ANTICANCER RESEARCH *40*: 1325-1334 (2020)
doi:10.21873/anticanres.14074

Gut Microbiota and Host Gene Mutations in Colorectal Cancer Patients and Controls of Iranian and Finnish Origin

VIRINDER SARHADI[1], LEO LAHTI[2], FARIDEH SABERI[3], OMAR YOUSSEF[1], ARTO KOKKOLA[4], TIINA KARLA[5], MILJA TIKKANEN[5], HILPI RAUTELIN[6], PAULI PUOLAKKAINEN[5], RASOUL SALEHI[3,7] and SAKARI KNUUTILA[1]

[1]*Department of Pathology, Faculty of Medicine, University of Helsinki, Helsinki, Finland;*
[2]*Department of Future Technologies, University of Turku, Turku, Finland;*
[3]*Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran;*
[4]*The HUCH Gastrointestinal Clinic, University Central Hospital of Helsinki, Helsinki, Finland;*
[5]*Thermo Fisher Scientific, Vantaa, Finland;*
[6]*Department of Medical Sciences, Clinical Microbiology, Uppsala University, Uppsala, Sweden;*
[7]*Pediatric Inherited Diseases Research Center, Isfahan University of Medical Sciences, Isfahan, Iran*

Abstract. *Background/Aim: Gut microbiota plays an important role in colorectal cancer (CRC) and its composition in CRC patients can be influenced by ethnicity and tumour genomics. Herein, the aim was to study the possible associations of ethnicity and gene mutations with the gut microbiota in CRC patients. Materials and Methods: Bacterial composition in stool samples of 83 CRC patients and 60 controls from Iran and Finland was studied by 16S rRNA gene sequencing. The association of gut microbiota composition with CRC, host mutations in KRAS, NRAS and TP53, and ethnicity analysed. Results: Beta diversity analysis indicated significant differences between the Iranian and Finnish gut microbiota composition, in both controls and patients' groups. The Iranian controls had higher abundance of Prevotella and lower abundance of Bacteroides compared to the Finnish controls, while the Finnish patients had higher abundance of Clostridium compared to Iranian patients. Abundance of Ruminococcus was higher in patients compared to the controls. Higher abundances of Herbaspirillum, Catenibacterium and lower abundances of Barnesiella were associated with mutations in NRAS, TP53, and RAS respectively. Conclusion: A possible link of host gene mutations with gut bacterial composition is suggested.*

Correspondence to: Sakari Knuutila, Department of Pathology, Faculty of Medicine, University of Helsinki, FI-00014, Finland. Tel: +358 504601554, e-mail: sakari.knuutila@helsinki.fi

Key Words: Gut microbiota, gene mutations, colorectal cancer, stool DNA, 16S rRNA gene sequencing.

The gut microbiota plays an integral role in the regulation of different gastrointestinal functions. There is increasing evidence that the gut microbiota not only contributes to the initiation and progression of colorectal cancer (CRC) but also influences the host response to cancer treatment, especially chemo- and immunotherapy (1). In addition to its regional effect in the gastrointestinal tract, the gut microbiota also affects distant organs *via* its role in shaping the immune system and host metabolism (2). Diet, lifestyle and ethnicity are among the main determinants of microbiota composition, and changes in diet associated with a modern lifestyle have been linked to microbiota-mediated cancer risk (3). The host genetic variation is also known to effect bacterial composition (4), although this area has not been investigated extensively. The host genetics is reported to strongly influence the abundance of gut Christensenellaceae (5), the bacteria associated with low body mass index. The *Bifidobacterium* abundance has been linked with a variation in the lactose tolerance gene LCT loci (6, 7). On the other hand, tumour genome is reported to be a factor modulating colonization of certain bacteria around colorectal tumour tissue (8, 9). A high abundance of *Fusobacterium* has been reported to be associated with CpG island methylator phenotype (CIMP), microsatellite instability (MSI) and mutations in *CHD7/8* (10). However, the associations between tumour mutations and gut microbiota composition remain poorly understood.

The association of the gut microbiota composition with different disease states makes them useful as potential diagnostic markers of health and disease (11). Recent large-scale studies have however demonstrated that ethnicity (12) and geography (13) are strong independent factors explaining

Table I. *Characteristics of Iranian and Finnish colorectal cancer (CRC) patients and controls examined in this study.*

	Iranian CRC (n=52)	Finnish CRC (n=31)	Iranian controls (n=47)	Finnish controls (n=13)
Gender (Male/Female)	33/19	20/11	26/21	3/10
Age, years (mean±SD)	62.4±10.6	71.9±8.3	62.8±10.1	44.2±13.6
Stage (Early/Late/Unknown)	52/0/0	21/7/3		
Location (Colon/ Rectum)	41/11	12/19		
Mutation (Yes/No)				
TP53	10/40	3/26		
KRAS	1/49	2/27		
NRAS	6/44	0/29		

CRC, Colorectal cancer.

inter-individual differences in microbiota composition, even more so than either diet or metabolic disease. It is therefore important that associations of gut microbiota with serious diseases, like cancer, should be tested in different populations. We have earlier described an association of gut microbiota with different types of gastrointestinal neoplasms in Finnish patients (14). In the present study, the gut microbiota composition of Iranian CRC patients and controls were studied and compared to that seen among Finnish CRC patients and controls to find bacterial profile related to ethnicity and that related to CRC. We also compared bacterial composition in CRC patients with and without host gene mutations in *KRAS, NRAS, HRAS* and *TP53* to examine the possible associations between gut microbiota and cancer gene mutations.

Materials and Methods

Patients and sample collection. Informed consent was obtained from all Iranian and Finnish patients and controls before collection of stool samples. The ethical permission (351/13/03/02/2014) for the study was obtained from the Hospital District of Helsinki and Uusimaa (HUS) review board.

All of the Iranian samples were collected in Isfahan University-Alzahra Hospital (Isfahan, Iran) from CRC patients (IrCRC) and controls (IrC) living in the Isfahan province located in the central part of Iran. The stool samples were collected from 52 patients diagnosed with CRC before starting any cancer treatment and 47 control samples from individuals who had undergone colonoscopy due to some gastrointestinal symptoms but had normal findings on colonoscopy (Table I). Among the CRC patients, 11 had rectal cancer and 41 had colon cancer. Both the patients and controls were mainly Muslims, following Iranian dietary habits and did not consume either pork or alcohol. Sixteen percent of the patients and 9% of the controls were smokers. All samples were collected between July, 2016 and December 2017. Specimens were collected in the hospital and immediately stored at -80˚C until DNA extraction.

Stool samples were also collected from 31 Finnish CRC patients (HefeCRC) between April 2015 to May 2017, in Helsinki Uusimaa Hospital District: Surgical, Meilahti, and Jorvi Hospitals in Finland. Details of CRC patients and control group (HefeC) of Finnish origin have been described previously (14). The Finnish patients and controls followed mainly non-vegetarian diet, and on an average week consumed 10-20 standard units of alcohol. Nearly thirty percent of the Finnish patients and controls were smokers.

DNA extraction. From Iranian patients and controls, DNA was extracted from stool samples (180-220 mg) using the QIAamp DNA stool mini kit (Qiagen GmbH, Hidden, Germany) according to manufacturer's guidelines. DNA was quantified by Qubit 2.0 Fluorimeter (Thermo Fisher Scientific, Waltham, MA, USA) using the Qubit dsDNA BR Assay Kit. Extracted DNA was stored at –20˚C. DNA extraction from Finnish stool samples was performed using PSP Spin Stool DNA Plus Kit (Stratec Molecular, Berlin, Germany), as previously described (14).

The results of 16S rRNA gene sequencing from 31 Finnish CRC patients and 13 healthy Finnish individuals published previously (14) were used to compare results with those of Iranian cohort. DNA samples of both Iranian and Finnish CRC patients and controls were checked for quality and quantity by Qubit 2.0 Fluorimeter and all the analysis were done in the same laboratory, by the same researchers in Helsinki, and following the same methodological procedures, as previously described (14).

Gene mutation analysis of host DNA from stool samples. Data regarding hotspot mutations in *KRAS, NRAS, HRAS*, and *TP53* genes in DNA from stool samples of Finnish patients was collected from previously published results (15). Gene mutation information was available for 79/83 CRC patients; *TP53* mutations were seen in 13, *NRAS* in 6, *KRAS* in 3 and no mutations in *HRAS* patients (all *RAS* mutations were found in 9 patients). DNA isolated from stool samples of both Finnish and Iranian CRC patients were studied for hotspot mutations in 22 cancer-related genes by targeted next generation sequencing using Ion AmpliSeq Colon and Lung Cancer panel v2 (Thermo Fisher Scientific, Waltham, MA, USA). DNA samples were sequenced on Ion Personal Genome Machine System (Thermo Fisher Scientific, Waltham, MA, USA) as previously described (15).

16S rRNA gene sequencing. Library preparation, template preparation and sequencing were performed on DNA isolated from Iranian CRC and controls as described earlier for Finnish CRC and control samples (14). For Finnish patients and controls, the results

of 16S rRNA gene sequencing published earlier (14) were used for comparison with Iranian samples and for examining the association with gene mutations in DNA from stool samples.

Briefly, sequencing libraries were prepared from 3 ng of DNA, using Ion 16S Metagenomics kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the vendor's instructions. Six hypervariable regions (Primer set V2, V4, V8 and Primer set V3, V6-7, V9) of 16S rRNA gene were amplified in two reactions/sample. After PCR, the samples were end-repaired, purified with Agencourt® AMPure® XP beads (Beckman Coulter, Brea, California, USA) and ligated to barcoded sequencing adapters according to the kit protocol. The libraries were quantified by the TapeStation (Agilent Technologies, Santa Clara, CA, USA) and samples were diluted to a 10 pM concentration.

The libraries were pooled and the template preparation was performed with either Ion OneTouch 2 system using the Ion PGM™ Hi-Q™ OT2 Kit (Thermo Fisher Scientific, Waltham, MA, USA) or Ion Chef system using the Ion PGM™ Hi-Q™ Chef Kit (Thermo Fisher Scientific, Waltham, MA, USA) following the kit protocols. The OneTouch 2 or the Ion Chef was used for emulsion PCR and the quality of resulting Ion Spheres were checked with Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). Sequencing was performed on the Ion PGM system using the Ion 318™ Chip (Thermo Fisher Scientific, Waltham, MA, USA) and Ion PGM Hi-Q Sequencing kit (Thermo Fisher Scientific, Waltham, MA, USA).

Data analysis. Operational taxonomic unit (OTU) tables were created for the 16S rRNA gene sequencing data of a total of 143 stool samples. The samples were classified according to nationality (Iranian/Finnish), cancer status (cancer/control), tumour location (colon/rectum), tumour stage (early/late), and gene mutation status of *RAS* (*KRAS, NRAS, HRAS*) and *TP53*. The between-sample normalization was performed by rarefaction of the read counts to an even depth and the rarefied read counts were converted into relative abundances using the *phyloseq* R package (16). A total of 163 unique genera were detected.

Alpha diversity and observed richness of the gut microbiota at the genus level was analysed with the *microbiome* (17) and *vegan* R packages (18). The Shannon index was used to estimate the bacterial diversity, with community richness being quantified by the number of unique taxa observed. Kruskal-Wallis test was applied to test for the significance of group-level differences and correction for multiple testing was performed for each group with the Benjamini-Hochberg FDR method (19).

Unsupervised principal coordinates analysis (PCoA), based on Euclidean distance between Hellinger-transformed abundance profiles (20) were performed with the *phyloseq* R package (16). Only the genera that were detected in at least 20% of all samples were included in the analysis. Community level differences between the groups were tested with PERMANOVA for CLR-transformed (centered log-ratio transformed) abundances (21) with the R package compositions (22) to remove compositionality bias. ANCOM was applied to assess the significance of the differences in the abundance of the individual genera (23).

Association of age and sex with taxa abundances was examined by Spearman correlation and Wilcoxon test respectively, both performed for CLR-transformed data in order to remove compositionality bias. Since no significant association of age or sex with taxa abundances was observed, we did not include these as covariates in our analysis.

Results

The characteristics of the CRC patients and controls groups included in the study are described in Table I.

Bacterial alpha-diversity. The most significant difference in gut microbiota alpha-diversity (Shannon index) was observed between the Iranian and the Finnish controls (p=0.004). Finnish controls had a higher bacterial richness than their Iranian counterparts (Figure 1A). However, no difference in alpha-diversity was observed between the Iranian and the Finnish CRC patients (p=1.0) (Figure1B); between the Iranian CRC patients and controls (p=0.51); and between Finnish CRC patients and controls (p=0.10). All p-values have been corrected for multiple testing as described in the Materials and Methods.

Bacterial beta-diversity. The beta diversity of gut microbiota composition (Figure 2; Table II) indicated that the Iranian controls were significantly different from the Finnish controls (p=0.006), and the Iranian CRC patients were significantly different from the Finnish CRC patients (p=0.006), indicating a strong ethnic influence on the gut microbiota composition. The CRC patients and controls were not significantly different for either nationality (p>0.1).

Bacterial genera with significant differences in relative abundances in different group comparisons. Significant differences were observed between the Iranian and Finnish controls also in the relative abundance of many individual genera (Table III). The most notable difference was the high relative abundance of *Prevotella* in the Iranian controls compared to the Finnish controls (Figure 3A). Other significant genera [*Bacteroides, Roseburia,* (*Ruminococcus*), *Eubacterium, Faecalibacterium* and *Lactococcus*] had higher relative abundances in Finnish control samples compared to the Iranian controls. On the other hand, a comparison of Iranian and Finnish cancer patients showed significant differences in the relative abundance of only *Clostridium*; that had a higher abundance in the Finnish CRC patients as compared to Iranian patients (Figure 3B).

The CRC patients had a higher relative abundance of *Ruminococcus* compared to healthy individuals, when samples were not classified according to ethnicity. Higher relative abundance of *Ruminococcus* and *Bifidobacterium* was also seen in Iranian CRC patients compared to Iranian controls. On the other hand, lower relative abundance of *Bifidobacterium*, *Phascolarctobacterium* and *Lachnoclostridium* was seen in Finnish CRC patients compared to the Finnish controls (Table III).

Association of host gene mutations and gut microbiota. CRC patients with *RAS (KRAS/NRAS)* gene mutations in the DNA

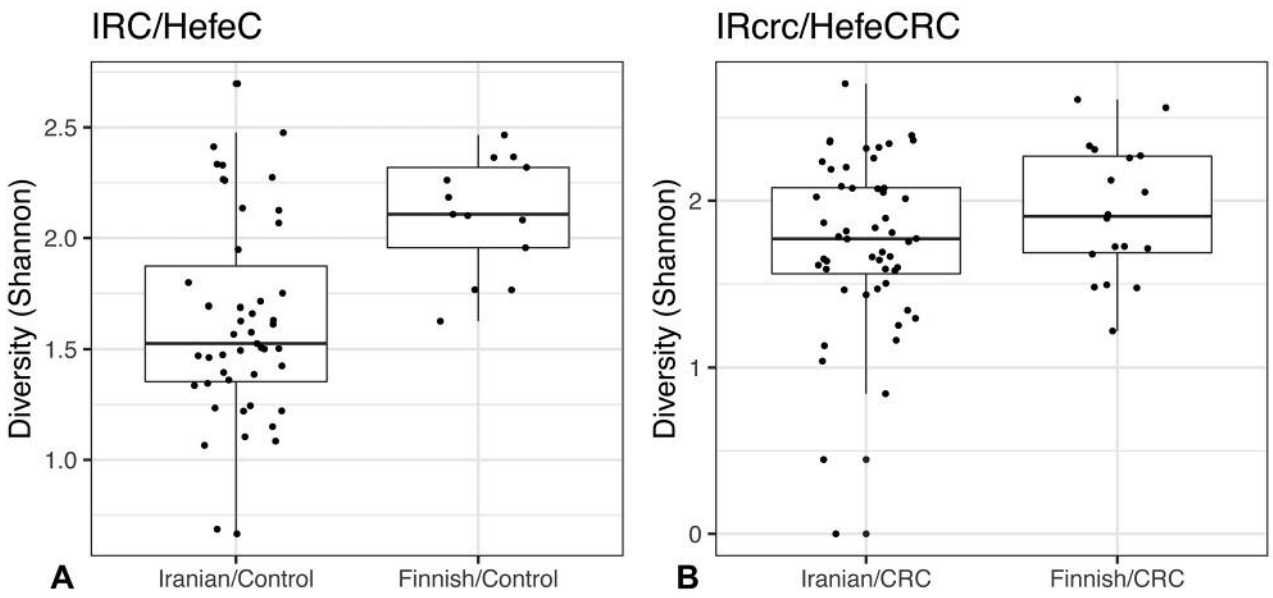

Figure 1. *Comparison of alpha diversity between the groups. (A) Iranian (IrC) and Finnish (HefeC) controls, and (B) Iranian (IrCRC) and Finnish (HefeCRC) colorectal cancer (CRC) patients.*

isolated from stool samples had lower abundance of *Barnesiella* in comparison to those patients without *RAS* gene mutation (Figure 3C; Table III). The relative abundance of *Herbaspirillum* was higher in *NRAS* mutated (Figure 3D); of *Catenibacterium* higher (Figure 3) and *Blautia* lower in *TP53* mutated patients compared to the patients with no mutation in the corresponding genes (Table III).

Fusobacterium among CRC and controls. We observed presence of *Fusobacterium* in 18.6% of the samples from CRC patients, which was significantly (p=0.01; Fisher's exact test) higher compared to 3.3% among controls.

Discussion

Composition of gut bacteria is affected by both environmental and genetic factors. Diet, lifestyle, ethnicity, age, sex and host genotype are among the main determinants of microbiota composition and the changes in gut bacterial composition are linked to microbiota-mediated cancer risk (24-27). We aimed to study gut microbiota and host gene mutations in stool samples from CRC patients and controls from Iran and Finland, in order to see the associations of ethnicity, CRC and cancer gene mutations on the gut microbiota.

Primarily, a strong influence of ethnicity on the gut bacterial composition was noted in beta-diversity analysis. Significant differences were seen between the Iranian and Finnish controls (p=0.006) and between the Iranian and Finnish CRC patients (p=0.006), while the differences were not significant between patients and controls from the same ethnic group (Table II; Figure 2). Thus, ethnicity seems to have a more significant effect on the overall community composition than the CRC status. The differences in alpha diversity were however significant only between healthy controls of Iranian and Finnish origin and non-significant between CRC patients of these ethnic groups (Figure 1). Finnish controls had higher bacterial diversity and richness (Figure 1A) than Iranian controls. Similar higher gut microbiota alpha-diversity among individuals of north European origin (Dutch) compared to individuals of Middle-Eastern and African ethnic groups living in the same city (Amsterdam) has been reported earlier (11). A higher bacterial alpha diversity has also been associated with a healthier metabolic state (28).

In the present study, the most striking difference observed between the controls was the high abundance of *Prevotella* in Iranian controls, whereas the levels of these genera were very low in Finnish controls (Figure 3A, Table III). A marked lower abundance of *Prevotella* has been similarly reported in Dutch individuals as compared to the high abundance found in Moroccans, Turks, and Ghanaians (12). It therefore seems that individuals of north European descent have a low *Provetella* abundance, while individuals from Middle-Eastern and other neighbouring countries harbour high amounts of these bacteria. Previous studies have described presence of three enterotypes driven by *Provetella*, *Bacteroides* and *Ruminococcus* (29). The higher abundance of *Bacteroides* among Finnish controls and *Provetella* among

Table II. *Group-level comparisons based on community composition. Analysed for beta diversity differences among groups by PERMANOVA with 999 permutations and Benjamini-Hochberg FDR correction for multiple testing. The associations with FDR<0.25 (adjusted p-value) are shown.*

Groups compared	R^2	Adj. *p*-Value
IrC and HefeC	0.151	0.006
IrCRC and HefeCRC	0.062	0.006
IrC and IrCRC	0.025	0.146
AllC and AllCRC	0.017	0.158
HefeC and HefeCRC	0.056	0.146
Location (colon and rectum)	0.021	0.190
Stage (early and late)	0.022	0.191

IrC, Iranian controls; HefeC, Finnish controls; IrCRC, Iranian colorectal cancer patients; HefeCRC, Finnish colorectal cancer patients; AllC, all controls; AllCRC, all colorectal cancer patients; FDR, false discovery rate.

Table III. *Genera which were significantly differentially abundant in the group-wise comparisons (False discovery rate, FDR<0.25). The Log10 fold change estimates are based on centered log-ratio (CLR)-transformed abundances between the group means.*

Genus	Log_{10} fold-change
Finnish controls/Iranian controls	
Prevotella	–3.10
Bacteroides	0.75
Roseburia	0.71
[Ruminococcus]	0.68
Eubacterium	0.55
Faecalibacterium	0.40
Lactococcus	0.33
Finnish CRC patients/Iranian CRC patients	
Clostridium	0.86
Iranian CRC patients/Iranian controls	
Bifidobacterium	0.86
Ruminococcus	0.77
Finnish CRC patients/Finnish controls	
Phascolarctobacterium	–1.41
Bifidobacterium	–1.10
Lachnoclostridium	–0.89
All CRC patients/All controls	
Ruminococcus	0.68
CRC patients with *RAS* mutation/no *RAS* mutation	
Barnesiella	–1.30
CRC patients with *NRAS* mutation/no *NRAS* mutation	
Herbaspirillum	0.98
CRC patients with *TP53* mutation/no *TP53* mutation	
Catenibacterium	0.99
Blautia	–0.66

CRC, Colorectal cancer; *RAS*, RAS type GTPase family (gene group including *HRAS, NRAS* and *KRAS*); *NRAS*, NRAS proto-oncogene, GTPase; *TP53*, tumor protein p53.

Iranian controls could be indicative of higher prevalence of *Bacteroides* –driven enterotype among Finnish controls and *Provetella* –driven enterotype among Iranian controls. Moreover, the enterotype driven by *Provetella* is associated with non-western or fibre-rich diet and enterotype driven by *Bacteroides* is associated with animal protein and saturated fat diet (30).

Although many genera had differential abundance among the controls of Iranian and Finnish origin, only *Clostridium* showed significant difference in abundance among the CRC patients of these two populations, with lower abundance among Iranian CRC patients. The differential abundances of the genera between Iranian and Finnish groups might have been influenced by diet. The major differences in the dietary habits between the two ethnic groups are the absence of pork meat and alcohol intake among Iranian patients and controls. Higher abundance of *Clostridium* (seen in Finnish CRC patients) and higher abundance of *Bacteroides* (seen in Finnish controls) has been previously associated with alcohol consumption (31).

Gut bacterial dysbiosis is commonly seen in patients with CRC. Although in literature there are discrepancies regarding the bacterial profiles associated with CRC, higher abundance of *Fusobacterium* is one of the most frequent observations associated with CRC status (32, 33). In line with previous studies, we observed significant differences (p=0.01) in the presence of *Fusobacterium* among CRC patients and controls. *Fusobacterium* has been found to be strongly associated with CRC (32) and involved in metastasis of CRC (33).

In order to identify which bacteria were associated with CRC, we compared the bacterial composition of all CRC patients with all controls, irrespective of ethnicity. Only, *Ruminococcus*, had a significantly higher abundance in cancer patients compared to the controls, as well as in Iranian CRC compared to Iranian controls. According to the literature, *Ruminococcus* is able to break down resistant starches in the gut and supply nutrients to the host and also provides growth substrates for other bacteria (35). Differential abundance of *Ruminococcus* has been previously reported in CRC (34) and a recent study revealed higher abundance of *Ruminococcus* in patients at stage-1 of CRC compared to healthy individuals (36). Similarly, in our study, all Iranian CRC patients and the majority of the Finnish patients had early stage of cancer (Table I). Therefore, the increased abundance of *Ruminococcus* might be associated with early stages of CRC.

The Finnish CRC had lower abundances of *Bifidobacterium*, *Phascolarctobacterium* and *Lachnoclostridium* than their healthy counterparts. Lower levels of both *Phascolarctobacterium* and Lachnospiraceae NK4A136 have been related to psychosocial stress in children (37). *Bifidobacterium* and *Phascolarctobacterium* confer beneficial health properties, and moreover, the presence of *Phascolarctobacterium* has been

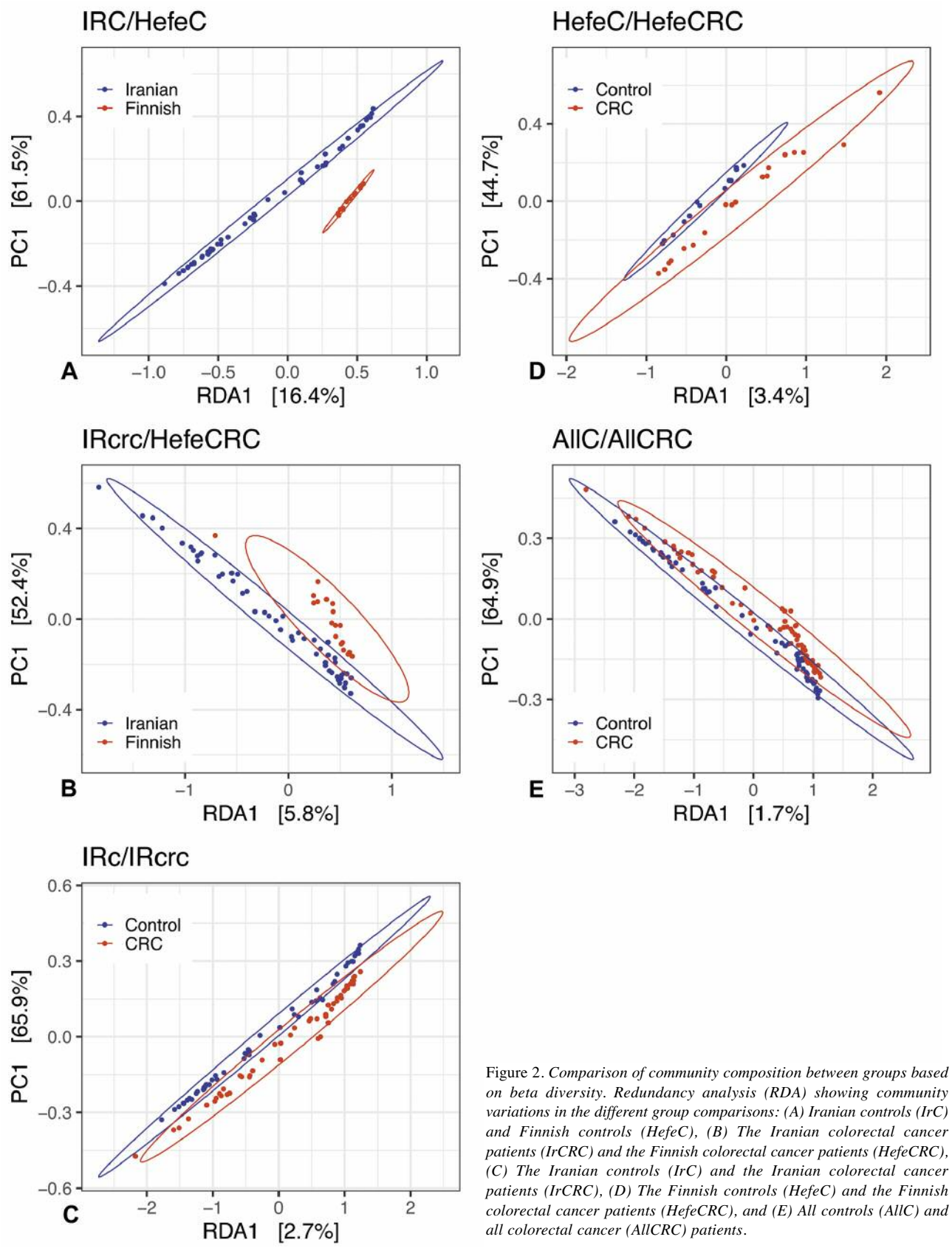

Figure 2. *Comparison of community composition between groups based on beta diversity. Redundancy analysis (RDA) showing community variations in the different group comparisons: (A) Iranian controls (IrC) and Finnish controls (HefeC), (B) The Iranian colorectal cancer patients (IrCRC) and the Finnish colorectal cancer patients (HefeCRC), (C) The Iranian controls (IrC) and the Iranian colorectal cancer patients (IrCRC), (D) The Finnish controls (HefeC) and the Finnish colorectal cancer patients (HefeCRC), and (E) All controls (AllC) and all colorectal cancer (AllCRC) patients.*

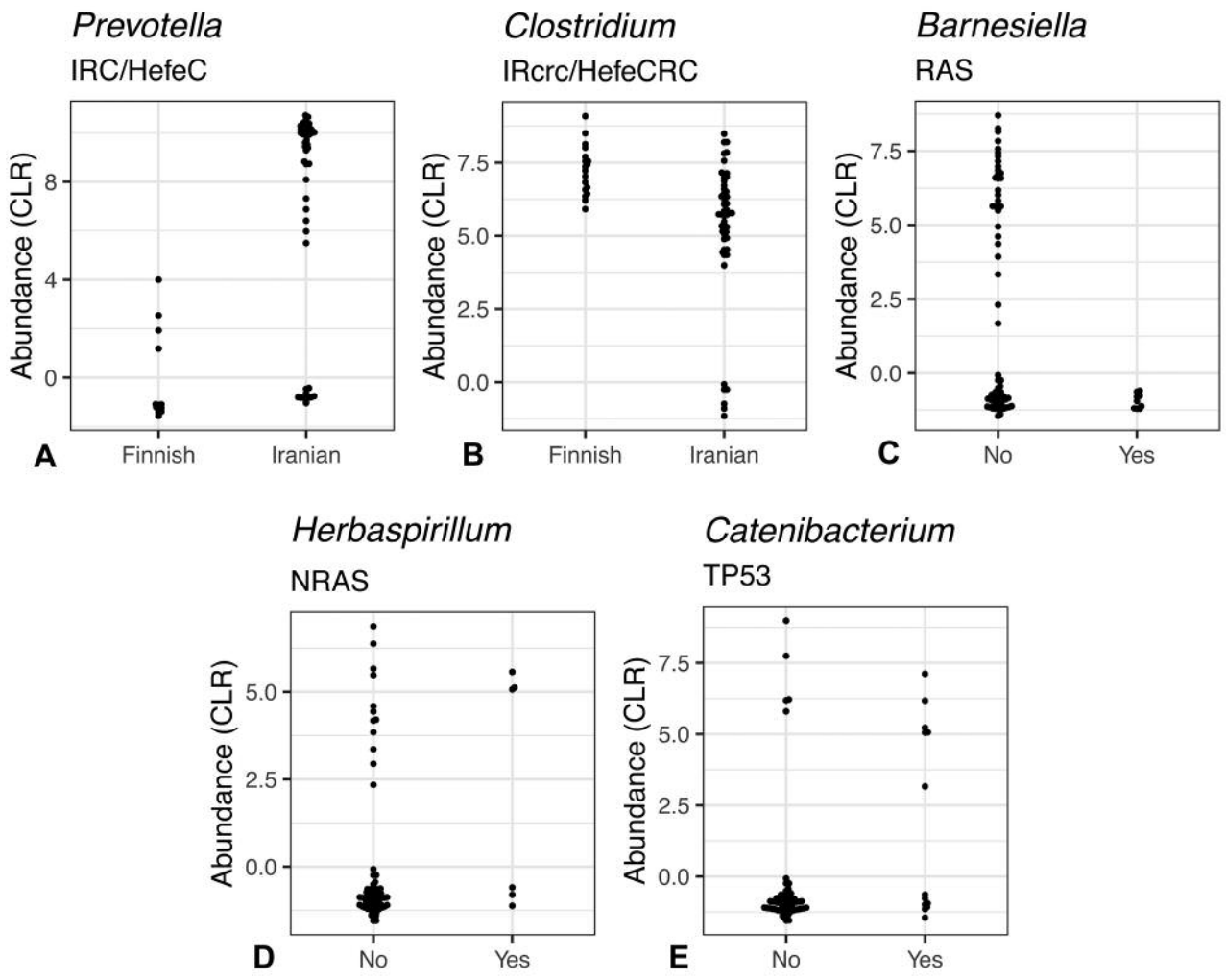

Figure 3. *Bacterial genera showing greatest differences in abundances (centered log-transformed; CLR) for various group comparisons. (A) Iranian controls (IrC) and Finnish controls (HefeC), (B) Iranian colorectal cancer patients (IrCRC) and Finnish colorectal cancer patients (HefeCRC), (C) colorectal cancer patients with and without the RAS mutations, (D) colorectal cancer patients with and without the NRAS mutations, and (E) colorectal cancer patients with and without the TP53 mutations.*

associated with a positive mood in the host (38). The higher abundances of these bacteria in Finnish controls might be linked to a healthier state.

The most significant association of host gene mutations with gut bacterial composition was a lower abundance of *Barnesiella* in patients with *RAS* mutations and higher abundance of *Herbaspirillum* in *NRAS* mutated patients compared to patients without these gene mutations. *TP53* mutated patients showed higher abundance of *Catenibacterium* and lower abundance of *Blautia* compared to patients without mutations in *TP53* gene. *Barnesiella* has been reported to protect against the colonization of the highly antibiotic – resistant *Enterococcus faecium* bacterium by restricting the growth of these pathogens (39). Importantly, *Barnesiella* has been associated with an anti-cancer immune response; these bacteria have been found to modulate the efficacy of the anti-cancer immunomodulatory agent, cyclophosphamide (40), by enhancing the Tc1 (cytotoxic T lymphocytes) and Th1 (T helper cell) response and restoring the function of intra-tumoral interferon gamma (IFN-γ)-producing T cells. Gene expression analysis from The Cancer Genome Atlas and the KFSYSCC (Koo Foundation Sun Yat-Sen Cancer Centre) data sets have revealed a suppressed Th1-/cytotoxicity in *KRAS* mutant colorectal cancer patients (41). The absence of *Barnesiella* in *RAS* mutated patients in our study, might also be associated with compromised tumour immune surveillance in these patients which would promote the development of the tumour.

Higher urinary abundance of *Herbaspirillum* has been linked to high risk of recurrence of bladder cancer (42). The abundance of *Catenibacterium*, seen in *TP53* mutated

patients, has been associated with animal fat diet (43) and diet rich in red meat, while food containing trans-fatty acids has been reported to be associated with *TP53* mutations in colon cancer (44).

Since age and gender are important factors affecting gut microbiota composition, a limitation of our study is that the Finnish controls were of relatively younger age than the other groups (Table I). However, no significant association of age and sex with the observed abundances was found in our data. Moreover, the small number of patients with *RAS* and *TP53* mutations in our study did not allow drawing any firm conclusions. A more in-depth study on larger number of patients with *RAS* and *TP53* mutations is warranted to confirm the associations of bacterial genera with *RAS* and *TP53* mutations found in this study.

We found that ethnicity is more significantly associated with gut microbiota composition than is the cancer status. Therefore, our results suggest that bacterial markers associated with cancer status should be further tested in patients from different ethnic groups, in order to assess their usefulness as cancer marker. Further studies on the association of bacterial genera with *RAS* and *TP53* gene mutations could elucidate the role of gut microbiota in the growth of mutated cells and moreover, the impact of cellular pathways that are dysregulated in mutated cells on the abundance of specific bacteria. This knowledge could be useful in identifying new therapeutic approaches.

Conflicts of Interest

The Authors declare no conflicts of interest.

Authors' Contributions

SK, VS, LL, PP and AK contributed to the study design. VS, SK, FS, OY, AK, TK, MT and RS contributed to sample collection and analysis. Bioinformatics and statistical data analysis was coordinated by LL. VS, LL, HR and SK contributed to manuscript preparation. All Authors approved the final version.

Acknowledgements

The Authors wish to thank Dr Ewen MacDonald for his linguistic correction of the manuscript. The study was supported by Sigrid Jusélius Foundation grant to SK. LL was supported by Academy of Finland (decision 295741). Rasoul Salehi greatly acknowledges financial support provided to him by the Iran National Scientific Foundation (INSF) through grant number 90002549.

References

1 Murphy CL, O'Toole PW and Shanahan F: The gut microbiota in causation, detection, and treatment of cancer. Am J Gastroenterol 114: 1036-1042, 2019. PMID: 30848738. DOI: 10.14309/ajg.0000000000000075

2 Kho ZY and Lal SK: The human gut microbiome - A potential controller of wellness and disease. Front Microbiol *9*: 1835, 2018. PMID: 30154767. DOI: 10.3389/fmicb.2018.01835

3 Vipperla K and O'Keefe SJ: Diet, microbiota, and dysbiosis: a 'recipe' for colorectal cancer. Food Funct *7*: 1731-1740, 2016. PMID: 26840037. DOI: 10.1039/c5fo01276g

4 Burns MB and Blekhman R: Integrating tumour genomics into studies of the microbiome in colorectal cancer. Gut Microbes *10*: 547-552, 2019. PMID: 30556775. DOI: 10.1080/19490976.2018.1549421

5 Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, Beaumont M, Van Treuren W, Knight R, Bell JT, Spector TD, Clark AG and Ley RE: Human genetics shape the gut microbiome. Cell *159*: 789-799, 2014. PMID: 25417156. DOI: 10.1016/j.cell.2014.09.053

6 Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowski R, Bell JT, Spector TD, Keinan A, Ley RE, Gevers D and Clark AG: Host genetic variation impacts microbiome composition across human body sites. Genome Biol *16*: 191, 2015. PMID: 26374288. DOI: 10.1186/s13059-015-0759-1

7 Goodrich JK, Davenport ER, Waters JL, Clark AG, Ley RE: Cross-species comparisons of host genetic associations with the microbiome. Science *352*: 532-535, 2016. PMID: 27126034. DOI: 10.1126/science.aad9379

8 Burns MB, Montassier E, Abrahante J, Priya S, Niccum DE, Khoruts A, Starr TK, Knights D and Blekhman R: Colorectal cancer mutational profiles correlate with defined microbial communities in the tumour microenvironment. PLoS Genet *14*: e1007376, 2018. PMID: 29924794. DOI: 10.1371/journal.pgen.1007376

9 Burns MB, Blekhman R: Integrating tumour genomics into studies of the microbiome in colorectal cancer. Gut Microbes *10*: 547-552, 2018. PMID: 30556775. DOI: 10.1080/19490976.2018.1549421

10 Tahara T, Yamamoto E, Suzuki H, Maruyama R, Chung W, Garriga J, Jelinek J, Yamano HO, Sugai T, An B, Shureiqi I, Toyota M, Kondo Y, Estécio MR and Issa JP: *Fusobacterium* in colonic flora and molecular features of colorectal carcinoma. Cancer Res *74*: 1311-1318, 2014. PMID: 24385213. DOI: 10.1158/0008-5472.CAN-13-1865

11 Lavelle A and Hill C: Gut microbiome in health and disease: Emerging diagnostic opportunities. Gastroenterol Clin North Am *48*: 221-235, 2019. PMID: 31046972. DOI: 10.1016/j.gtc.2019.02.003

12 Deschasaux M, Bouter KE, Prodan A, Levin E, Groen AK, Herrema H, Tremaroli V, Bakker GJ, Attaye I, Pinto-Sietsma SJ, van Raalte DH, Snijder MB, Nicolaou M, Peters R, Zwinderman AH, Bäckhed F and Nieuwdorp M: Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. Nat Med *24*: 1526-1531, 2018. PMID: 30150717. DOI: 10.1038/s41591-018-0160-1

13. He Y, Wu W, Zheng HM, Li P, McDonald D, Sheng HF, Chen MX, Chen ZH, Ji GY, Zheng ZD, Mujagond P, Chen XJ, Rong ZH, Chen P, Lyu LY, Wang X, Wu CB, Yu N, Xu YJ, Yin J, Raes J, Knight R, Ma WJ and Zhou HW: Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nat Med *24*: 1532-1535, 2018. PMID: 30150716. DOI: 10.1038/s41591-018-0164-x

14 Youssef O, Lahti L, Kokkola A, Karla T, Tikkanen M, Ehsan H, Carpelan-Holmström M, Koskensalo S, Böhling T, Rautelin H, Puolakkainen P, Knuutila S and Sarhadi V: Stool microbiota

composition differs in patients with stomach, colon, and rectal neoplasms. Dig Dis Sci *63*: 2950-2958, 2018. PMID: 29995183. DOI: 10.1007/s10620-018-5190-5

15 Youssef O, Sarhadi V, Ehsan H, Böhling T, Carpelan-Holmström M, Koskensalo S, Puolakkainen P, Kokkola A and Knuutila S: Gene mutations in stool from gastric and colorectal neoplasia patients by next-generation sequencing. World J Gastroenterol *23*: 8291-8299, 2017. PMID: 29307989. DOI: 10.3748/wjg.v23.i47.8291

16 McMurdie PJ and Holmes S: Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One *8*: e61217, 2013. PMID: 23630581. DOI: 10.1371/journal.pone.0061217

17 Lahti L and Shetty S: Tools for microbiome analysis in R. Microbiome package version 0.99.88; 2012-2017. Available at: https://www.bioconductor.org/packages/devel/bioc/html/microbiome.html

18 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E and Wagner H: vegan: Community ecology package. 2. R package version 2.4-4, 2017. Available at: https://cran.r-project.org/web/packages/vegan/index.html

19 Benjamini Y and Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc B *57(1)*: 289-300, 1995. DOI: 10.2307/2346101

20 Legendre P and Gallagher ED: Ecologically meaningful transformations for ordination of species data. Oecologia *129*: 271-280, 2001. PMID: 28547606. DOI: 10.1007/s004420100716

21 J Aitchison: The statistical analysis of compositional data. Monographs on statistics and applied probability. Chapman and Hall, London, 1986.

22 Compositions: compositional data analysis. 2019. Available at: https://CRAN.R-project.org/package=compositions.

23 Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R and Peddada SD: Analysis of composition of microbiomes: a novel method for studying microbial composition. Microb Ecol Health Dis *26*: 27663, 2015. PMID: 26028277. DOI: 10.3402/mehd.v26.27663

24 Vipperla K and O'Keefe SJ: Diet, microbiota, and dysbiosis: a 'recipe' for colorectal cancer. Food Funct *7*: 1731-1740, 2016. PMID: 26840037. DOI: 10.1039/C5FO01276G

25 Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, Beaumont M, Van Treuren W, Knight R, Bell JT, Spector TD, Clark AG and Ley RE: Human genetics shape the gut microbiome. Cell *159*: 789-799, 2014. PMID: 25417156. DOI: 10.1016/j.cell.2014.09.053

26 Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowski R, Bell JT, Spector TD, Keinan A, Ley RE, Gevers D and Clark AG: Host genetic variation impacts microbiome composition across human body sites. Genome Biol *16*: 191, 2015. PMID: 26374288. DOI: 10.1186/s13059-015-0759-1

27 Goodrich JK, Davenport ER, Waters JL, Clark AG and Ley RE: Cross-species comparisons of host genetic associations with the microbiome. Science *352*: 532-535, 2016. PMID: 27126034. DOI: 10.1126/science.aad9379

28 Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, Almeida M, Arumugam M, Batto JM, Kennedy S, Leonard P, Li J, Burgdorf K, Grarup N, Jørgensen T, Brandslund I, Nielsen HB, Agnieszka SJ, Bertalan M, Levenez F, Pons N, Rasmussen S, Sunagawa S, Tap J, Tims S, Zoetendal EG, Brunak S, Clément K, Doré J, Kleerebezem M, Kristiansen K, Renault P, Sicheritz-Ponten T, de Vos WM, Zucker JD, Raes J, Hansen T, MetaHIT consortium, Bork P, Wang J, Ehrlich SD and Pedersen O: Richness of human gut microbiome correlates with metabolic markers. Nature *500*: 541-546, 2013. PMID: 23985870. DOI: 10.1038/nature12506

29 Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Doré J; MetaHIT Consortium, Antolín M, Artiguenave F, Blottiere HM, Almeida M, Brechot C, Cara C, Chervaux C, Cultrone A, Delorme C, Denariaz G, Dervyn R, Foerstner KU, Friss C, van de Guchte M, Guedon E, Haimet F, Huber W, van Hylckama-Vlieg J, Jamet A, Juste C, Kaci G, Knol J, Lakhdari O, Layec S, Le Roux K, Maguin E, Mérieux A, Melo Minardi R, M'rini C, Muller J, Oozeer R, Parkhill J, Renault P, Rescigno M, Sanchez N, Sunagawa S, Torrejon A, Turner K, Vandemeulebrouck G, Varela E, Winogradsky Y, Zeller G, Weissenbach J, Ehrlich SD and Bork P: Enterotypes of the human gut microbiome. Nature *473*: 174-180, 2011. PMID: 21508958. DOI 10.1038/nature09944

30 Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD, de Vos WM, Ehrlich SD, Fraser CM, Hattori M, Huttenhower C, Jeffery IB, Knights D, Lewis JD, Ley RE, Ochman H, O'Toole PW, Quince C, Relman DA, Shanahan F, Sunagawa S, Wang J, Weinstock GM, Wu GD, Zeller G, Zhao L, Raes J, Knight R and Bork P: Enterotypes in the landscape of gut microbial community composition. Nat Microbiol *3*: 8-16, 2018. PMID: 29255284. DOI: 10.1038/s41564-017-0072-8

31 Queipo-Ortuño MI, Boto-Ordóñez M, Murri M, Gomez-Zumaquero JM, Clemente-Postigo M, Estruch R, Cardona Diaz F, Andrés-Lacueva C and Tinahones FJ: Influence of red wine polyphenols and ethanol on the gut microbiota ecology and biochemical biomarkers. Am J Clin Nutr *95*: 1323-1334, 2012. PMID: 22552027. DOI: 10.3945/ajcn.111.027847

32 Amitay EL, Werner S, Vital M, Pieper DH, Höfler D, Gierse IJ, Butt J, Balavarca Y, Cuk K and Brenner H: *Fusobacterium* and colorectal cancer: causal factor or passenger? Results from a large colorectal cancer screening study. Carcinogenesis *38*: 781-788, 2017. PMID: 28582482. DOI: 10.1093/carcin/bgx053

33 Chen S, Su T, Zhang Y, Lee A, He J, Ge Q, Wang L, Si J, Zhuo W and Wang L: *Fusobacterium nucleatum* promotes colorectal cancer metastasis by modulating *KRT7-AS*/KRT7. Gut Microbes: 1-15, 2020. PMID: 31910722. DOI: 10.1080/19490976.2019.1695494

34 Yang J, McDowell A, Kim EK, Seo H, Lee WH, Moon CM, Kym SM, Lee DH, Park YS, Jee YK and Kim YK: Development of a colorectal cancer diagnostic model and dietary risk assessment through gut microbiome analysis. Exp Mol Med *51:* 117, 2019. PMID: 31582724. DOI: 10.1038/s12276-019-0313-4

35 La Reau AJ and Suen G: The *Ruminococci*: key symbionts of the gut ecosystem. J Microbiol *56*: 199-208, 2018. PMID: 29492877. DOI: 10.1007/s12275-018-8024-4

36 Sheng Q, Du H, Cheng X, Cheng X, Tang Y, Pan L, Wang Q and Lin J: Characteristics of fecal gut microbiota in patients with colorectal cancer at different stages and different sites. Oncol Lett *18*: 4834-4844, 2019. PMID: 31611994. DOI: 10.3892/ol.2019.10841

37 Michels N, Van de Wiele T, Fouhy F, O'Mahony S, Clarke G and Keane J: Gut microbiome patterns depending on children's psychosocial stress: Reports versus biomarkers. Brain Behav Immun *80*: 751-762, 2019. PMID: 31112792. DOI 10.1016/j.bbi.2019.05.024

38 Li L, Su Q, Xie B, Duan L, Zhao W, Hu D, Wu R and Liu H: Gut microbes in correlation with mood: case study in a closed experimental human life support system. Neurogastroenterol Motil *28*: 1233-1240, 2016. PMID: 27027909. DOI: 10.1111/nmo.12822

39 Ubeda C, Bucci V, Caballero S, Djukovic A, Toussaint NC, Equinda M, Lipuma L, Ling L, Gobourne A, No D, Taur Y, Jenq RR, van den Brink MR, Xavier JB and Pamer EG: Intestinal microbiota containing *Barnesiella* species cures vancomycin-resistant *Enterococcus faecium* colonization. Infect Immun *81*: 965-973, 2013. PMID: 23319552. DOI: 10.1128/IAI.01197-12

40 Daillère R, Vétizou M, Waldschmitt N, Yamazaki T, Isnard C, Poirier-Colame V, Duong CPM, Flament C, Lepage P, Roberti MP, Routy B, Jacquelot N, Apetoh L, Becharef S, Rusakiewicz S, Langella P, Sokol H, Kroemer G, Enot D, Roux A, Eggermont A, Tartour E, Johannes L, Woerther PL, Chachaty E, Soria JC, Golden E, Formenti S, Plebanski M, Madondo M, Rosenstiel P, Raoult D, Cattoir V, Boneca IG, Chamaillard M and Zitvogel L: *Enterococcus hirae* and *Barnesiella intestinihominis* facilitate cyclophosphamide-induced therapeutic immunomodulatory effects. Immunity *45*: 931-943, 2016. PMID: 27717798. DOI: 10.1016/j.immuni.2016.09.009

41 Lal N, White BS, Goussous G, Pickles O, Mason MJ, Beggs AD, Taniere P, Willcox BE, Guinney J and Middleton GW: *KRAS* mutation and consensus molecular subtypes 2 and 3 are independently associated with reduced immune infiltration and reactivity in colorectal cancer. Clin Cancer Res *24*: 224-233, 2018. PMID: 29061646. DOI: 10.1158/1078-0432.CCR-17-1090

42 Wu P, Zhang G, Zhao J, Chen J, Chen Y, Huang W, Zhong J and Zeng J: Profiling the urinary microbiota in male patients with bladder cancer in China. Front Cell Infect Microbiol *8*: 167, 2018. PMID: 29904624. DOI: 10.3389/fcimb.2018.00167

43 Shin JH, Sim M, Lee JY and Shin DM: Lifestyle and geographic insights into the distinct gut microbiota in elderly women from two different geographic locations. J Physiol Anthropol *35*: 31, 2016. PMID: 27955701. DOI: 10.1186/s40101-016-0121-7

44 Slattery ML, Curtin K, Ma K, Edwards S, Schaffer D, Anderson K and Samowitz W: Diet activity, and lifestyle associations with p53 mutations in colon tumors. Cancer Epidemiol Biomarkers Prev *11*: 541-548, 2002. PMID: 12050095.

Received January 15, 2020
Revised February 4, 2020
Accepted February 11, 2020