Article

Global Plasma Profiling for Colorectal Cancer-Associated Volatile Organic Compounds: a Proof-of-Principle Study

Seongho Kim^{1,2}, Xinmin Yin³, Md Aminul Islam Prodhan³, Xiang Zhang³, Zichun Zhong⁴, and Ikuko Kato^{1,5,*}

¹Department of Oncology, Wayne State University School of Medicine, Detroit MI 48201, USA, ²Biostatistics Core, Karmanos Cancer Institute, Wayne State University, Detroit MI 48201, USA, ³Department of Chemistry, University of Louisville, Louisville, Kentucky 40208, USA, ⁴Department of Computer Science, College of Engineering, Wayne State University, Detroit MI 48202, USA, and ⁵Department of Pathology, Wayne State University School of Medicine, Detroit MI 48201, USA

*Author to whom correspondence should be addressed. Dr Ikuko Kato, Karmanos Cancer Institute, Wayne State University, 4100 John R Street, MM04EP, Detroit MI 48201, USA. Email katoi@karmanos.org

Received 19 July 2018; Revised 14 December 2018; Editorial Decision 22 January 2019

Abstract

Volatile organic compounds (VOCs) could reflect changes resulting from ongoing pathophysiological processes and altered body metabolisms, and thus have been studied for various types of cancers. We aimed to test an advanced global metabolomic technique to characterize circulating VOCs in patients diagnosed with colorectal cancer (CRC). We employed solid-phase microextraction (SPME) and comprehensive two-dimensional gas chromatography mass-spectrometry (GC × GC–MS). We analyzed 30 random plasma samples from incident cases of CRC. The 30 samples were from population controls enrolled in a large population-based case-control study. The number of metabolite peaks detected in the cases was significantly lower than that detected in the controls (median 1530 vs. 1694, $P = 0.02$). Partial least squares-discriminant analysis showed clear VOC profile differences between the CRC and the controls. After adjustment for multiple comparisons at the 5% false discovery rate level, five VOCs were differentially expressed between the cases and the controls. Among these five VOCs, 2,3,4-trimethyl-hexane (decreased) and 2,4-dimethylhept-1-ene (increased) were both lipid peroxidation products but not previously reported for CRC. In summary, this study pointed to an intriguing observation that the richness of volatile metabolites may be reduced in CRC cases and demonstrated the utility of SPME GC × GC–MS in discovery of candidate markers for further validation.

Introduction

Volatile organic compounds (VOCs) are organic compounds, which have high vapor pressure under room temperature, express changes resulting from ongoing pathophysiological processes and altered body metabolisms, and thus have been studied for various cancers (1, 2). Among more than 3000 VOCs emitted from the human body, approximately a little over 100 VOCs have been designated to be cancer-associated (2–4). Trained dogs in fact can distinguish

scents of cancer patients from those of non-cancer patients (5, 6). Furthermore, some VOCs found in human body fluids represent environmental carcinogens, such as nitrosamines and benzene, to which subjects are exposed (2, 7) and thus are potentially useful for cancer prevention research.

Depending on cancer-affected organs, cancer-associated VOCs can be directly excreted from the affected tissue/organ to feces, urine, saliva, semen, tear and vaginal, nasal and nipple discharges.

VOCs can also enter the blood circulation and then are exhaled in breath, excreted in urine or emitted from the skin (2). Although most commonly used human body matrix for VOC detection has been exhaled breath, which is a logical choice for lung cancer, breath is not a routinely collected clinical specimen. On the other hand, blood tests are used at any physician's office for the adult patient population and also used to monitor cancer patients with various tumor markers (8).

While one-dimensional gas chromatography mass-spectrometry (GC-MS) has been used for analysis of VOCs (9, 10), it has the limited separation power to undermine its utility in detecting low concentration VOCs emitted by a minuscule size of subclinical cancer due to the depredation from the long storage. Comprehensive two-dimensional gas chromatography mass-spectrometry (GC $\times$ GC-MS) uses two GC columns with different stationary phases to increase the separation power (11–14). Compounds eluted from the first-dimension column within a modulation period are collected by a thermal modulator and subjected to the second-dimension column for further separation at an increased temperature. Consequently, the GC $\times$ GC-MS system brings more increased separation power and more accurate information about metabolite retention times and mass spectra than those in a GC-MS system. So that, GC $\times$ GC-MS system is a powerful technique for analysis of metabolites in complex biological systems and particularly suitable for untargeted metabolomic studies.

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in both genders and the second and third leading cause of cancer deaths in males and females, respectively, in the USA (15). However, there have been few tumor markers for CRC, which have shown good diagnostic, predictive or prognostic values. Carcinoembryonic antigen has been routinely used to detect CRC recurrence after surgery, but has insufficient sensitivity to be used alone (16). Thus, new blood-based markers would greatly facilitate CRC prevention and patient care.

The objectives of the present proof-of-principal observational study were to explore the feasibility of using SPME GC $\times$ GC-MS platform to analyze comprehensively circulating VOCs associated with CRC that may warrant further validation. VOCs were first extracted by SPME from plasma samples of CRC ($n = 30$) and normal ($n = 30$), and then analyzed on a GC $\times$ GC-MS using the untargeted metabolomics approach.

Materials and methods

Materials

Sodium phosphate dibasic (Na_2HPO_4 , purity 100.2%, Cat. No. S374-500) and sodium phosphate monobasic (NaH_2PO_4 , purity 99%, Cat. No. 389872500), and sodium hydroxide (NOH, Cat. No. S399-500) were purchased from Fisher Scientific Inc. (Pittsburgh, PA, USA). Dichloromethane (CH_2Cl_2 , purity $\geq 99.8\%$, Cat. No. 34856-1 L) and a mixture containing C_7 – C_{30} n -alkanes (purity 95–99%, Cat. No. 49451U) was purchased from Sigma-Aldrich Corp. (St. Louis, MO, USA). The n -alkane mixture was dissolved in dichloromethane at a concentration of 25 $\mu\text{g/mL}$ per compound. diH_2O was prepared from a Millipore synergy system (Burlington, MA, USA).

Study samples

The samples used subjects for this study were plasma samples collected for a population-based CRC case-control study. It was conducted in Metropolitan Detroit during the period of 2003–2006 and

designed to enroll 1:1 frequency-matched cases and controls. The study was approved by the Institutional Review Board of Wayne State University. Details concerning the eligibility, recruitment and characteristics of the study subjects enrolled in the parent study were described elsewhere (17, 18). Blood was collected in EDTA tubes through home phlebotomy services at subjects' home or other preferred locations and delivered to a processing laboratory with cooling gels where plasma samples were separated, aliquoted and stored in -80°C freezers. Because the freeze-thaw cycle affects the stability of VOC (19–21), we used samples not previously thawed for this study. Information with respect to the clinical and histological characteristics of tumor, vital status (as of 6 July 2016) and cause of death was obtained through linkage with the Metropolitan Detroit Cancer Surveillance System. Dietary and other personal characteristics were obtained by use of structured questionnaires. For this study, from the original pool of 1582 subjects in the parent study, 30 cases and 30 controls were randomly selected, stratified by age ($< / \geq 65$) and smoking (never/ex/current), with over-sampling of cases with no cancer targeted therapies ($n = 11$) (surgery, radiation or chemotherapy).

SPME GC $\times$ GC-MS analysis

A SPME fiber (divinylbenzene/carboxen/polydimethylsiloxane (DVBCAR/PDMS), $d_f = 50/30 \mu\text{m}$, $L = 1 \text{ cm}$) was purchased from Sigma-Aldrich Co. (St. Louis, MO, USA). About 50 μL plasma was added into a 2-mL glass bottle followed by adding 50 μL water ($\text{pH}=7$). To increase the coverage of VOCs, each plasma sample (50 μL) was dissolved in 50 μL PBS ($\text{pH}=10$). The stock solution of PBS was prepared by mixing 94.7 mL of 0.2 M Na_2HPO_4 with 5.3 mL of 0.2 M NaH_2PO_4 followed by adjusting the pH to 10.0 with NaOH. All 60 samples were analyzed at pH 10, while 57 (29 cases and 28 controls) were at pH 7. The SPME fiber was inserted into the vial and extracted for 10 min after the sample was agitated at 60°C for 20 min. The SPME fiber was then transported in the inlet of GC $\times$ GC-MS at 250°C for 3 min.

The LECO Pegasus 4D GC $\times$ GC-TOF MS instrument was equipped with an Agilent 6890 gas chromatograph and a Gerstel MPS2 auto-sampler (GERSTEL Inc., Linthicum, MD, USA), featuring a LECO two-stage cryogenic modulator and secondary oven. The primary column was a 60 m $\times$ 0.25 mm $^1d_c \times 0.25 \mu\text{m}^1d_i$ DB-5ms GC capillary column (phenyl arylene polymer virtually equivalent to a (5%-phenyl)-methylpolysiloxane). A second GC column of 1 m $\times$ 0.25 mm $^1d_c \times 0.25 \mu\text{m}^2d_i$ DB17ms ((50%-phenyl)-methylpolysiloxane) was placed inside the secondary GC oven after the thermal modulator. Both columns were obtained from Agilent Technologies (Agilent Technologies J&W, Santa Clara, CA, USA). The helium carrier gas (99.999% purity) flow rate was set to 1.0 mL/min at a corrected constant flow via pressure ramps. The inlet temperature was set at 250°C . The primary column temperature was programmed with an initial temperature of 40°C for 2 min and then ramped at 7°C/min to 240°C and maintained for 3 min. The secondary column temperature program was set to an initial temperature of 50°C for 2 min and then ramped at the same temperature gradient employed in the first column to 250°C , accordingly. The thermal modulator was set to $+15^\circ\text{C}$ relative to the secondary oven, and a modulation time of $P_M = 2.0 \text{ s}$ was used. The mass range was set as 30–500 m/z with an acquisition rate of 200 mass spectra per second. The ion source chamber was set at 230°C with the transfer line temperature of 280°C , and the detector voltage was 1450 V with electron energy of 70 eV. The acceleration voltage was

turned on after a solvent delay of 250 s. VOCs were analyzed in the splitless mode.

Data analysis

The GC × GC–MS data were first processed using LECO’s software ChromaTOF (version 4.50.8.0) for the peak picking and the tentative metabolite identification, followed by the peak merging, the peak list alignment and the retention index filtering using MetPP (22). The peak abundances were normalized by the median and were transformed to log2 values, followed by an auto-scaling for the normalization. To compare the number of peaks detected in each sample between groups, a Mann–Whitney *U* test (MW test) was utilized (e.g., Figure 1). The presence rate was then calculated for each peak, which was defined as the ratio of the number of samples with

nonzero abundance to the total number of samples. The peaks with the presence rate < 10% were discarded and the rest of all data analyses (such as partial least squares-discriminant analysis (PLS-DA), Fisher’s exact test, *t*-test, etc.) were performed using the peaks with the presence rate ≥ 10%. Note that the presence rate was computed across all samples regardless of the group information.

PLS-DA was performed using all peaks with the presence rate ≥ 10% (e.g., Figure 2). Fisher’s exact presence/absence test was carried out on the peaks with the presence rate between 10% and 80%. An unpaired *t*-test along with false discovery rate (FDR) control was used to detect VOCs that were differentially expressed between groups (i.e., CRC vs. normal) only for the peaks with the presence rate > 80%. The comparisons of the number of detected peaks between groups were carried out using MW test. The unpaired *t*-test and Kruskal–Wallis test were utilized to investigate

Figure 1. Comparisons of the number of detected peaks between groups. The boxplots display the distributions of the number of detected peaks under pH 10 (left column), pH 7 (middle column), and both (right column) conditions. Group 1 (cases) and 2 (controls) represent colorectal cancer and normal groups. The *P*-values were estimated by a Mann–Whitney *U* test (MW test).

Figure 2. Score plots of the first and second components. The plot in the left column is for the data generated under pH 10 and that for the data under pH 7 is in the right column. Group 1 (cases) and 2 (controls) represent colorectal cancer and normal groups.

Table 1. Characteristics of Colorectal Cancer (CRC) Cases and Population Controls

Characteristics	CRC cases (<i>n</i> = 30)	Controls (<i>n</i> = 30)
Age (years)	62.5 (10.4)	62.9 (10.4)
Sex Male (%)	50%	40%
Non-White (%)	33.3%	43.3%
Body Mass Index	28.8 (5.2)	29.2 (8.9)
Total calorie intake/day (kcal)	2510 (1138)	2241 (1033)
% calorie from fat	38.3 (6.4)	37.4 (5.6)
n6PUFA	21.4 (5.9)	21.5 (5.0)
n3PUFA	6.7 (3.2)	7.1 (2.4)
Alcohol (drinks/week)	2.6 (6.1)	4.4 (8.7)
Smoking status (never/ex/current)	33.3%/33.3%/33.3%	33.3%/33.3%/33.3%
Pack-years of smoking	28.1 (34.4)	17.8 (21.1)
Case: Cancer site (right colon%)	46.70%	...
Case: Cancer stage (0,I,II %)	50%	...
Case: Vital status (% deceased)	43%	...

Numbers indicates means and standard deviations (parentheses) unless specified otherwise.

PUFA: polyunsaturated fatty acids; Vital status was as of 6 July 2016.

None of the differences in these characteristics were statistically significant.

the associations between a VOC metabolite and covariates, and the associations between the detected peak number and covariates of interest, respectively.

Metabolite identification was achieved in three sequential steps. Mass spectrum matching was first employed to identify the top five candidate compounds for each experimental mass spectrum, if the experimental mass spectrum and the best matched mass spectrum recorded in the NIST/EPA/NIH Mass Spectral Library 2011 (NIST11) have a spectral similarity score higher than 600. The tentatively assigned metabolites were then filtered by the retention index matching using iMatch software (23). The best retention index matched metabolite was preserved if multiple tentatively assigned metabolites had the correct retention index values. To verify the identification of metabolites detected with significant abundance difference between sample groups, the authentic standards of these metabolites, if commercially available, were analyzed on GC × GC–MS under the same experimental conditions as the biological samples analyzed. A correct metabolite identification was assigned only if the information of the authentic metabolite agrees with the information of the corresponding chromatographic peak in the biological samples. We used the following criteria for the match: difference of the first-dimension retention time $\Delta^1 t_R \leq 10$ s, difference of the second-dimension retention time $\Delta^2 t_R \leq 0.05$ s, and the mass spectral similarity $\Delta S \geq 700$ (12, 13).

Results

As shown in Table 1, the case (CRC) and control (normal) groups had similar demographic compositions. The control group consisted of slightly higher proportions of females and non-whites, and consumed more total calorie and alcoholic drinks per day, but smoked less cigarettes cumulatively. Half of the cases were at early stage and nearly half had cancer in the proximal colon. No molecular subtypes were available for these CRC cases.

Two sets of VOC metabolomics data were generated by SPME GC × GC–MS under two experimental conditions, pH 10 and pH 7, as mentioned before. The two VOC datasets were preprocessed for the peak detection, the VOC identification, and the cross-sample peak list alignment using MetPP software (22), resulting in two peak

tables. The number of metabolite peaks detected in each sample extracted at pH 7 ranged from 562 to 1378 (median 799), ranged from 113 to 1471 (median 1042) for samples extracted at pH 10, and ranged from 923 to 2183 (median 1601) overall. Note that this analysis was carried out using all peaks, i.e., no peak was filtered out using a threshold of the presence rate. When metabolite peak counts were compared between the groups, the CRC group had a lower number of metabolite peaks detected (Figure 1). Specifically, the difference in the number of metabolite peaks detected was statistically significantly lower in the CRC group at pH 10 (median 990.5 vs. 1085.5, $P = 0.04$) and overall (median 1530 vs. 1694, $P = 0.02$). Among all these detected metabolite peaks, 1041 metabolites were identified for the chromatographic peaks detected from the samples extracted at pH 7 by SPME and 1086 metabolites were identified from the samples extracted at pH 10.

After cross-sample peak list alignment, the pH 10 dataset had 1027 peaks and the pH 7 dataset had 958 peaks. In order to control peaks with excessive zero abundances in the peak tables, the presence rate was calculated for each peak and the peaks with the presence rate < 10% were removed for the analyses. As described in the previous section, the presence rate of each detected VOC was defined as the number of samples with nonzero abundance divided by the total number of samples regardless of the group. After filtering out the peaks in the peak tables using the presence rate < 10%, a total of 654 and 464 peaks were left for the pH 10 dataset and the pH 7 dataset, respectively.

The subsequent analyses were focused on these 654 and 464 metabolites present in at least 10% of the overall pH 10 and pH 7 samples, respectively. As illustrated in Figure 2, PLS-DA plots based on the two most discriminative components showed two clear clusters for the CRC and control samples. Among the samples extracted at pH 7, the largest loading in the first component was metabolite 2,4-dimethylhept-1-ene (27.98%) and the largest loading in the second component was metabolite 1-pentene,2,4,4-trimethyl- (18.76%). The corresponding metabolites in the samples extracted at pH 10 were 2,4-dimethylhept-1-ene (19.36%) and octane,3,5-dimethyl- (12.57%), respectively.

There were 83 metabolites in the samples extracted at pH 7 and 53 metabolites in the samples extracted at pH 10 that were found in

Table II. List of VOCs that have Significantly Different Abundances (Relative Intensities) Between Groups (*t*-test *P*-value < 0.01 and Presence Rate > 80%)

Name	CAS	Retention time (s)		Sample size		Total	Cases		Controls		Fold change Cases/Controls	<i>t</i> -Test <i>P</i> -value	FDR
		First	Second	Group 1	Group 2		Relative intensity	SD	Relative intensity	SD			
pH 7	2,4-Dimethylhept-1-ene	-	660.0005	28 (97)	28 (100)	56 (98)	0.8623	1.6636	1.6239	1.2529	0.5310	<0.0001	<0.0001
	Octane, 3,5-dimethyl-	15869-93-9	917.2976	27 (93)	28 (100)	55 (96)	0.8930	1.4045	1.2872	1.2147	0.6938	<0.0001	0.0004
	Hexane, 3-ethyl-	619-99-8	630.0000	29 (100)	28 (100)	57 (100)	1.0200	1.1783	0.7996	1.3173	1.2758	0.0002	0.0059
	Nonane	111-84-2	750.2500	25 (86)	23 (82)	48 (84)	3.1234	1.3091	2.5691	1.2786	1.2158	0.0059	0.1086
	Nonane, 4-ethyl-5-methyl-	1632-71-9	684.0000	28 (97)	26 (93)	54 (95)	0.7471	1.2368	0.6402	1.2161	1.1670	0.0075	0.1086
	Trichloromethane	67-66-3	378.0938	29 (100)	28 (100)	57 (100)	1.5479	1.7569	2.4795	2.0396	0.6243	0.0079	0.1086
pH 10	2,4,6-Trimethyl-1-nonene	55771-40-9	1014.4727	23 (77)	26 (87)	49 (82)	0.6240	1.2887	0.9742	1.3628	0.6405	<0.0001	0.0001
	2,4-Dimethylhept-1-ene	659-9995	1.2601	26 (87)	28 (93)	54 (90)	0.6624	1.8117	1.3012	1.4540	0.5091	<0.0001	0.0003
	Hexane, 2,3,4-trimethyl-	921-47-1	630.1787	26 (87)	28 (93)	54 (90)	0.8340	1.2017	0.5954	1.4075	1.4007	<0.0001	0.0008
	Hexane, 3-ethyl-	619-99-8	694.1110	28 (93)	22 (73)	50 (83)	0.7358	1.2150	0.5820	1.2746	1.2641	0.0007	0.0087
	Octane, 3,5-dimethyl-	15869-93-9	912.6165	25 (83)	29 (97)	54 (90)	0.8728	1.3769	1.1415	1.3649	0.7646	0.0025	0.0255

at least 80% of all samples (i.e., the presence rate >80%). Table I lists the eight VOCs that had significantly different abundances between the groups at the conventional 1% significance level at both pH levels ($n = 3$) or either pH level ($n = 5$). Among these eight, five metabolites remained significant at the 1% FDR level after adjustment for multiple comparisons. The abundance level of 2,3,4-trimethyl-hexane in the colorectal samples extracted at pH 10 was significantly increased compared to that in the control samples (intensity ratio (IR) = 1.40). The abundance level of 3-ethyl-hexane was significantly increased in the colorectal samples than in the control samples at both pH 7 and pH 10 (IR = 1.28 and IR = 1.26, respectively). However, the abundance level of 2,4-dimethylhept-1-ene was significantly decreased in the colorectal samples (IR = 0.53, at pH 7 and IR = 0.51 at pH 10). The abundance level of 3,5-dimethyl-octane was also significantly decreased in the colorectal samples (IR = 0.69 at pH 7 and IR = 0.76 at pH 10). Likewise, the abundance level of 2,4,6-trimethyl-1-nonene was decreased in the colorectal samples with the IR = 0.64 at pH 10 (Table II) (see Supporting Information Supplementary Tables S1–S2 for all detected metabolites).

Among the other metabolites less frequently present (i.e., presence rate $\leq 80\%$), after adjustment for multiple comparisons, we found that only one VOC, 2,4,6,8-tetramethyl-1-undecene, was differentially detected between the CRC samples and controls (odds ratio (OR) = 0.09 with pH 7 and 0.03 with pH 10). Among several other VOCs that yielded a significant difference at the conventional level ($P < 0.01$), a substantially increased OR was observed for oxetane, 2,2-dimethyl- (OR = 8.83) and cyclopentane, 1-methyl-2-(2-propenyl)-, *trans*- (OR = 7.25) at pH 7 (Table III and Supplementary Tables S3–S4).

We further examined the abundance of two major VOCs (2,3,4-trimethyl-hexane and 2,4-dimethylhept-1-ene) and total metabolite count that showed significance differences. The associations with various potential endogenous and exogenous modifiers were investigated among the controls and those with disease characteristics were assessed among the CRC samples. Among the controls, females had higher abundance of 2,3,4-trimethylhexane than males, obese subjects had higher abundance of 2,4-dimethylhept-1-ene than the non-obese non-overweight subjects, and alcohol intake was associated with higher total metabolite count, but with lower abundance of 2,3,4-trimethyl-hexane. No other dietary parameter, such as total energy intake, % calories from fat or types of fatty acids, showed an association. Among the CRC cases, there was no difference in these three measurements with stages of the disease or locations of the tumor. However, surviving cases had higher abundance of 2,3,4-trimethyl-hexane than deceased cases (Table IV).

Discussion

The primary aim of this study was to promote the applications of advanced MS technologies, specifically SPME GC $\times$ GC, for biomarker discovery by untargeted metabolomics on biological samples in the long-term storage. SPME is a powerful sample preparation technique widely used for sampling a wide spectrum of analytes present in media in different physiological state that are characterized by the complex matrix composition, e.g., environmental, biological and food samples. Among many extraction techniques, SPME is distinguished by several advantageous features, such as: (i) simplicity of operation, (ii) versatility, (iii) relatively low cost of equipment, (iv) short-extraction time, (v) complete elimination of organic solvents from the analytical protocol, (vi) possibility to

Table III. List of VOCs that have the Significantly Different Number of Samples Present between Groups (Fisher's Exact Test P -value < 0.01 and Presence Rate $\leq 80\%$)

Name	CAS	1t_R (s)	2t_R (s)	Number of samples present			Odds ratio	Fisher's exact test		FDR
				Cases N (%)	Controls N (%)	Total N (%)		Cases/Controls	P -value	
pH 7	59920-26-2	1324	1.305	4 (14)	18 (64)	22 (39)	0.0935	0.0001	0.0001	0.0446
	2,4,6,8-Tetramethyl-1-undecene	1300	1.3	13 (45)	24 (86)	37 (65)	0.1407	0.0020	0.0020	0.3756
	Oxalic acid, allyl nonyl ester	378	1.295	12 (41)	2 (7)	14 (25)	8.8326	0.0046	0.0046	0.4353
	Oxetane, 2,2-dimethyl-	2056	1.045	2 (7)	11 (39)	13 (23)	0.1189	0.0046	0.0046	0.4353
	7-Hydroxy-8-methylisoflavone	1100	1.31	6 (21)	16 (57)	22 (39)	0.2019	0.0067	0.0067	0.5084
	Oxalic acid, allyl heptyl ester	1254	1.275	4 (14)	13 (46)	17 (30)	0.1905	0.0096	0.0096	0.5402
	2-Propyl-1-pentanol	660	1.29	27 (93)	18 (64)	45 (79)	7.2456	0.0099	0.0099	0.5402
pH 10	50746-53-7	1314	1.32	4 (13)	25 (83)	29 (48)	0.0337	<0.0001	<0.0001	<0.0001
	Cyclopentane, 1-methyl-2-(2-propenyl)-, trans-	903.7179	1.133018	3 (10)	13 (43)	16 (27)	0.1502	0.0074	0.0074	>0.99
	2,4,6,8-Tetramethyl-1-undecene									
	Cyclotrisiloxane, hexamethyl-									

collect samples *in situ* and *in vivo*, (vii) possible automation and (viii) the direct desorption of the analytes from the fiber inside a measuring system (24). As a result, SPME is considered to be particularly useful for identifying biomarkers in the untargeted metabolomic studies (25). Three earlier studies applied SPME on different types of specimens from CRC patients and controls. Silva *et al.* identified 82 VOCs in urine samples from 12 cases and 21 controls. Among those 12 metabolites, it showed more than 2-fold differences (26). Wang *et al.* analyzed blood from 16 cases and 20 controls, identified 341 metabolites and found a significant difference in four metabolites, i.e., phenyl methylcarbamate, ethylhexanol, 1,1,4,4-tetramethyl-2,5-dimethylene-cyclohexane and 6-*t*-butyl-2,2,9,9-tetramethyl-3,5-decadiene-7-yne (27). Their subsequent study on breath samples from 20 cases and 20 controls detected 364 metabolites and nine metabolites showed a significant difference (28). Compared with these, the current study identified more metabolites, including metabolites not previously detected, and detected a similar number of significant metabolites. The present study cataloged over 1000 VOC metabolites, while providing a list of candidate biomarkers for further validation to improve CRC prevention and patient care. This study also suggested that richness of volatile metabolites might be reduced in CRC cases than those in controls. To our knowledge, no previous study on CRC utilized both SPME and GC $\times$ GC-MS to maximize the biomarker discovery.

Tables V and VI list the intraday and interday variations of GC $\times$ GC-MS in measuring a mixture of *n*-alkanes, respectively, where the variations of the first-dimension retention time (1t_R), the second-dimension retention time 2t_R and peak area were measured as relative standard deviation (RSD). Both the intraday and interday RSD of 1t_R is $< 0.2\%$, while the RSD of 2t_R is $< 0.4\%$ owing to its short-column length. The intraday variation of peak area is in a range of 0.8–6.3% with a mean of 3.2%. The interday variation of peak area is in a range of 3.0–8.8% with a mean of 5.2%. Overall, GC $\times$ GC-MS is very robust in compound detection. However, in general, the current compound identification accuracy in analysis of GC-MS and GC $\times$ GC-MS by mass spectrum matching is only 85% (29). For this reason, the VOC biomarkers identified in this work and by others need to be validated. In fact, it is our plan that the information of VOC biomarkers discovered in this work will be integrated with other biological data to generate some biological hypothesis for future experiments, in which the VOC validation will be performed.

Whereas applications of GC $\times$ GC-MS have been rather limited in studies of cancer biomarkers with a few publications on non-gastrointestinal cancers (30, 31), wealth of GC $\times$ GC-MS data have been published in various types of metabolomic studies in humans. Some of these results may be relevant to the current study, although they were not necessarily focused on VOCs. One of the largest studies involving more than 300 diabetic and non-diabetic Danish individuals revealed the link between insulin resistance and gut bacterial metabolite, specifically branched chain amino acids (32). The same group previously identified a range of metabolites associated with a low postprandial insulin response after sourdough fermented rye bread intake (33). Applying GC $\times$ GC, they also reported that mucosa from irritable bowel syndrome patients had distinct pro-inflammatory metabolic profile compared with that from control subjects (34). In their more recent intervention study with dietary polyphenols, they found significant correlations between *in vitro* fermentation metabolites with fecal anaerobic incubation and *in vivo* urinary metabolites for those with C1–C3 side chain (35).

Table IV. Analyses of Relative Intensities of the Two Major VOCs (2,3,4-trimethyl-hexane and 2,4-dimethylhept-1-ene) and Total Metabolite Count by Potential and Exogenous Modifiers and Disease Characteristics

2,4-Dimethylhept-1-ene (pH 7)				Hexane, 2,3,4-trimethyl (pH 10)				pH 7+pH 10 peak number						
<i>n</i>				Relative intensity (95% CI)	<i>P</i> - value*	<i>n</i>				Relative intensity (95% CI)	<i>P</i> - value*	<i>n</i>	Counts median (range)	<i>P</i> - value**
Controls	Age	<65	14	1.563 (1.350,1.810)	0.385	14	0.586 (0.462,0.745)	0.819	14	1804 (1185,2182)	0.066			
		>65	14	1.687 (1.555,1.830)		14	0.605 (0.549,0.666)		14	1604 (1117,1884)				
	Sex	Female	16	1.667 (1.495,1.859)	0.492	16	0.668 (0.570,0.782)	0.038	16	1655 (1185,1948)	0.486			
		Male	12	1.568 (1.373,1.791)		12	0.511 (0.427,0.611)		12	1718 (1117,2182)				
	Race	White	12	1.604 (1.451,1.773)	0.796	12	0.556 (0.453,0.682)	0.377	12	1756 (1117,1884)	0.546			
		Non-white	16	1.639 (1.442,1.863)		16	0.627 (0.533,0.736)		16	1650 (1185,2182)				
	Smoking	Never	9	1.722 (1.469,2.018)	0.511	9	0.589 (0.493,0.704)	0.719	9	1581 (1377,1948)	0.627			
		Current	9	1.593 (1.353,1.875)		9	0.622 (0.496,0.780)		9	1704 (1185,1879)				
	Alcohol (drinks)	None	20	1.645 (1.494,1.810)	0.673	20	0.658 (0.579,0.748)	0.027	20	1604 (1117,2182)	0.033			
		5≥/ week	8	1.573 (1.319,1.877)		8	0.464 (0.365,0.589)		8	1848 (1593,1948)				
	Body Mass index	<25	9	1.830 (1.602,2.090)	0.026	10	0.540 (0.409,0.714)	0.395	9	1819 (1185,1948)	0.409			
		≥30	6	1.493 (1.370,1.626)		7	0.635 (0.504,0.800)		6	1674 (1117,2182)				
	Total calorie (tertile)	Bottom	12	1.620 (1.455,1.804)	0.659	13	0.607 (0.550,0.670)	0.763	12	1734 (1117,1948)	0.817			
		Top	8	1.548 (1.312,1.826)		8	0.574 (0.411,0.802)		8	1660 (1431,1879)				
	% Fat (tertile)	Bottom	10	1.576 (1.329,1.868)	0.480	9	0.604 (0.521,0.701)	0.296	10	1718 (1377,1948)	0.790			
		Top	8	1.701 (1.510,1.917)		8	0.496 (0.361,0.682)		8	1702 (1397,2182)				
	n6 fat (tertile)	Bottom	9	1.522 (1.250,1.852)	0.339	7	0.630 (0.502,0.790)	0.386	9	1704 (1397,1948)	0.806			
		Top	10	1.703 (1.538,1.886)		11	0.542 (0.426,0.690)		10	1620 (1185,2182)				
	n3 fat (tertile)	Bottom	9	1.633 (1.505,1.772)	0.317	8	0.613 (0.482,0.780)	0.789	9	1737 (1431,1884)	0.346			
		Top	6	1.542 (1.437,1.655)		6	0.577 (0.405,0.824)		6	1604 (1185,2182)				
cases	Tumor location	Right	12	0.847 (0.645,1.113)	0.876	12	0.808 (0.713,0.916)	0.445	13	1502 (1236,1754)	>0.99			
		Left	16	0.874 (0.669,1.140)		14	0.857 (0.793,0.926)		16	1548 (923,2131)				
	Stage	Early	14	0.951 (0.747,1.211)	0.319	11	0.830 (0.728,0.946)	0.919	14	1450 (923,1746)	0.081			
		Other	14	0.782 (0.586,1.044)		15	0.837 (0.772,0.907)		15	1569 (1344,2131)				
	Vital status	Alive	16	0.905 (0.715,1.145)	0.585	16	0.886 (0.820,0.957)	0.044	17	1502 (964,1701)	0.308			
		Dead	12	0.809 (0.590,1.110)		10	0.757 (0.673,0.852)		12	1633 (923,2131)				
	* Unpaired <i>t</i> -test										**	Kruskal–Wallis test		

Table V. Intraday Variation of *n*-Alkanes Measured by GC × GC–MS in Five Consecutive Injections

<i>n</i> -Alkane (carbon no.)	¹ <i>t</i> _R (s)	RSD (%)	² <i>t</i> _R (s)	RSD (%)	Peak area	RSD (%)
8	362	0.0	0.966	0.2	1699953	0.9
9	494	0.0	1.023	0.2	2230219	6.3
10	658	0.0	1.070	0.3	1947898	0.8
11	834	0.0	1.102	0.2	1814149	1.2
12	1011	0.2	1.119	0.2	1827610	0.8
13	1182	0.0	1.133	0.2	1226468	3.1
14	1346	0.0	1.141	0.2	1342970	1.3
15	1498	0.0	1.156	0.2	1098484	1.6
16	1642	0.0	1.166	0.2	1095185	1.4
17	1782	0.0	1.175	0.0	892963	6.4
18	1914	0.0	1.184	0.2	821838	3.8
19	2038	0.0	1.195	0.0	741459	4.6
20	2154	0.0	1.210	0.0	684299	4.2
21	2270	0.0	1.222	0.2	707001	8.2
22	2378	0.0	1.236	0.2	652069	8.0
23	2482	0.0	1.250	0.0	575384	3.4
24	2582	0.0	1.266	0.2	582031	3.3
25	2678	0.0	1.282	0.2	566922	3.1
26	2773	0.1	1.362	0.2	535651	2.9
27	2882	0.0	1.496	0.1	499746	2.6
28	3006	0.0	1.653	0.2	470944	2.2
29	3158	0.0	1.850	0.2	452500	2.5
30	3346	0.0	2.091	0.3	437012	1.2

Table VI. Interday Variation of *n*-Alkanes Measured by GC × GC–MS. The Mixture of *n*-Alkanes was Analyzed Five Times in 5 Days

<i>n</i> -Alkane (carbon no.)	¹ <i>t</i> _R (s)	RSD (%)	² <i>t</i> _R (s)	RSD (%)	Peak area	RSD (%)
8	362	0.0	0.966	0.4	1491520	4.6
9	494	0.0	1.023	0.3	1514996	5.0
10	658	0.0	1.069	0.3	1466503	3.1
11	834	0.0	1.102	0.3	1201987	7.6
12	1011	0.1	1.119	0.3	1227809	3.9
13	1182	0.0	1.134	0.3	824941	7.1
14	1346	0.0	1.141	0.4	944611	4.2
15	1498	0.0	1.156	0.3	743075	7.9
16	1642	0.0	1.167	0.4	752672	3.1
17	1782	0.0	1.175	0.3	639675	4.3
18	1913	0.1	1.185	0.3	614181	3.8
19	2038	0.0	1.197	0.3	579009	6.7
20	2154	0.0	1.212	0.3	517067	3.6
21	2270	0.0	1.223	0.4	494432	4.8
22	2378	0.0	1.237	0.3	451337	4.8
23	2482	0.0	1.251	0.2	437198	4.0
24	2581	0.1	1.267	0.4	455025	8.2
25	2677	0.1	1.284	0.4	444927	8.1
26	2772	0.1	1.363	0.3	411515	8.8
27	2881	0.1	1.497	0.3	382151	4.7
28	3006	0.0	1.655	0.3	359866	3.8
29	3158	0.0	1.852	0.4	345654	3.0
30	3344	0.1	2.093	0.4	332238	3.8

We acknowledge that blood is a more challenging matrix than breath, urine or saliva for VOC analysis, due to constant homeostatic regulations, hemodynamic fluctuations, complex compartmental distributions of blood-borne VOCs and ubiquitous other confounders (e.g., foods and medications). While confounding factors of ambient air can be avoided via end-tidal/alveolar sampling of breath (36), many systemic confounders associated with blood are difficult to be ruled out completely due to inseparable and complex blood matrix.

To our knowledge, VOC richness in blood samples has not been previously tested for CRC cases and controls. Several investigators analyzed fecal VOCs and found that fecal VOC counts are lower in individuals at high risk for CRC, i.e., patients with inflammatory bowel disease than control subjects (37, 38). Although it was not limited to VOCs, the total fecal metabolite count was greatly reduced in mice fed with an unhealthy (high salt) diet, compared with those fed with a control diet (39). Our finding of reduced blood VOC counts in the CRC cases was intriguing. The majority of fecal

metabolites are presumably bacterial, and reduced microbial richness has been linked to development of high-grade premalignant colorectal adenoma in humans (40). Moreover, feeding fecal samples from CRC patients to conventional and germ-free mice led to significantly reduced microbial community richness compared with feeding fecal samples from controls (41). Besides, richness of gut microbiome and of microbial genes has been associated with favorable metabolic profiles and responses in human subjects (42, 43).

Cancer-associated VOCs have received growing interest recently as a new class of biomarkers. Combining both *in vivo* and *in vitro* studies, to date, at least 115 VOCs have shown some associations with one type of cancer (2). These metabolites span over a wide range of chemical classes, i.e., hydrocarbons, alcohols, aldehydes, acids, ketones, aromatic compounds, heterocycles, nitriles, terpenes, esters, ethers, sulfides, halogenated compounds and others (2, 3). In these studies, overall good discrimination of cancer cases from control subjects has been reported by using a combination of multiple metabolites (27, 28, 44, 45), and our PLS-DA based on the metabolites present in at least 10% of the samples also demonstrated an ability of identified VOCs to capture cancer-associated metabolomic signatures in this study.

However, as far as individual VOC is concerned, the evidence from global volatolomic studies was often based on single studies and the reproducibility of specific compounds or combinations has been poor between studies (2, 46). In case of CRC, in addition to breath, various bodily fluids have been used to identify cancer specific VOCs, including blood like ours, feces, and urine (46). Again, there have been few overlaps in candidate CRC specific VOCs confirmed in multiple studies (46). The shortlist includes acetone and 4-methyloctane (46), neither of which was discriminative between the cases and controls in our study. Another interesting observation by others is that resection of CRC does not normalize metabolic shifts seen in breath VOCs (47) or in blood total metabolites (48) preoperatively, implying that such metabolic markers may have played etiological roles rather than being just excreted from the tumor tissue. We are also aware that VOCs include many from a pure exogenous origin, being neither human nor bacterial metabolites. Such compounds may be related to smoking, foods and medicines ingested, household chemicals, occupational exposure, environmental pollutants; and emanate from medical and plastic materials, and fuel combustion (49). Consequently, biochemical pathways often cannot be determined to link these compounds to specific cellular and carcinogenic processes, as was the case in our study.

We found that 2,4-dimethylhept-1-ene and 2,3,4-trimethylhexane were differentially expressed between the cases and controls. While both aliphatic (acyclic) VOCs can be lipid peroxidation products, the former (alkene) was significantly lower and the latter (alkane) was significantly higher in CRC cases than in the controls. Interestingly, aliphatic alkanes and aldehydes have been postulated to serve as *in vivo* oxidative stress related to carcinogenesis (49–51). Their increased levels in the exhaled breath have been found in lung cancer patients (49, 51) and methylated alkanes have been suggested to be important discriminative metabolites between cancer and control subjects (51). However, the earlier study targeted only up to monomethylated alkanes and accordingly the potential utility of multi-methylated alkanes like 2,3,4-trimethylhexane remains to be elucidated. On the other hand, 2,4-dimethylhept-1-ene levels in the breath were significantly lower in the non-smoking Austrian lung cancer cases than in controls, with the mean concentration ratio of 0.62 (52), which was within the range of our own observations (0.51 and 0.53 for pH 7 and pH 10, respectively). Moreover,

in vitro culture studies support these clinical observations, demonstrating significantly reduced emission of 2,4-dimethylhept-1-ene from a cancer cell line compared with normal control cell lines (53).

In addition to being endogenous lipid oxidation products, both 2,4-dimethylhept-1-ene and 2,3,4-trimethylhexane are potentially ingested from foods as they are present in fats and oils of both animal and plant origins (54, 55). When we analyzed the associations of these metabolite levels with lipid related parameters in our control subjects (i.e., obesity, fat intake and its subfractions), we did not find any consistent trends. However, obese subjects had significantly lower 2,4-dimethylhept-1-ene levels compared with normal (BMI < 25) weight subjects. This corroborates the facts that obesity is an established risk factor for CRC (56) and that CRC cases had a decreased level of circulating 2,4-dimethylhept-1-ene in our study. Alternatively, 2,4-dimethylhept-1-ene can be produced by some bacteria, such as certain species of *Pseudomonas* and *Actinomyces*, which are also capable of synthesizing uncharacterized antimicrobial agents (57, 58). Thus, lower abundance of this metabolite may reflect potential underlying changes in gut microbiome.

Among less ubiquitously present VOCs which are likely to consist of more VOCs from exogenous origins, only 2,4,6,8-tetramethyl-1-undecene showed differential detection rates between the CRC cases and controls after adjustment for multiple comparisons, being significantly lower in the cases. There is the sparse information regarding this compound, but the presence has been reported in certain medicinal plants and algae (59, 60), while it could be a thermal degradation product from plastics (61). To date, there has been no solid data to link this VOC to any biological or clinical pathways. Despite the lack of statistical significance after multiple comparison adjustment, it is noteworthy that 2,2-dimethyloxetane, a cyclic ether, yielded the highest OR close to 9 with a conventional $P < 0.01$. Thus, this VOC seems rather cancer specific and a potential biomarker, although it is likely to be exogenous origins without any known *in vivo* metabolic pathways in humans. Interestingly, in the aforementioned cell culture study (53), significantly increased emission of this VOC from the lung cancer cell line, compared with control cell lines, was also observed. It remains to be studied whether this metabolite represents environmental etiological factors other than cigarette smoking, as smoking status was matched between the cases and controls in our study. Finally, as this compound was detected only for a little over 40% of the cases and as that the presence was not related to the tumor location, stage or prognosis, further studies are needed to identify specific molecular subtypes linked to this specific VOC.

We took sufficient precautions in analyzing high-dimensional data using appropriate statistical adjustment for multiple comparisons, which has been rarely enforced in earlier studies. Other strengths in the study design include the controls derived from the same general population that gave a rise to the cases, availability of an array of potential covariates, and a single-institutional study to assure standardized procedures. On the other hand, being an early stage of the study, the sample size was too small to obtain estimates with the high precision. Besides, as biological samples used in this study were collected after cancer treatment/resection from the cases, the direct effect of the presence of the primary tumor was deemed to have faded, although underlying metabolic deviations that fostered cancer development was expected to remain as discussed above. In addition, although we used two types of extraction conditions (pH 7 and pH 10) to maximize VOC recovery, our results are likely to underrepresent acidic analytes (50, 62) due to the omission of the low pH conditions. There is another possible concern about

metabolite stability over the long storage. Studies have demonstrated that stability is better at the lower temperature (e.g., a room temperature, -20°C vs. -80°C) (21, 63, 64), but it varies substantially with types of metabolites, i.e., a better stability of non-polar versus polar metabolites (63, 65), which may explain our observation that most of the metabolites showing a difference in the relative intensity were the non-polar VOC. There is a matrix effect on the stability of VOCs during the long-term storage. Gases/breath are more liable to the degradation (66), while liquid and solid masses are less liable (21, 64, 67). Despite the documented degeneration over the long-term storage (67), the storage at 80°C is considered appropriate for most studies. We assume that the degradation would occur similarly over the time between cases and controls, but acknowledge that the discriminating power may decline with the longer storage due to the degradation of low abundance metabolites. However, it is worth noting that serum/plasma samples stored at 80°C at for 10–15 years, just like our samples, have been successfully used in the untargeted metabolomic studies to discern disease groups (63, 68).

Conclusions

VOCs extracted by SPME from plasma samples of CRC ($n = 30$) and normal ($n = 30$) subjects were analyzed on a GC $\times$ GC-MS. It was shown that the number of metabolite peaks detected in the cases was significantly lower than that detected in the controls (median 1530 vs. 1694, $P = 0.02$). This may imply that the richness of volatile metabolites is reduced in CRC cases. PLS-DA revealed a difference in VOC profiles between the CRC samples and the control samples. After the adjustment for multiple comparisons, five VOCs were differentially expressed between the cases and the controls. Among these five VOCs, 2,3,4-trimethylhesane (decreased in CRC) and 2,4-dimethylhept-1-ene (increased in CRC) were both lipid peroxidation products but not previously reported for CRC. Given the limitations discussed above, a caution should be excised not to over-interpret our results and, due to the small sample size, these results need to be validated with a larger sample size. Nevertheless, this proof-of-principle study demonstrates the potential of this analytical platform to distinguish CRC cases from controls based on metabolite profiles in the future studies.

Supplementary Data

Supplementary material is available at *Journal of Chromatographic Science* online.

Acknowledgments

We thank Ms Barbara Rusin, Dr Maria Samerson and Ms Ann Bankowski from Wayne State University for technical support and/or collection/selection of blood samples from cases and controls.

Funding

The project was supported by grants R01-CA93817 from National Cancer Institute, Grant Boost Program from Wayne State University, 1P20GM113226 and P50AA024337 from National Institutes of Health (NIH). The Biostatistics Core in the Karmanos Cancer Institute at Wayne State University is supported in part by the NIH Cancer Center Support Grant P30 CA022453.

References

1. Nakhleh, M.K., Amal, H., Jeries, R., Broza, Y.Y., Aboud, M., Gharra, A., *et al.*; Diagnosis and classification of 17 diseases from 1404 subjects via pattern analysis of exhaled molecules; *ACS Nano*, (2017); 11: 112–125.
2. Haick, H., Broza, Y.Y., Mochalski, P., Ruzsanyi, V., Amann, A.; Assessment, origin, and implementation of breath volatile cancer markers; *Chemical Society Reviews*, (2014); 43: 1423–1449.
3. Amann, A., Mochalski, P., Ruzsanyi, V., Broza, Y.Y., Haick, H.; Assessment of the exhalation kinetics of volatile cancer biomarkers based on their physicochemical properties; *Journal of breath research*, (2014); 8: 016003.
4. Filipiak, W., Mochalski, P., Filipiak, A., Ager, C., Cumeras, R., Davis, C. E., *et al.*; A compendium of volatile organic compounds (VOCs) released by human cell lines; *Current Medicinal Chemistry*, (2016); 23: 2112–2131.
5. Sonoda, H., Kohnoe, S., Yamazato, T., Satoh, Y., Morizono, G., Shikata, K., *et al.*; Colorectal cancer screening with odour material by canine scent detection; *Gut*, (2011); 60: 814–819.
6. Tadeusz, J., Marta, W., Tomasz, L., Joanna, R., Bogusław, B.; Study of the art: canine olfaction used for cancer detection on the basis of breath odour. Perspectives and limitations; *Journal of Breath Research*, (2015); 9: 027001.
7. Hakim, M., Broza, Y.Y., Barash, O., Peled, N., Phillips, M., Amann, A., *et al.*; Volatile organic compounds of lung cancer and possible biochemical pathways; *Chemical Reviews*, (2012); 112: 5949–5966.
8. Nagpal, M., Singh, S., Singh, P., Chauhan, P., Zaidi, M.; Tumor markers: a diagnostic tool; *National Journal of Maxillofacial Surgery*, (2016); 7: 17–20.
9. Van den Velde, S., Nevens, F., Van Hee, P., van Steenberghe, D., Quirynen, M.; GC-MS analysis of breath odor compounds in liver patients; *J Chromatogr B*, (2008); 875: 344–348.
10. Poli, D., Goldoni, M., Corradi, M., Acampa, O., Carbognani, P., Internullo, E., *et al.*; Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatization SPME-GC/MS; *J Chromatogr B*, (2010); 878: 2643–2651.
11. Shi, X., Wahlang, B., Wei, X.L., Yin, X.M., Falkner, K.C., Prough, R.A., *et al.*; Metabolomic analysis of the effects of polychlorinated biphenyls in nonalcoholic fatty liver disease; *Journal of Proteome Research*, (2012); 11: 3805–3815.
12. Shi, X., Wei, X., Koo, I., Schmidt, R.H., Yin, X., Vaughn, A., *et al.*; Metabolomic analysis of the effects of chronic arsenic exposure in a mouse model of diet-induced fatty liver disease; *Journal of Proteome Research*, (2014); 13: 547–554.
13. Shi, X., Wei, X., Yin, X., Wang, Y., Zhang, M., Zhao, C., *et al.*; Hepatic and fecal metabolomic analysis of the effects of *Lactobacillus rhamnosus* GG on alcoholic fatty liver disease in mice; *Journal of Proteome Research*, (2015); 14: 1174–1182.
14. Winnike, J.H., Wei, X.L., Knagge, K.J., Colman, S.D., Gregory, S.G., Zhang, X.; Comparison of GC-MS and GCXGC-MS in the analysis of human serum samples for biomarker discovery; *Journal of Proteome Research*, (2015); 14: 1810–1817.
15. Seer cancer statistics review., 1975–2014; National Cancer Institute, Bethesda, MD (2017). http://seer.cancer.gov/csr/1975_2014/
16. Nicholson, B.D., Shinkins, B., Pathiraja, I., Roberts, N.W., James, T.J., Mallett, S., *et al.*; Blood CEA levels for detecting recurrent colorectal cancer; *Cochrane Database of Systematic Reviews*, (2015). doi: 10.1002/14651858.CD011134.
17. Kato, I., Land, S., Majumdar, A.P., Barnholtz-Sloan, J., Severson, R.K.; Functional polymorphisms to modulate luminal lipid exposure and risk of colorectal cancer; *Cancer epidemiology*, (2010); 34: 291–297.
18. Kato, I., Majumdar, A.P., Land, S.J., Barnholtz-Sloan, J.S., Severson, R. K.; Dietary fatty acids, luminal modifiers, and risk of colorectal cancer; *International Journal of Cancer. Journal International du Cancer*, (2010); 127: 942–951.
19. Bosch, S., el Manouni el Hassani, S., Covington, J.A., Wicaksono, A.N., Bomers, M.K., Benninga, M.A., *et al.*; Optimized sampling conditions for

- fecal volatile organic compound analysis by means of field asymmetric ion mobility spectrometry; *Analytical Chemistry*, (2018); 90: 7972–7981.
20. Berkhout, D., Benninga, M., van Stein, R., Brinkman, P., Niemark, H., de Boer, N., *et al.*; Effects of sampling conditions and environmental factors on fecal volatile organic compound analysis by an electronic nose device; *Sensors*, (2016); 16: 1967.
 21. Živković Semren, T., Brčić Karačonji, I., Safner, T., Brajenović, N., Tariba Lovaković, B., Pizent, A.; Gas chromatographic-mass spectrometric analysis of urinary volatile organic metabolites: optimization of the HS-SPME procedure and sample storage conditions; *Talanta*, (2018); 176: 537–543.
 22. Wei, X., Shi, X., Koo, I., Kim, S., Schmidt, R.H., Arteel, G.E., *et al.*; MetPP: a computational platform for comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry-based metabolomics; *Bioinformatics (Oxford, England)*, (2013); 29: 1786–1792.
 23. Zhang, J., Fang, A.Q., Wang, B., Kim, S.H., Bogdanov, B., Zhou, Z.X., *et al.*; iMatch: a retention index tool for analysis of gas chromatography-mass spectrometry data; *Journal of Chromatography. A*, (2011); 1218: 6522–6530.
 24. Spietelun, A., Marcinkowski, Ł., de la Guardia, M., Namieśnik, J.; Recent developments and future trends in solid phase microextraction techniques towards green analytical chemistry; *Journal of Chromatography. A*, (2013); 1321: 1–13.
 25. Roszkowska, A., Miękus, N., Bączek, T.; Application of solid-phase microextraction in current biomedical research; *Journal of Separation Science*, (2018).
 26. Silva, C.L., Passos, M., Camara, J.S.; Solid phase microextraction, mass spectrometry and metabolomic approaches for detection of potential urinary cancer biomarkers—a powerful strategy for breast cancer diagnosis; *Talanta*, (2012); 89: 360–368.
 27. Wang, C., Li, P., Lian, A., Sun, B., Wang, X., Guo, L., *et al.*; Blood volatile compounds as biomarkers for colorectal cancer; *Cancer Biology & Therapy*, (2014); 15: 200–206.
 28. Wang, C., Ke, C., Wang, X., Chi, C., Guo, L., Luo, S., *et al.*; Noninvasive detection of colorectal cancer by analysis of exhaled breath; *Analytical and Bioanalytical Chemistry*, (2014); 406: 4757–4763.
 29. Kim, S., Koo, I., Jeong, J., Wu, S., Shi, X., Zhang, X.; Compound identification using partial and semipartial correlations for gas chromatography-mass spectrometry data; *Analytical Chemistry*, (2012); 84: 6477–6487.
 30. Asiago, V.M., Alvarado, L.Z., Shanaiah, N., Gowda, G.A.N., Owusu-Sarfo, K., Ballas, R.A., *et al.*; Early detection of recurrent breast cancer using metabolite profiling; *Cancer Research*, (2010); 70: 8309–8318.
 31. Pasikanti, K.K., Esuvaranathan, K., Hong, Y., Ho, P.C., Mahendran, R., Raman Nee Mani, L., *et al.*; Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry; *Journal of Proteome Research*, (2013); 12: 3865–3873.
 32. Pedersen, H.K., Gudmundsdottir, V., Nielsen, H.B., Hyötyläinen, T., Nielsen, T., Jensen, B.A., *et al.*; Human gut microbes impact host serum metabolome and insulin sensitivity; *Nature*, (2016); 535: 376–381.
 33. Bondia-Pons, I., Nordlund, E., Mattila, I., Katina, K., Aura, A.-M., Kolehmainen, M., *et al.*; Postprandial differences in the plasma metabolome of healthy Finnish subjects after intake of a sourdough fermented endosperm rye bread versus white wheat bread; *Nutrition Journal*, (2011); 10: 116.
 34. Kajander, K., Myllyluoma, E., Kyrönpalo, S., Rasmussen, M., Sipponen, P., Mattila, I., *et al.*; Elevated pro-inflammatory and lipotoxic mucosal lipids characterise irritable bowel syndrome; *World Journal of Gastroenterology*, (2009); 15: 6068–6074.
 35. Vetrani, C., Rivellese, A.A., Annuzzi, G., Adiels, M., Borén, J., Mattila, I., *et al.*; Metabolic transformations of dietary polyphenols: comparison between *in vitro* colonic and hepatic models and *in vivo* urinary metabolites; *The Journal of Nutritional Biochemistry*, (2016); 33: 111–118.
 36. Phillips, M., Herrera, J., Krishnan, S., Zain, M., Greenberg, J., Cataneo, R.N.; Variation in volatile organic compounds in the breath of normal humans; *Journal of Chromatography B: Biomedical Sciences and Applications*, (1999); 729: 75–88.
 37. Garner, C.E., Smith, S., de Lacy Costello, B., White, P., Spencer, R., Probert, C.S.J., *et al.*; Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease; *The FASEB Journal*, (2007); 21: 1675–1688.
 38. De Preter, V., Machiels, K., Joossens, M., Arijis, I., Matthys, C., Vermeire, S., *et al.*; Faecal metabolite profiling identifies medium-chain fatty acids as discriminating compounds in ibd; *Gut*, (2015); 64: 447–458.
 39. Wilck, N., Matus, M.G., Kearney, S.M., Olesen, S.W., Forslund, K., Bartolomeus, H., *et al.*; Salt-responsive gut commensal modulates TH17 axis and disease; *Nature*, (2017); 551: 585.
 40. Peters, B.A., Dominianni, C., Shapiro, J.A., Church, T.R., Wu, J., Miller, G., *et al.*; The gut microbiota in conventional and serrated precursors of colorectal cancer; *Microbiome*, (2016); 4: 69.
 41. Wong, S.H., Zhao, L., Zhang, X., Nakatsu, G., Han, J., Xu, W., *et al.*; Gavage of fecal samples from patients with colorectal cancer promotes intestinal carcinogenesis in germ-free and conventional mice; *Gastroenterology*, (2017); 153: 1621–1633.e1626.
 42. Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., *et al.*; Richness of human gut microbiome correlates with metabolic markers; *Nature*, (2013); 500: 541–546.
 43. Cotillard, A., Kennedy, S.P., Kong, L.C., Prifti, E., Pons, N., Le Chatelier, E., *et al.*; Dietary intervention impact on gut microbial gene richness; *Nature*, (2013); 500: 585–588.
 44. Krilaviciute, A., Heiss, J.A., Leja, M., Kupcinskas, J., Haick, H., Brenner, H.; Detection of cancer through exhaled breath: a systematic review; *Oncotarget*, (2015); 6: 38643–38657.
 45. Batty, C.A., Cauchi, M., Lourenço, C., Hunter, J.O., Turner, C.; Use of the analysis of the volatile faecal metabolome in screening for colorectal cancer; *PLoS One*, (2015); 10: e0130301.
 46. Di Lena, M., Porcelli, F., Altomare, D.F.; Volatile organic compounds as new biomarkers for colorectal cancer: a review; *Colorectal Disease*, (2016); 18: 654–663.
 47. Altomare, D.F., Di Lena, M., Porcelli, F., Travaglio, E., Longobardi, F., Tutino, M., *et al.*; Effects of curative colorectal cancer surgery on exhaled volatile organic compounds and potential implications in clinical follow-up; *Annals of Surgery*, (2015); 262: 862–866. discussion 866–867.
 48. Chen, L., Zhang, C., Gui, Q., Chen, Y., Yang, Y.; Ultraperformance liquid chromatography coupled with quadrupole timeoflight mass spectrometrybased metabolic profiling of human serum prior to and following radical resection of colorectal carcinoma; *Molecular medicine reports*, (2015); 12: 6879–6886.
 49. Fuchs, P., Loeseken, C., Schubert, J.K., Miekisch, W.; Breath gas aldehydes as biomarkers of lung cancer; *International Journal of Cancer*, (2010); 126: 2663–2670.
 50. Rocha, S.M., Caldeira, M., Carrola, J., Santos, M., Cruz, N., Duarte, I.F.; Exploring the human urine metabolomic potentialities by comprehensive two-dimensional gas chromatography coupled to time of flight mass spectrometry; *Journal of Chromatography. A*, (2012); 1252: 155–163.
 51. Phillips, M., Cataneo, R.N., Cummin, A.R.C., Gagliardi, A.J., Gleeson, K., Greenberg, J., *et al.*; Detection of lung cancer with volatile markers in the breath; *Chest*, (2003); 123: 2115–2123.
 52. Wojciech, F., Anna, F., Andreas, S., Thomas, S., Bettina, Z., Clemens, A., *et al.*; Comparative analyses of volatile organic compounds (VOCs) from patients, tumors and transformed cell lines for the validation of lung cancer-derived breath markers; *Journal of Breath Research*, (2014); 8: 027111.
 53. Hanai, Y., Shimono, K., Oka, H., Baba, Y., Yamazaki, K., Beauchamp, G.K.; Analysis of volatile organic compounds released from human lung cancer cells and from the urine of tumor-bearing mice; *Cancer Cell International*, (2012); 12: 7.
 54. Yannai, S.; *Additives, flavors, and ingredients*. Chapman & Hall/CRC, Boca Raton, (2004).
 55. Park, S.Y., Yoo, S.S., Uh, J.H., Eun, J.B., Lee, H.C., Kim, Y.J., *et al.*; Evaluation of lipid oxidation and oxidative products as affected by pork meat cut, packaging method, and storage time during frozen storage (–10°C); *Journal of Food Science*, (2007); 72: C114–C119.

56. WCRF/AICR: 7.9 colon and rectum. In: Food, nutrition, physical activity and the prevention of cancer: Global perspective. Washington dc, aicr, (2007):280–288.
57. Cordero, P., Principe, A., Jofre, E., Mori, G., Fischer, S.; Inhibition of the phytopathogenic fungus *Fusarium proliferatum* by volatile compounds produced by *Pseudomonas*; *Archives of Microbiology*, (2014); 196: 803–809.
58. Selvin, J., Shanmughapriya, S., Gandhimathi, R., Seghal Kiran, G., Rajeetha Ravji, T., Natarajaseenivasan, K., *et al.*; Optimization and production of novel antimicrobial agents from sponge associated marine actinomycetes nocardiosis dassonvillei MAD08; *Applied Microbiology and Biotechnology*, (2009); 83: 435.
59. Jisha, M., Zeinul Hukuman, N.H., Leena, P.; GC–MS analysis of leaves and flowers of *Pogostemon quadrifolius* (benth.) F. Muell. (Lamiaceae); *World Journal of Pharmaceutical Research*, (2016); 5: 667–681.
60. Adhoni, S.A., Thimmappa, S.C., Kaliwal, B.B.; Phytochemical analysis and antimicrobial activity of *Chorella vulgaris* isolated from unkal lake; *Journal of Coastal Life Medicine*, (2016); 4: 368–373.
61. Koverzanova, E.V., Usachev, S.V., Shilkina, N.G., Lomakin, S.M., Gumargalieva, K.Z., Zaikov, G.E.; Specific features of thermal degradation of polypropylene in the presence of magnesium hydroxide; *Russian Journal of Applied Chemistry*, (2004); 77: 445–448.
62. Schmidt, K., Podmore, I.; Current challenges in volatile organic compounds analysis as potential biomarkers of cancer; *Journal of Biomarkers*, (2015); 2015: 981458.
63. Kyle, J.E., Casey, C.P., Stratton, K.G., Zink, E.M., Kim, Y.-M., Zheng, X., *et al.*; Comparing identified and statistically significant lipids and polar metabolites in 15-year old serum and dried blood spot samples for longitudinal studies; *Rapid communications in mass spectrometry: RCM*, (2017); 31: 447–456.
64. Hough, R., Archer, D., Probert, C.; A comparison of sample preparation methods for extracting volatile organic compounds (VOCs) from equine faeces using HS-SPME; *Metabolomics: Official journal of the Metabolomic Society*, (2018); 14: 19.
65. Sin, D.W., Wong, Y.C., Sham, W.C., Wang, D.; Development of an analytical technique and stability evaluation of 143 C3–C12 volatile organic compounds in summa canisters by gas chromatography-mass spectrometry; *The Analyst*, (2001); 126: 310–321.
66. Kang, S., Thomas, C.L.P.; How long may a breath sample be stored for at -80°C ? A study of the stability of volatile organic compounds trapped onto a mixed Tenax: Carbograph trap adsorbent bed from exhaled breath; *Journal of Breath Research*, (2016); 10: 026011.
67. Esfahani, S., Sagar, N., Kyrou, I., Mozdiak, E., O'Connell, N., Nwokolo, C., *et al.*; Variation in gas and volatile compound emissions from human urine as it ages, measured by an electronic nose; *Biosensors*, (2016); 6: 4.
68. Shu, X., Xiang, Y.-B., Rothman, N., Yu, D., Li, H.-L., Yang, G., *et al.*; Prospective study of blood metabolites associated with colorectal cancer risk; *International Journal of Cancer*, (2018); 143: 527–534.