm group (BF-bx, BF+T, BF-bx DC, BF+T PC, and BF+T DC) were detected using the lme function in the R nlme package, with the REML method (75) to fit a mixed linear model of the form:

$$\text{genus} \sim \text{variable} + 1|\text{cage} + \varepsilon$$

where genus indicates the $\log_{10}$ normalized abundance of a particular genera, variable indicates either the biofilm group or PCoA axis, and 1|cage indicates that we used the cage as a random effect. We then ran an analysis of variance (ANOVA) on the above model to generate *P* values for biofilm group or PCoA axis. We checked for possible cage effect by comparing the above model and a model with the cage removed (genus ∼ variable + ε) using ANOVA. The *P* values were adjusted for multiple hypothesis testing in R using the p.adjust function employing the method of Benjamini and Hochberg (76). The heatmaps were generated using the R function ggplot2 (77).

We performed two additional analyses on the 16S rRNA data, the first utilizing QIIME (70) v. 1.9.1 closed-reference at 97% similarity level using the Greengenes reference data set release 13_8 and the second employing Deblur (78) workflow v. 1.0.3 with the default parameters (using Deblur's default positive and negative reference filtering) and trim length set to 100 bases. Both pipelines showed no significant separation between the BF-bx and BF-bx DC samples (see Fig. S4C and D in the supplemental material).

RNA extraction, rRNA depletion, and RNA sequencing. Total RNA was extracted from frozen proximal colon tissue snips (Fig. 1AII) using the mirVana miRNA isolation kit, with phenol (ThermoFisher Scientific catalog no. AM1560) according to the manufacturer's instructions, with the addition of an ∼1:1 mix of 1-mm acid-washed glass beads and 0.1-mm zirconia beads and a Precellys24 (Bertin Instruments catalog no. EQ03119-200-RD000.0) bead beater for tissue disruption and lysis. Extracted RNA was treated with the Turbo DNA-free kit (ThermoFisher Scientific catalog no. AM1907) to remove DNA. Quality control, rRNA depletion, and cDNA library preparation were performed by the University of Florida's Interdisciplinary Center for Biotechnology Research (ICBR) Gene Expression and Genotyping core using the Agilent 2100 bioanalyzer (Agilent Genomics catalog no. G2939BA), Ribo-Zero Gold rRNA removal kit (Epidemiology) (Illumina catalog no. MRZE724) and ScriptSeq v2 RNASeq library preparation kit (Illumina catalog no. SSV21124) starting with 1 μg total RNA. Samples were sequenced by the University of Florida ICBR NextGen DNA Sequencing core on the Illumina HiSeq 3000 (2 × 100 run), multiplexing each sample into three lanes to avoid lane effect.

Mouse RNA-seq analysis. Reads were quality filtered at Q20 and trimmed to remove remaining adapters using Trimmomatic (79) version 0.36. The resulting reads were aligned to Illumina iGenome *Mus musculus* Ensembl GRCm38 reference genome using Tophat (80) version 2.1.1 utilizing Bowtie2 (81)

TABLE 2 Ratios of DE genes in BF+T and BF-bx groups

Data set	% of genes		
	DE genes (FDR < 0.05) (% of input transcripts)	Upregulated genes in BF+T group (% of DE transcripts)	Downregulated genes in BF+T group (% of DE transcripts)
Complete	36	93	7
Rarefied[a]	40	88	12

[a]The rarefied data set was based on 140,000 reads per sample.

version 2.3.0 following the approach of Gilad and Mizrahi-Man (82). The resulting alignments (averaging 34,079,158 concordant read pairs per sample) were processed using Cufflinks (83) version 2.2.1 along with Illumina iGenome *Mus musculus* Ensembl GRCm38 gene transfer format file, after masking rRNA features (82). We used cuffquant to perform transcript quantification and exported the raw counts (nonnormalized counts) to text files. The raw counts were then imported to edgeR (84) version 3.16.5 for detecting differentially expressed (DE) genes. A gene was considered for the differential expression test if it was present in at least 50% of the samples. We considered a gene DE if its edgeR false-discovery rate (FDR)-adjusted *P* value (P_{FDR}) was <0.05. Parallel analysis using featureCounts from the subread package version 1.5.3 for transcript quantification showed similar results (data not shown) (85). Pathway analysis was conducted through GAGE (86) version 2.24 using *Mus musculus* (mmu) Kyoto Encyclopedia of Genes and Genomes (KEGG) (87) pathways, and genes were mapped to KEGG pathways using Pathview (88). We considered a pathway significant if its GAGE false-discovery rate (*q* value) was less than 0.05. We tested the effect of sequencing lane on the clustering of the samples (Fig. 2A) and found it to be insignificant (*P* value > 0.05) (data not shown).

Metatranscriptome analysis. Quality-filtered and trimmed reads from above were aligned to iGenome *Mus musculus* Ensembl GRCm38 reference genome using BWA (89) version 0.7.16a, and reads with alignments were excluded from further analysis. The remaining reads were then filtered from rRNA and tRNA sequences by aligning (using BWA) to a collection of NCBI rRNA and tRNA sequences and SILVA database sequences, resulting in an average of 1,208,429 reads per sample, which were then submitted for *de novo* assembly using Trinity (90) version 2.4.0. The resulting assembly was annotated using Trinotate (91) version 3.0.1 (http://trinotate.github.io) with the following databases: uniprot_sprot (92), Pfam (93), and Virulence Factor Database (VFDB) (94). The resulting annotations were examined, and sequences annotated as nonbacterial were removed. Transcript abundance was determined using RNA-seq by expectation maximization (RSEM) (95) through Trinity's align_and_estimate_abundance.pl script, and the counts were imported to edgeR version 3.16.5 for differential expression analysis. A gene was considered for the differential expression test if it was present in at least 50% of the samples. We considered a transcript DE if its edgeR FDR-adjusted *P* value was <0.05. To account for normalization artifacts, we also examined the ratios of DE genes between the BF+T and BF-bx groups generated from a rarefied data set that was based on 140,000 reads per sample (Table 2). The similar ratios of DE genes from the complete and rarefied data sets suggest that our findings are not an artifact of normalization. We conducted a second analysis (reference-based analysis) by aligning the reads submitted for *de novo* assembly to the human gut microbiome integrated gene catalog (IGC) (96) using Bowtie2 (81) (v.2.3.4.2) followed by quantification using featureCounts (85) from the subread package (v.1.5.3) and obtained similar results (Fig. S2A).

Pathway analysis was conducted through GAGE version 2.24 using Kyoto Encyclopedia of Genes and Genomes (KEGG) reference pathways on the assembled transcript and The Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN) (97) on the unassembled reads. Genes were mapped to KEGG pathways using Pathview (88). We considered a pathway significant if its *q* value was <0.05. Weighted gene coexpression network analysis (WGCNA) version 1.68 (23) was utilized to detect modules in each biofilm status samples using the blockwiseConsensusModules function which performs the network construction and consensus module detection. The hub gene in each detected module was identified using the WGCNA function chooseTopHubInEachModule. The sequencing lane had no effect on the clustering of the samples in Fig. 1A (*P* value > 0.05, data not shown).

miRNA extraction and sequencing. Small RNAs were extracted from snap-frozen stool samples (Fig. 1AII) using the mirVana miRNA isolation kit. Because of the low amount of small RNA, GF stools were pooled from 13 $Apc^{Min\Delta 850/+}$;$Il10^{-/-}$ mice (20- to 44-week age range) total, or stools from 2 to 5 mice per sample ($n = 4$). BF-bx and BF+T small RNAs were extracted from the stools of 12-week-associated BF-bx ($n = 7$) and BF+T ($n = 10$) $Apc^{Min\Delta 850/+}$;$Il10^{-/-}$ mice. cDNA libraries were synthesized with the NEBNext Multiplex Small RNA Library Prep Set for Illumina kit (New England Biolabs catalog no. E7300) and small RNAs for each library (21- to 30-nucleotide size range) were gel purified. For the GF, BF-bx, and BF+T comparisons, a pool of 21 libraries (equivalent molar concentrations; 4 GF, 7 BF-bx, and 10 BF+T) were multiplexed and sequenced using the Illumina Miseq.

miRNA analysis. CAP-miRSeq (98) was used to process the miRNA sequences. We used the databases and reference sequences that ship with CAP-miRSeq for all the analyses. Briefly, sequences were filtered and trimmed using cutadapt (99). Quantification of miRNA was done using miRDeep2 (100), and DE miRNAs were detected using edgeR version 3.16.5. We considered a miRNA DE if its edgeR FDR-adjusted *P* value was <0.05. Principal-component analysis (PCA) was created using R's prcomp function from the normalized and $\log_{10}$-transformed miRNA counts according to the equation above.

Correlations with microbiota taxon abundance were done using R lm function, *P* values were determined using R's ANOVA function, and FDR correction was done using R's p.adjust function

employing the method of Benjamini and Hochberg (76) and only those with a FDR-adjusted *P* value of <0.05 were considered.

Mouse miRNA targets were predicted using miRDB (101), and bacterial target prediction for mice miRNA was done using PITA (102) on the assembled bacterial transcripts from the RNA-seq data described above. We considered a bacterial transcript a potential target for a particular mouse miRNA if its $\Delta\Delta G$ score was less than or equal to -15 kcal/mol. The bacterial and mouse gene targets of miRNAs significantly different between GF, BF-bx, and BF+T groups are listed in Tables S12 at https://figshare.com/s/33fb65da16763a5e3347 and S13 at https://figshare.com/s/be3f38607b2e9875e7f6.

For miRNA expression correlation with tumor numbers, two BF-bx samples were excluded because they were fixed without splaying the colon so tumor counts could not be generated. Correlation was done using Spearman's rank correlation through R's cor.test function.

Statistical analyses. For all sequencing analyses, a taxon or miRNA was considered only if it was present in at least 30% of the comparison samples, and statistics are described in above 16S rRNA sequencing, mouse RNA-seq, metatranscriptome and miRNA analysis sections. *P* values of <0.05 were considered statistically significant.

Ethics. All animal experiments were approved by the University of Florida Institutional Animal Care and Use Committee (protocol 201308038). All patient tissues were collected as previously described (11, 12, 103).

Data availability. The data supporting the results of this article have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) under accession number GSE108165 for the RNA-seq and miRNA data sets and the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) under BioProject identifier (ID) PRJNA422588 for 16S rRNA Run02 sequences. Run01 sequences are in the article by Tomkovich et al. (13).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 1.6 MB.

FIG S2, PDF file, 0.4 MB.

FIG S3, PDF file, 0.2 MB.

FIG S4, PDF file, 0.2 MB.

FIG S5, PDF file, 0.2 MB.

FIG S6, PDF file, 0.2 MB.

ACKNOWLEDGMENTS

We are grateful to the germfree services division of University of Florida's Animal Care Services for assistance with gnotobiotic experiments. We also thank the University of Florida Interdisciplinary Center for Biotechnology Research Gene Expression and Genotyping and NextGen DNA sequencing Cores for helping with the RNA sequencing.

We declare that we have no competing interests.

This research was supported by National Institutes of Health grants R01DK073338, R01DK47700, R21 CA195226 (all to University of Florida), R01CA151393, R01CA196845, P50 CA62924, P30 CA006973, P30 DK089502, and K99 CA230192 (all to Johns Hopkins University School of Medicine), University of Florida Health Cancer Center (UFHCC) Funds, and University of Florida Department of Medicine Gatorade Fund. Additional support was provided by the Bloomberg Philanthropies, the Johns Hopkins Discovery Award, the Commonwealth Foundation, and the Cancer Research Institute/Fight Colorectal Cancer Coalition (Johns Hopkins University School of Medicine). R.Z.G. is supported by UFHCC funds. Y.Y. is supported by the Crohn's & Colitis Foundation of America (CCFA) research fellowship award (CCFA reference number 409472).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions are as follows: study concept and design (S.T., C.L.S., and C.J.); acquisition of data (S.T., C.M.D., J.L.P., J.J., J.G., R.C.N., and Y.Y.); analysis and interpretation of data (S.T., R.Z.G., C.M.D., K.W., A.A.F., T.D.S., C.L.S., and C.J.); drafting of manuscript (S.T., R.Z.G., and C.J.); statistical analysis (S.T. and R.Z.G.); obtained funding (C.J. and C.L.S.); study supervision (C.L.S. and C.J.). All authors contributed to the revision of the manuscript.

REFERENCES

1. Borges-Canha M, Portela-Cidade JP, Dinis-Ribeiro M, Leite-Moreira AF, Pimentel-Nunes P. 2015. Role of colonic microbiota in colorectal carcinogenesis: a systematic review. Rev Esp Enferm Dig 107:659–671. https://doi.org/10.17235/reed.2015.3830/2015.

2. Wang J, Jia H. 2016. Metagenome-wide association studies: fine-mining the microbiome. Nat Rev Microbiol 14:508–522. https://doi.org/10.1038/nrmicro.2016.83.
3. Pope JL, Tomkovich S, Yang Y, Jobin C. 2017. Microbiota as a mediator of cancer progression and therapy. Transl Res 179:139–154. https://doi.org/10.1016/j.trsl.2016.07.021.
4. Chen J, Domingue JC, Sears CL. 2017. Microbiota dysbiosis in select human cancers: evidence of association and causality. Semin Immunol 32:25–34. https://doi.org/10.1016/j.smim.2017.08.001.
5. Yu T, Guo F, Yu Y, Sun T, Ma D, Han J, Qian Y, Kryczek I, Sun D, Nagarsheth N, Chen Y, Chen H, Hong J, Zou W, Fang J-Y. 2017. Fusobacterium nucleatum promotes chemoresistance to colorectal cancer by modulating autophagy. Cell 170:548–563.e16. https://doi.org/10.1016/j.cell.2017.07.008.
6. Amitay EL, Werner S, Vital M, Pieper DH, Höfler D, Gierse I-J, Butt J, Balavarca Y, Cuk K, Brenner H. 2017. Fusobacterium and colorectal cancer: causal factor or passenger? Results from a large colorectal cancer screening study. Carcinogenesis 38:781–788. https://doi.org/10.1093/carcin/bgx053.
7. Tomkovich S, Yang Y, Winglee K, Gauthier J, Mühlbauer M, Sun X, Mohamadzadeh M, Liu X, Martin P, Wang GP, Oswald E, Fodor AA, Jobin C. 2017. Locoregional effects of microbiota in a preclinical model of colon carcinogenesis. Cancer Res 77:2620–2632. https://doi.org/10.1158/0008-5472.CAN-16-3472.
8. Bullman S, Pedamallu CS, Sicinska E, Clancy TE, Zhang X, Cai D, Neuberg D, Huang K, Guevara F, Nelson T, Chipashvili O, Hagan T, Walker M, Ramachandran A, Diosdado B, Serna G, Mulet N, Landolfi S, Ramon y Cajal S, Fasani R, Aguirre AJ, Ng K, Élez E, Ogino S, Tabernero J, Fuchs CS, Hahn WC, Nuciforo P, Meyerson M. 2017. Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer. Science 358:1443–1448. https://doi.org/10.1126/science.aal5240.
9. Baxter NT, Zackular JP, Chen GY, Schloss PD. 2014. Structure of the gut microbiome following colonization with human feces determines colonic tumor burden. Microbiome 2:20. https://doi.org/10.1186/2049-2618-2-20.
10. Wong SH, Zhao L, Zhang X, Nakatsu G, Han J, Xu W, Xiao X, Kwong TNY, Tsoi H, Wu WKK, Zeng B, Chan FKL, Sung JJY, Wei H, Yu J. 2017. Gavage of fecal samples from patients with colorectal cancer promotes intestinal carcinogenesis in germ-free and conventional mice. Gastroenterology 153:1621–1633.e6. https://doi.org/10.1053/j.gastro.2017.08.022.
11. Drewes JL, White JR, Dejea CM, Fathi P, Iyadorai T, Vadivelu J, Roslani AC, Wick EC, Mongodin EF, Loke MF, Thulasi K, Gan HM, Goh KL, Chong HY, Kumar S, Wanyiri JW, Sears CL. 2017. High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. NPJ Biofilms Microbiomes 3:34. https://doi.org/10.1038/s41522-017-0040-3.
12. Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, Rossetti BJ, Peterson SN, Snesrud EC, Borisy GG, Lazarev M, Stein E, Vadivelu J, Roslani AC, Malik AA, Wanyiri JW, Goh KL, Thevambiga I, Fu K, Wan F, Llosa N, Housseau F, Romans K, Wu X, McAllister FM, Wu S, Vogelstein B, Kinzler KW, Pardoll DM, Sears CL. 2014. Microbiota organization is a distinct feature of proximal colorectal cancers. Proc Natl Acad Sci U S A 111:18321–18326. https://doi.org/10.1073/pnas.1406199111.
13. Tomkovich S, Dejea CM, Winglee K, Drewes JL, Chung L, Housseau F, Pope JL, Gauthier J, Sun X, Mühlbauer M, Liu X, Fathi P, Anders RA, Besharati S, Perez-Chanona E, Yang Y, Ding H, Wu X, Wu S, White JR, Gharaibeh RZ, Fodor AA, Wang H, Pardoll DM, Jobin C, Sears CL. 2019. Human colon mucosal biofilms from healthy or colon cancer hosts are carcinogenic. J Clin Invest 130:1699–1712. https://doi.org/10.1172/JCI124196.
14. Carethers JM, Jung BH. 2015. Genetics and genetic biomarkers in sporadic colorectal cancer. Gastroenterology 149:1177–1190.e3. https://doi.org/10.1053/j.gastro.2015.06.047.
15. Nakatsu G, Li X, Zhou H, Sheng J, Wong SH, Wu WKK, Ng SC, Tsoi H, Dong Y, Zhang N, He Y, Kang Q, Cao L, Wang K, Zhang J, Liang Q, Yu J, Sung J. 2015. Gut mucosal microbiome across stages of colorectal carcinogenesis. Nat Commun 6:8727. https://doi.org/10.1038/ncomms9727.
16. Feng Q, Liang S, Jia H, Stadlmayr A, Tang L, Lan Z, Zhang D, Xia H, Xu X, Jie Z, Su L, Li X, Li X, Li J, Xiao L, Huber-Schönauer U, Niederseer D, Xu X, Al-Aama JY, Yang H, Wang J, Kristiansen K, Arumugam M, Tilg H, Datz C, Wang J. 2015. Gut microbiome development along the colorectal adenoma-carcinoma sequence. Nat Commun 6:6528. https://doi.org/10.1038/ncomms7528.
17. Strubberg AM, Madison BB. 2017. MicroRNAs in the etiology of colorectal cancer: pathways and clinical implications. Dis Model Mech 10:197–214. https://doi.org/10.1242/dmm.027441.
18. Liu S, da Cunha AP, Rezende RM, Cialic R, Wei Z, Bry L, Comstock LE, Gandhi R, Weiner HL. 2016. The host shapes the gut microbiota via fecal microRNA. Cell Host Microbe 19:32–43. https://doi.org/10.1016/j.chom.2015.12.005.
19. Yuan C, Burns MB, Subramanian S, Blekhman R. 2018. Interaction between host microRNAs and the gut microbiota in colorectal cancer. mSystems 3:e00205-17. [Crossref] https://doi.org/10.1128/mSystems.00205-17.
20. Peck BCE, Mah AT, Pitman WA, Ding S, Lund PK, Sethupathy P. 2017. Functional transcriptomics in diverse intestinal epithelial cell types reveals robust microRNA sensitivity in intestinal stem cells to microbial status. J Biol Chem 292:2586–2600. https://doi.org/10.1074/jbc.M116.770099.
21. Seth P, Hsieh PN, Jamal S, Wang L, Gygi SP, Jain MK, Coller J, Stamler JS. 2019. Regulation of microRNA machinery and development by interspecies S-nitrosylation. Cell 176:1014–1025.e12. https://doi.org/10.1016/j.cell.2019.01.037.
22. Yang Y, Weng W, Peng J, Hong L, Yang L, Toiyama Y, Gao R, Liu M, Yin M, Pan C, Li H, Guo B, Zhu Q, Wei Q, Moyer M-P, Wang P, Cai S, Goel A, Qin H, Ma Y. 2017. Fusobacterium nucleatum increases proliferation of colorectal cancer cells and tumor development in mice by activating Toll-like receptor 4 signaling to nuclear factor-κB, and up-regulating expression of MicroRNA-21. Gastroenterology 152:851–866.e24. https://doi.org/10.1053/j.gastro.2016.11.018.
23. Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559. https://doi.org/10.1186/1471-2105-9-559.
24. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM. 2004. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 6:1–6. https://doi.org/10.1016/s1476-5586(04)80047-2.
25. Shah MS, DeSantis TZ, Weinmaier T, McMurdie PJ, Cope JL, Altrichter A, Yamal J-M, Hollister EB. 2018. Leveraging sequence-based faecal microbial community survey data to identify a composite biomarker for colorectal cancer. Gut 67:882–891. https://doi.org/10.1136/gutjnl-2016-313189.
26. Burns MB, Lynch J, Starr TK, Knights D, Blekhman R. 2015. Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. Genome Med 7:55. https://doi.org/10.1186/s13073-015-0177-8.
27. Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, Beghini F, Manara S, Karcher N, Pozzi C, Gandini S, Serrano D, Tarallo S, Francavilla A, Gallo G, Trompetto M, Ferrero G, Mizutani S, Shiroma H, Shiba S, Shibata T, Yachida S, Yamada T, Wirbel J, Schrotz-King P, Ulrich CM, Brenner H, Arumugam M, Bork P, Zeller G, Cordero F, Dias-Neto E, Setubal JC, Tett A, Pardini B, Rescigno M, Waldron L, Naccarati A, Segata N. 2019. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nat Med 25:667–678. https://doi.org/10.1038/s41591-019-0405-7.
28. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, Fleck JS, Voigt AY, Palleja A, Ponnudurai R, Sunagawa S, Coelho LP, Schrotz-King P, Vogtmann E, Habermann N, Niméus E, Thomas AM, Manghi P, Gandini S, Serrano D, Mizutani S, Shiroma H, Shiba S, Shibata T, Yachida S, Yamada T, Waldron L, Naccarati A, Segata N, Sinha R, Ulrich CM, Brenner H, Arumugam M, Bork P, Zeller G. 2019. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nat Med 25:679–689. https://doi.org/10.1038/s41591-019-0406-6.
29. Duran-Pinedo AE, Chen T, Teles R, Starr JR, Wang X, Krishnan K, Frias-Lopez J. 2014. Community-wide transcriptome of the oral microbiome in subjects with and without periodontitis. ISME J 8:1659–1672. https://doi.org/10.1038/ismej.2014.23.
30. Flynn KJ, Baxter NT, Schloss PD. 2016. Metabolic and community synergy of oral bacteria in colorectal cancer. mSphere 1:e00102-16. [Crossref] https://doi.org/10.1128/mSphere.00102-16.
31. Cornforth DM, Foster KR. 2013. Competition sensing: the social side of bacterial stress responses. Nat Rev Microbiol 11:285–293. https://doi.org/10.1038/nrmicro2977.
32. Parrow NL, Fleming RE, Minnick MF. 2013. Sequestration and scaveng-

ing of iron in infection. Infect Immun 81:3503–3514. https://doi.org/10.1128/IAI.00602-13.
33. Ribet D, Cossart P. 2015. How bacterial pathogens colonize their hosts and invade deeper tissues. Microbes Infect 17:173–183. https://doi.org/10.1016/j.micinf.2015.01.004.
34. Dannewitz B, Edrich C, Tomakidi P, Kohl A, Gabbert O, Eickholz P, Steinberg T. 2006. Elevated gene expression of MMP-1, MMP-10, and TIMP-1 reveal changes of molecules involved in turn-over of extracellular matrix in cyclosporine-induced gingival overgrowth. Cell Tissue Res 325:513–522. https://doi.org/10.1007/s00441-006-0200-x.
35. Kainulainen V, Korhonen TK. 2014. Dancing to another tune—adhesive moonlighting proteins in bacteria. Biology (Basel) 3:178–204. https://doi.org/10.3390/biology3010178.
36. Abed J, Emgård JEM, Zamir G, Faroja M, Almogy G, Grenov A, Sol A, Naor R, Pikarsky E, Atlan KA, Mellul A, Chaushu S, Manson AL, Earl AM, Ou N, Brennan CA, Garrett WS, Bachrach G. 2016. Fap2 mediates Fusobacterium nucleatum colorectal adenocarcinoma enrichment by binding to tumor-expressed Gal-GalNAc. Cell Host Microbe 20:215–225. https://doi.org/10.1016/j.chom.2016.07.006.
37. Kumar R, Herold JL, Schady D, Davis J, Kopetz S, Martinez-Moczygemba M, Murray BE, Han F, Li Y, Callaway E, Chapkin RS, Dashwood W-M, Dashwood RH, Berry T, Mackenzie C, Xu Y. 2017. Streptococcus gallolyticus subsp. gallolyticus promotes colorectal tumor development. PLoS Pathog 13:e1006440. https://doi.org/10.1371/journal.ppat.1006440.
38. Boleij A, Tjalsma H. 2012. Gut bacteria in health and disease: a survey on the interface between intestinal microbiology and colorectal cancer. Biol Rev Camb Philos Soc 87:701–730. https://doi.org/10.1111/j.1469-185X.2012.00218.x.
39. Pinho SS, Reis CA. 2015. Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer 15:540–555. https://doi.org/10.1038/nrc3982.
40. Nougayrède J-P, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, Buchrieser C, Hacker J, Dobrindt U, Oswald E. 2006. Escherichia coli induces DNA double-strand breaks in eukaryotic cells. Science 313:848–851. https://doi.org/10.1126/science.1127059.
41. Sicard J-F, Le Bihan G, Vogeleer P, Jacques M, Harel J. 2017. Interactions of intestinal bacteria with components of the intestinal mucus. Front Cell Infect Microbiol 7:387. https://doi.org/10.3389/fcimb.2017.00387.
42. Bäumler AJ, Sperandio V. 2016. Interactions between the microbiota and pathogenic bacteria in the gut. Nature 535:85–93. https://doi.org/10.1038/nature18849.
43. Martens EC, Chiang HC, Gordon JI. 2008. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe 4:447–457. https://doi.org/10.1016/j.chom.2008.09.007.
44. Porcheron G, Schouler C, Dozois CM. 2016. Survival games at the dinner table: regulation of enterobacterial virulence through nutrient sensing and acquisition. Curr Opin Microbiol 30:98–106. https://doi.org/10.1016/j.mib.2016.01.008.
45. Tronnet S, Garcie C, Rehm N, Dobrindt U, Oswald E, Martin P. 2016. Iron homeostasis regulates the genotoxicity of Escherichia coli that produces colibactin. Infect Immun 84:3358–3368. https://doi.org/10.1128/IAI.00659-16.
46. Garcie C, Tronnet S, Garénaux A, McCarthy AJ, Brachmann AO, Pénary M, Houle S, Nougayrède J-P, Piel J, Taylor PW, Dozois CM, Genevaux P, Oswald E, Martin P. 2016. The bacterial stress-responsive Hsp90 chaperone (HtpG) is required for the production of the genotoxin colibactin and the siderophore yersiniabactin in Escherichia coli. J Infect Dis 214:916–924. https://doi.org/10.1093/infdis/jiw294.
47. Arthur JC, Gharaibeh RZ, Mühlbauer M, Perez-Chanona E, Uronis JM, McCafferty J, Fodor AA, Jobin C. 2014. Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer. Nat Commun 5:4724. https://doi.org/10.1038/ncomms5724.
48. Li H, Limenitakis JP, Fuhrer T, Geuking MB, Lawson MA, Wyss M, Brugiroux S, Keller I, Macpherson JA, Rupp S, Stolp B, Stein JV, Stecher B, Sauer U, McCoy KD, Macpherson AJ. 2015. The outer mucus layer hosts a distinct intestinal microbial niche. Nat Commun 6:8292. https://doi.org/10.1038/ncomms9292.
49. Hecht AL, Casterline BW, Choi VM, Bubeck Wardenburg J. 2017. A two-component system regulates Bacteroides fragilis toxin to maintain intestinal homeostasis and prevent lethal disease. Cell Host Microbe 22:443–448.e5. https://doi.org/10.1016/j.chom.2017.08.007.
50. DeStefano Shields CE, Van Meerbeke SW, Housseau F, Wang H, Huso DL, Casero RA, O'Hagan HM, Sears CL. 2016. Reduction of murine colon tumorigenesis driven by enterotoxigenic Bacteroides fragilis using cefoxitin treatment. J Infect Dis 214:122–129. https://doi.org/10.1093/infdis/jiw069.
51. Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JI. 2008. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe 3:213–223. https://doi.org/10.1016/j.chom.2008.02.015.
52. Woting A, Pfeiffer N, Loh G, Klaus S, Blaut M. 2014. Clostridium ramosum promotes high-fat diet-induced obesity in gnotobiotic mouse models. mBio 5:e01530-14. https://doi.org/10.1128/mBio.01530-14.
53. O'Keefe S. 2016. Diet, microorganisms and their metabolites, and colon cancer. Nat Rev Gastroenterol Hepatol 13:691–706. https://doi.org/10.1038/nrgastro.2016.165.
54. Dubois V, Eeckhoute J, Lefebvre P, Staels B. 2017. Distinct but complementary contributions of PPAR isotypes to energy homeostasis. J Clin Invest 127:1202–1214. https://doi.org/10.1172/JCI88894.
55. Park J-I, Kwak J-Y. 2012. The role of peroxisome proliferator-activated receptors in colorectal cancer. PPAR Res 2012:876418. https://doi.org/10.1155/2012/876418.
56. Fabbiano S, Suárez-Zamorano N, Trajkovski M. 2017. Host-microbiota mutualism in metabolic diseases. Front Endocrinol (Lausanne) 8:267. https://doi.org/10.3389/fendo.2017.00267.
57. DeBerardinis RJ, Chandel NS. 2016. Fundamentals of cancer metabolism. Sci Adv 2:e1600200. https://doi.org/10.1126/sciadv.1600200.
58. Valle J, Da Re S, Schmid S, Skurnik D, D'Ari R, Ghigo J-M. 2008. The amino acid valine is secreted in continuous-flow bacterial biofilms. J Bacteriol 190:264–274. https://doi.org/10.1128/JB.01405-07.
59. Nakamura Y, Yamamoto N, Kino Y, Yamamoto N, Kamei S, Mori H, Kurokawa K, Nakashima N. 2016. Establishment of a multi-species biofilm model and metatranscriptomic analysis of biofilm and planktonic cell communities. Appl Microbiol Biotechnol 100:7263–7279. https://doi.org/10.1007/s00253-016-7532-6.
60. Johnson CH, Dejea CM, Edler D, Hoang LT, Santidrian AF, Felding BH, Ivanisevic J, Cho K, Wick EC, Hechenbleikner EM, Uritboonthai W, Goetz L, Casero RA, Pardoll DM, White JR, Patti GJ, Sears CL, Siuzdak G. 2015. Metabolism links bacterial biofilms and colon carcinogenesis. Cell Metab 21:891–897. https://doi.org/10.1016/j.cmet.2015.04.011.
61. Burrell M, Hanfrey CC, Murray EJ, Stanley-Wall NR, Michael AJ. 2010. Evolution and multiplicity of arginine decarboxylases in polyamine biosynthesis and essential role in Bacillus subtilis biofilm formation. J Biol Chem 285:39224–39238. https://doi.org/10.1074/jbc.M110.163154.
62. Pavlova NN, Thompson CB. 2016. The emerging hallmarks of cancer metabolism. Cell Metab 23:27–47. https://doi.org/10.1016/j.cmet.2015.12.006.
63. Abdulhussein AA, Wallace HM. 2014. Polyamines and membrane transporters. Amino Acids 46:655–660. https://doi.org/10.1007/s00726-013-1553-6.
64. Yin Y, Song M, Gu B, Qi X, Hu Y, Feng Y, Liu H, Zhou L, Bian Z, Zhang J, Zuo X, Huang Z. 2016. Systematic analysis of key miRNAs and related signaling pathways in colorectal tumorigenesis. Gene 578:177–184. https://doi.org/10.1016/j.gene.2015.12.015.
65. Liu S, Xiao Z, Ai F, Liu F, Chen X, Cao K, Ren W, Zhang X, Shu P, Zhang D. 2017. miR-142-5p promotes development of colorectal cancer through targeting SDHB and facilitating generation of aerobic glycolysis. Biomed Pharmacother 92:1119–1127. https://doi.org/10.1016/j.biopha.2017.05.134.
66. Izaurralde E. 2015. Breakers and blockers—miRNAs at work. Science 349:380–382. https://doi.org/10.1126/science.1260969.
67. Lee J, Park EJ, Yuki Y, Ahmad S, Mizuguchi K, Ishii KJ, Shimaoka M, Kiyono H. 2015. Profiles of microRNA networks in intestinal epithelial cells in a mouse model of colitis. Sci Rep 5:18174. https://doi.org/10.1038/srep18174.
68. Laukens D, Brinkman BM, Raes J, De Vos M, Vandenabeele P. 2016. Heterogeneity of the gut microbiome in mice: guidelines for optimizing experimental design. FEMS Microbiol Rev 40:117–132. https://doi.org/10.1093/femsre/fuv036.
69. Teng Y, Ren Y, Sayed M, Hu X, Lei C, Kumar A, Hutchins E, Mu J, Deng Z, Luo C, Sundaram K, Sriwastva MK, Zhang L, Hsieh M, Reiman R, Haribabu B, Yan J, Jala VR, Miller DM, Van Keuren-Jensen K, Merchant ML, McClain CJ, Park JW, Egilmez NK, Zhang H-G. 2018. Plant-derived exosomal microRNAs shape the gut microbiota. Cell Host Microbe 24:637–652.e8. https://doi.org/10.1016/j.chom.2018.10.001.
70. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA,

Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303.

71. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267. https://doi.org/10.1128/AEM.00062-07.
72. McCafferty J, Mühlbauer M, Gharaibeh RZ, Arthur JC, Perez-Chanona E, Sha W, Jobin C, Fodor AA. 2013. Stochastic changes over time and not founder effects drive cage effects in microbial community assembly in a mouse model. ISME J 7:2116–2125. https://doi.org/10.1038/ismej.2013.106.
73. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217. https://doi.org/10.1371/journal.pone.0061217.
74. R Core Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
75. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2016. nlme: linear and nonlinear mixed effects models.
76. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc Ser B (Methodological) 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
77. Wickham H. 2009. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, NY.
78. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, Kightley EP, Thompson LR, Hyde ER, Gonzalez A, Knight R, Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, Kightley EP, Thompson LR, Hyde ER, Gonzalez A, Knight R. 2017. Deblur rapidly resolves single-nucleotide community sequence patterns. mSystems 2:e00191-16. [Crossref] https://doi.org/10.1128/mSystems.00191-16.
79. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
80. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14:R36. https://doi.org/10.1186/gb-2013-14-4-r36.
81. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
82. Gilad Y, Mizrahi-Man O. 2015. A reanalysis of mouse ENCODE comparative gene expression data. F1000Res 4:121. https://doi.org/10.12688/f1000research.6536.1.
83. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562–578. https://doi.org/10.1038/nprot.2012.016.
84. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. https://doi.org/10.1093/bioinformatics/btp616.
85. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930. https://doi.org/10.1093/bioinformatics/btt656.
86. Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. 2009. GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinformatics 10:161. https://doi.org/10.1186/1471-2105-10-161.
87. Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27.
88. Luo W, Brouwer C. 2013. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 29:1830–1831. https://doi.org/10.1093/bioinformatics/btt285.
89. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595. https://doi.org/10.1093/bioinformatics/btp698.
90. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652. https://doi.org/10.1038/nbt.1883.
91. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494–1512. https://doi.org/10.1038/nprot.2013.084.
92. The UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res 45:D158–D169. https://doi.org/10.1093/nar/gkw1099.
93. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285. https://doi.org/10.1093/nar/gkv1344.
94. Chen L, Zheng D, Liu B, Yang J, Jin Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis–10 years on. Nucleic Acids Res 44:D694–D697. https://doi.org/10.1093/nar/gkv1239.
95. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323. https://doi.org/10.1186/1471-2105-12-323.
96. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, Arumugam M, Kultima JR, Prifti E, Nielsen T, Juncker AS, Manichanh C, Chen B, Zhang W, Levenez F, Wang J, Xu X, Xiao L, Liang S, Zhang D, Zhang Z, Chen W, Zhao H, Al-Aama JY, Edris S, Yang H, Wang J, Hansen T, Nielsen HB, Brunak S, Kristiansen K, Guarner F, Pedersen O, Doré J, Ehrlich SD, MetaHIT Consortium, Bork P, Wang J. 2014. An integrated catalog of reference genes in the human gut microbiome. Nat Biotechnol 32:834–841. https://doi.org/10.1038/nbt.2942.
97. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, White O, Kelley ST, Methé B, Schloss PD, Gevers D, Mitreva M, Huttenhower C. 2012. Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput Biol 8:e1002358. https://doi.org/10.1371/journal.pcbi.1002358.
98. Sun Z, Evans J, Bhagwate A, Middha S, Bockol M, Yan H, Kocher J-P. 2014. CAP-miRSeq: a comprehensive analysis pipeline for microRNA sequencing data. BMC Genomics 15:423. https://doi.org/10.1186/1471-2164-15-423.
99. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10. https://doi.org/10.14806/ej.17.1.200.
100. Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. 2012. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res 40:37–52. https://doi.org/10.1093/nar/gkr688.
101. Wang X. 2016. Improving microRNA target prediction by modeling with unambiguously identified microRNA-target pairs from CLIP-ligation studies. Bioinformatics 32:1316–1322. https://doi.org/10.1093/bioinformatics/btw002.
102. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. 2007. The role of site accessibility in microRNA target recognition. Nat Genet 39:1278–1284. https://doi.org/10.1038/ng2135.
103. Dejea CM, Fathi P, Craig JM, Boleij A, Taddese R, Geis AL, Wu X, DeStefano Shields CE, Hechenbleikner EM, Huso DL, Anders RA, Giardiello FM, Wick EC, Wang H, Wu S, Pardoll DM, Housseau F, Sears CL. 2018. Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. Science 359:592–597. https://doi.org/10.1126/science.aah3648.