DR MICHAEL MCFARLANE (Orcid ID : 0000-0002-8156-8014)

PROFESSOR RAMESH ARASARADNAM (Orcid ID : 0000-0002-2231-3062)

Article type : Original Article

Urinary volatile organic compounds and faecal microbiome profiles in colorectal cancer

Michael McFarlane[1], Andrew Millard[2], Holly Hall[3], Richard Savage[4], Chrystala Constantinidou[3], Ramesh Arasaradnam[1] and Chuka Nwokolo[1]

1. Department of Gastroenterology, University Hospitals Coventry, Clifford Bridge Road, Coventry. 2. Department of Infection, Immunity and Inflammation, University of Leicester, Leicester. 3. Department of Life Sciences, University of Warwick, Coventry. 4. Warwick Systems Biology Centre, University of Warwick, Coventry.

Running title: Urinary VOCs and faecal microbiomes in CRC

Corresponding author:

Michael McFarlane

Department of Gastroenterology

University Hospitals Coventry

Clifford Bridge Road

Coventry

CV2 2DX

mmcf1982@doctors.org.uk

07814886741

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/codi.14739

Author contributions:

MM: Recruited subjects, analysed samples, data analysis, manuscript writing

AM: data analysis

HH: Analysed samples

RS: data analysis

CC: data analysis

RA: data analysis, manuscript writing

CN: project co-ordinator, manuscript writing

Competing interests: the authors have no competing interest to declare

Funding: This work was supported by bursaries from Guts UK (formerly CORE), the Midlands Gastroenterology Society (MGS) and the Medical and Life Sciences Research Fund

Consent: All participants gave written informed consent.

Key words: Colorectal cancer, genetics, environment, VOCs, microbiome

Abbreviations:

ANOSIM: Analysis of Similarity

AUC: Area Under Curve

BCSP: Bowel cancer Screening Programme

CRC: colorectal cancer

FAIMS: Field Asymmetric Ion Mobility Spectrometry

FIT: Faecal Immunochemical Testing

FOBT: faecal occult blood test

LC-FAIMS-MS: Liquid Chromatography- Field Asymmetric Ion Mobility Spectrometry- Mass Spectrometer

MDT: Multi-Disciplinary Team

OTU: operational taxonomic units

PCR: polymerase chain reaction

PPV: positive predictive value

QIIME: Qualitative Insights Into Microbial Ecology

ROC: Receiver Operator Curves

TOF-MS: time of flight mass spectrometer

UHCW: University Hospitals Coventry and Warwickshire

UK: United Kingdom

VOCs: Volatile Organic Compounds

Abstract:

Background

Volatile organic compounds are potential biomarkers for diagnosing colorectal cancer (CRC). We characterised urinary VOCs from CRC patients, their spouses/co-habitors (spouses) and first-degree relatives (relatives) to determine any differences. Correlation with stool-derived microbiomes was also undertaken.

Methods

Urine from 56 CRC patients, 45 spouses and 37 relatives were assayed using Liquid Chromatography: Field Asymmetric Ion Mobility Spectrometry (FAIMS): Mass Spectrometer technology. Analysis was performed using 5-fold cross-validation and a Random Forrest classifier. Faecal microbiome 16s RNA was sequenced using Illumina Miseq protocols and analysed using UPARSE and QIIME pipelines.

VOC and microbiome profiles were also compared before, and after, cancer treatment.

Results

Urinary VOC profiles of CRC patients were indistinguishable from either spouses or relatives. When spouses and relatives were grouped together to form a larger non-cancer control group (n=82), their VOC profiles became distinguishable from CRC patients (n=56) with 69% sensitivity and specificity, area under curve 0.72 ($p<0.001$).

Microbiome analysis identified >1300 operational taxonomic units (OTUs) across all groups. Analysis of Similarity (ANOSIM) R value was 0.067 ($p=<0.001$), with significantly different bacterial abundances in 82 OTUs (6.2%) by Kruskal-Wallis testing.

CRC patients' VOC or stool microbiome profiles were unchanged after treatment.

Conclusion

Although CRC patients' urinary VOC profiles cannot be differentiated from spouses or relatives they can be differentiated from a larger non-cancer control group. Comparison of the groups' microbiomes confirmed differences in bacterial species abundance. The current FAIMS-based assay can detect a unique, but modest, signal in CRC patients' urinary VOCs, which remains unaltered after treatment.

What does this add?

This study adds to the growing evidence of the utility of urinary VOC testing as a non-invasive biomarker for the distinction of patients with colorectal cancer from healthy controls. Here we use crude genetic and environmental controls, whereas previous studies had utilised unconnected controls.

Introduction

There is a worldwide increase in both the incidence and mortality from Colorectal Cancer (CRC) (1). In the United Kingdom (UK), CRC is the fourth commonest cancer and second commonest cause of cancer-related death, with 15,900 deaths in 2014 (2,3). CRC is associated with genetic and environmental factors, including age, sex, family and past medical history, including colorectal polyps, smoking and diet (3, 4, 5, 6, 7, 8, 9, and 10).

Current non-invasive CRC screening tests have had their utility questioned. The UK Bowel cancer Screening Programme (BCSP) utilizes guaiac-based faecal occult blood testing (gFOBT), with a positive predictive value (PPV) of ~10% (11). gFOBT will be superseded by Faecal Immunochemical Testing (FIT). FIT is simpler to perform, both for patients and laboratory staff, with a higher PPV of 41% (12). FIT shows a sensitivity of 66–88%, depending on haemoglobin cut-off values (50–200 ng/ml), and a specificity of 87–96% (13 – 15). Screening based on faecal samples had a reported ~50% uptake rate after the first 1 million cases of the BCSP, although there was variation across the country (11). Uptake in France was reported as 34.3% after the first 2 years (16). Whilst Finland has a 71% uptake (17). This variable uptake reduces the overall efficacy of this strategy. Given this mixed performance, a more effective point-of-care CRC screening test involving non-stool samples is required.

An approach for non-invasive detection of cancers is volatile organic compounds (VOCs) analysis. The detection of VOC patterns by both invasive and non-invasive methods, and their utility as disease-specific gas phase biomarkers, has developed in recent years (18).

CRC has been distinguished from other cancers and healthy controls using exhaled VOC analysis in several studies (19-21). This has been replicated in urinary VOC analysis using Mass Spectrometry (MS) technology (22, 23), and subsequently Electronic Nose and Field Asymmetric Ion Mobility Spectrometry (FAIMS) technology (24, 25). Faecal VOCs have also been studied and shown promise, however, as the uptake of stool-based screening programmes is ~50%, this may well prove of limited clinical benefit long term (26). No studies have identified any consistently unique VOCs, suggesting the overall pattern or "smell print" allows differentiation, rather than individual chemicals. The presence of intestinally derived VOCs in urine is believed to occur via VOC migration into the blood steam and subsequent renal filtration allowing their presence, and detection, in urine (27).

There have been limited studies that have demonstrated a role for the detection of pre-cancerous polyps (15, 28). Indeed one study showed urinary VOCs could distinguish CRC from controls and CRC from advanced adenomas, but poor distinction of adenoma from control (28).The current stool-based assays have lower sensitivity and specificities for pre-malignant lesions (29).

VOCs themselves are a diverse group of carbon-based chemicals which are products of bodily metabolic processes. Pathological states are believed to alter VOC profiles, resulting in disease specific patterns, thus allowing their use as biomarkers (18). Within the context of CRC, this could include the interaction of malignant colonocytes altered metabolism and ingested dietary factor fermentation by the subject's intestinal microbiome (30). The true origin of VOCs remains uncertain.

Recent years has seen expanding interest into the role of gut micro-organisms in health and disease. Studies of the microbiome in CRC patients have demonstrated greatly reduced microbial diversity, termed dysbiosis (31, 32). Specific bacteria identified as over-represented in CRC include; *Streptococcus bovis, Helicobacter pylori, Bacteroides fragilis, Clostridium septicum,* and *Escherichia coli* strains (31, 33, and 34). Conversely, butyrate-producing bacteria, including *Roseburia* and *Fecalibacterium* are reduced (33, 35). It remains unclear whether these changes are a cause, or consequence, of CRC.

This study aimed to characterise urinary VOC and stool microbiome profiles of CRC patients, to ascertain whether these profiles could be distinguished from those of controls whom shared environment exposures (spouses/co-habitors) and genetic factors (first degree relatives). Previous studies have used healthy controls with no links to cancer patients (19-25). The objective was to determine whether VOCs and stool microbiomes remained distinguishable from control groups when the controls shared potential risk factors for CRC, and to provide evidence for the potential utility of urinary VOCs as a non-invasive CRC screening test.

Methods

Ethics:

Scientific and ethical approval was granted by the University Hospitals Coventry and Warwickshire (UHCW) Research and Development Office, and Solihull Ethics committee, ref: 13/WM/0136. Written, informed consent was obtained from all study participants.

Patient recruitment:

Patients were recruited between September 2015 and December 2016 from the Lower GI Multi-Disciplinary Team (MDT). Inclusion criteria were adult patients with confirmed CRC, with no family history to suggest hereditary CRC conditions, and relatives and spouses/co-habitors who had consented. Exclusion criteria were patients/relatives/spouses that had concurrent malignancy, non-malignant gastrointestinal conditions and urological/renal conditions requiring secondary hospital care. Patients with polyps showing advanced neoplasia were excluded.

Relatives and spouses were recruited after consent had been confirmed by the CRC patients. Relatives living with the CRC patients were excluded, as were spouses/partners not co-habiting with the CRC patients. Post-operative samples were collected 3 and 6 months after initial surgery. CRC patients who were 24-36 months post initial treatment and in endoscopic and radiological disease-free surveillance provided urine for analysis. All recruited subjects provided data regarding smoking, alcohol consumption, current medications and a dietary questionnaire. No patient had received bowel preparation or antibiotics within the 4 weeks prior to sample collection.

Sample Collection and storage:

Urine and stool samples were collected and stored at -80^{o}C within 72 hours of voiding/elimination. Urinary VOCs have been shown to be stable at room temperature during this period (unpublished data). They were defrosted in a laboratory fridge at 3^{o}C overnight, prior to analysis.

Urinary VOC and statistical analysis

Urine samples were analysed on the LC-FAIMS-MS equipment (Owlstone, UK) which is a bespoke device combining mass spectrometry and FAIMS technology, in series with an LC column to allow greater separation of chemicals and more sensitive assays of VOC concentrations. The LC column was pre-conditioned using a pooled sample of all 3 groups and samples applied to the column from a chilled autosampler (4^{o}C). The column was

washed with an increasing chromatographic gradient of acetonitrile (5-90%). The eluted fractions from the column were aerosolized by an electrospray and then passed into the FAIMS and time of flight mass spectrometer (TOF-MS). FAIMS and MS settings can be found in supplementary file 1.

Due to the hybrid instrument set-up acquired chromatograms were pseudo MS/MS files with each collision detected corresponding to one of 10 FAIMS compensation fields (CF). This meant it was not possible to extract features directly from raw LC-FAIMS-MS data. To resolve this, acquired data was split into individual chromatograms corresponding to FAIMS CF settings using a custom Python script and the data subjected to feature extraction using XCMS package in R. The output data contained a list of features, or peaks, for the sample. The peaks were force aligned, allowing direct comparison across samples and improved the signal-to-noise ratio. All features were then normalized to have zero mean and unit variance.

For subgroup analysis, the relevant sample subset was extracted, and the different groups defined. 5-fold cross-validation was then used to assess classification accuracy across these groups, using a Random Forest multi-class classifier. This analysis generated outputs of one-vs-all Receiver Operator Curves (ROC) i.e. comparing a single group vs all other groups, for example CRC vs relatives and spouses, relatives vs CRC and spouses etc. Other results generated included the Area-Under-Curve (AUC) statistic, sensitivity/specificity values, and a p-value, comparing the results to that expected by random chance (AUC=0.5), using a Wilcoxon rank-sum test. All analyses were carried out using R programming language.

Faecal 16S microbiome sequencing and statistical analysis

200mg of each stool underwent DNA extraction using a QIAamp Fast DNA stool mini-kit (QIAGEN). The 16S V3-V4 region was then amplified using V3-V4 specific primers and polymerase chain reaction (PCR). The DNA was purified using AMPure magnetic beads (Beckman Coulter) and further PCR performed using unique forward and reverse primers to allow identification of individual subject's sequences. The amplified products were purified using AMPure magnetic beads, pooled and diluted to 4nM concentration before sequencing on an Illumina Miseq platform (Illumina) using a Miseq V3 2 x 300bp paired end protocol.

The raw sequence data was merged and quality controlled using UPARSE software. Contiguous sequences were assigned to operational taxonomic units (OTUs) then clustered and filtered to a default 97% identity level before being linked to corresponding samples and assigned taxonomic classification. QIIME (Qualitative Insights Into Microbial Ecology) software was used to perform analysis using Analysis of Similarity (ANOSiM) and Kruskal-Wallis testing.

Results

Urinary VOC analysis

Seventy-two CRC patients 56 spouses/co-habitors (spouses) and 61 first degree relatives (relatives) and were approached. Urine samples were returned by 56 CRC patients, 45 spouses and 37 relatives.

There was no significant difference in the mean ages between the CRC patients and the spouses, although there was a significant difference between the CRC and spouse groups and the relative group, as a result of children of the CRC patients being included. There were more males in the cancer group and as expected this was reversed in the spouses group. The proportion of males and females in the relatives group was equal. The demographics are shown in table 1.

Urinary VOC analysis was performed using a 5-fold cross-validation, the Random Forest multi-class classifier and Wilcoxon rank-sum test. An analysis in which only CRC patients with a paired with spouse and relative (n=35 in each group) showed no significant differences in their VOC profiles. A larger analysis was performed using all recruited subjects (CRC n=56, Spouses n=45, relatives n=37). This analysis also could not distinguish CRC patients from either their spouses or first degree relatives. There was also no difference when the spouses were compared with the relatives. However when the spouses and relatives were grouped together (n=82) and compared to CRC patients (n=56) the technology was able to detect a difference with a sensitivity of 69% (95% confidence intervals: 54% – 81%), a specificity of 69% (57% – 79%) and Area under Curve (AUC) of 0.71 (0.62 – 0.8). The Bonferroni corrected p value was <0.001. The positive predictive value was 60%. The ROC curve is shown in figure 1.

Analysis of CRC subjects by stage of CRC, site of CRC and pathway for detection (2 week wait, BCSP etc) did not reveal any statistically significant differences.

Post-treatment urine sample analysis:

Of the original 56 CRC patients who provided urine samples before treatment, 23 and 9 returned samples at 3 and 6 months respectively.

Thirty patients who were 24-36 months post CRC resection, and in disease free surveillance, also provided urine samples. Demographics can be found in table 2.

An analysis was again performed using a 5-fold cross-validation and Random Forrest classifier. Comparison of the CRC patients VOC profiles with samples taken at 3 months, 6 months, combined 3 and 6 months, and 24-36 months after treatment revealed no difference.

Faecal 16S microbiome analysis

Seventy-two CRC patients 61 first degree relatives and 56 spouses were approached. Stool samples were returned by 44 CRC patients, 34 first degree relatives and 39 spouses.

The mean ages of the CRC patients, spouses and relatives and spouses were 65.3 years (SD 11.2), 59.9 years (SD 12.4) and 51.2 years (SD 12.6) respectively. Group demographics can be found in table 3.

The 16S V3-V4 region was sequenced using an Illumina Miseq platform. Once the raw sequence data had been merged, quality controlled and filtered to exclude low quality reads, with counts <5000, 5 samples were excluded. This left sample sizes of 41 for CRC, 33 relatives and 38 spouses.

There were 1346 OTUs identified across all samples. Relative abundance plots of the CRC samples (n=41) and the combined non-cancer control group (relatives and spouses: n=71) are in figure 2. The non-parametric analysis of similarity (ANOSiM) test was run using QIIME software. This returned an R value of 0.067 (p=<0.001), indicating CRC samples have a very similar microbiome profile to non-cancer controls. Principal component analysis of the overall diversity of CRC patients and non-cancer controls showed no statistically significant differences in microbiome composition (See figure 3).

The non-parametric Kruskal-Wallis test was performed to identify any significant differences between the CRC and non-cancer controls. It showed 82/1346 OTUs (6.2%) were significantly different between CRC and non-cancer control group, i.e. >93% were not significantly different. Of the 82 OTUs, 46 showed increased abundances and 36 decreased abundances in CRC samples relative to the control group, with 64/82 identified as clostridiales *sp*.

Comparison of the relative and spouse control sub-groups using Kruskal-Wallis testing identified 567 OTUs, with 25 (4.4%), showing statistically significant differences between the two groups.

Analysis of CRC subjects by stage of CRC, site of CRC and pathway for detection (2 week wait, BCSP etc) did not reveal any statistically significant differences.

Post-treatment stool sample analysis:

Of the 44 CRC patients, 15 and 14 returned additional samples at 3 and 6 months respectively. The demographics can be found in table 4.

Microbiome profiles of the CRC patients compared to the profiles at 3 months, 6 months and pooled 3 and 6 months after treatment revealed no significant differences.

The rarefaction curve plots of observed OTUs, plotted against average number of sample reads shows an increase in the number of observed OTUs in post-operative samples, though not to statistically significant levels. See Figure 4.

Discussion

An individual's urinary VOC profile is a sum of the different volatile compounds when exposed to assay methodology. The origin and role of VOCs in human metabolism remains unclear, but likely represents the complex interaction of malignant colonocytes, dietary factors and the individual's microbiome (30). While the specificity of each compound is unknown the VOC profile can give a unique identifier to diseased individuals.

The strength of the VOC signal detectable in urine is variable given the heterogeneity of the earlier pilot studies, which showed diversity of assay methodology, and the comparisons were between cancer and non-cancer controls. The number of subjects in groups studied in these pilot studies was about 20-30 subjects. In this study we aimed to have at least 20 in each group. A non-cancer control group included the spouses/co-habitors of CRC patients who shared their environment and diet and first degree relatives of CRC patients. The CRC patients shared a gene pool with their first degree relatives but did not have a hereditary CRC syndrome.

We found no difference in VOC profiles between the CRC patients and either their spouses or first degree relatives, either in a smaller analysis, with only paired CRC/spouse/relative samples or when all collected samples were included. Our interpretation of this result is that the cancer patients and the two non-cancer control groups shared enough aspects of the make-up of their VOC signals to mask any cancer-specific signal. When the non-cancer controls were grouped and then compared with the CRC patients a significant difference in the VOC signal was identified (AUC, sensitivity and specificity of 0.71, 0.69 and 0.69 respectively). Increasing the number of non-cancer controls raised the power of the study and allowed a modest cancer-specific signal to be detected. There is no validated database to allow identification of individual VOCs and so this was not possible.

These results though modest, conform to previous studies, which have shown distinct patterns of urinary VOCs in CRC patients compared to disease-free controls, (22-25). The PPV was 60%, which compares favourably to the ~8-10% demonstrated for FOBT and FIT (36-39), however, the specificity compares less favourably to FOBT ~98% (39). The relatively poor specificity represents a major obstacle for urinary VOCs entering clinical use for CRC screening. Another study by our group, comparing urinary VOC profiling to FIT testing has shown that FIT proved superior to VOCs in isolation but that a combination of FIT and VOCs raised sensitivity and specificity to 0.97 and 0.72 respectively (40).

There was no difference in urinary VOC profiles of the first-degree relatives and spouse/co-habitor groups. This suggests a heterogenous VOC composition among non-cancer individuals, and neither environmental nor genetic factors appear to significantly affect it.

The spouses/ co-habitors were all consuming the same food as the CRC patients and none of the main groups were consuming special diets (vegan, vegetarian, gluten free). This minimises the potential for VOC profile being confounded by dietary factors. None of the relatives group was consuming special diets. It is possible that each sub-group of the non-cancer control groups were underpowered to detect VOC profile suggesting that any group specific VOC signal is relatively weak, or that any variations have occurred by chance.

A potential weakness in the study design is the absence of endoscopic screening of the first-degree relatives, who are at increased risk of CRC, compared to the co-habitors group. Currently the risk of CRC in asymptomatic first-degree relatives of patients with sporadic CRC is not high enough to attract endoscopic screening in the UK.

This is the first study to include healthy control subjects that are connected to CRC subjects, by a shared environment, or shared gene pool. All previous studies had used unrelated healthy controls (19-25). Despite the relatedness of the non-cancer controls in this study, they could still be distinguished from the CRC group, with similar sensitivities, specificities and AUCs to those reported previously.

There was no significant difference in VOC signals between the CRC subjects when sub-analyses of stage of CRC, site of CRC and pathway for detection (2 week wait, BCSP etc) were performed. This is most likely due to under powering as a result of diminishing sample sizes.

There was no significant difference in VOC's and microbiome between pre-treatment and post-treatment CRC samples. This suggests that either the urinary VOC profile detected in CRC patients does not originate from the cancerous tissue or confirms the weakness of the signal. Either scenario would make VOC analysis unsuitable for CRC surveillance using the current FAIMs technology. Altomare et al found that exhaled breath analysis, 2 years after curative surgery, using GC-MS, distinguished pre- from post-operative samples with a sensitivity of 100%, specificity of 97%, accuracy of 98.8% and AUC 1 (41). The sample size of 32 was similar to our cohort, so may indicate GC-MS is better suited than LC-FAIMS-MS. Ma et al had previously demonstrated a significant reduction in 2 urinary VOCs post-operatively, using GC-MS, although they did not suggest its use for CRC surveillance (22).

The microbial 16s profiling of CRC samples and healthy non-cancer controls (relatives and spouses) was performed using Illumina 16s sequencing and bioinformatics analysis. It revealed no significant difference for >93% of identified OTUs between CRC patients and the non-cancer control group. Of 1346 identified OTUs, only 82 (6.2%) were significantly different between CRC patients and non-cancer controls, with 64 identified as clostridiales

sp. Other orders with higher representation among CRC patients included fusobacteroides and coriobacteroides, as found previously (31, 33, 34, 42). Given that most studies have identified bacterial differences to order or family, rather than species level, this would suggest a generalised alteration of microbial profile, possibly related to a more anaerobic environment, as clostridiales are obligate anaerobes. The relative abundance and PCA plots confirm that the overall microbiome composition is very similar between CRC and non-cancer control groups, with the ANOSIM analysis giving an R value of 0.067. These findings could be in-keeping with the theory of dysbiosis in CRC, or it could represent normal microbial population variation. Fewer stool samples were returned than urine samples, reflecting the lower uptake of faecal matter based tests, as seen in the BCSP (11).

The microbial 16S analysis of CRC patients samples before and after treatment revealed no significant differences for >95% of OTUs.

The microbiome role in VOC production has yet to be fully explored, although colonic fermentation is believed to play a role (30). The results of the 16S microbiome analysis suggest that any microbiome profile differences between the groups are minor. This would suggest they are not responsible for the VOC profile differences observed between the groups, although this cannot be completely excluded.

The current favoured theory for the microbiome role in CRC development is the “intestinal microbiota adaptions” model (43-45), which suggests CRC and dysbiosis have a symbiotic relationship. The CRC environment contains host-derived immune and inflammatory processes affecting microbial regulation. This could alter microbiome composition, favouring pro-carcinogenic bacteria proliferation, amplifying dysbiotic effects, and further promoting CRC. This appears more likely than the “alpha bug” and “driver/passenger” theories, as to-date, there have been no consistent bacterial species linked to CRC.

Conclusions

This study demonstrates that CRC patients have a unique urinary VOC profile different from non-cancer controls. This cancer-specific signal is modest, requiring the pooling on non-cancer controls, and becomes detectable at n=/> 50, however, this adds to growing evidence that urinary VOCs may have a potential future role in distinguishing CRC patients from non-cancer controls, whether they are independent controls as in previous studies, or related controls as in this study. Current technology is improving and should in future be able to detect cancer-specific signals in urine with greater specificity and sensitivity. Comparison with the stool-based screening tests will continue but the poor uptake of stool testing will hand an advantage to any urine-based screening tests developed.

References

1. Ferlay J, Shin HR, Bray F et al.Estimates of worldwide burden of cancer in 2008:GLOBOCAN 2008.Int J Cancer.2010;127:2893-917.
2. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Zero
3. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#ref-0
4. Murphy G, Devesa SS, Cross AJ, et al.Sex disparities in colorectal cancer incidence by anatomic subsite, race and age.Int J Cancer. 2011;128:1668-75.
5. Galiatsatos P, Foulkes WD. Familial adenomatous polyposis. Am J Gastroenterol.2006;101:385-98.
6. Gala M, Chung DC.Hereditary colon cancer syndromes. Semin Oncol.2011;38(4):490-9.
7. Butterworth AS, Higgins JP, Pharoah P.Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis.Eur J Cancer.2006;42(2):216-27.
8. Fonseca-Nunes A, Jakszyn P, Agudo A.Iron and cancer risk-a systematic review and meta-analysis of the epidemiological evidence.Cancer Epidemiol Biomarkers Prev.2014;23(1):12-31.
9. Parkin DM, Boyd L.Cancers attributable to dietary factors in the UK in 2010. III. Low consumption of fibre.Br J Cancer. 2011;105 Suppl 2:S27-30.
10. Huxley RR, Ansary-Moghaddam A, Clifton P, et al.The impact of dietary and lifestyle risk factors on risk of colorectal cancer:a quantitative overview of the epidemiological evidence. Int J Cancer.2009;125(1):171-80.
11. Logan RF, Patnick J, Nickerson C, et al.Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests.Gut. 2012;61(10):1439-46.
12. Chiang TH, Lee YC, Tu CH, et al.Performance of the immunochemical fecal occult blood test in predicting lesions in the lower gastrointestinal tract. CMAJ 2011.183(13):1474-81.
13. Brenner H, Tao S (2013)Superior diagnostic performance of faecal immunochemical tests for haemoglobin in a head-to-head comparison with guaiac based faecal occult blood test among 2235 participants of screening colonoscopy.Eur J Cancer 49(14)3049–54.

14. Imperiale TF (2012)Non-invasive screening tests for colorectal cancer.Dig Dis 30 Suppl 2: 16–26.
15. De Meij TG, Ben Larbi I, van der Schee MP, et al.(2014)Electronic nose can discriminate colorectal carcinoma and advanced adenomas by fecal volatile biomarker analysis: proof of principle study. Int J Cancer 134(5) 1132–1138.
16. Leuraud K, Jezewski-Serra D, Viguier J, Salines E. Colorectal cancer screening by guaiac faecal occult blood test in France: Evaluation of the programme two years after launching. Cancer Epidemiol. 2013 Dec;37(6):959-67.
17. Malila N, Oivanen T, Hakama M. Implementation of colorectal cancer screening in Finland: experiences from the first three years of a public health programme. Z Gastroenterol. 2008 Apr;46 Suppl 1:S25-8.
18. Schmidt K, Podmore I.Current Challenges in Volatile Organic Compounds Analysis as Potential Biomarkers of Cancer.J Biomark.2015;2015:981458.
19. Peng G, Tisch U, Adams O, et al.Diagnosing lung cancer in exhaled breath using gold nanoparticles. Nat Nanotechnol.2009;4(10):669-73.
20. Altomare DF, Di Lena M, Porcelli F, et al.Exhaled volatile organic compounds identify patients with colorectal cancer.Br J Surg.2013;100(1):144-50.
21. Amal H, Leja M, Funka K, et al.Breath testing as potential colorectal cancer screening tool. Int J Cancer.2016;138(1):229-36.
22. Ma YL, Qin HL, Liu WJ, et al.Ultra-high performance liquid chromatography-mass spectrometry for the metabolomic analysis of urine in colorectal cancer.Dig Dis Sci.2009;54(12):2655-62.
23. Silva CL, Passos M, Câmara JS.Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry.Br J Cancer.2011;105(12):1894-904.
24. Arasaradnam RP, McFarlane MJ, Ryan-Fisher C et al.Detection of colorectal cancer by urinary volatile organic compound analysis. PLoS One.2014;9(9):e108750.
25. Westenbrink E, Arasaradnam RP, O'Connell N, et al.Development and application of a new electronic nose instrument for the detection of colorectal cancer.Biosens Bioelectron.2015;67:733-8.
26. Bosch S, Berkhout DJ, Ben Larbi I, de Meij TG, de Boer NK. Fecal volatile organic compounds for early detection of colorectal cancer: where are we now? J
27. Cancer Res Clin Oncol. 2019 Jan;145(1):223-234.Haick H, Broza YY, Mochalski P, et al.Assessment, origin, and implementation of breath volatile cancer markers.Chem Soc Rev. 2014.43(5):1423-49.
28. Mozdiak E, Wicaksono AN, Covington JA, Arasaradnam RP. Colorectal cancer and adenoma screening using urinary volatile organic compound (VOC) detection: early results from a single-centre bowel screening population (UK BCSP). Tech Coloproctol. 2019 Apr 15.
29. Health Quality Ontario.Fecal occult blood test for colorectal cancer screening: an evidence-based analysis.Ont Health Technol Assess Ser.2009;9(10):1-40.

30. Arasaradnam RP, Covington JA, Nwokolo CU.Review article:next generation diagnostic modalities in gastroenterology--gas phase volatile compound biomarker detection. Aliment Pharmacol Ther.2014.39(8):780-9.
31. Gagnière J, Raisch J, Veziant J, et al.Gut microbiota imbalance and colorectal cancer.World J Gastroenterol.2016;22(2):501-18.
32. Bultman SJ.Interplay between diet, gut microbiota, epigenetic events, and colorectal cancer.Mol Nutr Food Res. 2016.
33. Bultman SJ.Emerging roles of the microbiome in cancer.Carcinogenesis.2014;35(2):249-55.
34. Sears CL, Garrett WS. Microbes, microbiota, and colon cancer.Cell Host Microbe.2014;15(3):317-28.
35. Burns MB, Lynch J, Starr TK et al.Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment.Genome Med.2015;7(1):55.
36. Rabeneck L, Rumble RB, Thompson F, et al. Fecal immunochemical tests compared with guaiac fecal occult blood tests for population-based colorectal cancer screening.Can J Gastroenterol.2012;26(3):131-47.
37. Steele RJ, McDonald PJ, Digby J, et al.Clinical outcomes using a faecal immunochemical test for haemoglobin as a first-line test in a national programme constrained by colonoscopy capacity.United European Gastroenterol J.2013;1(3):198-205.
38. van Rossum LG, van Rijn AF, Laheij RJ, et al.Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology.2008;135(1):82-90.
39. Kearns B, Whyte S, Patnick J.Guaiac faecal occult blood test performance at initial and repeat screens in the English Bowel Cancer Screening Programme.Br J Cancer.2014;111(9):1734-41.
40. Widlak MM, Neal M, Daulton E, Thomas CL, Tomkins C, Singh B, Harmston C, Wicaksono A, Evans C, Smith S, Savage RS, Covington JA, Arasaradnam RP. Risk stratification of symptomatic patients suspected of colorectal cancer using faecal and urinary markers. Colorectal Dis. 2018 Dec;20(12):O335-O342
41. Altomare DF, Di Lena M, Porcelli F, et al.Effects of Curative Colorectal Cancer Surgery on Exhaled VOCs and Potential Implications in Clinical Follow-up.Ann Surg. 2015;262(5):862-6;discussion 6-7.
42. Ohigashi S, Sudo K, Kobayashi D, et al.Changes of the intestinal microbiota, short chain fatty acids, and fecal pH in patients with colorectal cancer.Dig Dis Sci.2013;58(6):1717-26.
43. Consortium HMP.Structure, function and diversity of the healthy human microbiome.Nature.2012;486(7402):207-14.
44. Schwabe RF, Jobin C.The microbiome and cancer. Nat Rev Cancer.2013;13(11):800-12
45. Yu YN, Fang JY.Gut Microbiota and Colorectal Cancer. Gastrointest Tumors.2015;2(1):26-32.

Group	Pre-treatment CRC	Relative	Spouse	P value
Samples	56	37	45	
Mean age (SD)	65.4 (11.5)	50.0 (14.1)*	60.7 (12.1)	* <0.01
Sex (M:F)	33:23	17:20	15:30	0.016
Number of cigarettes smoked per day (SD)	1.5 (4.2)	2.5 (5.3)	1.6 (5.7)	0.26
Alcohol units per week (SD)	8.8 (11.6)	7.3 (8.1)	6.7 (10.1)	0.94
Mean BMI (SD)	27.5 (5.2)	25.5 (3.8)	26.7 (4.8)	0.45
Dukes stage (%)				
A	8 (14.2%)	-	-	
B	17 (30.4%)	-	-	
C1	20 (35.7%)	-	-	
C2	9 (16.1%)	-	-	
Site (%)				
Right	24 (42.8%)	-	-	
Left	17 (30.4%)	-	-	
Rectal	15 (26.8%)	-	-	
Referral Route (%)				
BCSP	18 (32.1%)	-	-	
2 Week Wait pathway	31 (55.4%)	-	-	
Other	7 (12.5%)	-	-	

Table 1. Demographic data of CRC patients, their first degree relatives and spouses who provided urine samples for VOC analysis.

Group	Pre-operative CRC	3 months post-operative	6 months post-operative	24-36 month post-operative	P values
Samples	56	23	9	30	
Mean age (SD)	65.4 (11.5)	66.5. (13.0)	64.7 (12.7)	66.5 (11.7)	0.57
Sex (M:F)	33:23	12:11	5:4	20:10	0.75
Number of cigarettes smoked per day (SD)	1.5 (4.2)	0.5 (2.1)	0 (0)	0.2 (0.9)	0.16
Alcohol units per week (SD)	8.1 (11.6)	6.5 (11.2)	9.9 (15.6)	6.6 (9.0)	0.83
Mean BMI (SD)	27.5 (5.2)	26.5 (7.7)	26.7 (11.6)	27.2 (4.5)	0.93
Dukes stage (%)					
A	8 (14.2%)	4 (17.4%)	1 (11.1%)	7 (23.3%)	0.93
B	17 (30.4%)	7 (30.4%)	2 (22.2%)	10 (33.3%)	
C1	20 (35.7%)	9 (39.1%)	5 (55.5%)	9 (30%)	
C2	9 (16.1%)	2 (8.7%)	0 (0%)	3 (10%)	
Site (%)					
Right	24 (42.8%)	13 (56.5%)	4 (44.4%)	14 (46.7%)	0.52
Left	17 (30.4%)	7 (30.4%)	4 (44.4%)	6 (20%)	
Rectal	15 (26.8%)	3 (13.4%)	1 (11.1%)	10 (33.3%)	
Referral Route (%)					
BCSP	18 (32.1%)	-	-	5 (16.7%)	0.20
2 Week Wait pathway	31 (55.4%)	-	-	17 (56.7%)	
Other	7 (12.5%)	-	-	7 (23.3%)	

Table 2. Demographic data of pre-operative, post-operative and 24-36 month disease free surveillance CRC patients who provided urine samples for VOC analysis.

Group	Pre-treatment CRC	Relative	Spouse	P value
Samples	44	34	39	
Mean age (SD)	65.3 (SD 11.2)	51.2 (SD 12.6)*	59.9 (SD 12.4)	*<0.01
Sex (M:F)	25:19	15:19	13:26	0.1
Number of cigarettes smoked per day (SD)	0.73 (2.6)	2.9 (5.6)	2 (6.4)	0.24
Alcohol units per week (SD)	7.8 (11.6)	6.8 (7.0)	8.1 (10.9)	0.94
Mean BMI (SD)	27.4 (4.3)	25.3 (4.1)	26.6 (4.8)	0.30
Dukes stage (%)				
A	9 (18.2%)	-	-	
B	15 (34.1%)	-	-	
C1	13 (29.5%)	-	-	
C2	5 (11.4%)	-	-	
Site (%)				
Right	18 (40.1%)	-	-	
Left	14 (31.2%)	-	-	
Rectal	12 (27.3%)	-	-	
Referral Route (%)				
BCSP	17 (40.5%)	-	-	
2 Week Wait pathway	21 (50%)	-	-	
Other	6 (9.5%)	-	-	

Table 3. Demographic data of CRC patients, their first degree relatives and spouses who provided stool samples for 16S microbiome analysis.

Group	Pre-operative CRC	3 months post-operative	6 months post-operative	P values
Samples	44	15	14	
Mean age (SD)	65.3 (11.2)	63.2 (12.1)	66.9 (11.4)	0.69
Sex (M:F)	25:19	8:7	7:7	0.90
Number of cigarettes smoked per day (SD)	0.73 (2.6)	0.67 (2.5)	0.7 (2.6)	1
Alcohol units per week (SD)	7.8 (11.6)	6.1 (11.2)	7.2 (12.5)	0.91
Mean BMI (SD)	27.4 (4.3)	28.5 (7.7)	26.8 (4.6)	0.61
Dukes stage (%)				
A	9 (21.4%)	3 (20.0%)	3 (21.4%)	0.99
B	15 (35.7%)	5 (33.3%)	5 (35.7%)	
C1	13 (30.1%)	5 (33.3%)	6 (42.9%)	
C2	5 (11.9%)	2 (13.3%)	0 (0%)	
Site (%)				
Right	18 (42.9%)	9 (60%)	9 (64.3%)	0.45
Left	13 (31.0%)	4 (26.7%)	4 (28.6%)	
Rectal	11 (26.2%)	2 (13.3%)	1 (7.1%)	

Table 4. Demographic data of pre-operative, post-operative and 24-36 month disease free surveillance CRC patients who provided stool samples for 16S microbiome analysis

Figure 1. ROC plot of sensitivity and specificity for Random Forrest analysis of urinary VOC data.

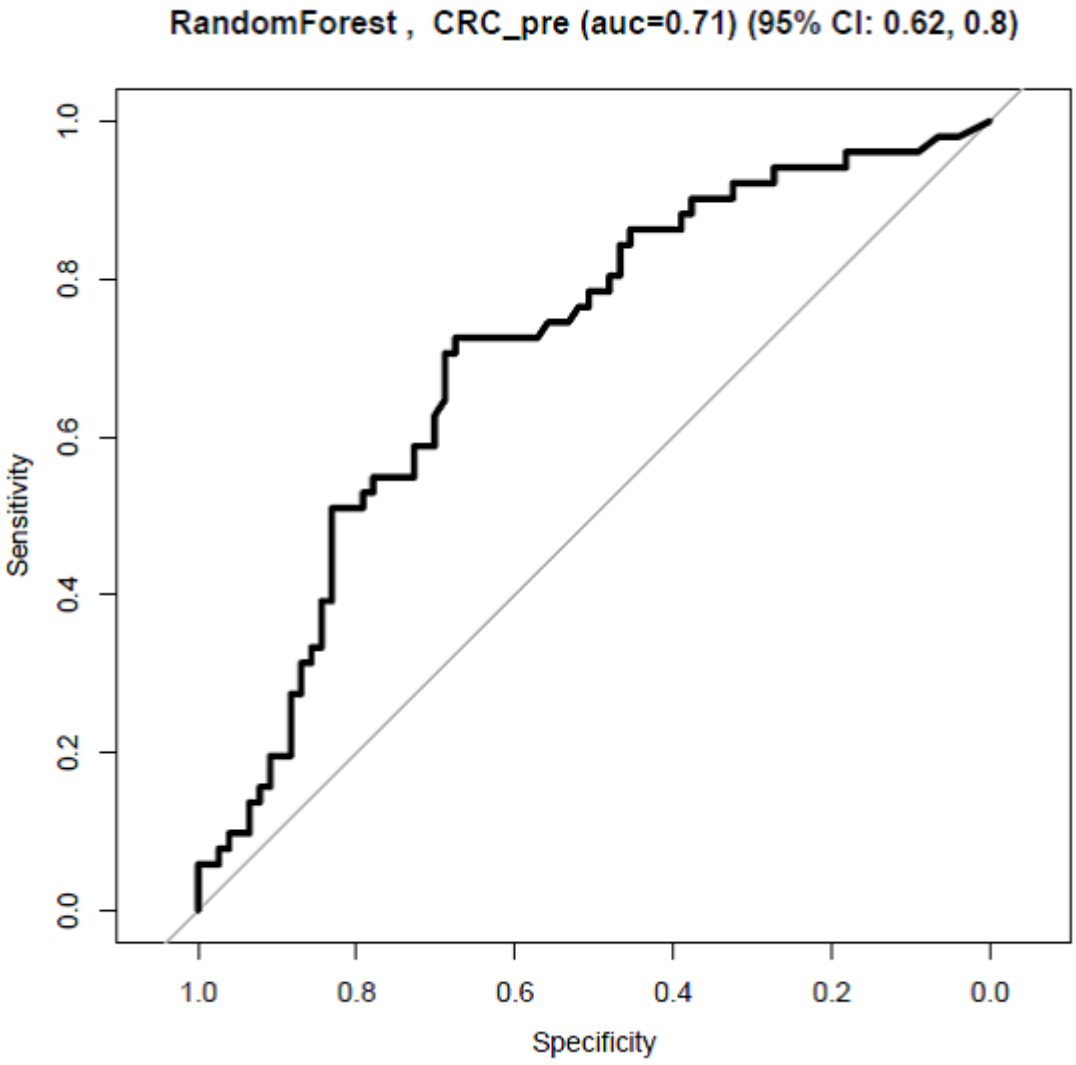

Figure 2. Relative abundance plots of pre-treatment CRC patients (left) and all healthy controls: relatives and spouses (right).

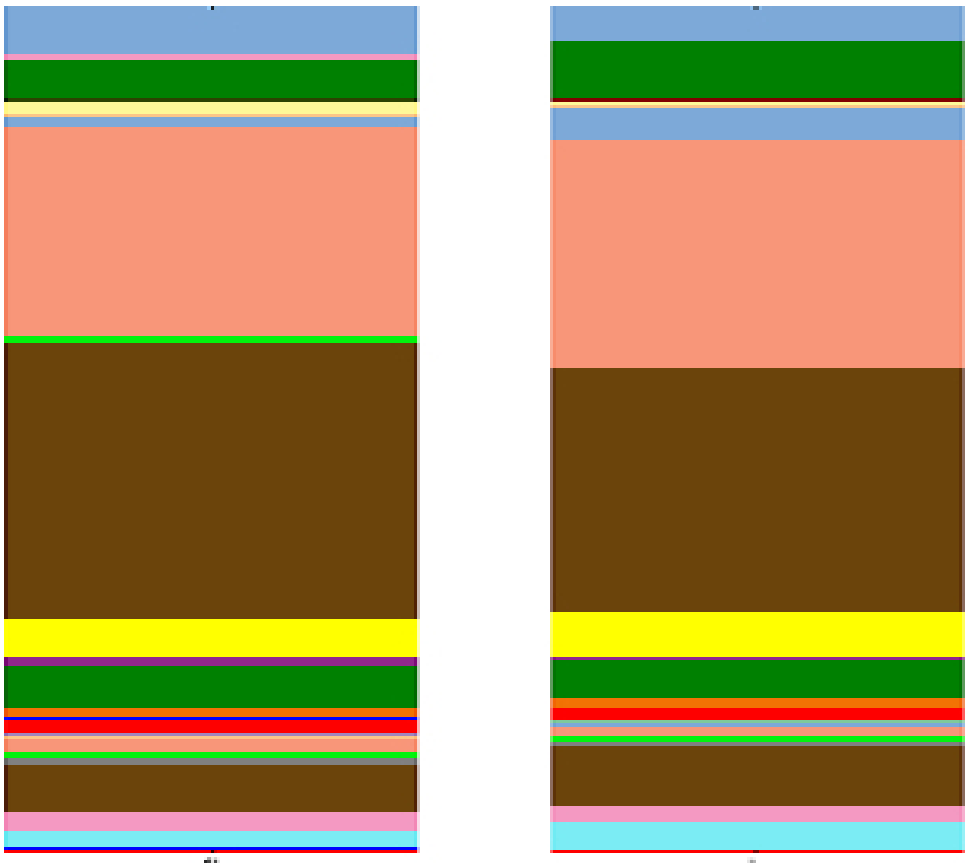

Figure 3. Principle component analysis of the overall diversity of pre-treatment CRC (red) and healthy controls (blue).

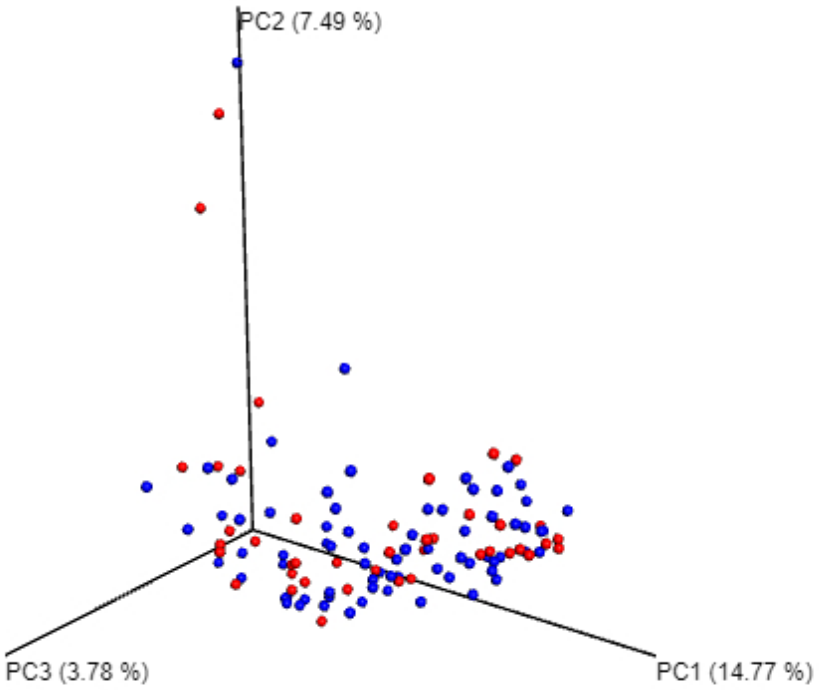

Figure 4. Rarefaction curve of observed number of OTUs against sequences per sample for pre-treatment samples (blue) and 6 month post-treatment samples (red).