Manipulation of the Gut Microbiota Reveals Role of Gut Microbiota in Colon Tumorigenesis

Joseph P. Zackular[1], Nielson T. Baxter[1], Grace Y. Chen[2*], and Patrick D. Schloss[1*]

* To whom correspondence should be addressed:

pschloss@umich.edu

gchenry@umich.edu

1 Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI

2 Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan, Ann Arbor, MI

3 Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, TN

Abstract

There is growing evidence that individuals with colonic adenomas and carcinomas harbor a distinct microbiota. Alterations to the gut microbiota may allow the outgrowth of bacterial populations that induce genomic mutations or exacerbate tumor-promoting inflammation. In addition, it is likely that the loss of key bacterial populations may result in the loss of protective functions that are normally provided by the microbiota. We explored the role of the gut microbiota in colon tumorigenesis using an inflammation-based murine model. We observed that perturbing the microbiota with different combinations of antibiotics did not change the bacterial load but reduced the number of tumors at the end of the model. Using the random forest machine learning algorithm we successfully modeled the number of tumors that developed over the course of the model based on the composition of the microbiota at the beginning. The timing of antibiotic treatment was an important determinant of tumor outcome as colon tumorigenesis was arrested with the use of antibiotics during the inflammation period of the murine model. Together, these results indicate that it is possible to predict colon tumorigenesis based on the composition of the microbiota and that altering the gut microbiota can alter the course of tumorigenesis.

Keywords: azoxymethane, dextran sodium sulfate, 16S rRNA gene sequencing, microbial ecology, microbiome

Introduction

The mammalian gastrointestinal tract is home to a complex and dynamic community of microorganisms, termed the gut microbiota, which is essential for maintaining host health [1]. There are complex interactions among bacterial populations in the gut that have an important effect on host health [2-4]. The number of diseases that are associated with abnormalities in the gut microbiota highlights the importance of these ecological interactions [5-7]. Over the last several years, it has been well documented that perturbations to this community are associated with colorectal cancer (CRC) in humans and mice [8-15]. We have previously shown that CRC-associated changes in the gut microbiota directly potentiate colon tumorigenesis in a mouse model of CRC [16]. In that study we observed clear shifts in the microbiota that were associated with a stepwise progression in the number of tumors that developed in the colon. In addition, we showed that transfer of the tumor-associated microbiota to germ-free mice resulted in increased tumor formation relative to germ-free mice that received the microbiota of healthy mice. These results were supported by a subsequent study where we colonized germ-free mice with the microbiota of human donors and observed that different starting communities yielded significant variation in the number of tumors that formed [17]. Combined, these results demonstrate that the microbiota interact with the host to affect tumor susceptibility. A critical question that remains unanswered is what factors and ecological principles mediate the gut microbiota's influence on tumor development. Deciphering how changes in microbial community composition and structure alters gut homeostasis, and subsequently modulate tumorigenesis, is an essential step in understanding the etiology of CRC.

Several bacterial populations including *E. coli*, *Bacteroides fragilis*, and *Fusobacterium nucleatum* have been shown to directly influence tumor development in the colon. The mechanisms by which bacteria potentiate these processes range from the production of

carcinogenic toxins [18,19] to direct manipulation of the inflammatory status of the tumor microenvironment [20,21]. Although individual bacterial populations undoubtedly modulate colorectal carcinogenesis, there are likely a myriad of commensal bacteria that work together to influence tumorigenesis in the colon. This is supported by several studies that have explored the gut microbiota associated with individuals with CRC [8-15,22]. With each study, the number of CRC-associated bacterial populations that likely play a role in tumorigenesis continues to grow. This is likely due to the fact that there is significant functional redundancy within the gut microbiota and various bacterial populations may fill similar roles in tumorigenesis [23-25]. Furthermore, some bacterial populations have been hypothesized to be protective against CRC [26,27]. This protective phenotype may be mediated through metabolite production, induction of immunotolerance, or an ability to outcompete pathogenic bacteria [28]. We hypothesize that multiple bacteria in the gut microbiota have the potential to play pro-tumorigenic or tumor-suppressive roles; thus, the gut microbiota's influence on CRC is likely to be driven by complex interactions within the microbiota and the colonic epithelium.

We have shown that conventionally-raised mice treated with a cocktail of metronidazole, streptomycin, and vancomycin in their drinking water had a significant decrease in tumor numbers using an inflammation-based model of CRC [22]. In the current study, we explored how differential alterations in the microbiota by different antibiotic treatments affected the composition of the microbiota and how changes in bacterial community structure affected tumor susceptibility. Our results confirmed our hypothesis that the microbiota is capable of driving tumorigenesis and that an antibiotic-based intervention during tumor induction can arrest tumorigenesis. Our analysis further supports the model that individual bacterial populations play an important role in CRC, but the ecological interactions and community structure of the gut microbiota mediate the capacity to modulate tumorigenesis.

Results

Antibiotic perturbation of the gut microbiota modulates tumorigenicity. We subjected specific pathogen-free (SPF) C57BL/6 mice to an inflammation-based model of colorectal cancer that utilizes azoxymethane (AOM) as a mutagen and dextran sodium sulfate (DSS) to induce inflammation [16,17,29] (Figure 1A). To determine how differential changes in the gut microbiota affected tumorigenesis, we manipulated the microbiota by administering seven different antibiotic combinations and then quantified the effects of the treatments on the number of tumors observed at the end of the model. Specifically, we treated mice with (i) no antibiotics, (ii) metronidazole, streptomycin, and vancomycin (all antibiotics), (iii) streptomycin and vancomycin (Δ metronidazole), (iv) metronidazole and vancomycin (Δ streptomycin), (v) metronidazole and streptomycin (Δ vancomycin), (vi) metronidazole, (vii) streptomycin, and (viii) vancomycin. Quantitative PCR targeting the 16S rRNA gene indicated that the number of copies of the 16S rRNA gene did not differ between treatment groups, indicating that the different antibiotic combinations did not affect total bacterial load. The three antibiotics were selected based on their reported ability to target general groups of bacteria including anaerobes (metronidazole), Gram-negatives (streptomycin), and Gram-positives (vancomycin). Sequencing the 16S rRNA genes that were present in the feces of conventional and antibiotic-treated mice demonstrated that the different antibiotic treatments generated different bacterial communities prior to AOM injection (Figure 1B); however, the composition of these communities could not have been predicted by the spectrum of the antibiotic that was used to treat the mice (Figure 1C,D). These results demonstrated that perturbation of the microbiota through the use of antibiotics yields a differential capacity for colon tumorigenesis.

Tumor burden can be predicted from the initial microbiota. Tumor burden can be predicted from the initial microbiota. Serial collection of fecal samples allowed us to ascertain the

composition of the microbiota for each mouse and associate it with the number of tumors that developed at the end of the model. Using the 16S rRNA gene sequence data generated from feces collected on the day of AOM injection, we assigned the sequences to operational taxonomic units (OTUs) that were defined as a group of sequences that, on average, were not more than 3% different from each other. We then used the regression-based random forest machine learning algorithm to identify OTUs that would enable us to predict the number of tumors that developed at the end of the model. The model that included OTUs that had an average relative abundance greater than 1.5% resulted in the greatest percentage of the variance explained (Supplementary Figure 1).This model included 15 OTUs and explained 62.6% of the variation in the tumor counts (Figure 2). The OTUs were ranked by their importance in the random forest model as measured by the percent the mean squared error increases when the OTU was removed. When the OTUs were sorted in decreasing order by the percent they contributed to increasing the mean squared error (MSE) of the model, there was a jump between the sixth and seventh OTUs (Figure 2A). In fact, when we reconstructed the model using only the six OTUs that provided the greatest change in the MSE, the model explained 67% of the variation in the observed tumor counts was indicating that the model based on the reduced dataset explained as much of the variation in tumor counts as the model based on all of the OTUs. These six OTUs included members of the Firmicutes (OTU 6), Bacteroidetes (OTUs 4 and 19), Proteobacteria (OTU 3), and Tenericutes (OTUs 34 and 35). Increased numbers of tumors were associated with decreases in the relative abundance of relatives of the Enterobacteriaceae (OTU 3), *Ureaplasma* (OTU 34), and *Lactobacillus* (OTU 6) and increases in the relative abundance of the *Anaeroplasma* (OTU 35), Porphyromonadaceae (OTU 4), and *Prevotella* (OTU 19) (Figure 3). Our random forest modeling demonstrated that it was possible to predict the number of tumors at the end of the model based on the composition of the microbiota at the beginning of the model.

Tumor burden can be predicted from the microbiota at the end of the model. Similar to our analysis using the initial composition of the microbiota, we developed a random forest regression model to predict the number of tumors in the mice based on the composition of the microbiota at the end of the model. The model included 11 OTUs after we again applied a filter requiring each OTU to have an average relative abundance of at least 1.5%. The model explained 52.9% of the variation in the tumor counts (Supplementary Figure 2), which is less than we observed when we modeled tumor counts based on the initial community composition. The seven most important OTUs in the model explained 55.4% of the variation and included *Odoribacter* (OTU 70), *Bacteroides* (OTU 5), *Lactobacillus* (OTU 6), Enterobacteriaceae (OTU 3), *Alloprevotella* (OTU 14), *Prevotella* (OTU 19), and Betaproteobacteria (OTU 17) (Supplementary Figure 3). Interestingly, of the OTUs that were predictive of the number of tumor counts using the baseline and final community composition data, only three of the OTUs overlapped. These included *Lactobacillus* (OTU 6), Enterobacteriaceae (OTU 3), and *Prevotella* (OTU 19).

The microbial community is dynamic during inflammation-associated tumorigenesis. Using mice that were colonized with human feces, we previously reported that tumor burden was associated with the amount of change in the community structure over the course of the AOM-DSS model [17]. In the current study, however, there was a non-significant association between the change in the community structure as measured by the θYC metric of community structure similarity and tumor burden (ρ=0.26, P=0.08; Figure 4A). We did observe that mice that did not receive antibiotics and those that received the Δvancomycin and Δmetronidazole treatments changed the most over the course of the model. Interestingly, when we investigated the temporal progression of the three OTUs that were most important for predicting the number of tumors based on the starting and final community structure (i.e. OTUs 3, 6, and 19; Figure 2B) we observed dynamic changes in relative abundance with time during the course of the

model. These data suggest that the magnitude of change that occurs in a microbial community during tumorigenesis does not influence tumor burden. Instead, specific changes in community structure and the abundance of tumor-associated bacterial populations dictate tumor burden.

Antibiotic intervention during inflammation reduces tumorigenesis. The AOM-DSS model reproduces certain characteristics observed with human CRC, but microbial contributions to tumorigenesis have not been elucidated [29]. To determine whether the gut microbiota modulates tumorigenesis by affecting AOM-induced mutagenesis or DSS-induced inflammation, we performed two antibiotic intervention experiments. We first treated mice with the vancomycin, metronidazole and streptomycin two weeks prior to the administration of AOM and up until the first round of DSS (Figure 1A). We found that these mice had a similar tumor burden to untreated mice (Figure 5). Next, we treated mice between the first and second round of DSS administration, when inflammatory responses were the greatest and aberrant changes in microbial community structure occurs [16] (Figure 1A). With this treatment, there was a significant decrease in the number of tumors (Figure 5). These results suggest that the gut microbiota-mediated effect on CRC is independent of AOM-mediated carcinogenesis. Furthermore, it shows that targeting the gut microbiota at later stages of tumor growth is a viable option for minimizing tumorigenesis and highlights microbiota manipulation as a potential therapeutic in CRC.

Discussion

In the present study, we established the importance of the microbial community structure in determining the extent of tumorigenesis. We demonstrated that manipulation of the murine gut microbiota with different antibiotic regimens resulted in non-overlapping community structures that were associated with disparate levels of tumorigenesis. Enrichment in the relative abundance of several bacterial populations was associated with high and low levels of colon tumors. We determined that the outgrowth of potentially inflammatory members of the gut microbiota was associated with increased tumorigenesis only when there was a corresponding decrease in potentially protective, butyrate producing bacteria. By perturbing the bacterial community at two different time points during the AOM/DSS model, we determined that the gut microbiota affects tumorigenesis via a mechanism that does not involve AOM-induced carcinogenesis. Our experiments also demonstrated that targeting the gut microbiota at the emergence of dysbiosis (i.e. after the first round of DSS in the AOM/DSS model) is a viable strategy for the amelioration of colon tumorigenesis.

In recent years, there has been a focus on identifying bacterial populations that are etiologic agents of CRC. Several commensal bacteria, including *E. coli*, *Fusobacterium nucleatum* and enterotoxigenic *Bacteroides fragilis* (ETBF) have been linked to CRC in humans [18,19,21]. F. nucleatum, which has been detected on the surface of over 50% adenomas in one study, can promote inflammation within the tumor microenvironment in multiple intestinal neoplasia mice [10,20]. ETBF increases tumor multiplicity in the colon of multiple intestinal neoplasia mice through the action of a secreted metalloprotease toxin. It has been estimated that between 5-35% of people carry ETBF [30]. Although there is substantial evidence for a role in potentiating tumorigenesis, the fact that each of these bacteria is only associated with a fraction of CRCs suggests that it is unlikely that there is a single microbial agent that causes cancer. Rather, the

role of the gut microbiota in CRC is likely polymicrobial in nature. The results in the present study support this hypothesis, as we demonstrated that non-overlapping community structures confer similar levels of tumorigenesis in mice. When we examined the relative abundance of bacterial populations associated with increased tumor burden, we never observed consistent enrichment of any one population in the three treatment groups that had the highest tumor levels (i.e., vancomycin only, streptomycin only, and ∆metronidazole). Similarly, potentially protective bacterial populations were not consistently depleted across treatment groups that developed the fewest tumors (All antibiotics, ∆vancomycin, ∆streptomycin, and metronidazole only). This suggests that there may be redundancy in tumor-modulating roles amongst different bacteria populations within the gut microbiota.

During tumor induction, we observed a marked increase in members of the Enterobacteriaceae associated with two antibiotic treatment groups (∆metronidazole and ∆vancomycin). Interestingly, one treatment group (∆vancomycin) developed fewer tumors despite a similar increase in this potentially tumor-modulating bacterial clade. A recent study by Arthur and colleagues [18] showed that in an IL-10-deficient colitis-associated mouse model of CRC, there was an enrichment of Enterobacteriaceae associated with inflammation. This led to an expansion of E. coli populations with genotoxic capabilities and a consequential increase in tumor multiplicity and invasion. Furthermore, members of the Enterobacteriaceae have been shown to perpetuate inflammation in several inflammatory diseases, including ulcerative colitis, which increase an individual's risk of developing CRC [31-33]. When we further examined the two antibiotic treatment groups, we observed that mice with an increased tumor burden had a corresponding decrease in several potentially anti-inflammatory and butyrate producing bacterial populations. These observations support a model by which the pathogenicity potential of individual members of the gut microbiota is ultimately determined by the overall community structure and ecological interactions within the gut microbiota. We hypothesize that

240 inflammatory and carcinogenic commensal bacteria, such as Enterobacteriaceae, can only
241 mediate a pathogenic phenotype if the context of the community structure is conducive.
242
243 One possible mechanism by which community structure mediates tumorigenicity is by shifting
244 the balance of immunomodulatory metabolites and signals. During health, the gut microbiota is
245 an important mediator of immunotolerance, but when the balance of pro- and anti-inflammatory
246 signals is disrupted, gut pathologies can arise [34]. In our mice, Enterobacteriaceae is likely acting
247 as an inflammatory member of the gut microbiota. However, we only observed an increase in
248 tumorigenesis when there was a corresponding depletion of potentially protective members of
249 the genera Clostridium, Enterococcus, and Streptococcus that have reported protective roles
250 against inflammation and tumorigenesis. For example, members of Clostridium are known
251 producers of short chain fatty acids (SCFA) in the colon [26]. SCFA, specifically butyrate, are
252 important nutrients for colonocytes and possess anti-inflammatory and anti-tumor properties
253 [26,35,36]. Furthermore, *Enterococcus* and *Streptococcus* species have been linked to down-
254 regulating inflammatory responses in the colon [37,38]. It is likely that these bacterial populations
255 have the ability to antagonize inflammatory clades (e.g. Enterobacteriaceae) and confer
256 protection; however, when perturbation to the microbial community structure disrupts this
257 homeostasis, opportunistic pathogens can potentiate tumorigenesis.
258
259 In our previous work, we demonstrated that dysbiosis of the gut microbiota generates a pro-
260 inflammatory environment which results in a self-reinforcing pathogenic cascade between the
261 gut microbiota and the host [16,17]. In this study, we demonstrated that antibiotic manipulation of
262 the gut microbiota during the onset of inflammation can significantly decrease tumorigenesis in
263 mice. This highlights the efficacy of targeting the gut microbiota in CRC. Additional studies are
264 needed to explore the viability of manipulating the gut microbiota in CRC with methods such as
265 diet, probiotics, and prebiotics.

Materials & Methods

Animals and animal care. Studies were conducted using adult (8 to 12 week old) age-matched C57BL/6 male mice that were maintained under SPF conditions. Mice were co-housed in groups of five and fed the same autoclaved chow diet. All animal experiments were approved by the University Committee on Use and Care of Animals at the University of Michigan and carried out in accordance with the approved guidelines.

Inflammation-induced colon tumorigenesis. Mice received a single intraperitoneal (i.p.) injection of azoxymethane (10 mg/kg). Water containing 2% DSS was administered to mice beginning on day 5 for 5 days followed by 16 days of water. This was repeated twice for a total of 3 rounds of DSS [16]. Mice were euthanized 3 weeks after the third round of DSS administration for tumor counting. At necropsy, all colons were harvested, flushed of luminal contents, and cut open longitudinally to count and measure tumors.

Antibiotic treatment. Mice were treated with all possible combinations of metronidazole (0.75 g/L), streptomycin (2 g/L), and vancomycin (0.5 g/L) to create eight treatment groups: no antibiotics (N=12), all antibiotics (n=9) (metronidazole, streptomycin, and vancomycin), Δ metronidazole (n=5) (streptomycin and vancomycin), Δ streptomycin (n=5) (metronidazole and vancomycin), Δ vancomycin (n=5) (metronidazole and streptomycin), metronidazole only (N=5), streptomycin only (N=5), and vancomycin only (N=3). Antibiotics were administered in mouse drinking water for 2 weeks prior to and throughout the duration of AOM/DSS administration, unless otherwise specified in Figure 1A. Tumors were enumerated at the end of the model.

16S rRNA quantitative PCR (qPCR) analysis. Relative bacterial loads were quantified by qPCR analysis of bacterial genomic DNA using KAPA SYBR-fast Master Mix (KAPA

291 biosciences) and universal 16S rRNA gene primers (F: ACTCCTACGGGAGGCAGCAGT; R:
292 ATTACCGCGGCTGCTGGC) [39]. Samples were normalized to fecal mass and relative fold
293 change was determined using untreated stool samples for each replicate mouse. Note that
294 qPCR measures relative fold change of 16S rRNA gene copy number, not actual bacterial
295 numbers.
296
297 **DNA extraction and 16S rRNA gene sequencing.** Fecal samples were collected daily from
298 the mice throughout the AOM/DSS protocol and immediately frozen for storage at -20°C. For
299 each mouse, 8 fecal samples distributed over the 73-day timeline of the AOM/DSS model were
300 selected for analysis (Figure 1A). Microbial genomic DNA was extracted using the PowerSoil-
301 htp 96 Well Soil DNA Isolation Kit (MO BIO laboratories) using an EpMotion 5075. The V4
302 region of the 16S rRNA gene from each sample was amplified, sequenced using the Illumina
303 MiSeq Personal Sequencing platform, and curated as described previously using the mothur
304 software package [40,41]. Briefly, we reduced sequencing and PCR errors by requiring reads to
305 fully overlap and in cases where base calls conflicted, we broke the conflict by requiring one
306 base call to have a PHRED quality score 6 units higher than the other otherwise the base call
307 was replaced with an ambiguous base call in the contig. Any reads containing ambiguous base
308 calls were culled. Sequences were aligned to a customized version of the SILVA 16S rRNA
309 sequence database [42] and were screened to insure that they correctly overlapped within the V4
310 region. Chimeric sequences were identified using the de novo implementation of UCHIME and
311 they were culled [43]. The resulting sequences had a median length of 253 nt and we rarefied to
312 2,500 sequences per sample to limit effects of uneven sampling. A mock community was
313 sequenced and processed in parallel to the fecal samples. Based on the mock community data
314 we observed a sequencing error rate of 0.05%. The complete analysis methods and this
315 document as an R-executable document are available at

https://github.com/SchlossLab/Zackular_AbAOMDSS_SciReports_2015. All FASTQ sequence data can be obtained from the Sequence Read Archive at NCBI (Accession SRP056144).

Statistical analysis. The microbiota data were analyzed using the R project for statistical computing. All R source code is available on our GitHub repository at https://github.com/SchlossLab/Zackular_AbAOMDSS_SciReports_2015. All random forest models were made using the randomForest package with 10,000 trees [44]. Diagnostic plots indicated that the percent of the variance explained had stabilized with this number of trees. Comparison of tumor counts were made by carrying out non-parametric pairwise Wilcoxon tests. The resulting p-values were corrected for multiple comparisons using the Benjamini-Hochberg procedure using an experiment-wide Type I error rate of 0.05.

Acknowledgements

This work was supported by grants from the National Institutes for Health to PDS (R01GM099514, R01HG005975, P30DK034933, University of Michigan GI SPORE) and GYC (University of Michigan GI SPORE, ARRA Supplement P30CA4659-22S3, and R01CA166879).

References

1. Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A. & Gordon, J. I. Host-bacterial mutualism in the human intestine. *Science* **307**, 1915-1920, (2005).
2. Levy, R. & Borenstein, E. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proc Natl Acad Sci U S A* **110**, 12804-12809, (2013).
3. Marino, S., Baxter, N. T., Huffnagle, G. B., Petrosino, J. F. & Schloss, P. D. Mathematical modeling of primary succession of murine intestinal microbiota. *Proc Natl Acad Sci U S A* **111**, 439-444, (2014).
4. Lepp, P. W. *et al.* Methanogenic Archaea and human periodontal disease. *Proc Natl Acad Sci U S A* **101**, 6176-6181, (2004).
5. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027-1031, (2006).
6. Tamboli, C. P., Neut, C., Desreumaux, P. & Colombel, J. F. Dysbiosis in inflammatory bowel disease. *Gut* **53**, 1-4, (2004).
7. Saulnier, D. M. *et al.* Gastrointestinal microbiome signatures of pediatric patients with irritable bowel syndrome. *Gastroenterology* **141**, 1782-1791, (2011).
8. Chen, H. M. *et al.* Decreased dietary fiber intake and structural alteration of gut microbiota in patients with advanced colorectal adenoma. *Am J Clin Nutr* **97**, 1044-1052, (2013).
9. Chen, W., Liu, F., Ling, Z., Tong, X. & Xiang, C. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLoS ONE* **7**, e39743, (2012).
10. Kostic, A. D. *et al.* Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res* **22**, 292-298, (2011).

11. Geng, J., Fan, H., Tang, X., Zhai, H. & Zhang, Z. Diversified pattern of the human colorectal cancer microbiome. *Gut Pathog* **5**, 2, (2013).
12. Shen, X. J. *et al.* Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. *Gut Microbes* **1**, 138-147, (2010).
13. Sobhani, I. *et al.* Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS ONE* **6**, e16393, (2011).
14. Wang, T. *et al.* Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J* **6**, 320-329, (2012).
15. Ahn, J. *et al.* Human gut microbiome and risk for colorectal cancer. *J Natl Cancer Inst* **105**, 1907-1911, (2013).
16. Zackular, J. P. *et al.* The gut microbiome modulates colon tumorigenesis. *MBio* **4**, e00692-00613, (2013).
17. Baxter, N. T., Zackular, J. P., Chen, G. Y. & Schloss, P. D. Structure of the gut microbiome following colonization with human feces determines colonic tumor burden. *MIcrobiome* **2**, 20, (2014).
18. Arthur, J. C. *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **338**, 120-123, (2012).
19. Sears, C. L. *et al.* Association of enterotoxigenic *Bacteroides fragilis* infection with inflammatory diarrhea. *Clin Infect Dis* **47**, 797-803, (2008).
20. Kostic, A. D. *et al. Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **14**, 207-215, (2013).
21. Rubinstein, M. R. *et al. Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell Host Microbe* **14**, 195-206, (2013).

22. Zackular, J. P., Rogers, M. A. M., Ruffin, M. T. & Schloss, P. D. The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prevention Research* **7**, 1112-1121, (2014).

23. Lepage, P. *et al.* A metagenomic insight into our gut's microbiome. *Gut* **62**, 146-158, (2013).

24. Turnbaugh, P. J. *et al.* The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* **1**, 6ra14, (2009).

25. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65, (2010).

26. Louis, P. & Flint, H. J. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett* **294**, 1-8, (2009).

27. Appleyard, C. B. *et al.* Pretreatment with the probiotic VSL#3 delays transition from inflammation to dysplasia in a rat model of colitis-associated cancer. *Am J Physiol Gastrointest Liver Physiol* **301**, G1004-1013, (2011).

28. Zhu, Y., Michelle Luo, T., Jobin, C. & Young, H. A. Gut microbiota and probiotics in colon tumorigenesis. *Cancer Lett* **309**, 119-127, (2011).

29. De Robertis, M. *et al.* The AOM/DSS murine model for the study of colon carcinogenesis: From pathways to diagnosis and therapy studies. *J Carcinog* **10**, 9, (2011).

30. Housseau, F. & Sears, C. L. Enterotoxigenic Bacteroides fragilis (ETBF)-mediated colitis in Min (Apc+/-) mice: a human commensal-based murine model of colon carcinogenesis. *Cell Cycle* **9**, 3-5, (2010).

31. Rolhion, N. & Darfeuille-Michaud, A. Adherent-invasive Escherichia coli in inflammatory bowel disease. *Inflamm Bowel Dis* **13**, 1277-1283, (2007).

32. Garrett, W. S. *et al.* Communicable ulcerative colitis induced by T-bet deficiency in the innate immune system. *Cell* **131**, 33-45, (2007).

33. Rooks, M. G. *et al.* Gut microbiome composition and function in experimental colitis during active disease and treatment-induced remission. *ISME J* **8**, 1403-1417, (2014).
34. Kelly, D., Conway, S. & Aminov, R. Commensal gut bacteria: mechanisms of immune modulation. *Trends Immunol* **26**, 326-333, (2005).
35. Hague, A., Elder, D. J., Hicks, D. J. & Paraskeva, C. Apoptosis in colorectal tumour cells: induction by the short chain fatty acids butyrate, propionate and acetate and by the bile salt deoxycholate. *Int J Cancer* **60**, 400-406, (1995).
36. Donohoe, D. R. *et al.* The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* **13**, 517-526, (2011).
37. Wang, S., Ng, L. H., Chow, W. L. & Lee, Y. K. Infant intestinal Enterococcus faecalis down-regulates inflammatory responses in human intestinal cell lines. *World J Gastroenterol* **14**, 1067-1076, (2008).
38. Kaci, G. *et al.* Inhibition of the NF-kappaB pathway in human intestinal epithelial cells by commensal Streptococcus salivarius. *Appl Environ Microbiol* **77**, 4681-4684, (2011).
39. Vaishnava, S. *et al.* The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255-258, (2011).
40. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* **79**, 5112-5120, (2013).
41. Schloss, P. D. *et al.* Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, 7537-7541, (2009).
42. Pruesse, E. *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**, 7188-7196, (2007).

43. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194-2200, (2011).

44. Breiman, L. Random forests. *Machine Learning* **45**, 5-32, (2001).

Contributions

All authors contributed to the design of the experiments. JPZ and NTB carried out the experiments and generated the data. JPZ and PDS analyzed the data. All authors participated in interpreting the results. JPZ and PDS wrote the manuscript and NTB and GYC helped with the final editing of the text.

Competing financial interests

The authors declare no competing financial interests.

Figure 1. Antibiotic perturbation drives changes in microbial community structure and final tumor burden. The AOM-DSS model was administered to C57BL/6 mice reared under standard pathogen free (SPF) conditions with different antibiotic perturbations; Black arrows indicate fecal samples that used for our analysis (A). The mice were treated with all possible combinations of metronidazole, streptomycin, and vancomycin to create eight treatment groups, which resulted in considerable differences in the taxonomic composition of the gut communities at the start of the model (Day 0) (B). These communities resulted in a continuum of tumor burden in the mice (C and D). The stars indicate which treatments yielded a significantly ($P<0.05$) different number of tumors when compared to the treatment with the vertical line.

Figure 2. A random forest model successfully predicted the number of tumors in the mice at the end of the model (A) based on their microbiota composition at the start of the model (B). The OTUs in B are ranked in decreasing order of their mean decrease in the mean squared error. The relationships between the first 6 OTUs and the number of tumors found in those mice are shown in Figure 3.

Figure 3. Relationship between the initial relative abundance of the most informative OTUs from the random forest model with the number of tumors found in the mice at the end of the model. The vertical gray line indicates the limit of detection. Panels are ordered in decreasing order of the percent increase in the mean squared error of the model when that OTU was removed.

Figure 4. The murine microbiota is dynamic but the amount of change is not associated with the final number of tumors. The structure of the gut microbiota associated with untreated and the Δ metronidazole and Δ vancoymcin-treated mice changed the most throughout the model as measured using the ΘYC distance metric (A). OTUs 3, 6, and 19 were among the

most salient features for predicting tumor burden at the beginning and end of the model (B). The plotting symbols and characters are the same as those used in Figure 1. In panel B, the median relative abundance is indicated by the plotting symbol and the range of observed relative abundances is plotted by the vertical bar. The vertical blue regions indicate when the DSS treatments were applied.

Figure 5. Antibiotic intervention prior to second administration of DSS alleviates tumor burden. Interventions with an antibiotic cocktail of metronidazole, vancomycin, and streptomycin were performed as depicted in Figure 1A with enumeration of tumors performed at the end point of the model (A). Representative images of tumors in the distal colon of mice from each treatment group (B).

Supplementary Figure 1. Quality of random forest regression fit as a function of the minimum average relative abundance an OTU must have to be included in the model. The integers displayed across the plot indicate the number of OTUs that were included in the model. Because a minimum average relative abundance of 1.5% yielded the best R2, it was used for the remainder of the analysis.

Supplemental Figure 2. A random forest model successfully predicted the number of tumors in the mice at the end of the model (A) based on their microbiota composition at the start end of the model (B). The OTUs in B are ranked in decreasing order of their mean decrease in the MSE. The relationships between the first 6 OTUs and the number of tumors found in those mice are shown in Supplemental Figure 3.

Supplemental Figure 3. Relationship between the initial relative abundance of the most informative OTUs from the random forest model with the number of tumors found in the mice at the end of the model. The vertical gray line indicates the limit of detection.