

Article

Volatile Markers for Cancer in Exhaled Breath—Could They Be the Signature of the Gut Microbiota?

Manohar Prasad Bhandari ^{1,*}, Inese Polaka ^{1,*}, Reinis Vangravs ¹, Linda Mezmale ^{1,2,3}, Viktors Veliks ¹, Arnis Kirshners ¹, Pawel Mochalski ^{4,5}, Emmanuel Dias-Neto ⁶ and Marcis Leja ^{1,7,8}

¹ Institute of Clinical and Preventive Medicine, University of Latvia, LV-1586 Riga, Latvia; reinis.vangravs@lu.lv (R.V.); linda.mezmale@lu.lv (L.M.); viktors.veliks@lu.lv (V.V.); arnis.kirshners@lu.lv (A.K.); marcis.leja@lu.lv (M.L.)

² Riga East University Hospital, LV-1038 Riga, Latvia

³ Faculty of Residency, Riga Stradins University, LV-1007 Riga, Latvia

⁴ Institute of Chemistry, Jan Kochanowski University of Kielce, PL-25406 Kielce, Poland; pawel.mochalski@ujk.edu.pl

⁵ Institute for Breath Research, University of Innsbruck, A-6850 Dornbirn, Austria

⁶ Laboratory of Medical Genomics, A.C. Camargo Cancer Center, Sao Paulo 01508-010, Brazil; emmanuel@accamargo.org.br

⁷ Digestive Diseases Center GASTRO, LV-1079 Riga, Latvia

⁸ Faculty of Medicine, University of Latvia, LV-1586 Riga, Latvia

* Correspondence: manohar_prasad.bhandari@lu.lv (M.P.B.); inese.polaka@lu.lv (I.P.)

Abstract: It has been shown that the gut microbiota plays a central role in human health and disease. A wide range of volatile metabolites present in exhaled breath have been linked with gut microbiota and proposed as a non-invasive marker for monitoring pathological conditions. The aim of this study was to examine the possible correlation between volatile organic compounds (VOCs) in exhaled breath and the fecal microbiome by multivariate statistical analysis in gastric cancer patients ($n = 16$) and healthy controls ($n = 33$). Shotgun metagenomic sequencing was used to characterize the fecal microbiota. Breath-VOC profiles in the same participants were identified by an untargeted gas chromatography–mass spectrometry (GC–MS) technique. A multivariate statistical approach involving a canonical correlation analysis (CCA) and sparse principal component analysis identified the significant relationship between the breath VOCs and fecal microbiota. This relation was found to differ between gastric cancer patients and healthy controls. In 16 cancer cases, 14 distinct metabolites identified from the breath belonging to hydrocarbons, alcohols, aromatics, ketones, ethers, and organosulfur compounds were highly correlated with 33 fecal bacterial taxa (correlation of 0.891, p -value 0.045), whereas in 33 healthy controls, 7 volatile metabolites belonging to alcohols, aldehydes, esters, phenols, and benzamide derivatives correlated with 17 bacterial taxa (correlation of 0.871, p -value 0.0007). This study suggested that the correlation between fecal microbiota and breath VOCs was effective in identifying exhaled volatile metabolites and the functional effects of microbiome, thus helping to understand cancer-related changes and improving the survival and life expectancy in gastric cancer patients.

Keywords: breath analysis; volatile organic compounds; VOCs; breath biomarker; fecal microbiota; canonical correlation analysis; gastric cancer



Citation: Bhandari, M.P.; Polaka, I.; Vangravs, R.; Mezmale, L.; Veliks, V.; Kirshners, A.; Mochalski, P.; Dias-Neto, E.; Leja, M. Volatile Markers for Cancer in Exhaled Breath—Could They Be the Signature of the Gut Microbiota? *Molecules* **2023**, *28*, 3488. <https://doi.org/10.3390/molecules28083488>

Academic Editors: Natalia Drabińska, Ben de Lacy Costello and Martyna Natalia Wiczorek

Received: 27 March 2023

Revised: 11 April 2023

Accepted: 13 April 2023

Published: 15 April 2023



Copyright: © 2023 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/>).

1. Introduction

The gut microbiota plays an important role in maintaining the healthy metabolism in the host. This regulation of the host health relies on the production of different classes of metabolites by the microbiota and their underlying molecular mechanisms in both healthy and diseased conditions [1,2]. Therefore, microbiota-derived metabolites are direct mediators of the microbe–microbe interactions and host–microbial interactions. The presence

and abundance of these diverse metabolites that can affect our health are subject to the microbiota composition, in addition to dietary intake, drug use, and environmental factors [3]. The volatile organic compounds (VOCs) released within the gut as metabolites are absorbed into the bloodstream, distributed amongst tissues and organs, transported to the lungs, and subsequently, excreted via the exhaled breath. This suggests that breath VOCs could provide a valuable insight into the gut microbiota and disease diagnosis [4].

The gut microbiome has been linked with a plethora of diseases, gastrointestinal disorders, and different types of cancer, including gastric cancer and colorectal cancer [5–7]. This may lead to the deviation of various metabolic pathways and altered concentrations of certain VOCs reflected in different bodily fluids, cell lines, and tissues, including exhaled breath.

The identification of volatile metabolites as potential biomarkers of a disease through non-invasive analysis still remains a challenge. Breath analysis is an easy, non-invasive, real-time, and innovative method used to analyze the chemical composition of exhaled breath [8,9]. In particular, VOCs are helpful in disease diagnosis and have great potential as indicators of metabolic disorders, disease onset, and progression [10–13]. Furthermore, breath VOCs provide a source of useful biomarkers with possible associations to the body's metabolism, along with the fecal microbes that contribute to numerous molecular mechanisms [14]. Therefore, the analysis of volatile metabolites in the breath could be used as a diagnostic and monitoring tool for diverse gastrointestinal diseases and cancers.

Several well-established volatile breath metabolites, such as aldehydes, ketones, and hydrocarbons; bacterial fermentation products, including alcohols; and short-chain fatty acids (SCFAs), such as acetate, butyrate, and propionate have been correlated with gut microbiota and are exhaled via the breath [15,16]. However, the origins of volatile metabolites and the intermediate products are not fully understood. The VOCs emitted by feces in the gastrointestinal tract could be transported into the blood and appear, finally, in the exhaled breath [17]. The human gut microbiome is composed of 10–100 trillion microorganisms with more than 400 bacterial species belonging to predominant phyla, such as *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *Verrucomicrobia*, *Fusobacteria*, and *Synergistetes* [18]. Gut microbes are exogenous sources of ethanol, which can be detected in the breath. An impaired gut microbiome and the implicated bacteria, such as *Fusobacterium nucleatum*, *Escherichia coli*, and *Bacteroides fragilis* were correlated with colorectal cancer [19]. Several facultative anaerobic species, such as *Dialister invisus*, *Parabacteroides* spp., *Lachnospiraceae*, and *Ruminococcaceae*, as well as strains such as *Faecalibacterium prausnitzii* and *Clostridium clostridiiforme*, have been associated with metabolite disturbances, including SCFAs, bile acids (BAs), choline, and specific pathways in several gastrointestinal and inflammatory disorders [20–22]. The gut microbiota is likely to play a significant role in the advancement and progression of gastric cancer.

In the last decade, the advancement and the application of metagenomics, metabolomics, breathomics, and omics data analysis (i.e., multi-omics approaches) have led to a better understanding of the fecal and breath volatilome and their contribution to pathogen virulence and overall gut health [23]. A number of efforts have been made to identify breath markers for pathogenic bacteria infecting the gut [24–26]. However, the main challenges in this context are the non-specificity of these VOCs and the confounding factors that prevent the use of these volatile markers in disease diagnostics. Understanding the complex and underlying biochemical mechanisms that link the fecal microbes to exhaled VOCs could help to illustrate the disease pathology and develop novel therapies and screening methods.

Studies have identified over 2700 VOCs in human breath [27]. Surprisingly, to date, a breath signature that reflects the interaction between fecal microbiota and breath metabolites within the context of gastric cancer has not been studied. In this article, we present an integrated statistical approach used to investigate possible correlations between breath VOCs, as identified by gas chromatography–mass spectrometry (GC–MS), and fecal microbiota, through sequencing, in gastric cancer patients and healthy controls. In general, some

of the specific breath metabolites were likely to have a direct or indirect link with the fecal microbiota composition and their presence or abundance varied between gastric cancer patients and healthy controls.

2. Results

2.1. Study Population

The analysis included a total of 49 participants. All participants provided both breath and fecal samples. Out of these participants, 16 were gastric cancer patients and 33 were healthy controls. The patient group showed different clinical stages of gastric cancer, starting at IA and going up to IV. An overview of clinical and demographic data is presented in Table 1. The data are presented as percentages or medians \pm .

Table 1. Demographic data of the study cohort.

		Cancer	Control	Total
Gender	Female	10 (62.5%)	24 (72.7%)	34 (69.4%)
	Male	6 (37.5%)	9 (27.3%)	15 (30.6%)
Age		62.2 \pm 11.5	51.4 \pm 14.3	54.9 \pm 14.3
Stage	IA	5 (31.2%)	-	-
	IB	0 (0.0%)	-	-
	IIA	0 (0.0%)	-	-
	IIB	4 (25.0%)	-	-
	IIIA	0 (0.0%)	-	-
	IIIB	1 (6.3%)	-	-
	IIIC	0 (0.0%)	-	-
	IV	4 (25.0%)	-	-
N/A ¹		2 (12.5%)	-	-

¹ N/A = not available.

2.2. Microbiome Analysis

A total of 6.2 billion raw reads were obtained from all the samples analyzed. After quality filtering, trimming (mean quality score of 30 and above 50 bp read length), and host sequence removal, 6 billion paired reads remained with an average of 123 million paired reads per sample, in a range between 86 and 178 million. The negative control added during the extraction did not provide enough DNA for library construction, thus indicating a low possibility of reagent and human contamination. Overall, there were 1324 bacterial species identified after the initial analysis of the microbiome composition using the MetaPhlAn4 marker database. At the phylum level, the fecal microbiota of cancer and the control patient samples were, on average, dominated by *Bacteroidetes* at 24.9% and 30.2%; *Firmicutes* at 67.5% and 60.7%; *Actinobacteria* at 3.3% and 3.5%; and *Proteobacteria* at 2.6% and 3%, respectively (Figure 1). After obtaining data from both Bracken and MetaPhlAn4, we compared the results and found them to be highly similar. As a result, we decided to perform further taxonomic analysis using the MetaPhlAn4 data exclusively. A permutational multivariate analysis of variance (PERMANOVA) was conducted to determine if the taxonomic data obtained from MetaPhlAn4 or Bracken revealed any significant differences between the fecal microbiomes in cancer patients and the control group. The results revealed that neither the MetaPhlAn4 nor the Bracken-obtained taxonomic data showed that the patient group (control or cancer) was the main driver of the observed microbiome variance among individual samples (p -value < 0.05), as in both cases only 4.7% and 3.0% of the variance could be explained, respectively.

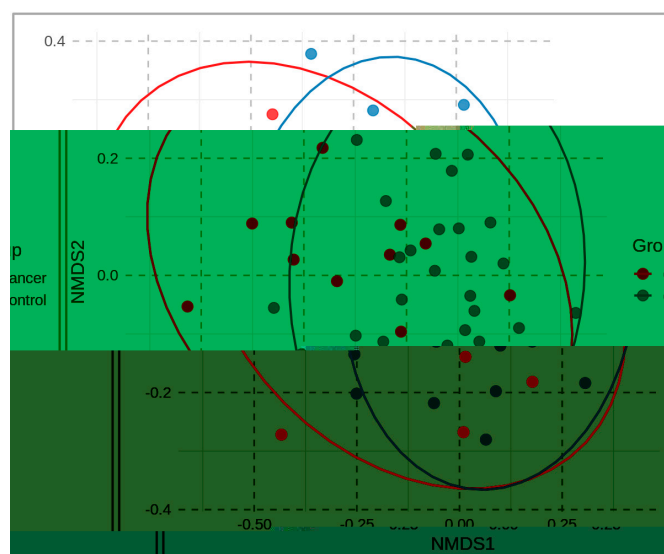


Figure 2. Two-dimensional NMDS plot of fecal microbial communities, based on beta-diversity (OTU level), derived from fecal samples of cancer patients (red) and healthy controls (blue).

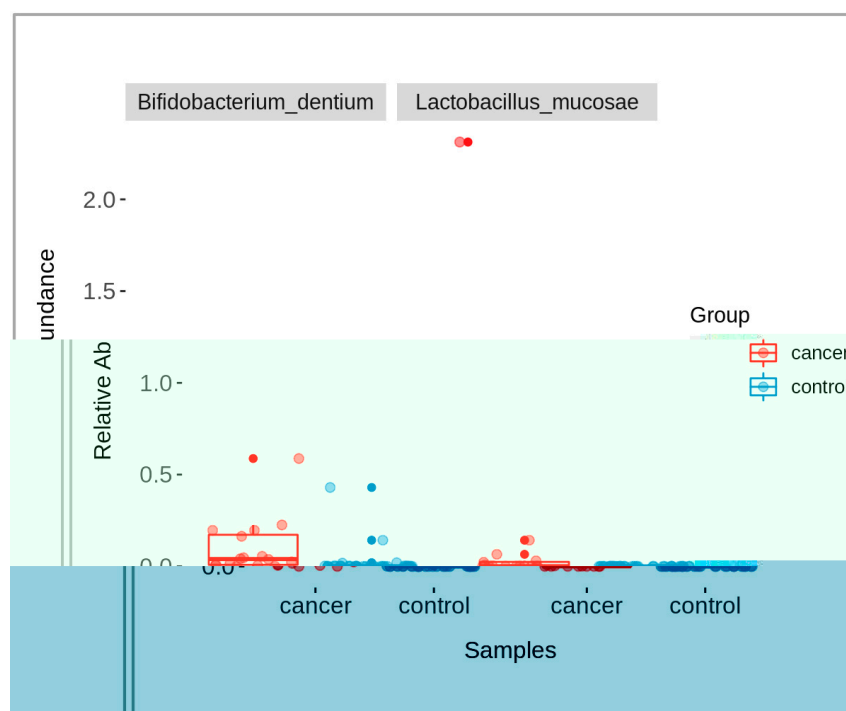


Figure 3. A box-plot showing the presence of significantly different bacterial taxa between cancer (red) and control (blue) groups (p -value < 0.05).

2.3. Analysis of Breath VOCs

Altogether, 378 VOCs were detected in the exhaled breaths of the participants using GC–MS. Since the majority exhibited low occurrence, only VOCs that were present in at least 20% of the samples were used for further analysis. Effectively, there were 116 breath volatile metabolites in the analysis dataset.

After a sensitivity analysis by backwards elimination of the least-contributing volatile metabolites, the most important metabolites identified in this study were selected, resulting in sets containing 14 and 7 breath VOCs for the cancer and control groups, respectively. These VOCs contributed the most to the overall correlation between the two datasets and

were designated as putatively relevant in the interaction with fecal microbiota. Interestingly, none of these reported metabolites were common across the two groups.

Figure 4 shows a cluster heatmap of the breath volatile metabolites identified in the breath samples of all 49 participants. The VOCs identified in the exhaled breath of the gastric cancer patients were the following: 1-dodecene; 1-octanol; 2-pentanone; 2-propanol, 1-(2-methoxy-1-methylethoxy)-; benzene; 1-methyl-3-(1-methylethyl)-benzene; n-decane; dimethyl disulfide (DMS); 5,8-diethyl-dodecane; 2-phenoxy-ethanol; n-nonane; 1,1'-oxybis-octane; toluene; and n-hexane. Similarly, the VOCs identified in the exhaled breath of the healthy controls were the following: 1-decanol; 3-methylcyclopentyl acetate; 4-ethylbenzamide; benzyl alcohol; methacrolein; phenol, 2-methyl-4-(1,1,3,3-tetramethylbutyl)-; and dl-erythro-1-phenyl-1,2-propanediol. The major compounds showing higher concentration levels and observed in most of the subjects were phenol, 2-methyl-4-(1,1,3,3-tetramethylbutyl)-; toluene; benzyl alcohol; benzene; 1,1'-oxybis-octane; 2-phenoxy-ethanol; 2-pentanone; and n-nonane.

Figure 4. Cluster heatmap of breath VOC levels for each study participant (21 VOCs and 49 participants). The *x*-axis depicts the respective breath VOCs and the *y*-axis, the patient IDs. Red color on the left vertical axis indicates gastric cancer cases, whereas blue color indicates healthy controls. The darker-colored cells indicate higher concentrations of metabolites, and lighter-colored cells represent lower concentration levels for each subject. The intensity of color corresponds to the concentration levels of VOCs.

The differences in the VOC concentrations between the two study groups are presented in the box-plots in Figure 5. Figure 5A shows the levels of breath metabolites found in the gastric cancer patients, whereas Figure 5B shows the concentration levels of the breath

metabolites in the healthy controls. Therefore, the plots showed the levels of breath VOCs that had been selected after the canonical correlation analysis (CCA) in the cancer and control groups, separately. Furthermore, this suggested that the VOCs were significant, as determined by the sensitivity analysis on each group.

() ()

Figure 5. Box-plots showing the VOC concentrations according to study groups: (A) gastric cancer patients, (B) healthy controls. Data are expressed as medians and interquartile range (IQR), minimum–maximum, and extreme values or outliers (p -value < 0.05 is considered as statistically significant).

A table with the descriptive statistics and p -values for the respective VOCs between the groups is presented in Table 2. We found that for most of the VOCs, there were no significant differences. However, two volatile metabolites were significantly different between the cancer and control groups, i.e., 1-octanol (p -value = 0.033) and octane, 1,1'-oxybis- (p -value = 0.027). We also identified that these two metabolites did not overlap between the study groups, even after a sensitivity analysis, and were present exclusively in cancer cases. These results demonstrated that 1-octanol and octane, 1,1'-oxybis- could be potential breath biomarkers for gastric cancer. Importantly, studies had previously associated these VOCs in exhaled breath with a number of bacteria-related metabolic pathways associated with cancer [4,28–30].

Table 2. Descriptive analysis of the VOC expression/level between cancer and control groups (p -value < 0.05 is considered significant).

VOCs	All	Cancer	Control	p -Value (Cancer vs. Control)
	Mean \pm Standard Deviation	Mean \pm Standard Deviation	Mean \pm Standard Deviation	
1-Decanol	374,885.551 \pm 658,847.43	353,430.063 \pm 615,367.78	385,288.212 \pm 687,948.44	0.848
1-Dodecene	230,929.612 \pm 551,407.085	123,288.563 \pm 360,970.491	283,119.212 \pm 621,606.394	0.356
1-Octanol	986,497.776 \pm 1.123 $\times 10^6$	530,659.625 \pm 525,202.942	1.208 $\times 10^6$ \pm 1.269 $\times 10^6$	0.033
2-Pentanone	1.225 $\times 10^6$ \pm 876,447.236	911,821.563 \pm 617,815.055	1.377 $\times 10^6$ \pm 948,989.101	0.078
2-Propanol, 1-(2-methoxy-1-methylethoxy)-	208,490.98 \pm 522,305.902	181,941.5 \pm 354,782.719	221,363.455 \pm 591,337.416	0.701

Table 2. Cont.

VOCs	All	Cancer	Control	<i>p</i> -Value (Cancer vs. Control)
	Mean \pm Standard Deviation	Mean \pm Standard Deviation	Mean \pm Standard Deviation	
3-Methylcyclopentyl acetate	853,658.306 \pm 1.142 $\times 10^6$	665,181.875 \pm 581,937.207	945,040.818 \pm 1.331 $\times 10^6$	0.923
4-Ethylbenzamide	816,527.918 \pm 1.398 $\times 10^6$	621,236.125 \pm 1.211 $\times 10^6$	911,214.848 \pm 1.488 $\times 10^6$	0.565
Benzene	2.426 $\times 10^6$ \pm 4.889 $\times 10^6$	2.240 $\times 10^6$ \pm 3.783 $\times 10^6$	2.516 $\times 10^6$ \pm 5.397 $\times 10^6$	0.677
Benzene, 1-methyl-3-(1-methylethyl)-	422,211.327 \pm 1.205 $\times 10^6$	138,661.938 \pm 390,489.767	559,689.818 \pm 1.430 $\times 10^6$	0.237
Benzyl alcohol	1.775 $\times 10^6$ \pm 3.207 $\times 10^6$	1.700 $\times 10^6$ \pm 2.987 $\times 10^6$	1.811 $\times 10^6$ \pm 3.352 $\times 10^6$	0.662
Decane	1.089 $\times 10^6$ \pm 1.918 $\times 10^6$	1.155 $\times 10^6$ \pm 2.254 $\times 10^6$	1.057 $\times 10^6$ \pm 1.770 $\times 10^6$	0.658
Dimethyl disulfide	105,279.388 \pm 257,353.642	102,375.75 \pm 143,024.5	106,687.212 \pm 299,585.267	0.342
Dodecane, 5,8-diethyl-	630,501.918 \pm 1.054 $\times 10^6$	423,613.313 \pm 912,549.372	730,811.545 \pm 1.115 $\times 10^6$	0.321
Ethanol, 2-phenoxy-	1.712 $\times 10^6$ \pm 3.064 $\times 10^6$	2.214 $\times 10^6$ \pm 4.897 $\times 10^6$	1.469 $\times 10^6$ \pm 1.628 $\times 10^6$	0.654
Methacrolein	239,398.347 \pm 420,052.987	204,225.125 \pm 271,535.237	256,452.03 \pm 478,734.464	0.862
Nonane	1.420 $\times 10^6$ \pm 1.805 $\times 10^6$	1.203 $\times 10^6$ \pm 1.168 $\times 10^6$	1.525 $\times 10^6$ \pm 2.052 $\times 10^6$	0.509
Octane, 1,1'-oxybis-	1.957 $\times 10^6$ \pm 6.918 $\times 10^6$	121,429.375 \pm 337,233.497	2.848 $\times 10^6$ \pm 8.320 $\times 10^6$	0.027
Phenol, 2-methyl-4-(1,1,3,3-tetramethylbutyl)-	3.848 $\times 10^6$ \pm 3.645 $\times 10^6$	3.574 $\times 10^6$ \pm 2.803 $\times 10^6$	3.980 $\times 10^6$ \pm 4.024 $\times 10^6$	0.966
Toluene	4.418 $\times 10^6$ \pm 7.580 $\times 10^6$	3.467 $\times 10^6$ \pm 6.208 $\times 10^6$	4.879 $\times 10^6$ \pm 8.213 $\times 10^6$	0.071
dl-Erythro-1-phenyl-1,2-propanediol	120,533.224 \pm 325,399.861	124,073.063 \pm 301,212.871	118,816.939 \pm 341,011.718	0.661
n-Hexane	286,432.816 \pm 842,786.914	105,393.813 \pm 225,226.503	374,209.303 \pm 1.009 $\times 10^6$	0.233

2.4. Canonical Correlation Analysis of VOCs and Microbiome Data

To reduce the dimensionality of the dataset and obtain a more similar number of features and samples, the feature space of both the breath-VOC and microbiome data subsets were transformed using a sparse principal component analysis (S-PCA).

To detect the possible relationships between the breath VOCs and the fecal microbiota specific to gastric cancer, separate analyses of the data related to the cancer patients and the data of the control group were carried out. The CCA analysis of the S-PCA transformed data showed a statistically significant correlation between the components (a correlation of 0.891, *p*-value = 0.045, in the cancer data subset and a correlation of 0.871, *p*-value = 0.0007, in the control data subset). Figure 6 depicts the resulting CCA score plots of the cancer patients (Figure 6A) and healthy controls (Figure 6B). In the figures, the *x*-axis represents the canonical variate for the breath VOCs, whereas the *y*-axis represents the canonical variate for the fecal microbiome data. Each point corresponds to integrated data from the exhaled breath and the fecal samples collected from the same participants.

After a sensitivity analysis by backwards elimination of least-contributing VOCs and bacteria (i.e., the removal of VOCs/bacteria that did not have a significant impact on the correlation and its statistical significance), there were 14 VOCs and 33 species of bacteria remaining in the cancer subset and 7 VOCs and 17 species of bacteria in the control subset. Similarly, we applied a threshold of ≥ 0.5 for the correlations in both datasets to obtain the final and most important exhaled VOCs and fecal bacterial taxa.

The analyses of the VOCs and fecal bacteria were continued by analyzing the significant correlations ≥ 0.5 and by investigating their biological relevance in the available literature, as well as in the metagenome sequencing results. The VOCs were evaluated according to the pathways of the bacteria with which they were correlated in order to determine whether they could explain the observations related to the specific VOCs and their potential roles in cancer.

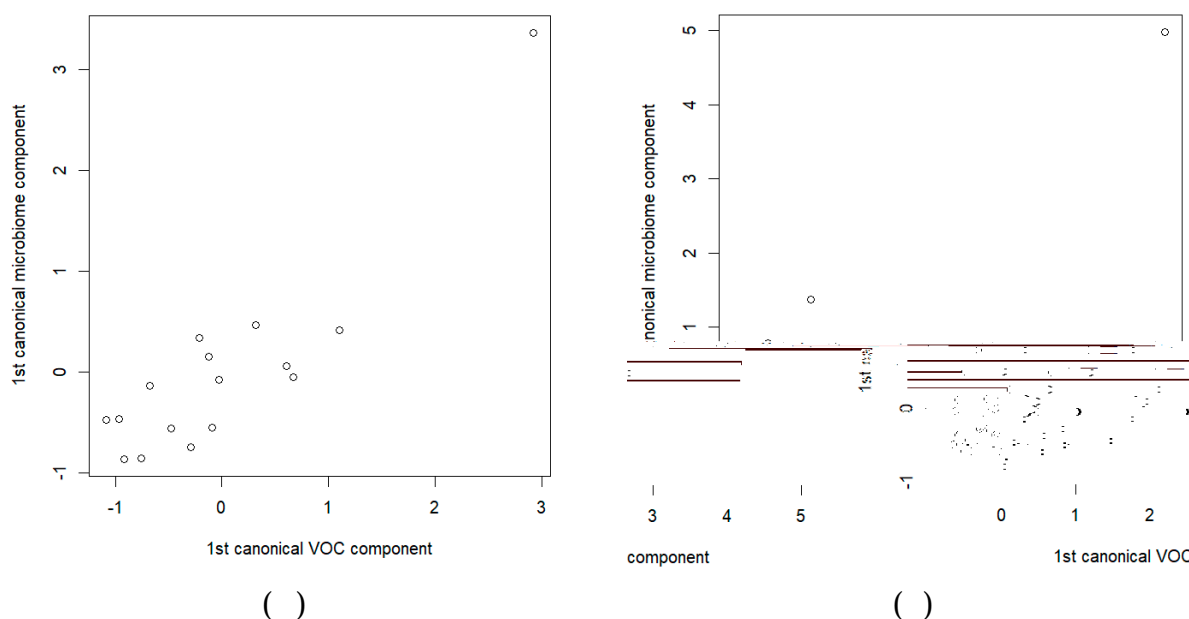


Figure 6. Canonical correlation score plot between the selected breath VOCs (the *x*-axis) and fecal microbiome composition (the *y*-axis) using the first canonical components for (A) cancer patients; (B) healthy controls. CCAs were performed on the sets of 14 and 7 breath VOCs, and 33 and 17 bacterial species, of the cancer patients and healthy controls, respectively.

The chemical identities of the 21 exhaled volatile metabolites that had the highest contribution to the overall correlation, found in both the cancer and control groups, are shown in Table 2. We found that none of the volatile metabolites were common between cancer and control groups, whereas only one fecal bacterial species (i.e., *Clostridium_SGB6179*) overlapped between the two groups.

The correlations between the most significant volatile components and the detected species of bacteria in the two study groups are represented in Figure 7 (cancer cases) and Figure 8 (controls), by correlation heat maps. Interestingly, in the cancer subset, we observed a higher number of bacterial species and exhaled volatile metabolites, as compared to those in the control subset. This outcome possibly indicated distinct microbiome compositions and subsequent breath metabolites across both disease and healthy states.

In Figure 7, an exhaled-breath-matrix dataset based on the fecal bacteria taxa in gastric cancer patients is shown. Based on the correlation heatmap, it was apparent that a vast majority of the individual correlations between the breath VOCs and the fecal bacteria were positive, which was indicated by the trend of lighter yellow, as compared to the other VOC columns. Of the 14 preselected breath VOCs in the 16 cancer patients, we observed that the VOCs, such as 2-propanol, 1-(2-methoxy-1-methylethoxy)-; 5,8-diethyl-dodecane; octane, 1,1'-oxybis-; 1-octanol; 1-dodecene; n-hexane; and n-nonane, in particular, were highly correlated with the abundance of microorganisms. Similarly, the levels of benzene; toluene; n-decane; and benzene, 1-methyl-3-(1-methylethyl)- in the breath correlated with the abundance of the *Bacteroides xylanisolvens* and *Gemmiger formicilis* bacteria. In addition, 2-pentanone, a ketone, was clearly linked to the *Dorea formicigenerans* and *Clostridiaceae bacterium_OM08_6BH* strains. We observed dimethyl disulfide correlated to the *Dorea longicatena*, *G. formicilis*, *D. formicigenerans*, *Clostridium_sp_AM22_11AC*, and *Oscillospiraceae bacterium_NSJ_64* species. Furthermore, ethanol, 2-phenoxy- in the exhaled breath of cancer patients was related to *Fusicatenibacter saccharivorans*, *Oscillibacter_sp_ER4*, *Clostridium_sp_AF20_17LB*, *Bilophila wadsworthia*, *Bittarella massiliensis*, *Blautia faecicola*, and *Clostridium_sp_AM22_11AC*.

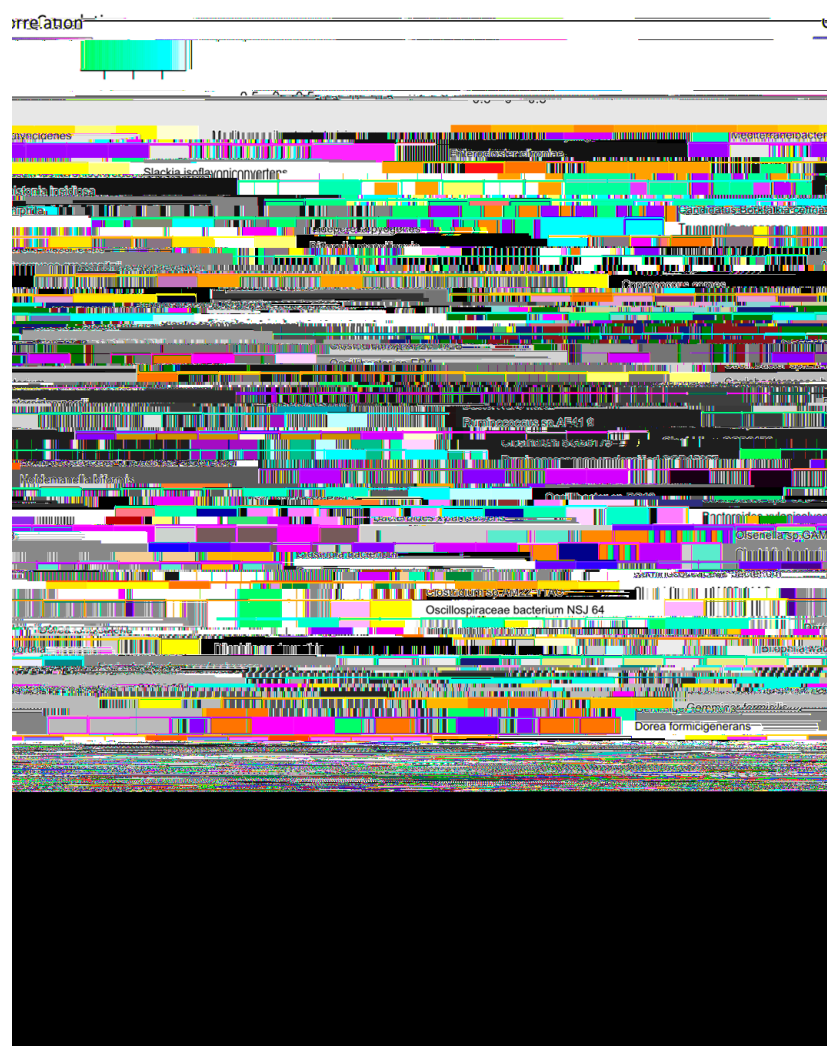


Figure 7. Correlation heatmap of the 14 VOCs and 33 fecal bacterial species in cancer patients. Red color indicates negative correlations, and white or light yellow color indicates positive correlations. The quantitative color scheme is shown at the top.

Furthermore, after comparing the rows, different bacterial species were significantly different from each other. Bacteria species, such as *B. xylanisolvens*, *Ruminococcaceae unclassified_SGB15257*, *Holdemanella biformis*, and *Olsenella_sp_GAM18*, trended towards the white color or had strong positive correlations with a majority of the breath volatile metabolites, as compared to the other bacteria.

The correlation heatmap for the breath VOCs and the fecal bacterial taxa that contributed most to the correlations in the control group is shown in Figure 8.

Interestingly, we found statistically significant correlations between the individual breath VOCs and fecal bacteria. Of the 7 preselected VOCs from the 33 healthy controls, benzyl alcohol; dl-erythro-1-phenyl-1,2-propanediol; 1-decanol; methacrolein; and 3-methylcyclopentyl acetate were strongly correlated with the abundance of most of the bacterial species under observation. Likewise, except for *Enterocloster aldensis* and *Actinomyces sp_ICM47*, all other bacterial species exhibited strong correlations with the levels of VOCs previously mentioned.

The correlation tables showing the magnitudes of the correlation coefficients between the breath metabolites and fecal bacterial taxa in the cancer and control subsets are provided in Tables S1 and S2, respectively, in the Supplementary Materials.

Figure 8. Correlation heatmap of the 7 VOCs and 17 fecal bacterial species in healthy controls. Red color indicates negative correlations, and white or light yellow color indicates positive correlations. The quantitative color scheme is shown at the top.

3. Discussion

The present study aimed to investigate the possible correlations between specific breath VOCs, as identified by GC–MS, and the gut microbiota, as evaluated by shotgun metagenomics sequencing, in relation to gastric cancer. The exhaled breath samples and the fecal samples from a total of 49 participants (16 cancer, 33 controls) were analyzed. We hypothesized that the breath VOCs would be related to the fecal microbiota composition, directly or indirectly, and their presence and levels would vary between the disease and healthy states.

Altogether, our results revealed a significant relation between the fecal microbiota and the volatile metabolites in exhaled breath, as they differed between the cancer patients and the healthy controls. The exhaled breath analysis was integrated with the fecal microbiome analysis, and together, they could serve as non-invasive diagnostic-and-screening tools for gastric cancer, as this disease has been associated with changes in the volatile metabolic profiles and dysregulation of the gut microbiota, its composition, and the related metabolic pathways [31,32]. However, the mechanisms underlying the presence of a specific breath VOCs detected in cancer patients, along with links to gut microbiota activities, have not yet been elucidated. Several studies have suggested hypotheses regarding the generation of various VOCs. Typically, these are produced by the gut microbes via enzymatic reactions, fermentation, and normal metabolism, giving rise to SCFAs, ketones, alcohols, phenols, branched-chain fatty acids (BCFAs), sulfides, bile acids, branched hydrocarbons, and aldehydes, amongst others [33–35].

While microorganisms are present throughout the human digestive tract, studying their presence in fecal samples is crucial to understand the composition of fecal microbiota

and their interactions with the host and their contributions to host phenotypes [36]. Although numerous studies have analyzed VOCs in different biological matrices for disease detection [37–39], the influence of fecal microbiota on human breath profiles and their complex interactions had not been previously investigated within the context of gastric cancer detection. In the current study, we analyzed the correlations between biologically relevant breath VOCs and gastric cancer-associated fecal microbiota for the first time in patients with gastric cancer, as compared to healthy controls. Interestingly, we discovered significant correlations between the profiles of breath VOCs and fecal microbes in the cancer and control subjects under study. It was evident that a majority of the exhaled breath VOCs could have been related to the fecal microbiota composition and abundance, for example, through the microbial metabolism of the host metabolites. Furthermore, the volatile metabolites in the exhaled air reflected the state of the gut homeostasis and the environmental alterations to the structure of the fecal microbiome community [40].

Recent studies have shown that the concentrations of VOCs in the breath and the abundance of fecal microbes differed according to the state of gastrointestinal diseases. Smolinska et al. highlighted a strong link between the specific VOCs in exhaled breath and the specific fecal bacterial strains in a study that included a cohort of patients with Crohn's disease, a type of inflammatory bowel disease [41]. The same authors found that there was a significant difference in the breath VOCs and the diversity of the fecal microbiota between the active and inactive stages of Crohn's disease [42] and ulcerative colitis [43]. This indicated that the relation was dependent on the disease stage. Profiling VOCs in exhaled breath to identify potential biomarkers for gastrointestinal diseases has been the focus of numerous studies. Previous research has associated the levels of different VOCs with the changes and the abundances of different fecal microbes [44]. Despite the large variations in the outcomes among studies, breath analyses have provided the most patient-friendly platform and achieved the highest participation rates [31].

In this study, we found distinct breath metabolites and microbiota compositions, as well as distinct correlations, in the analyzed samples of the cancer and control groups. This suggested different underlying metabolic processes or mechanisms could exist between the groups. The CCA analysis showed a slightly stronger correlation in cancer cases (correlation 0.891) than in the healthy controls (correlation 0.871). Similarly, a larger number of breath VOCs and bacterial species were observed in the cancer patients than in the controls, once again attesting to the distinct microbiome compositions and metabolisms associated with each group. Interestingly, the observed VOCs did not overlap between the groups, whereas a single bacterial taxon was present in both; therefore, the correlations found in this study were dependent on the state of disease or health. Long-term dietary patterns could alter microbiota composition and metabolic activity [45]. Herein, we found several ketones, branched hydrocarbons, alkanes, benzenes, sulfides, phenols (protein fermentation), alcohols, unsaturated aldehyde (methacrolein), aromatics, and bacterial fermentation products in the exhaled breath, contributing to the overall correlations with the fecal microbiota. All have been reported to be associated with several gastrointestinal disorders and cancers [46,47]. However, a shared presence among the related individual VOCs in gastric cancer was not found.

Similar to other studies of gastrointestinal disorders, we found statistically significant correlations among specific hydrocarbons (n-decane, n-nonane, n-hexane, 5,8-diethyl-dodecane), one ether (octane, 1,1'-oxybis), and bacteria *B. xylanisolvans* and *H. biformis*, in cancer patients. *B. xylanisolvans* has been linked with high-fiber dietary patterns [48]. The metabolic pathways responsible for the formation of these hydrocarbons remain unclear. Nevertheless, they might emanate from the lipid peroxidation processes or from environmental exposure, and they have been associated with *Helicobacter pylori* and *Mycobacterium avium*. Furthermore, 2-pentanone, commonly reported in the breath and fecal samples in numerous studies, was produced via the α -oxidation of fatty acids or the oxidation of 2-pentanol catalyzed by either alcohol dehydrogenases (ADHs) or cytochrome p450 (CYP2E1) [49,50]. The source of this compound may be the bacterial action in the gut. In our

study, anaerobes *D. formicigenerans* and *Clostridiaceae bacterium_OM08_6BH* of *Firmicutes* revealed significant correlations with 2-pentanone. The *D. formicigenerans* strain has been associated with obesity-related inflammation.

Prebiotic supplementation with pectin, a dietary fiber, did not have an impact on fecal microbiota composition, SCFAs, or exhaled VOCs in young adults versus elderly subjects [51]. Low-fiber, fat, and high-protein diets were related to an increase in proteolytic fermentation, resulting in shifts in the gut microbiota [52]. In turn, detrimental amino acids and proteolytic fermentation products (e.g., BCFAs) increased in concentration, while beneficial SCFAs were lowered. Host glycans and mucins can be degraded by adverse gut microbes in the intestinal lumen, which maintains the gut barrier integrity, the impairment of which may give rise to enhanced terpenes, aldehydes, and ketones. Furthermore, this caused low-grade systematic oxidative stress, and the microorganisms generated metabolites related to the tricarboxylic acid (TCA) cycle and the Warburg metabolism [31,53].

DMDS, a sulfur-containing compound, was most likely related to the proteolytic fermentation by human fecal microbiota and is an important breath biomarker [54]. Indeed, enhanced levels of sulfides were observed in fecal microbiome and VOC-related studies. DMDS could be synthesized by different microbes in the intestinal tract via common biochemical pathways [55]. The dominant species in the human gut, such as *D. longicatena*, *D. formicigenerans*, *G. formicilis*, and *Clostridium_sp*, showed a strong positive correlation to DMDS. However, the relation of DMDS to *G. formicilis* had not been demonstrated previously. Complex benzenes and toluene, as volatile aromatic hydrocarbons, were reported in higher concentrations in human breath, urine, and blood, and have been commonly attributed to environmental exposure, smoking, and pathogens, such as *H. pylori*, *Pseudomonas aeruginosa*, and *M. avium* [56,57]. In our study, 1-methyl-3-(1-methylethyl)-benzene was significantly correlated to *B. xylanisolvens*, whereas benzene and toluene were sparsely correlated to the *G. formicilis* species. *B. xylanisolvens* was shown to degrade polysaccharides in the gut and produce antibodies (IgM) against the carcinoma-associated Thomsen–Friedenreich (TF) antigen, in a dose- and time-dependent manner [58].

Several primary and branched alcohols were found to be present in both the cancer and control groups, in varying concentrations. They could be produced via the microbial fermentation of alcohols, including from methanol to heptanol [59]. Bacteria, such as *Pseudomonadota* (formerly *Proteobacteria*), generated alcohols, including *Ruminococcus*, *Prevotella*, and *Bifidobacterium*, among others, that could convert alcohols into aldehydes and SCFAs [60]. It is important to note that the alcohols in exhaled breath could be attributed to confounders, such as alcohol consumption. Furthermore, Phenol, 2-methyl-4-(1,1,3,3-tetramethylbutyl)- found in the control groups was linked to *E. aldensis*, a strain of *Lachnospiraceae*, which was metabolically less active and could be responsible for producing certain alcohols and SCFAs. However, the relation of Phenol, 2-methyl-4-(1,1,3,3-tetramethylbutyl)- to *E. aldensis* had not been investigated previously. We also found statistically significant correlations between the species *Candidatus Schneewindia gallinarum* and *Pseudoflavonifractor SGB15156*, and 4-ethylbenzamide (a benzamide derivative), in the healthy controls. This metabolite has been detected in the breath of smokers and emphysema patients, and has often been regarded as a possible inhibitor of *Mycobacterium tuberculosis* [61]. In sharp contrast with these findings, the association between 4-ethylbenzamide and the aforementioned bacterial taxa was not as clear. Similarly, 3-methylcyclopentyl acetate, a methyl ester resulting from the condensation of acetic acid and methanol, has been associated with Crohn's disease and ulcerative colitis in the literature [62]. In the current study, 3-methylcyclopentyl acetate was correlated significantly to 11 of the 17 selected bacterial taxa. However, these correlations did not necessarily indicate that these bacteria had synthesized this methyl ester, since it could have originated from other alternative sources, such as environmental exposure or diet. Methacrolein, an unsaturated aldehyde and a known carcinogen, was ubiquitously correlated to the majority of the bacterial species (12 out of 17). It has been associated with lung injury and acute respiratory failure [63]. In other studies, it was reported as

an oxidation product of isoprene, which has received widespread attention as a breath biomarker in several diseases, including gastric cancer.

Consequently, several classes of VOCs could be identified in exhaled breath associated with the fecal microbiome and related to different metabolic pathways. Including our results of the correlated exhaled breath metabolites in future research with a larger study cohort could further our understanding of the relationship and the origin of the metabolites. In addition to the sample size, a potential limitation in the present study could have been related to the confounding chemicals present in the ambient air, as a possible source of background bias. Similarly, the expression of certain breath VOCs in relation to the correlated bacteria is not yet fully understood. The presence of outliers in the fecal bacterial communities and their respective VOCs may suggest the presence of some rare and unclassified bacterial taxa.

We did not identify any noteworthy differences in the gut microbiome composition based solely on taxonomic composition between the gastric cancer and control groups that would enable us to distinguish between the two. This lack of differentiation was not unusual, given the limitations of the sample size, the variability of the tumor types, and the clinical stages represented in this study, as well as the restraints of microbiome DNA sequencing. While this method was not entirely quantitative, the presence of DNA alone did not necessarily indicate the production of a particular metabolite. To obtain more precise information about the processes occurring in the gut, one would need to use other approaches, such as RNA sequencing and metabolomics. Alternatively, breath analysis of VOCs has the potential to provide a less expensive, more direct, and non-invasive measure of the microbiome's status, which would be a more practical approach in clinical settings where rapid and cost-effective testing is necessary.

In line with our findings, the established link between exhaled metabolites and the fecal microbiota could provide information on the origin of the volatile compounds in the breath, indicate the state of health, and assess recent or long-term exposure to various contaminants. Additionally, this could generate new knowledge on the mechanisms and potential of breath VOCs for gastric cancer screening in exhaled breath analysis.

4. Materials and Methods

4.1. Study Participants

This study included a total of 49 individuals, aged 24–77 years, recruited at the Riga East University Hospital, Latvia Oncology Center (Riga, Latvia), who underwent the analyses. Among them, 16 were diagnosed with gastric cancer, and 33 were healthy controls. The patient group had different clinical stages of cancer, ranging from stage IA to stage IV [64]. The study was approved by the Ethics Committee of the Riga East University Hospital Support Foundation (approval No. 18A/19), and all patients provided written informed consent before participation.

The clinical data, breath, and fecal samples were collected during the hospital visit. Breath samples were collected in a dedicated room within the oncology department for the GC–MS analysis. The fecal samples from the same participants were taken for microbiome analysis.

4.2. Breath Sampling

Exhaled breath samples were taken randomly at room temperature, in a room free of potential sources of VOCs, such as food, kitchen waste, drugs, solvents, and disinfectants, to prevent the risk of background bias. All participants were exposed to the same test environment, temperature, and humidity conditions, without any medical intervention. Prior to the sampling, the study subjects were advised to:

1. Fast for at least 12 h;
2. Refrain from drinking coffee, tea, and soft drinks, for at least 12 h;
3. Refrain from smoking for at least two hours;
4. Avoid alcohol intake for at least 24 h;

5. Refrain from tooth-brushing, mouthwash, or flossing;
6. Avoid chewing gum and mouth fresheners for at least 12 h;
7. Refrain from using cosmetics/fragrances prior to the procedure on the day of the test;
8. Avoid excessive physical activity (gym, jogging, cycling, and intense physical work) at least two hours prior to the test.

Prior to the breath sample collection, the medical staff interviewed the study subjects using a clinical questionnaire containing questions on personal data, VOC exposure, confounding factors, and 24 h dietary recall. Breath samples were taken using an in-house custom-made breath sampler. The breath sampler consisted of a disposable mouthpiece (Intersurgical, Wokingham, UK, 1931000) installed on a disposable elbow (Intersurgical, Wokingham, UK, 2714000S), and the CO₂ sensor cell (Masimo, Irvine, CA, USA) was connected to the other end of the elbow. The elbow was equipped with a $\frac{1}{4}$ " port, enabling the connection of industry standard $\frac{1}{4}$ " Tenax TA sorbent tubes. Prior to the sample collection, the sampling end of the sorbent tube was inserted into the elbow, so it protruded 5–6 mm into the interior of the elbow and was fixed with a $\frac{1}{4}$ " PTFE nut. The other end of the sampling tube was connected to a 250 mL gas-tight glass syringe (Roth, Karlsruhe, Germany) using a $\frac{1}{8}$ " Teflon tube. The subject could freely inhale/exhale through a mouthpiece without pneumatic resistance. The sampling was achieved manually by drawing a volume of 500 mL during subsequent end-tidal exhalation segments, as identified via the CO₂ measurement. On average, 10–15 mL of breath was drawn per exhalation. Effectively, two breath samples (each 500 mL) were taken per subject. In parallel to the breath samples, one air sample of the room was collected via a manual draw of 500 mL of ambient air into a Tenax sorbent tube. Immediately after sampling, both ends of the sorbent tube were closed with brass $\frac{1}{4}$ " nuts, and the tubes were frozen at -80 °C. Samples were stored at -80 °C and transported on dry ice. An effort was made to reduce the storage time to 4 weeks. The volatile metabolites were measured by GC–MS.

4.3. Chemicals and Standards for GC–MS Analysis

All standard mixtures were produced using high-purity liquid substances. Reference chemicals with stated purities of 95–99.9% were purchased from Merck (Vienna, Austria) and Fluka (Buchs, Switzerland). Standards were prepared by injecting and evaporating several microliters of the liquid compound into evacuated and heated 1 L glass bulbs (Supelco, Bellefonte, PA, USA). Next, primary standards were diluted by transferring appropriate volumes from the bulb mixtures into 3–10 L Tedlar bags (SKC Inc., Eighty Four, PA, USA), that were pre-filled with purified and humidified air (Relative Humidity 100% at 34 °C). The stainless-steel industry-standard thermal desorption tubes ($\frac{1}{4}$ in. O.D., $3\frac{1}{2}$ inches long), pre-packed with Tenax TA, were purchased from Markes International Limited (Bridgend, UK). Prior to each sampling, sorbent tubes were reconditioned according to the manufacturer's recommendations. The calibration protocols were outlined elsewhere [65].

4.4. GC–MS Analysis of Exhaled Breath

A two-stage thermal desorption of breath VOCs was performed using a commercial thermal desorber (UNITY, Markes International Limited, Bridgend, UK) and an autosampler (TD100, Markes International Limited, Bridgend, UK). During the primary desorption, sorbent tubes were heated to 280 °C for 6 min under a flow rate of 6.0 (99.9999%) helium at 20 mL/min. VOCs desorbed during this phase were refocused into a cold trap packed with graphitized carbon-black and maintained at 5 °C. The injection of VOCs into the capillary column was performed via the rapid heating of the cold trap to 320 °C.

and up to 280 °C; and finally remaining at 280 °C for 5 min. The untargeted VOC analysis was performed using the mass spectrometer working in SCAN mode with the associated m/z ranging from 20 up to 250. The peak integration was based on the extracted m/z ratio chromatograms, and this approach allowed for a proper separation of the majority of the peaks of interest from their neighbors. The quadrupole, ion source, and transfer line were kept at 150 °C, 230 °C, and 280 °C, respectively.

Identification of breath VOCs was performed in two steps. Firstly, the peak spectrum was evaluated against the National Institute of Standards and Technology (NIST) mass spectral library database. Next, the NIST identification was confirmed by comparing the respective retention times with the retention times obtained for the reference standards prepared from pure compounds, as outlined previously.

4.5. DNA Isolation of Fecal Samples

The study utilized sample collection bottles (Eiken Chemical Co., Ltd., Tokyo, Japan) specifically designed for fecal immunochemical occult blood testing. These bottles were used to collect, transport, and store the fecal samples. After conducting a quantitative immunochemical fecal occult blood test, all the returned bottles were frozen at −80 °C and kept frozen during the transport and storage, until they were ready for further analysis for the microbiome research. The samples were processed within 60 days from the time of collection. Prior to DNA extraction, the entirety of the buffer content containing the fecal matter (~1.5 mL) was collected from the sample bottles using a sterile, disposable syringe and then concentrated through lyophilization. The resulting lyophilized sample was used to extract microbial DNA using the MagPure stool DNA LQ (MGI Tech, Shenzhen, China) and FastPrep 24 5G instrument. Additionally, DNA was extracted from a negative control.

4.6. Shotgun Metagenomic Sequencing and Data Handling

MGI Tech Latvia laboratories were responsible for the sequencing process. The extracted DNA was fragmented into roughly 400 bp fragments, and a sequencing library was created using the MGIEasy PCR-Free DNA (MGI Tech, Shenzhen, China) library preparation protocol, in accordance with the manufacturer's standard procedure. The DNBSEQ-T7 platform (MGI Tech, Shenzhen, China) was used to perform shotgun metagenomic sequencing with 150 bp paired-end reads.

The primary processing of the raw metagenomics sequencing data was conducted on high-performance computing clusters at Riga Technical University (Riga, Latvia) using an in-house Snakemake-based workflow. Adapter removal and base quality trimming was conducted with fastp tool. The host reads were extracted with Kraken2 (version 2.1.2) using the GRCh38 human reference genome. The taxonomic assignment was carried out using the MetaPhlAn4 (version 4.0.3) and vJan21_CHOCOPhAnSGB_202103 marker database, along with Kraken2, against a database built from all available bacteria full genomes in RefSeq (accessed July 2022). A confidence value of 0.2 of the Kraken2 results was re-estimated at species level with Bracken (version 2.8) in order to improve the taxonomic abundances. The functional annotation was performed using HUMAnN 3 (version 3.6) with a full UniRef90 database (accessed December 2022), and the results were normalized at all levels by community and expressed in copies-per-million.

The results were summarized and further transformed by in-house python scripts and analyzed in an R software environment (version 4.0.5.).

4.7. Statistical Analysis

The statistical analysis to evaluate the connection between the VOCs from exhaled breath and the bacteria species from the fecal microbiota was carried out using CCA. The CCA found the relationship between the two multivariate sets of variables that characterized the same study participants by creating linear combinations of variables from each set (in this study, one set of combinations described the breath VOCs, and the other, the microbiome composition), so they would have a maximum correlation between them [66,67].

The CCA score plot was used to visualize the relationships (correlations) between the VOCs found in the breath and the composition of the fecal microbiota.

The most significant breath VOCs and fecal microbiome species were identified using sensitivity analyses by backward elimination of the VOCs and bacteria, and then evaluating the changes in significance of correlation among the canonical components. The least significant VOCs and bacterial species had little impact on the correlation when removed. The final set of breath VOCs and bacteria species included only the significant markers, whose removal would lead to a significant drop in correlational significance.

Since the counts of unique VOCs and unique bacterial species were significantly larger than the number of study participants, the dimensionality was reduced using S-PCA [68].

The analysis and visualization were conducted using RStudio 2022.02.3 (Build 492) and the R libraries CCA, CCP, and sparsepca. A p -value < 0.05 was deemed to indicate statistical significance.

5. Conclusions

Overall, this study investigated the significant and informative correlations between breath volatile metabolites and fecal microbiota compositions. The production of microbe-associated breath VOCs could affect the disease pathology. Using GC–MS for the VOC measurements served to complement the fecal microbiome analysis using metagenomic sequencing, in addition to an integrated approach involving a multivariate statistical analysis and bioinformatics. The proposed integrated analyses could be beneficial for a better understanding of the underlying biological processes of microbiota composition and breath VOCs. Moreover, this prospective study offers potential directions for advanced screening and developing microbiome-directed novel therapeutic targets and biomarkers for gastric cancer. Our findings further highlighted how breath VOCs were an excellent medium for the investigation of microbial metabolic functions and promising biomarker candidates that could eventually deliver on the promise of non-invasive cancer screening.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules28083488/s1>, Table S1: Correlation coefficients (breath VOCs vs. fecal bacterial taxa) for gastric cancer patients, Table S2: Correlation coefficients (breath VOCs vs. fecal bacterial taxa) for healthy controls.

Author Contributions: Conceptualization, M.P.B., I.P. and M.L.; methodology, M.P.B., I.P., R.V. and P.M.; formal analysis, M.P.B., I.P., R.V. and P.M.; investigation, M.P.B., R.V., L.M., V.V. and P.M.; resources, I.P. and M.L.; data curation, M.P.B., I.P. and R.V.; writing—original draft preparation, M.P.B., I.P. and R.V.; writing—review and editing, L.M., V.V., A.K., P.M., E.D.-N. and M.L.; supervision, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: The development of the analysis approach and its evaluation and analysis were supported by a postdoctoral grant within the Activity 1.1.1.2 “Post-doctoral Research Aid” co-funded by the European Regional Development Fund (postdoctoral project number: 1.1.1.2/VIAA/3/19/495). L.M. was supported by the project ESF SAM 8.2.2.3. “Strengthening the doctoral capacity of the University of Latvia within the framework of the new doctoral model”, (project No. 8.2.2.0/20/1/006). E.D.-N. acknowledges the support received from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP—grant 2014/26897-0).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki. The study protocol was approved by the Research Ethics Committee of Riga East Clinical University Hospital (Riga, Latvia) on 5 December 2019, registration No. 18A/19.

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

Data Availability Statement: The data presented in this study are available upon request from the corresponding authors.

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

Sample Availability: Not applicable.

References

- 1.

27. Drabińska, N.; Flynn, C.; Ratcliffe, N.; Belluomo, I.; Myridakis, A.; Gould, O.; Fois, M.; Smart, A.; Devine, T.; Costello, B.D.L. A literature survey of all volatiles from healthy human breath and bodily fluids: The human volatilome. *J. Breath Res.* **2021**, *15*, 34001. [\[CrossRef\]](#)
28. Kneepkens, C.M.; Lepage, G.; Roy, C.C. The potential of the hydrocarbon breath test as a measure of lipid peroxidation. *Free Radic. Biol. Med.* **1994**, *17*, 127–160. [\[CrossRef\]](#)
29. Ratcliffe, N.; Wiczorek, T.; Drabińska, N.; Gould, O.; Osborne, A.; De Lacy Costello, B. A mechanistic study and review of volatile products from peroxidation of unsaturated fatty acids: An aid to understanding the origins of volatile organic compounds from the human body. *J. Breath Res.* **2020**, *14*, 034001. [\[CrossRef\]](#)
30. Hakim, M.; Broza, Y.Y.; Barash, O.; Peled, N.; Phillips, M.; Amann, A.; Haick, H. Volatile Organic Compounds of Lung Cancer and Possible Biochemical Pathways. *Chem. Rev.* **2012**, *112*, 5949–5966. [\[CrossRef\]](#)
31. van Vorstenbosch, R.; Cheng, H.R.; Jonkers, D.; Penders, J.; Schoon, E.; Masclee, A.; van Schooten, F.-J.; Smolinska, A.; Mujagic, Z. Systematic Review: Contribution of the Gut Microbiome to the Volatile Metabolic Fingerprint of Colorectal Neoplasia. *Metabolites* **2023**, *13*, 55. [\[CrossRef\]](#)
32. Liang, W.; Yang, Y.; Wang, H.; Wang, H.; Yu, X.; Lu, Y.; Shen, S.; Teng, L. Gut microbiota shifts in patients with gastric cancer in perioperative period. *Medicine* **2019**, *98*, e16626. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Visconti, A.; Le Roy, C.I.; Rosa, F.; Rossi, N.; Martin, T.C.; Mohny, R.P.; Li, W.; de Rinaldis, E.; Bell, J.T.; Venter, J.C.; et al. Interplay between the human gut microbiome and host metabolism. *Nat. Commun.* **2019**, *10*, 4505. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Pereira, J.; Porto-Figueira, P.; Cavaco, C.; Taunk, K.; Rapole, S.; Dhakne, R.; Nagarajaram, H.; Câmara, J.S. Breath Analysis as a Potential and Non-Invasive Frontier in Disease Diagnosis: An Overview. *Metabolites* **2015**, *5*, 3–55. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Lubes, G.; Goodarzi, M. GC-MS based metabolomics used for the identification of cancer volatile organic compounds as biomarkers. *J. Pharmaceut. Biomed. Anal.* **2018**, *147*, 313–322. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Lind, A.L.; Pollard, K.S. Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. *Microbiome* **2021**, *9*, 58. [\[CrossRef\]](#)
37. Amann, A.; Costello, B.D.L.; Miekisch, W.; Schubert, J.; Buszewski, B.; Pleil, J.; Ratcliffe, N.; Risby, T. The human volatilome: Volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J. Breath Res.* **2014**, *8*, 034001. [\[CrossRef\]](#)
38. Sharma, A.; Kumar, R.; Varadwaj, P. Smelling the Disease: Diagnostic Potential of Breath Analysis. *Mol. Diagn. Ther.* **2023**, 1–27. [\[CrossRef\]](#)
39. Gasparri, R.; Capuano, R.; Guaglio, A.; Caminiti, V.; Canini, F.; Catini, A.; Sedda, G.; Paolesse, R.; Di Natale, C.; Spaggiari, L. Volatolomic urinary profile analysis for diagnosis of the early stage of lung cancer. *J. Breath Res.* **2022**, *16*, 046008. [\[CrossRef\]](#)
40. Du, X.; Li, Q.; Tang, Z.; Yan, L.; Zhang, L.; Zheng, Q.; Zeng, X.; Chen, G.; Yue, H.; Li, J.; et al. Alterations of the Gut Microbiome and Fecal Metabolome in Colorectal Cancer: Implication of Intestinal Metabolism for Tumorigenesis. *Front. Physiol.* **2022**, *13*, 854545. [\[CrossRef\]](#)
41. Smolinska, A.; Tedjo, D.I.; Blanchet, L.; Bodelier, A.; Pierik, M.J.; Masclee, A.A.M.; Dallinga, J.; Savelkoul, P.H.M.; Jonkers, D.M.A.E.; Penders, J.; et al. Volatile metabolites in breath strongly correlate with gut microbiome in CD patients. *Anal. Chim. Acta* **2018**, *1025*, 1–11. [\[CrossRef\]](#)
42. Tedjo, D.I.; Smolinska, A.; Savelkoul, P.H.; Masclee, A.A.; Van Schooten, F.J.; Pierik, M.J.; Penders, J.; Jonkers, D.M.A.E. The fecal microbiota as a biomarker for disease activity in Crohn's disease. *Sci. Rep.* **2016**, *6*, 35216. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Smolinska, A.; Bodelier, A.G.; Dallinga, J.W.; Masclee, A.A.; Jonkers, D.M.; van Schooten, F.J.; Pierik, M.J. The potential of volatile organic compounds for the detection of active disease in patients with ulcerative colitis. *Aliment. Pharmacol. Ther.* **2017**, *45*, 1244–1254. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Belizário, J.E.; Faintuch, J.; Malpartida, M.G. Breath Biopsy and Discovery of Exclusive Volatile Organic Compounds for Diagnosis of Infectious Diseases. *Front. Cell. Infect. Microbiol.* **2021**, *10*, 564194. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Hills, R.D.; Pontefract, B.A.; Mishcon, H.R.; Black, C.A.; Sutton, S.C.; Theberge, C.R. Gut Microbiome: Profound Implications for Diet and Disease. *Nutrients* **2019**, *11*, 1613. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Altomare, D.F.; Picciariello, A.; Rotelli, M.T.; De Fazio, M.; Aresta, A.; Zambonin, C.G.; Vincenti, L.; Trerotoli, P.; De Vietro, N. Chemical signature of colorectal cancer: Case-control study for profiling the breath print. *BJS Open* **2020**, *4*, 1189–1199. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Xiang, L.J.; Wu, S.H.; Hua, Q.L.; Bao, C.Y.; Liu, H. Volatile organic compounds in human exhaled breath to diagnose gastrointestinal cancer: A meta-analysis. *Front. Oncol.* **2021**, *11*, 269. [\[CrossRef\]](#)
48. Adamberg, K.; Jaagura, M.; Aaspõllu, A.; Nurk, E.; Adamberg, S. The composition of faecal microbiota is related to the amount and variety of dietary fibres. *Int. J. Food Sci. Nutr.* **2020**, *71*, 845–855. [\[CrossRef\]](#)
49. Garner, C.E.; Smith, S.; de Lacy Costello, B.; White, P.; Spencer, R.; Probert, C.S.J.; Ratcliffe, N.M. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *FASEB J.* **2007**, *21*, 1675–1688. [\[CrossRef\]](#)
50. Mochalski, P.; Wiesenhofer, H.; Allers, M.; Zimmermann, S.; Güntner, A.T.; Pineau, N.J.; Lederer, W.; Agapiou, A.; Mayhew, C.A.; Ruzsanyi, V. Monitoring of selected skin- and breath-borne volatile organic compounds emitted from the human body using gas chromatography ion mobility spectrometry (GC-IMS). *J. Chromatogr. B* **2018**, *1076*, 29–34. [\[CrossRef\]](#)

51. An, R.; Wilms, E.; Smolinska, A.; Hermes, G.D.; Masclee, A.A.; de Vos, P.; Schols, H.A.; van Schooten, F.J.; Smidt, H.; Jonkers, D.M. Sugar Beet Pectin Supplementation Did Not Alter Profiles of Fecal Microbiota and Exhaled Breath in Healthy Young Adults and Healthy Elderly. *Nutrients* **2019**, *11*, 2193. [\[CrossRef\]](#)
52. Claesson, M.J.; Jeffery, I.B.; Conde, S.; Power, S.E.; O'Connor, E.M.; Cusack, S.; Harris, H.M.; Coakley, M.; Lakshminarayanan, B.; O'Sullivan, O.; et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature* **2012**, *488*, 178–184. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Huybrechts, I.; Zouliouich, S.; Loobuyck, A.; Vandenbulcke, Z.; Vogtmann, E.; Pisanu, S.; Iguacel, I.; Scalbert, A.; Indave, I.; Smelov, V.; et al. The human microbiome in relation to cancer risk: A systematic review of epidemiologic studies. *Cancer Epidemiol. Biomark. Prev.* **2020**, *29*, 1856–1868. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Moore, J.G.; Jessop, L.D.; Osborne, D.N. Gas-chromatographic and mass-spectrometric analysis of the odor of human feces. *Gastroenterology* **1987**, *93*, 1321–1329. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Filipiak, W.; Sponring, A.; Filipiak, A.; Baur, M.; Ager, C.; Wiesenhofer, H.; Margesin, R.; Nagl, M.; Troppmair, J.; Amann, A. Volatile organic compounds (VOCs) released by pathogenic microorganisms in vitro: Potential breath biomarkers. In *Volatile Biomarkers*; Elsevier: Amsterdam, The Netherlands, 2013; p. 463.
56. Buszewski, B.; Ulanowska, A.; Ligor, T.; Jackowski, M.; Klodzinska, E.; Szeliga, J. Identification of volatile organic compounds secreted from cancer tissues and bacterial cultures. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2008**, *868*, 88–94. [\[CrossRef\]](#)
57. Shestivska, V.; Spaněl, P.; Dryahina, K.; Sovová, K.; Smith, D.; Musílek, M.; Nemec, A. Variability in the concentrations of volatile metabolites emitted by genotypically different strains of *Pseudomonas aeruginosa*. *J. Appl. Microbiol.* **2012**, *113*, 701–713. [\[CrossRef\]](#)
58. Dahiya, D.K.; Renuka, D.A.K.; Shandilya, U.K.; Puniya, A.K.; Shukla, P. Chapter 44—New-Generation Probiotics: Perspectives and Applications. In *Microbiome and Metabolome in Diagnosis, Therapy, and other Strategic Applications*; Faintuch, J., Faintuch, S., Eds.; Academic Press: Cambridge, MA, USA, 2019; pp. 417–424.
59. Lourenço, C.; Turner, C. Breath analysis in disease diagnosis: Methodological considerations and applications. *Metabolites* **2014**, *4*, 465–498. [\[CrossRef\]](#)
60. Tsuruya, A.; Kuwahara, A.; Saito, Y.; Yamaguchi, H.; Tenma, N.; Inai, M.; Takahashi, S.; Tsutsumi, E.; Suwa, Y.; Totsuka, Y.; et al. Major Anaerobic Bacteria Responsible for the Production of Carcinogenic Acetaldehyde from Ethanol in the Colon and Rectum. *Alcohol Alcohol.* **2016**, *51*, 395–401. [\[CrossRef\]](#)
61. Cristescu, S.M.; Gietema, H.A.; Blanchet, L.; Kruitwagen, C.L.J.J.; Munnik, P.; Van Klaveren, R.J.; Lammers, J.W.J.; Buydens, L.; Harren, F.J.M.; Zanen, P. Screening for emphysema via exhaled volatile organic compounds. *J. Breath Res.* **2011**, *5*, 046009. [\[CrossRef\]](#)
62. Walton, C.; Fowler, D.P.; Turner, C.; Jia, W.; Whitehead, R.N.; Griffiths, L.; Dawson, C.; Waring, R.H.; Ramsden, D.B.; Cole, J.A.; et al. Analysis of volatile organic compounds of bacterial origin in chronic gastrointestinal diseases. *Inflamm. Bowel Dis.* **2013**, *19*, 2069–2078. [\[CrossRef\]](#)
63. Meehan-Atrash, J.; Luo, W.; Strongin, R.M. Toxicant formation in dabbing: The terpene story. *ACS Omega* **2017**, *2*, 6112–6117. [\[CrossRef\]](#)
64. Brierley, J.D.; Gospodarowicz, M.K.; Wittekind, C. *TNM Classification of Malignant Tumours*, 8th ed.; John Wiley & Sons: Weinheim, Germany, 2017; pp. 69–71.
65. Mochalski, P.; Leja, M.; Ślęfarska-Wolak, D.; Mezmale, L.; Patsko, V.; Ager, C.; Królicka, A.; Mayhew, C.A.; Shani, G.; Haick, H. Identification of Key Volatile Organic Compounds Released by Gastric Tissues as Potential Non-Invasive Biomarkers for Gastric Cancer. *Diagnostics* **2023**, *13*, 335. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Hotelling, H. Relations between two sets of variates. *Biometrika* **1936**, *28*, 321–377. [\[CrossRef\]](#)
67. Langworthy, B.W.; Stephens, R.L.; Gilmore, J.H.; Fine, J.P. Canonical correlation analysis for elliptical copulas. *J. Multivar. Anal.* **2021**, *183*, 104715. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Guerra-Urzola, R.; Van Deun, K.; Vera, J.C.; Sijtsma, K. A Guide for Sparse PCA: Model Comparison and Applications. *Psychometrika* **2021**, *86*, 893–919. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.