Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2026**; 31(11): 56530
https://doi.org/10.31083/FBL56530

Original Research

Compartment-Specific Microbiome Signatures Distinguish Colorectal Cancer From Adenomatous Polyps

Mohammad Fayaz[1], Alessio Rotelli[2], Edda Russo[3], Leandro Di Gloria[4], Giulia Nannini[2,3], Sara Bertorello[2,3], Amedeo Amedei[2,3,5,*]

[1]Department of Health and Social Medicine, School of Medicine, Shahed University, 3319118651 Tehran, Iran
[2]Laboratorio Congiunto MIA-LAB (Microbiome-Immunity Axis Research for a Circular Health), University of Florence, 50134 Florence, Italy
[3]Department of Experimental and Clinical Medicine, University of Florence, 50139 Florence, Italy
[4]Department of Biomedical, Experimental and Clinical "Mario Serio", University of Florence, 50134 Florence, Italy
[5]Network of Immunity in Infection, Malignancy and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), 50121 Florence, Italy
*Correspondence: amedeo.amedei@unifi.it (Amedeo Amedei)
Academic Editor: Wei-Lin Jin
Submitted: 30 July 2026 Revised: 27 September 2026 Accepted: 23 October 2026 Published: 20 November 2026

Abstract

Background: Colorectal cancer (CRC) arises through the adenoma–carcinoma sequence and is accompanied by remodeling of the intestinal microbiome. Dysbiosis may contribute to malignant transformation by sustaining inflammation, weakening epithelial barrier integrity, reshaping antitumor immunity, and altering microbial metabolism. However, reported microbial biomarkers remain inconsistent, and their biological and clinical significance is uncertain. We therefore sought to identify stable, compartment-specific microbial signatures that distinguish CRC from adenomatous polyps and to define their relevance to colorectal tumorigenesis. **Methods:** We analyzed a dataset including 46 patients with CRC and 15 with adenomatous polyps. Microbial profiles obtained by 16S rRNA sequencing from colorectal tumor biopsies, stool, and saliva were examined across five taxonomic levels. Complementary feature-selection approaches were integrated with Bayesian Model Averaging to retain taxa that remained informative across analytical assumptions. Differential abundance and model-interpretation analyses were used to assess the consistency and biological relevance of the selected microbial signals. **Results:** Colorectal tissue biopsies provided the strongest discrimination between CRC and adenoma, underscoring the biological relevance of the local tumor–microbiome interface. *Fusobacterium* was the most consistently identified CRC-associated genus, showing consistent evidence of inclusion across models and significant differential abundance. Additional signatures involved Actinobacteriota, Synergistota, Bacteroidota, Firmicutes, and Proteobacteria, together with genera associated with butyrate metabolism and inflammation. Collectively, these signatures converge on biological processes relevant to epithelial barrier dysfunction, immune modulation, chronic inflammation, and metabolic reprogramming. The marked differences among tissue, stool, and saliva further provide evidence that sample compartment strongly influences the biological signal captured. **Conclusions:** The tumor-associated microbiome emerges as a biologically relevant correlate of the differences between adenomatous polyps and CRC, with *Fusobacterium* representing the most robust candidate biomarker in this cohort. These findings support further development of compartment-aware microbial biomarkers to distinguish adenoma from CRC, while highlighting host–microbiome pathways that may inform future preventive or therapeutic strategies. Validation in larger, independent cohorts and functional studies is needed before clinical translation.

Keywords: gastrointestinal microbiome; colorectal neoplasms; adenomatous polyps; *Fusobacterium*; biomarkers; bayes theorem

1. Introduction

Colorectal cancer (CRC) is a major global health burden, ranking as the third most commonly diagnosed cancer worldwide. It is estimated to account for approximately 1.9 million new cases annually, representing around 10% of all newly diagnosed cancers [1]. The adenoma–carcinoma sequence is driven by the progressive accumulation of genetic and epigenetic alterations involving Adenomatous Polyposis Coli (APC)/Wingless-related integration site (Wnt), Kirsten Rat Sarcoma Viral Oncogene Homolog (KRAS), Tumor Protein 53 (TP53), and Transforming Growth Factor Beta (TGF-β) signaling pathways. Increasing evidence also indicates that these epithelial alterations occur within a dynamic tumor microenvironment characterized by immune activation, inflammatory signaling, stromal remodeling, and host–tumor interactions that accompany colorectal carcinogenesis [2,3].

In the healthy colon, resident microorganisms contribute to epithelial barrier integrity and immune homeostasis through multiple mechanisms, including the production of short-chain fatty acids (SCFAs), particularly butyrate, a major microbial metabolite that supports epithelial metabolism and mucosal immune regulation [4,5,6,7]. Consistent with this concept, one of the earliest microbiome studies in CRC demonstrated that patients with colorectal cancer exhibit a distinct microbial community compo-

Copyright: © 2026 The Author(s). Published by IMR Press.
This is an open access article under the CC BY 4.0 license.

Publisher's Note: IMR Press stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

sition compared with healthy individuals, supporting the existence of a cancer-associated dysbiosis characterized by alterations in dominant bacterial taxa rather than changes in total bacterial load [8]. Loss of this ecological balance (dysbiosis) promotes chronic inflammation, impairs epithelial barrier integrity, increases intestinal permeability, alters microbial metabolic activity, and remodels host immune responses, collectively creating a pro-inflammatory and genotoxic microenvironment that favors colorectal tumor initiation and progression [9,10].

Within this altered ecosystem, *Fusobacterium nucleatum* has emerged as a key mechanistic link between microbial dysbiosis and colorectal tumor progression. Its adhesin FadA binds to E-cadherin and activates β-catenin signaling, thereby promoting oncogenic and inflammatory responses, whereas the outer membrane protein Fap2 binds the inhibitory receptor TIGIT on natural killer and T cells, suppressing antitumor immunity and facilitating immune evasion [11,12]. Conversely, depletion of butyrate-producing commensals, including *Faecalibacterium prausnitzii* and members of the Lachnospiraceae and Ruminococcaceae families, may reduce butyrate availability, thereby impairing intestinal immune homeostasis and diminishing the anti-inflammatory functions normally supported by these microbial communities [13,14,15]. These contrasting changes, specifically, expansion of tumor-associated pathobionts and loss of protective microbial functions, provide a biologically coherent model of CRC-associated dysbiosis.

The microbial profile observed in CRC also depends on the anatomical compartment analyzed. Tissue biopsies characterize bacteria associated with the epithelial and tumor interface, whereas stool samples reflect the broader luminal microbiota and offer a readily accessible source for non-invasive biomarker discovery. Saliva provides complementary information on the oral microbiome, including taxa with the potential to translocate to or influence the intestinal environment. The recurrent detection of genera such as *Fusobacterium*, *Parvimonas*, *Prevotella*, *Porphyromonas*, and *Peptostreptococcus* across oral, fecal, and tumor samples is consistent with the proposed oral–gut microbial axis in CRC, although the mechanisms underlying this association remain under investigation [16,17,18,19].

Clinical translation nevertheless remains limited by the instability of reported microbial biomarkers. Microbiome data are sparse, compositional, highly dimensional, and sensitive to preprocessing and model choice; signatures derived from a single analytical pipeline may therefore be difficult to reproduce [20,21,22,23]. Analytical robustness is relevant, but its value lies in distinguishing analytically concordant biological signals from method-dependent findings.

In this study, we compared 16S rRNA profiles from colorectal tissue biopsies, stool, and saliva in patients with CRC or adenomatous polyps. We used complementary feature-selection and Bayesian integration strategies to identify taxa that remained informative across analytical assumptions, then interpreted these signatures in the context of epithelial barrier function, immune regulation, microbial metabolism, and the oral–gut axis.

The present study represents a secondary analysis of the cohort previously described by our group [16]. The original study characterized microbial differences between CRC and adenomatous polyps across saliva, tissue, and stool, reporting, among other findings, enrichment of *Fusobacterium* in CRC tissue, alterations in fecal microbial composition, and a correlation between tissue *Fusobacterium* and fecal *Parvimonas*. It also investigated predicted microbial functions, fecal metabolomic profiles, and microbiome variation according to CRC stage. The main aim of the present analysis is therefore not to re-establish these previously reported associations but to determine which microbial signals remain reproducibly informative when evaluated through complementary statistical and machine-learning frameworks. Specifically, by integrating multiple feature-selection approaches, consensus prioritization, Bayesian Model Averaging, differential-abundance testing, and sensitivity analyses across taxonomic levels and anatomical compartments, we assess the robustness and, notably, the discriminatory relevance of previously observed and additional microbial signals. In detail, this framework extends the original descriptive and association-based analysis by quantifying the stability of candidate biomarkers across analytical assumptions and by establishing a compartment-specific hierarchy of discriminatory information between tissue, stool, and saliva.

Our primary objective was to define the biological and potential clinical relevance of compartment-specific microbial changes associated with CRC.

2. Materials and Methods

2.1 Study Design and Analytical Overview

This study represents a secondary computational analysis of previously generated, and publicly available 16S rRNA gene-sequencing data to compare microbial communities in patients with CRC and adenomatous polyps (AP). No new patient recruitment, sample collection, DNA extraction, Polymerase Chain Reaction (PCR) amplification, or sequencing was performed as part of the present study. The design was cross-sectional: AP and CRC were treated as clinically distinct stages of colorectal neoplasia, but the study did not follow individuals longitudinally through progression. The primary endpoint was the identification of stable microbial features that discriminated CRC from AP across biological compartments.

Analyses were conducted in R (v4.3.1, R Foundation for Statistical Computing, Vienna, Austria). The main workflow comprised taxonomic preprocessing, complementary feature selection, consensus-based prioritization, Bayesian Model Averaging (BMA), performance assessment, differential abundance testing, and biological inter-

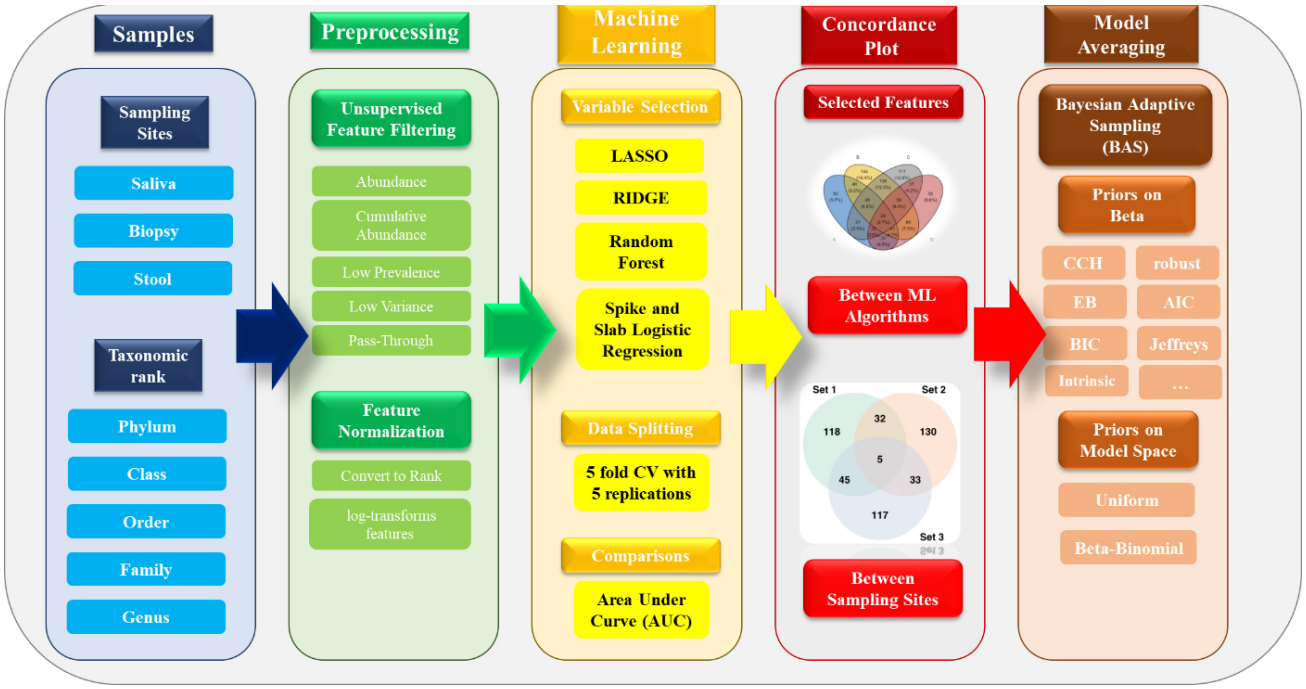


Fig. 1. Study design and analytical framework. Microbial profiles from tissue biopsy, stool, and saliva were compared between colorectal cancer (CRC) and adenomatous polyps. Complementary feature-selection methods, consensus prioritization, Bayesian Model Averaging (BMA), differential abundance testing, and biological interpretation were used to identify compartment-specific microbial signatures. Detailed analytical specifications are provided in **Supplementary A,B**.

pretation (Fig. 1). Analytical definitions, software packages, prior structures, and sensitivity results are provided in **Supplementary A,B**.

2.2 Patient Cohort, Sampling, and Sequencing

The cohort comprised 61 patients: 46 with CRC and 15 with adenomatous polyps (Table 1). All sequencing data analyzed in the present study were retrieved from the Gene Expression Omnibus (GEO) repository under accession number GSE217490, corresponding to the previously published study [16].

These compartments were strategically selected since they represent complementary microbial niches. Biopsies capture communities in direct contact with the colorectal mucosa and tumor microenvironment; stool reflects luminal dysbiosis across the colon and offers a non-invasive biomarker source; saliva represents the oral microbial reservoir and provides information on the oral–gut axis. Comparisons were performed separately by compartment to avoid treating these biologically distinct ecosystems as interchangeable.

Sample collection, DNA extraction, sequencing, and sequence-processing procedures were performed in the original study [16] and are summarized below to describe the provenance and preprocessing of the publicly available data. These procedures were not repeated for the present secondary analysis. In the original study, genomic DNA was extracted using the DNeasy PowerLyzer PowerSoil Kit (Qiagen), and DNA quality and quantity were assessed using NanoDrop ND-1000 and Qubit fluorometry, respectively. The V3–V4 region of the bacterial 16S rRNA gene was amplified using primers 341F and 805R and sequenced in paired-end mode (2 × 300 bp) on an Illumina MiSeq platform. Demultiplexed reads were processed in QIIME2 (v2021.4, QIIME 2 development team, open-source software); primers were removed using Cutadapt, and DADA2 was used for quality filtering, paired-end merging, trimming of low-quality nucleotides, chimera

Table 1. Baseline demographic and clinical characteristics of the study cohort.

Characteristic	CRC (n = 46)	AP (n = 15)
Sex, M/F	33/13	7/8
Median age, years	70.8	62.2
Age range, years	36–85	40–80
Proximal location	14	10
Distal location	32	5
Diet	Mediterranean	Mediterranean
CRC T stage, T0/T1/T2/T3/T4	1/13/21/10/1	NA
Adenoma characteristics	NA	Tubulovillous adenoma with dysplasia, not endoscopically removable

M, Male; F, Female; NA, Not Applicable.

removal, and Amplicon Sequence Variant (ASV) inference. Taxonomic assignment was performed against SILVA 138 using V-Search at a 99% identity threshold. For tissue samples, ASVs not assigned to the prokaryotic domain were excluded before downstream analysis because they were considered likely host-derived sequences in this low-biomass compartment. Rarefaction curves were additionally examined to assess sequencing saturation, as described in the original study. To reduce potential treatment-related microbiome confounding, the original study excluded patients who had received antibiotics, immunosuppressive drugs, or regular probiotics during the preceding two months, as well as patients with acute gastrointestinal infection during the preceding month, previous chemoradiotherapy, immunodeficiency, chronic inflammatory bowel disease, or other established malignancies. Samples were collected at the time of surgery and before treatment. Unstimulated saliva and stool were collected on the morning of surgery, whereas fresh tumor or adenoma tissue was obtained under sterile conditions during surgery. All samples were immediately frozen and stored at −80 °C until DNA extraction, as previously described [16].

2.3 Statistical Analysis and Biomarker Prioritization

Low-information taxa were filtered using abundance-, cumulative-abundance-, prevalence-, pass-, and variance-based criteria implemented in SIAMCAT, followed by log-unit or rank-unit normalization [24]. Multiple preprocessing combinations were retained to evaluate whether biological signals were robust to plausible analytical choices.

Feature selection and classification were performed using least absolute shrinkage and selection operator (LASSO) and ridge logistic regression, random forest, and spike-and-slab Bayesian logistic regression. Models were evaluated for each compartment and taxonomic level using five-fold cross-validation with five repetitions, and the area under the receiver operating characteristic curve (AUC) was averaged across repetitions. Different filtering and normalization methods were compared, and the configuration with the highest mean area under the receiver operating characteristic curve was selected for each model, compartment, and taxonomic level. All AUC values are provided in **Supplementary A**. These analyses were considered exploratory and complementary.

Taxa identified by at least two feature-selection methods were prioritized as consensus features and included in the logistic-regression Bayesian Model Averaging analysis. Posterior inclusion probabilities were evaluated across alternative coefficient and model-space priors to assess feature-selection uncertainty across different Bayesian model specifications. Detailed Bayesian Model Averaging results, including prior-specific predictive performance and calibration measures, are provided in **Supplementary A**, with the corresponding methodological definitions and procedures described in **Supplementary B.4**.

Differential abundance between CRC and AP was assessed with zero-inflated beta GAMLSS models [25], with false-discovery rate (FDR) control; p values and FDR-adjusted p values below 0.10 were considered significant. Variable-importance, partial-dependence, LIME, and SHAP analyses were used as secondary tools to examine the direction and consistency of selected features; **Supplementary A** provides the complete interpretability outputs, including all variable-importance plots, LIME outputs, and SHAP plots.

Zero-Inflated Generalized Dirichlet Multinomial (ZIGDM) tests were performed for each sampling site and taxonomic level, with an FDR threshold set at 0.05. Two sets of tests were conducted: (1) Non-Zero-Inflated ZIGDM, corresponding to the standard GDM model, and (2) Zero-Inflated ZIGDM, applied to taxa with a minimum of 10 zero counts. For each set, three types of outcomes were assessed: (i) differential mean, (ii) differential dispersion, and (iii) differential presence–absence frequency. All analyses were performed using the miLineage R package [26]. The results of ZIGDM analysis are provided in **Supplementary A**.

Analyses were conducted separately for tissue, stool, and saliva across five taxonomic levels, with additional integrated analyses combining all three compartments presented in **Supplementary A** to assess the consistency of microbial associations across sampling sites.

3. Results

3.1 Compartment-Specific Microbial Signals Distinguish CRC From Adenoma

First, the microbial discrimination between CRC and AP differed substantially by anatomical compartment (Table 2). Tissue biopsies contained the strongest disease-associated signal, stool showed intermediate discrimination, and saliva showed weaker but recurrent microbial differences. This gradient is consistent with increasing biological distance from the colorectal lesion.

In tissue biopsies, AUC values reached 0.94. Fusobacteriota and its lower taxonomic lineages (Fusobacteriia, Fusobacteriales, Fusobacteriaceae, and *Fusobacterium*) were repeatedly prioritized, together with Actinobacteriota, Synergistota, and selected genera including *Serratia*. The recurrence of the Fusobacterium lineage across taxonomic resolutions suggests a coherent tumor-associated signal rather than an isolated genus-level finding.

Stool models achieved moderate discrimination (AUC 0.76–0.86). Recurrent features included Bacteroidota, Firmicutes, Proteobacteria, Desulfovibrionales, Lachnospirales, Lachnospiraceae, and *Bacteroides*. Differential abundance also involved *Faecalibacterium*, the *Ruminococcus gnavus* group, *Dorea*, and other genera linked to butyrate metabolism, mucus utilization, and intestinal inflammation.

Finally, the salivary profiles were less discriminative (AUC 0.64–0.72) but retained signals involving Fir-

Table 2. Principal compartment-specific microbial signatures distinguishing CRC from adenomatous polyps.

Compartment	Discrimination	Key microbial signals	Literature based biological interpretation
Tissue biopsy	Strongest; AUC up to 0.94	Fusobacteriota/*Fusobacterium*; Actinobacteriota; Synergistota; Fusobacteriaceae; *Serratia*	Local tumor–microbiome interface; epithelial adhesion, immune modulation, inflammation, and community remodeling
Stool	Moderate; AUC 0.76–0.86	Bacteroidota; Firmicutes; Proteobacteria; *Bacteroides*; *Faecalibacterium*; *R. gnavus* group	Luminal dysbiosis involving butyrate metabolism, mucus utilization, and inflammatory ecology
Saliva	Lower; AUC 0.64–0.72	Firmicutes; Fusobacteriota; Proteobacteria; Clostridia-related taxa; Saccharimonadia; *Alloscardovia*	Oral microbial reservoir and possible oral–gut connection; lower specificity for the colorectal lesion

Detailed taxon lists, preprocessing combinations, posterior inclusion probabilities, and model-specific performance metrics are provided in **Supplementary A**.

micutes, Fusobacteriota, Proteobacteria, Clostridia-related taxa, Saccharimonadia, and *Alloscardovia*. These findings suggest that saliva captures an upstream oral ecological signal rather than the local tumor-associated community.

3.2 Fusobacterium Defines the Tumor-Associated Microbial Signature

Fusobacterium was the most consistently identified CRC-associated taxon in tissue biopsies (Table 2). Its lineage was detected at multiple nested taxonomic ranks, from phylum to genus. Because these taxonomic levels are hierarchically related and derived from the same sequencing data, this recurrence should be interpreted as taxonomic consistency rather than independent validation. Stronger evidence of analytical robustness was provided by the repeated prioritization of the Fusobacterium signal across distinct feature-selection approaches, preprocessing strategies, Bayesian model specifications, and differential-abundance analyses.

This cross-level and cross-model concordance is crucial since it links statistical reproducibility to a well-established biological pathway at the tumor surface, where *Fusobacterium* species can interact with epithelial and immune cells.

Additional tissue-associated features included Actinobacteriota, Synergistota, Proteobacteria, and taxa within Gemellaceae, Pseudomonadaceae, and several butyrate-associated genera. Their coexistence indicates that the CRC tissue microbiome is characterized by community remodeling rather than by a single organism. Extended consensus feature lists are provided in **Supplementary A**.

In Table 3, we summarized the features shared across selection methods, whereas Table 2 condenses the principal compartment-specific biological findings.

3.3 Stool Signatures Indicate Altered Microbial Homeostasis

The stool microbiome showed a broader, less lesion-specific pattern than tissue. Changes involving *Faecalibacterium* and Lachnospiraceae are consistent with altered butyrate-producing capacity, whereas signals involving the *R. gnavus* group, *Bacteroides*, Proteobacteria, and Desulfovibrionales are compatible with mucus degradation, inflammatory remodeling, and altered luminal metabolism. Since functional pathways were not measured directly, these interpretations represent biologically plausible associations rather than demonstrated metabolic effects.

3.4 Salivary Signatures Support an Oral–Gut Connection

Salivary associations included Clostridia-related taxa, Saccharimonadia, Firmicutes, Fusobacteriota, and Proteobacteria. Their lower predictive performance compared with biopsy and stool suggests limited disease specificity; nevertheless, recurrent oral taxa across compartments support the possibility that the oral cavity acts as a reservoir or marker of microbial changes associated with CRC.

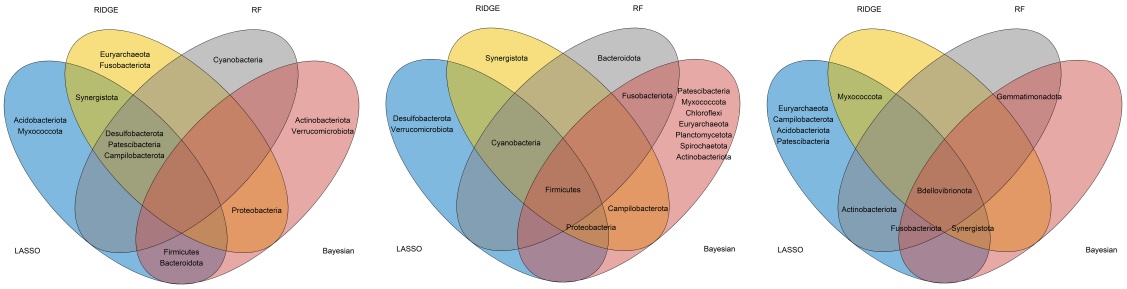
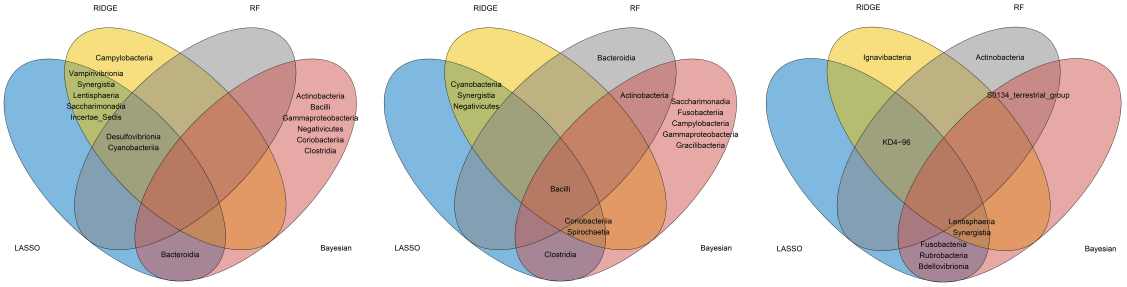
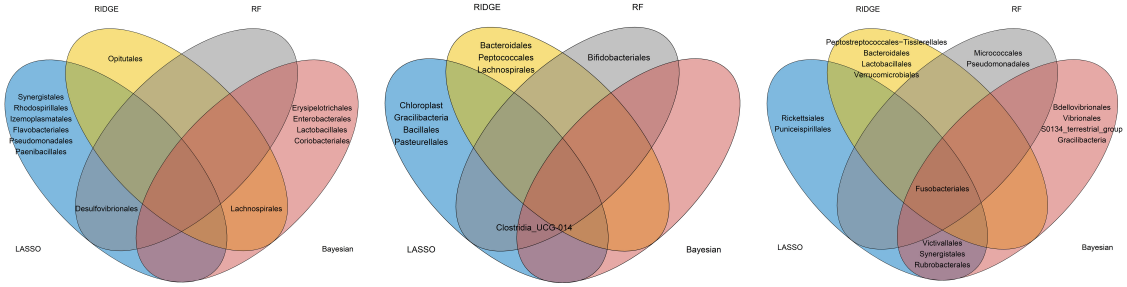
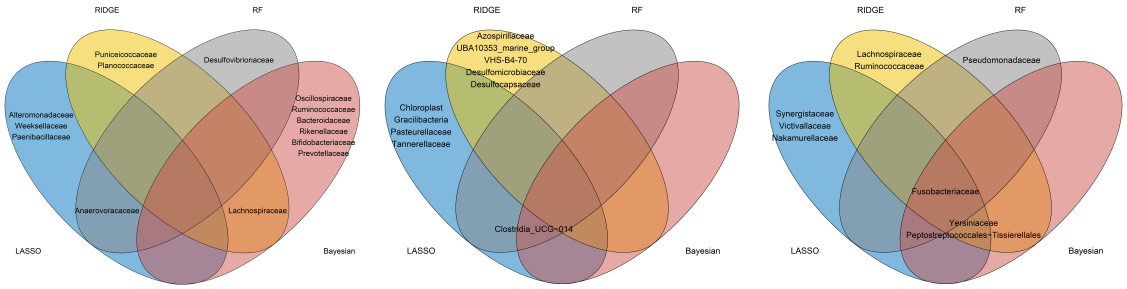
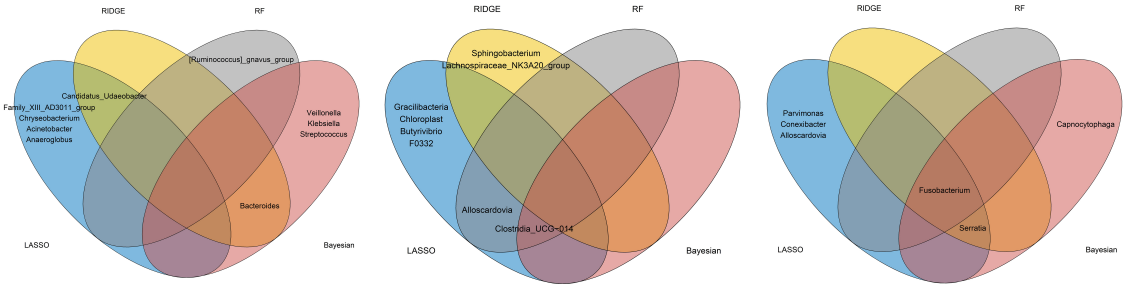
3.5 Differential Abundance and Analytical Robustness Converge on Biological Signals

Differential abundance analyses reinforced the compartment-specific pattern (Table 4). The clearest signal was the enrichment of the *Fusobacterium* lineage in tissue biopsies, while stool and saliva showed more distributed ecological changes. The consistency of these findings across feature selection, BMA, and differential abundance increased confidence in their biological relevance.

Most prior specifications produced similar discrimination, whereas a small number of settings performed less well. Prior-specific AUC, calibration, log-score, Brier-score, posterior inclusion results, and an interpretability summary are reported in **Supplementary A**. These analyses support the stability of the main findings but are not required to interpret the biological conclusions.

Together, the results define a hierarchy of microbial information: tumor tissue captures the most specific CRC-associated community, stool mirrors broader intestinal dysbiosis, and saliva provides a weaker but potentially informative oral signature.

Table 3. Consensus microbial signatures across biological compartments.

Taxa	Sampling sites		
	Stool	Saliva	Biopsy
Phylum			
Class			
Order			
Family			
Genus			

Overlap among complementary selection methods is shown by taxonomic level for tissue biopsy, stool, and saliva. Central intersections identify taxa that remained informative across analytical approaches and were prioritized for biological interpretation.

4. Discussion

This study identifies a compartment-dependent remodeling of the microbiome in CRC, with the strongest and most stable signal located at the colorectal tissue interface. The main biomedical finding is not simply that microbial classifiers can distinguish CRC from adenoma, but that the discriminating taxa converge on processes central to colorectal tumor biology: epithelial adhesion and barrier

Table 4. Differential microbial signals and their biological interpretation by sampling compartment.

Compartment	Differential signals	Biological context	Interpretive boundary
Tissue biopsy	Fusobacteriota, Fusobacteriia, Fusobacteriales, *Fusobacterium*; additional tissue-associated genera	Tumor-adherent community associated with epithelial interaction, immune evasion, and inflammation	16S data do not resolve species-specific causality
Stool	*Faecalibacterium*, *R. gnavus* group, *Dorea*, *Bacteroides*, Proteobacteria-related taxa	Loss of microbial homeostasis, altered butyrate-associated ecology, mucus degradation, and luminal inflammation	Metabolic functions were inferred, not directly measured
Saliva	Clostridia-related taxa, Saccharimonadia, Firmicutes, Fusobacteriota, Proteobacteria	Oral ecological changes compatible with an oral–gut microbial axis	Lower discrimination limits use as a tumor surrogate

disruption, immune evasion, chronic inflammation, altered microbial metabolism, and oral–gut microbial trafficking. Notably, different associations observed here, mainly the enrichment of *Fusobacterium* in CRC tissue, were already reported in the original analysis of the same cohort [16] and should not be considered independent biological discoveries. The key contribution of the present study is instead to assess the analytical stability and discriminatory value of these microbial signals across complementary feature-selection methods, preprocessing strategies, Bayesian models, taxonomic levels, and anatomical compartments.

In detail, the tissue samples showed the strongest discriminatory performance, followed by stool and saliva. Notably, the V3–V4 16S rRNA data support taxonomic identification primarily at the genus level and do not permit reliable species-level attribution. Therefore, the reported species, including *Fusobacterium nucleatum*, are cited solely on the basis of previous literature to provide biological context for the corresponding genus-level findings and were not directly identified in our analysis. Likewise, species-specific metabolic functions, virulence factors, and host-interaction mechanisms discussed herein are literature-based and not directly assessed in this study. *Fusobacterium* was the dominant tissue-associated signal, whereas stool mirrored a broader loss of ecological homeostasis and saliva captured a weaker oral signature. This hierarchy across compartments explains why biopsy-derived profiles were most accurate and why stool and saliva should be viewed as complementary, rather than equivalent, biomarker sources.

In our analyses, the Fusobacterium-associated signal was observed across multiple nested taxonomic ranks, from phylum to genus. This pattern mirrors taxonomic consistency within a hierarchically related lineage and should not be interpreted as independent replication across taxonomic levels. Of note, the signal was also repeatedly prioritized by distinct feature-selection methods, Bayesian integration, and differential-abundance analyses, providing complementary analytical support for its relevance in the analyzed cohort.

Previous studies have reported enrichment and accumulation of *Fusobacterium nucleatum* within colorectal adenomas and carcinomas relative to adjacent normal tissue [27,28]. As discussed above, experimental studies provide a plausible mechanistic framework linking *Fusobacterium nucleatum* to colorectal carcinogenesis. The FadA adhesin binds epithelial E-cadherin, activates β-catenin signaling, and induces oncogenic and inflammatory gene expression, thereby promoting colorectal cancer cell proliferation. In parallel, the Fap2 outer membrane protein binds the inhibitory receptor TIGIT on natural killer and cytotoxic T cells, attenuating antitumor immune responses and facilitating immune evasion [11,12]. More recently, structural studies demonstrated that Fap2 is a bifunctional adhesin with spatially distinct binding sites for TIGIT and the tumor-associated glycan Gal-GalNAc, providing a mechanistic basis for the simultaneous colonization of colorectal tumors and suppression of antitumor immunity [29]. Experimental studies indicate that *Fusobacterium nucleatum* reshapes the intestinal immune microenvironment through activation of inflammatory pathways and enhancement of T helper 17 cells/Interleukin-17 (Th17/IL-17) responses, whereas clinical studies associate *F. nucleatum*-positive tumors with distinct immune landscapes. Together, these observations support a role for *F. nucleatum* in establishing a tumor-permissive inflammatory microenvironment in colorectal cancer [30,31,32]. These mechanisms make the *Fusobacterium* lineage clinically interesting as a candidate biomarker and potential therapeutic target. However, our 16S rRNA data resolve the principal signal at genus and higher taxonomic levels; they do not demonstrate that *F. nucleatum* alone is responsible, nor do they establish microbial causality. Species-resolved and functional validation is therefore needed.

The tissue findings are also consistent with community-level microbial restructuring rather than enrichment of a single taxon. Alongside *Fusobacterium*, changes in butyrate-producing taxa are consistent with a broader dysbiotic ecosystem. In particular, depletion of *Faecalibacterium prausnitzii* and members of the Lachnospiraceae may reduce butyrate availability, potentially compromising epithelial homeostasis and immune regulation, processes in which butyrate plays a central role [6,13,14]. Signals involving *Ruminococcus gnavus* and *Bacteroides* point toward complementary inflammatory mechanisms. *R. gnavus* produces a glucorhamnan

polysaccharide that induces Tumor Necrosis Factor-alpha (TNF-α) secretion by dendritic cells through Toll-like Receptor 4 (TLR4), providing a potential link to mucosal inflammation. In contrast, the effects of *Bacteroides fragilis* are strain dependent: enterotoxigenic *B. fragilis* (ETBF) cleaves E-cadherin through Bacteroides fragilis toxin (BFT), activating β-catenin and Nuclear Factor kappa-light-chain-enhancer of activated B cells (NF-κB) signaling while disrupting epithelial barrier integrity, whereas non-toxigenic *B. fragilis* (NTBF) supports intestinal homeostasis by reinforcing barrier function and promoting anti-inflammatory immune responses [33,34,35]. These associations support a model in which barrier dysfunction, altered mucus utilization, decreased homeostatic SCFA production, and inflammatory microbial metabolism reinforce one another during colorectal neoplasia. Nevertheless, metabolic functions were inferred from taxonomic associations and prior literature, not measured directly in this cohort.

The superior performance of tissue biopsies is biologically plausible because adherent and biofilm-associated microorganisms are sampled in direct proximity to epithelial cells, immune infiltrates, hypoxia, altered nutrients, and extracellular-matrix remodeling. Biopsy profiles therefore provide high biological specificity for the tumor–microbiome interface, although their invasive acquisition limits screening use.

Stool integrates microbial material from the entire intestinal lumen. Its lower specificity is offset by non-invasive collection and greater clinical scalability, making fecal microbial signatures more suitable for future multi-marker approaches alongside fecal immunochemical testing, clinical variables, or host molecular markers.

Finally, saliva was the least discriminative compartment, but recurrent oral taxa remain biologically informative. Periodontal dysbiosis may increase the burden of oral pathobionts, which can reach the gastrointestinal tract through repeated swallowing and potentially contribute to intestinal microbial dysbiosis and inflammation. This proposed oral–gut axis provides a plausible mechanism linking oral microbial communities with colorectal carcinogenesis, although the underlying causal relationships remain to be fully established [18,36]. Oral microbial dysbiosis may facilitate the translocation of oral pathogens, including *Porphyromonas gingivalis*, and their microbial products to the gastrointestinal tract through repeated swallowing. This process has been proposed to contribute to gut microbial dysbiosis and intestinal barrier dysfunction, characterized by reduced expression of tight junction proteins, mucus layer disruption, and increased epithelial permeability [37]. The data therefore support an oral–gut connection while cautioning against treating salivary profiles as a direct surrogate for the tumor microbiome.

Taken together, the three compartments provide complementary views of CRC-associated dysbiosis: tissue reveals the local tumor ecosystem, stool captures luminal ecological remodeling, and saliva reflects an upstream oral reservoir or systemic correlate.

The most immediate translational implication is the potential development of compartment-aware biomarker panels. *Fusobacterium* may contribute to a tissue- or stool-based signature, whereas combinations of *Bacteroides*, *Ruminococcus*, *Faecalibacterium*, and other taxa may capture complementary inflammatory and metabolic dimensions. Such panels could eventually support discrimination between adenoma and CRC, but their performance must be tested in larger screening populations and independent cohorts.

Microbiome-targeted prevention or treatment is a longer-term possibility. Dietary modulation, prebiotics, probiotics, bacteriophages, selective antimicrobial approaches, and strategies directed against *Fusobacterium* are under investigation. The current study provides a biological rationale for these approaches but does not evaluate therapeutic efficacy. Because tissue microbiome profiling involves relatively low microbial biomass, contamination remains a critical consideration when interpreting low-abundance taxa. In detail, taxa with recognized environmental occurrence should be cautiously interpreted in the absence of a dedicated negative-control-based contaminant assessment. Accordingly, such taxa are considered exploratory signals in the present analysis, whereas greater interpretative emphasis is placed on signals supported by the broader biological context and by complementary analytical approaches. As reported in Table A3 in **Supplementary A**, the identified taxa were additionally classified according to their likely ecological origin, including "Human-related", "Possible Human", and "Possible Human/Environmental" categories, allowing potentially environmentally associated signals to be explicitly recognized and interpreted with appropriate caution. Furthermore, prior to feature selection, several abundance-, cumulative-abundance-, prevalence-, pass-through-, and variance-based filtering procedures were applied, followed by alternative log-unit and rank-unit normalization strategies. These preprocessing steps were designed to decrease the influence of sparse, low-information, or analytically unstable features; however, they should not be regarded as a substitute for dedicated negative-control-based contamination assessment.

Accordingly, the present signatures should be regarded as candidate biomarkers and mechanistic hypotheses, not as clinically validated diagnostic tests or treatment targets.

A strength of this study is the simultaneous analysis of three biologically distinct compartments across five taxonomic levels. Requiring support across complementary feature-selection methods and then integrating uncertainty through BMA reduced dependence on any single pipeline. The convergence with differential abundance analyses further prioritized signals that were both statistically stable and

biologically interpretable. The cross-sectional design of this study requires a clear distinction between observed microbial associations and potential biological mechanisms. Our data document compartment-specific differences in microbial profiles between CRC patients and those with adenomatous polyps, but they do not define the longitudinal transition from adenoma to carcinoma or establish a causal role for the identified taxa in malignant progression. Similarly, the detection of related taxa across oral and intestinal compartments does not demonstrate microbial translocation or establish the oral cavity as the source of colorectal microorganisms. In the absence of strain-resolved sequencing, ASV-level source tracking, or other direct evidence of microbial transmission between compartments, the oral–gut axis should therefore be regarded as a biologically plausible framework for interpreting the observed compartment-specific associations rather than as a mechanism documented by the present data. Proposed effects involving epithelial barrier function, host immunity, inflammation, or microbial metabolite production likewise represent mechanistic hypotheses supported by previously published studies rather than effects directly established in this cohort.

5. Limitations

Several limitations temper the conclusions. The cohort was modest, particularly the adenoma group, and no independent validation cohort was available. The cross-sectional comparison cannot demonstrate temporal progression from adenoma to CRC or establish whether dysbiosis is a cause, consequence, or amplifier of tumor development. Residual clinical confounding remains an important limitation. Although the original cohort applied exclusion criteria to minimize major treatment-related effects on the microbiome and sensitivity analyses accounted for available demographic covariates, individual-level information on potentially relevant factors, including BMI, smoking, detailed dietary variation, bowel preparation, and other medication use, was unavailable. Their contribution to the observed associations therefore cannot be excluded. The modest cohort size, especially the AP group, also limited multivariable adjustment. Contamination control represents an additional limitation for the relatively low-biomass tissue samples. Although the original workflow included DNA and sequence quality control, chimera removal, exclusion of non-prokaryotic ASVs, and sequencing-saturation assessment, dedicated extraction blanks, negative-control-based contaminant identification, and appropriate batch information were unavailable. Reagent contamination, extraction artifacts, and batch effects therefore cannot be fully excluded, warranting caution when interpreting low-abundance or potentially environmental taxa. Finally, 16S rRNA sequencing provides primarily taxonomic information and does not directly measure microbial metabolic activity or host–microbe functions. Thus, our speculative interpretations concerning butyrate production, mucus degradation, immune modulation, and related mechanisms are derived from previous literature rather than direct measurements in this study. Functional interpretations concerning butyrate production, genotoxins, or immune modulation therefore require confirmation by shotgun metagenomics, metatranscriptomics, metabolomics, spatial profiling, and experimental studies.

Future work should validate the compartment-specific signatures prospectively, assess their incremental value over established CRC screening tools, and determine how tumor site, stage, diet, medication exposure, and host immune features modify the observed associations.

6. Conclusions

In conclusion, CRC is associated with a compartment-specific restructuring of the microbiome rather than uniform dysbiosis across anatomical sites. The strongest signal occurs at the tumor surface, where *Fusobacterium* is accompanied by broader changes involving inflammatory and metabolically relevant taxa; stool captures a more diffuse loss of intestinal homeostasis, whereas saliva supports a weaker oral–gut signature. By separating biological conclusions from analytical detail, these findings position the tumor–microbiome interface as a clinically relevant feature of the differences between adenomatous polyps and CRC, while defining the validation required before microbiome biomarkers can enter practice.

Availability of Data and Materials

The 16S rRNA gene sequencing data analyzed in the present study are publicly available in the Gene Expression Omnibus (GEO) repository under accession number GSE217490. The processing and analysis scripts for the original study are available at https://github.com/LeandroD94, and the complete code and instructions for reproducing the present analyses are available at https://mohammad-fayaz.github.io/. The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

AA, ER, and MF conceived and designed the study. AA, ER, LDG, GN, AR and SB generated and curated the original data. MF and LDG performed the statistical analyses. All authors contributed to interpretation of the results and manuscript preparation. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

The authors gratefully acknowledge the patients who participated in this study and contributed biological samples for this research. We also thank the clinical and laboratory staff involved in patient recruitment, sample collection, processing, and microbiome analysis. We would like to express our gratitude to all those who helped us during the writing of this manuscript. We also thank all peer reviewers for their comments and suggestions.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest. Given his role as the Guest Editor and Editorial Board member, Amedeo Amedei had no involvement in the peer-review of this article and has no access to information regarding its peer review. Similarly, given her role as Guest Editor, Edda Russo had no involvement in the peer-review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Wei-Lin Jin.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) solely as a language-support tool to improve English fluency, grammar, clarity, and readability. The tool was not used to generate scientific data, perform statistical analyses, or formulate scientific conclusions. All AI-assisted revisions were critically reviewed and verified by the authors, who take full responsibility for the final content of the manuscript.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/FBL56530.

References

[1] Klimeck L, Heisser T, Hoffmeister M, Brenner H. Colorectal cancer: A health and economic problem. Best Practice & Research. Clinical Gastroenterology. 2023; 66: 101839. https://doi.org/10.1016/j.bpg.2023.101839

[2] Capuozzo M, Picone C, Sabbatino F, Santorsola M, Caraglia F, Iervolino D, et al. Genetic, Epidemiological, Clinical, and Therapeutic Trajectories in Colon and Rectal Cancers. Cancers. 2025; 17: 3438. https://doi.org/10.3390/cancers17213438

[3] Zhou B, Liu Q, Huang C, Chen H, Wang P, Lu Y, et al. Progressively altered genes in colorectal carcinogenesis link oncogenesis immune cycle and tumor microenvironment. Scientific Reports. 2025; 15: 37495. https://doi.org/10.1038/s41598-025-21401-y

[4] Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggiano GAD, Gasbarrini A, et al. What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. Microorganisms. 2019; 7: 14. https://doi.org/10.3390/microorganisms7010014

[5] Khosravi A, Mazmanian SK. Disruption of the gut microbiome as a risk factor for microbial infections. Current Opinion in Microbiology. 2013; 16: 221–227. https://doi.org/10.1016/j.mib.2013.03.009

[6] van der Hee B, Wells JM. Microbial Regulation of Host Physiology by Short-chain Fatty Acids. Trends in Microbiology. 2021; 29: 700–712. https://doi.org/10.1016/j.tim.2021.02.001

[7] Lange O, Proczko-Stepaniak M, Mika A. Short-Chain Fatty Acids-A Product of the Microbiome and Its Participation in Two-Way Communication on the Microbiome-Host Mammal Line. Current Obesity Reports. 2023; 12: 108–126. https://doi.org/10.1007/s13679-023-00503-6

[8] Sobhani I, Tap J, Roudot-Thoraval F, Roperch JP, Letulle S, Langella P, et al. Microbial dysbiosis in colorectal cancer (CRC) patients. PloS One. 2011; 6: e16393. https://doi.org/10.1371/journal.pone.0016393

[9] Li X, Wang Q, Yuan Q, Wang L. The intratumoral microbiome in colorectal cancer: origins, microenvironmental interactions, and new horizons in precision medicine. Frontiers in Immunology. 2026; 17: 1795736. https://doi.org/10.3389/fimmu.2026.1795736

[10] Bachir A, Altaie AM, Bendardaf R, Talaat IM, Hamoudi R. Microbial dysbiosis drives colorectal carcinogenesis via integrated inflammatory, metabolic, and biofilm pathways. Frontiers in Microbiology. 2026; 17: 1795882. https://doi.org/10.3389/fmicb.2026.1795882

[11] Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. Cell Host & Microbe. 2013; 14: 195–206. https://doi.org/10.1016/j.chom.2013.07.012

[12] Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, et al. Binding of the Fap2 Protein of Fusobacterium nucleatum to Human Inhibitory Receptor TIGIT Protects Tumors from Immune Cell Attack. Immunity. 2015; 42: 344–355. https://doi.org/10.1016/j.immuni.2015.01.010

[13] Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux JJ, et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105: 16731–16736. https://doi.org/10.1073/pnas.0804812105

[14] Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. Environmental Microbiology. 2017; 19: 29–41. https://doi.org/10.1111/1462-2920.13589

[15] Singh G, Chaudhry Z, Boyadzhyan A, Sasaninia K, Rai V. Dysbiosis and colorectal cancer: conducive factors, biological and molecular role, and therapeutic prospectives. Exploration of Targeted Anti-tumor Therapy. 2025; 6: 1002329. https://doi.org/10.37349/etat.2025.1002329

[16] Russo E, Gloria LD, Nannini G, Meoni G, Niccolai E, Ringressi MN, et al. From adenoma to CRC stages: the oral-gut microbiome axis as a source of potential microbial and metabolic biomarkers of malignancy. Neoplasia (New York, N.Y.). 2023; 40: 100901. https://doi.org/10.1016/j.neo.2023.100901

[17] Flemer B, Warren RD, Barrett MP, Cisek K, Das A, Jeffery IB, et al. The oral microbiota in colorectal cancer is distinctive and predictive. Gut. 2018; 67: 1454–1463. https://doi.org/10.1136/gutjnl-2017-314814

[18] Tortora SC, Agurto MG, Martello LA. The oral-gut-circulatory axis: from homeostasis to colon cancer. Frontiers in Cellular and Infection Microbiology. 2023; 13: 1289452. https://doi.org/10.3389/fcimb.2023.1289452

[19] Conde-Pérez K, Aja-Macaya P, Buetas E, Trigo-Tasende N, Nasser-Ali M, Rumbo-Feal S, et al. The multispecies microbial cluster of Fusobacterium, Parvimonas, Bacteroides and Faecalibacterium as a precision biomarker for colorectal cancer diagnosis. Molecular Oncology. 2024; 18: 1093–1122. https://doi.org/10.1002/1878-0261.13604

[20] Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. Frontiers in Microbiology. 2017; 8: 2224. https://doi.org/10.3389/fmicb.2017.02224

[21] Aldirawi H, Morales FG. Univariate and Multivariate Statistical Analysis of Microbiome Data: An Overview. Applied Microbiology. 2023; 3: 322–338. https://doi.org/10.3390/applmicrobiol3020023

[22] Rodriguez J, Hassani Z, Alves Costa Silva C, Betsou F, Carraturo F, Fasano A, et al. State of the art and the future of microbiome-based biomarkers: a multidisciplinary Delphi consensus. The Lancet. Microbe. 2025; 6: 100948. https://doi.org/10.1016/j.lanmic.2024.07.011

[23] Jiang L, Haiminen N, Carrieri AP, Huang S, Vázquez-Baeza Y, Parida L, et al. Utilizing stability criteria in choosing feature selection methods yields reproducible results in microbiome data. Biometrics. 2022; 78: 1155–1167. https://doi.org/10.1111/biom.13481

[24] Wirbel J, Zych K, Essex M, Karcher N, Kartal E, Salazar G, et al. Microbiome meta-analysis and cross-disease comparison enabled by the SIAMCAT machine learning toolbox. Genome Biology. 2021; 22: 93. https://doi.org/10.1186/s13059-021-02306-1

[25] Ho NT, Li F, Wang S, Kuhn L. metamicrobiomeR: an R package for analysis of microbiome relative abundance data using zero-inflated beta GAMLSS and meta-analysis across studies using random effects models. BMC Bioinformatics. 2019; 20: 188. https://doi.org/10.1186/s12859-019-2744-2

[26] Tang ZZ, Chen G. Zero-inflated generalized Dirichlet multinomial regression model for microbiome compositional data analysis. Biostatistics (Oxford, England). 2019; 20: 698–713. https://doi.org/10.1093/biostatistics/kxy025

[27] Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. Genome Research. 2012; 22: 299–306. https://doi.org/10.1101/gr.126516.111

[28] Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. Cell Host & Microbe. 2013; 14: 207–215. https://doi.org/10.1016/j.chom.2013.07.007

[29] Schöpf F, Marongiu GL, Milaj K, Sprink T, Kikhney J, Moter A, et al. Structural basis of Fusobacterium nucleatum adhesin Fap2 interaction with receptors on cancer and immune cells. Nature Communications. 2025; 16: 8104. https://doi.org/10.1038/s41467-025-63451-w

[30] Neuzillet C, Marchais M, Vacher S, Hilmi M, Schnitzler A, Meseure D, et al. Prognostic value of intratumoral Fusobacterium nucleatum and association with immune-related gene expression in oral squamous cell carcinoma patients. Scientific Reports. 2021; 11: 7870. https://doi.org/10.1038/s41598-021-86816-9

[31] Brennan CA, Clay SL, Lavoie SL, Bae S, Lang JK, Fonseca-Pereira D, et al. Fusobacterium nucleatum drives a pro-inflammatory intestinal microenvironment through metabolite receptor-dependent modulation of IL-17 expression. Gut Microbes. 2021; 13: 1987780. https://doi.org/10.1080/19490976.2021.1987780

[32] Borowsky J, Haruki K, Lau MC, Dias Costa A, Väyrynen JP, Ugai T, et al. Association of Fusobacterium nucleatum with specific T-cell subsets in the colorectal carcinoma microenvironment. Clinical Cancer Research. 2021; 27: 2816–2826. https://doi.org/10.1158/1078-0432.CCR-20-4009

[33] Crost EH, Tailford LE, Monestier M, Swarbreck D, Henrissat B, Crossman LC, et al. The mucin-degradation strategy of Ruminococcus gnavus: The importance of intramolecular trans-sialidases. Gut Microbes. 2016; 7: 302–312. https://doi.org/10.1080/19490976.2016.1186334

[34] Henke MT, Kenny DJ, Cassilly CD, Vlamakis H, Xavier RJ, Clardy J. Ruminococcus gnavus, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. Proceedings of the National Academy of Sciences of the United States of America. 2019; 116: 12672–12677. https://doi.org/10.1073/pnas.1904099116

[35] Jasemi S, Molicotti P, Fais M, Cossu I, Simula ER, Sechi LA. Biological Mechanisms of Enterotoxigenic Bacteroides fragilis Toxin: Linking Inflammation, Colorectal Cancer, and Clinical Implications. Toxins. 2025; 17: 305. https://doi.org/10.3390/toxins17060305

[36] Lam GA, Albarrak H, McColl CJ, Pizarro A, Sanaka H, Gomez-Nguyen A, et al. The Oral-Gut Axis: Periodontal Diseases and Gastrointestinal Disorders. Inflammatory Bowel Diseases. 2023; 29: 1153–1164. https://doi.org/10.1093/ibd/izac241

[37] Fu W, Yang N, Yan J, Han B, Niu Q, Li Z, et al. Oral-Gut Microbial Crosstalk and Therapeutic Applications of Bacterial Extracellular Vesicles. Biomolecules. 2026; 16: 26. https://doi.org/10.3390/biom16010026