Elucidating the Gut Microbiome of Colorectal Cancer: 40 Fecal Bacteria as Non-invasive Biomarkers

Biao Yuan

Shanghai East Hospital

SiPing Ma

Liaoning Cancer Institute and Hospital

Qingkai Meng

Liaoning Cancer Institute and Hospital

Tao Du

Shanghai East Hospital

Yueyan Zhu

Shanghai personalbio Co. Ltd.

Jing Yu

Shanghai Personalbio Co.Ltd.

Qinsi Liang

Shanghai personalbio Co. Ltd.

Peicheng Wang

Shanghai personalbio Co.Ltd.

Zikui Sun (✉ sunzikui@personalbio.cn)

Shanghai personalbio Co. Ltd. <https://orcid.org/0000-0001-6143-7263>

Chun Song

Shanghai East Hospital

Research

Keywords: Colorectal adenocarcinoma, non-invasive diagnosis, 16s rRNA sequencing, microbiome, biomarker, machine learning

Posted Date: November 25th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-113328/v1>

License: (cc) (i) This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Colorectal adenocarcinoma (CRC) ranks one of the 5 most lethal malignant tumors both in China and worldwide. Emerging evidences have revealed the importance of gut microbiome on CRC, thus microbial community could be termed as a potential screen for early diagnosis. Importantly, compared with the whole microbial community analysis, few numbers of bacteria genus as non-invasive biomarkers with high sensitivity and specificity causing less cost would benefit more in clinical.

Methods: Here we analyzed the gut microbiome from 226 CRC patients and 156 healthy people by 16s rRNA sequencing. We analyzed the microbiome diversity between CRC patients and healthy controls. We used ExtraTrees classifier to screening the biomarkers and took SVM (Support Vector Machine) model to test the specificity and sensitivity of our biomarkers.

Results: Compared with the healthy gut, the microbial composition are divergent in CRC, especially the increase of some bacteria related to CRC and the decrease of some healthy bacteria. 40 bacteria genus exhibiting high weight for the healthy and CRC microbiome classification were screened as biomarkers for CRC. In addition, the combination of 40 biomarkers and FOBT showed an outstanding sensitivity and specificity for discrimination CRC patients from healthy controls.

Conclusion: The method could be used as a non-invasive method for CRC early diagnosis.

Introduction

As one of the most common gastrointestinal tumors worldwide, colorectal cancer (CRC) ranks third in the world among men and second among women, affecting more than 1.36 million people every year [1]. Most of the CRC patients display no symptoms at early stages and the majority of CRCs develop slowly from adenomatous precursors [2]. It has been estimated that > 95% of colorectal cancer (CRC) would benefit from curative surgery if diagnosed at earlier or intermediate stages [3–6]. Thus, early detection is of vital importance for improving the survival of CRC patients. Conventional screening including barium enema, colonoscopy and sigmoidoscopy, are uncomfortable, invasive, time consuming and expensive [7, 8]. Fecal occult-blood testing (FOBT) and serum carcinoembryonic antigen (CEA) test are non-invasive methods, however, they are compromised by its low specificity [9–13]. More non-invasive screening methods with high specificity, high sensitivity, should be established for early detection of CRC.

Massive efforts in whole-genome sequencing and genome-wide association studies show that genetic factors only explain a small proportion of disease variance [14] and only about 5% cancers occur in the setting of a known genetic predisposition syndrome [15]. It has been established that epigenetic regulation altering gene expression alone, or in combination with inherited or somatic mutation plays important contribution to CRC [16]. As a result, an intensive effort has been undertaken on CRC early diagnosis, which is largely focus on the methylation detection of tumor DNA, or combined with the mutation detection on certain genes, however, methylation detection is compromised by its low sensitivity and specificity as well as the complicated detection [17–19]. More importantly, the epigenetic alteration

can be strongly affected by some environment aspect, including diet habits or chronic alcohol consumption, which also affect human gut microbiota [20].

The gut microbiota maintains survival and metabolism with nutrients in the human body, and works with the human body to respond to external environmental factors, carrying out metabolic and immune activities, as well as maintaining human health [21]. Studies have shown that the changes in the structure and quantity of gut microbiota or the dysfunction are closely related to the damage to human health and colorectal carcinogenesis [22, 23]. Studying the intestinal microbiome composition of colorectal cancer patients can open up new inspection methods for tumor screening. Recent studies have suggested that microbiota profiles determined by high-throughput sequencing may be effective in predicting CRCs [24]. It has been reported that *peptostreptococcus anaerobius*, an anaerobic bacterium enriched in the fecal and mucosal microbiota from CRC patients and promotes CRC [25]. In addition, a series of bacteria, including *Bacteroides fragilis* and a strain of *Escherichia coli* [26–31], *Streptococcus bovis* [32], *Clostridium septicum* [33] and *Fusobacterium nucleatum* [34, 35] have been reported the association with CRC. Furthermore, metagenomic analysis of fecal microbiome has been performed and a couple of gene markers have been identified and validated as biomarkers for early diagnosis of CRC [36]. Difference in gut microbiota between colorectal cancer patients and healthy people combined with other methods such as fecal immunochemical test (FIT)–CEA or other risks factors such as age and BMI index is required for improving accuracy [37, 38].

We evaluated differences in bacterial communities in stool samples of colorectal cancers and non-cancer controls through 16S rRNA high-throughput sequencing. In additions, 40 microbial biomarkers have been identified for CRC early detection. We also evaluated the performance of microbial as non-invasive markers in large cohorts and compared effectiveness between FOBT and microbial. The microbial biomarkers combined with FOBT could be used as non-invasive early diagnosis.

Results

The gut microbiome is dysbiosis in CRC patients

After quality filtering and primer trimming, a total of 5153 usable high-quality sequences reads were generated from 382 samples, the length of which was about 468 bp. In this study, a total of 4728 OTUs were obtained from the colon cancer group and 4331 OTUs from the healthy control group. 3906 OTUs were shared among two groups. Compared with 423 unique OTUs from healthy group, CRC group contained 822 unique OTUs (Figure S1). Rarefaction curves of CRC and control samples showed almost plateaued, suggesting the sequencing was sufficient (Figure S2).

Based on the total OTU statistical sequence, fecal microbial richness, as estimated by ACE and Chao1 (P-values < 0.001, respectively) was significantly decreased in CRC (Fig. 1a, b). The fecal microbial diversity, estimated by shannon and simpson, did not show significance between control and CRC samples (Figure S3). When compared microbiota composition between CRC and healthy gut, beta-diversity exhibited

difference between two group ($p = 0.001$) (Fig. 1c, d). These results suggested the dysbiosis gut microbiome in CRC patients.

The divergent taxonomic composition and functional performance of microbiota in CRC and healthy gut

After quality filtering, sequences at a 97% sequence similarity were selected for taxonomic composition analysis. 21 bacterial phyla, 34 microbial class, 56 microbial orders, 107 microbial families, 209 microbial genera and 268 microbial species have been identified (Table S1).

The LEfSe (Linear discriminant analysis effect size) analysis was performed to determine differences in bacterial taxonomy. The histogram with cladogram showed that phylum Bacteroidetes was overall highly accumulated in CRC while phylum Actinobacteria was overall less accumulated in health samples. Divergent alteration was observed at lower taxonomic levels from phylum Firmicutes and Proteobacteria.

We further compared the difference between control and CRCs microbiome at different level. At the phylum level, four phyla were detected with relatively high abundance. CRC samples increased abundance of phylum Bacteroidetes whilst healthy samples increased abundance of Firmicutes, Proteobacteria and Actinobacteria (Figure S4a). At the family level, for the family with relatively high abundance, CRC samples had higher abundances of Bacteroidaceae, Veillonellaceae and Prevotellaceae whereas healthy samples had higher abundances of Ruminococcaceae, Lachnospiraceae, Enterobacteriaceae and Bifidobacteriaceae (Figure S4b). At the genus level, the most abundant identified genus in healthy samples were Faecalibacterium (10.49%), Bifidobacterium (7.65%) and Bacteroides (7.33%), while in CRC gut, the most abundant genus shifted to Faecalibacterium (6.37%), Bacteroides (21.79%) and Prevotella (5.14%) (Fig. 3a). The most abundant gut microbes, Faecalibacterium, termed as the marker of healthy gut [39], decreased by 60%. Beneficial intestinal bacteria Bifidobacterium [40] decreased by 35% in CRC gut (Fig. 3b). Importantly, several genus has been reported the role in CRC, such as Peptostreptococcus [25], Fusobacterium [34, 35], Porphyromonas, Parvimonas, Gemella and Prevotella [41] were extremely upregulated in CRC gut (Fig. 3c, Table S2). These results strongly suggested the dysbiosis in CRC gut.

We further compared the functional capacity of the gut microbiota between CRC and healthy, transporter pathway, especially the ABC transporter pathway, are significantly increased in CRC gut, and large number of metabolism related pathways, such as vitamin B6 metabolism, energy metabolism, amino sugar and nucleotide sugar metabolism, fructose and mannose metabolism, phosphonate and phosphinate metabolism, pyruvate metabolism, phenylalanine metabolism, D-Glutamine and D-glutamate metabolism, sphingolipid metabolism, and nitrogen metabolism decreased in CRC gut, with the exception of glycerophospholipid metabolism. These results suggested the disorder of metabolism in CRC patients (Fig. 4).

Fecal microbial markers for CRC detection

The changes in the bacterial community between the two groups could be screened as biomarkers for colorectal cancer detection to assist in its diagnosis. To select the most relevant feature which could be term as biomarkers for CRC, the ExtraTrees classifier calculating feature importance score was performed. 40 significantly different features showing different abundance were selected for further analysis (Table 2).

Machine learning classification

To illustrate the diagnostic value of the selected biomarkers in the gut microbiome for colorectal cancer, we constructed a classifier established by SVM (Support Vector Machine) model to detect cancerous samples. We compared both radial basis function (RBF) kernel (Gaussian kernel) and a linear kernel function of SVM to get a better performance. For RBF kernel, the penalty parameter C was varied as {1e-4, 1e-3..., 1e4} and the gamma parameter G as {1e-5, 1e-3..., 1e3}. And for linear kernel, the penalty parameter C was varied as {1e-4, 1e-3..., 1e4}. In total, we examined 90 different parameter combinations to optimize SVM. The SVM model used the 30 biomarkers as shown in table 1 to distinguishing CRC and controls. The Receiver Operating Characteristic (ROC) curve was used to evaluate the diagnostic value of SVM model. The classifier could discriminate colorectal cancer patients from healthy controls with a sensitivity of 81.4%, specificity of 82.2%, precision of 87.7%, and accuracy of 81.7% under the para1 conditions ('C': 100, 'kernel': 'linear') (Table 1). For FOBT test from stool sample has been widely used in diagnosis, we also performed the classifier with the 40 biomarkers together with FOBT test result. The sensitivity, specificity, precision, and accuracy increased into 93.6%, 92.9%, 95.8% and 93.3% (Table 1). Combination of 40 biomarkers and FOBT showed improved diagnostic performance as compared with 40 biomarkers alone, with AUROC from 0.887 to 0.962 (Fig. 5).

Discussion

The gut microbiome plays a major role in protecting the host against the overgrowth of pathogens and sustaining the health of colon. There is intensive evidence revealing the close relationship between gut microbiome and colonic disease, such as colorectal cancer [42–44]. In addition to causing intestinal diseases, gut microbiome is also contribute to obesity, diabetes, allergic asthma and neuropsychiatric diseases [45–47], thus, clinical monitoring of fecal bacteria can assist in the diagnosis of other diseases related to gut microbiome. Furthermore, gut status could be improved by artificially guiding the intervention of diet or the intake of beneficial bacteria according to the changes of gut microbiome [48], and the improvement could be easily detect from the fecal bacteria. Thus, gut microbiome has become a hot spot in the clinical research.

We have performed high-throughput sequencing on the v3-v4 region of intestinal bacteria 16S rRNA gene in stool and described the patterns of gut microbiome relative to health and CRC patients. Fecal richness from colorectal cancer patients decreased, in addition, the proportion of various beneficial bacteria decreased, and the proportion of harmful bacteria significantly increased. A dozen of opportunistic pathogens including *Bacteroides* and *Prevotella* were significantly increased in patients with colorectal cancer. A couple of pathogens, including *Fusobacterium nucleatum* [34, 35], *Peptostreptococcus*

anaerobius and enterotoxigenic *Bacteroides fragilis* [25, 29] which have been established the role in CRC induction, were highly accumulated in CRC patients (Fig. 3; Supplemental table 1). The dysbiosis characteristics could be facilitated to term as taxonomic biomarkers for CRC screening.

We performed machine learning using SVM model between pairs of cohorts to conduct binary classification for classifying CRC and control. A variety of features including taxonomic, functional [49], and k-mer-based [50] classification schemes has been used for machine learning approaches. Here, we used 40 bacteria genus showing a great contribution to differ the CRC state versus control as features for machine learning. In addition, FOBT test result was selected as a feature as well for its importance on CRC diagnosis in clinical. Our machine learning results showed high performances in CRC versus control models (Table 1). The high performance of fecal bacteria and FOBT test from stool sample facilitates to establish a new non-invasive method for examination of colorectal cancer.

Colonoscopy and FOBT are widely used in CRC screening, however, for their low compliance or sensitivity, more non-invasive and painless methods with high sensitivity and specificity are required [7–13]. Circulating tumor DNA (ctDNA) is extracellular DNA originated from tumor cells and circulates in a number of bodily fluids, including blood, synovial fluid and cerebrospinal fluid [51]. For the similarity of genetic and epigenetic information provides by ctDNA to that of invasive tumor biopsies, ctDNA has been widely used to detected the gene mutation and termed as a non-invasive diagnostic tool for several cancers [52]. In many tumors, increased methylation of tumor suppressor genes occurs at an early stage, thus, ctDNA methylation profiling detection can be used for as an alternative non-invasive diagnostic tool [53–55]. Some specific DNA methylation sites, such as SEPT9 have been identified as biomarkers of CRC [56, 57]. However, the extremely low level in blood and the non-organ information of ctDNA gives a great challenge to early diagnosis.

In clinical application, changes of gut microbiome can be regularly monitored, early detection and treatment of CRC can improve the late survival rate and reduce the cost of late treatment. In this study, we monitored the gut microbiome and took 40 bacteria genus displaying high weight for classification between CRC and healthy gut as biomarkers for CRC early diagnose. Combined with FOBT test, our method showed an excellent performance on CRC early diagnose. The method benefit to those who cannot receive colonoscopy in a short time, and those who are not willing to use colonoscopy. Compared with the existing methods of CRC diagnosis, our method is non-invasive and painless, not only does it not require complex examination and preparation before sampling, but also improves the sensitivity and specificity of the test compared with the FOBT alone.

Conclusion

Based on the microbiome composition analysis from CRC and healthy controls, we have selected 40 bacteria genus for classification between CRC and health gut. Combined with the FOBT test, these 40 bacteria exhibited excellent sensitivity and specificity. We have proposed a non-invasive CRC early diagnosis method

Materials And Methods

Study participants and Stool samples collection

Stool samples were collected from 382 individuals undergoing colonoscopy at endoscopy center of Liaoning Cancer Hospital and Dongfang Hospital Affiliated to Tongji University, including 226 CRCs and 156 healthy controls. To avoid potential alternation of the gut microbiota, the exclusion criteria were: (1) a past history of any cancer; (2) use of antibiotics within the past 3 months; (3) had a surgery or an invasive procedure within the past 3 months; (4) had an inflammatory bowel disease. All enrolled subjects were asked to keep a steady dietary and lifestyle and leave fecal sample over 1.0 g in the special containment before bowel preparation for any endoscopy or surgery. After stool collection by the patients, samples were stored at -80°C for further analysis.

16S rRNA gene sequencing

DNA from stool samples was extracted using Qiagen QIAamp DNA Stool Mini Kit (Qiagen) according to manufacturer's instructions. Quality and quantity of extracted DNA were examined by electrophoretic separation in a 0.8% (wt/vol) agarose gel and NanoDrop 2000 spectrophotometer, respectively.

The hypervariable V3-V4 regions of the 16S rRNA gene were amplified using the primer set of 338F (5'-ACTCCTACGGGAGGCAGCA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3'). PCR amplification uses Pfu high-fidelity DNA polymerase from TransGen Biotech, and strictly controls the number of amplification cycles to keep the number of cycles as low as possible, while ensuring the same amplification conditions for the same batch of samples. PCR amplification, purification of amplified product, sequencing library preparation and pyrosequencing were performed at paired-end 250 bp on the Illumina MiSeq platform by Personal Biotechnology, Co., Ltd. (Shanghai, China).

Sequence data processing

Raw sequencing data was processed using Quantitative Insights into Microbial Ecology (QIIME) v1.8.0 [58], and filtered by removing tags and primers. A quality cut-off was applied to discarding the reads (1) that shorter than 150 bp, with (2) an average Phred score lower than 20, (3) with ambiguous bases. After that, the filtered reads were assembled using FLASH software v1.2.7 with overlapping between the paired-end reads > 10. Chimeric sequences were filtered using USEARCH v5.2.236. After quality filtering and chimera removal, clean reads were then clustered into Operational Taxonomic Units (OTUs) at 97% sequence identity using UCLUST. The taxonomic classification was performed with Greengenes database release 13.8. Alpha diversity indices of Chao1, ACE, Simpson and Shannon were estimated. Beta diversity analysis was performed with UniFrac in QIIME. Non-metric multi-dimensional scaling (NMDS) was generated by R language release package for analysis based on distance.

Fecal occult blood test (FOBT)

All enrolled subjects were asked to offer a valid fecal occult blood test report from a community hospital or a general hospital in recent 6 months. Stool samples with blank FOBT result would have to be

examined using Fecal Occult Blood Diagnostic Kit (Colloidal Gold) (Chemtrue®) which has been approved by the Chinese Food and Drug Administration Bureau. The cut-off value for positive FOBT is 200 ng/ml according to manufacturer's instructions.

Statistical analysis

Significant differences among treatments were identified through one-way analysis of variance (ANOVA) followed by Tukey's test. Typically, homogeneity of variance for the obtained data was tested and data of the test values > 0.05 were adopted for the ANOVA analysis. All statistical analyses were performed using SPSS 19.0 (IBM, New York, USA), and significant levels were reported at $p < 0.05$ and $p < 0.01$.

FOBT test results were recorded as positive or negative.

Declarations

Funding

This study was supported by Innovation Fund of Science and Technology Committee in Shanghai Pudong New Area (Nos. PKJ2016-Y60), National Natural Science Foundation of China (Nos. 81871953) and Jiangxi Youth Science Fund (Nos. 20171BAB215043).

Author Contributions

B.Y., S. M. and Q.M. prepared the samples, extracted the DNA for 16s rRNA sequencing and revised the manuscript. T.D. and Z.Y. performed the data analysis and organized the data. J.Y. wrote the manuscript. Q.L. performed the machine learning analysis. P.W. helped the 16s rRNA sequencing and data analysis. Z.S and C.S. designed and led the project.

Ethics approval

Patients and their families in this study have been fully informed. This study was approved by the Ethics Committee of Liaoning Cancer Hospital and Dongfang Hospital Affiliated to Tongji University, and conducted in accordance with the Declaration of Helsinki. All samples received written informed consent from the patients.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author information

¹Department of Gastroenterological Surgery, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 200123, China. ²Department of Colorectal Surgery, Cancer Hospital of China Medical University, Liaoning Cancer Hospital and Institute, Shenyang 110042, China. ³Shanghai Personal Biotechnology Co., Ltd, Shanghai 200231, China.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. 2015;136:E359-86.
2. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. *N Engl J Med*. 2000;343:162-8.
3. Gumaste VV. CT colonography can be an adjunct to optical colonoscopy in CRC screening. *Dig Dis Sci*. 2009;54:212-7.
4. Pawa N, Arulampalam T, Norton JD. Screening for colorectal cancer: established and emerging modalities. *Nat Rev Gastroenterol Hepatol*. 2011;8:711-22.
5. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal cancer statistics. *CA Cancer J Clin*. 2017;67:177-93.
6. Feng RM, Zong YN, Cao SM, Xu RH. Current cancer situation in China: Good or bad news from the 2018 Global Cancer Statistics? *Cancer Commun (Lond)*. 2019;39:22.
7. Sung JJ, Ng SC, Chan FK, Chiu HM, Kim HS, Matsuda T, et al. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. *Gut*. 2015;64:121-32.
8. Rex DK, Johnson DA, Anderson JC, Schoenfeld PS, Burke CA, Inadomi JM, et al. American College of Gastroenterology guidelines for colorectal cancer screening. *Am J Gastroenterol*. 2009;104:739–50.
9. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. *Cochrane Database Syst Rev*. 2007;CD001216.
10. Raginel T, Puvinel J, Ferrand O, Bouvier V, Levillain R, Ruiz A, et al. A population-based comparison of immunochemical fecal occult blood tests for colorectal cancer screening. *Gastroenterology*. 2013;144:918-25.
11. Nicholson BD, Shinkins B, Pathiraja I, Roberts NW, Mant D. Blood CEA levels for detecting recurrent colorectal cancer. *Cochrane Database Syst Rev*. 2015;CD011134.
12. Palmqvist R, Engaras B, Lindmark G, Hallmans G, Tavelin B, Nilsson O, et al. Prediagnostic levels of carcinoembryonic antigen and CA 242 in colorectal cancer: a matched case-control study. *Dis Colon rectum*. 2003;46:1538-44.
13. Macdonald JS. Carcinoembryonic antigen screening: pros and cons. *Semin Oncol*. 1999;26:556-60.

14. Galvan A, Ioannidis JP, Dragani TA. Beyond genome-wide association studies: genetic heterogeneity and individual predisposition to cancer. *Trends Genet.* 2010;26:132-41.
15. Foulkes WD. Inherited susceptibility to common cancers. *N Engl J Med.* 2008; 359:2143-53.
16. Yang T, Owen JL, Lightfoot YL, Kladde MP, Mohamadzadeh M. Microbiota impact on the epigenetic regulation of colorectal cancer. *Trends Mol Med.* 2013;19:714-25.
17. Muller HM, Oberwalder M, Fiegl H, Morandell M, Goebel G, Zitt M, et al. Methylation changes in faecal DNA: a marker for colorectal cancer screening? *Lancet.* 2004;363:1283-5.
18. Mitchell SM, Ho T, Brown GS, Baker RT, Thomas ML, McEvoy A, et al. Evaluation of methylation biomarkers for detection of circulating tumor DNA and application to colorectal cancer. *Genes (Basel).* 2016;7:125.
19. Carmona FJ, Azuara D, Berenguer-Llargo A, Fernandez AF, Biondo S, De Oca J, et al. DNA methylation biomarkers for noninvasive diagnosis of colorectal cancer. *Cancer Prev Res (Phila).* 2013;6:656-65.
20. Kim BC, Shin A, Hong CW, Sohn DK, Han KS, Ryu KH, et al. Association of colorectal adenoma with components of metabolic syndrome. *Cancer Causes Control.* 2012;23:727-35.
21. Jie Z, Xia H, Zhong SL, Feng Q, Li S, Liang S, et al. The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun.* 2017;8:845.
22. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med.* 2019;25:679-89.
23. Cougnoux A, Dalmasso G, Martinez R, Buc E, Delmas J, Gibold L, et al. Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut.* 2014;63:1932-42.
24. 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. *Mol Syst Biol.* 2014;10:766.
25. Long X, Wong CC, Tong L, Chu ESH, Ho Szeto C, Go MY, et al. *Peptostreptococcus anaerobius* promotes colorectal carcinogenesis and modulates tumour immunity. *Nat Microbiol.* 2019;4:2319-30.
26. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan TJ, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science.* 2013;338:120-3.
27. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrede JP. *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci USA.* 2010;107:11537-42.
28. Grivennikov S I, Wang K, Mucida D, Stewart CA, Schnabl B, Jauch D, et al. Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature.* 2012;491:254-8.
29. Toprak NU, Yagci A, Gulluoglu BM, Akin ML, Celenk T, Soyletir G. A possible role of *Bacteroides fragilis* enterotoxin in the aetiology of colorectal cancer. *Clin Microbiol Infect.* 2010;12:782-6.
30. Uronis JM, Mühlbauer M, Herfarth HH, Rubinas TC, Jones GS, Jobin C. Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS ONE.* 2009;4:e6026.

31. Wu S, Rhee KJ, Albesiano E, Rabizadeh S, Wu X, Yen HR, et al. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat Med*. 2009;15:1016-22.
32. Boleij A, Schaeps RM, Tjalsma H. Association between *Streptococcus bovis* and Colon Cancer. *J Clin Microbio*. 2009;47:516.
33. Seder CW, Kramer M, Long G, Uzieblo MR, Shanley CJ, Bove P. *Clostridium septicum* aortitis: Report of two cases and review of the literature. *J Vasc Surg*. 2009;49:1304-9.
34. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res*. 2012;22:299-306.
35. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res*. 2012;22:292-8.
36. Yu J, Feng Q, Wong SH, Zhang D, Liang QY, Qin Y, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut*. 2017;66:70-8.
37. 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. *Mol Syst Biol*. 2014;10:766.
38. Zackular JP, Rogers MA, Ruffin MT, Schloss PD. The Human Gut Microbiome as a Screening Tool for Colorectal Cancer. *Cancer Prev Res (Phila)*. 2014;7:1112-21.
39. Qin J, Li R, Raes J, Arumugam R, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010;464:59-65.
40. Gibson GR, Beatty ER, Wang X, Cummings JH. Selective stimulation of bifidobacteria in the human colon by oligofructose and inulin. *Gastroenterology*. 1995;108:975-82.
41. Baxter NT, Ruffin MT, Rogers MA, Schloss PD. Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. *Genome Med*. 2016;8:37.
42. Fong W, Li Q, Yu J. Gut microbiota modulation: a novel strategy for prevention and treatment of colorectal cancer. *Oncogene* 2020;39:4925-43.
43. Hold GL, Garrett WS. Microbiota organization—a key to understanding CRC development. *Nat Rev Gastroenterol Hepatol*. 2015;12:128-9.
44. Ray K. Gut microbiota: Oral microbiome could provide clues to CRC. *Nat Rev Gastroenterology Hepatol*. 2017;14:690.
45. Nielsen DS, Krych L, Buschard K, Hansen CH, Hansen AK. Beyond genetics. Influence of dietary factors and gut microbiota on type 1 diabetes. *FEBS Lett*. 2014;588:4234-43.
46. Serino M, Blasco-Baque V, Nicolas S, Burcelin R. Far from the eyes, close to the heart: dysbiosis of gut microbiota and cardiovascular consequences. *Curr Cardiol Rep*. 2014;16:540.
47. Arrieta MC, Finlay B. The intestinal microbiota and allergic asthma. *J Infect*. 2014;69:S53-5.
48. Lee JY, Chu SH, Jeon JY, Lee MK, Park JH, Lee DC, et al. Effects of 12 weeks of probiotic supplementation on quality of life in colorectal cancer survivors: A double-blind, randomized, placebo-controlled trial. *Dig Liver Dis*. 2014;46:1126-32.

49. Ning J, Beiko RG. Phylogenetic approaches to microbial community classification. *Microbiome*. 2015;3:47.
50. Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat* 2017;8:1784.
51. Cheng F, Su L, Qian C. Circulating tumor DNA: a promising biomarker in the liquid biopsy of cancer. *Oncotarget*. 2016;7:48832-41.
52. Berger AW, Schwerdel D, Welz H, Marienfeld R, Schmidt SA, Kleger A, et al. Treatment monitoring in metastatic colorectal cancer patients by quantification and KRAS genotyping of circulating cell-free DNA. *PLoS One*. 2017;12:e0174308.
53. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer*. 2011;11:726-34.
54. Baylin SB, Jones PA. Epigenetic determinants of cancer. *Cold Spring Harb. Perspect Biol*. 2016;8:a019505.
55. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet*. 2009;41:178-86.
56. Warren JD, Xiong W, Bunker AM, Vaughn CP, Furtado LV, Roberts WL, et al. Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. *BMC Med*. 2011;9:133.
57. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut*. 2014;63:317-25.
58. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7:335-6

Tables

Due to technical limitations, table 1 and table 2 are only available as a download in the Supplemental Files section.

Supplemental Note

Supplementary Figure S3 is not available in this version of the manuscript.

Figures

Figure 1

Comparison of gut microbiome between CRC and healthy gut. a, Alpha diversity communities based on observed OTUs by richness (Chao1, ACE). b, Beta diversity measured by unweighted unifrac. c, Beta diversity measured by weighted unifrac. Overall differences in the microbiome composition among groups were assessed by ANOSIM.

Figure 1

Comparison of gut microbiome between CRC and healthy gut. a, Alpha diversity communities based on observed OTUs by richness (Chao1, ACE). b, Beta diversity measured by unweighted unifrac. c, Beta diversity measured by weighted unifrac. Overall differences in the microbiome composition among groups were assessed by ANOSIM.

Figure 1

Comparison of gut microbiome between CRC and healthy gut. a, Alpha diversity communities based on observed OTUs by richness (Chao1, ACE). b, Beta diversity measured by unweighted unifrac. c, Beta diversity measured by weighted unifrac. Overall differences in the microbiome composition among groups were assessed by ANOSIM.

Figure 1

Comparison of gut microbiome between CRC and healthy gut. a, Alpha diversity communities based on observed OTUs by richness (Chao1, ACE). b, Beta diversity measured by unweighted unifrac. c, Beta diversity measured by weighted unifrac. Overall differences in the microbiome composition among groups were assessed by ANOSIM.

Figure 2

Linear discriminant analysis effect size (LEfSe) analysis of gut microbiota composition between two groups.

Figure 2

Linear discriminant analysis effect size (LEfSe) analysis of gut microbiota composition between two groups.

Figure 2

Linear discriminant analysis effect size (LEfSe) analysis of gut microbiota composition between two groups.

Figure 2

Linear discriminant analysis effect size (LEfSe) analysis of gut microbiota composition between two groups.

Figure 3

Analysis of genus level and selected genus between CRC and control. a, The distribution of two groups at genus level. b, Abundance of benefit intestinal bacteria between two groups. c, Abundance of CRC related intestinal bacteria between two groups.

Figure 3

Analysis of genus level and selected genus between CRC and control. a, The distribution of two groups at genus level. b, Abundance of benefit intestinal bacteria between two groups. c, Abundance of CRC related intestinal bacteria between two groups.

Figure 3

Analysis of genus level and selected genus between CRC and control. a, The distribution of two groups at genus level. b, Abundance of benefit intestinal bacteria between two groups. c, Abundance of CRC related intestinal bacteria between two groups.

Figure 3

Analysis of genus level and selected genus between CRC and control. a, The distribution of two groups at genus level. b, Abundance of benefit intestinal bacteria between two groups. c, Abundance of CRC related intestinal bacteria between two groups.

Figure 4

Pathways altered in CRC and healthy guts.

Figure 4

Pathways altered in CRC and healthy guts.

Figure 4

Pathways altered in CRC and healthy guts.

Figure 4

Pathways altered in CRC and healthy guts.

Figure 5

The comparison of ROC for fecal microbial markers alone and fecal microbial markers plus FOBT. Red line: AUC for fecal microbial markers alone is 0.911. Blue line: AUC for fecal microbial markers plus FOBT is 0.968.

Figure 5

The comparison of ROC for fecal microbial markers alone and fecal microbial markers plus FOBT. Red line: AUC for fecal microbial markers alone is 0.911. Blue line: AUC for fecal microbial markers plus FOBT is 0.968.

Figure 5

The comparison of ROC for fecal microbial markers alone and fecal microbial markers plus FOBT. Red line: AUC for fecal microbial markers alone is 0.911. Blue line: AUC for fecal microbial markers plus FOBT is 0.968.

Figure 5

The comparison of ROC for fecal microbial markers alone and fecal microbial markers plus FOBT. Red line: AUC for fecal microbial markers alone is 0.911. Blue line: AUC for fecal microbial markers plus FOBT is 0.968.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Figs1.pdf](#)
- [Figs1.pdf](#)
- [Figs1.pdf](#)

- Figs1.pdf
- Figs2.pdf
- Figs2.pdf
- Figs2.pdf
- Figs2.pdf
- Figs4.pdf
- Figs4.pdf
- Figs4.pdf
- Figs4.pdf
- renamed67b1f.xlsx
- renamed67b1f.xlsx
- renamed67b1f.xlsx
- renamed67b1f.xlsx
- renamed92f53.xlsx
- renamed92f53.xlsx
- renamed92f53.xlsx
- renamed92f53.xlsx
- Table2.pdf
- Table2.pdf
- Table2.pdf
- Table2.pdf
- Table1.pdf
- Table1.pdf
- Table1.pdf
- Table1.pdf