Integrating tumor genomics into studies of the microbiome in colorectal cancer

Michael B. Burns \textsuperscript{a} and Ran Blekhman \textsuperscript{b}

\textsuperscript{a}Department of Biology, Loyola University Chicago, Chicago, IL, USA; \textsuperscript{b}Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN, USA

ABSTRACT
Although the gut microbiome has been linked to colorectal cancer (CRC) development, associations of microbial taxa with CRC status are often inconsistent across studies. We have recently shown that tumor genomics, a factor that is rarely incorporated in analyses of the CRC microbiome, has a strong effect on the composition of the microbiota. Here, we discuss these results in the wider context of studies characterizing interaction between host genetics and the microbiome, and describe the implications of our findings for understanding the role of the microbiome in CRC.

ARTICLE HISTORY
Received 6 September 2018
Revised 12 October 2018
Accepted 30 October 2018

KEYWORDS
Microbiome; colorectal cancer; tumor genomics

Introduction
The gut microbiota has been consistently shown to be important for the development of colorectal cancer (CRC).	extsuperscript{1–6} However, recent meta-analyses of CRC-associated stool, mucosal, and tissue microbiota in the gut have demonstrated that there is a great deal of inter-study variability in the field, leading to mixed interpretations of which microbial taxa, if any, are consistently associated with CRC.	extsuperscript{7–11} Potential sources of cross-study differences can be variation in sample collection, DNA extraction, sequencing target region, sequencing technology used, as well as a wide array of other technical differences.	extsuperscript{12–16} Considering recent evidence that host genetic variation is correlated with the composition of the microbiome, another potential source of variability is the heterogeneity of tumors, and especially the genomic mutational profiles of the tumors. In our recent study, we assessed the effect of tumor mutational profiles on the microbiota in the tumor microenvironment.	extsuperscript{20} Our analysis shows that loss-of-function (LoF) mutations in cancer-related genes and pathways in the tumor are correlated with defined microbial communities, and that the microbiome can be used to statistically predict tumor mutational profiles. Here, we describe the results of our study in the context of the current knowledge in the field, discuss the implications, and explore next steps and remaining open questions.

Human genetic control of the microbiome
Several studies have investigated the role of human genetics in shaping host-associated microbial communities. Initial studies focused on candidate genes of interest, and found human genetic variants that can control microbiome composition in specific disease contexts.	extsuperscript{21,22} More recent studies have used a genome-wide approach to calculate the heritability of microbial taxa and identify variants and genes in the human genome that correlate with variation in the microbiome across body sites.	extsuperscript{17,18,23–26} In addition, several studies have used similar analysis techniques to assess host genetic effects on the microbiome in the context of various human diseases and conditions.	extsuperscript{26–29} Although results across cohorts have not had a high degree of overlap, several loci have been found to be associated across multiple studies,	extsuperscript{19,30,31} indicating that in some cases, host genes can affect microbiome composition across different contexts. In our recent study we aimed to assess the effect of host genetics on the microbiome in the context of CRC.

Although the approach used in our study is in some ways similar to GWAS studies of the microbiome, there are several important differences. First, our study investigated somatic mutations in the host, as opposed to host genetic variation from blood samples representing germline variation. To identify tumor somatic mutations, we compared whole
exome sequencing data from each tumors to matched normal tissue from the same patient (Figure 1). Second, our study analyzed the microbiome in colonic tissue samples, as opposed to fecal samples, which are more commonly used in studies of the CRC microbiome. Investigating the microbes in the microenvironment of tumors is informative, due to the ability to detect microbial signatures that may be lost in fecal samples. On the other hand, fecal samples are abundant, non-invasive, and likely more useful for development of microbial diagnostics. Third, our analysis of host genetic variants focused on a small number of genetic variants, namely LoF mutations in coding regions. While most studies of host genetic correlations with the microbiome used genome-wide techniques, we focused on coding mutations with severe effects, as these mutations are more likely to affect tumor physiology and act as tumor driver mutations. In the following sections we discuss these analyses in the context of previous studies of the microbiome in CRC.

**Tumor genomic profiles can influence the microbiome**

Our recent study aimed to test the hypothesis that different CRC tumor subtypes would harbor different microbial communities. This work builds on our previous report that included an evaluation of the differences between the microbiota found at tumor sites and those surrounding normal gut tissues from matched patients (Figure 1). The findings were consistent with other reports in that microbial communities at tumor sites were more diverse than those of normal tissues. Additionally, *Fusobacterium* and *Providencia* were two of the specific microbial taxa that exhibited elevated abundance at tumor sites relative to matched normal tissue sites. Predictive functional analyses indicated that these two genera harbor known virulence genes that may have a role in driving tumor development.

In our most recent study, we incorporated data on the mutational landscape of the tumors to information on their microbiome (Figure 1). This inter-tumor comparison using groupings defined by the LoF mutations in the tumor samples yielded promising results. Prevalent LoF mutations in 5 tumor genes were found to harbor distinct microbial communities, including *ANKRD36C*, *APC*, *CTBP2*, *KMT2C*, and *ZNF717*. In addition, when aggregating LoF mutations at the pathway level, we found sets of microbial taxa that discriminated between tumors with LoF mutations in 21 KEGG pathways and 15 PID pathways, including MAPK, Phosphatidylinositol, TP53, Wnt, and Notch signaling pathways. Correlations

![Figure 1](image-url)  
Figure 1. Illustration of the experimental design, comparing normal (left) and tumor (right) samples from the same individual. The analysis is based on using the normal sample as the baseline for each patient, and compared the changes from the patient-matched baseline in the microbiome and tumor genome.
between cancer phenotypes and these pathways are not surprising; however, these results suggest that the mucosal and tissue microbial communities at the tumor site themselves reflect genetic changes in these pathways. Of note, the accuracy of the statistical prediction was higher at the pathway level compared to mutations at the gene level. We hypothesize that this is due to the aforementioned potential for discrete mutations to have similar effects at the pathway level.

**Significance and implications**

The correlations found in this work are significant for several reasons. First, our results show that genetic heterogeneity among tumors is reflected in the mucosal and tissue microbial communities at the tumor sites. A recent meta-analysis found that studies of the CRC microbiome, taken as a whole, show few consistent findings with respect to identification of specific microbial taxa that track with cancer status. The interpretation of this finding is that the signals found in individual studies are lost due to the unique composition of the sets of confounding factors that track with each of the studies. In addition to this variability at the level of study design, experimental techniques, and analytical approaches, variation in the microbiome can potentially be caused by heterogeneity in CRC tumor genomes. Thus, future studies of the microbiota in CRC may benefit from incorporating tumor mutational information in the analysis.

A second implication of this work is the finding is that sets of microbial taxa, rather than individual microbes, correlate with tumor mutational status. While it is entirely possible that there are some specific microbes that alone are discriminatory, it is also likely, due to the ecological dynamics at play in the communities of interest, that a variety of microbes are acting in concert with one another and with the tumor microenvironment. One way to view this is to see individual taxa not just as discrete genera and species, but as placeholders for functional modules. As suggested in our previous work, it is possible that unrelated taxa exert similar functional effects. This suggests that in addition to using taxonomic information, studies that include functional information are likely to yield more holistic biological insight.

Lastly, we have demonstrated that it is possible to statistically predict tumor mutational status from microbiome composition alone, in this first instance, using mucosal/tissue-associated microbial communities. Using taxa abundances, we were able to accurately predict whether LoF mutations are found within several cancer-related genes, including APC. Moreover, a similar approach allowed us to use microbiome data to statistically predict whether cancer-related pathways, including MAPK and Wnt signaling, harbor LoF mutations in their protein-coding sequence. This ability to use the microbiome as a predictor for tumor mutational profiles has several implications. First, there is strong interest in using the gut microbiome as a diagnostic for CRC. Our results suggest that, in addition to the prediction of CRC status using gut microbiome profiles, these profiles may harbor information on the individual genes and pathways that are mutated in the tumor – information that can potentially be used for diagnostic purposes. Second, these results may be relevant for understanding why statistical prediction of CRC status from microbiome data has been challenging. Specifically, we find that different microbial taxa were important for the prediction of different tumor pathways. This may indicate that prediction of a single endpoint (CRC present or not) could be complicated by microbiome variation that is correlated with tumor genomic profiles, suggesting that incorporating this information in prediction models may be advantageous. Finally, and more generally, our results add to a growing body of literature showing that host genetics can affect microbiome composition, and suggest that knowledge of the microbial taxa and host genes that interact may be beneficial in the development of microbiome-based diagnostics.

**Next steps and open questions**

There are still many questions that need to be answered with respect to the effect of tumor genomics on the microbiome. For instance, we currently know little about the contribution of the microbiome to development of CRC and the interaction with different driver mutations during this process (Figure 2). Starting from the development of early lesions (Figure 2(a)), through early expansion, formation of adenoma (Figure 2(b)), and formation of...
the tumor (Figure 2(c)), there is a continuous cross-talk between the microbiota and lesion. Our results indicate that this cross-talk is likely affected by variation in tumor mutational profiles. Nevertheless, we still do not understand how the acquisition of new mutations affects the temporal interaction with the microbiota at the various tumor development stages. Importantly, our work only described correlations between tumor gene and pathway mutations and the microbiome. There are several, non-exclusive hypotheses that could explain potential mechanisms. For instance, host tissues could accumulate genomic mutations that drive changes in the local microenvironment, resulting in changes to the local community of microorganisms. At the same time, this community of microorganisms might contribute to a microenvironment that is more permissive for tumor growth and development, leading to accumulation of mutations. Future studies, utilizing model systems to study the directionality of host-microbiome interactions, will be able to pinpoint causal effects and characterize links between tumor genes and microbial taxa.

Our study focused entirely on the mucosal and tissue microbiome, namely, the bacteria that are in direct physical contact with the tumor. We still do not know to what extent tumor mutational profiles correlate with the microbiome of fecal samples, which are most commonly used for gut microbiome analyses. In addition, our study only highlighted one aspect of tumor genomics, specifically loss-of-function mutations in the DNA of coding genes. However, host-microbiome interactions of interest likely involve genetic variation in non-coding regions, as these may affect gene expression. Several recent studies have shown that the microbiome can affect gene regulation in interacting host cells, and in particular in colonic tumors. In addition, the effects of immune elements, tumor heterogeneity, epigenetic factors, and the developmental history of the tissue may also be important for interaction with the microbiome. It is also likely that the viral community present in the gut plays an important role, interacting with both host and microbial cells. Lastly, in every analysis of the microbiome it is critical to incorporate information on host environmental (non genetic) factors, such as diet, medication use, medical history, and other life history traits. Given the complexity of the biological system at hand, dissecting the mechanistic interplay between the tumor and its microbiome will require multi-level approaches to succeed.

Funding

This work was supported by the Minnesota Partnership for Biotechnology and Medical Genomics;

ORCID

Michael B. Burns http://orcid.org/0000-0001-9791-4359

References

1. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL, et al. Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host Microbe. 2013;14:207–215. doi:10.1016/j.chom.2013.07.007.
2. Schwabe RF, Jobin C. The microbiome and cancer. Nat Rev Cancer. 2013;13:800–812. doi:10.1038/nrc3610.
3. Wu S, Rhee KJ, Albesiano E, Rabizadeh S, Wu X, Yen HR, Huso DL, Brancati FL, Wick E, McAllister F, et al. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. Nat Med. 2009;15:1016–1022. doi:10.1038/nmm.2015.

4. Chung L, Thiele Orberg E, Geis AL, Chan JL, Fu K, DeStefano Shields CE, Dejea CM, Fathi P, Chen J, Finard BB, et al. Bacteroides fragilis toxin coordinates a pro-carcinogenic inflammatory cascade via targeting of colonic epithelial cells. Cell Host Microbe. 2018;23:203–14.e5. doi:10.1016/j.chom.2018.01.007.

5. Zackular JP, Baxter NT, Iverson KD, Sadler WD, Petrosino JF, Chen GY, Schloss PD. The gut microbiome modulates colon tumorigenesis. MBio. 2013;4: e00692–13. doi:10.1128/mBio.00692-13.

6. Ahn J, Sinha R, Pei Z, Dominiani C, Wu J, Shi J, Goedert JJ, Hayes RB, Yang L. Human gut microbiome and risk for colorectal cancer. J Natl Cancer Inst. 2013;105:1907–1911. doi:10.1093/jnci/djt300.

7. Sze MA, Schloss PD. Leveraging existing 16S rRNA gene surveys to identify reproducible biomarkers in individuals with colorectal tumors. MBio. [Internet] 2018;9. doi:10.1128/mBio.00630-18.

8. Shah MS, DeSantis TZ, Weinmaier T, McMurdie PJ, Cope JL, Altrichter A, Yamal J-M, Hollister EB. Leveraging sequence-based faecal microbial community survey data to identify a composite biomarker for colorectal cancer. Gut. 2018;67:882–891. doi:10.1136/gutjnl-2016-313189.

9. Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. Nat Commun. 2017;8:1784. doi:10.1038/s41467-017-01973-8.

10. Sinha R, Abu-Ali G, Vogtmann E, Fodor AA, Ren B, Amir A, Schwager E, Crabtree J, Ma S. Microbiome quality control project consortium, et al. assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. Nat Biotechnol. 2017;35:1077–1086.

11. Drewes JL, White JR, Dejea CM, Fathi P, Iyadorai T, Vadivelu J, Roslani AC, Wick EC, Mongodin EF, Loke MF, et al. High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. NPJ Biofilms Microbiomes. 2017;3:34. doi:10.1038/s41522-017-0040-3.

12. Rintala A, Pietilä S, Munukka E, Eerola E, Pursiheimo J-P, Laiho A, Pekkala S, Huovinen P. Gut microbiota analysis results are highly dependent on the 16S rRNA gene target region, whereas the impact of DNA extraction is minor. J Biomol Tech. 2017;28:19–30.

13. Human microbiome project consortium. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486:207–214. doi:10.1038/nature11234.

14. Knudsen BE, Bergmark L, Munk P, Lukjanenko O, Priemé A, Aarestrup FM, Pamp SJ. Impact of sample type and DNA isolation procedure on genomic inference of microbiome composition. mSystems. [Internet] 2016;1. doi:10.1128/mSystems.00095-16.

15. Panek M, Čipić Paljetak H, Barešić A, Perić M, Matijašić M, Lojkšić I, Vranesić Bender D, Krznarić Ž, Verbanac D. Methodology challenges in studying human gut microbiota - effects of collection, storage, DNA extraction and next generation sequencing technologies. Sci Rep. 2018;8:5143. doi:10.1038/s41598-018-23296-4.

16. Blekhman R, Tang K, Archie EA, Barreiro LB, Johnson ZP, Wilson ME, Kohn J, Yuan ML, Gesquiere L, Grieneisen LE, et al. Common methods for fecal sample storage in field studies yield consistent signatures of individual identity in microbiome sequencing data. Sci Rep. 2016;6:31519. doi:10.1038/srep31519.

17. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, Beaumont M, Van Treuren W, Knight R, Bell JT, et al. Human genomics shapes the gut microbiome. Cell. 2014;159:789–799. doi:10.1016/j.cell.2014.09.052.

18. Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowsk R, Bell JT, Spector TD, Keinan A, Ley RE, Gevers D, et al. Host genetic variation impacts microbiome composition across human body sites. Genome Biol. 2015;16:191. doi:10.1186/s13059-015-0667-4.

19. Goodrich JK, Davenport ER, Clark AG, Ley RE. The relationship between the human genome and microbiome comes into view. Annu Rev Genet. 2017;51:413–433. doi:10.1146/annurev-genet-110711-155532.

20. Burns MB, Montassier E, Abrahante J, Priya S, Niccum DE, Khoruts A, Starr TK, Knights D, Blekhman R. Colorectal cancer mutational profiles correlate with defined microbial communities in the tumor microenvironment. PLoS Genet. 2018;14:e1007367. doi:10.1371/journal.pgen.1007664.

21. Tong M, McHardy I, Ruegger P, Goudarzi M, Kashyap PC, Haritunians T, Li X, Graebeler TG, Schwager E, Huttenhower C, et al. Reprograming of gut microbiome energy metabolism by the FUT2 Crohn’s disease risk polymorphism. ISME J. 2014;8:2193–2206. doi:10.1038/ismej.2014.64.

22. Khachatryan ZA, Ktsiyan ZA, Manukyan GP, Kelly D, Ghazaryan KA, Aminov RI. Predominant role of host genetics in controlling the composition of gut microbiota. PLoS ONE. 2008;3:e3064. doi:10.1371/journal.pone.0003064.

23. Igartua C, Davenport ER, Gilad Y, Nicolae DL, Pinto J, Ober C. Host genetic variation in mucosal immunity pathways influences the upper airway microbiome. Microbiome. 2017;5:16. doi:10.1186/s40168-016-0227-5.
24. Zhermakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, Mujagic Z, Vila AV, Falony G, Vieira-Silva S, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science. 2016;352:565–569. doi:10.1126/science.aad3369.

25. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D, et al. Population-level analysis of gut microbiome variation. Science. 2016;352:560–564. doi:10.1126/science.aad3503.

26. Davenport ER, Cusanovich DA, Michelinii K, Barreiro LB, Ober C, Gilad Y. Genome-Wide association studies of the human gut microbiota. PLoS ONE. 2015;10:e0140301. doi:10.1371/journal.pone.0140301.

27. Le Roy CI, Beaumont M, Jackson MA, Stieves CJ, Spector TD, Bell JT. Heritable components of the human fecal microbiome are associated with visceral fat. Gut Microbes. 2018;9:61–67. doi:10.1080/19490976.2017.1356556.

28. Knights D, Silverberg MS, Weersma RK, Gevers D, Dijkstra N, Huang H, Tyler AD, van Sommenen S, Imhann F, Stempak JM, et al. Complex host genetics influence the microbiome in inflammatory bowel disease. Genome Med. 2014;6:107. doi:10.1186/s13073-014-0107-1.

29. Beaumont M, Goodrich JK, Jackson MA, Yet I, Davenport ER, Vieira-Silva S, Debelsius J, Pallister T, Mangino M, Raes J, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biol. 2016;17:189. doi:10.1186/s13073-016-1052-7.

30. Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N, et al. Environment dominates over host genetics in shaping human gut microbiota. Nature. 2018;555:210–215. doi:10.1038/nature25973.

31. Weissbrod O, Rothschild D, Barkan E, Segal E. Host genetics and microbiome associations through the lens of genome wide association studies. Curr Opin Microbiol. 2018;44:9–19.

32. Baxter NT, Ruffin MT, Rogers MAM, Schloss PD. Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. Genome Med. 2016;8:1–10. doi:10.1186/s13073-016-0290-3.

33. Pon JR, Marra MA. Driver and passenger mutations in cancer. Annu Rev Pathol. 2015;10:25–50. doi:10.1146/annurev-pathol-012414-040312.

34. Burns MB, Lynch J, Starr TK, Knights D, Blekhman R. Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. Genome Med. 2015;7:55. doi:10.1186/s13073-015-0177-8.

35. Mira-Pascual L, Cabrera-Rubio R, Ocon S, Costales P, Parra A, Suarez A, Moris F, Rodrigo L, Mira A, Collado MC. Microbial mucosal colonic shifts associated with the development of colorectal cancer reveal the presence of different bacterial and archaeal biomarkers. J Gastroenterol. 2015;50:167–179. doi:10.1007/s00535-014-0963-x.

36. Shen XJ, Rawls JF, Randall T, Burcal L, Mpande CN, Jenkins N, Jovob B, Abdo Z, Sandler RS, Keku TO. Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. Gut Microbes. 2010;1:138–147. doi:10.4161/gmic.1.3.12360.

37. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30. doi:10.1093/nar/28.1.27.

38. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014;42:D199–205. doi:10.1093/nar/gkt1076.

39. Schaefer CF, Anthony K, Krupa S, Buchhoff J, Day M, Hannay T, Buetow KH. PID: the pathway interaction database. Nucleic Acids Res. 2009;37:D674–9. doi:10.1093/nar/gkn653.

40. Zackular JP, Rogers MA, Ruffin MTT, Schloss PD. The human gut microbiome as a screening tool for colorectal cancer. Cancer Prev Res. [Internet] 2014. doi:10.1158/1940-6207.capr-14-0129.

41. Zitvogel L, Ma Y, Raoult D, Kroemer G, Gajewski TF. The microbiome in cancer immunotherapy: diagnostic tools and therapeutic strategies. Science. 2018;359:1366–1370. doi:10.1126/science.aar6918.

42. Luca F, Kupfer SS, Knights D, Khoruts A, Blekhman R. Functional genomics of host-microbiome interactions in humans. Trends Genet. 2018;34:30–40. doi:10.1016/j.tig.2017.10.001.

43. Sommer F, Nookaew I, Sommer N, Fogelstrand P, Bäckhed F. Site-specific programming of the host epithelial transcriptome by the gut microbiota. Genome Biol. 2015;16:62. doi:10.1186/s13059-015-0667-4.

44. Camp JG, Frank CL, Lickwar CR, Guturu H, Rube T, Wenger AM, Chen J, Bejerano G, Crawford GE, Rawls JF. Microbiota modulate transcription in the intestinal epithelium without remodeling the accessible chromatin landscape. Genome Res. 2014;24:1504–1516. doi:10.1101/gr.165845.113.

45. Davison JM, Lickwar CR, Song L, Breton G, Crawford GE, Rawls JF. Microbiota regulate intestinal epithelial gene expression by suppressing the transcription factor Hepatocyte nuclear factor 4 alpha. Genome Res. [Internet] 2017. doi:10.1101/gr.220111.116.

46. Yuan C, Burns MB, Subramanian S, Blekhman R. Interaction between host MicroRNAs and the gut microbiota in colorectal cancer. mSystems. [Internet] 2018;3. doi:10.1128/mSystems.00205-17.

47. Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA. The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. Cancer Cell. 2018;33:570–580. doi:10.1016/j.ccell.2018.03.015.