

RESEARCH ARTICLE

# Fecal Microbiota, Fecal Metabolome, and Colorectal Cancer Interrelations

Rashmi Sinha<sup>1</sup>, Jiyoung Ahn<sup>2</sup>, Joshua N. Sampson<sup>1</sup>, Jianxin Shi<sup>1</sup>, Guoqin Yu<sup>1</sup>, Xiaogin Xiong<sup>3</sup>, Richard B. Hayes<sup>2</sup>, James J. Goedert<sup>1\*</sup>

**1** Epidemiology and Biostatistics Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Bethesda, Maryland 20892–9704, United States of America, **2** Division of Epidemiology, Department of Population Health, New York University School of Medicine, 650 First Avenue, #518, New York, New York 10016, United States of America, **3** Information Management Services, 6110 Executive Boulevard, Rockville, Maryland 20852, United States of America

\* [goedertj@mail.nih.gov](mailto:goedertj@mail.nih.gov)



## Abstract

### Background and Aims

Investigation of microbe-metabolite relationships in the gut is needed to understand and potentially reduce colorectal cancer (CRC) risk.

### Methods

Microbiota and metabolomics profiling were performed on lyophilized feces from 42 CRC cases and 89 matched controls. Multivariable logistic regression was used to identify statistically independent associations with CRC. First principal coordinate-component pair (PCo1-PC1) and false discovery rate (0.05)-corrected P-values were calculated for 116,000 Pearson correlations between 530 metabolites and 220 microbes in a sex\*case/control meta-analysis.

### Results

Overall microbe-metabolite PCo1-PC1 was more strongly correlated in cases than in controls (Rho 0.606 vs 0.201,  $P = 0.01$ ). CRC was independently associated with lower levels of Clostridia, Lachnospiraceae, *p*-aminobenzoate and conjugated linoleate, and with higher levels of Fusobacterium, Porphyromonas, *p*-hydroxy-benzaldehyde, and palmitoyl-sphingomyelin. Through postulated effects on cell shedding (palmitoyl-sphingomyelin), inflammation (conjugated linoleate), and innate immunity (*p*-aminobenzoate), metabolites mediated the CRC association with Fusobacterium and Porphyromonas by 29% and 34%, respectively. Overall, palmitoyl-sphingomyelin correlated directly with abundances of Enterobacteriaceae (Gammaproteobacteria), three Actinobacteria and five Firmicutes. Only Parabacteroides correlated inversely with palmitoyl-sphingomyelin. Other lipids correlated inversely with Alcaligenaceae (Betaproteobacteria). Six Bonferroni-significant correlations were found, including low indolepropionate and threnoylvaline with Actinobacteria and high erythronate and an uncharacterized metabolite with Enterobacteriaceae.

## OPEN ACCESS

**Citation:** Sinha R, Ahn J, Sampson JN, Shi J, Yu G, Xiong X, et al. (2016) Fecal Microbiota, Fecal Metabolome, and Colorectal Cancer Interrelations. PLoS ONE 11(3): e0152126. doi:10.1371/journal.pone.0152126

**Editor:** Peh Yean Cheah, Singapore General Hospital, SINGAPORE

**Received:** September 15, 2015

**Accepted:** March 9, 2016

**Published:** March 25, 2016

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the National Cancer Institute Intramural Research Program and grants R03CA159414 and R01CA159036. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

**Abbreviations:** BMI, body mass index; CRC, colorectal cancer; CI, confidence interval; FDR, false

discovery rate; OR, odds ratio; PABA, *p*-aminobenzoate; PC, principal component; PCo, principal coordinate; SD, standard deviation.

## Conclusions

Feces from CRC cases had very strong microbe-metabolite correlations that were predominated by Enterobacteriaceae and Actinobacteria. Metabolites mediated a direct CRC association with *Fusobacterium* and *Porphyromonas*, but not an inverse association with Clostridia and Lachnospiraceae. This study identifies complex microbe-metabolite networks that may provide insights on neoplasia and targets for intervention.

## Introduction

The gut microbial population (microbiota) carries greater than 100-fold more genes than the human genome, through which it regulates numerous processes, such as energy harvesting, metabolism of dietary components, immunity, and activities of host or microbial derived chemicals.[1] Alteration or frank dysfunction of these processes is closely tied to inflammatory bowel disease, malnutrition and metabolic syndrome,[2–4] and it influences the risk for a wide range of diseases including colorectal cancer (CRC).[5–11] Whole-genome shotgun sequencing has provided insights on the metabolic potential of the gut microbiota, especially in studies that included transcriptomics.[1, 12–14] Targeted insights have come from studies of microbial consortia, dietary interventions, gnotobiotic mouse models, and transfer of fecal microbiota from diseased or healthy people.[3, 13, 15] Despite such progress, a comprehensive comparison of all detectable metabolites with all microbes in the distal human gut is lacking.

We have previously reported CRC associations with the fecal microbiota, specifically decreased relative abundance of Lachnospiraceae and other Clostridia and increased carriage of *Fusobacterium*, *Atopobium*, and *Porphyromonas*. [16] In the same population, CRC was associated with differences from the matched controls in levels of dozens of fecal metabolites. [17] Herein, we sought to uncover correlations between fecal microbes and metabolites and to identify statistically independent differences between CRC and matched controls.

## Materials and Methods

### Study participants and specimens

The study design has been described previously.[18, 19] Briefly, newly diagnosed cases with adenocarcinoma of the colon or rectum were recruited prior to surgery and treatment during 1985–1987.[18, 19] Controls were patients awaiting elective surgery for non-oncologic, non-gastrointestinal conditions at these hospitals during the same period. A median of 6 days (interquartile range, 3–13 days) prior to hospitalization and surgery, participants completed dietary and demographic questionnaires and provided two-day fecal samples that were frozen at home on dry ice and subsequently lyophilized. The two-day lyophilates were pooled, mixed and stored at -40°C. Participants provided written informed consent. The consent process and study procedures were reviewed and approved by an Institutional Review Board at the National Cancer Institute.[18, 19]

Of 69 cases and 114 controls in the original study,[18, 19] the case-control analysis included 48 cases and 102 controls for whom at least 100mg of lyophilized feces was available. Controls were frequency matched to cases by gender and body mass index (BMI). Microbiota and metabolomic analyses were conducted with these lyophilized fecal samples. As described previously, [16, 17] in both assays systems, the data were of excellent quality and highly reproducible. For the current analyses, there were 42 cases and 89 controls that had both metabolomics and microbiota data.

## Microbiota analyses

The details on the amplification, sequencing, classification and analysis of 16S rRNA genes are in Ahn et al.[16] Briefly, DNA was extracted using the Mobio PowerSoil DNA Isolation Kit (Carlsbad, CA). 16S rRNA amplicons covering variable regions V3 to V4 were generated, and the amplicons were sequenced with the 454 Roche FLX Titanium pyrosequencing system. Filtered sequences were binned into operational taxonomic units with 97% identity and aligned to fully-sequenced microbial genomes (IMG/GG Greengenes) using the QIIME pipeline.[20] The current analysis was restricted to the 220 microbes (across taxonomic levels, including 91 Firmicutes, 33 Bacteroidetes, 45 Proteobacteria, 11 Actinobacteria, 5 Fusobacteria, and 35 in other phyla) that were detected in at least 13 (10%) of the subjects.

## Metabolomics analyses

A range of small molecules (most <1000 Daltons) was detected in the lyophilized fecal specimens by high-performance liquid phase chromatography and gas chromatography coupled with tandem mass spectrometry (HPLC-GC/MS-MS, Metabolon, Inc., North Carolina, USA) as described previously.[21, 22] Briefly, non-targeted single methanol extraction was performed, followed by protein precipitation. Individual molecules and their relative levels were identified from the mass spectral peaks compared to a chemical reference library generated from 2,500 standards, based on mass spectral peaks, retention times, and mass-to-charge ratios. The molecules include, but are not limited to, amino acids, carbohydrates, fatty acids, androgens, and xenobiotics. Volatile molecules, such as short chain fatty acids, may be lost during lyophilization or extraction. However, such loss is generally equivalent across specimens, and lyophilization is optimal for fecal specimens to assure equal loading of dry weight. The current analysis was restricted to the 530 metabolites that were detected in at least 118 (90%) of the subjects.

## Statistical analyses

The overall objective was to identify covariation and possible interactions between fecal metabolites and fecal microbes, either associated with CRC or not. For the CRC association, we used unconditional logistic regression to calculate the odds ratio (OR) and 95% confidence interval (CI), with case status as the dependent variable and with each CRC-associated microbe as the primary independent variable;[16] age, sex, and BMI were included for empiric adjustment of potential confounding. Including race in the models had no substantive impact on the estimates. To each microbe model, metabolites were added in a forward stepwise logistic regression, and metabolites associated with CRC at  $P \leq 0.15$  were retained. Change in OR with addition of metabolites was calculated as  $(OR_{no\ metabolites} - OR_{metab}) / (OR_{no\ metabolites} - 1)$ . For standardized estimates, relative abundance of the microbes and natural-log levels of the metabolites were normalized to mean 0 and standard error 1. We also report the Pearson correlation coefficients between the metabolites and microbes that were associated with CRC.

For the global objective irrespective of CRC, we considered all 530 metabolites and 220 microbes, and used linear regression, stratified by sex and case status, to identify associations between metabolites and microbes. For each of the 530 x 220 regressions, we adjusted for age, race (White vs Other), BMI, and hospital. We illustrate the overall extent of associations by plotting the  $-\log_{10}(P\text{-values})$  for each metabolite-microbe pairing in a “Manhattan” plot. We also calculated correlations of the top principal component (PC1) and principal coordinate (PCo1) of the metabolites and microbes, respectively. The principal components and principal coordinates were obtained from the residual matrix of linear regression models, adjusted for age, race, BMI, and hospital, for each metabolite or microbe, respectively. To compare the

correlation  $\rho_1$ , between PC1 and PCo1 in cases against the correlation  $\rho_2$  in controls, we applied Fisher's Z-transformation  $Z(\rho) = 0.5\ln((1+\rho)/(1-\rho))$  to each correlation and then tested whether  $Z(\rho_1) - Z(\rho_2)$  was significantly different from 0. We assumed  $Z(\rho_1) - Z(\rho_2)$  was normally distributed with mean 0 and variance  $1/(N_1-3) + 1/(N_2-3)$  under the null hypothesis where  $N_1$  and  $N_2$  are the number of cases and controls, respectively. The 2-sided P-value is  $2 * (1 - P_{\text{norm}}[(Z(\rho_1) - Z(\rho_2))/\text{sqrt}(\text{var})])$ . Statistical analyses were performed in R version 3.1.0 (R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/>).

## Results

Complete fecal microbiota and metabolome data were analyzed for 42 CRC cases and 89 age- and BMI-matched controls.[2] These 131 subjects had a mean age of 60 years (SD 13.2) and a mean BMI of 25.6 (SD 4.2); they were predominantly white and male (Table 1). Cases did not differ from controls on age, BMI, smoking or hospital, but a higher proportion of cases were African American and male (Table 1).

## Joint associations of fecal microbiota and fecal metabolites with CRC

In logistic regression models that included age, sex, and BMI, four microbes were significantly associated with CRC in separate models: Fusobacterium (OR 10.17, CI 2.95–35.0), Porphyromonas (OR 5.32, CI 1.76–16.05), Clostridia (OR 0.57, CI 0.38–0.85), and Lachnospiraceae (OR 0.61, CI 0.40–0.92). Table 2 presents these and the addition of fecal metabolites associated with

**Table 1. Selected characteristics of colorectal cancer cases and controls with fecal metabolite data.**

Characteristic	Colorectal cancer status	
	Cases (n = 42)	Controls (n = 89)
<b>Age</b> , mean (SD), years	63.4 (13.1)	58.4 (13.0)
<b>BMI</b> , mean (SD), kg/m <sup>2</sup>	25.8 (3.9)	25.5 (4.6)
<b>Sex</b>		
Male	59.5%	62.9%
Female	40.5%	37.1%
<b>Race</b>		
White	73.8%	85.4%
Black	23.8%	12.4%
Other	2.4%	2.2%
<b>Smoking</b>		
Never	54.8%	44.9%
Former	31.0%	46.1%
Current	11.9%	9.0%
Missing	2.4%	0.0
<b>Hospital</b>		
National Naval Medical Center	33.3%	41.6%
Walter Reed Army Medical Center	50.0%	37.1%
George Washington University Hospital	16.7%	21.3%
<b>Cancer stage</b>		
Non-invasive	21.4%	Not applicable
Invasive, no known metastases	42.9%	Not applicable
Metastatic	33.3%	Not applicable
Missing	2.4%	Not applicable

doi:10.1371/journal.pone.0152126.t001

**Table 2. Multivariable logistic regression models of fecal microbes and metabolites independently associated with colorectal cancer (CRC).\***

Microbe alone	Microbe plus metabolites	Odds Ratio	95% Confidence Limits	
g_Fusobacterium	-----	10.17	2.95	35.00
	g_Fusobacterium	7.53	1.40	40.45
	Palmitoyl_Sphingomyelin	2.56	1.32	4.96
	p_Hydroxybenzaldehyde	2.75	1.34	5.63
	Conjugated linoleic acid (CLA)	0.47	0.24	0.93
	p_Aminobenzoate (PABA)	0.58	0.29	1.16
	Alpha_Tocopherol	0.57	0.29	1.12
g_Porphyromonas	-----	5.32	1.76	16.05
	g_Porphyromonas	3.83	1.03	14.22
	Palmitoyl_Sphingomyelin	3.28	1.70	6.33
	p_Hydroxybenzaldehyde	2.34	1.17	4.69
	Conjugated linoleic acid (CLA)	0.51	0.27	0.98
	p_Aminobenzoate (PABA)	0.56	0.29	1.11
	Alpha_Tocopherol	0.61	0.31	1.20
c_Clostridia	-----	0.57	0.38	0.85
	c_Clostridia	0.58	0.32	1.06
	Palmitoyl_Sphingomyelin	2.77	1.48	5.17
	p_Hydroxybenzaldehyde	2.40	1.19	4.86
	Conjugated linoleic acid (CLA)	0.59	0.31	1.11
	p_Aminobenzoate (PABA)	0.39	0.20	0.77
	Mandelate	1.64	0.93	2.90
f_Lachnospiraceae	-----	0.61	0.40	0.92
	f_Lachnospiraceae	0.61	0.36	1.06
	Palmitoyl_Sphingomyelin	2.82	1.50	5.33
	p_Hydroxybenzaldehyde	2.24	1.09	4.61
	Conjugated linoleic acid (CLA)	0.56	0.30	1.05
	p_Aminobenzoate (PABA)	0.35	0.17	0.72
	Mandelate	1.60	0.92	2.78

\* Age, body mass index (BMI), sex, and one microbe were included in each model. Metabolites associated with colorectal cancer (CRC) at a criterion of  $P \leq 0.15$  were added to each microbe model. Microbe and metabolite levels were standardized to mean 0, standard error 1.

doi:10.1371/journal.pone.0152126.t002

CRC at a criterion of  $P \leq 0.15$ . In these models, the OR for CRC was approximately 2.8 with palmitoyl-sphingomyelin, 2.4 with *p*-hydroxy-benzaldehyde, 0.5 with *p*-aminobenzoate (PABA), and 0.5 with conjugated-linoleate-18-2N7 (CLA). Alpha tocopherol (OR 0.6) contributed to the Fusobacterium and Porphyromonas models, and mandelate (OR 1.6) contributed to the Clostridia and Lachnospiraceae models. With metabolites in the microbe models, the high OR of CRC with Fusobacterium was reduced by 29% (from 10.17 to 7.53), and the high OR with Porphyromonas was reduced by 34% (from 5.32 to 3.83). Attenuation of the low ORs with Clostridia and Lachnospiraceae was less marked (3.9% and 1.5%, respectively).

For further insight on the microbes and four metabolites that contributed to all of the logistic regression models, pairwise Pearson correlation coefficients were calculated by case-control status. In cases, strong correlations ( $|r| \geq 0.30$ ) were found for three metabolite pairs: direct for linoleate-PABA, inverse for benzaldehyde-sphingomyelin and benzaldehyde-CLA (Table 3). Cases also had strong correlations of sphingomyelin with microbes, which were inverse with Clostridia and Lachnospiraceae and direct with Fusobacterium. Also in cases, Fusobacterium was directly correlated with Porphyromonas. Controls had few strong correlations:

**Table 3. Pearson correlation coefficients for metabolites and microbes associated with colorectal cancer case status in multivariate analyses.\***

Metabolites and microbes	Correlation coefficients in cases							
	Sphingomyelin	Benzaldehyde	Linoleate	PABA	Clostridia	Lachnospiraceae	Fusobacterium	Porphyromonas
Sphingomyelin	1	<b>-0.36</b>	0.03	0.08	<b>-0.30</b>	<b>-0.33</b>	<b>0.44</b>	0.10
Benzaldehyde		1	<b>-0.45</b>	-0.14	0.14	0.09	-0.24	0.08
Linoleate			1	<b>0.44</b>	0.11	0.11	0.06	0.02
PABA				1	0.13	-0.11	-0.15	0.16
Clostridia					1	<b>0.70</b>	-0.20	0.17
Lachnospiraceae						1	-0.27	-0.09
Fusobacterium							1	<b>0.42</b>
Porphyromonas								1
Metabolites and microbes	Correlation coefficients in controls							
	Sphingomyelin	Benzaldehyde	Linoleate	PABA	Clostridia	Lachnospiraceae	Fusobacterium	Porphyromonas
Sphingomyelin	1	-0.02	0	-0.15	-0.12	-0.14	0.12	0.05
Benzaldehyde		1	-0.16	<b>0.30</b>	0.04	-0.08	-0.03	0.04
Linoleate			1	-0.02	-0.02	-0.08	0.05	-0.04
PABA				1	-0.12	<b>-0.36</b>	0.02	-0.07
Clostridia					1	<b>0.55</b>	0.07	-0.04
Lachnospiraceae						1	-0.16	-0.12
Fusobacterium							1	0.13
Porphyromonas								1

\* Relative abundance of class-level Clostridia and Lachnospiraceae; carriage (detection/non-detection) of genus-level Fusobacterium and Porphyromonas. PABA indicates *p*-aminobenzoate. **Bold italic**  $P < 0.05$ . *Italic*  $P < 0.10$ .

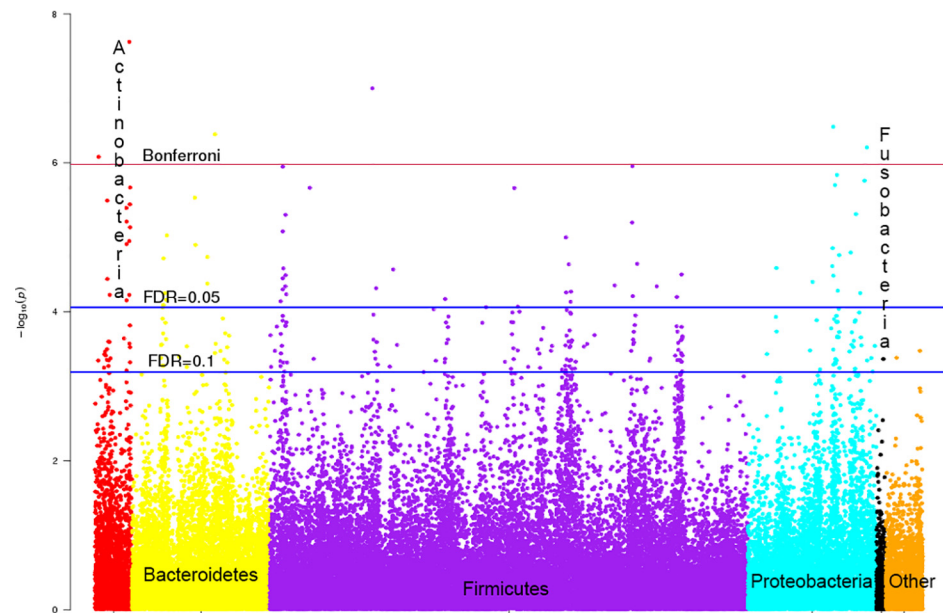
doi:10.1371/journal.pone.0152126.t003

benzaldehyde-PABA ( $\rho = 0.30$ ), Lachnospiraceae-PABA ( $\rho = -0.36$ ), and Lachnospiraceae-Clostridia ( $\rho = 0.55$ ). [S1 Table](#) presents, for cases and controls separately, the 20 metabolites that were most strongly correlated with each of the CRC-associated microbes.

## Associations of the fecal microbiota with fecal metabolites

To further explore the association between the microbiota and metabolites, we conducted the principal component/coordinate analysis of all 530 metabolites and 220 microbes. We found that the correlation between metabolite PC1 and microbial PCo1 was much stronger in CRC cases than in controls (Rho 0.606 vs 0.201,  $P = 0.01$ ). For an overall view, we used a 4-group meta-analysis (sex\*case/control), further adjusted for age, BMI, race, and enrollment hospital. [Fig 1](#) presents all 116,600 (530\*220) meta-analyzed P-values by microbial phylum. At the FDR = 0.1 threshold, there were 263 significant metabolite correlations, including 32 (12%) with Actinobacteria, 54 (20%) with Proteobacteria, 141 (54%) with Firmicutes, 33 (13%) with Bacteroidetes, 1 (0.3%) with Fusobacteria, and 2 (0.7%) with microbes in other phyla. At the FDR = 0.05 threshold, there were 72 significant metabolite correlations, including 14 (19%) with Actinobacteria, 15 (21%) with Proteobacteria, 31 (43%) with Firmicutes, 12 (17%) with Bacteroidetes, and none with Fusobacteria or microbes in other phyla. [S5 Table](#) presents exploratory associations of CRC with these 72 FDR = 0.05-significantly correlated microbe-metabolite pairs. In these 72 logistic regression models, CRC had a nominal direct association with 2-aminobutyrate (OR 1.60, CI 1.07–2.39) and *g\_Arcobacter* (phylum Proteobacteria, OR 1.94, CI 1.17–3.22), and it had a nominal inverse association with unknown metabolite X\_17626 (OR 0.59, CI 0.39–0.91) and *g\_Ruminococcus* (phylum Firmicutes, OR 0.59, CI 0.35–0.99).





**Fig 1. Observed inverse ( $-\log_{10}$ ) meta-analyzed P-values for associations between 530 fecal metabolites and 220 fecal microbes.** Microbes are color coded by phylum (Actinobacteria, red; Bacteroidetes, yellow; Firmicutes, purple; Proteobacteria, cyan; Fusobacteria, black; other phyla, orange) and sorted by genus. Bonferroni and false discovery rate (FDR) 0.05 and 0.1 threshold lines are presented.

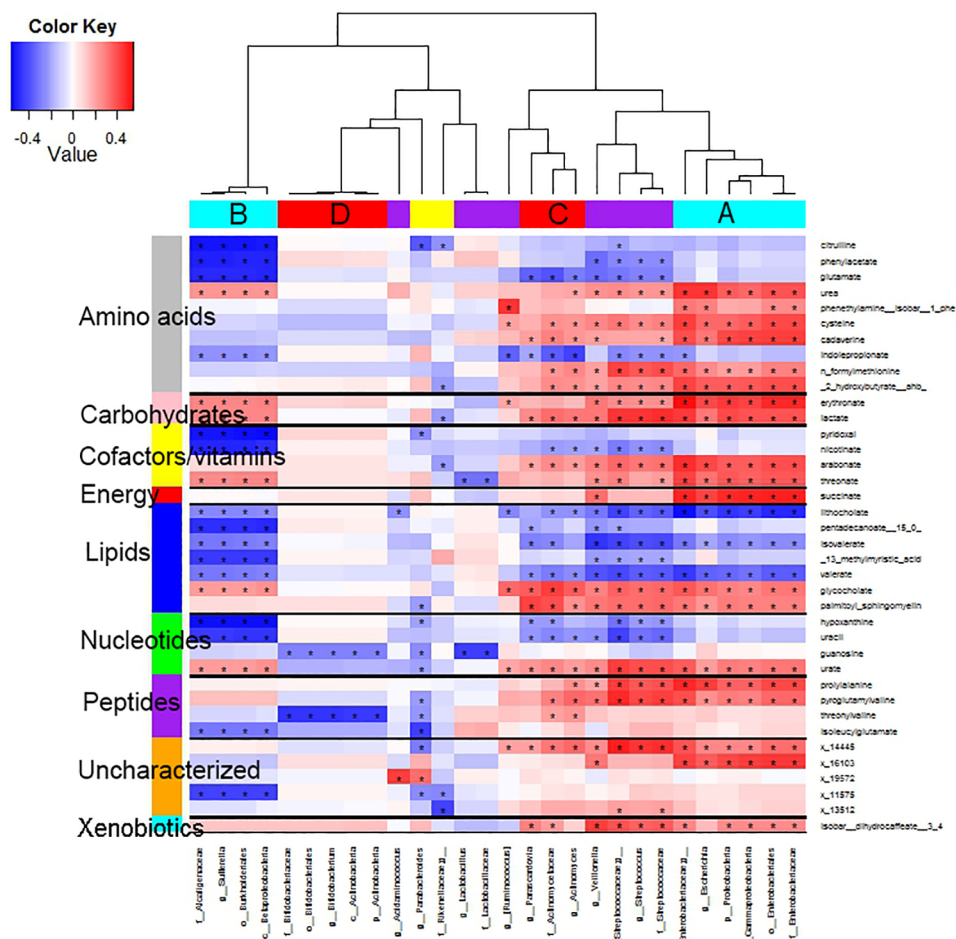
doi:10.1371/journal.pone.0152126.g001

The 4 Bonferroni-significant inverse correlations were indolepropionate with Actinomyces (Actinobacteria), threnoylvaline with Bifidobacterium (Actinobacteria), alanylalanine with Catabacteriaceae (Firmicutes), and 2-aminobutyrate with Butyricimonas (Bacteroidetes); the 2 Bonferroni-significant direct correlations were erythronate with Enterobacteriaceae (Proteobacteria) and an uncharacterized metabolite with Klebsiella (Proteobacteria).

Two clusters of Proteobacteria had distinct metabolite correlations (Fig 2, cyan bars A and B). Cluster A (Gammaproteobacteria, particularly Enterobacteriaceae) had inverse correlations with three lipids (lithocholate, isovalerate, and valerate), and this cluster had strong direct correlations with six amino acids, two carbohydrates (erythronate and lactate), two cofactors/vitamins (arabonate and threonate), one energy (succinate), two lipids (glycocholate and palmitoyl-sphingomyelin), one nucleotide (urate), two peptides, two uncharacterized metabolites, and a xenobiotic (dihydrocaffeate). Cluster B (Betaproteobacteria, particularly Alcaligenaceae) had direct correlations with the same carbohydrates, erythronate and lactate; but most of the other correlations differed for clusters A and B (Enterobacteriaceae and Alcaligenaceae). Cluster C comprised three Actinobacteria and five Firmicutes; and it differed from cluster A predominantly by inverse correlations with three amino acids, one cofactor/vitamin, and two nucleotides. Cluster D included five Actinobacteria (particularly Bifidobacteriaceae) that were inversely correlated with guanosine and threnoylvaline. Parabacteroides, the only microbe inversely correlated with palmitoyl-sphingomyelin, was also inversely correlated with three dipeptides and three nucleotides.

## Discussion

By comparing a comprehensive profile of the microbiota to a comprehensive panel of metabolites in the same specimens, the current study revealed microbe-metabolite correlations in human feces. It also revealed statistically independent microbe-metabolite differences between



**Fig 2. Heat map of the 100 strongest values of Pearson correlation coefficients of the residuals of all 530\*220 fecal metabolite-microbe pairs.** Asterisk (\*) indicates correlation significant at false discovery rate (FDR) 0.2. Bars at the top are color coded by phylum, as in Fig 1 (Proteobacteria, cyan; Actinobacteria, red; Firmicutes, purple; Bacteroidetes, yellow). Clusters are arbitrarily labeled A, B, C, and D. Bars on the left are color coded by metabolite pathway, as indicated.

doi:10.1371/journal.pone.0152126.g002

CRC cases and matched controls. These findings complement the metagenomic and animal-model studies that have identified characteristics of the distal human gut microbiota that are associated with CRC, inflammatory bowel disease, metabolic syndrome, obesity or malnutrition.[1–4, 11, 13, 14, 16] Overall, in 131 individuals we found 72 correlations between fecal metabolites and microbes that were significant at the FDR 0.05 level, of which six were significant at the Bonferroni level. The highly diverse Firmicutes phylum had 43% of the FDR-significant correlations, whereas the highly conserved Fusobacteria and other rare phyla had none. Microbe-metabolite correlations were significantly stronger in CRC cases than in controls. Directly comparing cases to controls, CRC was associated with significantly lower levels of Clostridia, Lachnospiraceae, PABA and CLA, and with higher levels of Fusobacterium, Porphyromonas, palmitoyl-sphingomyelin and *p*-hydroxy-benzaldehyde.

Our Bonferroni-significant microbe-metabolite pairs should be noted. Butyricimonas, a butyrate-producing genus in the family Porphyromonadaceae (Bacteroidetes), was inversely correlated with 2-aminobutyrate and apparently caused septic shock in a recently reported CRC patient.[23] Four other septic patients yielded the discovery of *Catabacter hongkongensis*, [24, 25] which is the sole member of the new Catabacteriaceae (Firmicutes) that we found to be



inversely correlated with a fecal dipeptide. Proteobacteria and Actinobacteria were correlated with several metabolites. Enterobacteriaceae (Gammaproteobacteria), which includes *Klebsiella*, *Escherichia*, *Shigella*, *Salmonella*, *Serratia*, and other pathogens, were directly correlated with an uncharacterized metabolite and with erythronate, a product of hyaluronic acid metabolism and oxidative stress.[26, 27] Of the Actinobacteria, some *Streptomyces* species produce a wide range of commonly used antimicrobial medications and other metabolites;[28] and *Actinomyces* had a Bonferroni-significant inverse association with indolepropionate in our study.

Of 11 fecal metabolites associated with CRC in univariate analysis,[17] only four were independently associated with the malignancy when adjusted for each other and for a CRC-associated microbe (Table 2). This reflects, at least in part, the correlations of several metabolites with each other (Table 3) and perhaps shared pathways.[17] Nonetheless, the CRC associations with these four metabolites (PABA, CLA, palmitoyl-sphingomyelin, and *p*-hydroxy-benzaldehyde) were only modestly attenuated when they were mutually adjusted for each other. Similarly, these metabolites minimally attenuated the CRC association with two low-risk microbes (*Clostridia* and *Lachnospiraceae*). In contrast, CRC association with the high-risk microbes (*Fusobacterium* and *Porphyrromonas*) was attenuated 40–53% by the metabolites, suggesting that these metabolites mediate, in part, the association of *Fusobacterium* and *Porphyrromonas* with CRC.

As reviewed elsewhere,[29] the microbiota produces thousands of chemically diverse molecules that potentially affect human health. How such microbial metabolites, including those in Table 2, affect or mark CRC risk is unknown. Possible mechanisms include shedding of cell membranes due to microbial invasion;[30–32] modulation of bacterial replication, inflammation, and cancer;[33–37] and synthesis of PABA and antibiotic precursors.[38–40]

This study had important limitations. First, the representativeness of the metabolites detected in our 20 year-old specimens is unknown, although they were stored in a lyophilized state at or below -40°C. Second, our study did not formally dissect the interactions of the highlighted metabolites and microbes. This might be accomplished by study of systematically constructed microbial consortia.[15] Third, while the microbe-metabolite correlations considered the multiplicity of comparisons, the associations with CRC did not. Despite this, both the microbe-metabolite and the CRC associations present hypotheses for independent or joint effects that can be examined in future studies. Fourth, we lacked an additional set of specimens for external validation. However, by focusing on a fixed set of the top metabolites, we obtained an estimate of the upper bound of the effect of the metabolites on each of the CRC-associated bacteria. Fourth, although our study considered 530 small molecules, it did not employ state-of-the-art holistic platforms that detect up to 10-fold more fecal metabolites,[41, 42] nor did it specifically probe immunologic and inflammatory pathways that are centrally involved in CRC pathogenesis.[10, 43, 44] Finally, we have not identified functions of the fecal microbes that we detected. Previously, we noted that the activities of two important enzymes in feces,  $\beta$ -glucuronidase and  $\beta$ -glucosidase, were directly correlated with microbiota alpha diversity and abundance of *Clostridia*, and inversely correlated with abundances of *Streptococcus* and *Alistipes*. [45] Others have shown that the microbiota of specific pathogen-free mice can generate anti-inflammatory regulatory T cells, which moderate systemic immunity, through the production of butyrate.[46]

In summary, this study uncovered a complex network of microbes and molecules in human feces. In this network, CRC cases had strong microbe-metabolite correlations that were predominated by Proteobacteria and Actinobacteria. To obtain insights on disease and to identify targets for intervention, functional studies will be needed. Ultimately, innovative prospective human studies, including clinical trials, will be required.[15, 47]

## Supporting Information

**S1 Table. Metabolites most strongly correlated with colorectal cancer-associated taxa, by case-control status.**

(XLSX)

**S2 Table. Standardized levels of all metabolites among study participants.**

(CSV)

**S3 Table. Relative abundance of all microbial taxa among study participants.**

(CSV)

**S4 Table. Colorectal cancer (CRC) case-control status, covariates, and logistic regression beta values for associations of candidate metabolites and microbial taxa with CRC.**

(CSV)

**S5 Table. Logistic regression association with colorectal cancer (CRC) for the 72 metabolite-microbial taxon pairs that were correlated with each other at false discovery rate (FDR) 0.05.**

(XLSX)

## Acknowledgments

Funding. Supported by the National Cancer Institute Intramural Research Program and grants R03CA159414 and R01CA159036.

## Author Contributions

Conceived and designed the experiments: RS JA JNS GY RBH JJG. Performed the experiments: GY XX. Analyzed the data: RS JA JNS GY XX RBH JJG. Contributed reagents/materials/analysis tools: RS JNS. Wrote the paper: RS JA JNS JS GY XX RBH JJG.

## References

1. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, et al. Richness of human gut microbiome correlates with metabolic markers. *Nature*. 2013; 500(7464):541–6. Epub 2013/08/30. doi: [10.1038/nature12506](https://doi.org/10.1038/nature12506) PMID: [23985870](https://pubmed.ncbi.nlm.nih.gov/23985870/).
2. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009; 457(7228):480–4. Epub 2008/12/02. doi: [10.1038/nature07540](https://doi.org/10.1038/nature07540) PMID: [19043404](https://pubmed.ncbi.nlm.nih.gov/19043404/); PubMed Central PMCID: PMC2677729.
3. Smith MI, Yatsunenko T, Manary MJ, Trehan I, Mkakosya R, Cheng J, et al. Gut Microbiomes of Malawian Twin Pairs Discordant for Kwashiorkor. *Science*. 2013. Epub 2013/02/01. doi: [10.1126/science.1229000](https://doi.org/10.1126/science.1229000) PMID: [23363771](https://pubmed.ncbi.nlm.nih.gov/23363771/).
4. McHardy IH, Goudarzi M, Tong M, Ruegger PM, Schwager E, Weger JR, et al. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome*. 2013; 1(1):17. doi: [10.1186/2049-2618-1-17](https://doi.org/10.1186/2049-2618-1-17) PMID: [24450808](https://pubmed.ncbi.nlm.nih.gov/24450808/); PubMed Central PMCID: PMC3971612.
5. Tjalsma H, Boleij A, Marchesi JR, Dutilh BE. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nature reviews Microbiology*. 2012; 10(8):575–82. Epub 2012/06/26. doi: [10.1038/nrmicro2819](https://doi.org/10.1038/nrmicro2819) PMID: [22728587](https://pubmed.ncbi.nlm.nih.gov/22728587/).
6. Scanlan PD, Shanahan F, Clune Y, Collins JK, O'Sullivan GC, O'Riordan M, et al. Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. *Environmental microbiology*. 2008; 10(3):789–98. doi: [10.1111/j.1462-2920.2007.01503.x](https://doi.org/10.1111/j.1462-2920.2007.01503.x) PMID: [18237311](https://pubmed.ncbi.nlm.nih.gov/18237311/).
7. Jobin C. Colorectal cancer: looking for answers in the microbiota. *Cancer discovery*. 2013; 3(4):384–7. doi: [10.1158/2159-8290.CD-13-0042](https://doi.org/10.1158/2159-8290.CD-13-0042) PMID: [23580283](https://pubmed.ncbi.nlm.nih.gov/23580283/); PubMed Central PMCID: PMC3625977.
8. Uronis JM, Muhlbauer M, Herfarth HH, Rubinas TC, Jones GS, Jobin C. Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PloS one*. 2009; 4(6):e6026. doi: [10.1371/journal.pone.0006026](https://doi.org/10.1371/journal.pone.0006026) PMID: [19551144](https://pubmed.ncbi.nlm.nih.gov/19551144/); PubMed Central PMCID: PMC2696084.

9. Arthur JC, Perez-Chanona E, Muhlbauer M, Tomkovich S, Uronis JM, Fan TJ, et al. Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. *Science*. 2012. Epub 2012/08/21. doi: [10.1126/science.1224820](https://doi.org/10.1126/science.1224820) PMID: [22903521](https://pubmed.ncbi.nlm.nih.gov/22903521/).
10. Allen-Vercoe E, Jobin C. *Fusobacterium* and *Enterobacteriaceae*: important players for CRC? *Immunology letters*. 2014; 162(2 Pt A):54–61. doi: [10.1016/j.imlet.2014.05.014](https://doi.org/10.1016/j.imlet.2014.05.014) PMID: [24972311](https://pubmed.ncbi.nlm.nih.gov/24972311/); PubMed Central PMCID: PMC4259803.
11. Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Molecular systems biology*. 2014; 10:766. Epub 2014/11/30. doi: [10.15252/msb.20145645](https://doi.org/10.15252/msb.20145645) PMID: [25432777](https://pubmed.ncbi.nlm.nih.gov/25432777/).
12. Turnbaugh PJ, Quince C, Faith JJ, McHardy AC, Yatsunenkov T, Niazi F, et al. Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(16):7503–8. Epub 2010/04/07. doi: [10.1073/pnas.1002355107](https://doi.org/10.1073/pnas.1002355107) PMID: [20363958](https://pubmed.ncbi.nlm.nih.gov/20363958/); PubMed Central PMCID: PMC2867707.
13. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science translational medicine*. 2009; 1(6):6ra14. Epub 2010/04/07. doi: [10.1126/scitranslmed.3000322](https://doi.org/10.1126/scitranslmed.3000322) PMID: [20368178](https://pubmed.ncbi.nlm.nih.gov/20368178/); PubMed Central PMCID: PMC2894525.
14. Feng Q, Liang S, Jia H, Stadlmayr A, Tang L, Lan Z, et al. Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nature communications*. 2015; 6:6528. doi: [10.1038/ncomms7528](https://doi.org/10.1038/ncomms7528) PMID: [25758642](https://pubmed.ncbi.nlm.nih.gov/25758642/).
15. Yen S, McDonald JA, Schroeter K, Oliphant K, Sokolenko S, Blondeel EJ, et al. Metabolomic analysis of human fecal microbiota: a comparison of feces-derived communities and defined mixed communities. *Journal of proteome research*. 2015; 14(3):1472–82. doi: [10.1021/pr5011247](https://doi.org/10.1021/pr5011247) PMID: [25670064](https://pubmed.ncbi.nlm.nih.gov/25670064/).
16. Ahn J, Sinha R, Pei Z, Dominianni C, Wu J, Shi J, et al. Human gut microbiome and risk for colorectal cancer. *Journal of the National Cancer Institute*. 2013; 105(24):1907–11. Epub 2013/12/10. doi: [10.1093/jnci/djt300](https://doi.org/10.1093/jnci/djt300) PMID: [24316595](https://pubmed.ncbi.nlm.nih.gov/24316595/); PubMed Central PMCID: PMC3866154.
17. Goedert JJ, Sampson JN, Moore SC, Xiao Q, Xiong X, Hayes RB, et al. Fecal metabolomics: assay performance and association with colorectal cancer. *Carcinogenesis*. 2014; 35(9):2089–96. Epub 2014/07/20. doi: [10.1093/carcin/bgu131](https://doi.org/10.1093/carcin/bgu131) PMID: [25037050](https://pubmed.ncbi.nlm.nih.gov/25037050/); PubMed Central PMCID: PMC4146421.
18. Schiffman MH, Andrews AW, Van Tassel RL, Smith L, Daniel J, Robinson A, et al. Case-control study of colorectal cancer and fecal mutagenicity. *Cancer research*. 1989; 49(12):3420–4. Epub 1989/06/15. PMID: [2655896](https://pubmed.ncbi.nlm.nih.gov/2655896/).
19. Schiffman MH, Van Tassel RL, Robinson A, Smith L, Daniel J, Hoover RN, et al. Case-control study of colorectal cancer and fecapentaene excretion. *Cancer research*. 1989; 49(5):1322–6. Epub 1989/03/01. PMID: [2917361](https://pubmed.ncbi.nlm.nih.gov/2917361/).
20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nature methods*. 2010; 7(5):335–6. Epub 2010/04/13. doi: [10.1038/nmeth.f.303](https://doi.org/10.1038/nmeth.f.303) PMID: [20383131](https://pubmed.ncbi.nlm.nih.gov/20383131/); PubMed Central PMCID: PMC3156573.
21. Evans AM, DeHaven CD, Barrett T, Mitchell M, Milgram E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Analytical chemistry*. 2009; 81(16):6656–67. doi: [10.1021/ac901536h](https://doi.org/10.1021/ac901536h) PMID: [19624122](https://pubmed.ncbi.nlm.nih.gov/19624122/).
22. Sha W, da Costa KA, Fischer LM, Milburn MV, Lawton KA, Berger A, et al. Metabolomic profiling can predict which humans will develop liver dysfunction when deprived of dietary choline. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*. 2010; 24(8):2962–75. doi: [10.1096/fj.09-154054](https://doi.org/10.1096/fj.09-154054) PMID: [20371621](https://pubmed.ncbi.nlm.nih.gov/20371621/); PubMed Central PMCID: PMC2909293.
23. Toprak Ulger N, Bozan T, Birkan Y, Isbir S, Soyletir G. *Butyrivibrio* virosa: the first clinical case of bacteraemia. *New Microbes New Infections*. 2015; 4:7–8. doi: [10.1016/j.nmni.2014.12.004](https://doi.org/10.1016/j.nmni.2014.12.004) PMID: [25830028](https://pubmed.ncbi.nlm.nih.gov/25830028/)
24. Lau SK, McNabb A, Woo GK, Hoang L, Fung AM, Chung LM, et al. *Catabacter hongkongensis* gen. nov., sp. nov., isolated from blood cultures of patients from Hong Kong and Canada. *Journal of clinical microbiology*. 2007; 45(2):395–401. doi: [10.1128/JCM.01831-06](https://doi.org/10.1128/JCM.01831-06) PMID: [17122022](https://pubmed.ncbi.nlm.nih.gov/17122022/); PubMed Central PMCID: PMC1829005.
25. Codony F, Adrados B, Perez LM, Fittipaldi M, Morato J. Detection of *Catabacter hongkongensis* in polluted European water samples. *Journal of Zhejiang University Science B*. 2009; 10(12):867–9. doi: [10.1631/jzus.B0920218](https://doi.org/10.1631/jzus.B0920218) PMID: [19946949](https://pubmed.ncbi.nlm.nih.gov/19946949/); PubMed Central PMCID: PMC2789520.
26. Jahn M, Baynes JW, Spiteller G. The reaction of hyaluronic acid and its monomers, glucuronic acid and N-acetylglucosamine, with reactive oxygen species. *Carbohydrate research*. 1999; 321(3–4):228–34. PMID: [10614067](https://pubmed.ncbi.nlm.nih.gov/10614067/).

27. Ellis JK, Athersuch TJ, Thomas LD, Teichert F, Perez-Trujillo M, Svendsen C, et al. Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population. *BMC medicine*. 2012; 10:61. doi: [10.1186/1741-7015-10-61](https://doi.org/10.1186/1741-7015-10-61) PMID: [22713677](https://pubmed.ncbi.nlm.nih.gov/22713677/); PubMed Central PMCID: PMC3391181.
28. Donadio S, Sosio M, Lancini G. Impact of the first *Streptomyces* genome sequence on the discovery and production of bioactive substances. *Applied microbiology and biotechnology*. 2002; 60(4):377–80. doi: [10.1007/s00253-002-1143-0](https://doi.org/10.1007/s00253-002-1143-0) PMID: [12466876](https://pubmed.ncbi.nlm.nih.gov/12466876/).
29. Donia MS, Fischbach MA. HUMAN MICROBIOTA. Small molecules from the human microbiota. *Science*. 2015; 349(6246):1254766. doi: [10.1126/science.1254766](https://doi.org/10.1126/science.1254766) PMID: [26206939](https://pubmed.ncbi.nlm.nih.gov/26206939/).
30. Linardic CM, Hannun YA. Identification of a distinct pool of sphingomyelin involved in the sphingomyelin cycle. *The Journal of biological chemistry*. 1994; 269(38):23530–7. PMID: [8089120](https://pubmed.ncbi.nlm.nih.gov/8089120/).
31. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome research*. 2012; 22(2):292–8. Epub 2011/10/20. doi: [10.1101/gr.126573.111](https://doi.org/10.1101/gr.126573.111) PMID: [22009990](https://pubmed.ncbi.nlm.nih.gov/22009990/); PubMed Central PMCID: PMC3266036.
32. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome research*. 2012; 22(2):299–306. Epub 2011/10/20. doi: [10.1101/gr.126516.111](https://doi.org/10.1101/gr.126516.111) PMID: [22009989](https://pubmed.ncbi.nlm.nih.gov/22009989/); PubMed Central PMCID: PMC3266037.
33. Storey A, Rogers JS, McArdle F, Jackson MJ, Rhodes LE. Conjugated linoleic acids modulate UVR-induced IL-8 and PGE2 in human skin cells: potential of CLA isomers in nutritional photoprotection. *Carcinogenesis*. 2007; 28(6):1329–33. Epub 2007/03/29. doi: [10.1093/carcin/bgm065](https://doi.org/10.1093/carcin/bgm065) PMID: [17389616](https://pubmed.ncbi.nlm.nih.gov/17389616/).
34. Zulet MA, Marti A, Parra MD, Martinez JA. Inflammation and conjugated linoleic acid: mechanisms of action and implications for human health. *Journal of physiology and biochemistry*. 2005; 61(3):483–94. Epub 2006/01/31. PMID: [16440602](https://pubmed.ncbi.nlm.nih.gov/16440602/).
35. Belury MA. Conjugated dienoic linoleate: a polyunsaturated fatty acid with unique chemoprotective properties. *Nutrition reviews*. 1995; 53(4 Pt 1):83–9. Epub 1995/04/01. PMID: [7624062](https://pubmed.ncbi.nlm.nih.gov/7624062/).
36. Huang CB, George B, Ebersole JL. Antimicrobial activity of n-6, n-7 and n-9 fatty acids and their esters for oral microorganisms. *Archives of oral biology*. 2010; 55(8):555–60. doi: [10.1016/j.archoralbio.2010.05.009](https://doi.org/10.1016/j.archoralbio.2010.05.009) PMID: [20541177](https://pubmed.ncbi.nlm.nih.gov/20541177/); PubMed Central PMCID: PMC2902640.
37. Gholami Z, Khosravi-Darani K. An overview of conjugated linoleic acid: microbial production and application. *Mini reviews in medicinal chemistry*. 2014; 14(9):734–46. PMID: [25138090](https://pubmed.ncbi.nlm.nih.gov/25138090/).
38. Ottesen AR, Gonzalez Pena A, White JR, Pettengill JB, Li C, Allard S, et al. Baseline survey of the anatomical microbial ecology of an important food plant: *Solanum lycopersicum* (tomato). *BMC microbiology*. 2013; 13:114. doi: [10.1186/1471-2180-13-114](https://doi.org/10.1186/1471-2180-13-114) PMID: [23705801](https://pubmed.ncbi.nlm.nih.gov/23705801/); PubMed Central PMCID: PMC3680157.
39. Rastogi G, Sbodio A, Tech JJ, Suslow TV, Coaker GL, Leveau JH. Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *The ISME journal*. 2012; 6(10):1812–22. doi: [10.1038/ismej.2012.32](https://doi.org/10.1038/ismej.2012.32) PMID: [22534606](https://pubmed.ncbi.nlm.nih.gov/22534606/); PubMed Central PMCID: PMC3446804.
40. Cociancich S, Pesic A, Petras D, Uhlmann S, Kretz J, Schubert V, et al. The gyrase inhibitor albicidin consists of p-aminobenzoic acids and cyanoalanine. *Nature chemical biology*. 2015; 11(3):195–7. doi: [10.1038/nchembio.1734](https://doi.org/10.1038/nchembio.1734) PMID: [25599532](https://pubmed.ncbi.nlm.nih.gov/25599532/).
41. Vanden Bussche J, Marzorati M, Laukens D, Vanhaecke L. Validated High Resolution Mass Spectrometry-Based Approach for Metabolomic Fingerprinting of the Human Gut Phenotype. *Analytical chemistry*. 2015; 87(21):10927–34. doi: [10.1021/acs.analchem.5b02688](https://doi.org/10.1021/acs.analchem.5b02688) PMID: [26451617](https://pubmed.ncbi.nlm.nih.gov/26451617/).
42. Xu W, Chen D, Wang N, Zhang T, Zhou R, Huan T, et al. Development of high-performance chemical isotope labeling LC-MS for profiling the human fecal metabolome. *Analytical chemistry*. 2015; 87(2):829–36. doi: [10.1021/ac503619q](https://doi.org/10.1021/ac503619q) PMID: [25486321](https://pubmed.ncbi.nlm.nih.gov/25486321/).
43. Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, Rossetti BJ, et al. Microbiota organization is a distinct feature of proximal colorectal cancers. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(51):18321–6. doi: [10.1073/pnas.1406199111](https://doi.org/10.1073/pnas.1406199111) PMID: [25489084](https://pubmed.ncbi.nlm.nih.gov/25489084/); PubMed Central PMCID: PMC4280621.
44. Brown EM, Sadarangani M, Finlay BB. The role of the immune system in governing host-microbe interactions in the intestine. *Nature immunology*. 2013; 14(7):660–7. doi: [10.1038/ni.2611](https://doi.org/10.1038/ni.2611) PMID: [23778793](https://pubmed.ncbi.nlm.nih.gov/23778793/).
45. Flores R, Shi J, Gail MH, Gajer P, Ravel J, Goedert JJ. Association of fecal microbial diversity and taxonomy with selected enzymatic functions. *PloS one*. 2012; 7(6):e39745. Epub 2012/07/05. doi: [10.1371/journal.pone.0039745](https://doi.org/10.1371/journal.pone.0039745) PMID: [22761886](https://pubmed.ncbi.nlm.nih.gov/22761886/); PubMed Central PMCID: PMC3386201.

46. Arpaia N, Campbell C, Fan X, Dikiy S, van der Veeken J, deRoos P, et al. Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature*. 2013; 504(7480):451–5. doi: [10.1038/nature12726](https://doi.org/10.1038/nature12726) PMID: [24226773](https://pubmed.ncbi.nlm.nih.gov/24226773/); PubMed Central PMCID: PMC3869884.
47. Mai V, Morris JG Jr. Need for prospective cohort studies to establish human gut microbiome contributions to disease risk. *Journal of the National Cancer Institute*. 2013; 105(24):1850–1. doi: [10.1093/jnci/djt349](https://doi.org/10.1093/jnci/djt349) PMID: [24316594](https://pubmed.ncbi.nlm.nih.gov/24316594/).