A prospective cohort analysis of gut microbial co-metabolism in Alaska Native and rural African people at high and low risk of colorectal cancer

Soeren Ocvirk,[1,2] Annette S Wilson,[1] Joram M Posma,[3] Jia V Li,[4,5] Kathryn R Koller,[6] Gretchen M Day,[6] Christie A Flanagan,[6] Jill Evon Otto,[6] Pam E Sacco,[6] Frank D Sacco,[6] Flora R Sapp,[6] Amy S Wilson,[6] Keith Newton,[7] Faye Brouard,[8] James P DeLany,[9,10] Marissa Behnning,[1] Corynn N Appolonia,[1] Devavrata Soni,[1] Faheem Bhatti,[1] Barbara Methé,[11] Adam Fitch,[11] Alison Morris,[11] H Rex Gaskins,[12] James Kinross,[13] Jeremy K Nicholson,[13,5] Timothy K Thomas,[6] and Stephen JD O'Keefe[1]

[1]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [2]Department of Gastrointestinal Microbiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; [3]Section of Bioinformatics, Division of Systems Medicine, Department of Metabolism, Digestion and Reproduction, Imperial College, London, United Kingdom; [4]Section of Nutritional Research, Division of Digestive Diseases, Department of Metabolism, Digestion, and Reproduction, Imperial College, London, United Kingdom; [5]Centre for Digestive and Gut Health, Institution of Global Health Innovation, Imperial College, London, United Kingdom; [6]Clinical & Research Services, Community Health Services, Alaska Native Tribal Health Consortium, Anchorage, AK, USA; [7]Division of Gastroenterology, University of KwaZulu-Natal, Durban, South Africa; [8]Manguzi Hospital, Manguzi, KwaZulu-Natal, South Africa; [9]Division of Endocrinology, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [10]Translational Research Institute for Metabolism and Diabetes, Advent Health, Orlando, FL, USA; [11]Center for Medicine and the Microbiome, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [12]Division of Nutritional Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA; and [13]Department of Surgery and Cancer, Faculty of Medicine, Imperial College, London, United Kingdom

ABSTRACT

Background: Alaska Native (AN) people have the world's highest recorded incidence of sporadic colorectal cancer (CRC) (~91:100,000), whereas rural African (RA) people have the lowest risk (<5:100,000). Previous data supported the hypothesis that diet affected CRC risk through its effects on the colonic microbiota that produce tumor-suppressive or -promoting metabolites.

Objectives: We investigated whether differences in these metabolites may contribute to the high risk of CRC in AN people.

Methods: A cross-sectional observational study assessed dietary intake from 32 AN and 21 RA healthy middle-aged volunteers before screening colonoscopy. Analysis of fecal microbiota composition by 16S ribosomal RNA gene sequencing and fecal/urinary metabolites by ^{1}H-NMR spectroscopy was complemented with targeted quantification of fecal SCFAs, bile acids, and functional microbial genes.

Results: Adenomatous polyps were detected in 16 of 32 AN participants, but not found in RA participants. The AN diet contained higher proportions of fat and animal protein and less fiber. AN fecal microbiota showed a compositional predominance of *Blautia* and *Lachnoclostridium*, higher microbial capacity for bile acid conversion, and low abundance of some species involved in saccharolytic fermentation (e.g., *Prevotellaceae*, *Ruminococcaceae*), but no significant lack of butyrogenic bacteria. Significantly lower concentrations of tumor-suppressive butyrate (22.5 ± 3.1 compared with 47.2 ± 7.3 SEM μmol/g) coincided with significantly higher concentrations of tumor-promoting deoxycholic acid (26.7 ± 4.2 compared with 11 ± 1.9 μmol/g) in AN fecal samples. AN participants had lower quantities of fecal/urinary metabolites than RA participants and metabolite profiles correlated with the abundance of distinct microbial genera in feces. The main microbial and metabolic CRC-associated markers were not significantly altered in AN participants with adenomatous polyps.

Conclusions: The low-fiber, high-fat diet of AN people and exposure to carcinogens derived from diet or environment are associated

Supported by NIH grants R01 CA135379 (to SJDOK) and R01 CA204403 (to SJDOK). This work was performed with infrastructure support provided by the National Institute for Health Research Imperial Biomedical Research Centre. SO is supported by German Cancer Aid (Deutsche Krebshilfe) Mildred-Scheel-Scholarship 70112121 and by German Research Foundation (Deutsche Forschungsgemeinschaft) grant 338582098. JMP is supported by Health Data Research UK Rutherford Fund Fellowship MR/S004033/1. JVL is supported by Medical Research Council New Investigator Grant MR/L009803/1 and European Research Council Starting Grant 715662. SJDOK is supported by NIH grant R01 CA204403.

Supplemental Methods, Supplemental Table 1, and Supplemental Figures 1–5 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Data described in the article will be made available upon request pending (e.g., application and approval).

Address correspondence to SJDOK (e-mail: sjokeefe@pitt.edu).

Abbreviations used: AN, Alaska Native; CA, cholic acid; CRC, colorectal cancer; DCA, deoxycholic acid; OTU, operational taxonomic unit; RA, rural African; rRNA, ribosomal RNA.

Received July 16, 2019. Accepted for publication November 14, 2019.

First published online December 18, 2019; doi: https://doi.org/10.1093/ajcn/nqz301.

with a tumor-promoting colonic milieu as reflected by the high rates of adenomatous polyps in AN participants. *Am J Clin Nutr* 2020;111:406–419.

Keywords: colorectal cancer, gut microbiota, dietary fiber, short-chain fatty acids, butyrate, bile acids, deoxycholic acid, Alaska Native people, rural African people

Introduction

Colorectal cancer (CRC) is the fourth leading cause of cancer deaths worldwide and shows rising incidence in most countries (1, 2). The majority of sporadic CRC cases results from accumulated genetic alterations in response to detrimental environmental factors, in particular diet (3, 4). Meta-analyses have provided convincing evidence for red and processed meat (involving high amounts of fat) increasing and fiber reducing CRC risk (5, 6). Dietary fiber is indigestible for the host and fermented by the gut microbiota to SCFAs, particularly butyrate, that demonstrate potent antiproliferative and antineoplastic effects in the colon (7). A high consumption of fat stimulates hepatic synthesis of bile acids and their delivery to the colon, where they undergo complex biotransformation performed by gut bacteria. This results in high concentrations of secondary bile acids such as deoxycholic acid (DCA), which shows experimental tumorigenic activity (8). A series of studies by our group has demonstrated that healthy rural African (RA) people, who rarely get CRC (<5:100,000) and consume a diet rich in fiber and low in fat, show significantly higher concentrations of SCFAs and lower concentrations of bile acids in feces than African Americans, who consume a low-fiber, high-fat diet and have the highest CRC incidence in the contiguous United States (49:100,000) (9). A diet switch among volunteers from these 2 populations led to reciprocal changes in fecal SCFAs, bile acids, and related functional microbial genes of the gut microbiota (10). Critically, this correlated with changes in mucosal markers of CRC risk (10).

Although CRC rates are high in industrialized countries with "Western diets," the world's highest risk of CRC was paradoxically revealed in a population that still pursues a traditional lifestyle in remote communities: Alaska Native (AN) people have the world's highest reported risk of CRC (~91:100,000), but their frequency of genetic mutations predisposing for CRC seems to be similar to other populations in the United States (11–14). We aimed to investigate how the diet of AN people affects the composition and function of the gut microbiota and related metabolic markers that are associated with increased CRC risk. Because the diet of AN people is rich in fat and low in fiber, we evaluated our results by comparison with RA people from South Africa, who have a low-fat, high-fiber diet.

Methods

Study population and sample collection

Healthy AN people and healthy RA individuals were recruited for this study. Inclusion criteria were AN or RA heritage, respectively, and regional resident for at least the past 5 y. Exclusion criteria were history of cancer, medical treatment of bowel conditions (e.g., ulcers, diarrhea), history of bowel disease (e.g., Crohn's disease, ulcerative colitis), previous bowel surgery resulting in chronic diarrhea, diagnosis of HIV, antibiotics within the past 12 wk, taking insulin, taking pre-/probiotics, or taking steroids (**Supplemental Figure 1**). Informed consent was acquired before enrollment of study participants. A colonoscopy was performed for every study participant, where the number of polyps was quantified and polyps were removed. Research nurses collected participant diet data using an annual FFQ containing traditional AN or RA foods, respectively (15). The average dietary intake of all macro- and micronutrients was calculated using values normalized to a 1000-kcal intake. Study participants were asked to collect a sample from the first stool and urine passed in the morning into sterile plastic containers. Samples were stored at 4°C, collected, and transported on ice to be stored frozen at −70°C. The Alaska Area Institution Review Board (#2012-07-029), University of Pittsburgh Institution Review Board (PRO 08100243), and the University of KwaZulu-Natal Medical Ethics and Research Committee (REF: BE006/01) provided ethical review and approval for the study. Tribal approval was received from Alaska Native Tribal Health Consortium, Southcentral Foundation, Artic Slope Native Association, and Samuel Simmonds Memorial Hospital Administration.

16S ribosomal RNA gene sequencing and taxonomic analysis

DNA was extracted from frozen fecal samples using a DNA Stool Mini Kit (Qiagen) according to the manufacturer's recommendations. A bead-beating step was included to increase the yields of DNA extracted from gram-positive bacteria. Extracted DNA was PCR-amplified using the method/primers of Caporaso et al. (16) and Q5 HS High-Fidelity polymerase (NEB). See the **Supplemental Methods** for details on sample processing and taxonomic and statistical analysis.

^{1}H-NMR spectroscopy

Urine and fecal water samples were analyzed using 600 MHz ^{1}H-NMR spectroscopy according to established protocols (17, 18). Spectra were referenced to an internal standard, baseline- and phase-corrected using in-house software, and imported into MATLAB (R2014a; MathWorks). Urine spectra were digitized to the chemical shift range δ 0.3–10, excluding the water resonance at δ 4.71–4.91 and urea at δ 5.48–6.10. Fecal water spectra were digitized to the chemical shift range δ 0.3–10, excluding the water resonance at δ 4.62–5.18 and polyethylene glycol signals at δ 3.63–3.81. Metabolites were identified using different statistical and analytical techniques (19). See the Supplemental Methods for details on data analysis.

Quantification of SCFAs

Fecal SCFAs were analyzed by GC as described previously (9).

TABLE 1 Demographic data of study participants

	Total	Female	Male	Geographic area (*n*)	Age, y[1]	Polyps detected[2]
Alaska Native	32	24	8	Anchorage (22), Utqiagvik (10)	51.0 ± 8.9	16 of 32
Rural African	21	12	9	Rural KwaZulu-Natal (21)	53.3 ± 11.5	0 of 21

[1] Values are mean ± SEM.
[2] Polyps detected (participants with polyps out of total number of participants) and removed during colonoscopy.

Quantification of bile acids

Fecal bile acids were quantified by HPLC coupled with MS using electrospray ionization in negative ion mode by monitoring the $(M\text{-}H)^-$ ion as described earlier (9).

Analysis of functional microbial genes

As described before (9, 10), we targeted functional microbial genes involved in butyrate synthesis, bile acid conversion, sulfite reduction, and methane production (**Supplemental Table 1**). See the Supplemental Methods for details.

Statistical analysis

The predeclared primary outcome variables were fecal concentrations of tumor-suppressive butyrate and tumor-promoting DCA (9, 10). Secondary outcome variables were fecal SCFAs and bile acids, composition and metabolome of fecal microbiota, functional microbial genes involved in butyrogenesis, and bile acid conversion (9, 10). In this context, CRC risk was defined as incidence of CRC previously identified for the respective ethnic group (7, 14). Based on our previous studies (9, 10), we expected a sample size of 20 participants per group to be adequate to detect differences in primary outcome measures (with an α set to 0.05 and β set to 0.1), especially considering the great disparity in CRC incidence of the 2 populations. Statistical analysis of data sets with 1 variable was performed by an unpaired *t* test (data normally distributed) or a nonparametric Mann–Whitney *U* test (data not normally distributed) using Prism 7.0 software (GraphPad Software Inc.). Data for SCFAs, bile acids, functional microbial genes, and relative taxonomic abundances are shown as box-and-whisker plots according to Tukey unless stated otherwise. 16S ribosomal RNA (rRNA) gene sequencing data were analyzed using the Mann–Whitney *U* test and corrected for multiple testing by the Benjamini–Hochberg method using the R pipeline Rhea (20). ^{1}H-NMR data were analyzed using Monte-Carlo cross-validated projections to latent structures (21) and metabolic reactions networks using MetaboNetworks (22). See the Supplemental Methods for details.

Results

Diets of AN and RA participants are associated with diverse compositions of the gut microbiota

We recruited a total of 32 AN participants and 21 healthy RA individuals (**Table 1**). Among the AN participants, the majority ($n = 22$) were from the largest concentration of AN people, who live in the city of Anchorage (Alaska, United States), whereas the remaining volunteers ($n = 10$) were from Utqiagvik (formerly called Barrow, located in Alaska's Arctic North Slope Borough). All RA participants ($n = 21$) were recruited in rural KwaZulu-Natal, located in the northeast of South Africa. In 16 of 32 AN participants polyps were detected during colonoscopy, whereas no polyps were observed in RA individuals, reflecting the known difference in CRC risk.

AN participants in this study consumed a diet that significantly differed from the diet of RA participants with greater total intake

TABLE 2 Mean dietary intake of macro- and micronutrients in AN and RA study populations[1]

Diet component	AN	RA	*P* value
Energy, kcal	3414 ± 314	2404 ± 140	0.0739
Total fat, g/1000 kcal	45.15 ± 1.56	18.39 ± 2.87	<0.0001
Total carbohydrate, g/1000 kcal	101.6 ± 5.5	164.6 ± 5.9	<0.0001
Total protein, g/1000 kcal	47.89 ± 2.10	28.75 ± 2.32	<0.0001
Total animal protein, g/1000 kcal	37.51 ± 2.41	9.38 ± 2.60	<0.0001
Total vegetable protein, g/1000 kcal	10.38 ± 0.74	19.38 ± 1.53	<0.0001
Total dietary fiber, g/1000 kcal	8.83 ± 0.72	27.91 ± 1.60[2]	<0.0001
Folate, μg/1000 kcal	180.2 ± 15.4	148.8 ± 40.0	0.0562
Calcium, mg/1000 kcal	317 ± 22.0	118.7 ± 29.8	<0.0001
β-Carotene equivalents, μg/1000 kcal	2288 ± 840	2206 ± 1205	0.0390
Retinol, μg/1000 kcal	1041 ± 371	2.74 ± 1.79	<0.0001
Calciferol, μg/1000 kcal	6.87 ± 0.80	0.73 ± 0.25	<0.0001
Ascorbic acid, mg/1000 kcal	42.55 ± 7.0	20.83 ± 6.25	0.0207

[1] Values are mean ± SEM. Statistical analysis by an unpaired *t* test (data normally distributed) or a nonparametric Mann–Whitney *U* test (data not normally distributed) with AN $n = 32$, RA $n = 21$ samples. AN, Alaska Native; RA, rural African.
[2] Includes resistant starch + 10% carbohydrates, which is predominantly resistant starch in the cooked African cornmeal diet [as described before (10)].

of fat, protein, calcium, β-carotene, calciferol, and ascorbic acid (**Table 2**). The AN diet contained significantly higher amounts of animal-derived protein and retinol, likely reflecting a diet rich in meat, fish, and seafood. In contrast, RA diets had a significantly greater intake of total carbohydrates, dietary fiber, and vegetable-derived protein, indicating a plant-based diet (Table 2). The RA diet was also characterized by a low consumption of retinol, reflecting the vitamin A deficiency prevalent in South African populations (23).

The analysis of AN fecal microbiota by 16S rRNA gene sequencing revealed a significantly lower α-diversity and a discrete clustering compared with the RA fecal microbiota on the supervised multidimensional scaling plot (**Figure 1**A, B). Among the dominant gut-related phyla, levels of operational taxonomic units (OTUs) belonging to Actinobacteria and Verrucomicrobia were enriched in the AN fecal microbiota (Figure 1C, **Supplemental Figure 2**A). At family level, Ruminococcaceae and Prevotellaceae, both covering major genera involved in saccharolytic fermentation and SCFA synthesis, showed significantly lower abundance in the AN fecal microbiota (Supplemental Figure 2B). Surprisingly, OTUs belonging to the family Lachnospiraceae, which includes butyrate-producing species from the *Clostridium* XIVa cluster, were more abundant in the AN samples (Supplemental Figure 2B). This may be explained by the large abundance of OTUs related to *Blautia* and *Lachnoclostridium*, which belong to the Lachnospiraceae family (Figure 1D). Within other most abundant genera, the AN fecal microbiota exhibited significantly greater numbers of OTUs belonging to *Bifidobacterium* and *Escherichia-Shigella* (Figure 1D). In contrast, the RA fecal microbiota had significantly greater abundance of OTUs related to *Prevotella 9*, Ruminococcaceae, *Succinivibrio*, and the *Eubacterium coprostanoligenes* group (Figure 1D), covering complex carbohydrate-fermenting and butyrogenic species. OTUs related to *Bilophila* were significantly more abundant in the AN samples, but this did not correlate with higher concentrations of *Bilophila wadsworthia* in feces from AN participants (Supplemental Figure 2C, D). Similarly, no difference was detected in the relative abundance of *Fusobacteria* in feces (Supplemental Figure 2E).

AN participants exhibit different metabolite profiles in feces and urine compared with RA subjects

Consistent with the significant compositional differences in diet and fecal microbiota, we detected distinct metabolite profiles in AN and RA feces and urine using ^{1}H-NMR spectroscopy (**Figure 2**A, B). These differences were significant and can be clearly visualized in the scores plots and the corresponding kernel density estimates of the predictions (21) (Figure 2C, D), which showed a distinct separation between AN and RA samples in fecal and urinary metabolic profiles ($Q^2_Y = 0.65$ and 0.56, respectively). The only colonic metabolites present in greater quantities in AN participants were choline and formate (**Table 3**). In the RA fecal samples, we detected significantly higher relative concentrations of amino acids, purine and pyrimidine derivatives, and SCFAs (Table 3), indicating a high functional diversity of the gut microbiota (24). With regard to urinary metabolites, bile acids and compounds involved in nicotinamide/niacin metabolism were detected in significantly greater amounts in AN samples (Table 3), likely indicating a diet rich in fat and meat or fish. AN urine samples contained higher concentrations of 2-furoylglycine and *N*-methylpyridinium, both associated with coffee intake, whereas RA samples had significantly higher concentrations of compounds associated with amino acid metabolism, consumption of cruciferous vegetables (*S*-methyl-cysteine sulfoxide and related metabolites), and legumes (dimethylglycine) (25).

Differences in colonic metabolism among healthy AN and RA participants are closely linked to divergent abundance of microbial genera

We plotted the fecal/urinary metabolites that showed significantly different concentrations between the AN and RA samples in a metabolic reaction network to analyze their interdependence as linked to host and microbial metabolic pathways (22). The AN samples demonstrated a fecal/urinary metabolite profile suggestive of an increased turnover in nicotinamide metabolism (**Figure 3**A). In contrast, RA samples exhibited greater metabolic activity related to SCFAs, sulfur, purine and pyrimidine, and aromatic and indole compounds, indicating the high diversity of metabolic reaction networks in that cohort (Figure 3A). We performed a correlation analysis between all significantly changed metabolites in urine or fecal water and the abundance of the microbial genera detected by 16S rRNA gene sequencing to explore potential associations of metabolites and specific microbial genera in the gut (Figure 3B). With regard to CRC risk, SCFAs, which were present in higher concentrations in RA feces, showed a significant correlation with the abundance of various genera that cover SCFA producers or microbes (potentially) involved in fermentation of fiber (e.g., *Prevotella* genera and *Butyrivibrio*). These genera were also significantly more abundant in the RA fecal microbiota (Figure 3B). The high fecal SCFA concentrations detected in RA feces also showed a significant negative correlation with genera significantly enriched in the AN fecal microbiota (e.g., *Blautia*), suggesting a distinct assembly of the fecal microbiota with respect to metabolic function (SCFA synthesis) and substrate (dietary fiber). In contrast, choline and formate, identified to be more abundant in AN fecal water, exhibited fewer significant correlations with microbial genera, suggesting that the presence of these 2 metabolites may be linked to multiple factors in addition to bacterial metabolism, e.g., dietary intake. As expected, there were fewer significant correlations detected between urinary metabolites and fecal microbial genera than for fecal samples. The bile acids detected in greater amounts in AN urine samples correlated significantly with high abundance of genera involved in bile acid biotransformation (*Blautia*, *Lachnoclostridium*) and low presence of saccharolytic bacteria (Prevotellaceae, Ruminococcaceae), suggesting an inverse correlation of bile acid biotransformation and SCFA synthesis at the microbial metabolism level.

Lower levels of SCFAs, but not butyrogenic bacteria, in AN than in RA fecal samples

Because the metabonomic analysis indicated lower concentrations of SCFAs in feces of AN people, we performed a targeted quantification of SCFA levels. The concentrations of acetate, propionate, butyrate, and total SCFAs were significantly lower in AN than in RA fecal samples (**Figure 4**A–D). The ratio of these 3 SCFAs in feces remained similar between AN

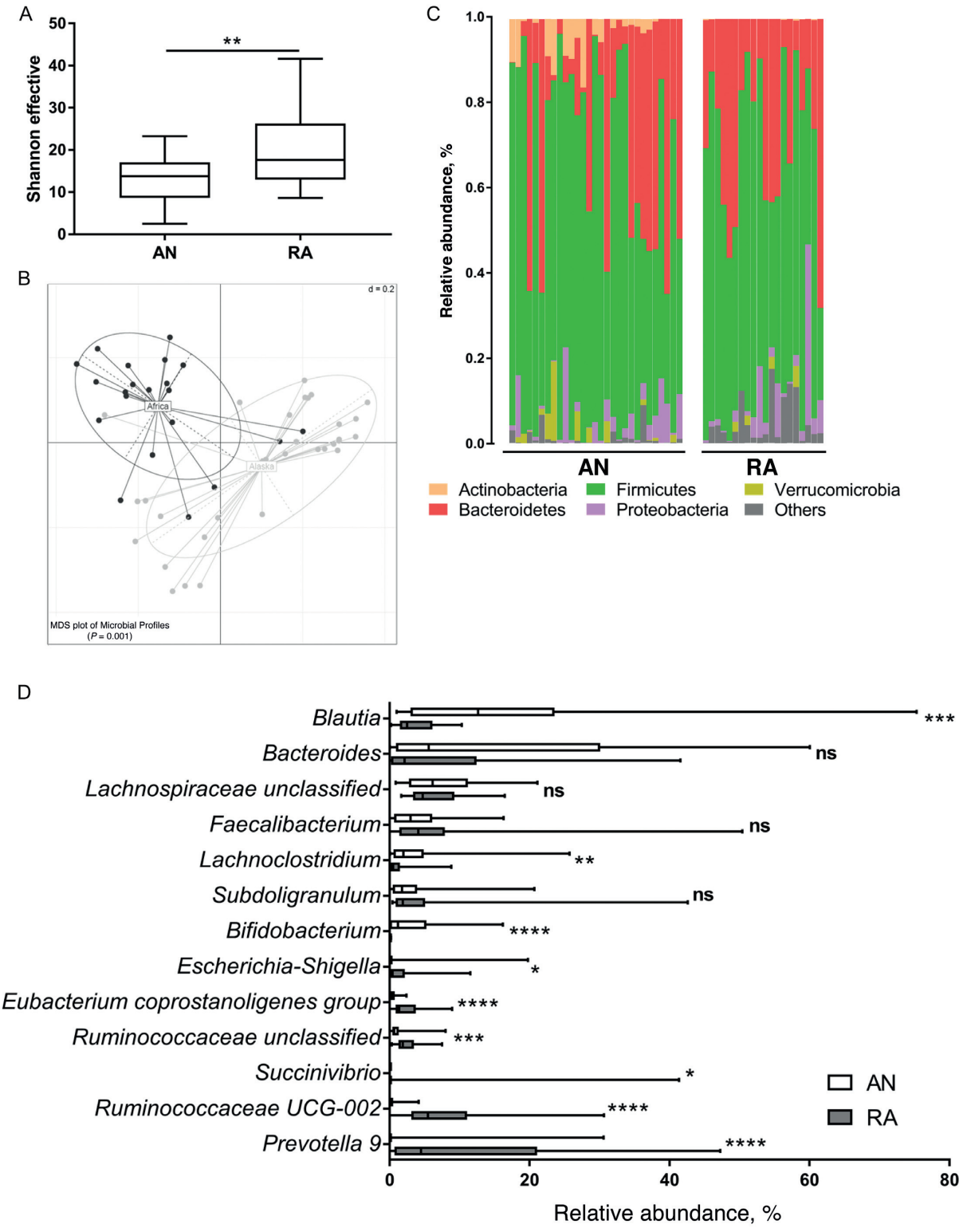

FIGURE 1 The fecal microbiota of AN participants shows a separate clustering and abundance pattern at genus level compared with RA participants. (A) Shannon effective index and (B) MDS plot of generalized UniFrac distances (including permutational multivariate ANOVA test to determine group separation) showing distinct clustering of fecal microbiota profiles detected in the fecal microbiota of AN and RA study participants by 16S rRNA gene sequencing. (C) Cumulative relative abundance of the major gut-related phyla acquired by taxonomic binning of OTUs detected in the fecal microbiota of AN and RA participants by 16S rRNA gene sequencing; each column represents 1 individual. (D) Relative abundance of most abundant genera assigned by taxonomic binning of OTUs detected in the AN and RA fecal microbiota by 16S rRNA gene sequencing in box-and-whisker plots showing minimum to maximum. Statistical analysis used the nonparametric Mann–Whitney *U* test corrected for multiple testing by the Benjamini–Hochberg method with AN $n = 29$ and RA $n = 20$ or 21 samples, respectively. $P < 0.05$ was considered statistically significant: $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; ns = nonsignificant. AN, Alaska Native; MDS, multidimensional scaling; OTU, operational taxonomic unit; RA, rural African; rRNA, ribosomal RNA.

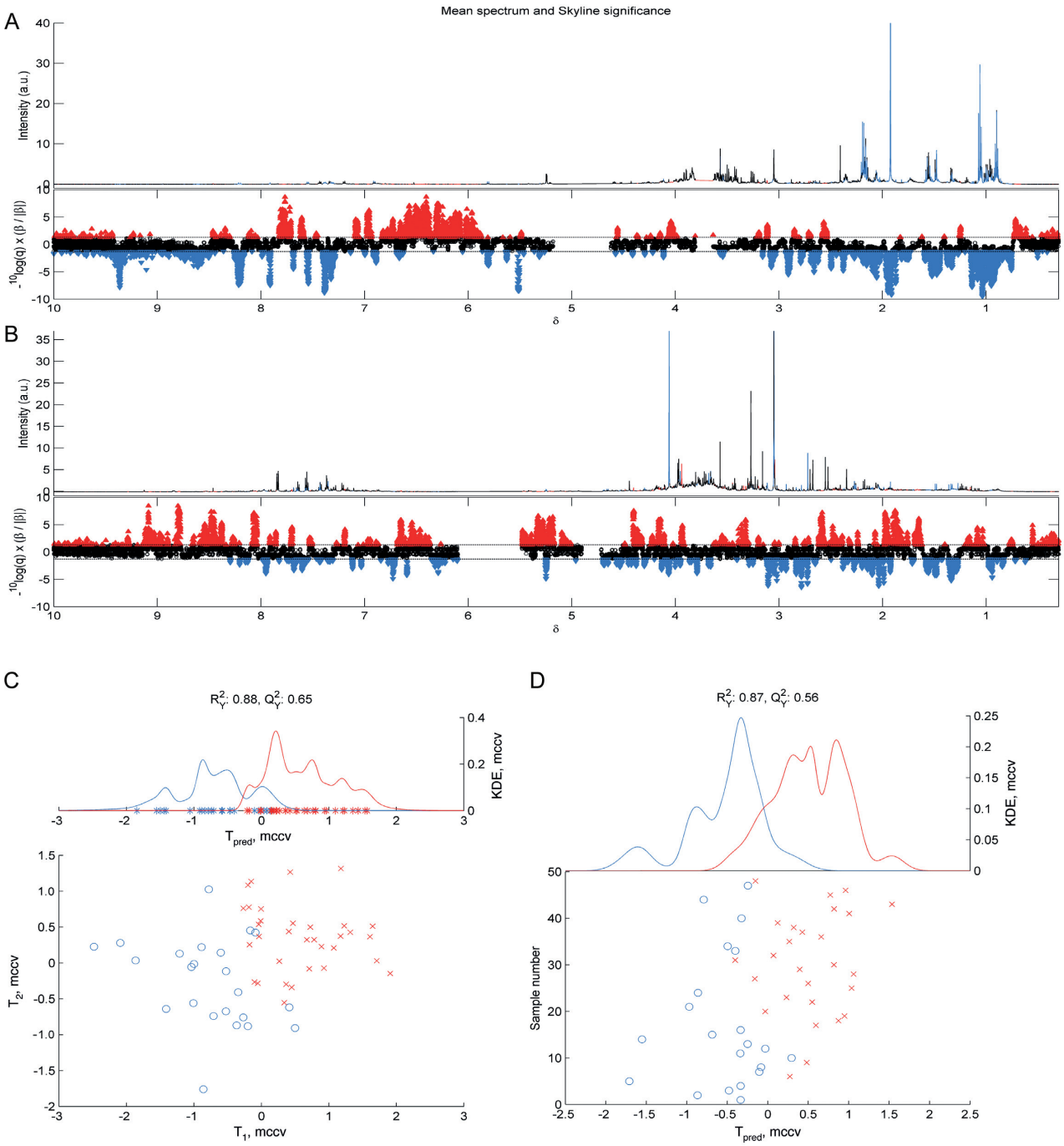

FIGURE 2 Fecal water and urine of samples from AN participants show separate metabolite profiles and are less diverse compared with RA samples. (A, B) Average ^{1}H-NMR spectrum of AN and RA (A) fecal water or (B) urine samples with significant peaks shown in red (higher in AN participants) and blue (higher in RA participants) in top panels (see Table 3 for metabolites and their chemical shifts). The corresponding bottom panels show the Skyline significance ($-\log_{10}$ Q value $\times$ sign of regression coefficient β). The horizontal dashed lines indicate the false discovery rate of 5% on the $\log_{10}$ scale. (C, D) Scores plots of the Monte-Carlo cross-validated partial least-squares discriminant analysis model (21) with KDE and model statistics displayed on top of (C) fecal water and (D) urinary data. Average score is displayed in a red cross (for AN participants) and blue circle (for RA participants). The x axis shows the average scores of the first component of the model, the y axis the second component, and the x axis of the top panel shows the predicted overall score with the KDE (distribution of scores across all models). For both data sets the groups show distinct separation [model goodness of prediction (Q^2_Y) is (C) 0.65 with a high RCV of 0.74 or (D) 0.56 with a high RCV of 0.64]. Statistical analysis was performed with $n = 48$ urine (27 AN, 21 RA) and $n = 53$ fecal water samples. AN, Alaska Native; KDE, kernel density estimate; RA, rural African, RCV, robustness of cross-validation.

and RA samples (Figure 4E). Consistently, AN fecal samples had significantly lower concentrations of isobutyrate, isovalerate, valerate, and hexanoate (**Supplemental Figure 3**A–D). Targeted bacterial genes that are involved in the main pathways of butyrate production, butyryl CoA:acetate-CoA transferase (*bcoA*) and butyrate kinase (*buk*), were present at lower levels in the AN than in the RA fecal samples, but not reaching a statistically significant level (Figure 4F, G). Further analysis of selected

TABLE 3 Identified metabolites in feces and urine with statistically significant differences between AN and RA participants[1]

Compound name	Chemical shift in ppm (multiplicity)	*P* value	*Q* value
Significantly more abundant in AN than in RA fecal samples			
Choline	3.20 (s)	5.44×10^{-3}	1.61×10^{-2}
Formate	8.46 (s)	7.05×10^{-4}	3.26×10^{-3}
Significantly more abundant in RA than in AN fecal samples			
Valerate	0.89 (t) 1.30 (m) 1.53 (m) 2.17 (t)	4.71×10^{-8}	1.42×10^{-6}
Butyrate	0.90 (t) 1.56 (tq) 2.16 (t)	1.18×10^{-5}	1.25×10^{-4}
Propionate	1.06 (t) 2.19 (q)	2.71×10^{-8}	9.26×10^{-7}
Acetate	1.92 (s)	9.98×10^{-5}	6.95×10^{-4}
Glutamate	2.36 (m)	2.14×10^{-4}	1.27×10^{-3}
Glycine	3.57 (s)	3.56×10^{-3}	1.16×10^{-2}
Uracil	5.81 (d) 7.54 (d)	6.95×10^{-10}	6.41×10^{-8}
Tyrosine	6.91 (d) 7.20 (d)	1.11×10^{-3}	4.66×10^{-3}
Phenylalanine	7.31 (d) 7.38 (m)	1.33×10^{-12}	1.44×10^{-9}
Nicotinate	7.53 (dd) 8.61 (dd) 8.94 (d) 8.26 (dt)	5.30×10^{-6}	6.68×10^{-5}
Xanthine	7.91 (s)	2.69×10^{-10}	3.27×10^{-8}
Hypoxanthine	8.21 (d)	3.15×10^{-10}	3.62×10^{-8}
Significantly more abundant in AN than in RA urinary samples			
Unknown bile acids	0.75 (s) 0.76 (s)	9.60×10^{-4}	5.51×10^{-3}
Creatine	3.04 (s) 3.94 (s)	5.00×10^{-4}	3.43×10^{-3}
N-methyl-2-pyridone-carboxamide	3.65 (d) 6.67 (d) 7.83 (dd) 8.34 (d)	1.01×10^{-7}	4.85×10^{-6}
2-furoylglycine	3.93 (s) 7.71 (d) 7.19 (dd) 6.65 (dd)	2.26×10^{-8}	1.52×10^{-6}
N-methylpyridinium	4.40 (s) 8.79 (d) 8.55 (td) 8.05 (td)	5.54×10^{-13}	3.60×10^{-9}
N-methylnicotinamide	4.48 (s) 8.19 (t) 8.90 (d) 8.97 (d) 9.28 (s)	1.09×10^{-5}	1.90×10^{-4}
Significantly more abundant in RA than in AN urinary samples			
Fatty acids (C5–C10)	0.84 (d) 1.3 (s) 1.56 (m)	4.21×10^{-6}	9.46×10^{-5}
3-hydroxyisovalerate	1.27 (s) 2.37 (s)	7.63×10^{-5}	8.37×10^{-4}
2-hydroxyisobutyrate	1.36 (s)	3.98×10^{-4}	2.87×10^{-3}
Alanine	1.48 (d)	1.71×10^{-6}	4.81×10^{-5}
Acetate	1.92 (s)	9.81×10^{-6}	1.76×10^{-4}
Phenylacetylglutamine	1.93 (m) 2.13 (m) 2.27 (t) 3.68 (q) 4.19 (dd) 7.36 (t) 7.43 (t)	2.31×10^{-6}	5.95×10^{-5}
Dimethylamine	2.72 (s)	2.00×10^{-8}	1.38×10^{-6}
N-acetyl-*S*-methyl-cysteine sulfoxide	2.78 (s)	3.22×10^{-9}	3.39×10^{-7}
S-methyl-cysteine sulfoxide metabolite	2.80 (s)	2.92×10^{-4}	2.28×10^{-3}
S-methyl-cysteine sulfoxide	2.83 (s)	3.72×10^{-6}	8.56×10^{-5}
Dimethylglycine	2.93 (s)	2.63×10^{-5}	3.70×10^{-4}
Lysine	3.02 (t)	1.17×10^{-7}	5.51×10^{-6}
Creatinine	3.05 (s) 4.06 (s)	3.48×10^{-6}	8.10×10^{-5}

[1]AN, Alaska Native; d, doublet; dd, doublet of doublets; dt, doublet of triplets; m, multiplet; q, quartet; Q value, *P* value adjusted for false discovery rate (FDR); RA, rural African; s, singlet; t, triplet; td, triplet of doublets; tq, triplet of quartets.

genera belonging to major groups of butyrogenic bacteria (e.g., *Roseburia*, *Faecalibacterium*) (26, 27) confirmed that most levels were not significantly different between AN and RA samples (**Supplemental Figure 4**A–C; Figure 1D). *Alistipes*, the *Eubacterium halli* group, and *Flavonifractor* were significantly more abundant in the AN fecal samples, whereas *Butyrivibrio* and *Oscillibacter* were present in greater numbers in the RA fecal microbiota (Supplemental Figure 4A–C). However, these genera tended to be present at low numbers in both populations. Being a critical prerequisite for efficient microbial fermentation of fiber, we investigated the metabolic capacity of the fecal microbiota to remove hydrogen from the intestinal lumen through sulfate-reduction and methane synthesis. Gene copies of the dissimilatory sulfite reductase subunit A (*dsrA*), which represents sulfate-reducing bacteria, were significantly less abundant in the AN than in the RA fecal samples (Figure 4H). In contrast, gene copy numbers of the methyl coenzyme-M reductase A (*mcrA*), representative of methanogenesis by Archaea, were not different between AN and RA samples (Figure 4I).

High concentrations of fecal bile acids in, and a high capacity for, bile acid conversion by the AN gut microbiota

The concentrations of the main unconjugated primary bile acids, cholic acid (CA) and chenodeoxycholic acid, were significantly higher in AN than in RA samples (**Figure 5**A, B). Fecal concentrations of the main unconjugated secondary bile acid, DCA, were >2-fold higher in AN than in RA participants (Figure 5C). Consistently, the abundance of 7α-dehydroxylating bacteria [indicated by gene copy numbers of the *CD* cluster of the bile acid-inducible operon (*baiCD*)] was significantly greater in the fecal microbiota of AN participants relative to RA participants (Figure 5D). Fecal concentrations of lithocholic acid and ursodeoxycholic acid, both showing inconsistent effects on CRC risk, were similar in both groups (Figure 5E, F). Although not significant for every single bile acid, AN participants showed higher amounts of fecal bile acids conjugated with glycine than did their RA counterparts (Figure 5G). Similarly, taurine-conjugated bile acids were present in higher amounts in the bile acid pool of AN participants (Figure 5H).

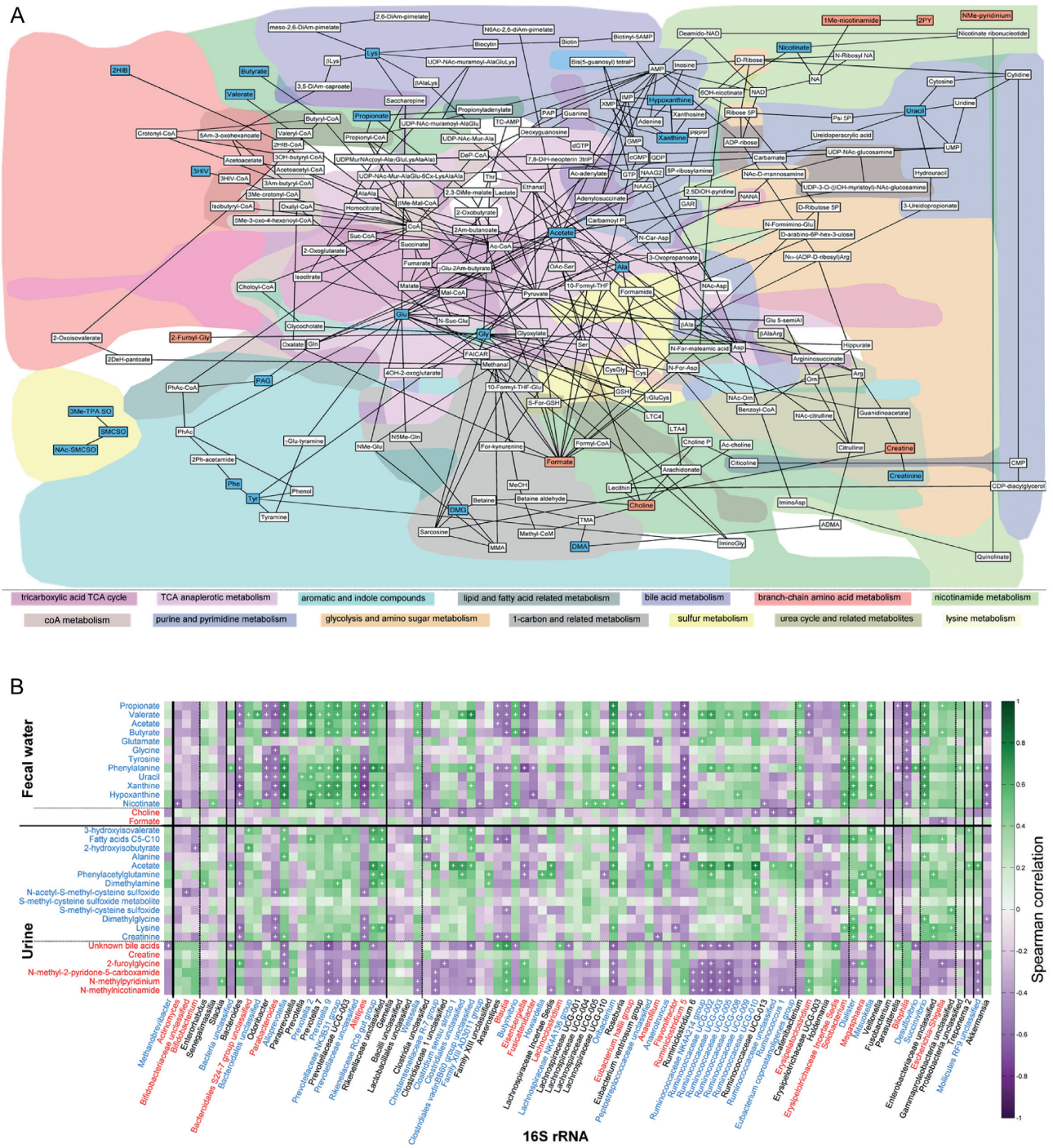

FIGURE 3 The fecal water/urinary metabolite profiles of AN participants cluster in separate metabolic networks and correlate with distinct microbial genera compared with RA participants. (A) Metabolic reaction networks of metabolites found differentially expressed between AN and RA participants' urine and feces. The network shows links between metabolites, if the reaction entry in Kyoto Encyclopedia of Genes and Genomes indicates a main reactant pair and the reaction is either mediated by *1*) an enzyme linked to human genes, *2*) an enzyme linked to genes from identified bacterial groups, or *3*) it is part of a spontaneous process. The color of the metabolites indicates whether the metabolite is found in higher concentrations in the AN (red) or RA (blue) urine/feces, or whether the metabolites are not significantly associated with any comparison but are part of the metabolic network (white). (B) Heat map of Spearman correlations between fecal and urinary metabolites associated with differences in AN and RA participants with 16S rRNA gene sequencing data. Correlations were adjusted for multiple testing using the Benjamini–Hochberg FDR. Significant correlations (at FDR of 1%) are indicted by a white '+'. A thick black line indicates a split between either kingdom (*x* axis) or biofluid (*y* axis), a solid line indicates different phyla, and a dashed line indicates different class (*x* axis) and a different sign of association between AN and RA samples (*y* axis). The color of the metabolites/microbial genera indicates whether it is found in greater concentrations in AN (red) or RA (blue) participants or if there is no significant difference between both groups (black). Statistical analysis was performed with sample numbers as listed for previous analyses in Figures 1 and 2. ADMA, asymmetric dimethyl arginine; AN, Alaska Native; DMA, dimethylamine; FAICAR, 5-Formamido-1-(5-phosphoribosyl)imidazole-4-carboxamide; FDR, false discovery rate; GAR, glycinamide ribonucleotide; GSH, glutathione; 2HIB, 2-hydroxyisobutyrate; IMP, inosine monophosphate; LTA4, leukotriene A4; LTC4, leukotriene C4; NA, nicotinamide; NAAG, N-Acetylaspartylglutamate; NANA, N-Acetylneuraminic acid; PAG, phenylacetylglutamine; PAP, phosphoadenosine phosphate; PRPP, 5-Phosphoribosyl 1-pyrophosphate; RA, rural African; SMCSO=S-methylcysteine sulfoxide; TCA, tricarboxylic acid; TMA, Trimethylamine, XMP, Xanthosine monophosphate.

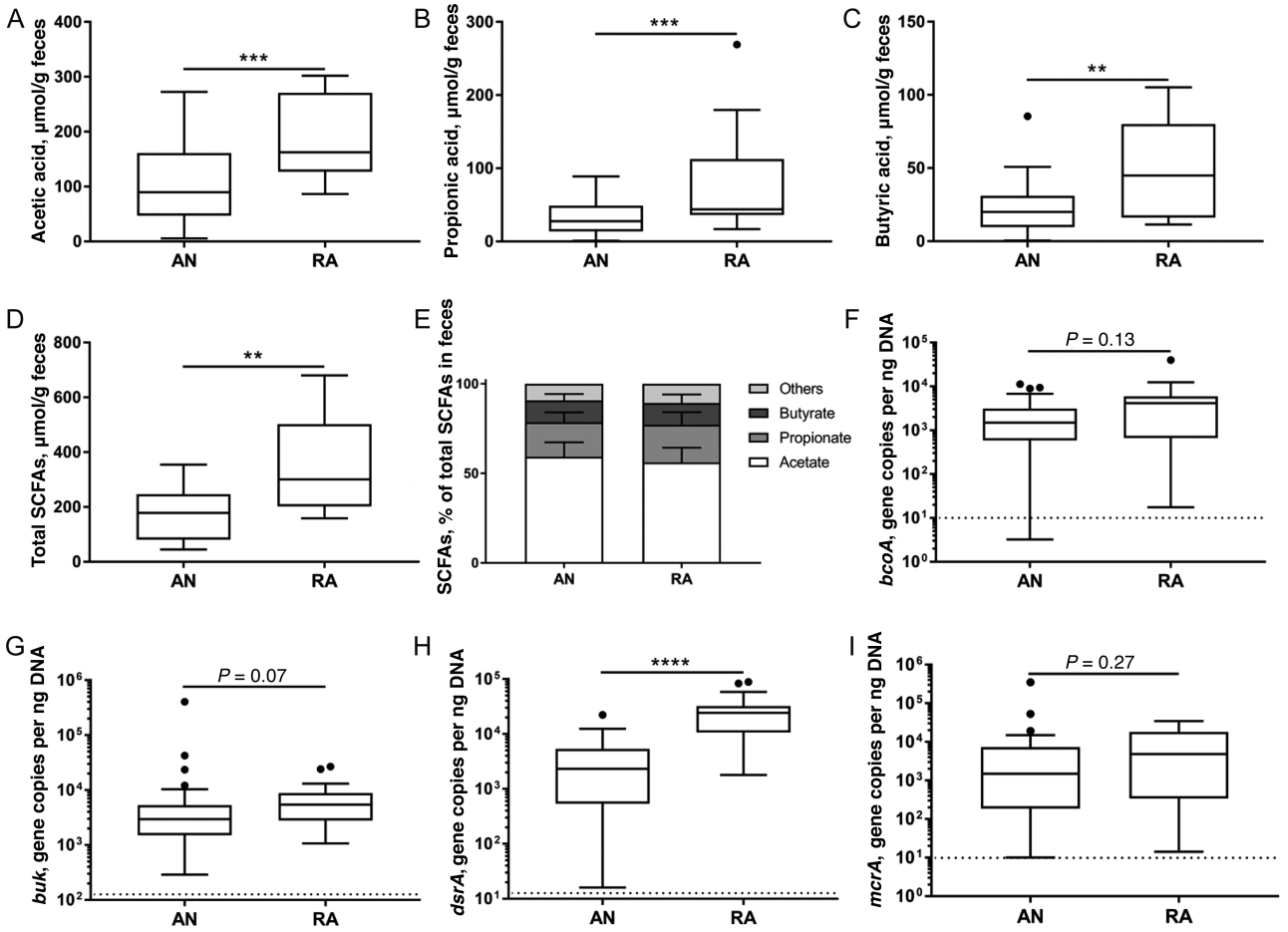

FIGURE 4 Concentrations of SCFAs are significantly lower in AN than in RA fecal samples, not correlating with the abundance of bacterial genes involved in butyrate synthesis. Concentrations of (A) acetic acid, (B) propionic acid, (C) butyric acid, and (D) total SCFAs detected by GC in feces of AN and RA participants. (E) Percentaged ratio of major SCFAs to total concentrations of SCFAs in feces of AN and RA participants. Gene copy numbers of (F) *bcoA* and (G) *buk* representative for butyrate-producing bacteria, (H) *dsrA* representative for sulfate-reducing bacteria, and (I) *mcrA* involved in production of methane by Archaea in the AN and RA fecal microbiota using qPCR. Dashed line indicates the detection limit of qPCR. Statistical analysis was performed by an unpaired *t* test (data normally distributed) or a nonparametric Mann–Whitney *U* test (data not normally distributed) with AN $n = 32$, RA $n = 21$ samples. $P < 0.05$ was considered statistically significant: $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$. AN, Alaska Native; *bcoA*, butyryl CoA:acetate-CoA transferase; *buk*, butyrate kinase; *dsrA*, dissimilatory sulfite reductase subunit A; *mcrA*, methyl coenzyme-M reductase A; RA, rural African.

Tumor-promoting capacity of the AN gut microbiota is not linked to adenomatous polyps in the colon

The gut microbiota composition of AN individuals with or without adenomatous polyps did not differ significantly (**Figure 6**A) and showed only minor alterations at genus level, with *Erysipelotrichaceae incertae sedis* being significantly more and *Fusicatenibacter* being less abundant in AN participants with polyps (Figure 6B, C). All quantified fecal SCFAs or bile acids did not show significant differences according to polyp status of AN individuals (**Supplemental Figure 5**A–C, F, G shows the main representative SCFAs and bile acids). Consistently, presence of adenomatous polyps did not affect the expression of functional microbial genes involved in butyrate synthesis or bile acid conversion (Supplemental Figure 5D, E, H), sulfite reduction, or methanogenesis and the abundance of *Fusobacteria* or *B. wadsworthia* in feces (data not shown). AN participants with polyps had a higher intake of protein, which was statistically nonsignificant (mean ± SEM: 51 ± 2.7 compared with 45.2 ± 3.1 g/1000 kcal; $P = 0.17$), and of calciferol (7.6 ± 0.9 compared with 6.2 ± 1.3 µg/1000 kcal; $P = 0.04$) than AN participants without polyps.

Discussion

Here, we have provided novel evidence of how interactions of dietary and microbial factors may contribute to the high CRC risk observed in AN people: the AN low-fiber diet results in low levels of saccharolytic fermentation and low concentrations of tumor-suppressive butyrate in the colon, despite a conserved capacity for butyrate synthesis by the gut microbiota. This is complemented by a metabolic shift of the AN microbiota toward bile acid transformation characterized by high abundance of 7α-dehydroxylating bacteria and a compositional predominance of *Blautia* and *Lachnoclostridium*, leading to high concentrations of tumor-promoting DCA. Together, low butyrate and high DCA concentrations create a colonic microenvironment that opens susceptibility to carcinogenic stimuli such as *N*-nitrosamines or

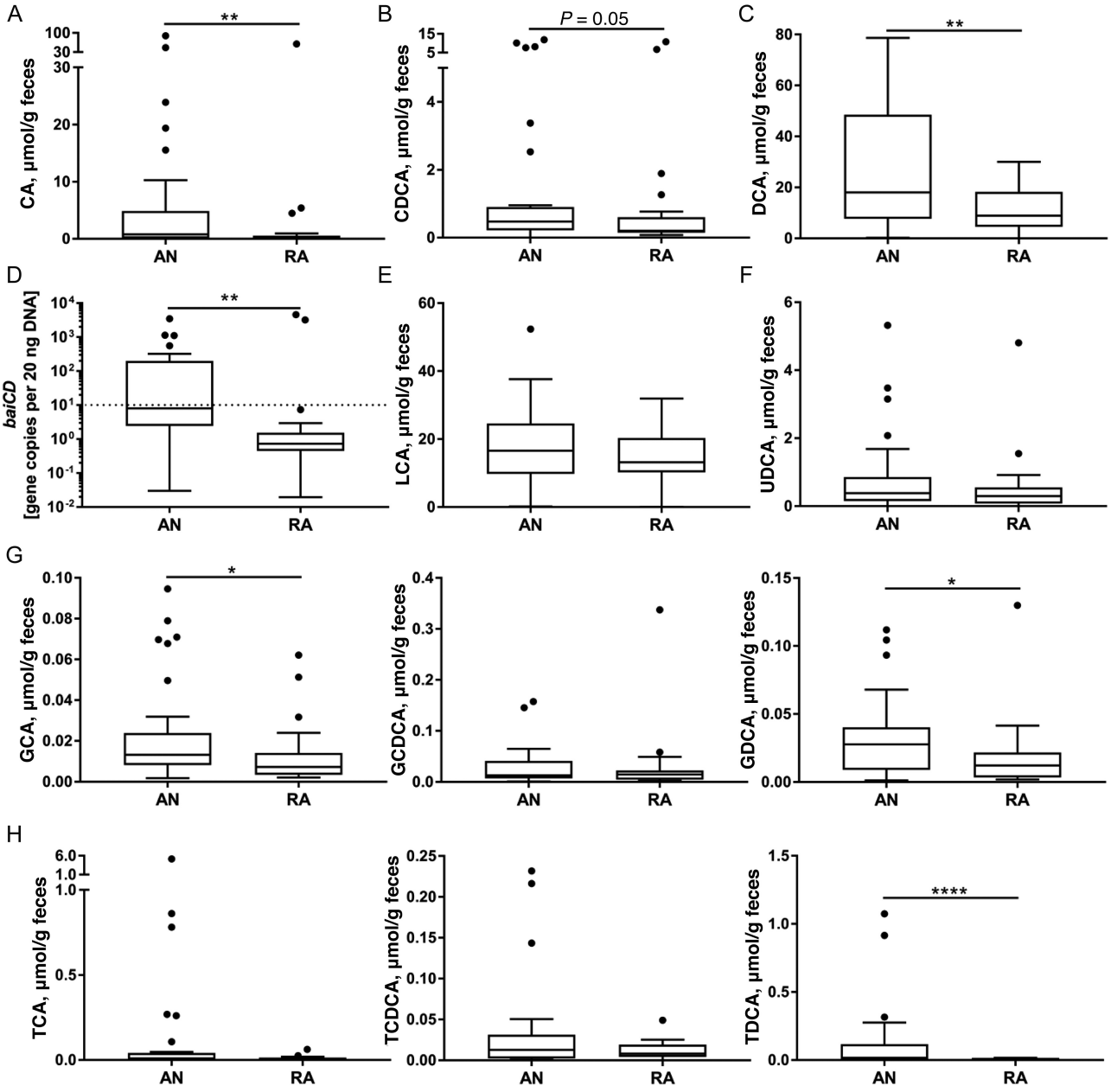

FIGURE 5 Concentrations of major primary and secondary bile acids are higher in AN than in RA fecal samples, consistent with a greater capacity for bile acid conversion and independent of conjugation status. Concentrations of (A) CA, (B) CDCA, and (C) DCA detected by LC-MS in feces of AN and RA participants. (D) Gene copy numbers of *baiCD* representative for 7α-dehydroxylating bacteria detected in the fecal microbiota of AN and RA participants by qPCR. Dashed line indicates the detection limit of qPCR. Concentrations of (E) LCA, (F) UDCA, and bile acids conjugated to (G) glycine (GCA, GCDCA, GDCA) or (H) taurine (TCA, TCDCA, TDCA) detected by LC-MS in AN and RA fecal samples. Statistical analysis was performed by an unpaired *t* test (data normally distributed) or a nonparametric Mann–Whitney *U* test (data not normally distributed) with AN $n = 32$, RA $n = 21$ samples. $P < 0.05$ was considered statistically significant: $^{*}P < 0.05$; $^{**}P < 0.01$; $^{****}P < 0.0001$. AN, Alaska Native; *baiCD*, bile acid-inducible operon; CA, cholic acid; CDCA, chenodeoxycholic acid; DCA, deoxycholic acid; GCA, glycocholic acid; GCDCA, glycochenodeoxycholic acid; GDCA, glycodeoxycholic acid; LCA, lithocholic acid; RA, rural African; TCA, taurocholic acid; TCDCA, taurochenodeoxycholic acid; TDCA, taurodeoxycholic acid; UDCA, ursodeoxycholic acid.

polycyclic aromatic hydrocarbons, derived from the high intake of meat or tobacco smoking (7, 28, 29).

Numerous studies demonstrate antineoplastic effects of butyrate in the colon (30–32). It is tempting to speculate that AN people face a fundamental deficiency in butyrogenic bacteria due to low fiber intake, especially considering that a low-fiber diet can promote an irreversible loss of taxa in the microbiota (33). To our surprise and inconsistent with previous studies in high-risk cohorts (9, 10), we did not detect in the AN fecal microbiota significantly lower copy numbers of enzymes that are involved in the major pathways of microbial butyrate synthesis. This may be, in part, explained by differences in distinct environmental factors other than diet that contribute to the composition of the gut microbiota (34) and may also suggest the low intake of fiber

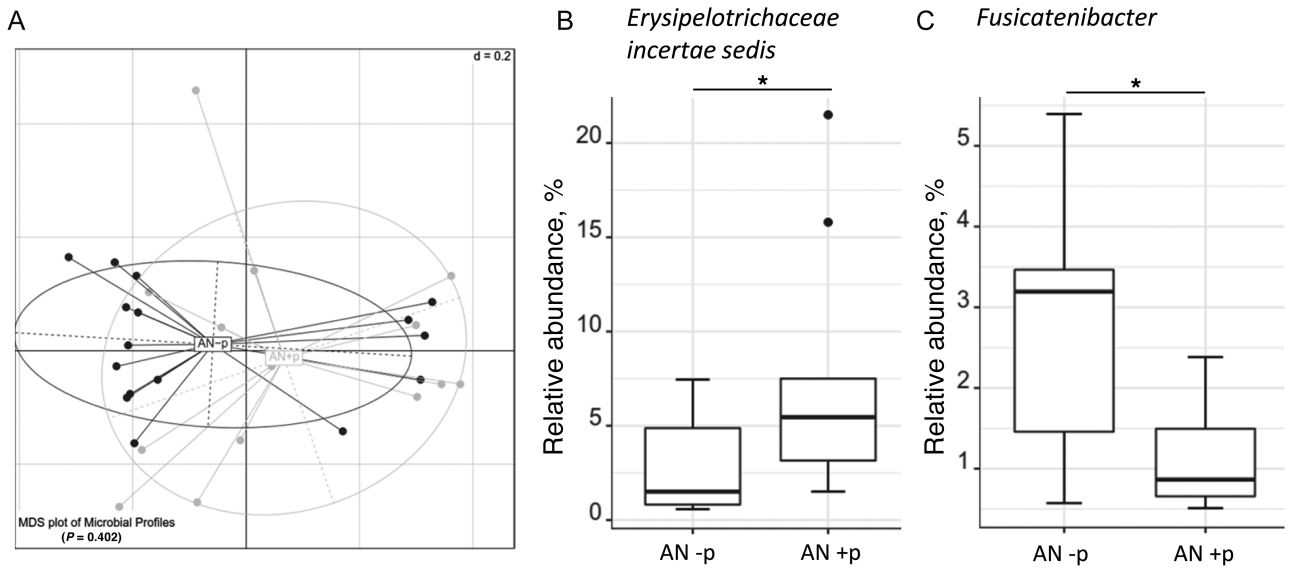

FIGURE 6 Adenomatous polyps in the colon have minor impact on the fecal microbiota of AN participants. (A) MDS plot of generalized UniFrac distances (including permutational multivariate ANOVA test to determine group separation) showing no distinct clustering of fecal microbiota profiles detected in the fecal microbiota of AN study participants with (AN +p) or without (AN -p) adenomatous polyps in the colon by 16S rRNA gene sequencing. (B, C) Relative abundance of genera of the fecal microbiota of AN participants that are significantly differently abundant between AN study participants with (AN +p) or without (AN -p) adenomatous polyps in the colon assigned by taxonomic binning of operational taxonomic units detected in the AN and RA fecal microbiota by 16S rRNA gene sequencing. Statistical analysis used a nonparametric Mann–Whitney *U* test corrected for multiple testing by the Benjamini–Hochberg method with AN $n = 29$ samples. $P < 0.05$ was considered statistically significant: $^{*}P < 0.05$. AN, Alaska Native; MDS, multidimensional scaling.

to be the "metabolic bottleneck." The similar levels of genera that cover major butyrogenic bacteria (e.g., *Faecalibacterium*, *Roseburia*) in the AN and RA fecal microbiota indicate a conserved abundance of butyrate-producing bacteria in the AN microbiota. Genera able to utilize protein-dependent pathways for butyrate synthesis were similarly abundant in AN and RA fecal microbiota (e.g., *Odoribacter*, *Pseudoflavonifractor*) or enriched in AN samples (*Alistipes* and *Flavonifractor*, both able to utilize lysine for butyrate production) (27). However, the overall contribution of these genera to colonic butyrate production is low and may not compensate for the detrimental effects of nitrogenous metabolites resulting from high protein fermentation. The persistent levels of butyrogenic bacteria in AN participants with a low-fiber diet indicate the conserved potential for fiber fermentation and suggest that the increased consumption of fiber-rich foods may be well tolerated in AN people. The reactivation of this metabolic capacity would support fiber supplementation as a promising approach to reduce CRC risk in healthy AN individuals.

The high fiber intake among RA participants augments the saccharolytic fermentation capacity of the gut microbiota characterized by high abundance of Prevotellaceae, Ruminococcaceae, some genera of Lachnospiraceae, and *Succinivibrio* as demonstrated previously (9, 10, 35). Besides the large quantities of indigestible carbohydrates coming from legumes and cruciferous vegetables, the plant-based diet of RA individuals is also likely to provide high amounts of phytochemicals, such as polyphenols and glucosinolates (as confirmed in the RA urinary metabolite profiles). It has been shown that phytochemicals released from cell walls during plant residue breakdown may augment the antiproliferative effects of butyrate (7, 36). Beyond the effect on butyrate synthesis, recent experimental studies demonstrated that low-fiber diets stimulate the activity and growth of mucus-degrading bacteria (e.g., *Akkermansia muciniphila*), leading to an impaired intestinal mucus barrier (37, 38). Despite the large abundance of Verrucomicrobia in the AN fecal microbiota, we did not detect significantly greater levels of *Akkermansia* than in RA participants (data not shown).

High concentrations of butyrate in the colon suppress intestinal tumor formation or inflammation in mice stimulated by high-fat or bile acid–supplemented diets (32, 39). Aggravating the effects of low butyrate concentrations with regard to CRC risk, the high-fat diet of AN participants stimulated hepatic synthesis of bile acids and their delivery to the colon, where 7α-dehydroxylating bacteria promote high concentrations of DCA demonstrating tumor-promoting activity (40–42). The metabolic shift toward bile acid metabolism is further supported by the high abundance of *Blautia* and *Lachnoclostridium* in the AN gut microbiota. Both genera belong to the Lachnospiraceae family, which showed a positive association with secondary bile acids in murine cecum (43). They cover species comprised of or closely linked to 7α-dehydroxylating bacteria of the *Clostridium* genus, such as *Clostridium scindens*, that mediate the conversion of bile acids to DCA (44). The high abundance of *Blautia* in the AN fecal microbiota confirms experimental studies that demonstrated an expansion of *Blautia* and 7α-dehydroxylating bacteria when CA was added to the diet of rodents (45, 46). Lower concentrations of CA correlated with lower numbers of *Blautia* and 7α-dehydroxylating bacteria in feces of patients with cirrhosis (47, 48). Though at a low relative abundance level, the genus *Blautia* was also found to be enriched in the fecal microbiota of CRC patients (49), suggesting a potential contribution of *Blautia* species to colonic tumorigenesis in terms of bile acid metabolism, in particular

generation of DCA. A recent meta-analysis confirmed a higher capacity of CRC metagenomes to produce secondary bile acids than those in healthy controls (50). The AN fecal microbiota also showed greater numbers of *Escherichia-Shigella*, a genus that includes several opportunistic pathogens and belongs to Gammaproteobacteria. These are highly resistant to bile salts and were also enriched under oral administration of CA (45, 51). In contrast to prior studies (10), we did not detect a significantly different abundance of *B. wadsworthia*, a pathobiont that grows when exposed to high amounts of saturated dietary fat and that is associated with experimental colitis and CRC risk (52, 53), or *Fusobacteria* that are associated with CRC (54) in fecal samples from both groups. For *Fusobacteria* most studies show a higher abundance in CRC patients and not in healthy individuals (54, 55), supporting an association with CRC rather than CRC risk. However, *B. wadsworthia* was shown to be associated with CRC risk markers, suggesting that distinct AN environmental factors other than diet may affect the abundance of these 2 gut bacteria in healthy individuals.

In addition, the high amount of meat in the AN diet may promote detrimental alterations in a colonic environment preconditioned by low butyrate concentrations. A metabolic shift of microbial metagenomes from carbohydrate fermentation to protein degradation was shown to be associated with CRC (50). A diet rich in meat promotes high concentrations of heterocyclic amines, *N*-nitrosamines, or polycyclic aromatic hydrocarbons in the colon that have procarcinogenic effects (7). Critically, their tumor-promoting activity is ameliorated by supplementation with fiber or butyrate (56, 57). Another source of carcinogens, tobacco use, in particular smoking cigarettes, is high in prevalence in AN communities (58, 59) and a link between cigarette smoking and CRC risk has been demonstrated (28). Increased amounts of dietary meat may also limit the beneficial effects from the high intake of fish/seafood by AN people. The AN diet combines the increased abundance of microbial genera found in mice fed a fish oil– or lard-based diet; the latter is associated with higher body weight, a risk factor for CRC (7, 60). Finally, high concentrations of choline, as detected in AN fecal samples, and a stimulated choline metabolism were demonstrated in healthy individuals and rodents fed a Western diet (10, 61). Choline can be converted by the gut microbiota to trimethylamine, which is then oxidized in the liver to produce trimethylamine *N*-oxide. Plasma concentrations of trimethylamine *N*-oxide are positively associated with and genetically related to CRC (62, 63) or adverse cardiovascular events (64, 65).

In contrast to previous studies, we did not find major differences in microbiota signatures between AN participants with or without polyps. In a study by Zackular et al. (66), adenoma polyp individuals had a lower abundance of *Bacteroides*, Lachnospiraceae, Clostridiales, and some *Clostridium* genera than healthy controls, which was not confirmed in our AN study cohort. Colonic polyps have been reported to be associated with higher abundance of the genus *Bilophila*, *Fusobacterium nucleatum*, increased microbial bile acid metabolism, and lower abundance of butyrate-producing bacteria (66–68). We detected most of these polyp-associated changes in AN individuals even when no polyps were present, highlighting the tumor-promoting colonic milieu in healthy AN individuals. This also suggests that compositional shifts in the gut microbiota may not serve as a screening marker for polyps in this population, where other factors may determine whether this colonic environment triggers neoplastic change. Because diet was not analyzed or reported in the previous studies (66–68), our results also suggest that polyp-associated alterations in the (fecal) gut microbiota may be not primarily related to polyp presence but diet patterns associated with CRC risk.

There are limitations to this observational study, which may have benefited from more participants, several time points of sampling, and the collection of mucosal biopsies during colonoscopy. Being a general limitation of observational studies, we described associations of diets and metabolites linked to CRC risk, but not causation. However, this study provides an initial baseline characterization for a population at very high CRC risk and signifies which approaches, especially with regard to larger-scale follow-up intervention studies, might best reduce CRC risk.

In conclusion, our results strongly support the hypothesis that a diet low in fiber and rich in fat promotes a colonic environment that contributes to the high rate of CRC in AN people. The loss of colonic butyrate may predispose the colonic epithelium to a high susceptibility toward tumor-promoting DCA, triggered by high-fat diet, and procarcinogenic metabolites, derived from high meat intake. Critically, this is driven by altered metabolic activity in the AN microbiota, characterized by impaired metabolic and microbial diversity in the colon. At the other end of the CRC risk chart, the results from RA participants highlight the importance of a high-fiber, low-fat and -meat diet for colonic health. The RA diet provides a blueprint for efficient CRC risk reduction by improving butyrate synthesis of the gut microbiota, while containing biotransformation of bile acids, despite its relative deficiency in some micronutrients (e.g., vitamin A). Owing to the very high CRC rates, AN people present a unique opportunity for proof-of-concept dietary modifications to prevent CRC and for mechanistic analysis of the antagonistic relation of butyrate and DCA. Supplementation of the AN diet with fiber to increase luminal butyrate concentrations, followed by a reduction in saturated animal-derived fat to limit colonic DCA concentrations, is warranted to reduce the high rates of CRC in the AN population.

We thank the administration and outpatient clinic and pharmacy staffs at the Alaska Native Medical Center in Anchorage and Samuel Simmonds Memorial Hospital in Utqiagvik, Alaska for their support of this project. We thank Timothy Stevenson for dietary data transformation and entry, enabling nutrient data analysis. We thank Chadd and Kate Bain, and the iZulu Lodge and Orphanage Projects team in KwaZulu-Natal, for help with the recruitment of the African volunteers.

The authors' responsibilities were as follows—KRK, TKT, JKN, and SJDOK: designed the research; SO, ASW, JMP, JVL, KRK, GMD, CAF, JEO, PES, FDS, FRS, ASW, KN, F Brouard, JPD, MB, CNA, DS, F Bhatti, BM, AF, AM, and HRG: conducted the research; SO, JK, and SJDOK: wrote the manuscript; SJDOK: had primary responsibility for the final content; and all authors: read and approved the final manuscript. The authors report no conflicts of interest.

References

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer 2019;144:1941–53.
2. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut 2017;66:683–91.

3. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology 2010;138:2044–58.
4. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst 1981;66:1191–308.
5. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Report. Food, nutrition, physical activity, and the prevention of colorectal cancer [Internet]. London: World Cancer Research Fund/American Institute for Cancer Research; 2011. Available from: http://www.wcrf.org/sites/default/files/Colorectal-Cancer-2011-Report.pdf (accessed 17 July 2017).
6. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Morenga LT. Carbohydrate quality and human health: a series of systematic reviews and meta-analyses. Lancet 2019;393(10170):434–45.
7. O'Keefe SJD. Diet, microorganisms and their metabolites, and colon cancer. Nat Rev Gastroenterol Hepatol 2016;13:691–706.
8. Ocvirk S, O'Keefe SJ. Influence of bile acids on colorectal cancer risk: potential mechanisms mediated by diet-gut microbiota interactions. Curr Nutr Rep 2017;6:315–22.
9. Ou J, Carbonero F, Zoetendal EG, DeLany JP, Wang M, Newton K, Gaskins HR, O'Keefe SJ. Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. Am J Clin Nutr 2013;98:111–20.
10. O'Keefe SJD, Li JV, Lahti L, Ou J, Carbonero F, Mohammed K, Posma JM, Kinross J, Wahl E, Ruder E, et al. Fat, fibre and cancer risk in African Americans and rural Africans. Nat Commun 2015;6:6342.
11. Brown MO, Lanier AP, Becker TM. Colorectal cancer incidence and survival among Alaska Natives, 1969–1993. Int J Epidemiol 1998;27:388–96.
12. Boardman LA, Lanier AP, French AJ, Schowalter KV, Burgart LJ, Koller KR, McDonnell SK, Schaid DJ, Thibodeau SN. Frequency of defective DNA mismatch repair in colorectal cancer among the Alaska Native people. Cancer Epidemiol Biomarkers Prev 2007;16:2344–50.
13. Kelly JJ, Alberts SR, Sacco F, Lanier AP. Colorectal cancer in Alaska Native people, 2005–2009. Gastrointest Cancer Res 2012;5:149–54.
14. Perdue DG, Haverkamp D, Perkins C, Daley CM, Provost E. Geographic variation in colorectal cancer incidence and mortality, age of onset, and stage at diagnosis among American Indian and Alaska Native people, 1990–2009. Am J Public Health 2014;104(Suppl 3):S404–14.
15. Johnson JS, Nobmann ED, Asay E, Lanier AP. Developing a validated Alaska Native food frequency questionnaire for western Alaska, 2002–2006. Int J Circumpolar Health 2009;68:99–108.
16. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 2012;6:1621–4.
17. Dona AC, Jiménez B, Schäfer H, Humpfer E, Spraul M, Lewis MR, Pearce JTM, Holmes E, Lindon JC, Nicholson JK. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. Anal Chem 2014;86:9887–94.
18. Gratton J, Phetcharaburanin J, Mullish BH, Williams HRT, Thursz M, Nicholson JK, Holmes E, Marchesi JR, Li JV. Optimized sample handling strategy for metabolic profiling of human feces. Anal Chem 2016;88:4661–8.
19. Posma JM, Garcia-Perez I, Heaton JC, Burdisso P, Mathers JC, Draper J, Lewis M, Lindon JC, Frost G, Holmes E, et al. Integrated analytical and statistical two-dimensional spectroscopy strategy for metabolite identification: application to dietary biomarkers. Anal Chem 2017;89:3300–9.
20. Lagkouvardos I, Fischer S, Kumar N, Clavel T. Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. PeerJ 2017;5:e2836.
21. Posma JM, Garcia-Perez I, Ebbels TMD, Lindon JC, Stamler J, Elliott P, Holmes E, Nicholson JK. Optimized phenotypic biomarker discovery and confounder elimination via covariate-adjusted projection to latent structures from metabolic spectroscopy data. J Proteome Res 2018;17:1586–95.
22. Posma JM, Robinette SL, Holmes E, Nicholson JK. MetaboNetworks, an interactive Matlab-based toolbox for creating, customizing and exploring sub-networks from KEGG. Bioinformatics 2014;30:893–5.
23. Wirth JP, Petry N, Tanumihardjo SA, Rogers LM, McLean E, Greig A, Garrett GS, Klemm RDW, Rohner F. Vitamin A supplementation programs and country-level evidence of vitamin A deficiency. Nutrients 2017;9(3):190.
24. Yap IKS, Li JV, Saric J, Martin F-P, Davies H, Wang Y, Wilson ID, Nicholson JK, Utzinger J, Marchesi JR, et al. Metabonomic and microbiological analysis of the dynamic effect of vancomycin-induced gut microbiota modification in the mouse. J Proteome Res 2008;7:3718–28.
25. Madrid-Gambin F, Brunius C, Garcia-Aloy M, Estruel-Amades S, Landberg R, Andres-Lacueva C. Untargeted ^{1}H NMR-based metabolomics analysis of urine and serum profiles after consumption of lentils, chickpeas, and beans: an extended meal study to discover dietary biomarkers of pulses. J Agric Food Chem 2018;66:6997–7005.
26. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E. Microbial degradation of complex carbohydrates in the gut. Gut Microbes 2012;3:289–306.
27. Vital M, Karch A, Pieper DH. Colonic butyrate-producing communities in humans: an overview using omics data. mSystems 2017;2:e00130–17.
28. Hannan LM, Jacobs EJ, Thun MJ. The association between cigarette smoking and risk of colorectal cancer in a large prospective cohort from the United States. Cancer Epidemiol Biomarkers Prev 2009;18:3362–7.
29. Chu KM, Cho CH, Shin VY. Nicotine and gastrointestinal disorders: its role in ulceration and cancer development. Curr Pharm Des 2013;19(1):5–10.
30. Clarke JM, Topping DL, Bird AR, Young GP, Cobiac L. Effects of high-amylose maize starch and butyrylated high-amylose maize starch on azoxymethane-induced intestinal cancer in rats. Carcinogenesis 2008;29:2190–4.
31. Donohoe DR, Holley D, Collins LB, Montgomery SA, Whitmore AC, Hillhouse A, Curry KP, Renner SW, Greenwalt A, Ryan EP, et al. A gnotobiotic mouse model demonstrates that dietary fiber protects against colorectal tumorigenesis in a microbiota- and butyrate-dependent manner. Cancer Discov 2014;4:1387–97.
32. Schulz MD, Atay C, Heringer J, Romrig FK, Schwitalla S, Aydin B, Ziegler PK, Varga J, Reindl W, Pommerenke C, et al. High-fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity. Nature 2014;514:508–12.
33. Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL. Diet-induced extinctions in the gut microbiota compound over generations. Nature 2016;529:212–15.
34. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, Chehoud C, Albenberg LG, Nessel L, Gilroy E, et al. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. Gut 2016;65:63–72.
35. De Filippo C, Di Paola M, Ramazzotti M, Albanese D, Pieraccini G, Banci E, Miglietta F, Cavalieri D, Lionetti P. Diet, environments, and gut microbiota. A preliminary investigation in children living in rural and urban Burkina Faso and Italy. Front Microbiol 2017;8:1979.
36. Beyer-Sehlmeyer G, Glei M, Hartmann E, Hughes R, Persin C, Böhm V, Rowland I, Schubert R, Jahreis G, Pool-Zobel BL. Butyrate is only one of several growth inhibitors produced during gut flora-mediated fermentation of dietary fibre sources. Br J Nutr 2003;90:1057–70.
37. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M, Pudlo NA, Kitamoto S, Terrapon N, Muller A, et al. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. Cell 2016;167:1339–53.e21.
38. Schroeder BO, Birchenough GMH, Ståhlman M, Arike L, Johansson MEV, Hansson GC, Bäckhed F. Bifidobacteria or fiber protects against diet-induced microbiota-mediated colonic mucus deterioration. Cell Host Microbe 2018;23:27–40.e7.
39. Sheng L, Jena PK, Hu Y, Liu H-X, Nagar N, Kalanetra KM, French SW, French SW, Mills DA, Wan Y-JY. Hepatic inflammation caused by dysregulated bile acid synthesis is reversible by butyrate supplementation. J Pathol 2017;243:431–41.
40. Narisawa T, Magadia NE, Weisburger JH, Wynder EL. Promoting effect of bile acids on colon carcinogenesis after intrarectal instillation of *N*-methyl-*N* ′-nitro-*N*-nitrosoguanidine in rats. J Natl Cancer Inst 1974;53:1093–7.
41. Flynn C, Montrose DC, Swank DL, Nakanishi M, Ilsley JNM, Rosenberg DW. Deoxycholic acid promotes the growth of colonic aberrant crypt foci. Mol Carcinog 2007;46:60–70.

42. Cao H, Xu M, Dong W, Deng B, Wang S, Zhang Y, Wang S, Luo S, Wang W, Qi Y, et al. Secondary bile acid-induced dysbiosis promotes intestinal carcinogenesis. Int J Cancer 2017;140:2545–56.
43. Theriot CM, Bowman AA, Young VB. Antibiotic-induced alterations of the gut microbiota alter secondary bile acid production and allow for *Clostridium difficile* spore germination and outgrowth in the large intestine. mSphere 2016;1(1):e00045–15.
44. Ridlon JM, Harris SC, Bhowmik S, Kang D-J, Hylemon PB. Consequences of bile salt biotransformations by intestinal bacteria. Gut Microbes 2016;7:22–39.
45. Islam KBMS, Fukiya S, Hagio M, Fujii N, Ishizuka S, Ooka T, Ogura Y, Hayashi T, Yokota A. Bile acid is a host factor that regulates the composition of the cecal microbiota in rats. Gastroenterology 2011;141:1773–81.
46. Ridlon JM, Alves JM, Hylemon PB, Bajaj JS. Cirrhosis, bile acids and gut microbiota. Gut Microbes 2013;4:382–7.
47. Kakiyama G, Pandak WM, Gillevet PM, Hylemon PB, Heuman DM, Daita K, Takei H, Muto A, Nittono H, Ridlon JM, et al. Modulation of the fecal bile acid profile by gut microbiota in cirrhosis. J Hepatol 2013;58:949–55.
48. Bajaj JS, Heuman DM, Hylemon PB, Sanyal AJ, White MB, Monteith P, Noble NA, Unser AB, Daita K, Fisher AR, et al. Altered profile of human gut microbiome is associated with cirrhosis and its complications. J Hepatol 2014;60:940–7.
49. Baxter NT, Zackular JP, Chen GY, Schloss PD. Structure of the gut microbiome following colonization with human feces determines colonic tumor burden. Microbiome 2014;2:20.
50. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, Fleck JS, Voigt AY, Palleja A, Ponnudurai R, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nat Med 2019;25:679–89.
51. Begley M, Gahan CGM, Hill C. The interaction between bacteria and bile. FEMS Microbiol Rev 2005;29:625–51.
52. Devkota S, Wang Y, Musch MW, Leone V, Fehlner-Peach H, Nadimpalli A, Antonopoulos DA, Jabri B, Chang EB. Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in *Il10*$^{-/-}$ mice. Nature 2012;487:104–8.
53. Yazici C, Wolf PG, Kim H, Cross T-WL, Vermillion K, Carroll T, Augustus GJ, Mutlu E, Tussing-Humphreys L, Braunschweig C, et al. Race-dependent association of sulfidogenic bacteria with colorectal cancer. Gut 2017;66:1983–94.
54. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, Barnes R, Watson P, Allen-Vercoe E, Moore RA, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. Genome Res 2012;22:299–306.
55. Mehta RS, Nishihara R, Cao Y, Song M, Mima K, Qian ZR, Nowak JA, Kosumi K, Hamada T, Masugi Y, et al. Association of dietary patterns with risk of colorectal cancer subtypes classified by *Fusobacterium nucleatum* in tumor tissue. JAMA Oncol 2017;3(7): 921–7.
56. Winter J, Nyskohus L, Young GP, Hu Y, Conlon MA, Bird AR, Topping DL, Le Leu RK. Inhibition by resistant starch of red meat-induced promutagenic adducts in mouse colon. Cancer Prev Res (Phila) 2011;4:1920–8.
57. Le Leu RK, Winter JM, Christophersen CT, Young GP, Humphreys KJ, Hu Y, Gratz SW, Miller RB, Topping DL, Bird AR, et al. Butyrylated starch intake can prevent red meat-induced O^6-methyl-2-deoxyguanosine adducts in human rectal tissue: a randomised clinical trial. Br J Nutr 2015;114:220–30.
58. Redwood D, Lanier AP, Renner C, Smith J, Tom-Orme L, Slattery ML. Differences in cigarette and smokeless tobacco use among American Indian and Alaska Native people living in Alaska and the Southwest United States. Nicotine Tob Res 2010;12:791–6.
59. Dilley JA, Peterson E, Hiratsuka VY, Rohde K. Discovering unique tobacco use patterns among Alaska Native people. Int J Circumpolar Health 2013;72:21208.
60. Caesar R, Tremaroli V, Kovatcheva-Datchary P, Cani PD, Bäckhed F. Crosstalk between gut microbiota and dietary lipids aggravates WAT inflammation through TLR signaling. Cell Metab 2015;22: 658–68.
61. Lin H, An Y, Hao F, Wang Y, Tang H. Correlations of fecal metabonomic and microbiomic changes induced by high-fat diet in the pre-obesity state. Sci Rep 2016;6:21618.
62. Bae S, Ulrich CM, Neuhouser ML, Malysheva O, Bailey LB, Xiao L, Brown EC, Cushing-Haugen KL, Zheng Y, Cheng T-YD, et al. Plasma choline metabolites and colorectal cancer risk in the Women's Health Initiative Observational Study. Cancer Res 2014;74: 7442–52.
63. Xu R, Wang Q, Li L. A genome-wide systems analysis reveals strong link between colorectal cancer and trimethylamine N-oxide (TMAO), a gut microbial metabolite of dietary meat and fat. BMC Genomics 2015;16(Suppl 7):S4.
64. Tang WHW, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, Wu Y, Hazen SL. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med 2013;368:1575–84.
65. Heianza Y, Ma W, Manson JE, Rexrode KM, Qi L. Gut microbiota metabolites and risk of major adverse cardiovascular disease events and death: a systematic review and meta-analysis of prospective studies. J Am Heart Assoc 2017;6(7):e004947.
66. Zackular JP, Rogers MAM, Ruffin MT, Schloss PD. The human gut microbiome as a screening tool for colorectal cancer. Cancer Prev Res (Phila) 2014;7:1112–21.
67. Hale VL, Chen J, Johnson S, Harrington SC, Yab TC, Smyrk TC, Nelson H, Boardman LA, Druliner BR, Levin TR, et al. Shifts in the fecal microbiota associated with adenomatous polyps. Cancer Epidemiol Biomarkers Prev 2017;26:85–94.
68. Mangifesta M, Mancabelli L, Milani C, Gaiani F, de'Angelis N, de'Angelis GL, van Sinderen D, Ventura M, Turroni F. Mucosal microbiota of intestinal polyps reveals putative biomarkers of colorectal cancer. Sci Rep 2018;8:13974.