





Article

Serum Metabolomics Profiling Reveals Metabolic Alterations Prior to a Diagnosis with Non-Small Cell Lung Cancer among Chinese Community Residents: A Prospective Nested Case-Control Study

Yu Xiang ¹, Qi Zhao ¹, Yilin Wu ², Xing Liu ¹, Junjie Zhu ^{1,3}, Yuting Yu ¹, Xuyan Su ², Kelin Xu ⁴, Yonggen Jiang ^{2,*} and Genming Zhao ^{1,*}

¹ Key Laboratory of Public Health Safety of Ministry of Education, Department of Epidemiology, School of Public Health, Fudan University, Shanghai 200032, China

² Songjiang District Center for Disease Control and Prevention, Shanghai 201600, China

³ Department of Epidemiology and Health Statistics, School of Public Health, Dali University, Dali 671000, China

⁴ Department of Biostatistics, School of Public Health, Key Laboratory of Public Health Safety of Ministry of Education, Fudan University, Shanghai 200032, China

* Correspondence: sjkzx1106@126.com (Y.J.); gmzhao@shmu.edu.cn (G.Z.)



Citation: Xiang, Y.; Zhao, Q.; Wu, Y.; Liu, X.; Zhu, J.; Yu, Y.; Su, X.; Xu, K.; Jiang, Y.; Zhao, G. Serum Metabolomics Profiling Reveals Metabolic Alterations Prior to a Diagnosis with Non-Small Cell Lung Cancer among Chinese Community Residents: A Prospective Nested Case-Control Study. *Metabolites* **2022**, *12*, 906. <https://doi.org/10.3390/metabo12100906>

Academic Editors: Yaohua Tian and Jixuan Ma

Received: 18 August 2022

Accepted: 22 September 2022

Published: 27 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/>).

Abstract: The present high mortality of lung cancer in China stems mainly from the lack of feasible, non-invasive and early disease detection biomarkers. Serum metabolomics profiling to reveal metabolic alterations could expedite the disease detection process and suggest those patients who are harboring disease. Using a nested case-control design, we applied ultra-high-performance liquid chromatography/mass spectrometry (LC-MS)-based serum metabolomics to reveal the metabolomic alterations and to indicate the presence of non-small cell lung cancer (NSCLC) using serum samples collected prior to disease diagnoses. The studied serum samples were collected from 41 patients before a NSCLC diagnosis (within 3.0 y) and 38 matched the cancer-free controls from the prospective Shanghai Suburban Adult Cohort. The NSCLC patients markedly presented cellular metabolism alterations in serum samples collected prior to their disease diagnoses compared with the cancer-free controls. In total, we identified 18 significantly expressed metabolites whose relative abundance showed either an upward or a downward trend, with most of them being lipid and lipid-like molecules, organic acids, and nitrogen compounds. Choline metabolism in cancer, sphingolipid, and glycerophospholipid metabolism emerged as the significant metabolic disturbance of NSCLC. The metabolites involved in these biological processes may be the distinctive features associated with NSCLC prior to a diagnosis.

Keywords: non-small cell lung cancer (NSCLC); metabolomics; cancer cells metabolism; nested case-control study

1. Introduction

Lung cancer is currently the leading cause of death in China [1]. According to the International Agency for Research on Cancer, in 2020, the estimated number of incidents of lung cancer was over three million in China, accounting for nearly 19.4% of deaths in both sexes and all ages [2]. Moreover, the survival rate of lung cancer remains dismal. Most early-stage lung cancers are asymptomatic, contributing to a delayed diagnosis and resulting in an overall 5 y survival rate of 19% over all stages; however, the detection of lung cancer at the early stage could increase the 5 y survival rates, rising to 57% [3]. This significant difference indicates the importance of detecting lung cancer at an early stage for improving the overall survival rate. With the available treatments having increased in recent years, the outcomes with an early detection would also be improved.

At present, low-dose spiral computerized tomography (LDCT) is considered an effective imaging tool for detecting early-stage lung lesions and for screening high-risk populations of lung cancer [4]. However, the screening recommendations have been subjectively assessed based on smoking behavior, resulting in a potential reluctance on the part of patients. Alongside concerns about the economic burden to the nation, the potential radiation hazards and psychological burdens to patients limit the feasibility of widespread LDCT screening [5]. In addition, the false-positive rate in a low-dose group, for example, was estimated to be 96.4%, with most positive results having been withdrawn after further imaging examinations [6]. These reasonable arguments support the urgent need for a less invasive screening test, either without or with less harmful side effects, to provide an alert for the suspicious signs of early neoplasm. Such tests could help medical practitioners triage patients with malignancy, support an advanced treatment efficiency, and improve treatment outcomes to lower lung cancer-associated mortality.

Molecular biology, genomics, proteomics, and metabolomics are interrelated in the probing and identifying of biological processes. Genomics measures genetic mutations and predicts the possibility of disease development at the individual level; however, proteomics and metabolomics could identify ongoing biological activities, including alterations due to the presence of disease [7]. In cancer, metabolomics captures complex physiological and pathological characteristics, and detects oncological developments via measurable metabolic profiles from the metabolic pathways through global metabolite variations.

Although metabolomic studies of cancer are carried out frequently, most of these studies are conducted with case-control study designs, and the results might be prone to reverse causation [8–11]. Currently, the serum metabolomic profiling revealing metabolic alterations prior to a diagnosis with lung cancer in the Chinese population is scarce and limited. The Shanghai Suburban Adult Cohort and Biobank (SSACB) study is a population-based prospective cohort that included 36,404 participants from the Songjiang district, in Shanghai, China, who were predominantly community residents aged 20–74 years old. Here, we report the metabolic alterations established through blood serum samples collected before the diagnosis of non-small cell lung cancer (NSCLC) and compare the patients' results with the serum samples obtained from healthy, cancer-free control subjects in the Shanghai Suburban Adult Cohort. Our study provides an opportunity to prospectively investigate the role of metabolites and their alteration in serum samples collected before the diagnosis of NSCLC to explore its associated biological processes to prevent disease precisely. To our knowledge, our study was the first prospective study to investigate the pre-diagnostic serum metabolic biomarkers and pathway alterations with the liquid chromatography/mass spectrometry (LC-MS) technique among the community residents in Shanghai. Our results demonstrate the potential of LC-MS-based serum metabolites in conjunction with baseline characteristics to indicate the presence of NSCLC, and its potential to be used as an effective tool for screening high-risk patients of NSCLC.

2. Materials and Methods

2.1. Ethics

This study was performed in conformity with the Declaration of Helsinki and was approved by the Institutional Review Board of Fudan University, School of Public Health, China (Authorization number: IRB#2016-04-0586). Informed consent was obtained from all participants.

2.2. Design and Subjects

From April 2016 to September 2017, the SSACB study recruited 36,404 participants aged 20–74 years old, primarily community residents living in the Songjiang District, in Shanghai, China. Blood samples of the participants were drawn during the baseline investigation. Detailed descriptions of our participants' recruitment criteria and phases are illustrated in the cohort profile [12]. Sociodemographic information, health status (self-evaluation), lifestyle, and the history of chronic diseases were collected using a structured questionnaire at the baseline investigation for all the participants included in the SSCAB

study. Fasting blood serum samples were collected from the participants enrolled in the baseline investigation. Afterward, the serum samples were transported to the biobank on ice at approximately 0 °C to −4 °C and stored at −80 °C until the analysis.

We have depicted the study overview, subjects, and analysis with a flow chart in Figure 1. All cohort members received periodic follow-ups based on the health information system (HIS) linked to cancer, stroke, death, and other vital report registry systems at different local disease control and prevention centers. In this study, the NSCLC cases were identified through the National Cancer Registry System implemented by hospitals in Shanghai. The information of local community residents was pooled by the Songjiang District Center for Disease Control and Prevention (Songjiang CDC) from the cancer registry data sources. The cancer-free controls were matched according to age and gender within the same district. A diagnosis of incident NSCLC was confirmed by the medical record reviewed by the staff from the department of non-communicable disease of the Songjiang CDC. In our study, the incident NSCLC cases were coded as C34.9 according to the *International Classification of Diseases, tenth revision (ICD-10)*, and the information on the pathological diagnoses was collected from the cancer report system.

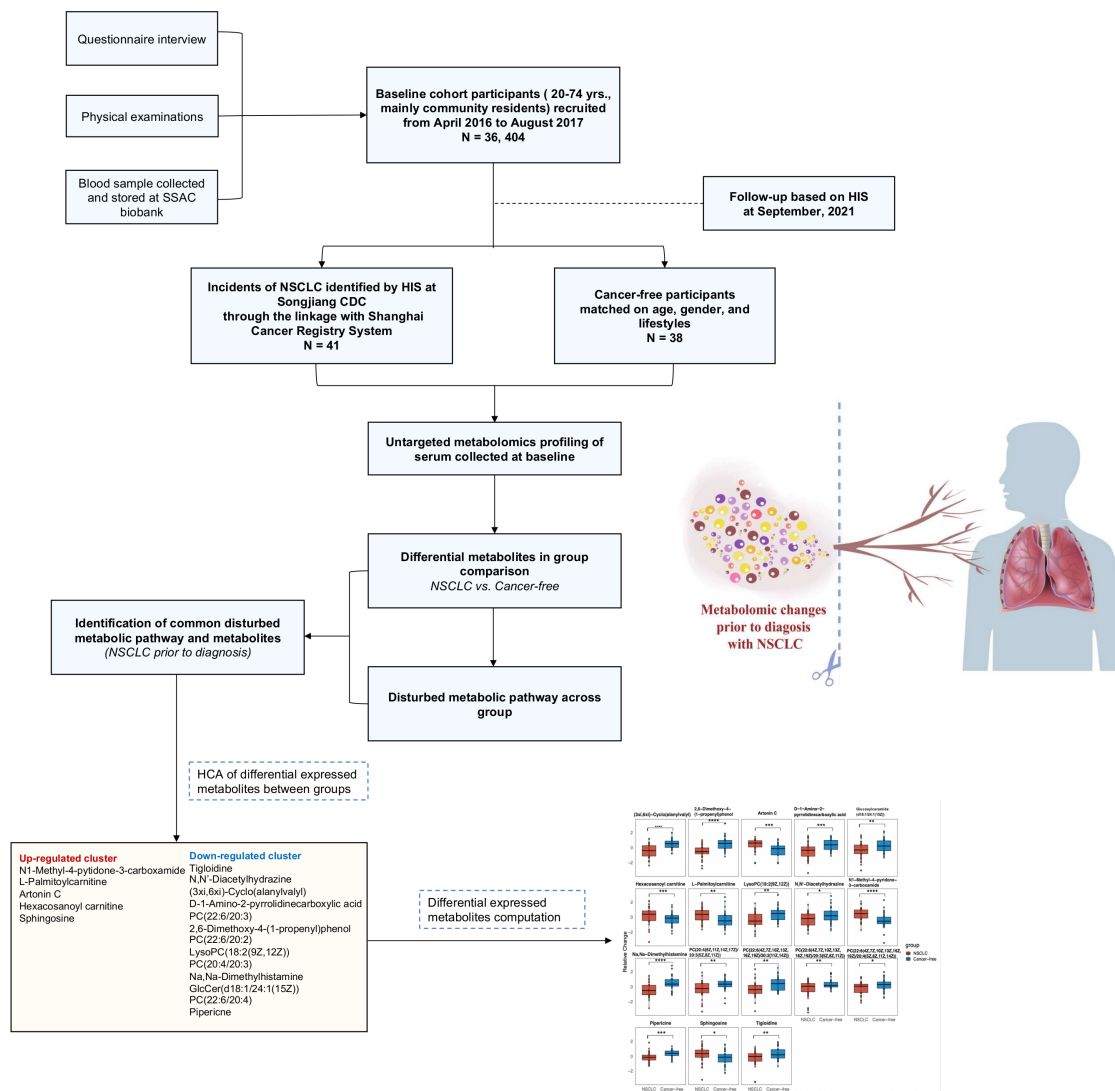


Figure 1. Flow chart of the study overview, subjects, and analysis.

In this present study, the inclusion criteria of the NSCLC cases were: an incident of NSCLC and serum samples collected with sufficient volume without hemolysis; the

exclusion criteria were: a participant with an alcohol-drinking habit, a prior diagnosis of another type of cancer at the baseline investigation, and being diagnosed with NSCLC after entering the cohort within a half year to exclude the subjects with preclinical stages of cancer. Participants in the cohort without having major health events such as a stroke, diabetes, cancer, and other chronic diseases self-reported in the baseline investigation were determined as being healthy, cancer-free participants. Among those, we randomly selected 40 subjects matched in age (± 2 years) and gender as the cancer-free control group, but two serum samples were excluded because of hemolysis. Finally, 41 NSCLC cases and 38 cancer-free controls were included for further metabolomic profiling.

2.3. Data Collection and Measurement

At the baseline investigation, a structured questionnaire was conducted face-to-face by trained healthcare staff at the local community healthcare centers. Information on age, gender, education, socio-demographic information, lifestyles, and self-reported disease history was collected. Physical examinations, including height, weight, waist and hip circumference were carried out for every participant. Fasting blood was collected for a classical lipid profile, serum alanine aminotransferase (ALT), aspartate amino transferase (AST), and creatinine (the analysis was finished by Dian Diagnostic Co. Ltd. (Hangzhou, China) (a medical laboratory center). At follow-up, details of the NSCLC cases were extracted from the non-communicable disease department at the Songjiang CDC, which included the international classification (tenth revision) of diseases code, pathological diagnosis, the TNM classification of malignant tumor, etc.

2.4. Metabolomics Profiling and Data Preprocessing

The fasting serum samples were collected at the baseline investigation and were transported on ice and stored at -80 °C for each subject enrolled in our study. The serum samples were processed by the following steps: first, the frozen serum samples were thawed at 4 °C on ice. A total of 400 μ L of the extraction solution (acetonitrile/methanol = $1/1$, v/v) containing the isotopically-labelled internal standard mixture was added into an aliquot of serum samples (100 μ L). Next, the samples were mixed on a vortex for 30 s, sonicated in the ice-watered bath for 10 min, and incubated for an hour to precipitate. Afterward, the samples were centrifuged at $12,000$ rpm for 15 mins at 4 °C. Then, the supernatant was transferred to the glass vial for analysis through an ultra-high performance liquid chromatography-mass spectrometry (UPLC-MS) using the UHPLC system (Vanquish, Thermo Fisher Scientific) with a UPLC BEH Amide column coupled to a Q-Exactive HFX mass spectrometer (Orbitrap MS, Thermo). Details about the metabolite extraction and quality control process are presented in Appendix A. During the data acquisition process, one quality control (QC) sample was run after every ten samples. The LC-MS raw data were converted to the mzXML file using the ProteoWizard software (Pala Alto, California, USA). The peak detection, extraction, alignment, and integration were performed by the XCMS package (La Jolla, California, USA). The metabolite annotation was performed by the in-house MS2 database (BiotreeDB). The cut-off value for an annotation was set at 0.4 .

Missing values were replaced with half of the minimum. A relative standard deviation of metabolites $> 25\%$ in the QC samples was considered unproducibile and, thus, removed from further analysis. The data were normalized by dividing the peak area of the metabolite by the peak area of the internal standards, which is a common technique for untargeted metabolomics. We identified 3868 metabolite features (273 named and 3595 unnamed) under the positive ion mode of the Q-Exactive HFX mass spectrometer. After finishing the preprocessing, a data frame consisting of the retention time, mass-to-charge ratio values, and normalized peak intensity was subjected to further analysis.

2.5. Statistical Analysis

Continuous variables were presented as the mean \pm (standard deviation [SD]) or median (interquartile range [IQR]). Categorical variables were expressed as a number

(percentage). In univariate statistical comparisons, we used a chi-square or Fisher's exact test for the categorical variables, a Student's *t*-test or ANOVA for the normally distributed continuous variables, and a Wilcoxon signed-rank test or Kruskal–Wallis test for the skewed distributed continuous variables.

To minimize the impact of noise and the high variance of the variables, we scaled and logarithmically transformed the peak intensity. Then, the principal component analysis (PCA), an unsupervised analysis that could reduce the dimension of the data, was performed to visualize the distribution and grouping of the samples, including the QC samples. To visualize the separation of the groups, we performed a partial least squares-discriminant analysis (PLS-DA). The results for the PLS-DA with a positive ion mode are presented in Section 3.2. We performed an orthogonal projections to latent structures discriminant analysis (OPLS-DA) to find the significant different metabolites between the two groups. The metabolites with a variable importance in projection (VIP) > 1 were included for further analysis. The results of the PCA and OPLS-DA are presented in Appendix B.

We applied the Benjamini–Hochberg procedure to control the false discovery rate (FDR). A *p*-value < 0.05 (using a *t*-test or Kruskal–Wallis test) and FDR corrected *p*-value (*q*-value) ≤ 0.2 were considered as significantly changed metabolites. We provide the named differential metabolites in the Supplementary Table S1: Part A. We set a strict selection criterion to select the differentially-expressed metabolites to generate the biologically valuable metabolites, given this study's exploratory and hypothesis-generating nature. The selected 18 differentially-expressed metabolites were compared using the Student's *t*-test to measure the relative abundance between the groups. The data were normalized by a log-transformation and UV scaling. A *p*-value ≤ 0.5 was considered statistically significant. The volcano plot of the overall metabolites identified across groups and the relative abundance of the differentially-expressed metabolites were presented in Section 3.2.

We used the R version 4.0.1 for the pathway analysis. The significantly changed metabolites (Supplementary Table S1) were mapped into their biochemical pathway using the Kyoto Encyclopedia of Genes and Genomes (KEGG) (<https://www.kegg.jp>) [13]. The pathway significance was determined based on the total number of metabolites mapped into the biochemical pathway and the total number of compounds in the involved pathway was calculated based on the hypergeometric distribution. The pathway was considered disturbed if the number of significant metabolites (i.e., hits) was ≥ 1 and the raw *p* of the hypergeometric distribution was < 0.10. The raw *p* was calculated based on the number of hits and the total number of compounds in a pathway. We used this lenient criterion because the sample size for each comparison group was small. The disturbed metabolic pathways enriched by the KEGG database are presented in Section 3.3. We provide the results of the KEGG enrichment analysis in Supplementary Table S2. Furthermore, a HCA was performed among the significantly changed metabolites in the Human Metabolome Database (HMDB) (<https://hmdb.ca/>) [14], using the “complete” method and the “Euclidean” distance measure. Finally, a Spearman correlation analysis was performed to evaluate the relationships between the significantly changed metabolites, baseline parameters, and classical lipids among all the participants.

We also performed a hierarchical clustering analysis (HCA) among the significantly expressed metabolites to identify whether types of metabolites tended to cluster together. In addition, we conducted a least absolute shrinkage and selection operator (LASSO) regression to optimize the variable selection and to identify whether any combinations of serum metabolite signatures could reduce the dimension in the data. Then, we applied a multivariable logistic regression analysis to assess the association between each metabolite and the risk of having NSCLC. The peak intensity of the metabolites was log-transformed in the regression analysis. Age and BMI were adjusted as continuous variables. Sex, education level, the family history of NSCLC, smoking status, drinking status, and exercise levels were retained as categorical covariates in the logistic regression models. Additionally, to assess the performance of the model, we conducted an internal validation with a three-fold cross validation, with the results presented in Section 3.4. Finally, we performed a Spearman

correlation matrix between the significantly expressed metabolites, classical lipids, and baseline characteristics among all the subjects. The results are shown in Section 3.5. All the statistical analyses mentioned above were performed using the R version 4.0.1.

2.6. Role of the Funding Sponsors

The sponsors of the study declare no role in the study design, data collection, analysis, interpretation, or writing of the manuscript. The corresponding author had full access to all the data in this study and was responsible for the final decision to submit this paper for publication.

3. Results

3.1. Participant Characteristics

Table 1 presents the baseline characteristics of the study population. Of 79 participants, the NSCLC and the cancer-free controls were comparable in age and sex. The participants of our cohort primarily consisted of the elderly and middle-aged adults, leading to the majority of our study subjects being over 55 years old.

Table 1. Characteristics of the study participants.

Characteristic	NSCLC (n = 41)	Cancer-Free Control (n = 38)	Total (n = 79)	p-Value
Age, years				0.371
41–55	9 (21.95%)	8 (21.05%)	17 (21.52%)	
56–65	18 (43.90%)	22 (57.89%)	40 (50.63%)	
65–75	14 (34.15%)	8 (21.05%)	22 (27.85%)	
Mean ± SD (years)	61.24 ± 6.97	60.21 ± 6.74	60.75 ± 6.84	0.505
Gender				0.411
Male	21 (51.22%)	15 (39.47%)	36 (45.57%)	
Female	20 (48.78%)	23 (60.53%)	43 (54.43%)	
Education				0.558
Middle school or below	20 (48.78%)	22 (57.89%)	42 (53.16%)	
High school or above	21 (51.22%)	16 (42.11%)	37 (46.84%)	
History of Respiratory Diseases				0.241
Yes	3 (7.32%)	0 (0.00%)	3 (3.80%)	
No	38 (92.68%)	38 (100.00%)	76 (96.20%)	
Smoking Status				0.160
Never	27 (65.85%)	31 (81.58%)	58 (73.42%)	
Former	1 (2.44%)	0 (0.00%)	1 (1.27%)	
Current	13 (31.71%)	7 (18.42%)	20 (25.32%)	
Second-hand Exposure				0.026
Yes	6 (14.63%)	0 (0.00%)	6 (7.59%)	
No	35 (85.37%)	38 (100.00%)	73 (92.41%)	
Alcohol drinking				0.228
Yes	0 (0.00%)	2 (5.26%)	2 (2.53%)	
No	41 (100.00%)	36 (94.74%)	77 (97.47%)	
Exercise				0.185
Yes	14 (34.15%)	7 (18.42%)	21 (26.58%)	
No	27 (65.85%)	31 (81.58%)	58 (73.42%)	
BMI, kg/m²	25.08 (22.88–27.29)	21.73 (20.72–22.55)	22.81(21.35–25.28)	<0.001
Waist to hip circumference ratio	0.88 ± 0.06	0.85 ± 0.05	0.88 ± 0.06	<0.001
HDL cholesterol, mmol/L	1.31 (1.12–1.56)	1.40 (1.31–1.49)	1.40 (1.21–1.52)	0.156
LDL cholesterol, mmol/L	2.8 (2.27–3.28)	2.78 (2.50–3.12)	2.79 (2.39–3.12)	0.910
TG, mmol/L	1.4 (1.17–1.76)	1.16 (0.90–1.79)	1.29 (1.07–1.79)	0.016
TC, mmol/L	4.87 ± 1.12	4.76 ± 0.57	4.82 ± 0.90	0.591
Time to diagnosis	1.44 (1.17–1.76)	NA	1.44 (1.17–1.76)	
Histological subtypes				
Adenocarcinoma	32 (78.05%)	NA	32 (78.05%)	
Squamous cell carcinoma	4 (9.76%)	NA	4 (9.76%)	
Other subtypes	5 (12.20%)	NA	5 (12.20%)	

Values are mean ± standard deviation (SD) for normally-distributed variables, median (interquartile range, (IQR)) for skewed-distributed variables, or n (%) for categorical variables. Abbreviations: BMI: body mass index; HDL: high-density lipo-protein; LDL: low-density lipo-protein; TG: triglyceride; TC: total cholesterol; NA: not applicable. *p*-values were calculated from a Wilcoxon signed-rank test for skewed continuous variables, a Student's *t*-test for variables with normal distributions, and a Chi-square test or Fisher's exact test for categorical variables.

3.2. Significantly Changed Metabolites and Their relative Changes across NSCLC Cases and Cancer-Free Controls

We first performed a PCA to assess the stability of the quality control (QC) samples and to visualize the metabolic separations across the NSCLC, cancer-free controls, and QC samples (Figure A1). Then, we analyzed the PLS-DA results and observed a significant separation between the NSCLC and cancer-free groups (Figure 2). The OPLS-DA model was applied to characterize the metabolic disturbances with the calculated VIP index. The OPLS-DA score plot and the result of the permutation test are shown in Figure A2. Afterward, we performed a Student's *t*-test and Benjamini–Hochberg procedure to control the FDR due to multiple testing. Furthermore, we also calculated the fold change between the groups for each metabolite. In total, 3868 peaks were identified under the positive mode, with 273 named (the cut-off value for annotation was 0.4) and 3595 unnamed. The volcano plot of the overall characteristics of the metabolite peaks is shown in Figure 3. Using the selection criterion of a VIP > 1, *p*-value < 0.05 and corrected *p*-value (*q*-value) < 0.2, we identified 18 metabolites and considered them to be significantly-changed (or expressed) metabolites. Here, we conducted stepwise comparisons for the identified significantly-changed metabolites for identifying the common metabolic pathways underlying the cases with NSCLC. The relative abundance of 18 metabolites between the groups is shown in Figure 4.

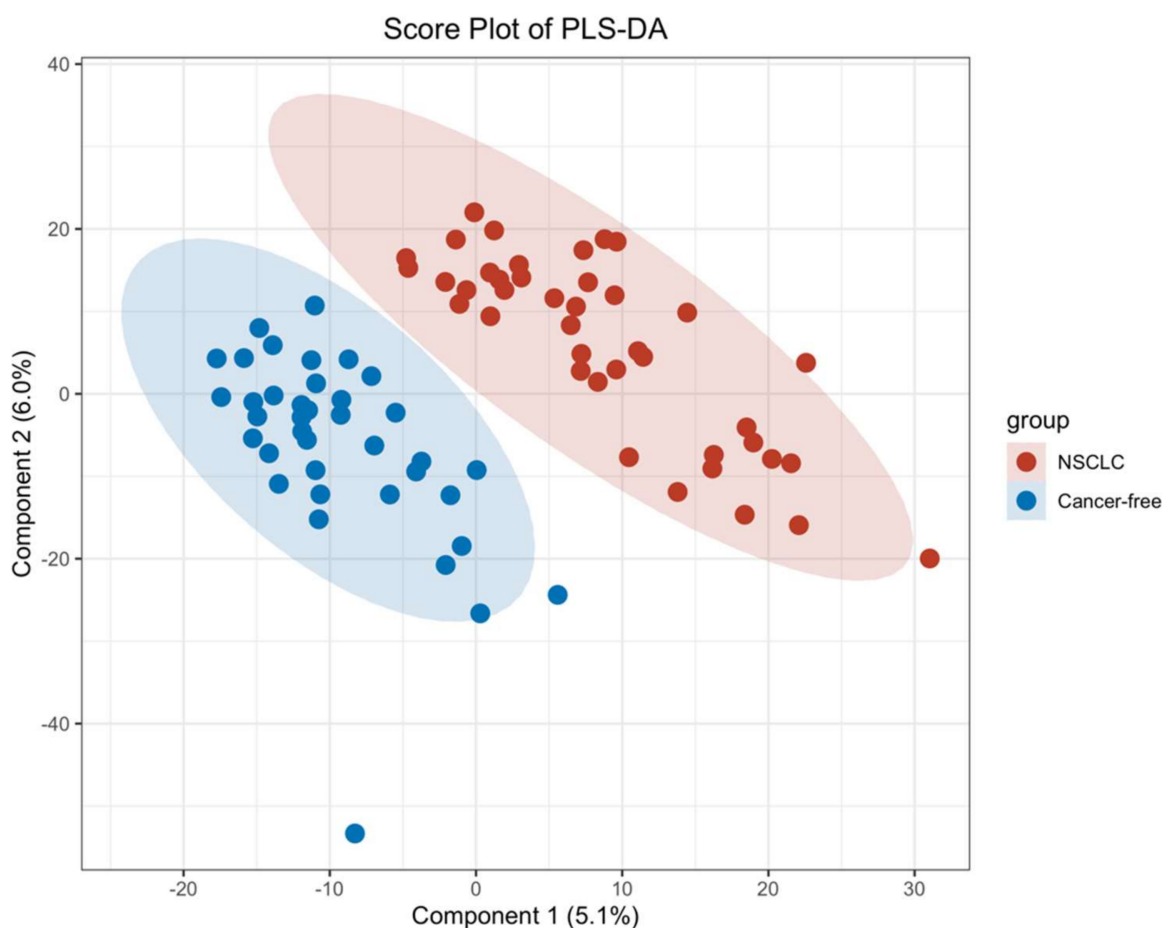


Figure 2. PLS-DA score plot between the NSCLC group and cancer-free control group.

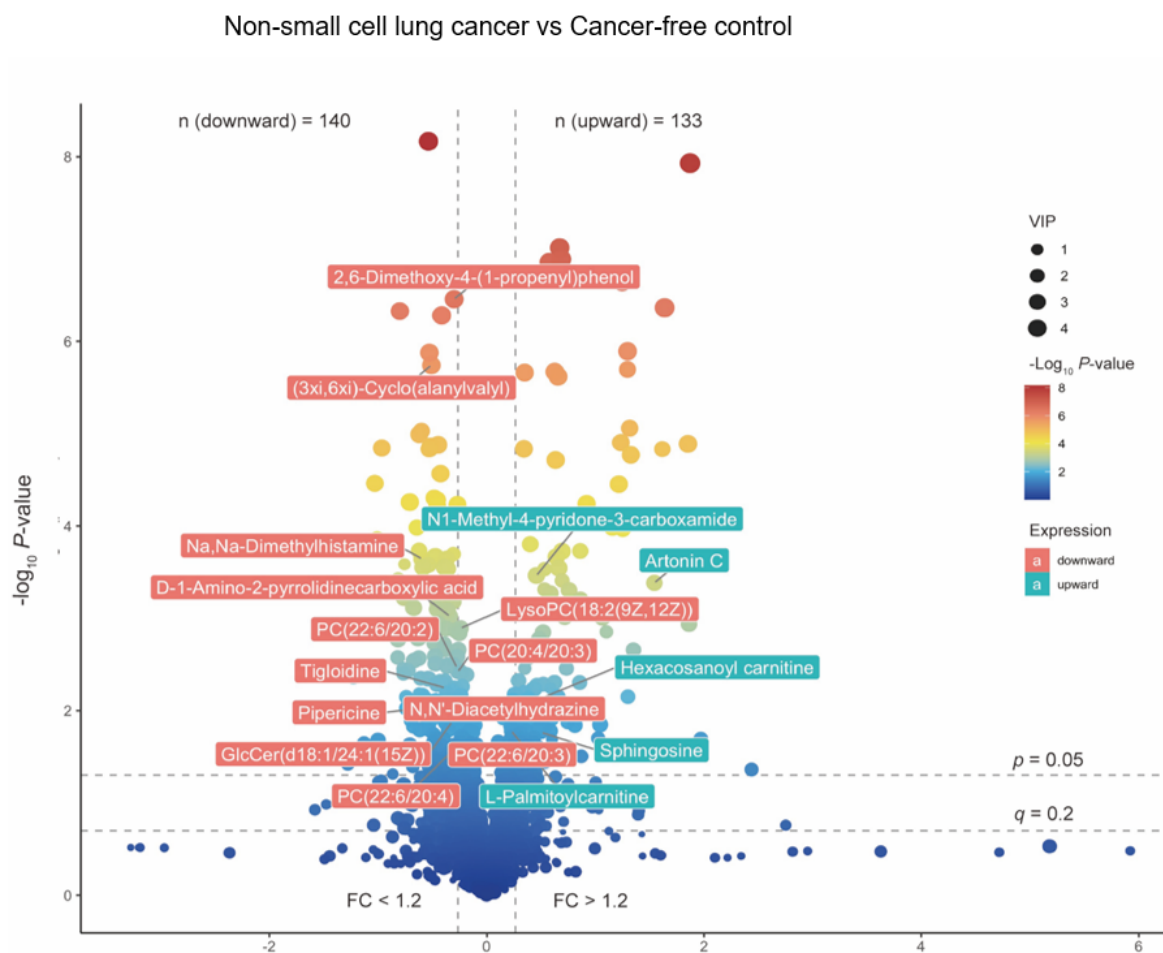


Figure 3. The volcano plot of overall metabolites identified between NSCLC cases and cancer-free controls. The p -value was calculated based on Student's t -test or Kruskal–Wallis test; The q -value was calculated with Benjamini–Hochberg procedure to control the FDR. FC was calculated based on the relative abundance across groups. The 18 significantly altered metabolites were labeled. Red labels stand for NSCLC significantly depleted metabolites, and green labels stand for NSCLC significantly enriched metabolites. Abbreviations: Lysophospholipid: LysoPC; PC: Phosphatidylcholine; GlcCer: Glucosylceramide.

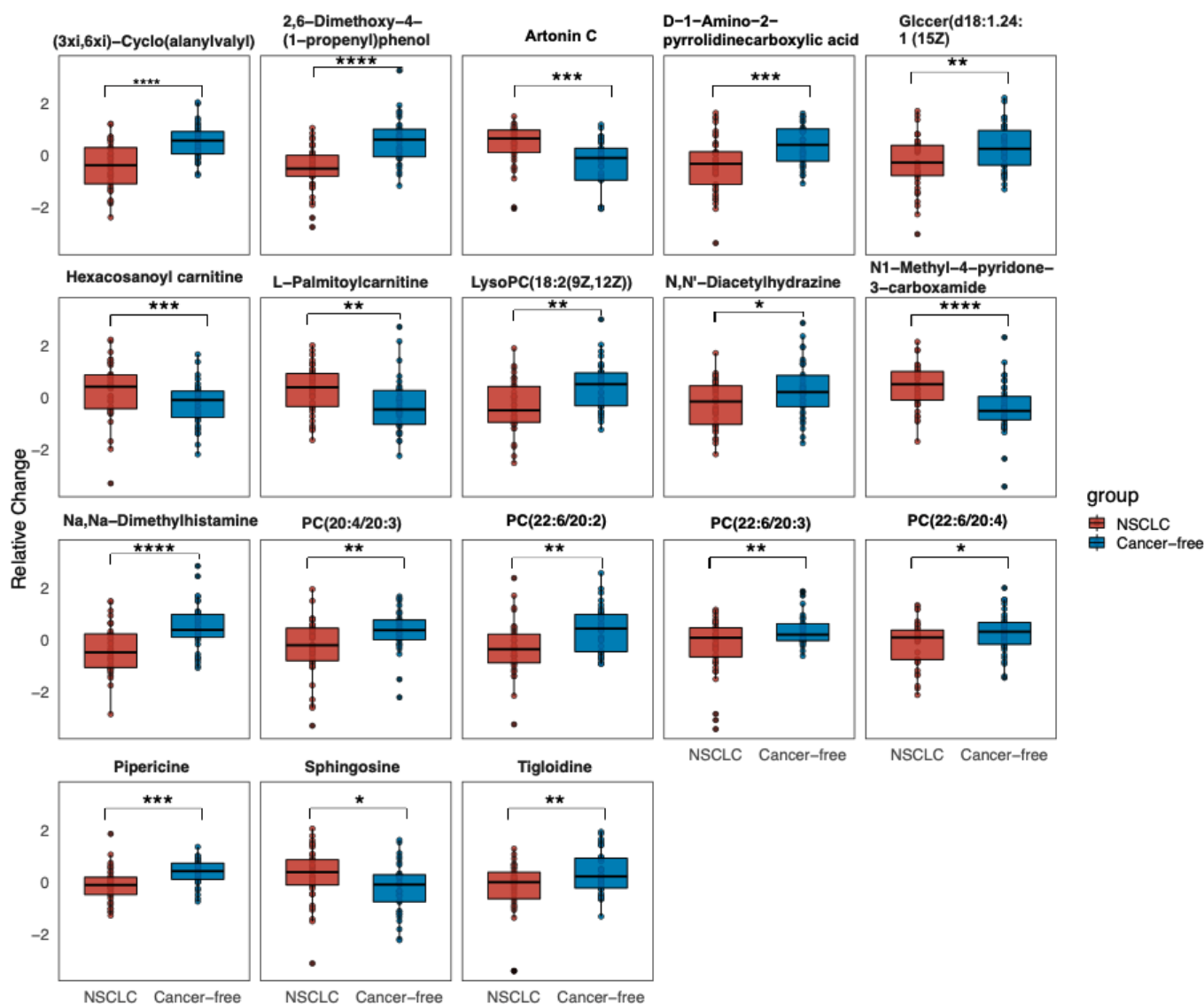


Figure 4. Relative abundance of 18 significantly changed metabolites across the groups (the relative expressed values were compared using a Student's *t*-test). Data are the normalized values by log-transformation and UV scaling. The presented box shows the 25th and 75th percentiles. The horizontal line represents the mean values of each metabolite. $p < 0.05$, *; $0.001 < p < 0.01$, **; $0.0001 < p < 0.001$, ***; $p < 0.0001$, ****. Abbreviation: GlcCer: Glucosylceramide; PC: Phosphatidylcholine; LysoPC: Lysophospholipid.

3.3. Altered Metabolic Pathways across NSCLC Cases and Cancer-Free Controls

The significantly-changed metabolites were mapped into their biological metabolic pathways underlying NSCLC. The perturbed metabolic pathways are presented in Figure 5.

For the participants before a diagnosis with NSCLC with a median time of 1.44 (1.17–1.76) years, compared with the cancer-free controls, the choline metabolism in the cancer was the most significant disturbed pathway. Notably, the sphingolipid and glycerophospholipid (GPL) metabolisms were the metabolic disturbances between the two groups, with two hits matched in the pathways and with *p*-values of an enrichment analysis being 0.003 and 0.009, respectively. Additionally, apoptosis, linoleic acid and necroptosis were also the disturbed pathways with a *p*-value < 0.05 but with only one hit matched. The Sphingolipid signaling pathway, nicotinate and nicotinamide, and the retrograde endocannabinoid signaling were

the matched pathways with a p -value < 0.1 , but they did not meet the criteria for a disturbed metabolic pathway, possibly due to the small sample size.

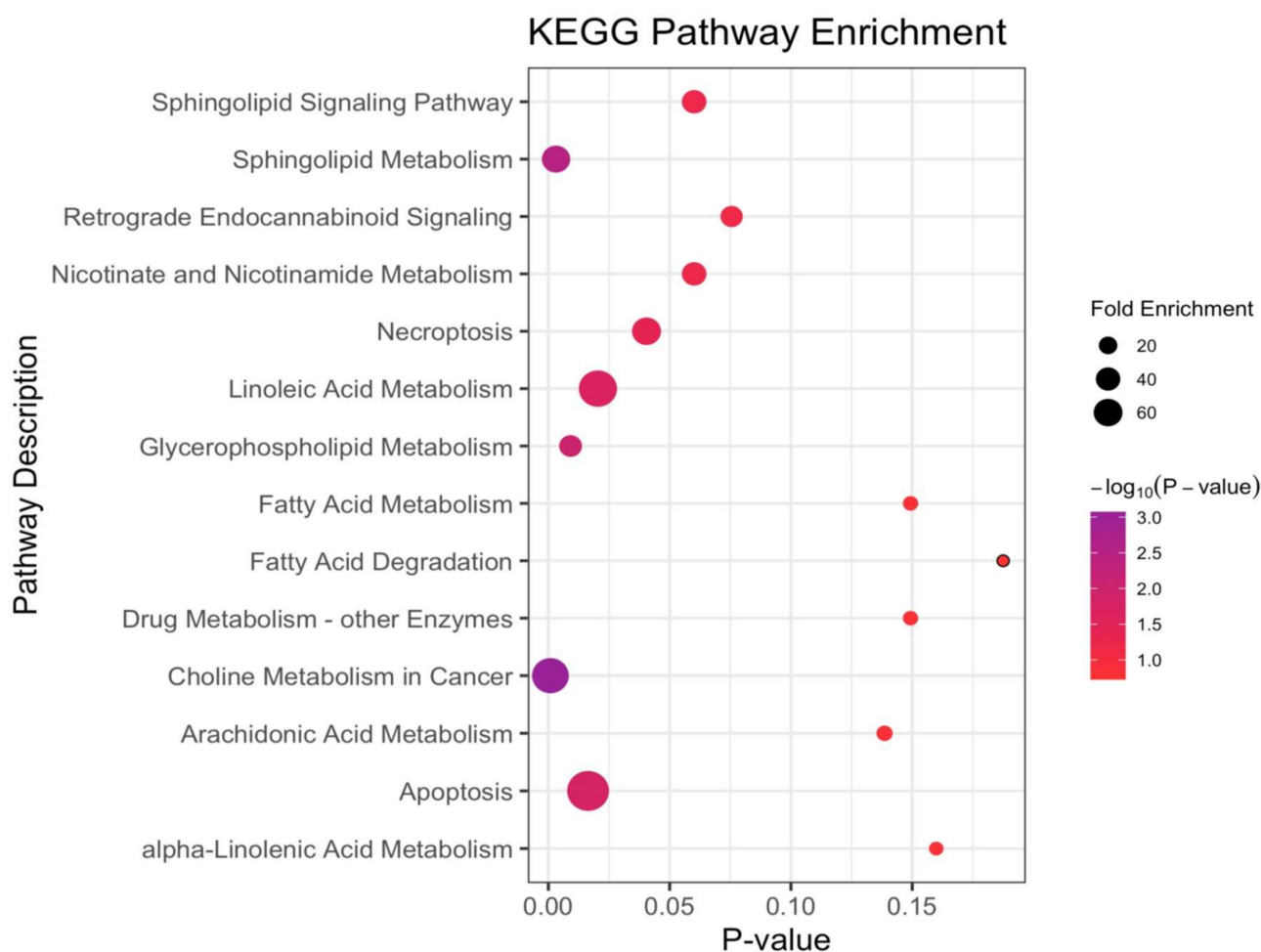


Figure 5. The disturbed metabolic pathways when comparing the NSCLC cases versus cancer-free controls. Various metabolism changes before a diagnosis with NSCLC was manifested (the p -value was calculated based on the hypergeometric distribution; fold enrichment was calculated based on the total number of metabolites mapped into the biochemical pathway; and the total number of compounds in the involved pathway).

3.4. Serum Metabolite Signatures for NSCLC

We explored whether any combinations of the serum metabolites could reduce the dimension in the data and serve as signatures of NSCLC. A hierarchical clustering analysis (HCA) was based on the 18 differentially-expressed metabolites, which provided intuitive visualizations of the trends in metabolites between the two groups (Figure 6a). We focused on the clusters that showed an upward or downward trend between the NSCLC and cancer-free controls. The first composition of 12 metabolites (mainly lipids and lipid-like molecules, organic acids and derivatives, and organic nitrogen compounds) showed a decreasing trend. The second composition of three metabolites (mainly lipids and lipid-like molecules and organoheterocyclic compounds) showed an increasing trend.

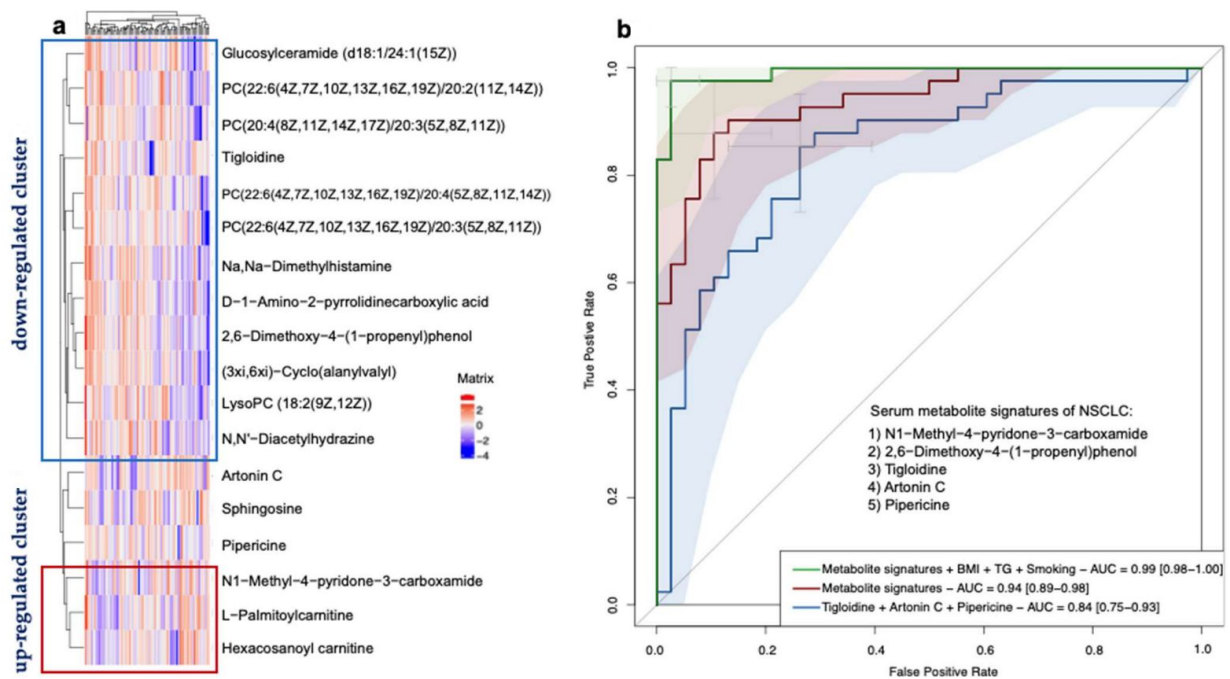


Figure 6. Identification of NSCLC-associated serum metabolite signatures. (a) Hierarchical clustering analysis of 18 significantly differentially-expressed metabolites according to OPLS-DA VIP values. The definition of significantly-expressed metabolites was the same as in Figure 4. (b) Receiver operating characteristic curve based on the logistic regression of metabolites selected from the LASSO regression and characteristics from the baseline investigation to identify the metabolite signatures and assess their probabilities in identifying NSCLC. The x-axis represents the false-positive rate of the risk prediction. The y-axis represents the true-positive rate of the risk prediction. The lines represent the performance of the model, and the shades represent the confidence intervals of the area under the curve.

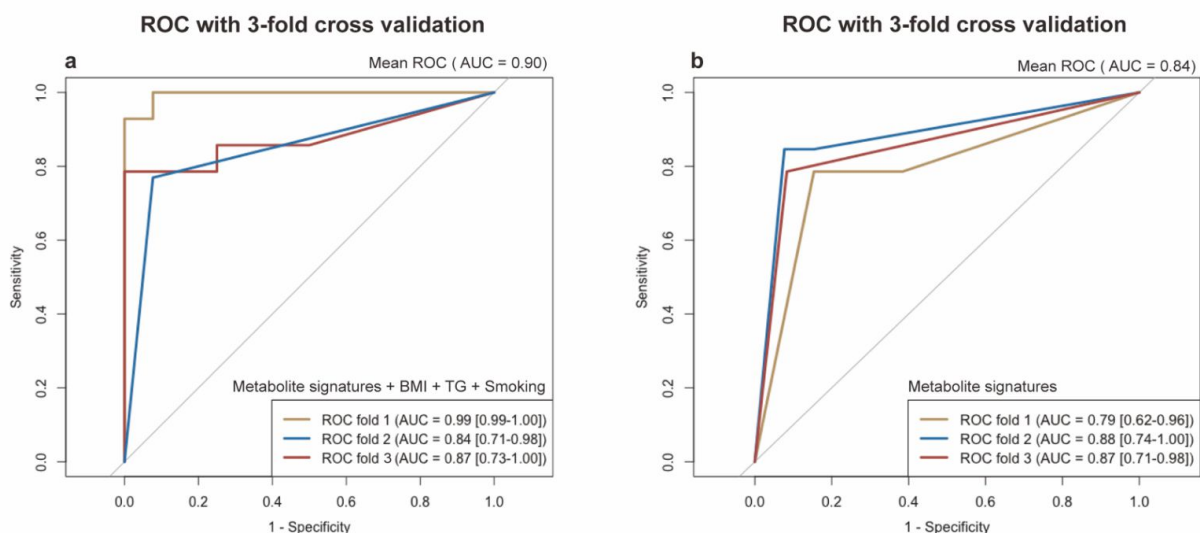


Figure 7. Receiver operating characteristic (ROC) curve with internal 3-fold cross validation. (a) ROC curve of logistic model based on the metabolite signatures, e.g., BMI, TG, and smoking. (b) ROC curve of logistic model only based on the metabolite signatures. The x-axis represents the false-positive rate. The y-axis represents the true-positive rate. The lines represent the performance of each fold. Metabolite signatures include N1-Methyl-4-pyridone-3-carboxamide, 2,6-Dimethoxy-4-(1-propenyl)phenol, Tigloidine, Artonin C, and Pipericine.

3.5. Correlation between Significantly Changed Metabolites, Baseline Characteristics, and Classical Lipids

We conducted a Spearman correlation matrix between the significantly-changed metabolites, baseline clinical characteristics, and classical lipid levels among all the participants (Figure 8). The metabolites that increased or decreased tended to cluster together in the correlation matrix. The body mass index (BMI), waist-to-hip ratio, triglyceride (TG), and years since a NSCLC diagnosis, were clustered together with the upward metabolites including N1-methyl-pyridone-3-carboxamide, Sphingosine, L-Palmitoylcarnitine, Hexacosanoyl carnitine, and Artonin C. The total cholesterol (TC), high-density lipoprotein (HDL) cholesterol, and low-density lipoprotein (LDL) cholesterol were clustered together with the downward metabolites, e.g., mostly the lipids and lipid-like molecules, organic acids and derivatives, and organic nitrogen compounds. In addition, age and second-hand exposure were clustered together with tigloidine, the alkaloid and derivative, which is the downward-expressed metabolite.

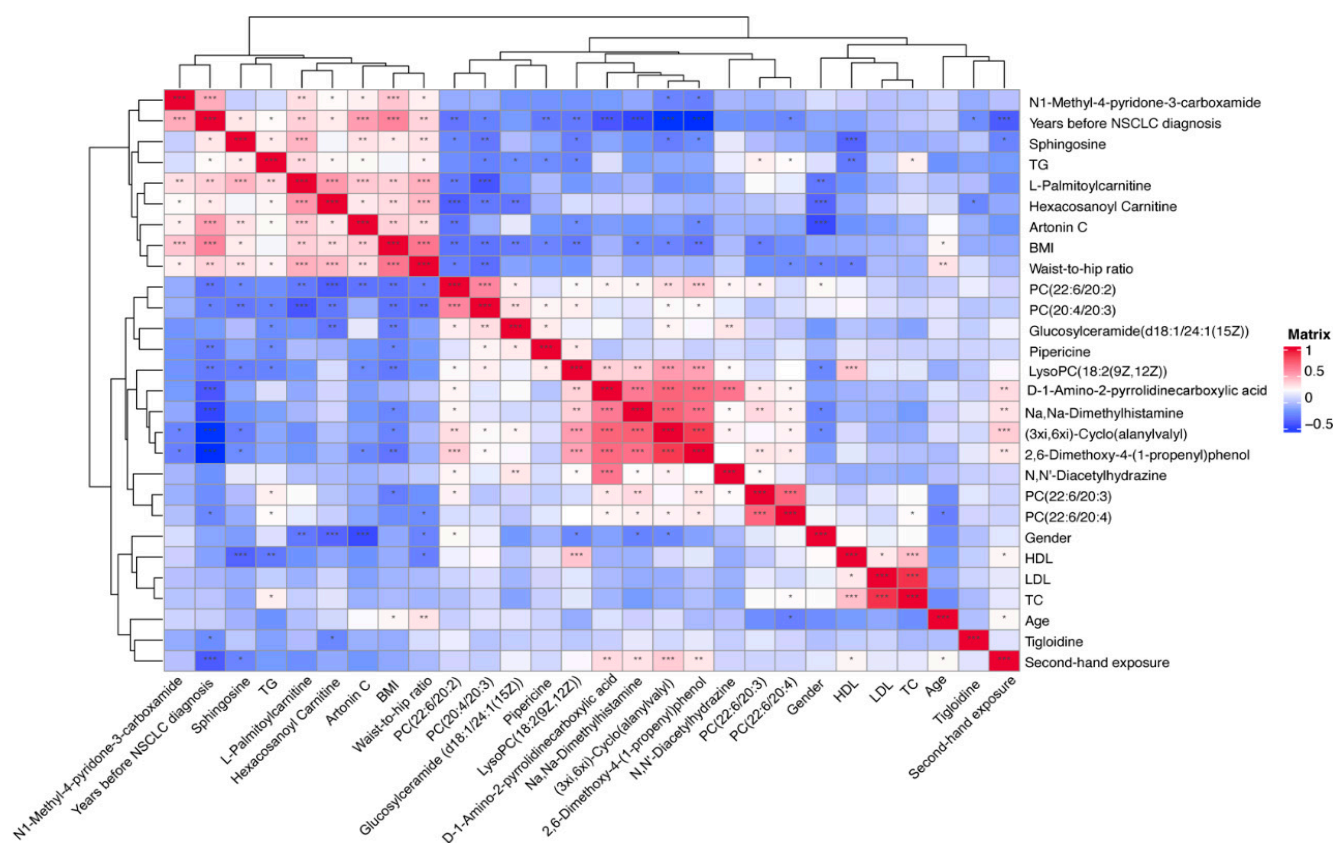


Figure 8. Spearman correlation matrix between significantly changed metabolites, classical lipids, and baseline characteristics in all subjects. Shades of blue represent a negative correlation coefficient; shades of red represent a positive correlation coefficient. Variables were ordered using a hierarchical clustering analysis. Threshold for significance testing of the Spearman correlation: $0.01 < p < 0.05$, *; $0.001 < p < 0.01$, **; $p < 0.001$, ***.

4. Discussion

To our knowledge, we report the first comprehensive serum metabolomics characteristics of NSCLC patients before the time of a disease diagnosis among Chinese community residents. In the PLS-DA, we observed that the first and second component scores were relatively low due to the possible reasons that the blood serum metabolomics data were highly multi-dimensional with high noise, while the sample size in our study was relatively small. However, we observed that the group separation was optimized with the PLS-DA compared with the PCA. We further performed an OPLS-DA to assist in identifying im-

portant metabolites. Using a strict standard ($VIP > 1$, a p -value < 0.05 and a q -value ≤ 0.2), we found 18 metabolites significantly-expressed between the groups after accounting for multiple comparisons. Among them, two were positively-associated with a risk of NSCLC and showed an upward trend in the NSCLC group, while four had negative associations and showed downward trends. We performed a LASSO regression and identified that N1-Methyl-4-pyridone-3-carboxamide, 2,6-Dimethoxy-4-(1-proponyl) phenol, tigloidine, artonin C, and pipericine could serve as the serum metabolite signatures of NSCLC with the AUC reaching 0.94 (a 95% CI: 0.89–0.98). Metabolic alterations identified in the KEGG pathway enrichment analysis imply that pathways including the choline metabolism, sphingolipid and glycerophospholipid metabolism, linoleic acid metabolism, and cell apoptosis and necroptosis were associated with a NSCLC risk before diagnosis.

Metabolic reprogramming is one of the hallmarks of cancer. Tumoral cells exhibit significant metabolic alterations and change their capability to metabolize carbohydrates, lipids and proteins to support cell proliferation [15]. The abnormality of lung tumors occurs in vessel structures that could limit the nutrient supply to the tumoral cells and produce hypoxia, which could induce multiple metabolic alterations.

In our study, we identified that D-1-Amino-2-pyrrolidinecarboxylic acid, a compound that belongs to proline and its derivatives, was significantly lowered in the NSCLC group. Proline has a unique role in metabolic regulation. Recent research has shown that proline metabolism plays a critical role in cancer metabolic reprogramming [16–18]. Proline has its α -amino group within a pyrrolidine ring and, thus, it is the sole proteinogenic secondary amino acid with its own metabolic pathways [12]. Proline dehydrogenase/proline oxidase (PRODH/POX) is encoded by tumor protein 53 (TP-53)-induced gene 6 (PIG-6), while TP-53 is the most frequently mutated gene in NSCLC [19,20]. In cancer cells, proline functions as a source for cancer cellular energy production and as an intermediate between the tricarboxylic acid (TCA) cycle and the urea cycle [21]. Researchers have already identified the critical role of proline catabolism in NSCLC. PRODH is activated to reduce the proline levels by the chromatin remodeling factor lymphoid-specific helicase (LSH), an epigenetic driver of NSCLC. PRODH promotes NSCLC tumorigenesis and could promote the expression of inflammatory genes and tumor cell proliferation [22]. In our study, the relative abundance of (3xi,6xi)-Cyclo(alanylvalyl) was significantly lowered in the NSCLC group. Moreover, (3xi,6xi)-Cyclo(alanylvalyl) belongs to the class of organic compounds known as alpha amino acids and derivatives and amino acids, such as methionine and tryptophan, have been reported to be associated with lung cancer [23]. Methionine is found to be involved in nucleotide biosynthesis via the one-carbon metabolism pathway and protein synthesis, which are the critical activities in cancer cells [24]. Additionally, methionine is also involved in the formation of glutathione, a biomarker of oxidative stress and it might contribute to chronic inflammation and cancer development [23].

We found that the relative concentration of N1-Methyl-4-pyridone-3-carboxamide was significantly higher among the NSCLC group. N1-Methyl-4-pyridone-3-carboxamide belongs to the class of organic compounds known as nicotinamides. Nicotinamide (NAM) is the water-soluble form of Vitamin B3 (niacin) and a precursor of nicotinamide-adenine dinucleotide (NAD^+), which takes part in several redox and non-redox reactions that regulates cellular energy metabolism [25]. The uncontrolled proliferation of cancer cells requires the reprogramming of their metabolism, which is critical for their survival and growth. In dehydrogenase reactions, NAD^+ acts as a co-enzyme to produce adenosine triphosphate (ATP). High NAD^+ could suppress the production of reactive oxygen species (ROS) and increase the mitochondrial quality; thus, it could protect against oxidative stress and improve cell survival [26]. In addition, we also identified nine lipids and lipid-like molecules that were significantly-expressed across the groups. Lipid metabolism is highly altered among proliferating cells because cancer cells increase de novo adipogenesis for membrane biosynthesis and signaling molecules instead of relying on the uptake of exogenous fatty acids (FA). We identified four types of phosphatidylcholine (PC) species. Phospholipids are the primary component of cell membranes and work as the second messengers in

intracellular signal transduction. Researchers have found that the phospholipids were altered in NSCLC. For example, Marien et al. used 167 NSCLC patients' tissue samples and profiled 179 phospholipids species that were altered between malignant and matched controls by a mass spectrometry-based approach. Their most important findings included a decrease in the sphingomyelins (SM) and an increase in several phosphatidylethanolamines (PEs) and PC species [27]. In our study, we also found phospholipid alterations among the NSCLC group; however, phospholipid alterations have also been described in other tumor types such as lymphoma [28]. Changes in phospholipid metabolism are not distinctive in NSCLC because it occurs early in carcinogenesis, irrespective of the cancer subtype and cancer stage [15].

Sphingosine (Sph), a metabolite of sphingomyelin, was over-expressed in the NSCLC group within our study. The metabolites of sphingomyelin, such as Sph, ceramide (Cer) and sphingosine-1-phosphate (S1P), form the lipid bilayers of cell membranes and play critical roles in the development of cancer [29]. Sph and Cer induce cell apoptosis, while S1P promotes cancer cell growth, survival, angiogenesis, and inflammation [30]. When the physiological condition is normal, the expression of Sph, Cer, and S1P should maintain a dynamic balance through enzymatic reactions that are described as forming a 'sphingolipid-rheostat' that is crucial for cell survival [31]. Sphingosine kinases (SPHKs) could phosphorylate Sph to S1P, which are the critical mediators of the 'sphingolipid-rheostat.' SPHK1 and SPHK2 are the two isoenzymes of sphingosine kinase, while SPHK1 has been reported as an effective pharmacologic target in anticancer therapy by inhibiting its activity of promoting the transition from Sph/Cer to S1P [32]. Song et al. reported that the expression of SPHK1 was markedly increased in NSCLC and associated with cancer progression and a poor survival of patients with NSCLC. The upregulation of SPHK1 significantly inhibited apoptosis, it was associated with the induction of antiapoptotic proteins, and it promoted the proliferation and migration of NSCLC cells *in vitro* and *in vivo*. In contrast, silencing SPHK1 expression can induce the apoptosis of NSCLC cells and increase the chemosensitivity of NSCLC to cytotoxic drugs [33]. Our results further proved that the metabolic balance of sphingolipids is of paramount importance in the cellular activities in NSCLC, which could be detected in the prior-to-diagnosis serum samples. In conjunction with the baseline characteristics, the metabolites identified by the LC-MS techniques could be used as a less-invasive, early disease detection tool among Chinese community residents.

Previous findings have showed that the metabolic pathways altered in lung cancer patients included those of glycolysis (the "Warburg effect"), lipids, choline phospholipid (the Kennedy pathway), one-carbon, amino acids, nucleotide metabolism, oxidative stress, and inflammation [34,35]. Consistent with the previous literature, our study proved that NSCLC patients notably presented alterations in the choline metabolism compared with the cancer-free participants. The GPL metabolism was also one disturbed pathway presented in comparisons between the NSCLC and cancer-free participants, followed by the sphingolipid metabolism. Moreover, we identified metabolomic characteristics specific for NSCLC prior to diagnosis that reflected 13 metabolites whose relative abundance showed a downward trend (mainly lipids and lipid-like molecules, organic acids and derivatives, organic nitrogen compounds, alkaloids and derivatives, and benzenoids). Additionally, five metabolites showed an upward trend (mainly the organoheterocyclic compound, phenylpropanoid and polyketide, lipids and lipid-like molecules, and the organic nitrogen compound). Among these 18 significantly-expressed metabolites, besides the metabolites involved in the KEGG metabolic pathway, we also identified a few metabolites, such as (3xi,6xi)-Cyclo(alanylvalyl) and Na₂Na-Dimethylhistamine, that were the potential biomarkers for the consumption of some specific foods. Moreover, a Spearman correlation matrix further indicated that the metabolites that showed an upward or downward trend tended to cluster together across the groups and related to each other with the baseline characteristics in the expected direction.

In cancer, metabolomics detects oncological developments through measurable metabolic profiles from the metabolic pathways through global metabolite variations [36].

A growing body of literature shows that abnormal choline metabolism is a distinguishing feature of carcinogenesis and tumor progression [37,38]. In the NSCLC group, we identified four types of phosphatidylcholine (PC), such as PC (22:6/20:3), PC (20:4/20:3), and PC (22:6/20:4), that showed a downward trend. Phosphatidylcholine (PC) and other phospholipids such as phosphatidylethanolamine (PE) are neutral lipids, forming the bilayer structure of cellular membranes and regulating membrane integrity, and they are the most abundant phospholipids in the cell membrane [39]. To synthesize PC and PE, cytidine diphosphate (CDP)-choline and CDP-ethanolamine are required. The biosynthesis and hydrolysis of PC could mediate the mitogenic signal transduction events in cells. Phosphocholine (PCho), diacylglycerol (DAG), and arachidonic acid metabolites, which are the products of choline phospholipid metabolism, might function as second messengers that are important to the mitogenic activity [40,41]. Growth factor stimulation, cytokines, oncogenes, or the requirements for eicosanoid production also regulate choline phospholipid metabolism [42–44]. A network of transporter systems and enzymes are involved in the choline phospholipid metabolism that is deregulated in cancer cells. In addition, the tumor microenvironment also affects choline metabolism. Hypoxia and an acidic extracellular pH could lead to the abnormal physiological environments caused by tumors. Studies have found that an acidic extracellular pH could significantly increase the glycerophosphocholine (GPC) level and decrease the PCho levels in perfused mammalian cells [45].

Sphingolipid metabolism was also a significantly altered pathway between the groups, apart from the choline metabolism. Sphingolipids, the structural molecules of cell membranes, play an important role in regulating cancer cell death and survival by controlling the cell-signaling functions. Our data showed that the critical metabolites involved in this pathway were sphingosine and glucosylceramide, and the abundance was significantly different between the groups. In the NSCLC group, the relative abundance of sphingosine was significantly higher than the controls, while glucosylceramide showed a downward trend. In cells, cellular stress could induce the generation of sphingosine and/or ceramide to mediate cancer cell death by activating the *de novo* synthesis pathways, sphingomyelin hydrolysis or the salvage pathway, which could explain why we observed an increased abundance of sphingosine in NSCLC. Many tumors could increase the ceramide metabolism by increasing the activities of glucosyl-ceramide synthase (GCS), sphingomyelin synthase (SMS), and ceramide kinase (CERK), etc., which could increase the generation of sphingolipids with pro-survival functions [46]. In addition, the accumulation of ceramide in response to cellular stress could mediate cancer cell death through the induction of apoptosis, necroptosis, autophagy, and ER stress. Endogenously generated ceramide is an inducer of apoptosis regulated by various mechanisms in a cell-type-dependent and/or context-dependent manner [47–49]. In contrast, there are studies that have demonstrated that ceramide could protect some cancer cells from cell death [50,51]. Overall, ceramide mediates apoptosis in most cancer cells via mitochondrial membrane perturbation and/or influencing the cell-death signaling; however, it might have protective effects against apoptosis depending on the downstream ceramide targets, the subcellular localization of ceramide and the type of stress stimuli. Our data showed that aberrant sphingolipid metabolism with an abnormal accumulation of sphingolipids and ceramides could be observed among the Chinese residents' serum samples before their diagnosis.

In our study, we found four types of PC that showed a downward trend in the NSCLC group when compared with the cancer-free controls. Underlying these phenotypic alterations in the choline metabolism might indicate the degradation of PC, resulting in the dysplasia of the surrounding substances of the cell membrane. Same as with the previous findings, the aberrant expression of PC or its related compounds could be detected by a non-invasive LC-MS to detect NSCLC at an early stage; however, abnormal choline phospholipid metabolism has also been observed in many other types of cancer, not heterogeneous to the NSCLC [38,52]. We observed that the AUC of the serum metabolite signatures combined with the smoking history, BMI, and TG indicators for identifying NSCLC could reach 0.99 (a 95% CI: 0.98–1.00). This proved that the serum metabolite

fingerprints combined with the baseline characteristics could be an effective tool to identify an NSCLC presence within our study. However, future validations are warranted to test whether our results could be applicable to populations from other districts in China.

Several research groups have recently applied metabolomic techniques to unveil the metabolic alterations associated with lung cancer, with most of the studies being case-control studies. Hori et al. detected 58 metabolites in serum using chromatography/mass (GC/MS) and found 23 differentially-expressed metabolites in the Japanese population [8]. Maeda et al. studied 21 amino acids in NSCLC patients' plasma with LC/MS and proved that the amino acid profiles could be used for screening NSCLC among the Japanese population [53]. Jordan et al. applied a nuclear magnetic resonance (NMR) to measure 21 metabolites and showed that serum metabolomics could differentiate lung cancer between patients and healthy controls [54]. Neither the study design nor the sample population was different to ours. To avoid reverse causation, we used a prospective nested case-control study to illustrate the serum metabolomic alterations before diagnosing NSCLC. Schult TA et al. established the magnetic resonance spectroscopy (MRS)-based metabolomics predictive models to identify the lung cancer presence in the U.S. population [7]. However, serum metabolomic profiling prior to a diagnosis with NSCLC is scarce in China and, thus, the metabolomic predictive models could not be established. Our findings could have implications for lung cancer screening in China. Considering that the metabolite fingerprints might improve the identifying of participants with the highest risk of NSCLC, who could benefit the most from the screening, a metabolomic predictive tool validated in other cohorts is warranted in future studies. The public health implication of this study is the importance and the feasibility of the non-invasive and early-detection of NSCLC with LC-MS among the Chinese community residents. Healthcare centers throughout the nation could evaluate the use of spectroscopic imaging to assist in preventing and diagnosing NSCLC. Additionally, exploring choline phospholipid metabolism and sphingolipid metabolism might help to identify new therapeutic targets for NSCLC treatments. Because the LC-MS technique is relatively expensive and qualified practitioners are in short supply, the cost-effectiveness of its ability for early-detecting NSCLC could be evaluated before being widely applied.

The strength of our study is this prospective setting. The reverse causation could be mitigated with our study design; however, several limitations should be considered. First, we did not validate the results of the identified metabolites with an untargeted LC-MS technique to reconfirm the robustness of the metabolites that we measured. Although we applied a standard LC-MS data acquisition operation to identify the significantly-expressed metabolites, the validation of the results should be paramount and must be improved in our future studies. Second, from the perspective of statistics, the sample sizes for the NSCLC and cancer-free controls were relatively small. Third, although we identified metabolomic alterations before a diagnosis with NSCLC among our study participants, and conducted an internal cross validation of our risk prediction model, the extrapolation has not been established through external validation in other participants from our cohort and this should be improved in future research. Last but not least, there is a possibility that the results could be affected by other confounders that we did not consider. For example, the serum metabolite profiles could have been affected by diet composition and circadian rhythm. We did identify a few metabolites that might be the potential biomarkers for the consumption of certain foods; however, we did not assess the impact of dietary composition in our study, which should be improved in our future research.

5. Conclusions

In summary, we reported the metabolome profile of NSCLC patient blood serum samples collected before disease diagnosis and compared the results with those observed in serum samples obtained from healthy control subjects. We identified 18 significantly-expressed metabolites, with most of them being lipid and lipid-like molecules, organic acids, and nitrogen compounds. Additionally, choline metabolism exhibits a markedly

metabolic disturbance in NSCLC patients, followed by sphingolipid metabolism and glycerophospholipid metabolism. Our data point to the metabolisms mentioned above and the related metabolites as the potential critical targets for cancer interventions that could help with the early diagnosis of NSCLC patients. Our results demonstrate the potential feasibility of applying LC-MS-based serum metabolomics to predict the presence of NSCLC combined with baseline characteristics, such as smoking habits and the exposure history to second-hand smoking. With future validation, metabolomic techniques might be used as an effective and less-invasive tool to supplement the screening programs of NSCLC in the Chinese population.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/metabo12100906/s1>, Table S1: NSCLCs vs Cancer-free controls: Differentially expressed metabolites; Table S2: NSCLCs vs Cancer-free controls: KEGG Pathway Enrichment Analysis.

Author Contributions: G.Z. and Y.J. contributed to the study concept. Y.X. contributed to the data analysis, manuscript draft, and revision. Q.Z. and X.L. contributed to supervising the enrollment of cohort participants and blood sample collection. Y.W. and X.S. contributed to the follow-up data collection through the Health Information System. J.Z. and Y.Y. contributed to data cleaning and quality control. K.X. contributed to the review of statistical analysis. G.Z. contributed to the paper review and editing. Y.X. and G.Z. verified the underlying data. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Key Research and Development Program of China (2017YFC0907000), the Local High Level Discipline Construction Project of Shanghai, and the Shanghai New Three-year Action Plan for Public Health (Grant No. GWV-10.1-XK16).

Institutional Review Board Statement: This study was performed in conformity with the Declaration of Helsinki and was approved by the Institutional Review Board of Fudan University, School of Public Health, China (Authorization number: IRB#2016-04-0586).

Informed Consent Statement: Written informed consent was obtained from all participants.

Data Availability Statement: Raw data in this study are available from the corresponding author upon reasonable request. All the data that support the findings of this study are free to obtain in the paper or in the supplementary materials.

Acknowledgments: We thank the staff of the non-communicable disease department at Songjiang CDC for the provision and correction of the cancer registry data. We thank the members of the SSCAB research group who participated and devoted themselves during the baseline investigation of SSACB from the department of epidemiology (School of Public Health, Fudan University) to organize the data collection work and help construct the biobank. We thank all the participants in our cohort and the investigators at healthcare centers who enrolled the participants. We also thank the staff of Shanghai Biotree Biotech Co., Ltd. for the LC-MS data analysis.

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

Appendix A

Appendix A.1. Untargeted Metabolomics Profiling

Appendix A.1.1. LC-MS Metabolites Extraction

A total of 100 μ L of a sample was transferred to an EP tube. A 600 μ L amount of the extract solution (e.g., acetonitrile: methanol = 1:1; containing isotopically-labelled internal standard) was added into the EP tube. The mixed samples were vortexed for 30 s, sonicated in the ice-water bath for 10 min, incubated at -40 °C for 1 h to precipitate the proteins, and centrifuged at 120,000 rpm at 4 °C for 15 min. The supernatant was transferred to a fresh glass vial for the LC/MS analysis. The quality control (QC) samples were prepared by using an equal aliquot of the supernatant from all the samples.

Appendix A.1.2. LC-MS/MS Analysis

The liquid chromatography-mass spectrometry (LC-MS) analyses were performed using a UHPLC system (Vanquish, Thermo Fisher Scientific) with a Waters ACQUITY UPLC BEH Amide column (2.1 mm × 100, 1.7 μm) coupled to a Q Exactive HFX mass spectrometer (Orbitrap MS, Thermo Fisher Scientific).

Amounts of 25 mmol/L of ammonium acetate and 25 mmol/L of ammonia hydroxide in water (pH = 9.75) (A), and acetonitrile (B), consisted of the mobile phase. The temperature of the auto-sampler was set at 4 °C. The injection volume was 2 μL. The Q Exactive HFX mass spectrometer was used to acquire MS/MS spectra on the information-dependent acquisition (IDA) mode and was controlled by the acquisition software (Xcalibur, Thermo). In this mode, the Xcalibur continuously evaluated the full scan MS spectrum. The ESI source conditions were set as follows: sheath gas flow rate: 30 Arb, aux gas flow rate: 25 Arb, capillary temperature: 350 °C, full ms resolution: 60,000, MS/MS resolution: 7500, collision energy: 10/30/60 in the NCE mode, and spray voltage: 3.6 kV (positive).

To preprocess the raw data and its annotation, ProteoWizard was used to convert the raw data to a mzXML format. The data were processed with the Biotree in-house program that was developed using R packages and based on XCMS for peak detection, extraction, alignment, and integration. Next, the in-house compound database (BiotreeDB) was applied for the metabolite annotation. The cutoff value for an annotation was set at 0.3.

Appendix A.1.3. Quality Control Process

One quality control (QC) sample was run after every ten samples during the data acquisition process. The stability and reliability of the instruments were assessed by the difference in the response peak intensity in the internal standard between the QC samples. In addition, the peak detected in the internal standard of the blank sample was assessed to evaluate the compound residue. No peak was detected among the internal standard of the blank samples, indicating the substance residue was controlled successfully, and the cross contamination between the samples was under control. For the data quality control, we assessed the distribution of the QC samples in the score plot of a principal component analysis. As shown in the Supplementary Figure S1, different QC samples were clustered together, suggesting the stability of the metabolite extraction process. Additionally, we performed a correlation analysis between the different QC samples. The results showed that the QC samples were highly correlated with each other, suggesting the data quality was acceptable.

Appendix B

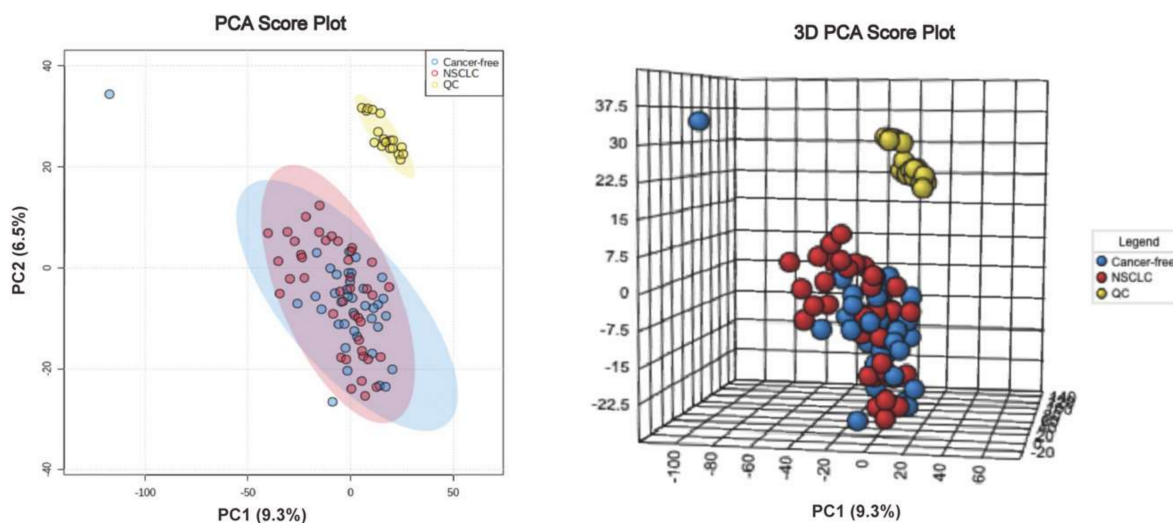


Figure A1. PCA score plot and 3D PCA score plot between NSCLC and cancer-free controls.

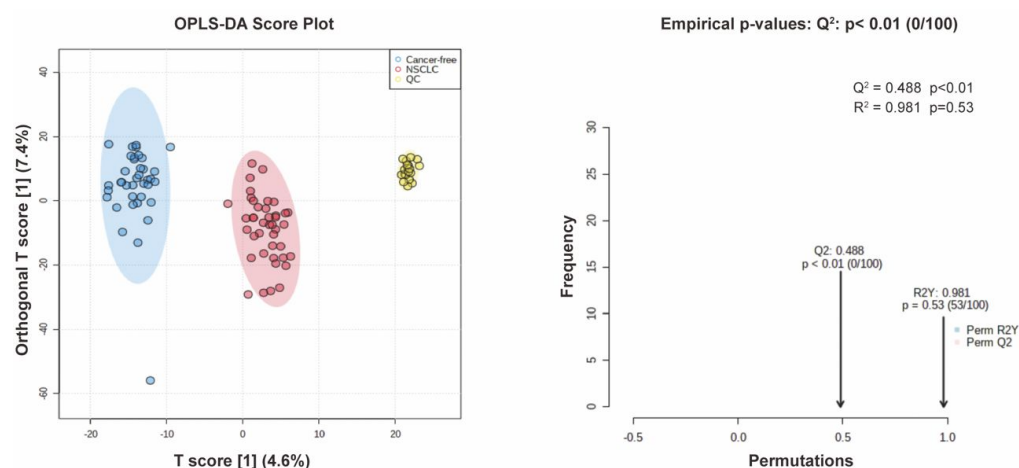


Figure A2. OPLS-DA score plot between NSCLC and cancer-free controls and the permutation test of OPLS-DA.

References

- Chen, W.; Zheng, R.; Baade, P.D.; Zhang, S.; Zeng, H.; Bray, F.; Jemal, A.; Yu, X.Q.; He, J. Cancer statistics in China, 2015. *CA Cancer J. Clin.* **2016**, *66*, 115–132. [[CrossRef](#)] [[PubMed](#)]
- Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
- Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2020. *CA Cancer J. Clin.* **2020**, *70*, 7–30. [[CrossRef](#)] [[PubMed](#)]
- Oudkerk, M.; Liu, S.; Heuvelmans, M.A.; Walter, J.E.; Field, J.K. Lung cancer LDCT screening and mortality reduction—Evidence, pitfalls and future perspectives. *Nat. Rev. Clin. Oncol.* **2021**, *18*, 135–151. [[CrossRef](#)]
- Toumazis, I.; Bastani, M.; Han, S.S.; Plevritis, S.K. Risk-Based lung cancer screening: A systematic review. *Lung Cancer* **2020**, *147*, 154–186. [[CrossRef](#)]
- Aberle, D.R.; Adams, A.M.; Berg, C.D.; Black, W.C.; Clapp, J.D.; Fagerstrom, R.M.; Gareen, I.F.; Gatsonis, C.; Marcus, P.M.; Sicks, J.D. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* **2011**, *365*, 395–409. [[PubMed](#)]
- Schult, T.A.; Lauer, M.J.; Berker, Y.; Cardoso, M.R.; Vandergrift, L.A.; Habbel, P.; Nowak, J.; Taupitz, M.; Aryee, M.; Mino-Kenudson, M.A.; et al. Screening human lung cancer with predictive models of serum magnetic resonance spectroscopy metabolomics. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2110633118. [[CrossRef](#)] [[PubMed](#)]
- Hori, S.; Nishiumi, S.; Kobayashi, K.; Shinohara, M.; Hatakeyama, Y.; Kotani, Y.; Hatano, N.; Maniwa, Y.; Nishio, W.; Bamba, T.; et al. A metabolomic approach to lung cancer. *Lung Cancer* **2011**, *74*, 284–292. [[CrossRef](#)]
- Wittmann, B.M.; Stirdivant, S.M.; Mitchell, M.W.; Wulff, J.E.; McDunn, J.E.; Li, Z.; Dennis-Barrie, A.; Neri, B.P.; Millburn, M.V.; Lotan, Y.; et al. Bladder cancer biomarker discovery using global metabolomic profiling of urine. *PLoS ONE* **2014**, *9*, e115870. [[CrossRef](#)]
- Yu, L.; Jiang, C.; Huang, S.; Gong, X.; Wang, S.; Shen, P. Analysis of urinary metabolites for breast cancer patients receiving chemotherapy by CE-MS coupled with on-line concentration. *Clin. Biochem.* **2013**, *46*, 1065–1073. [[CrossRef](#)]
- Playdon, M.C.; Ziegler, R.G.; Sampson, J.N.; Stolzenberg-Solomon, R.; Thompson, H.J.; Irwin, M.L.; Mayne, S.T.; Hoover, R.N.; Moore, S.C. Nutritional metabolomics and breast cancer risk in a prospective study. *Am. J. Clin. Nutr.* **2017**, *106*, 637–649. [[CrossRef](#)]
- Zhao, Q.; Chen, B.; Wang, R.; Zhu, M.; Shao, Y.; Wang, N.; Liu, X.; Zhang, T.; Jiang, F.; Wang, W.; et al. Cohort profile: Protocol and baseline survey for the Shanghai Suburban Adult Cohort and Biobank (SSACB) study. *BMJ Open* **2020**, *10*, e035430. [[CrossRef](#)]
- The Kyoto Encyclopedia of Genes and Genomes Database. Available online: <http://www.kegg.jp> (accessed on 5 December 2021).
- The Human Metabolome Database. Available online: <https://hmdb.ca/> (accessed on 25 December 2021).
- Merino Salvador, M.; Gómez de Cedrón, M.; Moreno Rubio, J.; Falagán Martínez, S.; Sánchez Martínez, R.; Casado, E.; Ramírez de Molina, A.; Sereno, M. Lipid metabolism and lung cancer. *Crit. Rev. Oncol. Hematol.* **2017**, *112*, 31–40. [[CrossRef](#)] [[PubMed](#)]
- Elia, I.; Broekaert, D.; Christen, S.; Boon, R.; Radaelli, E.; Orth, M.F.; Verfaillie, C.; Grünewald, T.G.P.; Fendt, S.M. Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nat. Commun.* **2017**, *8*, 15267. [[CrossRef](#)]
- Liu, W.; Hancock, C.N.; Fischer, J.W.; Harman, M.; Phang, J.M. Proline biosynthesis augments tumor cell growth and aerobic glycolysis: Involvement of pyridine nucleotides. *Sci. Rep.* **2015**, *5*, 17206. [[CrossRef](#)]
- Sahu, N.; Dela Cruz, D.; Gao, M.; Sandoval, W.; Haverty, P.M.; Liu, J.; Stephan, J.P.; Haley, B.; Classon, M.; Hatzivassiliou, G.; et al. Proline Starvation Induces Unresolved ER Stress and Hinders mTORC1-Dependent Tumorigenesis. *Cell Metab.* **2016**, *24*, 753–761. [[CrossRef](#)]

19. Polyak, K.; Xia, Y.; Zweier, J.L.; Kinzler, K.W.; Vogelstein, B. A model for p53-induced apoptosis. *Nature* **1997**, *389*, 300–305. [[CrossRef](#)]
20. Canale, M.; Andrikou, K.; Priano, I.; Cravero, P.; Pasini, L.; Urbini, M.; Delmonte, A.; Crinò, L.; Bronte, G.; Ulivi, P. The Role of TP53 Mutations in EGFR-Mutated Non-Small-Cell Lung Cancer: Clinical Significance and Implications for Therapy. *Cancers (Basel)* **2022**, *14*, 1143. [[CrossRef](#)] [[PubMed](#)]
21. Adams, E. Metabolism of proline and of hydroxyproline. *Int. Rev. Connect. Tissue Res.* **1970**, *5*, 1–91. [[PubMed](#)]
22. Liu, Y.; Mao, C.; Wang, M.; Liu, N.; Ouyang, L.; Liu, S.; Tang, H.; Cao, Y.; Liu, S.; Wang, X.; et al. Cancer progression is mediated by proline catabolism in non-small cell lung cancer. *Oncogene* **2020**, *39*, 2358–2376. [[CrossRef](#)]
23. Lee, K.B.; Ang, L.; Yau, W.P.; Seow, W.J. Association between Metabolites and the Risk of Lung Cancer: A Systematic Literature Review and Meta-Analysis of Observational Studies. *Metabolites* **2020**, *10*, 362. [[CrossRef](#)] [[PubMed](#)]
24. Locasale, J.W. Serine, glycine and one-carbon units: Cancer metabolism in full circle. *Nat. Rev. Cancer* **2013**, *13*, 572–583. [[CrossRef](#)]
25. Nikas, I.P.; Paschou, S.A.; Ryu, H.S. The Role of Nicotinamide in Cancer Chemoprevention and Therapy. *Biomolecules* **2020**, *10*, 477. [[CrossRef](#)]
26. Song, S.B.; Jang, S.Y.; Kang, H.T.; Wei, B.; Jeoun, U.W.; Yoon, G.S.; Hwang, E.S. Modulation of Mitochondrial Membrane Potential and ROS Generation by Nicotinamide in a Manner Independent of SIRT1 and Mitophagy. *Mol. Cells* **2017**, *40*, 503–514. [[PubMed](#)]
27. Marien, E.; Meister, M.; Muley, T.; Fieuws, S.; Bordel, S.; Derua, R.; Spraggins, J.; Van de Plas, R.; Dehairs, J.; Wouters, J.; et al. Non-small cell lung cancer is characterized by dramatic changes in phospholipid profiles. *Int. J. Cancer* **2015**, *137*, 1539–1548. [[CrossRef](#)] [[PubMed](#)]
28. Eberlin, L.S.; Gabay, M.; Fan, A.C.; Gouw, A.M.; Tibshirani, R.J.; Felsher, D.W.; Zare, R.N. Alteration of the lipid profile in lymphomas induced by MYC overexpression. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10450–10455. [[CrossRef](#)]
29. Saddoughi, S.A.; Song, P.; Ogretmen, B. Roles of bioactive sphingolipids in cancer biology and therapeutics. *Subcell. Biochem.* **2008**, *49*, 413–440.
30. Haddadi, N.; Lin, Y.; Simpson, A.M.; Nassif, N.T.; McGowan, E.M. "Dicing and Splicing" Sphingosine Kinase and Relevance to Cancer. *Int. J. Mol. Sci.* **2017**, *18*, 1891. [[CrossRef](#)]
31. Newton, J.; Lima, S.; Maceyka, M.; Spiegel, S. Revisiting the sphingolipid rheostat: Evolving concepts in cancer therapy. *Exp. Cell Res.* **2015**, *333*, 195–200. [[CrossRef](#)]
32. Zheng, X.; Li, W.; Ren, L.; Liu, J.; Pang, X.; Chen, X.; Kang, D.; Wang, J.; Du, G. The sphingosine kinase-1/sphingosine-1-phosphate axis in cancer: Potential target for anticancer therapy. *Pharmacol. Ther.* **2019**, *195*, 85–99. [[CrossRef](#)]
33. Song, L.; Xiong, H.; Li, J.; Liao, W.; Wang, L.; Wu, J.; Li, M. Sphingosine kinase-1 enhances resistance to apoptosis through activation of PI3K/Akt/NF- κ B pathway in human non-small cell lung cancer. *Clin. Cancer Res.* **2011**, *17*, 1839–1849. [[CrossRef](#)]
34. Zhang, X.; Zhu, X.; Wang, C.; Zhang, H.; Cai, Z. Non-targeted and targeted metabolomics approaches to diagnosing lung cancer and predicting patient prognosis. *Oncotarget* **2016**, *7*, 63437–63448. [[CrossRef](#)]
35. Seow, W.J.; Shu, X.O.; Nicholson, J.K.; Holmes, E.; Walker, D.I.; Hu, W.; Cai, Q.; Gao, Y.T.; Xiang, Y.B.; Moore, S.C.; et al. Association of Untargeted Urinary Metabolomics and Lung Cancer Risk Among Never-Smoking Women in China. *JAMA Netw. Open* **2019**, *2*, e1911970. [[CrossRef](#)]
36. Vandergrift, L.A.; Decelle, E.A.; Kurth, J.; Wu, S.; Fuss, T.L.; DeFeo, E.M.; Halpern, E.F.; Taupitz, M.; McDougal, W.S.; Olumi, A.F.; et al. Metabolomic Prediction of Human Prostate Cancer Aggressiveness: Magnetic Resonance Spectroscopy of Histologically Benign Tissue. *Sci. Rep.* **2018**, *8*, 4997. [[CrossRef](#)] [[PubMed](#)]
37. Sonkar, K.; Ayyappan, V.; Tressler, C.M.; Adelaja, O.; Cai, R.; Cheng, M.; Glunde, K. Focus on the glycerophosphocholine pathway in choline phospholipid metabolism of cancer. *NMR Biomed.* **2019**, *32*, e4112. [[CrossRef](#)] [[PubMed](#)]
38. Glunde, K.; Bhujwala, Z.M.; Ronen, S.M. Choline metabolism in malignant transformation. *Nat. Rev. Cancer* **2011**, *11*, 835–848. [[CrossRef](#)] [[PubMed](#)]
39. Mountford, C.E.; Wright, L.C. Organization of lipids in the plasma membranes of malignant and stimulated cells: A new model. *Trends Biochem. Sci.* **1988**, *13*, 172–177. [[CrossRef](#)]
40. Cuadrado, A.; Carnero, A.; Dolfi, F.; Jiménez, B.; Lacal, J.C. Phosphorylcholine: A novel second messenger essential for mitogenic activity of growth factors. *Oncogene* **1993**, *8*, 2959–2968. [[PubMed](#)]
41. Price, B.D.; Morris, J.D.; Marshall, C.J.; Hall, A. Stimulation of phosphatidylcholine hydrolysis, diacylglycerol release, and arachidonic acid production by oncogenic ras is a consequence of protein kinase C activation. *J. Biol. Chem.* **1989**, *264*, 16638–16643. [[CrossRef](#)]
42. Besterman, J.M.; Duronio, V.; Cuatrecasas, P. Rapid formation of diacylglycerol from phosphatidylcholine: A pathway for generation of a second messenger. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 6785–6789. [[CrossRef](#)] [[PubMed](#)]
43. Bogin, L.; Papa, M.Z.; Polak-Charcon, S.; Degani, H. TNF-induced modulations of phospholipid metabolism in human breast cancer cells. *Biochim. Biophys. Acta* **1998**, *1392*, 217–232. [[CrossRef](#)]
44. Aboagye, E.O.; Bhujwala, Z.M. Malignant transformation alters membrane choline phospholipid metabolism of human mammary epithelial cells. *Cancer Res.* **1999**, *59*, 80–84. [[PubMed](#)]
45. Galons, J.P.; Job, C.; Gillies, R.J. Increase of GPC levels in cultured mammalian cells during acidosis. A 31P MR spectroscopy study using a continuous bioreactor system. *Magn. Reson. Med.* **1995**, *33*, 422–426. [[CrossRef](#)]
46. Ogretmen, B. Sphingolipid metabolism in cancer signalling and therapy. *Nat. Rev. Cancer* **2018**, *18*, 33–50. [[CrossRef](#)] [[PubMed](#)]

47. White-Gilbertson, S.; Mullen, T.; Senkal, C.; Lu, P.; Ogretmen, B.; Obeid, L.; Voelkel-Johnson, C. Ceramide synthase 6 modulates TRAIL sensitivity and nuclear translocation of active caspase-3 in colon cancer cells. *Oncogene* **2009**, *28*, 1132–1141. [[CrossRef](#)] [[PubMed](#)]
48. Siskind, L.J.; Feinstein, L.; Yu, T.; Davis, J.S.; Jones, D.; Choi, J.; Zuckerman, J.E.; Tan, W.; Hill, R.B.; Hardwick, J.M.; et al. Anti-apoptotic Bcl-2 Family Proteins Disassemble Ceramide Channels. *J. Biol. Chem.* **2008**, *283*, 6622–6630. [[CrossRef](#)]
49. Blom, T.; Li, S.; Dichlberger, A.; Bäck, N.; Kim, Y.A.; Loizides-Mangold, U.; Riezman, H.; Bittman, R.; Ikonen, E. LAPT4B facilitates late endosomal ceramide export to control cell death pathways. *Nat. Chem. Biol.* **2015**, *11*, 799–806. [[CrossRef](#)]
50. Schiffmann, S.; Sandner, J.; Birod, K.; Wobst, I.; Angioni, C.; Ruckhäberle, E.; Kaufmann, M.; Ackermann, H.; Lötsch, J.; Schmidt, H.; et al. Ceramide synthases and ceramide levels are increased in breast cancer tissue. *Carcinogenesis* **2009**, *30*, 745–752. [[CrossRef](#)]
51. Senkal, C.E.; Ponnusamy, S.; Manevich, Y.; Meyers-Needham, M.; Saddoughi, S.A.; Mukhopadhyay, A.; Dent, P.; Bielawski, J.; Ogretmen, B. Alteration of ceramide synthase 6/C16-ceramide induces activating transcription factor 6-mediated endoplasmic reticulum (ER) stress and apoptosis via perturbation of cellular Ca²⁺ and ER/Golgi membrane network. *J. Biol. Chem.* **2011**, *286*, 42446–42458. [[CrossRef](#)]
52. Podo, F. Tumour phospholipid metabolism. *NMR Biomed.* **1999**, *12*, 413–439. [[CrossRef](#)]
53. Maeda, J.; Higashiyama, M.; Imaizumi, A.; Nakayama, T.; Yamamoto, H.; Daimon, T.; Yamakado, M.; Imamura, F.; Kodama, K. Possibility of multivariate function composed of plasma amino acid profiles as a novel screening index for non-small cell lung cancer: A case control study. *BMC Cancer* **2010**, *10*, 690. [[CrossRef](#)] [[PubMed](#)]
54. Jordan, K.W.; Adkins, C.B.; Su, L.; Halpern, E.F.; Mark, E.J.; Christiani, D.C.; Cheng, L.L. Comparison of squamous cell carcinoma and adenocarcinoma of the lung by metabolomic analysis of tissue-serum pairs. *Lung Cancer* **2010**, *68*, 44–50. [[CrossRef](#)] [[PubMed](#)]