










Article

Blood Bacterial DNA Load and Profiling Differ in Colorectal Cancer Patients Compared to Tumor-Free Controls

Massimiliano Mutignani ^{1,†}, Roberto Penagini ^{2,3,†} , Giorgio Gargari ⁴ , Simone Guglielmetti ⁴ , Marcello Cintolo ¹, Aldo Airoidi ⁵, Pierfrancesco Leone ⁶, Pietro Carnevali ⁷, Clorinda Ciafardini ², Giulio Petrocelli ^{1,8}, Federica Mascaretti ^{2,9}, Barbara Oreggia ⁶, Lorenzo Dioscoridi ¹, Federica Cavalcoli ¹⁰, Massimo Primignani ¹¹ , Francesco Pugliese ¹, Paola Bertuccio ^{9,12}, Pietro Soru ¹³, Carmelo Magistro ⁷, Giovanni Ferrari ⁷, Michela C. Speciani ⁹ , Giulia Bonato ¹, Marta Bini ¹, Paolo Cantù ², Flavio Caprioli ^{2,3}, Marcello Vangeli ⁵, Edoardo Forti ¹ , Stefano Mazza ^{2,14} , Giulia Tosetti ¹¹, Rossella Bonzi ⁹, Maurizio Vecchi ^{2,3}, Carlo La Vecchia ⁹  and Marta Rossi ^{9,*} 

- ¹ Digestive and Interventional Endoscopy Unit, ASST Grande Ospedale Metropolitano Niguarda, 20162 Milan, Italy; massimiliano.mutignani@ospedaleniguarda.it (M.M.); marcello.cintolo@ospedaleniguarda.it (M.C.); giulio.petrocelli@ospedaleniguarda.it (G.P.); lorenzo.dioscoridi@ospedaleniguarda.it (L.D.); francesco.pugliese@ospedaleniguarda.it (F.P.); giulia.bonato@ospedaleniguarda.it (G.B.); marta.bini@ospedaleniguarda.it (M.B.); edoardo.forti@ospedaleniguarda.it (E.F)
- ² Gastroenterology and Endoscopy Unit, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, 20122 Milan, Italy; roberto.penagini@unimi.it (R.P.); clorinda.ciafardini@policlinico.mi.it (C.C.); federica.mascaretti@unimi.it (F.M.); paolo.cantu@policlinico.mi.it (P.C.); flavio.caprioli@unimi.it (F.C.); stefano.mazza@asst-cremona.it (S.M.); maurizio.vecchi@unimi.it (M.V.)
- ³ Department of Pathophysiology and Transplantation, University of Milan, 20122 Milan, Italy
- ⁴ Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20133 Milan, Italy; giorgio.gargari@unimi.it (G.G.); simone.guglielmetti@unimi.it (S.G.)
- ⁵ Hepatology and Gastroenterology Unit, ASST Grande Ospedale Metropolitano Niguarda, 20162 Milan, Italy; aldo.airoidi@ospedaleniguarda.it (A.A.); marcello.vangeli@ospedaleniguarda.it (M.V.)
- ⁶ General Surgery Unit, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, 20122 Milan, Italy; pierfrancesco.leone@policlinico.mi.it (P.L.); barbara.oreggia@policlinico.mi.it (B.O.)
- ⁷ Division of Minimally-Invasive Surgical Oncology, Niguarda Cancer Center, ASST Grande Ospedale Metropolitano Niguarda, 20162 Milan, Italy; pietro.carnevali@ospedaleniguarda.it (P.C.); carmelomagistro@ospedaleniguarda.it (C.M.); giovanni.ferrari@ospedaleniguarda.it (G.F.)
- ⁸ Associazione Nazionale Operatori Tecniche Endoscopiche (ANOTE), 80061 Massa Lubrense, Italy
- ⁹ Department of Clinical Science and Community Health, University of Milan, 20133 Milan, Italy; paola.bertuccio@unipv.it (P.B.); michela.speciani@unimi.it (M.C.S.); rossella.bonzi@unimi.it (R.B.); carlo.lavecchia@unimi.it (C.L.V.)
- ¹⁰ Diagnostic and Therapeutic Endoscopy Unit, Fondazione IRCCS Istituto Nazionale Tumori, 20133 Milan, Italy; federica.cavalcoli@istitutotumori.mi.it
- ¹¹ Division of Gastroenterology and Hepatology, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, 20122 Milan, Italy; massimo.primignani@policlinico.mi.it (M.P.); giulia.tosetti@policlinico.mi.it (G.T.)
- ¹² Department of Public Health, Experimental and Forensic Medicine, University of Pavia, 27100 Pavia, Italy
- ¹³ Division of Endoscopy, IRCCS Istituto Europeo di Oncologia, 20141 Milan, Italy; pietro.soru@ieo.it
- ¹⁴ Gastroenterology and Digestive Endoscopy Unit, ASST Cremona, Cremona 26100, Italy
- * Correspondence: marta.rossi@unimi.it
- † These authors contributed equally.



Citation: Mutignani, M.; Penagini, R.; Gargari, G.; Guglielmetti, S.; Cintolo, M.; Airoidi, A.; Leone, P.; Carnevali, P.; Ciafardini, C.; Petrocelli, G.; et al. Blood Bacterial DNA Load and Profiling Differ in Colorectal Cancer Patients Compared to Tumor-Free Controls. *Cancers* **2021**, *13*, 6363. <https://doi.org/10.3390/cancers13246363>

Academic Editor: Athanasios G. Papavassiliou

Received: 7 November 2021

Accepted: 15 December 2021

Published: 18 December 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: In colorectal cancer patients, epithelial barrier dysfunction can lead to increased intestinal permeability, and gut microbiome was found to vary compared to healthy subjects. We conducted a study to investigate whether bacterial translocation from gastrointestinal tract to bloodstream is associated to intestinal adenoma and/or colorectal cancer. In particular, an epidemiological and metagenomic approach was used to evaluate the relation of the bacterial DNA load and the bacterial taxonomic groups—assessed by 16S rRNA profiling—in blood with the risks of intestinal adenoma and colorectal cancer. These findings can confirm the presence of bacterial DNA in blood in healthy adults and serve as a basis to evaluate new non-invasive techniques for an early CRC diagnosis through the analyses of bacterial DNA circulating in peripheral blood.

Abstract: Inflammation and immunity are linked to intestinal adenoma (IA) and colorectal cancer (CRC) development. The gut microbiota is associated with CRC risk. Epithelial barrier dysfunction can occur, possibly leading to increased intestinal permeability in CRC patients. We conducted a case-control study including 100 incident histologically confirmed CRC cases, and 100 IA and 100 healthy subjects, matched to cases by center, sex and age. We performed 16S rRNA gene analysis of blood and applied conditional logistic regression. Further analyses were based on negative binomial distribution normalization and Random Forest algorithm. We found an overrepresentation of blood 16S rRNA gene copies in colon cancer as compared to tumor-free controls. For high levels of gene copies, community diversity was higher in colon cancer cases than controls. Bacterial taxa and operational taxonomic unit abundances were different between groups and were able to predict CRC with an accuracy of 0.70. Our data support the hypothesis of a higher passage of bacteria from gastrointestinal tract to bloodstream in colon cancer. This result can be applied on non-invasive diagnostic tests for colon cancer control.

Keywords: bacterial translocation; bacterial 16S rRNA gene; case-control study; colon cancer diagnoses; microbiome

1. Introduction

Colorectal cancer (CRC) is the 3rd most common cancer, and ranks second in terms of mortality, worldwide [1]. Although mortality trends have been favorable in Europe during the last decades, with a predicted rate of 15.4/100,000 in men and 8.6/100,000 in women in 2020, in most eastern European countries CRC mortality is still increasing [2,3].

CRC derives from a sequential accumulation of genetic alterations that involves the transition from normal mucosa to pre-malignant lesions with progression to intestinal adenoma (IA) and invasive CRC [4]. Inflammation and immunity are inextricably linked to all phases of CRC development [5]. Numerous studies identified chronic intestinal inflammation as a risk factor for CRC, as also confirmed by an increased incidence of this tumor in inflammatory bowel disease patients [6]. IA and CRC have also been associated with increased circulating inflammation [7,8], and more recently with dysfunction of the gut mucosal barrier [9].

The community structure of the intestinal microbial ecosystem influences the risk of IA and CRC [9–11]. Gut mucosal microbiota *Fusobacterium* spp. were increased in CRC patients [10,11]; *Bacteroides fragilis* and the genus *Porphyromonas* have been also associated with an increased risk of CRC [10]. Numerous studies analyzed fecal microbiome, not always reporting consistent results [11–13], but recent meta-analyses including geographically and technically different shotgun metagenomic studies showed a higher fecal microbiota richness in CRC cases as compared to controls [14], and identified a set of bacterial taxa significantly enriched in CRC cases [14,15].

Increasing evidence has pointed out the presence of bacterial DNA in blood [16,17]. Microbial signatures have been reported for gut dysbiosis among diabetic subjects and for liver fibrosis in obese patients [18,19]. Microbiome analysis of blood has also been proposed as a tool to discriminate between cancer patients and healthy subjects [20]. In particular, differences in circulating bacterial factors can occur in CRC patients, in whom epithelial barrier dysfunction can lead to increased intestinal permeability, plausibly resulting into a greater bacterial translocation from the gastrointestinal tract to bloodstream in IA and/or CRC. Some differences in relative abundance of the bacterial DNA in plasma between healthy, IA and CRC subjects have been reported in a case-control study from China, including 57 participants, but differences in terms of total bacterial load have not been analyzed yet [21].

In this study, we aimed to compare the load of bacterial DNA in blood and taxonomic profile between CRC, IA and healthy controls, using data from a case-control study.

2. Results

Table 1 gives the distribution of 100 healthy controls, 100 IA subjects, and 100 cases of CRC according to sex, age, study center and education. By design, the three groups had the same sex and center distributions and were similar in terms of age. Cases tended to be less educated than IA subjects and controls in absence, however, of a significant difference (χ^2 test $p = 0.155$).

Table 1. Distribution of 100 healthy controls, 100 intestinal adenoma (IA) patients and 100 colorectal cancer (CRC) cases by sex, age, study center and years of education. Italy 2017–2019.

Characteristic	Controls	IA	CRC
Sex			
Male	62 (62%)	62 (62%)	62 (62%)
Female	38 (38%)	38 (38%)	38 (38%)
Age group (years)			
<50	7 (7%)	4 (4%)	10 (10%)
50–59	23 (23%)	20 (20%)	19 (19%)
60–69	26 (26%)	36 (36%)	29 (29%)
70–79	33 (33%)	29 (29%)	31 (42%)
≥80	11 (11%)	11 (11%)	11 (11%)
χ^2 test, $p = 0.76$			
Mean (SD) age (years) *	66.0 (11.8)	65.9 (10.9)	66.1 (11.6)
Center			
Niguarda	65 (65%)	65 (65%)	65 (65%)
Policlinico	35 (35%)	35 (35%)	35 (35%)
Education (years) †			
<7	12 (12%)	19 (19%)	25 (25%)
7–11	24 (24%)	26 (26%)	25 (25%)
≥12	64 (64%)	55 (55%)	49 (50%)
χ^2 test, $p = 0.155$			

SD: standard deviation (SD). * Anova test for heterogeneity $p = 1.00$. † The sum does not add up to the total because of one missing value.

2.1. 16S rRNA Gene Copies

We found an overall mean of 7687 16S rRNA gene copies per μL of blood, with a mean of 7628 among controls, 7586 among IA and 8387 among CRC subjects (9145 in colon and 7629 in rectal cancers), with no significant differences between the three groups (p for heterogeneity = 0.482) (Supplementary Figure S1). Since 16S rRNA gene copy distribution was very similar in control and IA subjects (p for heterogeneity = 0.95), we grouped them as reference group and compared their 16S rRNA gene copy distribution with that of CRC, colon and rectal cancer cases. We found significant differences between colon cancer and control/IA ($p = 0.025$; Figure 1).

Table 2 shows the distribution of control/IA subjects, CRC, colon and rectal cancer cases, the odds ratios (ORs) and the corresponding 95% confidence intervals (CIs) according to quintiles of 16S rRNA gene copies, as well as the continuous ORs. We found a direct association of 16S rRNA gene copies with the risk of colon cancer. The OR of colon cancer for the highest (≥ 9707) versus the lowest first three (< 7618) quintiles of gene copies was 2.62 (95%CI = 1.22–5.65). The association significantly increased after the fourth quintile (p for trend = 0.013) and became stronger for levels higher than the fifth quintile cut-off: the OR was 7.22 (95%CI = 2.18–23.9) after the 90th centile ($> 11,265$ copies) and 17.08 (95%CI = 3.36–86.87) after the 95th centile ($> 13,000$ copies). The continuous OR indicated a two-fold increased risk for an increment equal to 4328 copies (OR = 2.02; 95%CI = 1.26–3.25). No association was found between 16S rRNA gene copies and rectal cancer risk. The OR for the highest versus the lowest three quintiles was 0.81 (95%CI = 0.32–2.03). The test for heterogeneity between colon and rectal results was significant ($p = 0.021$).

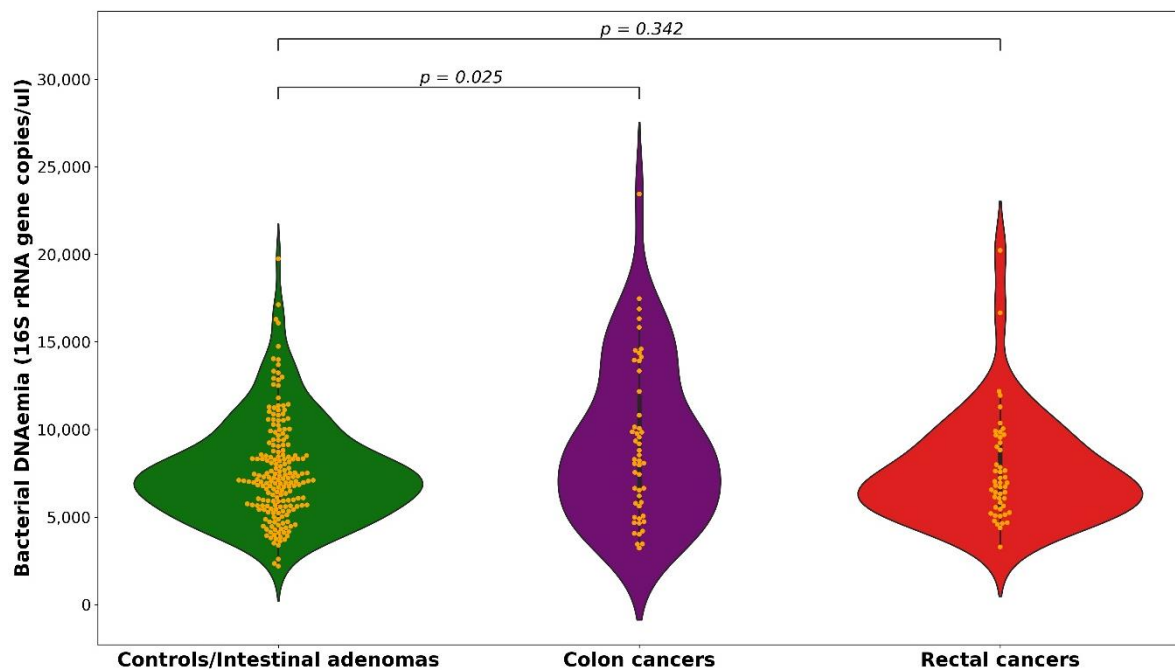


Figure 1. Distribution of 16S rRNA gene copies per μL of whole blood among controls/intestinal adenomas, colon and rectal cancers.

Table 2. Odds ratios (OR) * and corresponding 95% confidence intervals (CI) according to quintiles of 16S rRNA gene copies in whole blood/ μL of 100 control and 100 intestinal adenoma (IA) patients and of 100 colorectal cancer (CRC) cases (50 colon and 50 rectal cancers). Italy 2017–2019.

	Mean (SD)	Quintile of Number of Gene Copies †, OR (95% CI)			χ^2 Trend (p Value) across the 3 Categories	Continuous OR §
		1–3 ‡	4	5		
Upper cutpoints (n copies/ μL)		7617.5	9707.4	-		
Control/IA, n (%)	7606.6 (3895.8)	120 (60%)	40 (20%)	40 (20%)		
Total CRC, n (%)	8387.1 (2865.4)	52 (52%)	20 (20%)	28 (28%)		
Colon cancer, n (%)	9145.4 (4476.2)	1 ‡	1.16 (0.60–2.22)	1.59 (0.89–2.82)	2.40 (0.121)	1.39 (1.00–1.92)
Rectal cancer, n (%)	7628.8 (3075.0)	1 ‡	1.96 (0.75–5.08)	2.62 (1.22–5.65)	6.21 (0.013)	2.02 (1.26–3.25)
χ^2 interaction (p value) between colon and rectum		1 ‡	0.73 (0.29–1.84)	0.81 (0.32–2.03)	0.358 (0.549)	0.86 (0.51–1.42)
				5.30 (0.021)		4.34 (0.037)

SD: standard deviation (SD); * Estimated from logistic regression models, conditioned on age, sex, and study center; † Computed among control and IA distribution; ‡ Reference category; § Estimated for an increment equal to the difference between the upper cut-points of 4th and the 1st quintiles (=4328 copies).

When the association with 16S rRNA gene copies was further examined according to colon subsites, a positive association appeared to be more pronounced for right colon, with an OR for the highest versus the lowest three quintiles of 10.75 (95%CI = 2.16–53.42) as compared to 1.25 (95%CI = 0.46–3.38) for other colon sites, in absence, however, of het-

erogeneity ($p = 0.123$). Among 21 right colon cancers, 11 (53%) were in the highest quintile of gene copies, whereas among 29 other colon cancer subsites, 8 (26%) were in the highest quintile, in comparison to 40 out of 200 (20%) among control/IA subjects. In particular, among 11 cancers of the ascending colon, 7 (64%) were in the highest quintile (Table 3).

Table 3. Distribution of control and intestinal adenoma (IA) subjects, and colon and rectal cancer cases by cancer subsites and quintiles of 16S rRNA gene copies. Italy 2017–2019.

	Total	Quintile of Number of Gene Copies *, <i>n</i> (%)		
		1–3	4	5
Control/IA	200	120 (60%)	40 (20%)	40 (20%)
Tumor subsite				
Right colon	21	7 (33%)	3 (14%)	11 (53%)
Cecum	4	2 (50%)	0 (0%)	2 (50%)
Ascending	11	2 (18%)	2 (18%)	7 (64%)
Hepatic flexure	6	3 (50%)	1 (17%)	2 (33%)
Other than right colon	29	14 (48%)	7 (24%)	8 (38%)
Transverse colon	2	1 (50%)	1 (50%)	0 (0%)
Splenic flexure	3	1 (33%)	1 (33%)	1 (33%)
Descending colon	7	4 (57%)	2 (29%)	1 (14%)
Sigmoid colon	17	8 (47%)	3 (18%)	6 (35%)
Rectum	50	31 (62%)	10 (20%)	9 (18%)
Rectosigmoid junction	3	3 (100%)	0 (0%)	0 (0%)
Rectum	47	28 (60%)	10 (21%)	9 (19%)

* Computed among control and IA distribution.

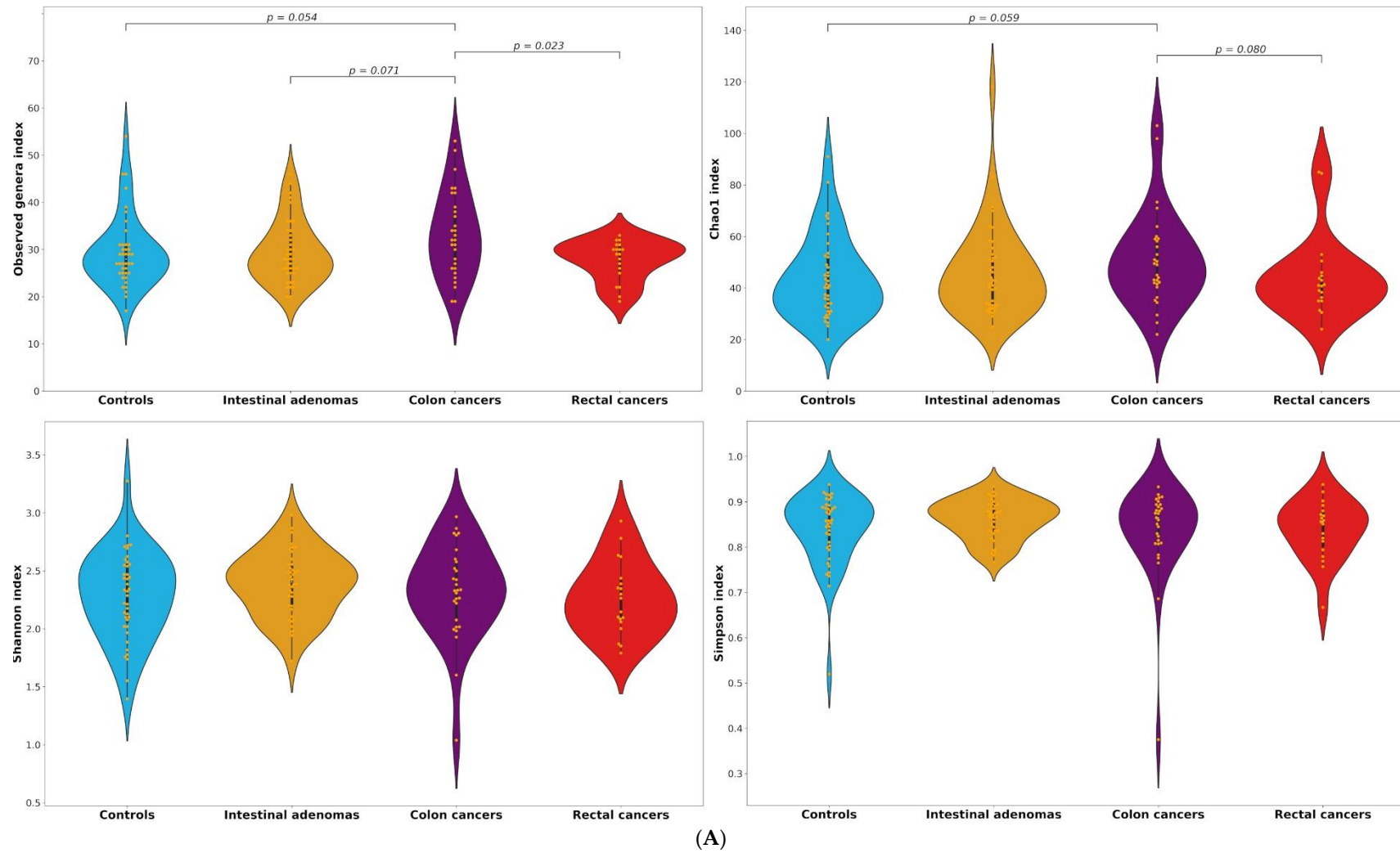
2.2. Alpha and Beta Diversity

No differences between groups were found in terms of α -diversity indices (Supplementary Table S1). When we restricted the analyses to subjects in the two highest quintiles of 16S rRNA gene copies, we found a higher diversity in colon cancer cases as compared to controls in terms of observed taxa and Chao1 indices for both genera (median of 32 vs. 28, $p = 0.054$ and median of 49 vs. 40.6, $p = 0.059$, respectively) and operational taxonomic units (OTUs) (median of 40 vs. 34, $p = 0.039$ and median of 71.1 vs. 53.4, $p = 0.067$), and as compared to rectal cancer cases in terms of both observed genera and OTUs (median of 32 vs. 29, $p = 0.023$ and median of 40 vs. 37, $p = 0.029$) (Table 4, Figure 2). Colon cancers also appeared to be higher than IA in terms of observed genera (median of 32 vs. 28, $p = 0.071$), and when we compared the observed genera index in colon cancer cases versus control/IA subjects together, the p value for heterogeneity decreased to 0.035.

Table 4. Distributions of observed, Chao1, Shannon and Simpson alpha-diversity indices of control, intestinal adenoma (IA), colon and rectal cancer subjects for bacterial genera and OTUs among the highest two quintiles of 16S rRNA gene copies. Italy 2017–2019.

	<i>Alpha-Diversity</i>	Median (I–III Quartiles)				<i>p</i> *	<i>p</i> *	<i>p</i> *	<i>p</i> *	<i>p</i> *	<i>p</i> *
		Controls	IA	Colon Cancer	Rectal Cancer	Controls vs. IA	Colon Cancers vs. IA	Colon Cancer vs. Controls	Rectal Cancer vs. IA	Rectal Cancer vs. controls	Colon Cancer vs. Rectal Cancer
Genera	<i>Observed</i>	28 (25–31)	28 (25–33.5)	32 (26–39)	29 (25–31)	0.942	0.071	0.054	0.714	0.968	0.023
	<i>Chao1</i>	40.6 (31.7–52.3)	42.5 (33.4–52.5)	49 (41.5–59)	41 (35–46)	0.476	0.148	0.059	0.703	0.789	0.080
	<i>Shannon</i>	2.33 (2.05–2.55)	2.41 (2.18–2.56)	2.33 (2.08–2.58)	2.26 (2.06–2.44)	0.254	0.535	0.614	0.119	0.715	0.442
	<i>Simpson</i>	0.86 (0.80–0.89)	0.87 (0.84–0.89)	0.84 (0.81–0.89)	0.86 (0.79–0.88)	0.394	0.496	0.846	0.186	0.697	0.513
OTUs	<i>Observed</i>	34 (30–39)	35 (32.4–43.5)	40 (33–51)	37 (30–38)	0.413	0.154	0.039	0.570	0.820	0.029
	<i>Chao1</i>	53.4 (47.8–70.5)	66 (51.9–92.8)	71.1 (52–87)	56 (51–73)	0.070	0.981	0.067	0.233	0.565	0.278
	<i>Shannon</i>	2.52 (2.27–2.74)	2.63 (2.46–2.84)	2.60 (2.46–2.73)	2.50 (2.27–2.65)	0.154	0.662	0.473	0.089	0.727	0.149
	<i>Simpson</i>	0.90 (0.86–0.93)	0.91 (0.88–0.93)	0.90 (0.87–0.92)	0.87 (0.86–0.91)	0.233	0.473	0.653	0.062	0.403	0.229

* *p* for heterogeneity estimated from the Wilcoxon rank-sum.



(A)
Figure 2. Cont.

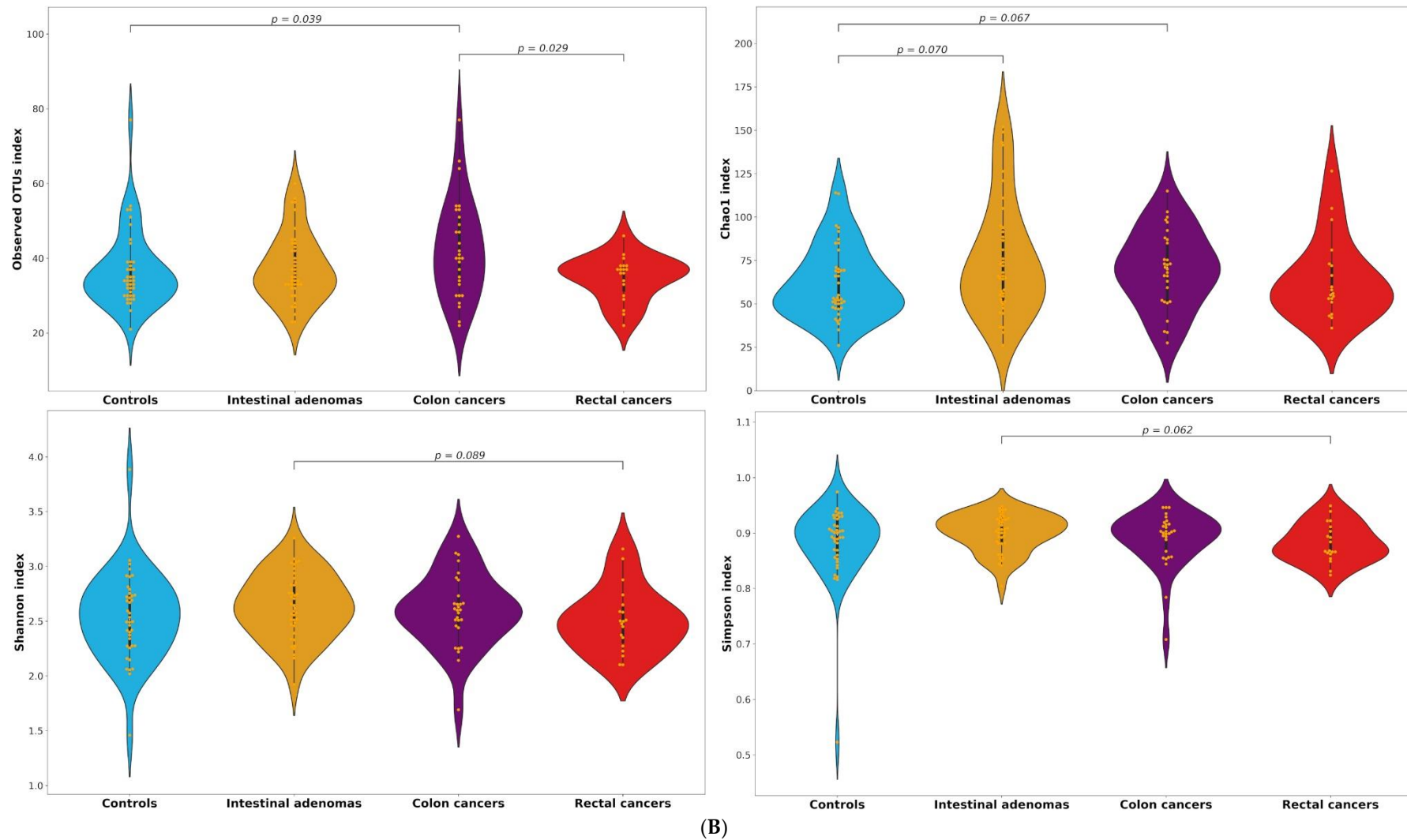


Figure 2. (A) Distributions of observed, Chao1, Shannon and Simpson alpha-diversity indices among controls, intestinal adenomas, colon and rectal cancers for genera among the highest two quintiles of 16S rRNA gene copies. (B) Distributions of observed, Chao1, Shannon and Simpson alpha-diversity indices among controls, intestinal adenomas, colon and rectal cancers for the operational taxonomic units (OTUs) among the highest two quintiles of 16S rRNA gene copies.

Concerning the beta diversity, no differences between groups were found overall and among subjects into the two highest quintiles of 16S rRNA gene copies. When we restricted the analyses to the highest quintile of 16S rRNA, we observed significant differences between control/IA, colon and rectal cancer patients (Weighted UniFrac, $p = 0.026$; Unweighted UniFrac, $p = 0.051$; Generalized UniFrac, $p = 0.031$) (Figure 3A–C). Post-hoc analyses splitting the groups 2 by 2 showed a trend between controls/IA and colon cancers in Weighted UniFrac ($p = 0.073$) (Figure 3D).

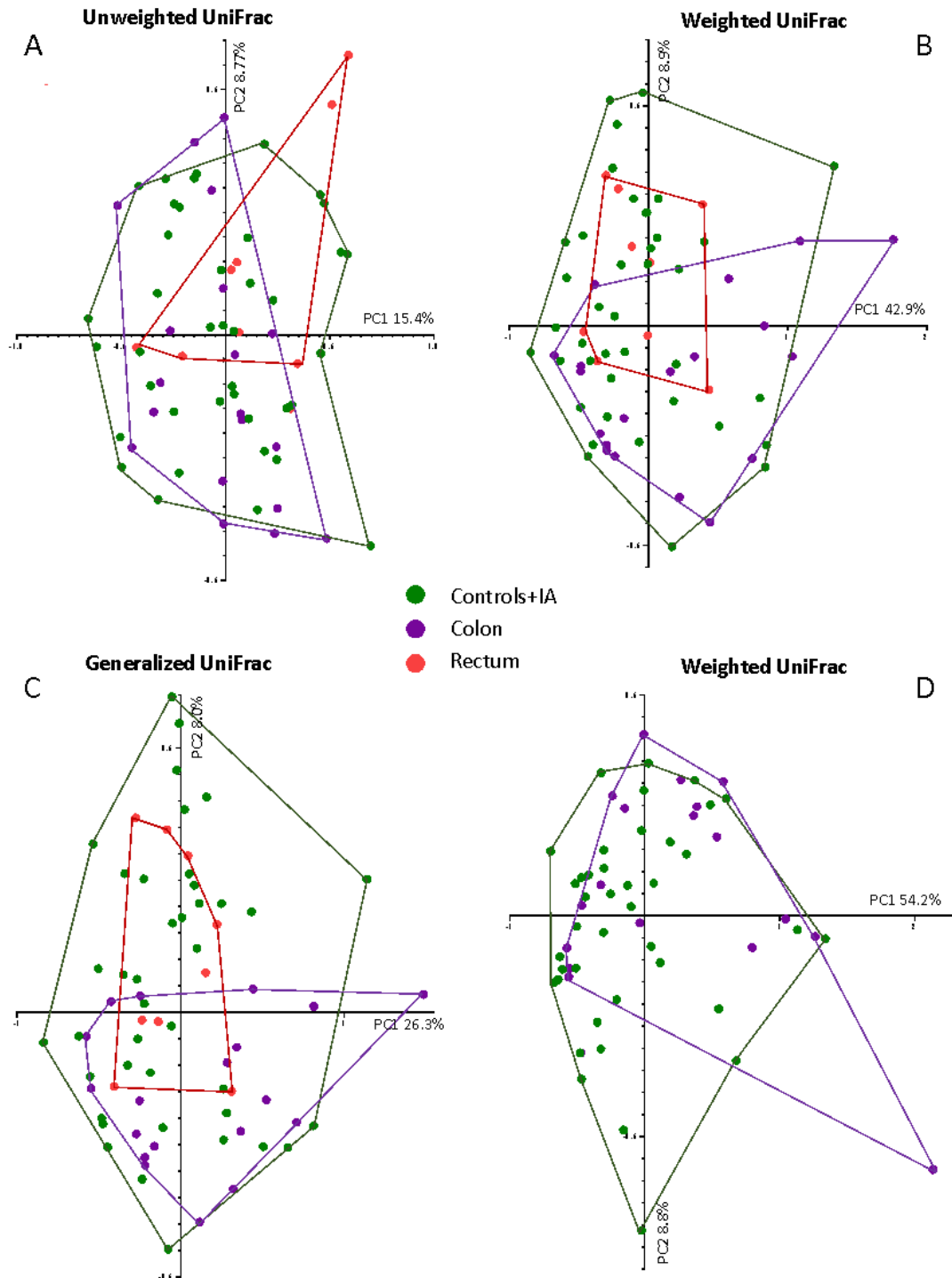


Figure 3. β -diversity of controls and intestinal adenoma (Controls + IA) group, colon cancer (Colon) and rectal cancer (Rectum) among subjects into the fifth quintile of 16S rRNA gene copies. The figure shows the UniFrac[®]-diversity in all the three variants: (A) Unweighted UniFrac; (B) Weighted UniFrac; (C) generalized UniFrac for all the three groups: Controls + IA, Colon and Rectum. Panel (D) shows the Weighted UniFrac for Controls + IA and colon cancers after post-hoc analyses.

2.3. Taxonomic Profiling of Blood Bacterial DNA between Groups

We detected a total of 1081 OTUs that were taxonomically classified into 15 phyla, 34 classes, 87 orders, 164 families and 325 genera.

Pseudomonadaceae, Micrococcaceae, Burkholderiaceae, Caulobacteraceae, Moraxellaceae and Flavobacteriaceae were the six most represented families, which together accounted for more than 50% of all reads assigned to bacterial taxa (Supplementary Figure S2).

The mean of DESeq2 normalized data and the adjusted p values comparing CRC versus control/IA on all the taxonomy levels and OTUs are shown in Figure 4. CRC samples were characterized by the increase of sequencing reads assigned to the bacterial families Peptostreptococcaceae and Acetobacteriaceae, together with a lower representation of the bacterial families Bacteroidaceae, Lachnospiraceae, and Ruminococcaceae. Comparisons between every two groups (CRC versus control, CRC versus IA and IA versus control) are also shown in Supplementary Figure S3.

Through the Random Forest supervised method, we found a set of variables that predict a CRC case from control/IA subjects with an accuracy of 0.70 (Sensitivity = 0.45; Specificity = 0.87) and another model that inferred the location of CRC discriminating colon from rectal cancer with an accuracy of 0.77 (Sensitivity = 0.71; Specificity = 0.82) (Figure 5). The first model inferred that the families Acetobacteraceae, Peptostreptococcaceae and Oligoflexaceae, the genus *Melittangium*, and the OTUs belonging to the genera *Acinetobacter*, *Pelomonas*, *Novosphingobium* and *Pajaroellobacter* were the most important variables to predict between CRC or control/IA group (Figure 5A). The biplot in Figure 5A shows a separation between CRC and control/IA subjects due to a higher dispersion of CRC cases. The second model inferred that the families *Peptostreptococcaceae*, *Streptococcaceae* and *Ruminococcaceae*, the genera *Arthrobacter*, *Clostridium* sensu stricto and *Kocuria*, and the OTUs belonging to the genera *Legionella*, *Kocuria* and *Lepisosteus oculatus* were the most important variables to predict the group between colon and rectal cancer, together with other variables that contributed to increase the accuracy of the model, including 16S rRNA gene copies, the phylum *Proteobacteria*, the order *Rhizobiales* and the genus *Bacteroides* (Figure 5B).

When the Random Forest algorithm was applied to the sample belonging to the highest quintile of 16S rRNA gene copies, the accuracy to predict between CRC and control/IA group increased to 0.79.

CRC	Controls+IA	OTU#	Taxonomy; CRC versus Controls+IA	padj	log2FoldChange
			p_Actinobacteria;c_Actinobacteria;o_Propionibacteriales;f_Nocardioidaceae	3E-05	-5.93
			p_Actinobacteria;c_Actinobacteria;o_Propionibacteriales;f_Nocardioidaceae;g_Nocardioides	3E-03	-4.97
		OTU_6	p_Actinobacteria;c_Actinobacteria;o_Propionibacteriales;f_Propionibacteriaceae;g_Cutibacterium;s_Multi-affiliation	4E-02	0.87
			p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae	2E-14	-8.13
			p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides	2E-14	-8.13
		OTU_52	p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;s_Multi-affiliation	1E-03	4.37
			p_Bacteroidetes;c_Bacteroidia;o_Chitinophagales;f_Chitinophagaceae	3E-08	-7.72
			p_Bacteroidetes;c_Bacteroidia;o_Cytophagales;f_Spirosomaceae	4E-07	-8.58
		OTU_28	p_Bacteroidetes;c_Bacteroidia;o_Flavobacteriales;f_Crocinitomicaceae;g_Fluviicola;s_Unknown	1E-02	-4.09
			p_Bacteroidetes;c_Bacteroidia;o_Sphingobacteriales	1E-06	-7.14
			p_Bacteroidetes;c_Bacteroidia;o_Sphingobacteriales;f_Sphingobacteriaceae	1E-05	-6.43
			p_Dependentiae	5E-04	-5.52
			p_Dependentiae;c_Babeliae	5E-04	-5.52
			p_Dependentiae;c_Babeliae;o_Babeliales	5E-04	-5.52
		OTU_50	p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus;s_Multi-affiliation	4E-02	-4.27
		OTU_62	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Clostridiaceae;g_Clostridium;sensu stricto;f1;s_Multi-affiliation	2E-04	-5.05
			p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae	9E-03	-4.26
			p_Firmicutes;c_Clostridia;o_Clostridiales;f_Peptostreptococcaceae	1E-17	9.51
			p_Firmicutes;c_Clostridia;o_Clostridiales;f_Peptostreptococcaceae;g_Romboutsia	4E-05	5.81
		OTU_49	p_Firmicutes;c_Clostridia;o_Clostridiales;f_Peptostreptococcaceae;g_Romboutsia;s_metagenome	5E-11	8.41
			p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae	1E-03	-4.84
			p_Proteobacteria;c_Alphaproteobacteria;o_Acetobacterales	4E-03	4.39
			p_Proteobacteria;c_Alphaproteobacteria;o_Acetobacterales;f_Acetobacteraceae	4E-03	4.39
		OTU_5	p_Proteobacteria;c_Alphaproteobacteria;o_Caulobacterales;f_Caulobacteraceae;g_Unknown;s_Unknown	3E-02	1.14
		OTU_42	p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g_Sphingomonas;s_Multi-affiliation	7E-06	-5.61
		OTU_51	p_Proteobacteria;c_Gammaproteobacteria;o_Betaproteobacteriales;f_Burkholderiaceae;g_Ralstonia;s_Multi-affiliation	3E-02	3.89
			p_Proteobacteria;c_Gammaproteobacteria;o_Betaproteobacteriales;f_Neisseriaceae;g_Neisseria	9E-04	4.24
			p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales;f_Enterobacteriaceae;g_Pantoea	4E-07	7.85
		OTU_46	p_Proteobacteria;c_Gammaproteobacteria;o_Pseudomonadales;f_Moraxellaceae;g_Acinetobacter;s_Multi-affiliation	3E-15	9.11
		OTU_31	p_Proteobacteria;c_Gammaproteobacteria;o_Pseudomonadales;f_Moraxellaceae;g_Acinetobacter;s_Multi-affiliation	1E-06	-6.53



Figure 4. Different taxa between colorectal cancer (CRC) cases and controls and intestinal adenomas (IA) group by DESeq2 analyses. The taxonomic lineage of each taxon is shown: p, phylum; c, class; o, order; f, family; g, genus; OTU#, Operational Taxonomic Unit. The first two columns show the logarithmic transformation of normalized base mean value for each group. The “padj” column shows the p-value for heterogeneity between groups adjusted for multi-testing analyses. Positive fold changes (shown on a green background) designate taxon overrepresentation in the CRC group. Negative fold changes (shown on a yellow background) designate taxon underrepresentation in the CRC group.

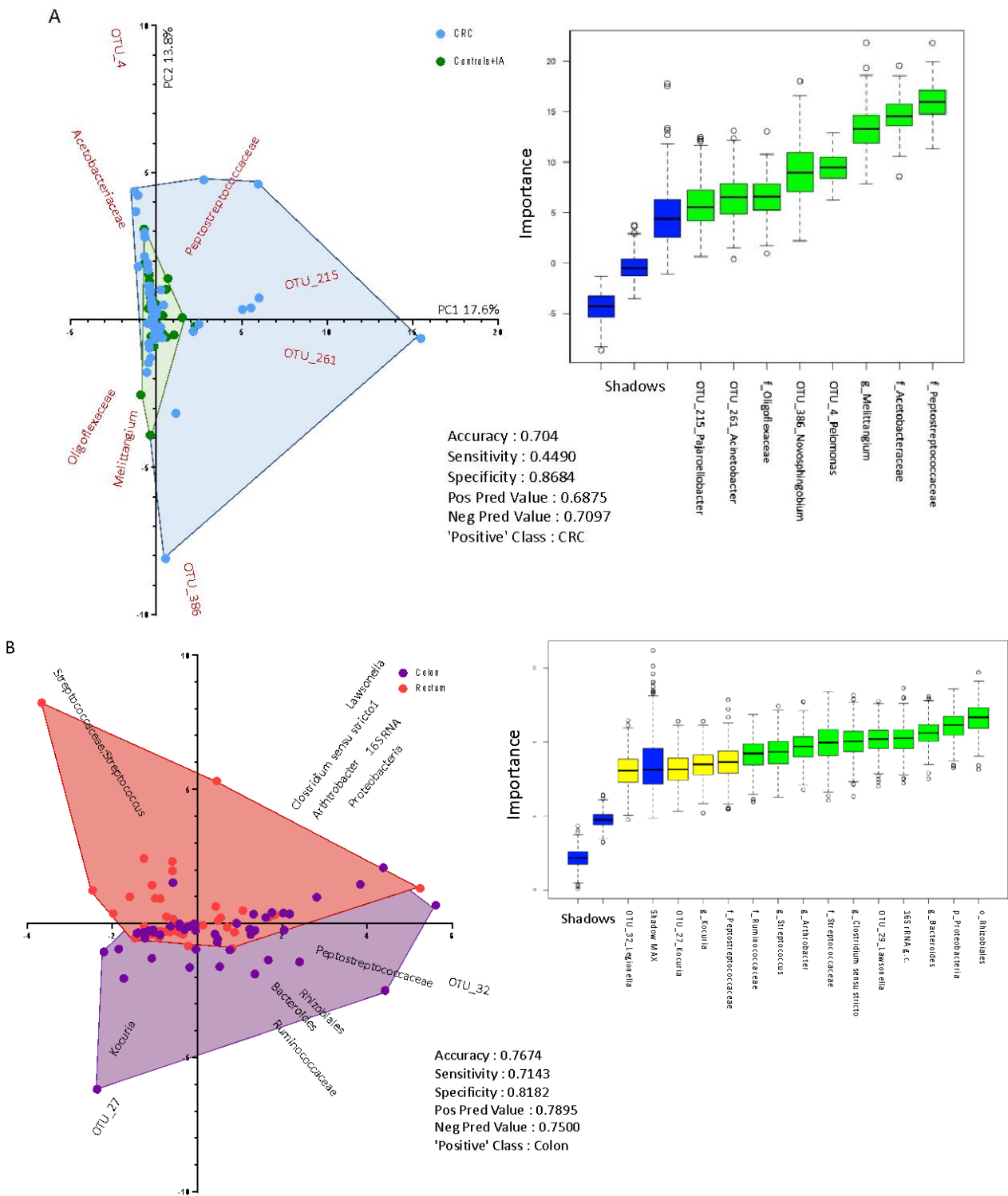


Figure 5. Biplot of predictive variables discriminating (A) colorectal cancer (CRC) versus control and intestinal adenoma (IA) subjects; (B) colon versus rectal cancer cases, using Random Forest algorithm. The boxplot on the right side of each figure shows the importance (based on mean accuracy level) of the variables by Boruta feature selection. The shadows are part of the Boruta algorithm and show the max, medium and lowest level of mean accuracy, using the same dataset with group labels shuffled. The table in the middle part of each figure shows the Random Forest results considering the 'Positive' Class as indicated.

3. Discussion

This study shows that colon cancer patients have an overrepresentation of bacterial DNA in blood as compared to tumor-free subjects, including IA or healthy controls. This result appeared stronger for cancers in right colon, whereas no difference in terms of bacterial load was found for rectal cancer. For high levels of 16S rRNA gene copies (>7618), colon cancers had increased community diversity but did not differ on community evenness from controls.

To our knowledge, no other study conducted an ad hoc data collection to systematically investigate blood bacterial DNA load and profiling in relation to the risk of CRC and/or IA to date. In a Chinese study, circulating bacterial DNA of 25 CRC, 10 IA and 22 healthy subjects was analyzed through whole genome sequencing techniques on plasma samples [21], suggesting that the *Flavobacterium* DNA relative abundance was reduced in CRC and IA (<1%) as compared to control subjects (9.4%); on the contrary, there was a 10-fold increase DNA abundance of genus *Ruminococcus* in CRC (0.2%) as compared to controls (0.02%). In the same study, various attempts to identify bacterial biomarkers of CRC or IA through Random Forest algorithms were proposed, reporting a set of 28 species as important features to discriminate between CRC/IA group and controls, but results were based on small samples (23 and 34 subjects for discovery and validation cohort, respectively). Messaritakis et al. used PCR for the amplification of genomic DNA on blood in order to compare 397 adjuvant or metastatic CRC patients with 32 healthy controls in terms of the presence of 3 bacterial genes [22]. Higher rates of the glutamine synthase gene of *Bacteroides fragilis* and the 5.8S rRNA of *Candida albicans* were observed in CRC patients ($p < 0.001$), especially in metastatic disease. No association was found for the genus *E. coli*.

In our data, CRC patients had an enrichment of Peptostreptococcaceae and Acetobacteriaceae and a reduction of Bacteroidaceae, Lachnospiraceae, and Ruminococcaceae. The latter families are the most represented in the fecal and intestinal microbiota [23], while Peptostreptococcaceae and Acetobacteriaceae are less represented in the human intestine. These two families were also found to be more abundant in chronic kidney disease (CKD) than healthy subjects in a case-control study on blood microbiome including 20 CKD cases and 20 controls [24]. In a study involving 99 CRC cases and 103 controls, high abundance of Lachnospiraceae was negatively correlated with colonic colonisation by oral bacteria, including oral pathogens associated with CRC, suggesting a protective role of Lachnospiraceae, potentially influenced by Western dietary patterns [25]. Ruminococcaceae abundance was found to decrease in CRC tissue as compared to tumor-adjacent biopsies and stool samples from the same case in a study including 294 subjects [26]. Moreover, the members of the families Lachnospiraceae and Ruminococcaceae were recognized as the most active members of the human colonic microbiota [27], able to efficiently convert fibers into butyrate [28], a bacterial catabolite widely demonstrated to regulate T-reg lymphocyte priming preventing CRC [29,30].

Gut microbiota, inflammation and nutrition play an important role on intestinal permeability which may influence the risk of CRC [31,32]. Bacteria have shown capacity to interact directly with immune system cells and to impact in multiple host functions [33,34], but it is unclear which one comes first between local inflammation, intestinally permeability and changes in resident microbiota [35,36]. In this context, it has also been shown that bacteria can disseminate to liver through the disruption of gut vascular barrier in colorectal cancer with hepatic metastasis [37].

Our data corroborate the hypothesis of a greater bacterial translocation from gastrointestinal tract to bloodstream in colon cancer, especially in right colon cancer, but not in rectal cancer patients. Although this result should be interpreted with caution given the reduced number of cases, various etiological factors and biological aspects are different according to CRC sites. Physical activity, antibiotic use and family history of CRC are relevant in colon but not in rectal cancer [37–40], and associations of some dietary components are stronger for some colon subsites cancers [41,42]. Moreover, the proximal and distal colon have a different embryological origin, also resulting in a distinct vascular supply [43].

Local differences in underlying genetics, including genetic expression and immunological activity have been highlighted [44], along with a negative gradient of immune cells from proximal colon to distal colon and rectum [43]. These features can play a role in changing the extent of bacterial translocation along the lower intestinal tract, leading to differences in blood 16S rRNA gene copies according to CRC localization. Moreover, gut microbiota and mechanisms of carcinogenesis also vary from left to right colon and to rectum [40,45]. Another important factor that could affect regional differences in bacterial translocation and blood 16S rRNA gene copies is the organization and thickness of mucus, which appears to be looser and less organized in the proximal colon, facilitating bacterial contact with the epithelium and, thus, translocation [46,47]. Mucus layers were found to vary along the colon in mice in terms of O-glycosylated entities of Muc2 and of the host-microbiota symbiosis regulation [48].

Recent meta-analyses on fecal microbiome suggested universal, validated predictive taxonomic and functional microbiome CRC signatures, revealing potential mechanisms behind the intestinal carcinogenesis processes, and putting the basis for non-invasive CRC diagnosis through metagenomic analysis of fecal microbiome [14,15]. CRC screening programs can suffer from limited sensitivity and specificity of the tests used and from possible low adherence with CRC screening recommendations—mainly due to the refusal of fecal test and colonoscopy—in some countries [49]. Innovative, non-invasive diagnostic tests would be of support for CRC control [49]. Various tests based on blood analyses have been proposed, including techniques estimating the presence of *Streptococcus bovis* [50] or evaluating the antibody level against *Fusobacterium nucleatum* [51] in blood, and, more recently, the use of blood microbiome profile has been suggested [20].

Our data showed various taxa associated with CRC and identified a set of covariates of taxa and OTUs that is able to predict CRC from control/IA subjects and colon from rectal cancers with an accuracy around 0.7. When restricting the analyses to subjects with high levels of 16S rRNA gene copies, we found more accurate models, but we were not able to separate analysis for colon cancer only and results should be interpreted with caution given the small numbers. However, these findings can serve as a basis to conceive new non-invasive techniques for an early diagnosis of CRC based on bacterial DNA circulating in peripheral blood. In particular, they can be relevant for the detection of right colon cancers, which often have a subtle presentation and a more advanced stage at diagnosis, partly because right colon is more difficult to be explored, when compared to rectal and distal colon [52].

Our data show a non-trivial dysfunction of the intestinal epithelial barrier permeability also for patients without cancer, in line with previous studies that characterized blood microbiota among healthy subjects [17]. It is not easy to explain the lack of difference between rectal cancer and tumor-free subjects, since the bacterial translocation can be different due to the related inflammatory condition in rectal cancer patients [53]. A possible role of size or shape selectivity for bacteria transiting from the intestine to the bloodstream in different locations could explain at least in part this result but a limitation of this study is the lack of fecal samples on our subjects, which prevented us from further analyzing this issue. One of the strengths of this study is the conduction of an ad-hoc data collection, which includes a new developed standardized protocol, fully observed by the recruitment centers. CRC and the corresponding IA and control subjects were comparable in terms of setting since they derived from the same catchment area and recruitment procedures. Moreover, interviewers and investigators were blinded to the group assignment, as data collection was performed before endoscopy and diagnosis. Most cases were detected at the first CRC-diagnosing colonoscopy, allowing us to recruit truly incident cases, characterized by a minimal time between participant's recruitment and cancer diagnosis and by available clinical data from the very beginning of the diagnostic process. Moreover, the presence of healthy controls allowed a clean comparison with CRC and IA, and the inclusion of IA allowed to investigate an important phase on the mechanisms behind the process of the adenoma-carcinoma sequence. However, the lack of an ad hoc developed validation

cohort remains a limitation of our work. We were able to adjust for study center, sex and age, eliminating possible confounding effects of these covariates on 16s rRNA gene copies results.

In conclusion, our data confirm the presence of bacterial DNA in blood in healthy adults and indicate that colon cancer patients had a higher DNA bacterial load and a different bacterial profiling as compared to healthy, IA and rectal cancer subjects, revealing a higher passage of bacteria from gastrointestinal tract to bloodstream in colon cancer. Further studies are needed to confirm this result and possibly exploit it for the development of innovative early techniques for colon cancer diagnosis.

4. Materials and Methods

We conducted an ad hoc developed case-control study between May 2017 and November 2019 in the metropolitan area of Milan, Italy. Recruitment centers included two general hospitals of Milan: the Digestive and Interventional Endoscopy Unit, Azienda Socio Sanitaria Territoriale (ASST) Grande Ospedale Metropolitano Niguarda, the coordinator center, and the Gastroenterology and Endoscopy Unit, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Ca' Granda Ospedale Maggiore Policlinico. Both hospitals included a colonoscopy screening referral center of the CRC screening program, managed by Health Protection Agency.

CRC cases were enrolled together with non-cancer adenomatous polyps and healthy controls, frequency-matched with cases by center, age ± 5 years and sex. Trained interviewers recruited study participants among eligible outpatients or inpatients who were scheduled for a colonoscopy, including patients referred for the CRC screening program. Excluded criteria were: (1) colonoscopy in the last 5 days; (2) reported previous cancers; (3) inflammatory chronic bowel diseases, (4) liver or kidney failure (creatinine ≥ 1.7 mg/dL, dialysis); (5) NYHA grade III or IV heart failure; (6) primary or secondary immunodeficiency; (7) recent hospitalization (1 month) for immune, inflammatory, autoimmune diseases, or bacterial/viral infections, (8) blood transfusions in the previous year; (9) celiac disease and a relevant diet modification during the last month. IA and control subjects with previous colonoscopy/sigmoidoscopy with endoscopic resection of a colonic lesion were also excluded.

A total of 620 patients were contacted by the trained interviewers. Of these, around 25% did not meet the eligibility criteria and less than 2% refused to participate in the study. Furthermore, 49 subjects were excluded due to some inaccuracies in the enrolment procedures and 42 due to previous cancers or to other ineligible conditions that were discovered after further data check. From the remaining 347 patients, the final sample after matching included 300 subjects: 100 CRC cases, 100 IA patients and 100 healthy controls.

Colonoscopy and histological examinations were revised by two pathologists who determined CRC cases and their clinical characteristics (e.g., stage, lymph node), as well as IA and their major features (e.g., morphology, measure), and healthy controls.

Cases included 100 incident, histologically confirmed CRC: 62 men and 38 women (mean age: 67, range 31–85 years). Of these, 21 were in the right colon (International Classification of Diseases, 10th Edition, ICD-10, C18.0, C18.2, C18.3), 12 in the transverse colon, in the splenic flexure, and in the descending colon (ICD-10, C18.4, C18.5, C18.6), 17 in the sigmoid colon (ICD-10, C18.7), and 50 in the rectum, including the rectosigmoid junction (ICD-10, C20, C19.9). According to the TNM system, 2 CRC were stage I, 49 stage II, 26 stage III, and 19 stage IV (4 had missing information).

One hundred IA patients (mean age: 66, range 34–84 years) and 100 healthy controls (mean age: 66, range 26–85 years) were included.

The study protocol was revised and approved by the ethical committees of the hospitals involved in data recruitment: ASST Grande Ospedale Metropolitano Niguarda (No. 477-112016) and Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico (No. 742-2017).

4.1. Interview

After written consent, a face-to-face interview was performed. The questionnaire included information on socio-demographics, smoking habits, physical activity, anthropometric measures, occupational exposures, medical history, selected drug and supplement use, family history of cancer, sleeping habits and dental care. A food frequency questionnaire was used to assess the past patients' usual diet.

4.2. Blood Collection

Blood samples were collected before the colonoscopy in order to avoid possible bacteria contamination after colonoscope insertion and to keep the same setting for each participant. An aliquot of 7 mL of blood was collected in a tube with EDTA and an aliquot of 3 mL in a blank (without anticoagulant) tube. Three microvials of 1 mL from EDTA tube were immediately stored at $-180\text{ }^{\circ}\text{C}$ for the microbiomic analysis. The remaining blood was processed and centrifuged, and then stored at $-80\text{ }^{\circ}\text{C}$. At the end of data recruitment, blood samples were sent to Vaiomer SAS, Labège, France, for the analysis of the microbiome. The operators were blind to the group assignment and the samples were analyzed in the same experiment, with the same reagent batches and manipulator, in order to keep the signal to noise ratio optimal and to reduce technical variability.

4.3. DNA Extraction, qPCR Experiments and Sequencing of 16S rRNA Gene Amplicons

Bacterial DNA quantification and sequencing reactions were performed by Vaiomer SAS using an optimized blood-specific technique [17,54,55]. DNA was extracted from 0.25 mL of whole blood and collected in a final 50 μL extraction volume. Real-time polymerase chain reaction (PCR) amplification was performed using panbacterial primers EUBF 5'-TCCTACGGGAGGCAGCAGT-3' and EUBR 5'-GGACTACCAGGTATCTAATCCTGTT-3' [56], which target the V3-V4 hypervariable regions of the bacterial 16S rRNA gene with 100% specificity (i.e., no eukaryotic, mitochondrial, or Archaea DNA is targeted) and high sensitivity (16S rRNA of more than 95% of bacteria in Ribosomal Database Project database are amplified). The abundance of the 16S rRNA gene in blood samples was measured by qPCR in triplicate and normalized using a plasmid-based standard range. The results were reported as number of copies of 16S rRNA gene per μL of blood. DNA from whole blood was also used for 16S rRNA gene taxonomic profiling applying MiSeq Illumina technology (2×300 paired-end MiSeq kit V3). The samples 20,056, 10,251, 10,248 and 20,086 (referring to 2 IA and 2 control subjects) were excluded from the diversity analyses as they did not reach the threshold of 5000 reads.

Then, sequences were analyzed using Vaiomer bioinformatic pipeline to determine bacterial community profiles. Briefly, after demultiplexing of the bar-coded Illumina paired reads, single read sequences were trimmed (reads R1 and R2 to 290 and 240 bases, respectively) and paired for each sample independently into longer fragments, non-specific amplicons were removed and remaining sequences were clustered into OTUs using FROGS v1.4.0 [57] with default parameters. A taxonomic assignment was performed by Blast+ v2.2.30 against the Silva 132 Parc database. OTUs were clustered based on 97% sequence similarities by two steps through swarm algorithm v2.1.6. The first step consisted of a clustering with an aggregation distance equal to 1. The second step consisted of a clustering with an aggregation distance equal to 3. OTUs with relative abundance lower than 0.005% of the whole dataset of reads were removed. All the reads are publicly available in the European Nucleotide Archive (ENA) with the accession number: PRJEB46474.

4.4. Bacterial DNA Contamination Assessment

To assess the potential bacterial DNA contamination from environment and reagents, several negative controls were included in the analyses (Supplementary Table S2, Supplementary Figure S4 and Supplementary Methods) showing that background noise and blood contamination did not impact the results of this study.

4.5. Statistical Analyses

Two-tailed Wilcoxon signed-rank tests and Friedman tests were used to compare 16S rRNA gene between groups. Since 16S rRNA gene distribution was very similar in control and IA subjects, we grouped the tumor-free subjects as comparison group and increased the study power. ORs of CRC and their corresponding 95% CI were estimated through logistic regression models conditioned on the matching variable. The number of 16S rRNA gene copies was included in the models as quintiles (categorically) based on the distribution of controls and IAs, and as continuous variables, with the measurement unit sets to the difference between the upper cut-points of the 4th and 1st quintile (equal to 4328). Multinomial logistic regression was used to estimate separate ORs for colon and rectal cancer and to test for heterogeneity between the two sites.

Analysis on alpha-diversity and beta-diversity indices, as well as on taxonomic variables were computed among 296 subjects (because of 4 missing data due to technical reasons described above).

To assess samples diversity in terms of richness and evenness, various alpha-diversity indices, including Observed, Chao1, Shannon, Simpson and InvSimpson, were calculated by R PhyloSeq v1.14.0 package. Two-tailed Mann–Whitney tests were used to determine differences in terms of alpha-diversity between groups.

To estimate beta-diversity, Permutational Multivariate Analysis of Variance Using Distance Matrices (PERMANOVA) was applied based on the UniFrac distances, and Principal Coordinates Analysis (PCoA) was applied to visualize possible differences between groups.

Differences in terms of bacterial taxa and OTUs were evaluated through Welch test after DESeq2 normalization of data, based on negative binomial distribution (R package “DESeq2” v1.26.0).

For each statistical analysis, a post-hoc *p*-value adjustment was performed using Hochberg–Benjamini correction, when appropriate.

Random Forest (R libraries: “randomforest”, “caret”, “Boruta”) was used to infer whether there was a set of variables able to discriminate which group the samples belonged to, by the use of training and test sets randomly selected from our samples. In order to decrease the background noise due to the different library size of the samples sequenced, data were normalized: the relative abundance of taxa was multiplied by 16S rRNA gene abundance (determined by qPCR).

5. Conclusions

Data confirm the presence of bacterial DNA in blood in healthy adults and indicate that colon cancer patients had a higher DNA bacterial load as compared to healthy, adenoma and rectal cancer subjects. This result supports the hypothesis of a higher passage of bacteria from gastrointestinal tract to bloodstream, with increased community diversity, in colon cancer patients but not in rectal cancer and IA patients.

Moreover, this study found a set of bacterial taxa and OTUs able to discriminate CRC from IA and healthy subjects with an accuracy of 0.70. The confirmation and suitability of this finding in non-invasive diagnostic tests for colon cancer control should be evaluated in larger studies.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/cancers13246363/s1>, Table S1: Distributions of Observed, Chao1, Shannon and Simpson alpha-diversity indices of control, intestinal adenoma (IA), colon and rectal cancer subjects for bacterial genera and OTUs; Table S2: PERMANOVA and PERMDISP on beta-diversity Bray–Curtis index for the 10 randomized comparisons between negative controls (CTRL) and whole blood samples (WB). For all randomized comparisons, the sample size is equal to 20 (10 Neg CTRL vs. 10 WB) at the exception of RandExt16 (10 Neg CTRL vs. 9 WB). The statistical testing used were pseudo-F and F-value for PERMANOVA and PERMDISP, respectively. For both tests, 2000 permutations were performed. *p*-values < 0.05 in yellow, < 0.005 in orange and < 0.0005 in red; Figure S1: Distribution of 16S rRNA gene copies per μ L of whole blood among controls, intestinal adenomas and colorectal

cancers; Figure S2: Mean of the taxonomy composition in terms of relative abundance per group of samples at family taxonomic level. The panel on the left shows the taxonomic composition of controls, intestinal adenoma subjects (IA) and both controls/IA, respectively (as per legend to the right side). The panel on the right shows the taxonomic composition of colorectal cancer (CRC), and rectal and colon cancer cases separately (as per legend to the right side); Figure S3: Different taxa between: (A) colorectal cancers (CRC) and controls; (B) CRC and intestinal adenoma (IA); (C) IA and controls by DESeq2 analyses. The taxonomic lineage of each taxon is shown: p, phylum; c, class; o, order; f, family; g, genus; OTU#, Operational Taxonomic Unit. The first two columns show the logarithmic transformation of normalized base mean value for each group. The “padj” column shows the *p*-value for heterogeneity between groups adjusted for multi-testing analyses. Positive fold changes (shown on a green background) designate taxon overrepresentation in the CRC group. Negative fold changes (shown on a yellow background) designate taxon underrepresentation in the CRC group; Figure S4: Multidimensional scaling (MDS) differentiates bacterial patterns of study blood samples and DNA extraction negative controls on Bray-Curtis distance matrices for all 10 randomized comparisons. Study blood samples (in Red) separated from negative controls (in Blue) in all of the randomized comparisons.

Author Contributions: Conceptualization: M.R., S.G. and C.L.V. Methodology and Data curation: M.R., S.G., C.L.V., M.P., R.P., A.A., M.C., P.B., P.C. (Pietro Carnevali), R.B., M.V. (Maurizio Vecchi) and M.M. Formal analysis of data: M.R., G.G. and S.G. Interpretation of data: M.R., S.G., C.L.V., R.P., A.A., M.M., M.C., G.G. and R.B. Drafting the manuscript: M.R., S.G., G.G., C.L.V., M.C.S. and R.B. Data collection, critical revision and final approval of the manuscript: M.M., R.P., G.G., S.G., M.C., A.A., P.L., P.C. (Pietro Carnevali), C.C., G.P., F.M., B.O., L.D., F.C. (Federica Cavalcoli), M.P., F.P., P.B., P.S., C.M., G.F., M.C.S., G.B., M.B., P.C. (Paolo Cantù), F.C. (Flavio Caprioli), M.V. (Marcello Vangeli), E.F., S.M., G.T., R.B., M.V. (Maurizio Vecchi), C.L.V. and M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Italian Foundation for Cancer Research (AIRC) (My First AIRC grant No. 17070).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committees of ASST Grande Ospedale Metropolitano Niguarda (No. 477-112016; 25 November 2016) and Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico (No. 742-2017; 14 December 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: All the reads are publicly available in the European Nucleotide Archive (ENA) with the accession number: PRJEB46474.

Acknowledgments: The authors would like to express their most sincere gratitude and appreciation to all participants and collaborators to this study, without whose effort this work would not have been feasible. A special thanks to Margherita Cozzi, Elena Tansi, Cinzia Della Noce, Rosa Restieri, Nadia Zaretti for their valuable involvement in this study. We thank all the nursing staff at the Digestive and Interventional Endoscopy Unit, ASST Grande Ospedale Metropolitano Niguarda, Milan, and at the Gastroenterology and Endoscopy Unit, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milan, and Valentina Taverniti, Guido Basilisco, Luca Elli, Francesca Ferretti, Gian Lorenzo Scacchi, Gian Eugenio Tontini, Giuseppe Torgano for their contribution in the data collection. A thankful mention to Annalisa Pascarella for the help in the graphic preparation and to Patrizia Riso for participating in the concept of the project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. Ait Ouakrim, D.; Pizot, C.; Boniol, M.; Malvezzi, M.; Boniol, M.; Negri, E.; Bota, M.; Jenkins, M.A.; Bleiberg, H.; Autier, P. Trends in colorectal cancer mortality in europe: Retrospective analysis of the who mortality database. *BMJ* **2015**, *351*, h4970. [[CrossRef](#)] [[PubMed](#)]

3. Carioli, G.; Bertuccio, P.; Boffetta, P.; Levi, F.; La Vecchia, C.; Negri, E.; Malvezzi, M. European cancer mortality predictions for the year 2020 with a focus on prostate cancer. *Ann. Oncol.* **2020**, *31*, 650–658. [[CrossRef](#)] [[PubMed](#)]
4. Vogelstein, B.; Papadopoulos, N.; Velculescu, V.E.; Zhou, S.B.; Diaz, L.A.; Kinzler, K.W. Cancer genome landscapes. *Science* **2013**, *339*, 1546–1558. [[CrossRef](#)] [[PubMed](#)]
5. Terzic, J.; Grivennikov, S.; Karin, E.; Karin, M. Inflammation and colon cancer. *Gastroenterology* **2010**, *138*, 2101–2114.e5. [[CrossRef](#)]
6. Herrinton, L.J.; Liu, L.; Levin, T.R.; Allison, J.E.; Lewis, J.D.; Velayos, F. Incidence and mortality of colorectal adenocarcinoma in persons with inflammatory bowel disease from 1998 to 2010. *Gastroenterology* **2012**, *143*, 382–389. [[CrossRef](#)]
7. Godos, J.; Biondi, A.; Galvano, F.; Basile, F.; Sciacca, S.; Giovannucci, E.L.; Grosso, G. Markers of systemic inflammation and colorectal adenoma risk: Meta-analysis of observational studies. *World J. Gastroenterol.* **2017**, *23*, 1909–1919. [[CrossRef](#)]
8. Huang, W.Y.; Berndt, S.I.; Shiels, M.S.; Katki, H.A.; Chaturvedi, A.K.; Wentzensen, N.; Trabert, B.; Kemp, T.J.; Pinto, L.A.; Hildesheim, A.; et al. Circulating inflammation markers and colorectal adenoma risk. *Carcinogenesis* **2019**, *40*, 765–770. [[CrossRef](#)]
9. Liu, X.; Cheng, Y.; Shao, L.; Ling, Z. Alterations of the predominant fecal microbiota and disruption of the gut mucosal barrier in patients with early-stage colorectal cancer. *Biomed. Res. Int.* **2020**, *2020*, 2948282. [[CrossRef](#)]
10. Ahn, J.; Sinha, R.; Pei, Z.; Dominianni, C.; Wu, J.; Shi, J.; Goedert, J.J.; Hayes, R.B.; Yang, L. Human gut microbiome and risk for colorectal cancer. *J. Natl. Cancer Inst.* **2013**, *105*, 1907–1911. [[CrossRef](#)]
11. Song, M.; Chan, A.T.; Sun, J. Influence of the gut microbiome, diet, and environment on risk of colorectal cancer. *Gastroenterology* **2020**, *158*, 322–340. [[CrossRef](#)] [[PubMed](#)]
12. Shah, M.S.; DeSantis, T.Z.; Weinmaier, T.; McMurdie, P.J.; Cope, J.L.; Altrichter, A.; Yamal, J.M.; Hollister, E.B. Leveraging sequence-based faecal microbial community survey data to identify a composite biomarker for colorectal cancer. *Gut* **2018**, *67*, 882–891. [[CrossRef](#)] [[PubMed](#)]
13. Yu, J.; Feng, Q.; Wong, S.H.; Zhang, D.; Liang, Q.Y.; Qin, Y.; Tang, L.; Zhao, H.; Stenvang, J.; Li, Y.; et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* **2017**, *66*, 70–78. [[CrossRef](#)]
14. Thomas, A.M.; Manghi, P.; Asnicar, F.; Pasolli, E.; Armanini, F.; Zolfo, M.; Beghini, F.; Manara, S.; Karcher, N.; Pozzi, C.; et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* **2019**, *25*, 667–678. [[CrossRef](#)]
15. Wirbel, J.; Pyl, P.T.; Kartal, E.; Zych, K.; Kashani, A.; Milanese, A.; Fleck, J.S.; Voigt, A.Y.; Palleja, A.; Ponnudurai, R.; et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat. Med.* **2019**, *25*, 679–689. [[CrossRef](#)]
16. Castillo, D.J.; Rifkin, R.F.; Cowan, D.A.; Potgieter, M. The healthy human blood microbiome: Fact or fiction? *Front. Cell. Infect. Microbiol.* **2019**, *9*, 148. [[CrossRef](#)]
17. Paise, S.; Valle, C.; Servant, F.; Courtney, M.; Burcelin, R.; Amar, J.; Lelouvier, B. Comprehensive description of blood microbiome from healthy donors assessed by 16s targeted metagenomic sequencing. *Transfusion* **2016**, *56*, 1138–1147. [[CrossRef](#)] [[PubMed](#)]
18. Lelouvier, B.; Servant, F.; Paise, S.; Brunet, A.C.; Benyahya, S.; Serino, M.; Valle, C.; Ortiz, M.R.; Puig, J.; Courtney, M.; et al. Changes in blood microbiota profiles associated with liver fibrosis in obese patients: A pilot analysis. *Hepatology* **2016**, *64*, 2015–2027. [[CrossRef](#)]
19. Sato, J.; Kanazawa, A.; Ikeda, F.; Yoshihara, T.; Goto, H.; Abe, H.; Komiya, K.; Kawaguchi, M.; Shimizu, T.; Ogihara, T.; et al. Gut dysbiosis and detection of “live gut bacteria” in blood of Japanese patients with type 2 diabetes. *Diabetes Care* **2014**, *37*, 2343–2350. [[CrossRef](#)]
20. Poore, G.D.; Kopylova, E.; Zhu, Q.; Carpenter, C.; Fraraccio, S.; Wandro, S.; Kosciulek, T.; Janssen, S.; Metcalf, J.; Song, S.J.; et al. Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **2020**, *579*, 567–574. [[CrossRef](#)] [[PubMed](#)]
21. Xiao, Q.; Lu, W.; Kong, X.; Shao, Y.W.; Hu, Y.; Wang, A.; Bao, H.; Cao, R.; Liu, K.; Wang, X.; et al. Alterations of circulating bacterial DNA in colorectal cancer and adenoma: A proof-of-concept study. *Cancer Lett.* **2021**, *499*, 201–208. [[CrossRef](#)]
22. Messaritakis, I.; Vogiatzoglou, K.; Tsantaki, K.; Ntretaki, A.; Sfakianaki, M.; Koulouridi, A.; Tsiaoussis, J.; Mavroudis, D.; Souglakos, J. The prognostic value of the detection of microbial translocation in the blood of colorectal cancer patients. *Cancers* **2020**, *12*, 1058. [[CrossRef](#)]
23. King, C.H.; Desai, H.; Sylvetsky, A.C.; LoTempio, J.; Ayanyan, S.; Carrie, J.; Crandall, K.A.; Fochtman, B.C.; Gasparyan, L.; Gulzar, N.; et al. Baseline human gut microbiota profile in healthy people and standard reporting template. *PLoS ONE* **2019**, *14*, e0206484. [[CrossRef](#)]
24. Shah, N.B.; Allegretti, A.S.; Nigwekar, S.U.; Kalim, S.; Zhao, S.; Lelouvier, B.; Servant, F.; Serena, G.; Thadhani, R.I.; Raj, D.S.; et al. Blood microbiome profile in CKD: A pilot study. *Clin. J. Am. Soc. Nephrol.* **2019**, *14*, 692–701. [[CrossRef](#)] [[PubMed](#)]
25. Flemer, B.; Warren, R.D.; Barrett, M.P.; Cisek, K.; Das, A.; Jeffery, I.B.; Hurley, E.; O’Riordain, M.; Shanahan, F.; O’Toole, P.W. The oral microbiota in colorectal cancer is distinctive and predictive. *Gut* **2018**, *67*, 1454–1463. [[CrossRef](#)]
26. Shah, M.S.; DeSantis, T.; Yamal, J.M.; Weir, T.; Ryan, E.P.; Cope, J.L.; Hollister, E.B. Re-purposing 16s rRNA gene sequence data from within case paired tumor biopsy and tumor-adjacent biopsy or fecal samples to identify microbial markers for colorectal cancer. *PLoS ONE* **2018**, *13*, e0207002. [[CrossRef](#)]
27. Peris-Bondia, F.; Latorre, A.; Artacho, A.; Moya, A.; D’Auria, G. The active human gut microbiota differs from the total microbiota. *PLoS ONE* **2011**, *6*, e22448. [[CrossRef](#)]
28. Vital, M.; Karch, A.; Pieper, D.H. Colonic butyrate-producing communities in humans: An overview using omics data. *mSystems* **2017**, *2*, e00130-17. [[CrossRef](#)] [[PubMed](#)]

29. Singh, N.; Gurav, A.; Sivaprakasam, S.; Brady, E.; Padia, R.; Shi, H.; Thangaraju, M.; Prasad, P.D.; Manicassamy, S.; Munn, D.H.; et al. Activation of gpr109a, receptor for niacin and the commensal metabolite butyrate, suppresses colonic inflammation and carcinogenesis. *Immunity* **2014**, *40*, 128–139. [[CrossRef](#)] [[PubMed](#)]
30. Venkataraman, A.; Sieber, J.R.; Schmidt, A.W.; Waldron, C.; Theis, K.R.; Schmidt, T.M. Variable responses of human microbiomes to dietary supplementation with resistant starch. *Microbiome* **2016**, *4*, 33. [[CrossRef](#)]
31. Bhat, A.A.; Uppada, S.; Achkar, I.W.; Hashem, S.; Yadav, S.K.; Shanmugakonar, M.; Al-Naemi, H.A.; Haris, M.; Uddin, S. Tight junction proteins and signaling pathways in cancer and inflammation: A functional crosstalk. *Front. Physiol.* **2018**, *9*, 1942. [[CrossRef](#)]
32. Yu, L.C. Microbiota dysbiosis and barrier dysfunction in inflammatory bowel disease and colorectal cancers: Exploring a common ground hypothesis. *J. Biomed. Sci.* **2018**, *25*, 79. [[CrossRef](#)]
33. Taverniti, V.; Stuknyte, M.; Minuzzo, M.; Arioli, S.; De Noni, I.; Scabiosi, C.; Cordova, Z.M.; Junttila, I.; Hamalainen, S.; Turpeinen, H.; et al. S-layer protein mediates the stimulatory effect of lactobacillus helveticus mimlh5 on innate immunity. *Appl. Environ. Microbiol.* **2013**, *79*, 1221–1231. [[CrossRef](#)]
34. Zheng, D.; Liwinski, T.; Elinav, E. Interaction between microbiota and immunity in health and disease. *Cell Res.* **2020**, *30*, 492–506. [[CrossRef](#)]
35. Chelakkot, C.; Ghim, J.; Ryu, S.H. Mechanisms regulating intestinal barrier integrity and its pathological implications. *Exp. Mol. Med.* **2018**, *50*, 1–9. [[CrossRef](#)]
36. Wasinger, V.C.; Lu, K.; Yau, Y.Y.; Nash, J.; Lee, J.; Chang, J.; Paramsothy, S.; Kaakoush, N.O.; Mitchell, H.M.; Leong, R.W.L. Spp24 is associated with endocytic signalling, lipid metabolism, and discrimination of tissue integrity for ‘leaky-gut’ in inflammatory bowel disease. *Sci. Rep.* **2020**, *10*, 12932. [[CrossRef](#)] [[PubMed](#)]
37. Bertocchi, A.; Carloni, S.; Ravenda, P.S.; Bertalot, G.; Spadoni, I.; Lo Cascio, A.; Gandini, S.; Lizier, M.; Braga, D.; Asnicar, F.; et al. Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver. *Cancer Cell* **2021**, *39*, 708–724.e711. [[CrossRef](#)] [[PubMed](#)]
38. Paschke, S.; Jafarov, S.; Staib, L.; Kreuser, E.D.; Maulbecker-Armstrong, C.; Roitman, M.; Holm, T.; Harris, C.C.; Link, K.H.; Kornmann, M. Are colon and rectal cancer two different tumor entities? A proposal to abandon the term colorectal cancer. *Int. J. Mol. Sci.* **2018**, *19*, 2577. [[CrossRef](#)] [[PubMed](#)]
39. Wei, E.K.; Giovannucci, E.; Wu, K.; Rosner, B.; Fuchs, C.S.; Willett, W.C.; Colditz, G.A. Comparison of risk factors for colon and rectal cancer. *Int. J. Cancer* **2004**, *108*, 433–442. [[CrossRef](#)]
40. Zhang, J.; Haines, C.; Watson, A.J.M.; Hart, A.R.; Platt, M.J.; Pardoll, D.M.; Cosgrove, S.E.; Gebo, K.A.; Sears, C.L. Oral antibiotic use and risk of colorectal cancer in the United Kingdom, 1989–2012: A matched case-control study. *Gut* **2019**, *68*, 1971–1978. [[CrossRef](#)]
41. Hjartaker, A.; Aagnes, B.; Robsahm, T.E.; Langseth, H.; Bray, F.; Larsen, I.K. Subsite-specific dietary risk factors for colorectal cancer: A review of cohort studies. *J. Oncol.* **2013**, *2013*, 703854. [[CrossRef](#)] [[PubMed](#)]
42. Rossi, M.; Mascaretti, F.; Parpinel, M.; Serraino, D.; Crispo, A.; Celentano, E.; Giacosa, A.; La Vecchia, C. Dietary intake of branched-chain amino acids and colorectal cancer risk. *Br. J. Nutr.* **2021**, *126*, 22–27. [[CrossRef](#)]
43. Lee, G.H.; Malietzis, G.; Askari, A.; Bernardo, D.; Al-Hassi, H.O.; Clark, S.K. Is right-sided colon cancer different to left-sided colorectal cancer?—A systematic review. *Eur. J. Surg. Oncol.* **2015**, *41*, 300–308. [[CrossRef](#)]
44. Glebov, O.K.; Rodriguez, L.M.; Nakahara, K.; Jenkins, J.; Cliatt, J.; Humbyrd, C.J.; DeNobile, J.; Soballe, P.; Simon, R.; Wright, G.; et al. Distinguishing right from left colon by the pattern of gene expression. *Cancer Epidemiol. Biomark. Prev.* **2003**, *12*, 755–762.
45. Tamas, K.; Walenkamp, A.M.; de Vries, E.G.; van Vugt, M.A.; Beets-Tan, R.G.; van Etten, B.; de Groot, D.J.; Hospers, G.A. Rectal and colon cancer: Not just a different anatomic site. *Cancer Treat. Rev.* **2015**, *41*, 671–679. [[CrossRef](#)] [[PubMed](#)]
46. Herath, M.; Hosie, S.; Bornstein, J.C.; Franks, A.E.; Hill-Yardin, E.L. The role of the gastrointestinal mucus system in intestinal homeostasis: Implications for neurological disorders. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 248. [[CrossRef](#)]
47. Kamphuis, J.B.J.; Mercier-Bonin, M.; Eutamene, H.; Theodorou, V. Mucus organisation is shaped by colonic content; a new view. *Sci. Rep.* **2017**, *7*, 8527. [[CrossRef](#)]
48. Bergstrom, K.; Shan, X.; Casero, D.; Batushansky, A.; Lagishetty, V.; Jacobs, J.P.; Hoover, C.; Kondo, Y.; Shao, B.; Gao, L.; et al. Proximal colon-derived o-glycosylated mucus encapsulates and modulates the microbiota. *Science* **2020**, *370*, 467–472. [[CrossRef](#)] [[PubMed](#)]
49. Ransohoff, D.F. Evaluating a new cancer screening blood test: Unintended consequences and the need for clarity in policy making. *J. Natl. Cancer Inst.* **2021**, *113*, 109–111. [[CrossRef](#)]
50. Boleij, A.; Tjalsma, H. The itinerary of streptococcus gallolyticus infection in patients with colonic malignant disease. *Lancet Infect. Dis.* **2013**, *13*, 719–724. [[CrossRef](#)]
51. Wang, H.F.; Li, L.F.; Guo, S.H.; Zeng, Q.Y.; Ning, F.; Liu, W.L.; Zhang, G. Evaluation of antibody level against fusobacterium nucleatum in the serological diagnosis of colorectal cancer. *Sci. Rep.* **2016**, *6*, 33440. [[CrossRef](#)] [[PubMed](#)]
52. Franco, D.L.; Leighton, J.A.; Gurudu, S.R. Approach to incomplete colonoscopy: New techniques and technologies. *Gastroenterol. Hepatol.* **2017**, *13*, 476–483.
53. Genua, F.; Raghunathan, V.; Jenab, M.; Gallagher, W.M.; Hughes, D.J. The role of gut barrier dysfunction and microbiome dysbiosis in colorectal cancer development. *Front. Oncol.* **2021**, *11*, 626349. [[CrossRef](#)] [[PubMed](#)]

-
54. Gargari, G.; Mantegazza, G.; Taverniti, V.; Del Bo, C.; Bernardi, S.; Andres-Lacueva, C.; Gonzalez-Dominguez, R.; Kroon, P.A.; Winterbone, M.S.; Cherubini, A.; et al. Bacterial dnaemia is associated with serum zonulin levels in older subjects. *Sci. Rep.* **2021**, *11*, 11054. [[CrossRef](#)]
 55. Lluch, J.; Servant, F.; Paise, S.; Valle, C.; Valiere, S.; Kuchly, C.; Vilchez, G.; Donnadieu, C.; Courtney, M.; Burcelin, R.; et al. The characterization of novel tissue microbiota using an optimized 16s metagenomic sequencing pipeline. *PLoS ONE* **2015**, *10*, e0142334. [[CrossRef](#)]
 56. Nadkarni, M.A.; Martin, F.E.; Jacques, N.A.; Hunter, N. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* **2002**, *148*, 257–266. [[CrossRef](#)] [[PubMed](#)]
 57. Escudie, F.; Auer, L.; Bernard, M.; Mariadassou, M.; Cauquil, L.; Vidal, K.; Maman, S.; Hernandez-Raquet, G.; Combes, S.; Pascal, G. Frogs: Find, rapidly, otus with galaxy solution. *Bioinformatics* **2018**, *34*, 1287–1294. [[CrossRef](#)] [[PubMed](#)]