

Article

LC-MS/MS Based Volatile Organic Compound Biomarkers Analysis for Early Detection of Lung Cancer

Shuaibu Nazifi Sani ¹, Wei Zhou ², Balarabe B. Ismail ³ , Yongkui Zhang ⁴, Zhijun Chen ⁴, Binjie Zhang ⁴, Changqian Bao ⁵, Houde Zhang ^{6,*} and Xiaozhi Wang ^{1,*}

¹ College of Information Science & Electronic Engineering, Zhejiang University, Hangzhou 310027, China

² Biochemical Analysis Laboratory, Breath (Hangzhou) Technology Co., Ltd., Hangzhou 310000, China

³ College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China

⁴ Zhejiang Zhoushan Hospital, Zhoushan 316021, China

⁵ Department of Hematology, The Second Affiliated Hospital, College of Medicine Zhejiang University, Hangzhou 310009, China

⁶ Department Gastroenterology, Nanshan Hospital, Guandong Medical University, Shenzhen 518052, China

* Correspondence: szkjk@126.com (H.Z.); xw224@zju.edu.cn (X.W.)

Simple Summary: In our work, we described and proposed a novel LC-MS/MS based approach for early lung cancer screening based on the VOC marker. Through this approach, two key volatile organic compounds (VOCs) used in the early screening and diagnosis of lung cancer were determined as 3-hydroxy-2-butanone and 2-pentanone. 3-hydroxy-2-butanone in the human body was found for the first time to be derived from the sugar metabolism of oral bacteria and we found a correlation between the bacteria that metabolize 3-hydroxy-2-butanone and lung cancer.

Abstract: (1) Background: lung cancer is the world's deadliest cancer, but early diagnosis helps to improve the cure rate and thus reduce the mortality rate. Annual low-dose computed tomography (LD-CT) screening is an efficient lung cancer-screening program for a high-risk population. However, LD-CT has often been characterized by a higher degree of false-positive results. To meet these challenges, a volatolomic approach, in particular, the breath volatile organic compounds (VOCs) fingerprint analysis, has recently received increased attention for its application in early lung cancer screening thanks to its convenience, non-invasiveness, and being well tolerated by patients. (2) Methods: a LC-MS/MS-based volatolomics analysis was carried out according to P/N 5046800 standard based breath analysis of VOC as novel cancer biomarkers for distinguishing early-stage lung cancer from the healthy control group. The discriminatory accuracy of identified VOCs was assessed using subject work characterization and a random forest risk prediction model. (3) Results: the proposed technique has good performance compared with existing approaches, the differences between the exhaled VOCs of the early lung cancer patients before operation, three to seven days after the operation, as well as four to six weeks after operation under fasting and 1 h after the meal were compared with the healthy controls. The results showed that only 1 h after a meal, the concentration of seven VOCs, including 3-hydroxy-2-butanone (TG-4), glycolaldehyde (TG-7), 2-pentanone (TG-8), acrolein (TG-11), nonaldehyde (TG-19), decanal (TG-20), and crotonaldehyde (TG-22), differ significantly between lung cancer patients and control, with the invasive adenocarcinoma of the lung (IAC) having the most significant difference. (4) Conclusions: this novel, non-invasive approach can improve the detection rate of early lung cancer, and LC-MS/MS-based breath analysis could be a promising method for clinical application.



Citation: Sani, S.N.; Zhou, W.; Ismail, B.B.; Zhang, Y.; Chen, Z.; Zhang, B.; Bao, C.; Zhang, H.; Wang, X. LC-MS/MS Based Volatile Organic Compound Biomarkers Analysis for Early Detection of Lung Cancer. *Cancers* **2023**, *15*, 1186. <https://doi.org/10.3390/cancers15041186>

Academic Editors: Jelena Knezevic, Luka Brcic and András Khoór

Received: 5 January 2023

Revised: 2 February 2023

Accepted: 6 February 2023

Published: 13 February 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/>).

Keywords: volatile organic compounds; biomarker analysis; early screening; lung cancer; LC-MS/MS analysis; exhaled breath

1. Introduction

Lung cancer is among the top three cancers affecting individuals irrespective of gender and is responsible for cancer-related deaths globally. However, more than two-thirds of lung cancer incidences are often discovered at late and incurable stages due to early-stage unrecognizable symptoms. Nevertheless, only 4.7% of patients with metastatic stage IV lung cancer survive for five years, compared to 56.3% of patients with stage I cancer [1,2]. Early diagnosis and prompt surgery help to improve the cure rate and thus reduce the mortality rate. Annual low-dose computed tomography (LD-CT) screening is an efficient lung cancer-screening program for a high-risk population [3,4]. However, LD-CT has often been characterized by a higher degree of false-positive results and an elevated risk of misdiagnosis. To meet these challenges, volatolomics, which encompass the analysis of volatile organic compounds (VOCs) emanating from cancer cells that can be detected in various body fluids, including blood, urine, exhaled breath, and sweat, presents an auspicious development in cancer biomarker-based approaches. The main characteristic of cancerous tumor is the uncontrolled growing of cells inside the organism, changes in the body's metabolism, and the outrush of VOC emissions. Likely, these uncontrolled growths are, consequently, changes in the metabolism that could be analytically monitored [5]. Depending on where the cancer cells are located, they provide different characteristic profiles. These profiles, as fingerprints, are used for differentiation in a comparison to normal cells [6]. In particular, the breath VOC fingerprint analysis has recently received increased attention for its application in early lung cancer screening thanks to its convenience, non-invasiveness, and being well tolerated by patients [7]. Inspired by the unusual smell of diabetes, liver cirrhosis, renal failure, and other diseases, S. M. Gordan successfully detected VOCs in exhaled breath to diagnose lung cancer in 1985. The results showed that the VOCs in the exhaled breath of 12 patients with advanced lung cancer were different from those of 17 healthy people, and the diagnostic model constructed by acetone, methyl ethyl ketone, and *n*-propanol can completely distinguish the lung cancer group from the normal group [8]. This pioneering research using VOC fingerprint analysis to diagnose lung cancer has gradually attracted more researchers' attention. Besides, it has been explored in diagnosing other cancers and infectious diseases, such as breast cancer, COPD, asthma, etc. [9–13].

So far, about 83-lung cancer VOCs have been identified [14,15]. However, no single VOC or VOC fingerprint model has been approved for clinical application yet due to several shortcomings, including poor reproducibility, making large-scale verification very difficult. However, previous results suggested that, for the same illness, the repetitiveness of VOC markers screened between different laboratories or even in the same laboratory at different periods was relatively low [16,17]. This resulted in the absence of a unified standard for sample processing, instrumentation, data analysis, and quantitative verification method [18]. Mass spectrometry (MS) has emerged as the most common detection technique widely used to detect VOCs. The LC-MS/MS system, in particular, is a powerful analytical tool for the detection, confirmatory identification, and quantitation of various molecules with myriad applications in numerous research fields [19]. The LC-MS/MS system comprised three key parts: an ion source, mass analyzer, and a detector with electrospray ionization (ESI), atmospheric pressure ionization, and atmospheric pressure chemical ionization (APCI) being the commonly used ion sources. Compared with GC-MS, the LC-MS-based analysis is affordable, and it requires less complicated sample preparations, and it possesses high resolution and fast analyte detection speed. Besides, the number of metabolites identified by the GC-MS study are limited to about 100, while LC-MS/MS could identify nearly 500 metabolites. Besides, LC-MS/MS techniques has demonstrated high separation efficiency for thousands of metabolites in various samples without loss of sensitivity [20,21]. Despite the wide application of mass spectrometry, for detecting VOCs, its application in breath VOCs fingerprint analysis for early detection of lung cancer biomarkers has not been reported thus far.

In the present study, we hypothesized that clarifying the source of VOCs and confirming their correlation with tumors is a prerequisite for establishing repeatable and standardized testing and diagnosis. A standard approach could be used to screen early lung cancer patients based on VOC biomarkers using a novel sensing technique, liquid chromatography with tandem mass spectrometry (LC-MS/MS), to enhance the sensitivity and reliability of qualitative results. The detection objects are 25 carbonyl VOCs with stable derivatization reactions and LC-MS/MS analysis methods, and 4 carbonyl VOCs cannot be detected and were not displayed (Table S1). In addition to looking for differences in the exhaled breath of lung cancer patients, we also studied the possible metabolic sources and metabolic mechanisms of 3-Hydroxy-2-butanone and 2-Pentanone (TG-4 and TG-8) in the exhaled breath of lung cancer patients, and we found that TG-4 is mainly derived from the sugar metabolism of oral bacteria in lung cancer patients, and TG-8 may be derived from lipid metabolism in the human body. The changes in oral bacteria in lung cancer patients were analyzed by 16SrRNA".

2. Materials and Methods

2.1. Materials

2,4-Dinitrophenylhydrazine (DNPH) was purchased from Rhawn (Shanghai, China), with purity $\geq 99\%$. Methanol and acetonitrile of liquid chromatographic grade (purity $\geq 98\%$) were purchased from Sigma-Aldrich (Shanghai, China). Hydrochloric acid and silica gel powder of 300–400 mesh were acquired from Sinopharm Chemical Reagent (Beijing, China). The standard solution was composed of: 2-butanone-DNPH, butyraldehyde-DNPH, pentanal-DNPH, pentanone-DNPH, acrolein-DNPH, acetaldehyde-DNPH, butenal-DNPH, furfural-DNPH, formaldehyde-DNPH, hexaldehyde-DNPH, benzaldehyde-DNPH, 2-crotonaldehyde-DNPH, formylfuran-DNPH, heptanal-DNPH, octanal-DNPH, nonanal-DNPH, decanal-DNPH, and cyclohexanone-DNPH, which were purchased from A Chemtek; additionally, acetoin-DNPH, 2-pentanone-DNPH, and 4-heptanone-DNPH, with a purity of above 98%, were acquired from TanMo Quality Testing Co., Ltd. (Beijing, China). A blank solid-phase extraction (SPE) column without any other filler applied was purchased from Biocomma Limited Co., Ltd. (Shenzhen, China). A 5 L Tedlar bag was purchased from Sigma-Aldrich (Shanghai, China), and prior to their use, bags were cleaned with N_2 at a high temperature ($80^\circ C$) until carbonyl compounds were free. Ultrapure water was used for all aqueous solution preparation. A detailed procedure for the solid phase column extraction is presented in Supplementary material.

2.2. Study Participants

Breath samples were collected from the cardiothoracic surgery department of Zhoushan hospital (Zhoushan city, Zhejiang Province, China). Approval was obtained from the ethics committee of the Zhoushan hospital (2019/208). All individuals involved in the study were formally enrolled, signed written documents, and interviewed regarding health records and smoking habits. All lung cancer patients were in the early stage of lung cancer and were diagnosed by imaging techniques (computed tomography and positron emission tomography) and histology (biopsy). Participants with a history of malignancy and ventilation dysfunction were excluded. Breath samples (fasting and 1 h after a meal) were collected from lung cancer patients before surgery, three to seven days after surgery, and four to six weeks after surgery. The control group consisted of healthy volunteers without pulmonary disease history. Subjects were asked not to engage in vigorous physical activity before sample collection for two days. After dinner on the night before collection, subjects were asked to abstain from food and drink (except water) and smoking. During sampling, the volunteers blew air into the Tedlar bags (>3 L) while sitting in a specific chamber.

2.3. Breath Sample Processing and Analysis

During the sampling process, the volunteer sits in the sampling room and blows into the sampling bag (>3 L). The bags were stored at ambient temperature ($22 \pm 4^\circ C$) in the

dark, and the analysis was conducted within 24 h after collection, minimizing contamination or loss of compounds. During treatment, the sampling bag and DNPH-coated silica gel solid-phase extraction column were placed in an oven at 50 °C for 30 min. Then, the air outlet of the sampling bag was connected with the lower end of the adsorption tube using a catheter, and the upper end of the DNPH-coated silica gel solid-phase extraction column was connected to the suction device. The pumping rate was set at 50 mL/min, and the pumping volume was set at 3 L. After the extraction, 2 mL acetonitrile was added from the adsorption tube's upper end in successive drops. After natural or pressurized elution, the eluent was collected, diluted to 2 mL with acetonitrile, and filtered using a 0.22 µm PTFE needle filter to prepare the sample.

2.4. Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis

The LC-MS/MS analysis was carried out according to P/N 5046800 standard using (AB SCIEX 4500MD Mass Spectrometer, Framingham, MA, USA). The separation and analysis conditions were as follows: Thermo column-Acclaim™ Carbonyl C18 (3 µm, 120 Å, 3.0 × 150 mm); mobile phase: 5 mmol tendency for ammonium acetic acid aqueous solution (A)—acetonitrile (B); gradient elution program: 0–3 min, 50% B, 3–5 min, 50–60% B, 5–8 min, 60–70% B, 8–9.5 min, 70% B, 9.5–10 min, 70–80% B, 10–11 min, 80% B, 11–11.5 min, 80–85% B, 11.5–12.5 min, 85% B, 12.5–13 min, 85–90% B, 13–14.5 min, 90% B, 14.5–15.5 min, 90–95% B, 15.5–16 min, 95% B, 16–18 min, 95–50% B, 18–20 min. 50% B. Total flow rate: 0.4 mL/min; column temperature: constant temperature 40 °C; sample room temperature: constant temperature 10 °C.

A negative mode was used for operating the electrospray source. The curtain, nebulizer, auxiliary, and collision gas (nitrogen) were used. The ion transfer voltage and the probe temperature parameter were set to −4500 V and 500 °C, respectively. Sample analysis was carried out in MRM mode, with a dwell time of 25 ms per channel. The infusion of standard solutions and source-dependent parameter optimizations were performed. The most and second-most sensitive fragment signals were selected as the quantification and confirmation ion pairs. The detailed parameters of the electrospray source for optimization analysis are summarized in Table S2.

2.5. Quality Control

Experimental materials, including columns, were sourced in bulk from a single lot. Reagents and consumables were tested prior to use. The quality control samples were spaced equally among the injections for each day, and all experimental samples were arbitrarily distributed throughout each day's run. To monitor data quality and process variation, a plurality of quality control points was arranged in parallel and evenly distributed among the experimental samples. Through injection of an authentic standard, limit of detection (LOD), limit of quantification (LOQ), and Relative Standard Deviation (RSD) analyses, the proposed High-performance liquid chromatography (HPLC) quantitative analysis method was verified. Linear regression with the peak area (Y) of each analyte and the corresponding mass concentration (X, µg/L) was performed to obtain the standard curve and correlation coefficient (R) of 21 VOCs. A good linearity of each calibration curve was achieved (R > 0.99). The LOD and LOQ are shown in Table S3 based on the content of the target compound when the characteristic ion chromatographic peak S/N ≥ 3. The RSD values of intra-day and inter-day variability, repeatability, and stability of 21 VOCs were all less than 5%. The results show that the established method can be used for the determination of 21 VOCs in breath samples.

2.6. 16S rRNA

Saliva samples for genomic DNA extraction were randomly collected from some participants before breakfast, gargled with normal saline at 15 mL, collected into a 50 mL conical tube, and stored at −80 °C until further processing. Genomic DNA was extracted for HiSeq PE250 sequencing. Based on the Illumina HiSeq sequencing platform, according

to the amplified 16S region characteristics, paired-end sequencing was used to construct a small fragment library for paired-end sequencing. Genome analysis was performed based on sequencing results.

2.7. Statistical Analysis

GraphPad Prism 7 and SPSS 22 were used for graphic presentation and statistical analysis. *t*-tests were used to compare the characteristic of the study population of lung cancer patients with healthy controls. The data are presented as the mean $\pm$ standard error, and the *p*-value of <0.05 was used to assign statistical significance. The differences in VOC concentrations between the different groups were shown using box and whisker plots, and the statistical significance of the differences was assessed using a non-parametric Kruskal-Wallis test to count the non-normality of the distribution. To evaluate the differentiation performance (specificity and sensitivity), the area under the curve (AUC) of the receiver operating characteristic (ROC) curves was used. A diagnostic prediction model was created for early lung cancer versus controls using a decision tree—random forest analysis based on the subset of VOCs that are statistically relevant for discrimination.

The lung cancer prediction model was established based on the random forest algorithm, and indicators, such as gender, age, and VOCs, were selected as model features. The data set is divided into a training set (1/2) and a validation set (1/2). Among them, the out-of-bag data generated during the random forest training process, which account for about 1/3 of the training set, are used as the test sets. Input the training set data into the random forest model modeling (based on the sklearn framework); the prediction accuracy of the output of random forest out-bag test set is used as input for the training set data in the random forest model modelling. Input the test set data into the built model, and output the test set model confusion matrix to predict its sensitivity and specificity.

3. Results

3.1. Effect of Ambient Air

In this study, the long-term dynamic changes in the environment and the effect of respiration on volunteers were investigated. All human exhaled air samples in this experiment are from oral exhalation. A total of 21 carbonyl VOCs were unambiguously identified and quantified by LC-MS/MS analysis. The differences between exhaled breath and ambient air were shown in Figure S1. Among the 21 carbonyl VOCs, 10 carbonyl VOCs in the environment show higher concentrations than the human exhaled breath, which may impact the discriminant analysis of cancer patients' exhaled breath.

Moreover, the concentrations of TG-1, TG-24, and TG-25 in different locations were different during the experiment. The concentrations of TG-1 and TG-24 were also different in the same place at different times (Figure S2). Besides, TG-24 at varying concentrations was found in different brands of syringes and organic solvents used in the experiment. This was consistent with previous reports that TG-24 and TG-25 are released in some medical equipment and pipelines [22]. Therefore, TG-1, TG-24, and TG-25 were excluded to avoid interference from ambient air and medical instruments. The concentration of 11 carbonyl VOCs in the environment is similar to or lower than that of human exhaled breath. Among them, TG-4, TG-8, TG-10, and TG-13 in human exhaled breath were significantly higher than in ambient air. Therefore, the 11 carbonyl VOCs were analyzed and finally assessed for their potential relationship with lung cancer.

3.2. Carbonyl VOCs in Exhaled Breath of Early Lung Cancer Patients

Through the pre-experiment involving 40 lung cancer patients and 40 control individuals, exhaled breath was collected at fasting, 1 h after the meal, and 2 h after the meal. The results showed that compared with the control group, TG-4 in the exhaled breath 1 h after the meal was significantly higher among lung cancer patients, and TG-8 was significantly higher at any time (Figure S3). Therefore, subsequent experiments sampled exhaled breath for VOC analysis after morning fasting and 1 h after the meal from early

lung cancer patients before surgery, three to seven days after surgery, and four to six weeks after surgery. A total of 386 people participated in the 17-month study: 291 early lung cancer patients and 95 healthy controls. Table 1 summarizes the clinical characteristics of this population. The lung cancer group and the control group differ significantly ($p < 0.05$) according to age and gender. However, the smoking habit was not significantly different among the sample population.

Table 1. Characteristics of the study population.

	Lung Cancer Patients (n = 291)	Healthy Controls (n = 95)	χ^2/t	<i>p</i> Value
Mean age (years)	60 (19–89)	50 (23–81)	6.917	0.000
Proportion (man)	116 (40%)	27 (29%)	4.02	0.044
Smoking behavior			0.793	0.370
Active	19 (7%)	6 (7%)		
Former	22 (7%)	4 (4%)		
Never	250 (86%)	85 (89%)		
Histologic classification				
Adenocarcinoma	251 (86%)			
Squamous carcinoma	11 (4%)			
Others	29 (10%)			

χ^2/t : Chi-square value divided by *t*-test value.

The difference in carbonyl VOCs in the exhaled breath between the control group and early lung cancer patients before surgery, three to seven days after surgery, and four to six weeks after surgery was mainly at 1 h after the meal. The statistical data are summarized in Table S7 and the results in Figure 1 showed significant differences in the concentrations of 7 VOCs. TG-4 and TG-8 in the exhaled breath of patients with early lung cancer were significantly higher before surgery but decreased after surgery. Nevertheless, both before and after surgery, TG-7, TG-11, TG-19, TG-20, and TG-22 were continuously low. Figure 1 depicts the ROC curve analysis results of the 7 VOCs. It is clear that TG-4 was the VOC with the most significant difference, and its sensitivity and specificity were 70% and 76%, respectively. However, the sensitivity and specificity of a single marker for screening are not yet sufficient for clinical applications. Therefore, through the comprehensive analysis of seven markers (decision tree-random forest discriminant analysis), the sensitivity and specificity were 83% and 78%, respectively. By ignoring the environmental impact, the sensitivity and specificity for the comprehensive analysis of 21 detectable VOCs were 97% and 100%, respectively.

Early lung cancer manifestations are divided into different stages: adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA), or invasive adenocarcinoma (IAC). The AUC for the ROC curves to discriminate AIS/MIA/IAC from controls were shown in Table S4 and Figure 2. Approximately 86% of patients with early lung cancer have lung adenocarcinoma, and early lung adenocarcinoma was also divided into different stages of AIS/MIA/IAC. The AUC value of ROC curve analysis to distinguish AIS/MIA/IAC from the control group is shown in Table S4. Compared with the control group, the difference in IAC was the most significant. The sensitivity and specificity of the random forest discriminant analysis of seven markers were 96% and 73%, respectively. Ignoring the environmental impact, the sensitivity and specificity of the random forest discriminant analysis of 21 detectable VOCs were 96% and 86%, respectively.

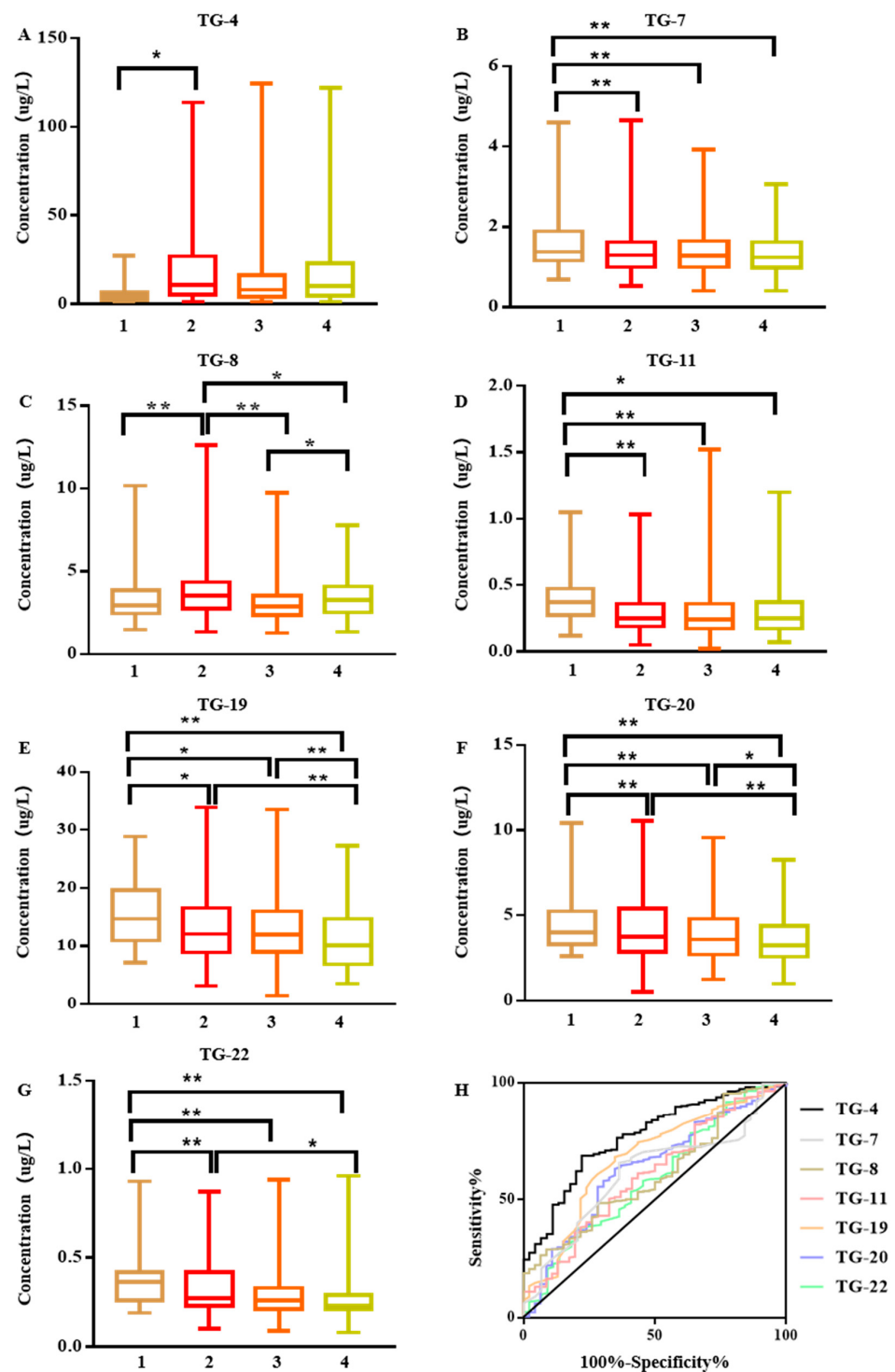


Figure 1. (A–G), 1, control; 2, lung cancer before the operation; 3, lung cancer after the operation after three to seven days; 4, lung cancer after the operation after four to six weeks. Box and whisker plots of the measured concentrations of TG-4, TG-7, TG-8, TG-11, TG-19, TG-20, and TG-22 in the exhaled breath 1 h after meals of the study groups, (least significant difference test) LSD-t analysis of variance. * $p < 0.05$; ** $p < 0.01$. (H), ROC curve analysis was performed on the exhaled air 1 h after the meal in the control group and the lung cancer group before the operation. ROC curves for the TG-4, TG-7, TG-8, TG-11, TG-19, TG-20, and TG-22 breath model in the diagnosis of lung cancer. The area under the curve of TG-4, TG-7, TG-8, TG-11, TG-19, TG-20, and TG-22 were 0.766, 0.611, 0.636, 0.682, 0.604, 0.611, and 0.625, respectively.

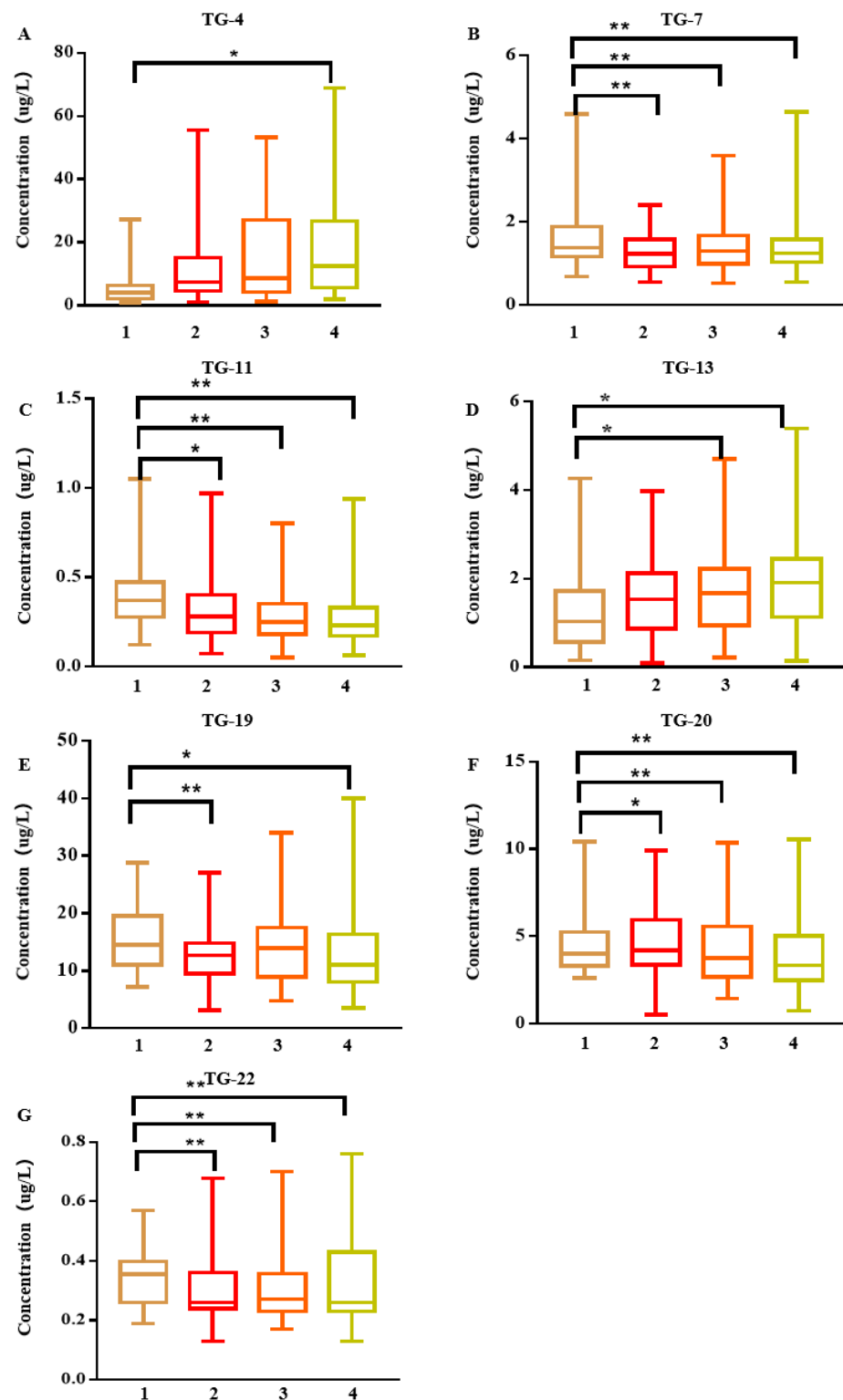


Figure 2. (A–G), 1, control; 2 AIS; 3, MIA; 4, IAC. Box and whisker plots of the measured concentrations of TG-4, TG-7, TG-11, TG-13, TG-19, TG-20, and TG-22 in the exhaled breath 1 h after meals of the study groups, (least significant difference test) LSD-t analysis of variance. * $p < 0.05$; ** $p < 0.01$. (NB: All the data from AIS, MIA, and IAC are before surgery).

3.3. Effects of Diet on Exhaled Breath Carbonyl VOCs

Diet specifically affects carbonyl VOC concentration in human exhaled breath, but studies on how it affects the carbonyl VOCs in human exhaled breath are scarce. Therefore,

in this study, changes in carbonyl VOC concentration in human exhaled breath after eating some common foods (the methods and a list of foods are shown in Supplementary Data and Table S5) were investigated. The results showed that eating protein and lipids does not affect the concentration of 11 carbonyls VOCs in the exhaled breath. However, eating sugar-containing foods causes a significant change in the concentration of TG-4 in the exhaled breath (Figure 3A–C). Besides, after eating a banana, the concentration of TG-8 in the exhaled breath also increased, with the same trend as TG-4, returning to normal levels within 1 h. After drinking 50 mL of a characteristic Chinese liquor, TG-8 and TG-10 in the exhaled breath can reach high concentrations within 5 min and fail to return to normal levels within 2 h.

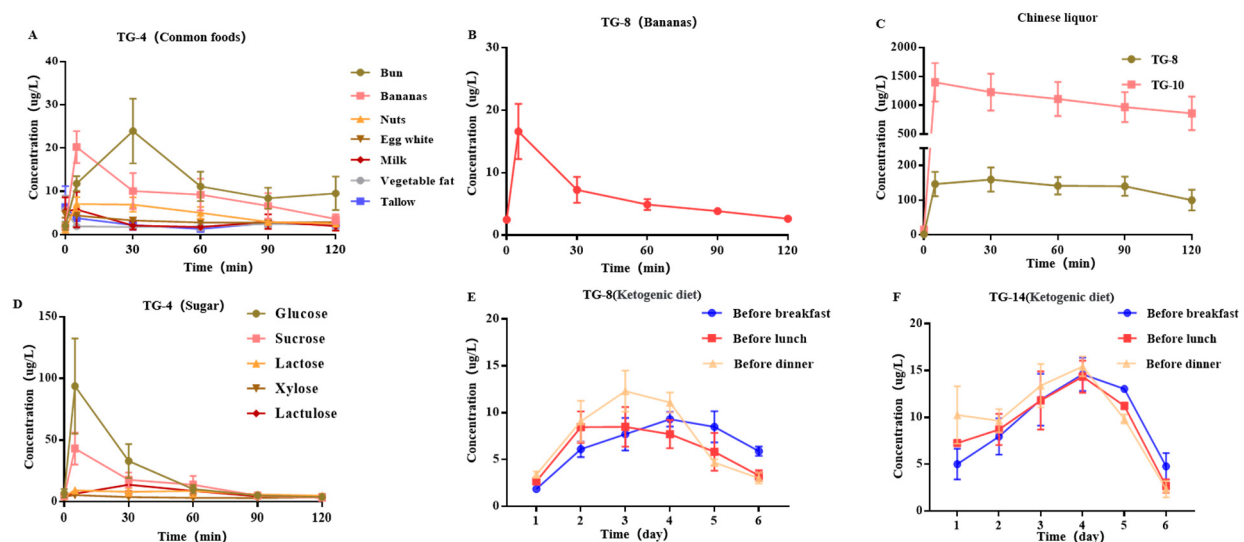


Figure 3. Changes in VOC concentration in exhalation after different diets. (A–C). Changes in carbonyl VOCs' concentration in human exhaled breath after eating some common foods. (D) Changes in the concentration of TG-4 in the exhaled breath after drinking different sugar solutions (E,F). Values and error bars represent mean $\pm$ SE.

Because of the change in TG-4 in exhaled breath after eating sugar-containing foods, the changes in exhaled breath after drinking a variety of sugar solutions were studied (Figure 3D). The results showed that, after drinking different sugar solutions, the TG-4 in the exhaled breath reached a high concentration at 5 min, decreased significantly at 30 min, and recovered to the fasting level within 1 h. According to Pearson correlation analysis, the concentration of TG-4 from the exhaled breath after 5 min of drinking sugar solution was positively correlated with the glycemic index (GI) ($r = 0.975$, $p = 0.005$, Table S6). It is worth noting that short-term protein and lipid diets have no effect on exhaled breath, but after long-term protein and lipid diets (ketogenic diet), TG-8 and TG-14 concentrations in exhaled breath remained high and decreased only after carbohydrate intake (Figure 3E,F, see the Supplementary File for the experimental method of dietary influence). Therefore, volunteers whose breath was collected needed to regulate their diet.

3.4. Influence of Exhalation Mode on Exhalation Collection

Based on the earlier results on the influence of diet on exhaled air, we speculated that the carbonyl VOCs in the exhaled breath might come from the oral cavity. Therefore, we chose the exhalation collection of the oral cavity and nasal cavity after drinking a 5% glucose solution for verification. The results (Figure 4) suggest that TG-4 and TG-10 concentrations in the oral exhaled breath were significantly higher than in the nasal cavity, regardless of status before or after drinking sugar, indicating that carbonyl VOCs in oral metabolism were the main components in exhaled breath. Besides, the significantly higher

TG-4 in oral exhalation after drinking sugar or oral sugar indicated that TG-4 largely comes from the oral cavity.

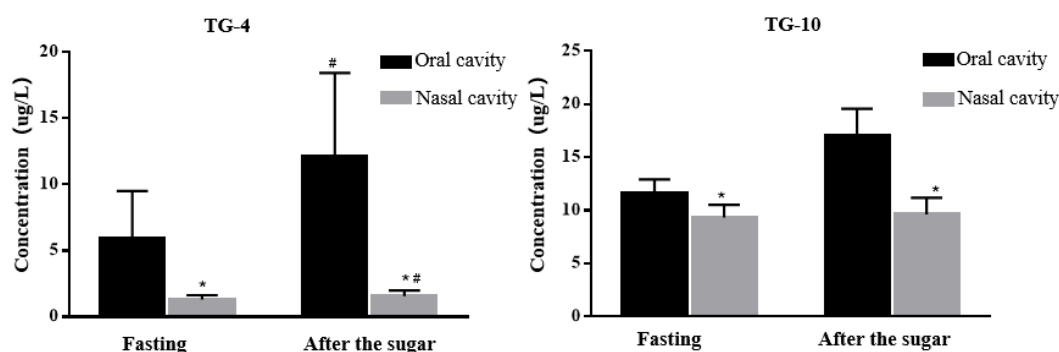


Figure 4. Difference between oral and nasal exhalation. Eight healthy volunteers drank the 5% glucose solution and immediately blew into two sampling bags in parallel with each other in the mouth and nose. Values and error bars represent mean $\pm$ SE. * nasal cavity compared with the oral cavity, $p < 0.05$, and # compared fasting after sugar.

3.5. Mouth Glucose Metabolism

The human mouth houses more than 700 bacterial species, some of which can be metabolized to produce TG-4. The variation in characteristics of TG-4 concentration in the exhaled air was observed after taking sugar in the mouth for 30 s (Figure S4). The concentration of TG-4 in the exhaled breath for all sample populations had an initial increase and then decreased after sugar intake for 30 s, reaching the highest concentration between 5 and 20 min (Figure S4). Compared with direct drinking of glucose solution, the concentration of TG-4 in exhaled breath increased significantly after sugar intake for 30 s, indicating that TG-4 came from the sugar metabolism of oral microorganisms.

3.6. 16S rRNA Detection

To investigate the bacteria in the oral cavity, we randomly sampled the oral saline gargle of 20 lung cancer patients and performed gene sequencing and Spearman correlation coefficient analysis. By testing the composition of microorganisms in oral saliva, the study found that, compared with the healthy control group at the phylum level, the relative abundance of *Elusimicrobia* decreased significantly (Figure 5). At the genera level, lung cancer patients have a significantly higher relative abundance of *Desulfobulbus*, *Family_xiII_UCG-001*, *Gemella*, *Halomonas*, *Klebsiella*, *Roseburia*, *Ruminiclostridi_5*, *Synergistes*, and *Alloscardovia* in their oral saliva than the control group. In contrast, *Alloprevotella*'s relative abundance was significantly lower (Figure 5).

The results showed that the relative abundance of *Streptococcus pharyngitis*, *Porphyromonas pulposus*, *Lactobacillus mucosa*, and *Alloscardovia_omnicolens* positively correlated with TG-4 in the exhaled breath. Moreover, the relative abundance of TM7 animal phylum oral clone DR034 was negatively correlated with exhaled breath TG4 (Table 2), which signaled that TG-4 has a specific correlation with bacterial abundance in the oral saliva. However, how oral bacteria affect the concentration of TG-4 in exhaled breath remains to be determined.

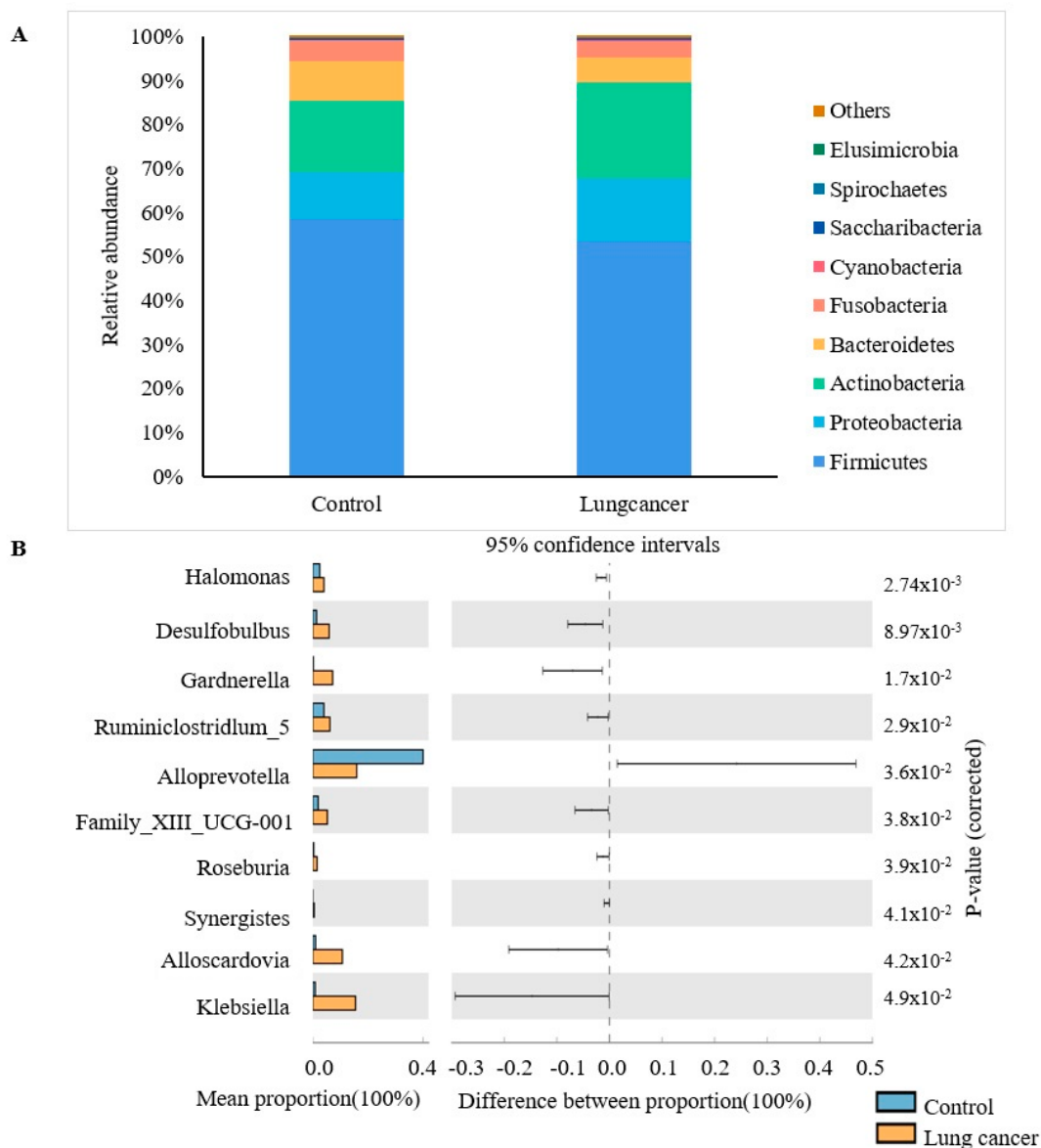


Figure 5. The differences in oral salivary flora between the control group and the lung cancer group. **(A)** At the phylum level, the human oral salivary flora is mainly *Firmicutes*, *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, and *Fusobacteria*, accounting for about 99% of the proportion. Compared with the control group, the proportion of *Elusimicrobia* in the lung cancer group decreased ($p = 0.038$). **(B)** At the genus level, STAMP analyzed the difference in oral saliva between the control group and the lung cancer group.

Table 2. Correlation between TG-4 and oral saliva bacterial abundance.

Positive Correlation	Negative Correlation
Phylum: <i>Spirochetes</i> , <i>Intertrophic bacteria</i> , <i>Pericarpium</i> , <i>Verruca microflora</i>	Phylum: <i>Proteobacteria</i>
Class: <i>Some spirochetes</i> , <i>Mutual nutrient bacteria</i> , <i>Flexible membrane bacteria</i> , <i>Verruca microbe</i>	Class: <i>Saccharibacteria</i>
Order: <i>Spirochetes</i> , <i>Ynergistales</i> , <i>Mycoplasmas</i> , <i>Microflora</i>	Order: <i>Saccharibacteria</i>
Family: <i>Streptococcus</i> , <i>Spirochetes</i>	Family: <i>Saccharibacteria</i>
Genus: <i>Streptococcus</i> , <i>Treponema Pallidum</i>	Genus: <i>Digestive streptococcus</i> , <i>Actinomycetes</i> , <i>Saccharibacteria</i>
Species: <i>Streptococcus pharyngitis</i> , <i>Porphyrin monomonas dental medulla</i> , <i>Lactobacillus mucous</i> , <i>Alloscardovia_omnicolens</i>	Species: <i>TM7 animal phylum oral clone DR034</i>

4. Discussion

The use of breath analysis for cancer diagnosis has been around for more than 30 years since S. M. Gordan's pioneering research in 1985. The lack of standard procedures for breath analysis is responsible for the results' variability and poor repeatability. Table 3 compares the most commonly used VOC biomarker detection techniques from the exhaled breath of lung cancer patients. Gas chromatography coupled mass spectrometry (GC-MS) has been widely used in many research groups [23–29]. Despite being a powerful tool for VOC analyses, its application has been limited by the complex sample preparation and time-consuming analysis, the requirement of trained personnel, and high cost. The other commonly used technique, the headspace temperature programmed vaporization mass spectrometer (HS-PTV-MS), provides a lower limit of detection besides spectral overlap, making compounds' identification difficult [30]. Electronic noses, such as gold nanoparticles (GNP), are based on chemical reactions with sensor materials. They provide highly sensitive results that are easy to interpret, but quantitative results cannot be obtained due to the inability to identify unknown substances [31]. Another promising method [32] for the detection of early lung cancer is using FT-ICR-MS, which significantly elevated the concentration of the detected VOC in lung cancer patients; a statistical classification model was used to analyze how the measured concentration of the carbonyl compound can be used to determine the compounds that are best for distinguishing between lung cancer patients from healthy controls. FT-ICR-MS were used to capture and clarify the molecular formulae of carbonyl compounds (hydroxyl-acetaldehyde and 3-hydroxy-2-butanone) of the trace VOC from exhaled breath in combination with GC-MS for the analysis of the detected VOC [33]. Liquid chromatography and tandem mass spectrometry (LC-MS/MS) are sequential analysis platforms that utilize high-performance liquid chromatography as the separating technique and high-resolution mass spectrometry as the detection system. LC-MS/MS is more suitable for analyzing metabolites with low volatility or poor thermal stability than the above and other analytical techniques.

Table 3. Previous studies on VOC biomarkers in patients with lung cancer.

Study	Method	VOCs	Sensitivity	Specificity	References
Anton et al.	HS-PTV-MS	2-butanone	40%	100%	[23]
Hanai et al.	GC-TOFMS	2-Pentanone	85%	70%	[24]
G. Song et al.	GC-MS	3-Hydroxy-2-butanone	93%	92.7%	[19]
J. Rudnicka et al.	GC-MS	Acetone	74%	73%	[20]
M. Phillips et al.	GC-MS	Pentane	89.6%	82.9%	[21]
N. Peled et al.	GNP sensor	Toluene	70%	100%	[22]
Mingxiao Li et al.	FT-ICR-MS	2-butanone	97%	84%	[32]
Ralph J. Knipp	FT-ICR-MS and GC-MS	Hydroxyl acetaldehyde and 3-hydroxy-2-butanone	94.2 ± 2.5%	-	[33]
This study	LC-MS/MS	3-Hydroxy-2-butanone	96%	73%	-

GC-MS = gas chromatography mass spectrometry, GC-TOFMS = gas chromatography time of flight mass spectrometry, HS-PTV-MS, headspace-programmed temperature vaporizer-mass spectrometer, GNP = gold nanoparticles, LC-MS/MS = liquid chromatography with tandem mass spectrometry.

Multiple candidate markers have been discovered, but they have not yet entered clinical applications. Their common shortcoming was low repeatability, and the underlying mechanism for the observed difference has not been analyzed and discussed in depth. Exhalation analysis is affected by many factors, such as the exhalation environment, exhalation method, daily diet, and internal factors [34]. The current study, therefore, uniquely examined their influence in different ways, analyzed the source of their differences, and devised means to avoid the influence of these differences. Ambient air pollution has a major impact on respiratory analysis. As such, a number of approaches were proposed to reduce the local background: inhaling clean air, balancing ambient air, collecting background samples, and calculating alveolar gradients or setting cut-offs [35–37]. We chose to collect the air in the sampling environment for a long time and found that the environmental changes in the

same place are in a fluctuating stable balance. However, TG-1, TG-24, and TG-25 in ambient air varied greatly and were previously used as markers for disease diagnosis [38]. TG-1 could come from tobacco burning, building materials, and some unknown sources. TG-24 and TG-25 could originate from organic solvents, disinfectants, syringes, hoses, and other experimental and medical supplies. For example, the variation in TG-24 concentration for acetonitrile solvents of different brands can reach 1.5–3.4 fold, while the difference in acetonitrile solvents of the same brand, but being parts of different batches, could be up to 2–7.5 fold. Therefore, their high concentrations for different reasons may lead to different test results. Therefore, to ensure the reliability of the results, we have excluded VOCs in the exhaled breath that are lower than the environment, but these VOCs are not unchanged in the exhaled breath. The high sensitivity and specificity of the random forest discriminant analysis results of 21 VOCs indicated that the VOCs in the exhaled breath lower than the environment are also meaningful. In order to ensure the stability of the exhalation collection, the exhalation collection method is also very important. Some studies used nasal intubation for exhaling, and one-lung ventilation was used during lung tumor resection, but such collection methods are invasive and can cause physical discomfort [39,40]. A variety of non-invasive and convenient sampling methods were tried in this study, such as nasal/oral breath sampling, back-end air sampling, clean room sampling, breath-holding sampling, etc. Interestingly, no significant difference between the selected markers was found, which differs from the conclusions in some papers [38]. It is speculated that the larger number of exhalations collected (>3 L) in the current study was useful in keeping the total amount of markers relatively high, as well as reducing the differences between several exhalation collection methods and the different populations' breathing patterns. Currently, there is little data on respiratory carbonyl VOCs associated with dietary regimes [41]. The results on the effects of the daily diet show that the complexity of the diet can also cause changes in some VOCs in the exhaled breath, and a long-term protein and lipid diet can cause changes in the TG-8 and TG-14 in the exhaled breath. Eating high-sugar foods can also cause changes in some VOCs in the exhaled breath, but they return to normal levels within an hour. There is a significant difference between the exhaled breath of early lung cancer patients and the control group 1 h after eating. Therefore, changes in VOCs in the exhaled breath of lung cancer patients after a unified diet are observed. Strenuous exercise also has a certain impact on the analysis of exhaled breath. After strenuous aerobic or anaerobic exercise, the metabolism of glycogen and fat in the body accelerates, and its metabolites can be excreted through urine, sweat, and exhaled breath. Therefore, strenuous exercise was prohibited two days before sampling [42].

The reliable data of the breath detection analysis were obtained through a series of experiments by eliminating the interfering factors and their influence. Compared with the control group, the seven VOCs with the most significant differences in early lung cancer patients' exhaled breath were TG-4, TG-7, TG-8, TG-11, TG-19, TG-20, and TG-22. TG-4 and TG-8 in the exhaled breath of patients with early-stage lung cancer before surgery were significantly higher but decreased after surgery. However, regardless of before or after surgery, TG-7, TG-11, TG-19, TG-20, and TG-22 were all low. TG-7, TG-11, TG-19, TG-20, and TG-22 belong to the aldehyde group of VOCs derived from several source and pathways, including metabolized alcohols, hydroperoxide reduction by cytochrome p450 as a secondary product of lipid peroxidation, tobacco smoke (saturated and unsaturated aldehydes, e.g., formaldehyde) and detoxification processes of tobacco by-products, and dietary sources [43]. The low level of these aldehydes in lung cancer patients' exhaled breath may be because the tumor inhibits the metabolic pathway of certain aldehyde substances, while the inhibitory effect is sustained and long, and surgical resection of the tumor cannot restore its normal metabolism.

TG-4 and TG-8 are members of the ketone group of VOCs produced in the liver from fatty acids, such as acetone, and they are then oxidized in the Krebs cycle by peripheral tissues. The liver produces acetone, a common human VOC, by decarboxylating surplus acetoacetate from acetyl-CoA. Diet also has an impact on ketone levels in the blood, in-

creasing with the increase in protein or fat metabolism (e.g., notably in cachexia). Based on the possible causes of ketones and the results of ketogenic experiments, we speculated that TG-8 might be mainly produced by fat metabolism in the human body, which is closely related to occurrence and metastasis of tumors. Metabolism in the tumor cell metabolism is far more vigorous than a normal cell. Cancer-related fat cells can produce fatty acids through lipolysis, produced by decomposition. They are delivered to tumor cells as energy sources, and, at the same time, adipokines are secreted to promote tumor cells' adhesion and metastasis. Therefore, it is not surprising to observe an increased TG-8 concentration in the exhaled breath of lung cancer patients. The related metabolic mechanism can return to normal after tumor resection, and the concentration of TG-8 in the exhaled breath also begins to recover.

Compared with the control group, TG-4 was the most different VOC marker in the exhaled breath of the lung cancer group. There was no obvious change in the ketogenic diet experiment, and fat metabolism could not very well explain its primary source. According to the experimental results of oral glucose metabolism, we speculated that TG-4 came from the glucose metabolism of oral bacteria. The results of 16SrRNA analysis showed that the relative abundance of *Desulfobulbus*, *Family_xiii_UCG-001*, *Gemella*, *Halomonas*, *Klebsiella*, *Roseburia*, *Ruminiclostridi_5*, *Synergistes*, and *Alloscardovia* in the oral saliva of lung cancer patients were significantly higher than that of the control group. Therefore, the increased concentration of TG-4 in the exhaled breath of lung cancer patients was probably due to obvious changes in TG-4 metabolizing bacteria in the oral cavity of lung cancer patients. According to existing reports, the microorganisms producing acetoin in the body comprise mainly bacteria, such as *Enterobacter*, *Bacillus*, *Klebsiella*, *Lactococcus*, *Paenibacillus*, *Serratia*, *Acetobacter*, *Aeromonas*, etc. [44–46]. Hence, we speculate that *Klebsiella* may be the main genus that metabolizes TG-4 in the oral cavity of lung cancer patients. The TG-4 in the exhaled breath of lung cancer patients was higher than that in the control group, and it was statistically significant. It may be due to tumor resection and drug treatment to eliminate some related bacteria, so the TG-4 concentration in the exhaled breath of lung cancer patients decreased after surgery. Many studies have proven that oral bacteria in cancer patients are related to cancer occurrence and development. *Porphyromonas gingivalis*, a pathogen of periodontitis, has been proven to promote the occurrence and development of colorectal cancer. The bacteria have also been found in the oral cavity of lung cancer patients. Besides, its abundance positively correlated with TG-4 [47]. Many studies have proven that the structure of the salivary flora of lung cancer patients is similar to that of the lungs [48,49]. Due to time and resource constraints, the sample size of this analysis is small, and it was not possible to analyze whether the same changes have occurred in the lung flora of patients with lung cancer. Therefore, whether the bacteria in the mouth and lungs of lung cancer patients (especially the TG-4 metabolizing bacteria) have the same changes remains to be determined.

The limitations of the current study include fewer samples in the control group, age and gender distribution between control group and lung, and the high detection rate of early lung cancer by breath analysis, which may all be related to the small sample size of the control group. Nevertheless, what cannot be disputed is the difference in VOC trends observed in the exhaled breath between the control group and the early lung cancer group. In comparing the different stages of early lung adenocarcinoma and the control group, the difference in the IAC stage was the most obvious. Using seven significantly different VOC screening analysis methods, IAC has higher sensitivity than all early lung cancer group analyses, indicating that the concentration of VOCs in the exhaled breath is related to the development of lung lesions. Moreover, due to the degree of lung adenocarcinoma infiltration, the change in VOCs in the exhaled breath will become more apparent, and the breath analysis has a higher detection rate for early lung adenocarcinoma.

5. Conclusions

The current study presented a standardized approach that could be used to screen early lung cancer patients based on VOC markers, overcoming the sensitivity and reproducibility challenges associated with conventional testing methods. Using a developed LC-MS/MS acquisition method, the VOC biomarker test may, in the future, be complementary to the imaging, sputum cytology, and biopsy tests following validation using prospective studies. For each lung cancer patient, VOC analysis could be expressed as a chance of discovering the disease early. Nevertheless, although breath VOC analysis could represent a highly probable method to discover lung cancer, yet, it is not conclusive and would require an optimum combination of sensitivity and specificity. In this article, in addition to looking for differences in the exhaled breath of lung cancer patients, we also studied the possible metabolic sources and metabolic mechanisms of TG-4 and TG-8 in the exhaled breath of lung cancer patients. We found that TG-4 is mainly derived from the sugar metabolism of oral bacteria in lung cancer patients, and TG-8 may be derived from lipid metabolism in the human body. The changes in oral bacteria in lung cancer patients were analyzed by 16SrRNA. The sensitivity and specificity of TG-4, with the most significant difference in receiver operating characteristic (ROC) curve analysis, were 70% and 76%, respectively. This novel, non-invasive approach can improve the detection rate of early lung cancer, and further work to improve the suitability of this technique and to assess its limitations in clinical practice would be required.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers15041186/s1>, Figure S1: Differences in the concentration of VOCs in the environment and human breath. (A–D). Among the 21 VOCs, there are 10 VOCs in the environment that were higher than the human exhaled breath, the concentration of the remaining 11 VOCs in the environment is similar to or lower than that of human exhaled breath. Values and error bars represent mean $\pm$ SE. * $p < 0.05$; Figure S2: VOCs concentrations at different locations and at different times. (A). Ambient air was collected multiple times at the same location in different hospitals. Compared with the laboratory environment, TG-1, TG-24, and TG-25 showed different concentrations. (B). The ambient air was collected at the same location in the laboratory at 9:00 and 15:00, and different VOCs in the ambient air also showed different concentrations. Values and error bars represent mean $\pm$ SE. * means compared with laboratory, $p < 0.05$; Figure S3: The preliminary experiment comparing the concentration of VOCs in the exhaled breath between the control group and the lung cancer group. 1, fasting; 2, 1h after the meal; 3, 2 h after the meal; Among the 11 carbonyls VOCs, only TG-4 and TG-8 are different. Compared with the control group, TG-4 in the exhaled breath of lung cancer patients is higher at 1 h after a meal, and TG-8 is higher at 3-time points. * $p < 0.05$; Figure S4: Changes in the concentration of TG-4 in exhaled breath after sugar in the mouth for the 30 s, the color indicates TG-4 concentration pattern of 8 healthy volunteers A. Changes in the concentration of TG-4 in the oral exhaled breath of 8 healthy volunteers after sugar in the mouth for the 30 s. B. The concentration of TG-4 in the exhaled breath of 8 healthy volunteers reached the highest value within 20 min after sugar in the mouth for the 30 s; Table S1: Potential biomarkers for lung cancer; Table S2: MS conditions; Table S3: Linear range, linear equation, correlation coefficient, LOD, LOQ and RSD of 21 VOCs; Table S4: The AUC for the ROC curves to discriminate AIS/MIA/IAC of early lung cancer and control group; Table S5: Different foods and intake quantity; Table S6: The level of TG-4 after the intake of different sugars; Table S7: VOCs Statistical data for control, Lung cancer before the operation, Lung cancer 3–7 days after surgery, and Lung cancer 4–6 weeks after surgery.

Author Contributions: Conceptualization, S.N.S., W.Z. and Y.Z.; methodology, S.N.S., X.W., H.Z. and W.Z.; validation, C.B.; formal analysis, S.N.S., C.B. and W.Z.; data curation, Z.C., B.Z.; writing—original draft preparation, S.N.S. and B.B.I.; writing—review and editing, S.N.S., B.B.I. and H.Z.; supervision, X.W. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work received support from the following programs: National Natural Science Foundation of China (62174147, 21927810), Key Research and Development Project of Zhejiang (2022C01141), Dr. Li Dak Sum & Yip Yio Chin Development Fund for Regenerative Medicine, Zhejiang University.

Institutional Review Board Statement: Approvals for human studies were obtained from the Zhoushan Hospital Ethics Committee (Approval number: 2019-208 and date of approval is 20 November 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Prior to inclusion, all patients and controls of the aforementioned study gave written consent to participate in data collection according to the criteria raised by the regional legal authorities.

Data Availability Statement: The datasets generated and/or analyzed during the current study are not publicly available due to hospital guidelines and legislation regarding personal data. Data will be available from the corresponding author upon reasonable request and with permission of the Zhoushan Hospital Legal Department.

Acknowledgments: The authors would like to thank Xingbin Duan for his technical assistance in Breath (Hangzhou) Technology Co., Ltd.

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

Abbreviations

AIS	Adenocarcinoma in situ
APCI	Atmospheric pressure chemical ionization
AUC	Area under the curve
DNPH	2,4-Dinitrophenylhydrazine
ESI	Electrospray ionization
GC-MS	Gas chromatography mass spectrometry
GC-TOFMS	Gas chromatography time of flight mass spectrometry
GI	Glycemic index
GNP	Gold nanoparticles
HPLC	High-performance liquid chromatography
HS-PTV-MS	Headspace temperature programmed vaporization mass spectrometer
IAC	Invasive adenocarcinoma
LC-MS	Liquid chromatography–mass spectrometry
LC-MS/MS	Liquid chromatography and tandem mass spectrometry
LD-CT	Low dose computed tomography
LOD	Limit of detection
LOQ	Limit of quantification
LSD-t	Least significant difference test
MIA	Minimally invasive adenocarcinoma
MRM	Multiple reaction monitoring
MS	Mass spectrometry
PTFE	Polyfluortetraethylene
R	Correlation coefficient
ROC	Receiver operating characteristic
RSD	Relative standard deviation
S/N	Signal-to-noise ratio
SPE	Solid-phase extraction
VOC	Volatile organic compound

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [[CrossRef](#)] [[PubMed](#)]
2. Henley, S.J.; Ward, E.M.; Scott, S.; Ma, J.; Anderson, R.N.; Firth, A.U.; Thomas, C.C.; Islami, F.; Weir, H.K.; Lewis, D.R.; et al. Annual report to the nation on the status of cancer, part I: National cancer statistics. *Cancer* **2020**, *126*, 2225–2249. [[CrossRef](#)] [[PubMed](#)]
3. Tanoue, L.T.; Tanner, N.T.; Gould, M.K.; Silvestri, G.A. Lung cancer screening. *Am. J. Respir. Crit. Care Med.* **2015**, *191*, 19–33. [[CrossRef](#)]
4. Karachaliou, N.; Sosa, A.E.; Rosell, R. Annual or biennial lung cancer CT screening? *J. Thorac. Dis.* **2016**, *8*, 2424–2426. [[CrossRef](#)] [[PubMed](#)]
5. Farnum, A.; Parnas, M.; Hoque Apu, E.; Cox, E.; Lefevre, N.; Contag, C.H.; Saha, D. Harnessing insect olfactory neural circuits for detecting and discriminating human cancers. *Biosens. Bioelectron.* **2023**, *219*, 114814. [[CrossRef](#)]

6. Saidi, T.; Moufid, M.; de Jesus Beleño-Saenz, K.; Welearegay, T.G.; El Bari, N.; Jaimes-Mogollon, A.L.; Ionescu, R.; Bourkadi, J.E.; Benamor, J.; El Ftouh, M.; et al. Non-invasive prediction of lung cancer histological types through exhaled breath analysis by UV-irradiated electronic nose and GC/QTOF/MS. *Sens. Actuators B Chem.* **2020**, *311*, 127932. [[CrossRef](#)]
7. Wang, C.; Dong, R.; Wang, X.; Lian, A.; Chi, C.; Ke, C.; Guo, L.; Liu, S.; Zhao, W.; Xu, G.; et al. Exhaled volatile organic compounds as lung cancer biomarkers during one-lung ventilation. *Sci. Rep.* **2014**, *4*, 7312. [[CrossRef](#)] [[PubMed](#)]
8. Gordon, S.M.; Szldon, J.P.; Krotoszynski, B.K.; Gibbons, R.D.; O'Neill, H.J. Volatile Organic Compounds in Exhaled Air from Patients with Lung Cancer. *Clin. Chem.* **1985**, *318*, 1278–1282. [[CrossRef](#)]
9. 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](#)] [[PubMed](#)]
10. Markar, S.R.; Wiggins, T.; Antonowicz, S.; Chin, S.T.; Romano, A.; Nikolic, K.; Evans, B.; Cunningham, D.; Mughal, M.; Lagergren, J.; et al. Assessment of a noninvasive exhaled breath test for the diagnosis of oesophagogastric cancer. *JAMA Oncol.* **2018**, *4*, 970–976. [[CrossRef](#)] [[PubMed](#)]
11. Skallefold, H.E.; Vallenari, E.M.; Sapkota, D. Salivary Biomarkers in Lung Cancer. *Mediators Inflamm.* **2021**, *2021*, 6019791. [[CrossRef](#)] [[PubMed](#)]
12. Sholl, L.M. Biomarkers of response to checkpoint inhibitors beyond PD-L1 in lung cancer. *Mod. Pathol.* **2022**, *35*, 66–74. [[CrossRef](#)] [[PubMed](#)]
13. Otálora-Otálora, B.A.; Osuna-Garzón, D.A.; Carvajal-Parra, M.S.; Cañas, A.; Montecino, M.; López-Kleine, L.; Rojas, A. Identifying General Tumor and Specific Lung Cancer Biomarkers by Transcriptomic Analysis. *Biology* **2022**, *11*, 1082. [[CrossRef](#)] [[PubMed](#)]
14. Montes, A.G.; Molins López-Rodó, L.; Ramón Rodríguez, I.; Dequigiovanni, G.S.; Nolas Segarra, V.; Sicart, R.M.M.; Hernández Ferrández, J.; José Fibla Alfara, J.; García-Navarro, Á.A. Lung cancer diagnosis by trained dogs. *Eur. J. Cardiothorac. Surg.* **2017**, *52*, 1206–1210. [[CrossRef](#)]
15. De Vries, R.; Muller, M.; Van Der Noort, V.; Theelen, W.S.M.E.; Schouten, R.D.; Hummelink, K.; Muller, S.H.; Wolf-Lansdorf, M.; Dagelet, J.W.F.; Monkhurst, K.; et al. Prediction of response to anti-PD-1 therapy in patients with non-small-cell lung cancer by electronic nose analysis of exhaled breath. *Ann. Oncol.* **2019**, *30*, 1660–1666. [[CrossRef](#)] [[PubMed](#)]
16. Buszewski, B.; Ulanowska, A.; Kowalkowski, T.; Cieliski, K. Investigation of lung cancer biomarkers by hyphenated separation techniques and chemometrics. *Clin. Chem. Lab. Med.* **2012**, *50*, 573–581. [[CrossRef](#)]
17. Fu, X.-A.; Li, M.; Knipp, R.J.; Nantz, M.H.; Bousamra, M. Noninvasive detection of lung cancer using exhaled breath. *Cancer Med.* **2014**, *3*, 174–181. [[CrossRef](#)]
18. Hanna, G.B.; Boshier, P.R.; Markar, S.R.; Romano, A. Accuracy and Methodologic Challenges of Volatile Organic Compound-Based Exhaled Breath Tests for Cancer Diagnosis: A Systematic Review and Meta-analysis. *JAMA Oncol.* **2019**, *5*, e182815. [[CrossRef](#)]
19. Ismail, B.B.; Pu, Y.; Guo, M.; Ma, X.; Liu, D. LC-MS/QTOF identification of phytochemicals and the effects of solvents on phenolic constituents and antioxidant activity of baobab (*Adansonia digitata*) fruit pulp. *Food Chem.* **2019**, *277*, 279–288. [[CrossRef](#)]
20. Williams, M.D.; Reeves, R.; Resar, L.S.; Hill, H.H. Metabolomics of colorectal cancer: Past and current analytical platforms. *Anal. Bioanal. Chem.* **2013**, *405*, 5013–5030. [[CrossRef](#)]
21. Weitzel, K.-M. Bond-Dissociation Energies of Cations—Pushing the limits to quantum state resolution. *Mass Spectrum. Res.* **2011**, *30*, 221–235. [[CrossRef](#)] [[PubMed](#)]
22. Walker, V.; Mills, G.A. Urine 4-heptanone: A β -oxidation product of 2-ethylhexanoic acid from plasticisers. *Clin. Chim. Acta* **2001**, *306*, 51–61. [[CrossRef](#)] [[PubMed](#)]
23. Horváth, I.; Lázár, Z.; Gyulai, N.; Kollai, M.; Losonczy, G. Exhaled biomarkers in lung cancer. *Eur. Respir. J.* **2009**, *34*, 261–275. [[CrossRef](#)] [[PubMed](#)]
24. Song, G.; Qin, T.; Liu, H.; Xu, G.B.; Pan, Y.Y.; Xiong, F.X.; Gu, K.S.; Sun, G.P.; Chen, Z.D. Quantitative breath analysis of volatile organic compounds of lung cancer patients. *Lung Cancer* **2010**, *67*, 227–231. [[CrossRef](#)]
25. Rudnicka, J.; Walczak, M.; Kowalkowski, T.; Jezierski, T.; Buszewski, B. Determination of volatile organic compounds as potential markers of lung cancer by gas chromatography-mass spectrometry versus trained dogs. *Sens. Actuators B Chem.* **2014**, *202*, 615–621. [[CrossRef](#)]
26. Phillips, M.; Cataneo, R.N.; Cummin, A.R.C.; Gagliardi, A.J.; Gleeson, K.; Greenberg, J.; Maxfield, R.A.; Rom, W.N. Detection of lung cancer with volatile markers in the breath. *Chest* **2003**, *123*, 2115–2123. [[CrossRef](#)]
27. Peled, N.; Barash, O.; Tisch, U.; Ionescu, R.; Broza, Y.Y.; Ilouze, M.; Mattei, J.; Bunn, P.A.; Hirsch, F.R.; Haick, H. Volatile fingerprints of cancer specific genetic mutations. *Nanomed. Nanotechnol. Biol. Med.* **2013**, *9*, 758–766. [[CrossRef](#)]
28. Pérez Antón, A.; Ramos, Á.G.; del Nogal Sánchez, M.; Pavón, J.L.P.; Cordero, B.M.; Pozas, Á.P.C. Headspace-programmed temperature vaporization-mass spectrometry for the rapid determination of possible volatile biomarkers of lung cancer in urine. *Anal. Bioanal. Chem.* **2016**, *408*, 5239–5246. [[CrossRef](#)]
29. 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 Int.* **2012**, *12*, 7. [[CrossRef](#)]
30. Pérez Antón, A.; del Nogal Sánchez, M.; Crisolino Pozas, A.Á.P.; Pérez, J.L.P.; Cordero, B.M. Headspace-programmed temperature vaporizer-mass spectrometry and pattern recognition techniques for the analysis of volatiles in saliva samples. *Talanta* **2016**, *160*, 21–27. [[CrossRef](#)]

31. Rocco, R.; Incalzi, R.A.; Pennazza, G.; Santonico, M.; Pedone, C.; Bartoli, I.R.; Vernile, C.; Mangiameli, G.; La Rocca, A.; De Luca, G.; et al. BIONOTE e-nose technology may reduce false positives in lung cancer screening programmes. *Eur. J. Cardio Thoracic Surg.* **2016**, *49*, 1112–1117. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Li, M.; Yang, D.; Brock, G.; Knipp, R.J.; Bousamra, M.; Nantz, M.H.; Fu, X.-A. Breath carbonyl compounds as biomarkers of lung cancer. *Lung Cancer* **2015**, *90*, 92–97. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Knipp, R.J.; Li, M.; Fu, X.A.; Nantz, M.H. A versatile probe for chemoselective capture and analysis of carbonyl compounds in exhaled breath. *Anal. Methods* **2015**, *7*, 6027–6033. [\[CrossRef\]](#)
34. Horváth, I.; Barnes, P.J.; Loukides, S.; Sterk, P.J.; Högman, M.; Olin, A.C.; Amann, A.; Antus, B.; Baraldi, E.; Bikov, A.; et al. A european respiratory society technical standard: Exhaled biomarkers in lung disease. *Eur. Respir. J.* **2017**, *49*, 1600965. [\[CrossRef\]](#)
35. Martin, A.N.; Farquar, G.R.; Jones, A.D.; Frank, M. Human breath analysis: Methods for sample collection and reduction of localized background effects. *Anal. Bioanal. Chem.* **2010**, *396*, 739–750. [\[CrossRef\]](#)
36. Risby, T.H. Critical issues for breath analysis. *J. Breath Res.* **2008**, *2*, 030302. [\[CrossRef\]](#)
37. Slingers, G.; Goossens, R.; Janssens, H.; Spruyt, M.; Goelen, E.; Vanden, E.M.; Raes, M.; Koppen, G. Real-time selected ion flow tube mass spectrometry to assess short- and long-term variability in oral and nasal breath. *J. Breath Res.* **2020**, *14*, 036006. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Li, B.; Zou, X.; Wang, H.; Lu, Y.; Shen, C.; Chu, Y. Standardization study of expiratory conditions for on-line breath testing by proton transfer reaction mass spectrometry. *Anal. Biochem.* **2019**, *581*, 113344. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Kischkel, S.; Miekisch, W.; Patricia, F.; Schubert, J.K. Breath analysis during one-lung ventilation in cancer patients. *Eur. Respir. J.* **2012**, *40*, 706–713. [\[CrossRef\]](#)
40. Capuano, R.; Santonico, M.; Pennazza, G.; Ghezzi, S.; Martinelli, E.; Roscioni, C.; Lucantoni, G.; Galluccio, G.; Paolesse, R.; Di Natale, C.; et al. The lung cancer breath signature: A comparative analysis of exhaled breath and air sampled from inside the lungs. *Sci. Rep.* **2015**, *5*, 16491. [\[CrossRef\]](#)
41. Krilaviciute, A.; Leja, M.; Kopp-Schneider, A.; Barash, O.; Khatib, S.; Ama, H.; Broza, Y.Y.; Polaka, I.; Parshutin, S.; Rudule, A.; et al. Associations of diet and lifestyle factors with common volatile organic compounds in exhaled breath of average-risk individuals. *J. Breath Res.* **2019**, *13*, 026006. [\[CrossRef\]](#)
42. Zou, X.; Wang, H.; Ge, D.; Lu, Y.; Xia, L.; Huang, C.; Shen, X.; Chu, Y. On-line monitoring human breath acetone during exercise and diet by proton transfer reaction mass spectrometry. *Future Sci.* **2018**, *11*, 33–40. [\[CrossRef\]](#)
43. Vishinkin, R.; Haick, H. Nanoscale Sensor Technologies for Disease Detection via Volatolomics. *Small* **2015**, *11*, 6142–6164. [\[CrossRef\]](#)
44. Yang, T.; Rao, Z.; Zhang, X.; Xu, M.; Xu, Z.; Yang, S.T. Metabolic engineering strategies for acetoin and 2,3-butanediol production: Advances and prospects. *Crit. Rev. Biotechnol.* **2017**, *37*, 990–1005. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Xiao, Z.; Lu, J.R. Strategies for enhancing fermentative production of acetoin: A review. *Biotechnol. Adv.* **2014**, *32*, 492–503. [\[CrossRef\]](#)
46. Sun, J.A.; Zhang, L.Y.; Rao, B.; Shen, Y.L.; Wei, D.Z. Enhanced acetoin production by *Serratia marcescens* H32 with expression of a water-forming NADH oxidase. *Bioresour. Technol.* **2012**, *119*, 94–98. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Yan, X.; Yang, M.; Liu, J.; Gao, R.; Hu, J.; Li, J.; Zhang, L.; Shi, Y.; Guo, H.; Cheng, J.; et al. Discovery and validation of potential bacterial biomarkers for lung cancer. *Am. J. Cancer Res.* **2015**, *5*, 3111–3122.
48. Zhang, W.; Luo, J.; Dong, X.; Zhao, S.; Hao, Y.; Peng, C.; Shi, H.; Zhou, Y.; Shan, L.; Sun, Q.; et al. Salivary microbial dysbiosis is associated with systemic inflammatory markers and predicted oral metabolites in non-small cell lung cancer patients. *J. Cancer* **2019**, *10*, 1651–1662. [\[CrossRef\]](#)
49. Poli, D.; Goldoni, M.; Corradi, M.; Acampa, O.; Carbognani, P.; Internullo, E.; Casalini, A.; Mutti, A. Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME-GC/MS. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2010**, *878*, 2643–2651. [\[CrossRef\]](#) [\[PubMed\]](#)

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.