Research Paper

Metabolomic profiling of dried blood spots reveals gender-specific discriminant models for the diagnosis of small cell lung cancer

Li Yu[1], Kefeng Li[2], Xiangmin Li[1], Chao Guan[1], Tingting Sun[1], Xiaoye Zhang[1]

[1]Department of Clinical Oncology, Shengjing Hospital of China Medical University, Shenyang, Liaoning 110004, China
[2]School of Medicine, University of California, San Diego, CA 92103, USA

Correspondence to: Xiaoye Zhang; **email:** zhangxy1@sj-hospital.org
Keywords: small cell lung cancer, dried blood spot, metabolomics, differential diagnosis, gender differences
Received: September 27, 2019 **Accepted:** December 24, 2019 **Published:** January 12, 2010

ABSTRACT

The accurate diagnosis of small cell lung cancer (SCLC) at initial presentation is essential to ensure appropriate treatment. No validated blood biomarkers that could distinguish SCLC from non-small cell lung cancer (NSCLC) has yet been developed. Dried blood spot (DBS) microsampling has gained increasing interest in biomarkers discovery. In this study, we first performed metabolomic profiling of DBS samples from 37 SCLC, 40 NSCLC, and 37 controls. Two gender-specific multianalyte discriminant models were established for males and females, respectively to distinguish SCLC from NSCLC and controls. The receiver operator characteristic (ROC) curve analysis showed the diagnostic accuracy of 95% (95% CI: 83%-100%) in males SCLC using five metabolites in DBS and 94% (95% CI: 74%-100%) for females using another set of five metabolites. The robustness of the models was confirmed by the random permutation tests ($P < 0.01$ for both). The performance of the discriminant models was further evaluated using a validation cohort with 78 subjects. The developed discriminant models yielded an accuracy of 91% and 81% for males and females, respectively, in the validation cohort. Our results highlighted the potential clinical utility of the metabolomic profiling of DBS as a convenient and effective approach for the diagnosis of SCLC.

INTRODUCTION

Lung cancer is the most common malignant tumor that threatens human health and one of the leading causes of cancer-related death in both men and women [1]. There are two major subtypes of lung cancer including small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLC accounts for about 10–15% of all lung cancer diagnosis. The two subtypes of lung cancer behave differently and thus are treated very differently [2]. Therefore, it is essential for the accurate classification of the subtypes of lung cancer at initial presentation to ensure the selection of appropriate treatment approaches.

Currently, the diagnosis of SCLC primarily relies on the histological and cytological results from tumor tissue biopsies. In many cases, the cytologic features are equivocal due to the sampling issues, fixation artifacts and morphological variability of SCLC tumor cells [3]. Immunohistochemistry (IHC) might be helpful to increase the confidence of the pathologists in SCLC classification. However, a recent study reported that the interobserver agreement for the interpretation of IHC slides was as low as 40% [4]. Moreover, SCLC often originates in the submucosal tissue, which makes the accurate collection of tumor samples challenging [5]. The analysis of tumor markers in biological fluids such as the serum, plasma, and urine has remarkable advantages over histologic differentiation.

The serum level of neuron-specific enolase (NSE) is the tumor biomarker currently used as the reference for the diagnosis, prognosis, and follow-up of SCLC [6]. Unfortunately, recent studies reported that NSE has low sensitivity and the elevated levels of NSE were also

observed in about 25% of patients with NSCLC [7]. There is an urgent need for a simple, specific and accurate approach to increase the diagnostic certainty for SCLC.

Metabolites are the small molecules (<2000 Da) with a wide range of functions in cells and organisms such as energy production, signal transduction, and apoptosis. Metabolites are the direct "readout" of the overall physiological status and very sensitive to pathological changes [8]. Metabolomics, the profiling of the metabolites in biofluids, has been increasingly used as a tool for the biomarker discovery and understanding of the pathogenesis of the diseases [9]. However, previous metabolomic studies in lung cancer primarily focused on discriminating lung cancer patients from non-cancer participants [10]. Most of the potential metabolic biomarkers identified in these studies were energy-related metabolites such as glucose and glutamate. These biomarkers were actually reported to be abnormal in other cancers, not limited to lung cancer [10]. There is little information regarding the classification of lung cancer subtypes using circulating metabolite profiles. Recently, with the innovative development of analytical techniques and machine learning-based bioinformatic tools, next-generation metabolomics is now more robust and sensitive allowing the accurate detection of subtle differences between groups [9, 11]. In our previous study, using next-generation metabolomics, we revealed the unique metabolic features in SCLC cells compared to normal human bronchial epithelial cells (HBECs) and NSCLC cells [12].

DBS is a minimally invasive microsampling strategy that involves the collection of a drop of blood on a filter card. Compared to the conventional sampling procedures, DBS has several advantages including the low amount of blood collected, easiness of blood collection, the possibility to store and ship collected samples at room temperature and low cost and logistics [13]. The most common clinical use of DBS is for the screening of inborn errors of metabolism in newborns. In recent years, there is a growing interest in biomarker development and validation based on omic-based analysis of DBS. Metabolomic analysis of DBS has been shown high sensitivity and accuracy for rapid breast [14] and colorectal cancer detection [15]. However, the application of DBS metabolomic approach for the differentiation of lung cancer subtypes has not been reported.

In this study, we conducted a comparative metabolomic analysis of DBS collected from patients with SCLC, NSCLC and healthy controls using liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). We identified the unique metabolic features in the DBS of both male and female SCLC patients. The selected metabolites biomarkers for distinguishing SCLC from NSCLC and controls were further validated using another cohort.

RESULTS

Subjects characteristics

The study flow was shown in Figure 1. In the discovery phase, a total of 114 subjects were recruited including 40 patients with NSCLC, 37 patients with SCLC and 37 noncancer controls (Table 1). In the validation set, 78 cases were enrolled including 20 patients with NSCLC, 31 patients with SCLC and 27 noncancer controls (Supplementary Table 1). The subjects were age and sex-matched and there were no significant differences for gender, smoking status and alcohol drinking record between groups. All the lung cancer patients were at the initial diagnosis prior to any treatment. The details for clinical characteristics were listed in Table 1 and Supplementary Table 1.

The metabolomic profiling of SCLC in the DBS of the discovery set

To investigate the unique metabolic features of SCLC, we first performed comparative metabolomic profiling of DBS for the discovery cohort. After quality control analysis, we detected and identified 372 metabolites with less than 20% miss values in all the samples. The representative chromatogram from one of the DBS specimens is shown in Supplementary Figure 1. Multivariate analysis showed the metabolomic profile of SCLC in DBS was not separated from NSCLC and noncancer controls by PLS-DA when all the male and female samples were combined for analysis (Supplementary Figure 2A). In addition, ROC curve analysis revealed low diagnostic accuracy for SCLC if the males and females were analyzed together (Supplementary Figure 2B). Previous studies reported that women have a different incidence of SCLC compared to men [16, 17]. These results suggested that there might be gender differences in metabolomic profiles between male and female SCLC patients. We then analyzed the male and female samples separately.

Both 2-D and 3-D PLS-DA plots showed that male SCLC had a unique metabolic signature that was distinct from patients with NSCLC and noncancer controls (Figure 2A and 2B). Since PLS-DA is a supervised model, we then performed leave-one-out cross-validation (LOOCV) to evaluate the level of overfitting for the PLS-DA model (Supplementary Figure 3A). The Q^2 value of 0.76 indicated the PLS-DA model is a good predictive model for the dataset. The top 15 most discriminant

metabolites revealed by variance in projection (VIP) analysis after PLS-DA was listed in Figure 2C. The VIP metabolites were further confirmed by univariate one-way ANOVA analysis. Pathway analysis revealed that sphingolipids metabolism, phospholipids, purine, urea cycle, acylcarnitines, amino acids, and endocannabinoids metabolism were significantly altered in patients with SCLC compared to NSCLC and controls (Figure 2D).

Similarly, there was a clear separation of metabolomic features in DBS among female SCLC, NSCLC and control subjects revealed by both 2-D and 3-D plots (Figure 3A and 3B). The PLS-DA classification model for the female dataset was confirmed by LOOCV (Supplementary Figure 3B). The top 7 most differential metabolites in the VIP analysis were from the biochemical pathways of phospholipids (Phosphatidylethanolamine, PE (18:2/20:4)), one-carbon metabolism (5-methyl-tetrahydrofolic acid), eicosanoids (9-HETE), sphingo-lipids (Ceramide (d18:1/20:4)), acylcarnitines (Oleoylcarnitine), branch chain amino acids (Arginine) and urea cycle (Glutamate) (Figure 3C and 3D).

The unique metabolic signatures in the DBS of male and female SCLC

We then compared the shared and unique metabolic features in male and female SCLC. As shown in the Venn diagram (Figure 4A), endocannabinoids, purine and bile acids metabolism were only significantly

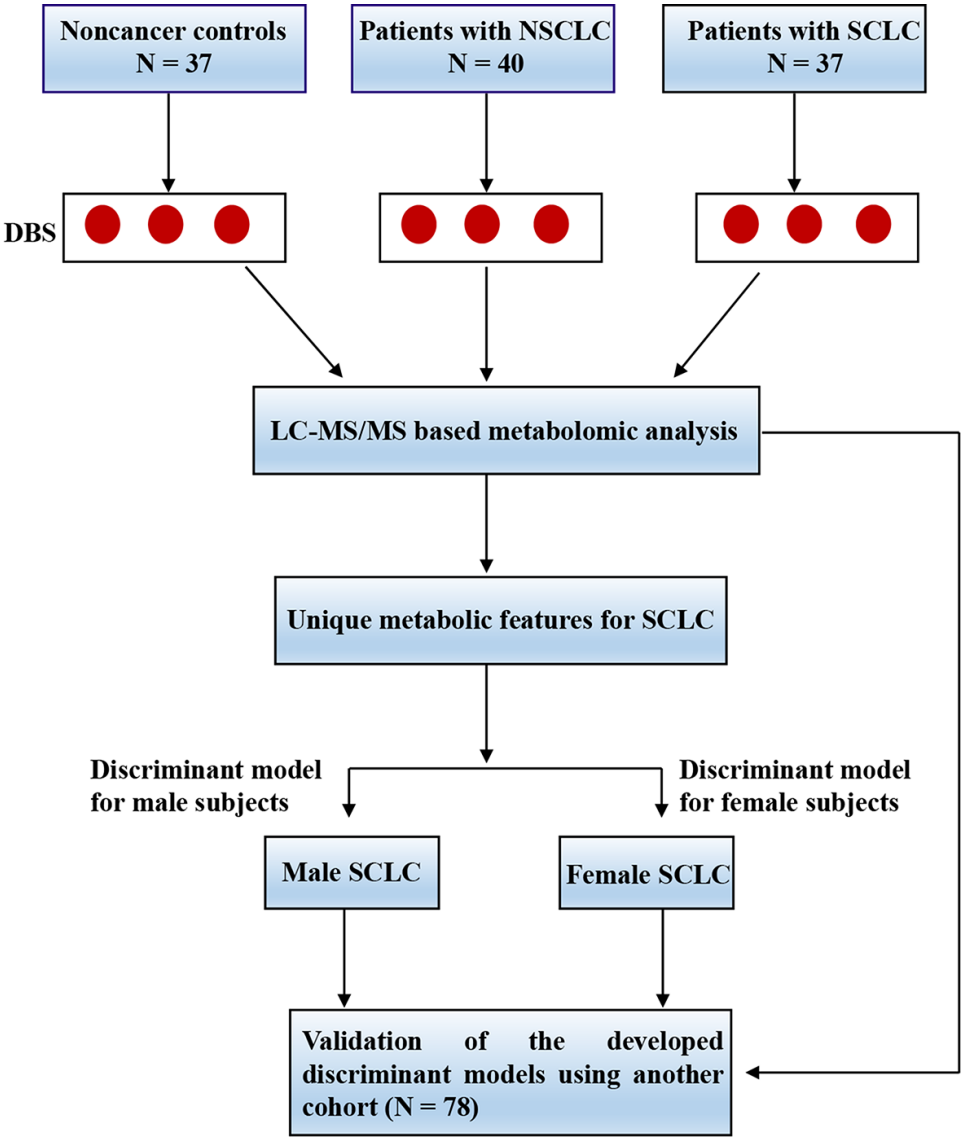


Figure 1. The experimental flow for the study. Abbreviations: DBS = dried blood spot; LC-MS/MS = liquid chromatography coupled with tandem mass spectrometry; NSCLC = non-small cell lung cancer; SCLC = small cell lung cancer.

Table 1. The characteristics of the subjects in the discovery stage.

Variables	Control	NSCLC	SCLC	P value
N	37	40	37	NA
Age, mean (SD)	62.8 (7.6)	62.6 (9.2)	63.7 (7.6)	0.81
Male (%)	23 (62.6%)	24 (60.0%)	23 (62.6%)	0.73
Smoking Status (%)				
Never	22 (59.5%)	24 (60.0%)	20 (54.1%)	0.95
Former	1 (2.7%)	2(5%)	4 (10.1%)	0.95
Current	14 (37.8%)	14(35%)	18 (48.62%)	0.95
Alcohol drinking status				
Never	29 (78.4%)	32(80.0%)	27 (73.0%)	0.76
Former	0 (0%)	2(5.1%)	4 (17.4%)	0.07
Current	8 (21.6%)	6(15.0%)	6 (16.2%)	0.75

Abbreviations: SCLC = small cell lung cancer; SD = standard deviation; NSCLC = non-small cell lung cancer.

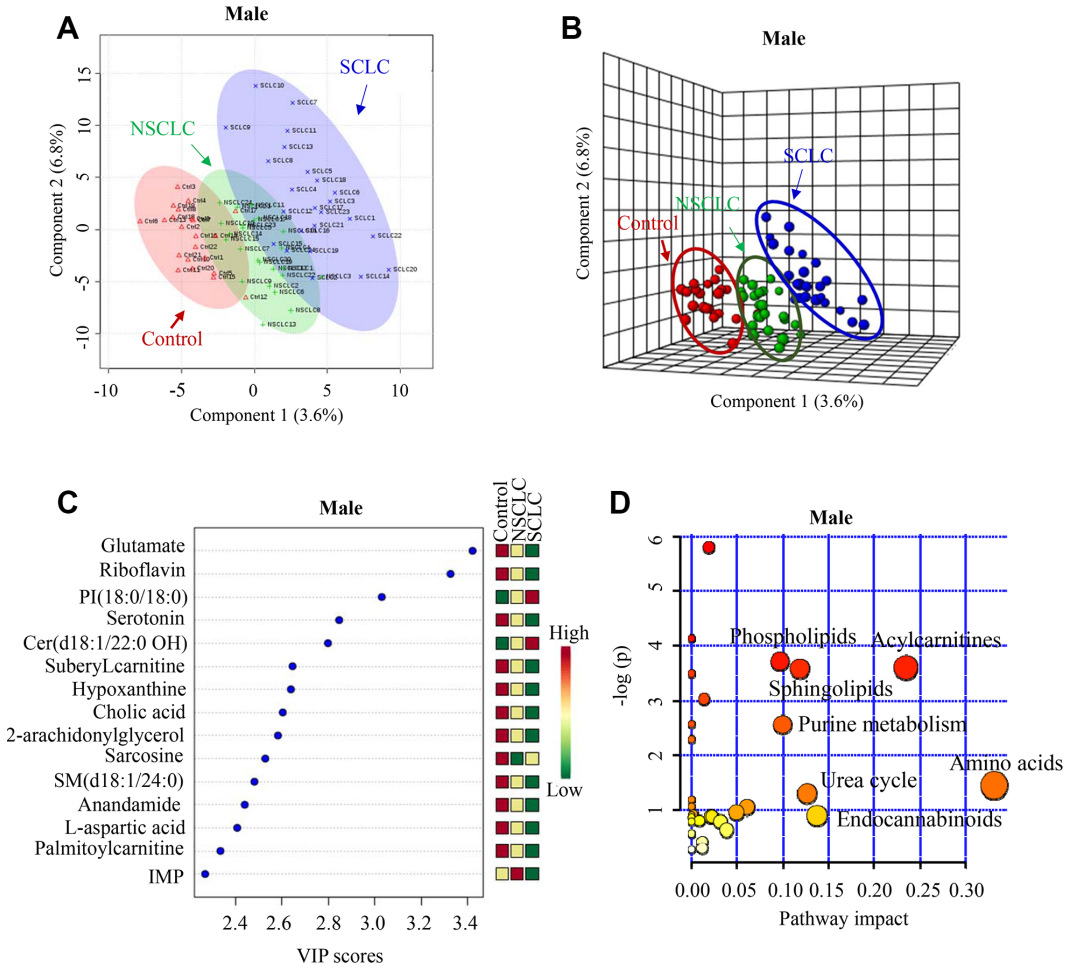


Figure 2. The unique metabolic features of male SCLC compared to NSCLC and controls. (**A**, **B**) Multivariate analysis of metabolomic data using PLS-DA resulted in a clear separation of metabolic features among SCLC, NSCLC and the control group in males. (**A**) 2-D plot. (**B**) 3-D plot. (**C**) The top 15 most differential metabolites in male patients with SCLC revealed by VIP analysis. VIP>1.5 was considered as statistically significant. The VIP results were also verified by univariate ANOVA analysis. (**D**) The top pathways disturbed in male SCLC patients. Abbreviations: IMP = inosine monophosphate; Cer = Ceramide; SM = Sphingomyelin; PI = Phosphatidylinositol; NSCLC = non-small cell lung cancer; PLS-DA = partial least square discriminant analysis; SCLC = small cell lung cancer; VIP = variance in projection.

altered in male SCLC, while, eicosanoids, steroids, pyrimidines, and one-carbon metabolism were disturbed in female SCLC patients. Five disturbed metabolic pathways including sphingolipids, phospholipids, acylcarnitines, amino acids, and urea cycle were common to both males and females with SCLC (Figure 4A).

Even though sphingolipids and phospholipids were both altered in male and female SCLC patients, further analysis showed gender-specific differences in the subclasses of these two classes of lipids. In detail, the levels of total 2-hydroxy ceramides in DBS of male SCLC were significantly higher than male NSCLC and controls (Figure 4B). But there were no significant differences in 2-hydroxy ceramides between female subjects. The changes in the DBS of female SCLC were mainly unsaturated ceramides such as Ceramide (d18:1/20:4) (Figure 3C). Additionally, in the subclasses of phospholipids, the unsaturated PE lipids decreased dramatically only in the DBS of female SCLC (Figure 4C), while, the saturated phosphatidylinositol lipids increased significantly in male SCLC (Figure 2C).

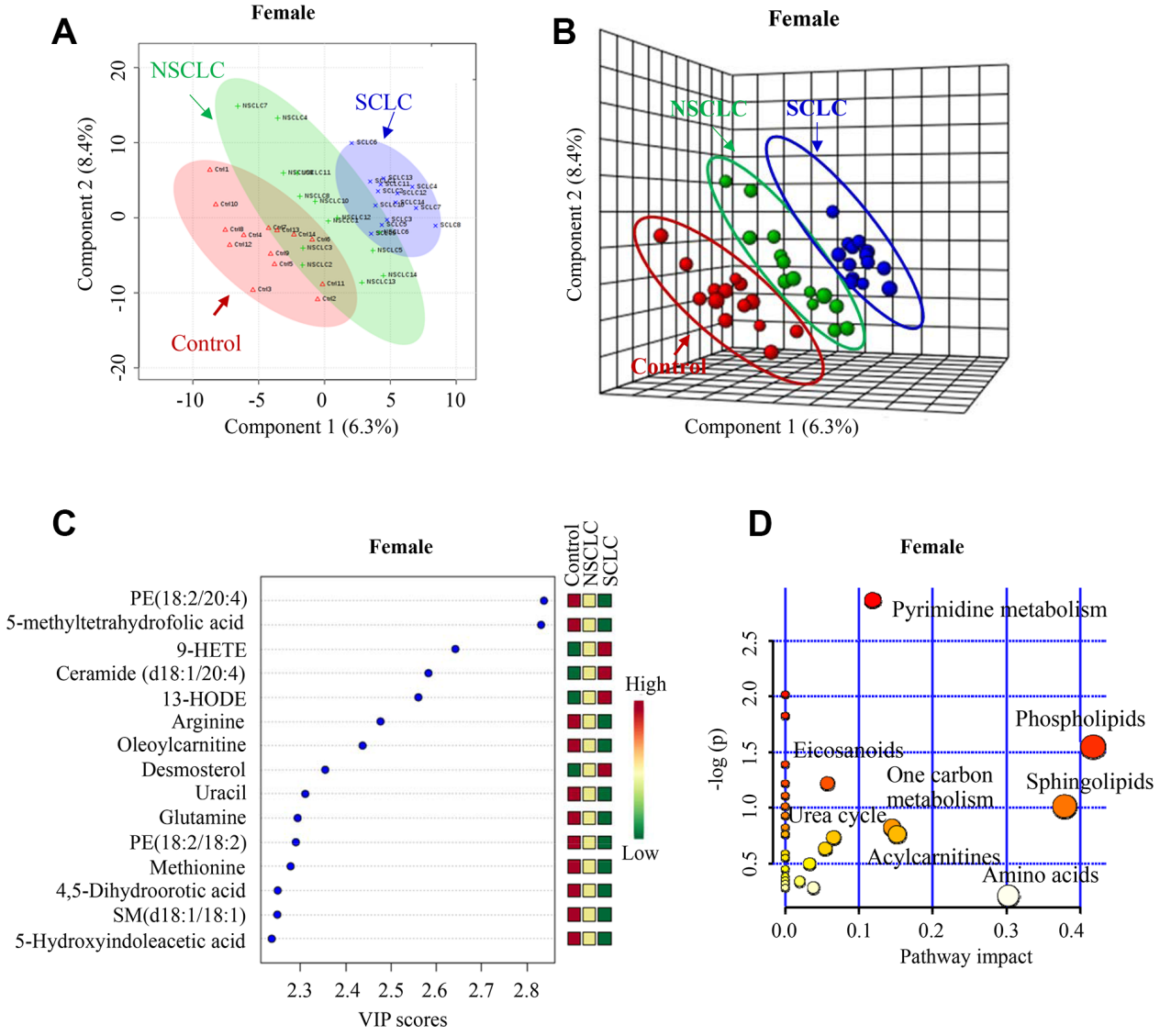


Figure 3. The unique metabolic features in the DBS of female SCLC patients compared to NSCLC and noncancer controls. (**A**, **B**) Multivariate analysis of metabolomic data using PLS-DA resulted in a clear separation of metabolic features among SCLC, NSCLC and the control group in females. (**A**) 2-D plot. (**B**) 3-D plot. (**C**) The top 15 most differential metabolites in female patients with SCLC revealed by VIP analysis. VIP>1.5 was considered as statistically significant. The VIP results were also verified by univariate ANOVA analysis. (**D**) The top pathways disturbed in female SCLC patients. Abbreviations: DBS = dried blood spot; NSCLC = non-small cell lung cancer; PE = phosphatidylethanolamine; 9-HETE = 9-hydroxyeicosatetraenoic acid; 13-HODE = 13-Hydroxyoctadecadienoic acid; SM = Sphingomyelin; PLS-DA = partial least square discriminant analysis; SCLC = small cell lung cancer; VIP = variance in projection.

The discriminant models of SCLC for male and female patients based on the discovery set

We next aimed to build the classification models that could be used for the identification of SCLC. We first evaluated the feasibility of the univariate model using male samples. Surprisingly, we found that a single analyte (2-arachidonylglycerol, 2-AG) had an accuracy of 85.6% (95% CI: 72.4%-95.1%) for the diagnosis of SCLC in the discovery cohort (Supplementary Figure 4A). However, it performed poorly in the validation cohort (Supplementary Figure 4B). This suggested that single metabolic biomarker might be biologically implausible for complex diseases like cancer.

We then explored the utility of the multiple variables classification. A set of 3-11 metabolites were selected and tested according to the following criteria: (1) Top 30 VIP metabolites in PLS-DA model (2) Covering as many as different biochemical pathways and k-mean clusters (3) High frequency in the least absolute shrinkage and selection operator (LASSO) model. By using this approach, the biological variation across the subjects is more readily accommodated. The performance of each classifier set of metabolites was evaluated by ROC analysis. The optimal variable combinations were obtained using the forward selection and backward elimination methods [18].

As a result, a 5-metabolite panel including 2-AG, cholic acid, PI (18:0/18:0), IMP and Cer (d18:1/22:0 OH) was selected as the best variable model for male SCLC. The biochemical pathways, VIP scores in the PLS-DA model, k-mean clusters and frequency of LASSO model of the selected metabolites were listed in Supplementary Table 2. The relative levels of PI (18:0/18:0) and Cer (d18:1/22:0 OH) were significantly upregulated in male SCLC (Figure 5A, 5B). In contrast, 2-arachidonylglycerol, IMP and cholic acid decreased significantly in the DBS of male SCLC subjects

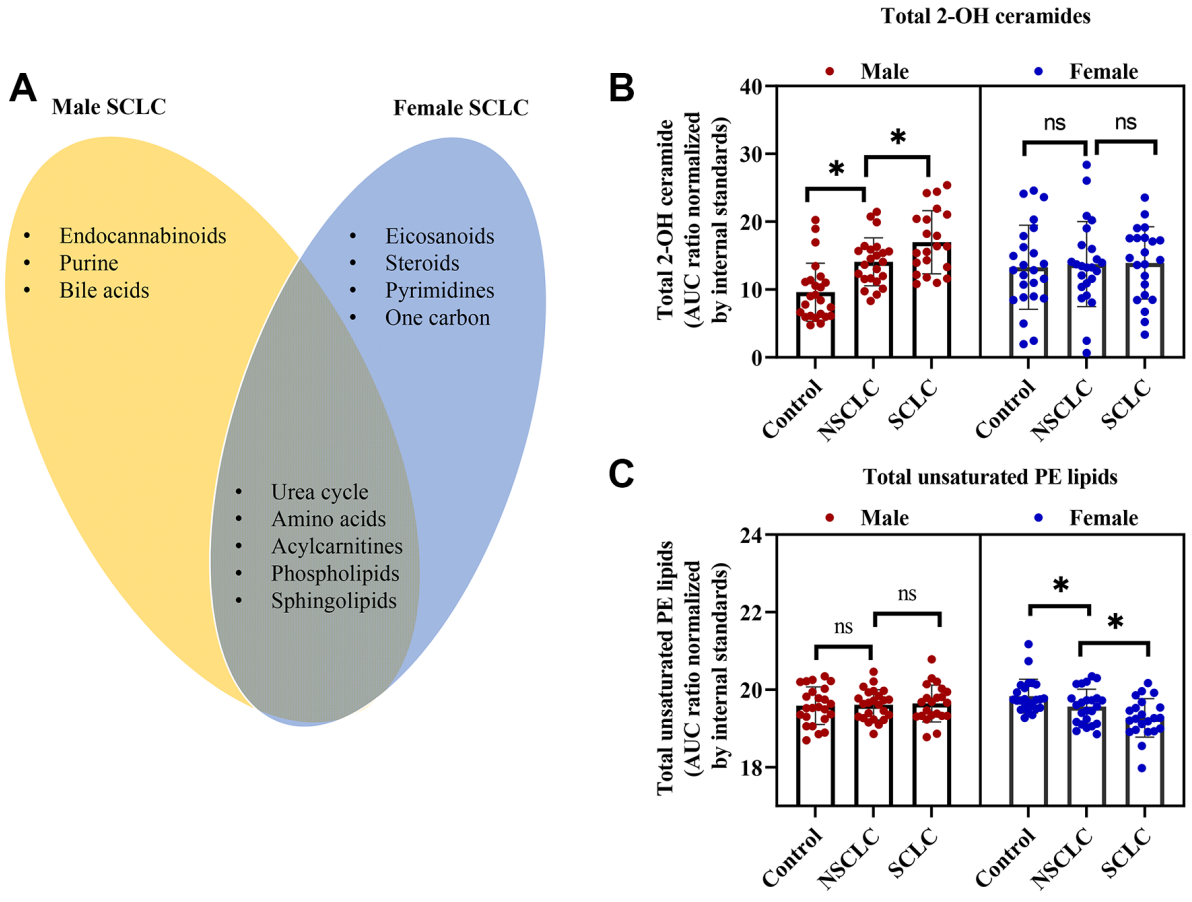


Figure 4. The shared and unique metabolic signature between male and female patients with SCLC. (**A**) The common and unique disturbed pathways in male and female SCLC compared to NSCLC and controls. (**B**) Total 2-hydroxy (2-OH) ceramides were significantly higher in male SCLC, while, no differences were observed in female patients with SCLC. (**C**) Total unsaturated PE was dramatically lower in female SCLC without significant changes in male SCLC. One-way ANOVA followed by Tukey's test. * $P < 0.05$ and ns: nonsignificant. Abbreviations: NSCLC = non-small cell lung cancer; PE = phosphatidylethanolamine; SCLC = small cell lung cancer.

(Figure 5C–5E). The ROC analysis using the 5-metabolite panel yielded the diagnostic accuracy of 0.95 (95% CI: 0.83-1) for the diagnosis of male SCLC patients in the discovery set (Figure 5F). The selected 5-metabolite classifier was further validated by a random permutation test (500 times). A P value of 0.0084 indicated the classifier was robust (Figure 5G). The relative importance of each metabolite in the classification model of SCLC calculated by multinomial logistic regression analysis is listed in Supplementary Table 4. Among 5 selected metabolites in the model, Cholic acid has the highest odds ratio of 2.48 (95% CI: 2.19 - 2.76) for SCLC over the controls and NSCLC. In addition, multiple linear regression analysis showed a significant correlation between the serum levels of NSE and the relative levels of 5 diagnostic metabolites in the DBS of male SCLC subjects (R^2 = 0.63, P < 0.01, Figure 5H).

The discriminant model for female SCLC was established on the basis of another set of metabolites including PE (18:1/20:4), 5-Methyltetrahydrofolic acid, Desmosterol, 4, 5-Dihydroorotic acid, and 9-HETE. There biochemical pathways, VIP scores in the PLS-DA model, k-mean clusters and frequencies of the LASSO model were shown in Supplementary Table 3. The levels of PE (18:1/20:4), 5-methyltetrahydrofolic acid, Desmosterol, and 4, 5-Dihydroorotic acid were dramatically reduced in female SCLC compared to the controls and NSCLC (Figure 6A–6D). In contrast, 9-HETE was significantly higher in female SCLC (Figure 6E). In the discovery set, the ROC analysis using the selected analytes in DBS had the diagnostic accuracy of 0.94 (95% CI: 0.74-1) for the diagnosis of female SCLC patients (Figure 6F). Classifier robustness for the model was verified with a random permutation test with a P value of 0.0066 (Figure 6G). We then

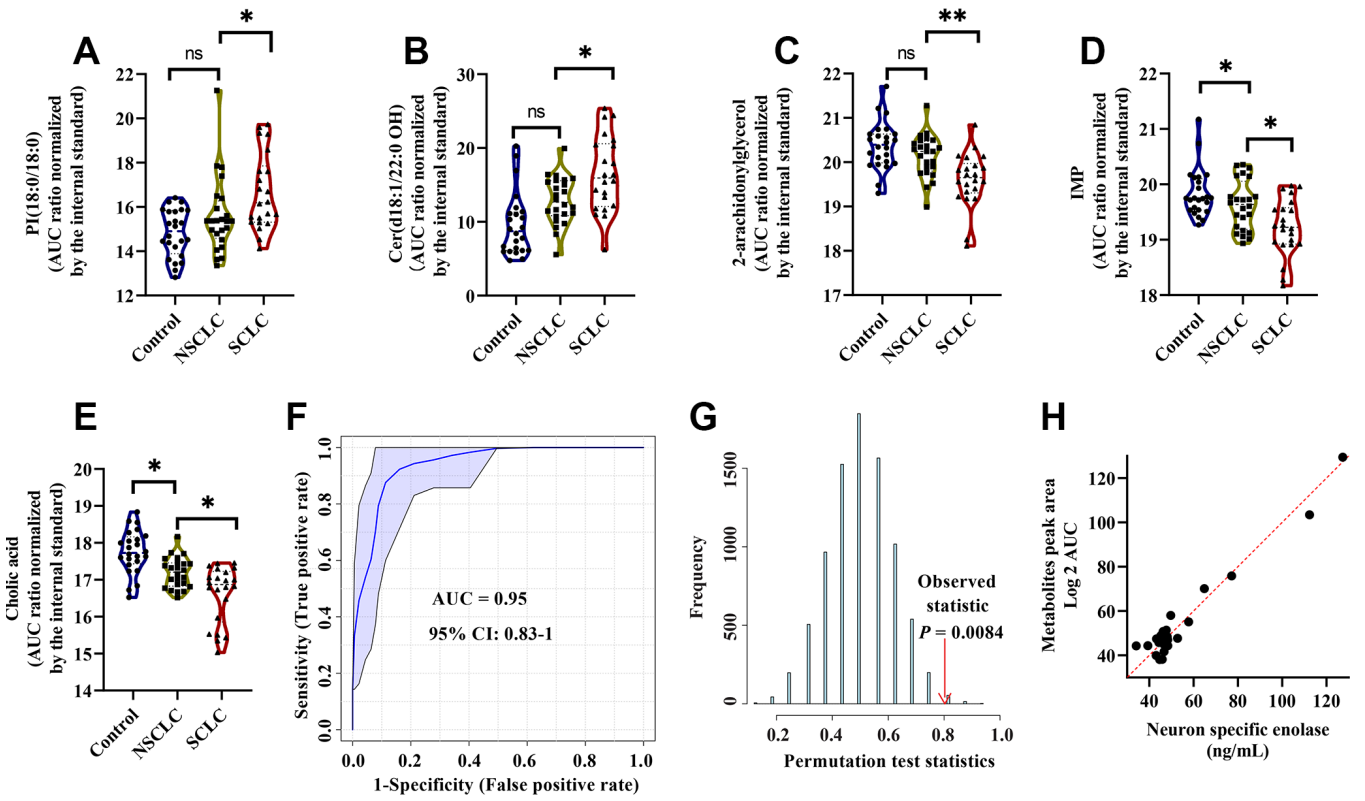


Figure 5. The diagnostic performance of the developed multianalyte discriminant model for male SCLC in the discovery set. (**A**–**E**) Violin plots showed the relative levels of five selected metabolites in the model. (**A**) PI(18:0/18:0); (**B**) Cer(d18:1/22:0 OH); (**C**) 2-Arachidonoylglycerol; (**D**) IMP; and (**E**) Cholic acid. The AUC ratio for each metabolite was calculated by the AUC of the corresponding internal standard. One-way ANOVA followed by Tukey's test was used. *P < 0.05 and ns: nonsignificant. (**F**) ROC curve analysis for distinguishing male SCLC patients from NSCLC and controls. (**G**) The random permutation test to examine the robustness of the model. Permutation P-value represented the probability that the classification of SCLC and NSCLC and controls using the selected metabolites could be obtained by chance. (**H**) Multiple linear regression analysis showed a significant correlation between the five selected metabolites and serum neuron-specific enolase (NSE) in male SCLC patients. R^2 = 0.63, P < 0.01. Abbreviations: IMP = inosine monophosphate; PI = Phosphatidylinositol; Cer = Ceramide; NSCLC = non-small cell lung cancer; NSE = neuron-specific enolase; ROC = receiver operator characteristic; SCLC = small cell lung cancer.

performed the multinomial logistic regression analysis to elucidate the relative importance of each metabolite for the classification of SCLC in females (Supplementary Table 5). We found that among five selected metabolites in the female model, the level of Desmosterol in DBS of SCLC has the highest an odds ratio (3.35, 95% CI: 3.06 - 3.63) over the controls and NSCLC. In addition, similar to the male classifier, a significant correlation with NSE levels in the serum was found by multiple linear regression analysis ($R^2 = 0.56$, $P < 0.05$, Figure 6H).

Validation of the discriminant models

To validate the accuracy of the developed models, we conducted the metabolomic profiling of another 78 DBS specimens (Supplementary Table 1). For the male subjects, the discriminant model established from the discovery set distinguished male SCLC group from those of in the validation set with the sensitivity of 86.6% (95% CI: 59.5%-98.3%) and the specificity of 84.1% (95% CI: 63.9%-95.5%) (Supplementary Table 6). Similarly, for the female subjects, the selected 5 metabolites from the discovery cohort had the sensitivity of 86.4% (95% CI: 65.1%-97.1%) and the specificity of 84.2% (95% CI: 68.8%-97.1%) for distinguishing female SCLC from the controls and NSCLC in the validation set (Supplementary Table 6).

DISCUSSION

Even though a few protein biomarkers in the serum have been used for lung cancer diagnosis, biomarkers capable of distinguishing lung cancer subtypes are yet to be discovered [19]. In this study, we demonstrated the unique metabolomic characteristics in DBS of SCLC patients compared to NSCLC and controls. We developed two highly sensitive and specific discriminant models for the diagnosis of male and female SCLC, respectively. The models were further

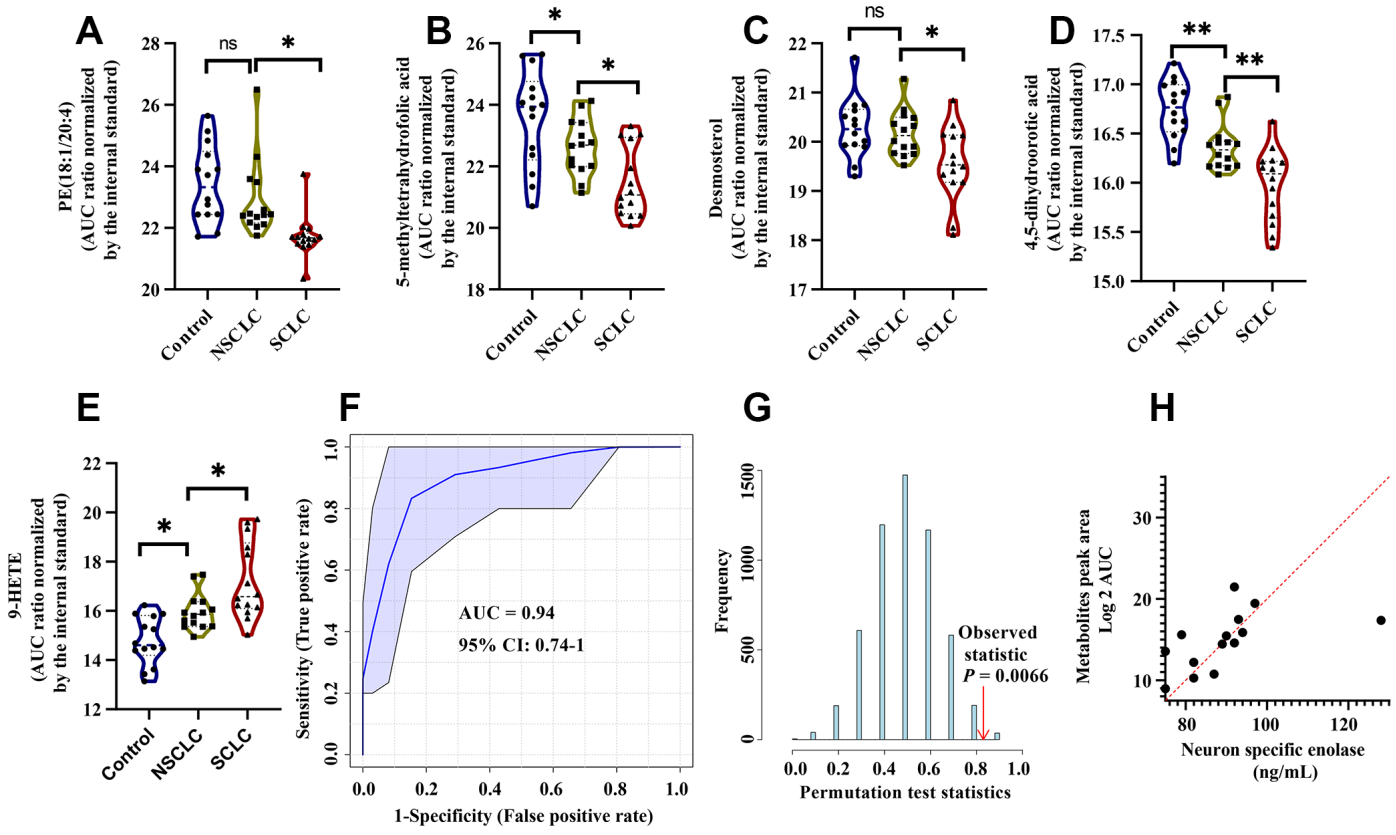


Figure 6. The diagnostic performance of the developed multianalyte discriminant model for female SCLC in the discovery set. (**A**–**E**) Violin plots showed the relative levels of five selected metabolites in the female discriminant model. (**A**) PE (18:1/20:4); (**B**) 5-Methyltetrahydrofolic acid; (**C**) Desmosterol; (**D**) 4, 5-Dihydroorotic acid and (**E**) 9-HETE. (**F**) ROC curve analysis for distinguishing female SCLC patients from female NSCLC and controls. (**G**) The random permutation test to examine the robustness of the female discriminant model. (**H**) Multiple linear regression analysis showed a significant correlation between the five selected metabolites and NSE in female SCLC patients. $R^2 = 0.56$, $P < 0.05$. Abbreviations: NSCLC = non-small cell lung cancer; PE = phosphatidylethanolamine; 9-HETE = 9-hydroxyeicosatetraenoic acid; ROC = receiver operator characteristic; SCLC = small cell lung cancer.

verified with an independent validation cohort using the same analytic method as that in the discovery set. The selected five-metabolite panels of biomarkers from DBS offer a simple and minimally invasive approach that may facilitate the accurate diagnosis and potentially lead to the effective and appropriate treatment of SCLC.

We identified the clear separation of DBS metabolomic profiles by PLS-DA between SCLC, NSCLC and the controls in both male and female subjects. This phenomenon was similar to our previous findings in the metabolomic profiling of SCLC tumor cells [12]. Out of the top 30 discriminating metabolites detected in DBS, only 6 were significantly altered in SCLC tumor cells. This might be due to the differences in the compositions between the whole blood and cells. In addition to red, white cells and platelets, the whole blood contains extracellular metabolites from the plasma. In contrast, only intracellular metabolites were measured in SCLC tumor cell lines. However, we found that even though the discriminating metabolites were largely different, the differential metabolic pathways in SCLC were highly similar in DBS and tumor cell lines. This highlighted the presence of "metabostasis" or metabolic balance that exists between metabolites in cells, plasma, and tissues and DBS could potentially reflect the abnormalities in lung cancer tissues [20].

The present study reported for the first time the gender-specific biomarkers for distinguishing SCLC from controls and NSCLC. This suggested that gender differences might need to be considered in the diagnosis and treatment of SCLC patients. Similarly, metabolic biomarker profiles were reported to be significantly different between men and women in cardiovascular diseases (CVD) [21] and chronic obstructive pulmonary disease (COPD) [22]. A previous study had also reported some sex-related differences in plasma proteins in NSCLC patients [23]. For example, the plasma level of soluble Fas (sFas) was significantly elevated in men with NSCLC but not in women. The reasons for the presence of the gender differences in SCLC remain largely unknown and might be associated with the differences in sex hormones and the proportion of fat tissue. In this study, we found a dramatic decrease in PE lipids in female SCLC. It was reported that estrogen induces the upregulation of phosphatidylethanolamine N-methyltransferase (PEMT), a key enzyme in the conversion of PE lipids to phosphatidylcholine [24]. Another study showed that estradiol was closely related to the enzymes and proteins linked to enzyme metabolism pathways and lipids biosynthesis [25]. Additionally, testosterone production was suppressed by 2-hydroxy ceramides [26]. The role of sex hormones, however, does not fully explain the metabolic differences between men and women. There are fundamental differences in the regulation of metabolic homeostasis between the sexes [27]. Further studies are needed on the metabolic differences between men and women with SCLC.

Through the selection and validation in two cohorts, we identified five metabolites for SCLC diagnosis in males including PI (18:0/18:0), Cer (d18:1/22:0 OH), 2-AG, IMP and Cholic acid. We found a significant increase of PI lipids and 2-hydroxy ceramides in the DBS of male SCLC patients compared to NSCLC and the noncancer controls. PI lipids are not only key cell membrane constituents but also involved in essential metabolic processes such as cell proliferation, growth, and survival via the activation of PI3K/AKT/mTOR signaling pathway [28]. Interestingly, an increase of PI lipids was also observed in the tumor tissue of NSCLC and lymphomas compared to normal tissue [29, 30]. The roles of ceramides and 2-hydroxy ceramides in cancer are not fully understood. Studies in HeLa cells revealed that overaccumulation of C24 ceramide protected cells from ionizing radiation (IR)-induced apoptosis [31]. A strong negative correlation between bile acids and membrane receptor TGR5 was recently reported and the overexpression of TGR5 was found to be associated with cell growth and migration in NSCLC [32]. Endo-cannabinoids including 2-AG and anandamide were known to be involved in cancer tumor growth and progression and associated pain [33]. However, the direction of the changes for endocannabinoids in different tumors was not consistent. Sailler and the co-authors found the elevated level of circulating 2-AG in the plasma of patients with brain tumor and breast cancer [34]. In our study, we found a decrease of 2-AG in the DBS of SCLC patients. This might be due to the differential roles of endocannabinoids in different cancers and different cell types [35].

In female SCLC patients, the circulating metabolites in DBS including PE (18:1/20:4), 5-methyltetrahydrofolic acid, Desmosterol, 4, 5-dihydroorotic acid, and 9-HETE were significantly altered. 5-methyltetrahydrofolic acid plays an important role in DNA synthesis, repair, methylation, and chromosomal integrity. Previous studies showed that the polymorphisms of 5-methylenetetrahydrofolate reductase (MTHFR) gene were associated with the risk of lung cancer in the female population [36, 37]. Beyond the literature, in this study, we found a significant reduction of 5-methyltetrahydrofolic acid in female SCLC patients. 9-HETE is an eicosanoid metabolite of arachidonic acid and the elevated level of 9-HETE was observed in DBS of women SCLC patients in our study. The link between 9-HETE and SCLC was unknown. However, the activation of 20-HETE biosynthesis was reported to promote the proliferation of cancer cells [38].

We have not tested the prognostic value of two developed discriminant models for small cell lung cancer. Some of the selected biomarker metabolites have been reported to be associated with the prognosis of several other diseases in the previous studies. For example, plasma ceramides are the potential prognostic biomarkers for acute myocardial infarction (MI) and sepsis [39, 40]. Plasma 5-Methyltetrahydrofolic acid is associated with the overall survival of postmenopausal breast cancer patients [41]. The circulating PE lipids were significantly correlated with poor prognostic genotypes of human papillomavirus in cervical cancer [42]. To our knowledge, the prognostic role of the multi-biomarker models in the whole has not been tested in the literature.

CONCLUSIONS

In summary, we demonstrated the presence of unique metabolic biosignature in the DBS of SCLC patients compared to NSCLC and controls. The gender-specific discriminant models with high accuracy for SCLC were developed and validated. Our study highlighted the potential clinical utility of the metabolomic profiling of DBS as a convenient and feasible approach to help the better diagnosis of SCLC. Further validation in larger cohorts will be needed through multi-regional and multi-center studies.

MATERIALS AND METHODS

Subjects and sample collection

This study was registered at the China Clinical Trial Registration Center (ID: ChiCTR-DDD-17010893) (http://www.chictr.org.cn/showproj.aspx?proj=18317). The study protocol was approved by the Ethics Committee of Shengjing Hospital of China Medical University on February 22, 2017. The study was conformed to the rules of the Declaration of Helsinki of 1957: Ethical Principles for Medical Research Involving Human Subjects. Written informed consent was obtained from all the subjects after full disclosure of the study details.

Inclusion criteria: volunteers in the control group met the following criteria: (1) no significant diseases (cancer, diabetes, cardiovascular disease, etc.), or recovery for more than 3 months; (2) no long-term medication history; (3) no surgery in the last 3 months; (4) age and gender-matched with SCLC patients. Patients in the SCLC and NSCLC group met the following criteria: (1) at the initial diagnosis prior to the treatment; (2) a pathological diagnosis of SCLC or NSCLC. In the discovery stage, a total of 114 subjects were enrolled and in the validation stage, another 78 subjects were recruited. The subjects' characteristics were listed in Table 1 and Supplementary Table 1.

Blood collection and DBS preparation: the finger-prick blood was directly spotted onto Whatman 903 Specimen Collection Paper Cards (GE Healthcare, USA).Three spots were taken from each subject and the spots were dried completely and stored at room temperature with desiccant before analysis. All the specimens were fasting blood samples.

Extraction of metabolites from DBS

The extraction of metabolites from DBS was conducted according to a previous study with modifications [43]. Briefly, a 3 mm disc per subject was obtained using a GE Healthcare Uni-Core punch and transferred to a 1.5 mL tube. A blank disc from the adjacent of DBS from the same card was used as the analytical blank. One hundred ninety (195) μL of extraction buffer (isopropanol/acetonitrile/water, 3:3:2, v/v/v) and 5 μL of internal standard mix (custom-synthesized stable isotope cocktail) [44] were added to each tube. The mixture was vortexed thoroughly and incubated on ice for 10 min. Samples were centrifuged at 16,000 g for 10 min at 4 °C. The supernatant was stored at -80 °C prior to metabolomic analysis.

LC-MS/MS-based metabolomic analysis

The metabolomic analysis was performed on a Shimadzu LC-20AD (Shimadzu, Japan) coupled with a Qtrap 5500 mass spectrometer (SCIEX, USA) as described before with slight modifications [44, 45]. Briefly, the chromatographic separation was achieved on a Luna NH2 column (250 × 2 mm, 5 μm, Phenomenex, USA) with the following conditions: mobile phase A: 95% H_2O + 5% acetonitrile + 20 mM ammonium hydroxide + 20 mM ammonium acetate (pH 9.4); B: 100% acetonitrile; the flow rate: 0.35 mL/min, the column temperature: 25 °C. The injection volume was 10 μL. The gradient was: 0-3 min 95% B, 3-6 min 75% B, 6-7 min 0% B, 7-15 min 0% B, and 15-17 min, 95% B. MS/MS analysis was performed by scheduled multiple reaction monitoring (sMRM) with electrospray ionization (ESI) and positive and negative switch mode. The polarity switch time was set at 50 ms. The ion source temperature was 500 °C. A total of 420 metabolites covering the major metabolomic pathways were targeted.

The chromatographic peaks were integrated using MultiQuant 3.0 (SCIEX, USA) and confirmed by manual inspection. All the detected metabolites were normalized using the corresponding internal standards and the AUC ratios (AUC of the detected

metabolites/AUC of the internal standards) prior to the statistical analysis.

Statistical analysis

The general statistical analysis was performed using GraphPad Prism 8.0 (GraphPad, USA). Data were mean and standard deviation (SD) for continuous variables and actual values with percentages for categorical variables. One-way ANOVA followed by Tukey's test was used for multiple comparisons of continuous variables and two proportion z test was used for the categorical variables. Multiple linear regression analysis was performed to analyze the relationship between serum NSE and the selected metabolites biomarkers. Significance was set at $P < 0.05$.

Metabolomic analysis was performed using metaboanalyst 4.0 (https://www.metaboanalyst.ca). The k-nearest neighbor method was used to deal with missing data due to individual differences, and the metabolites with missing data > 50% were removed. The data was log 2 transformed and auto-scaled. Multivariate analysis was conducted using PLS-DA. Since PLS-DA is a supervised model, the accuracy of the PLS-DA model was verified by LOOCV and Q^2 values were reported which represented the estimate of the predictive ability for the models. The top 30 most influential metabolites were obtained by VIP. VIP scores > 1.5 were considered statistically significant. The metabolites were also verified by a one-way ANOVA test. The significant metabolites identified by VIP analysis was then used for biochemical pathway analysis. In addition to the PLS-DA model, both the unsupervised k-mean clustering model and the LASSO frequencies were also performed.

Sets of 1-11 metabolites were manually selected by bootstrap resampling using the VIP lists and pathways obtained PLS-DA model, k-mean clusters and LASSO frequencies as an example of one possible multianalyte diagnostic classifier. The diagnostic performance of the selected classifiers was evaluated using ROC implemented in Metaboanalyst 4.0. In order to produce a smooth ROC curve, 100 cross-validations (CVs) were performed and the results were averaged to generate the plots. Classifier robustness was estimated by the permutation test (500 times). The sensitivity, specificity, accuracy, positive predictive value (PPV) and negative predictive value (NPV) of the selected biomarkers were analyzed by 2 × 2 contingency table analysis. The relative importance of each metabolite in the predictive models was evaluated using multinomial logistic regression analysis. The multinomial logistic regression coefficients, P values, and odds ratios were reported.

Abbreviations

2-AG: 2-arachidonylglycerol; 2-OH: 2-hydroxy; 9-HETE: 9-hydroxyeicosatetraenoic acid; 13-HODE: 13-Hydroxyoctadecadienoic acid; Cer: Ceramide; CI: confidence interval; COPD: chronic obstructive pulmonary disease; CVD: cardiovascular diseases; CVs: cross-validations; DBS: dried blood spot; ESI: electrospray ionization; HBECs: human bronchial epithelial cells; IHC: immunohistochemistry; IMP: inosine monophosphate; IR: ionizing radiation; LASSO: least absolute shrinkage and selection operator; LC-MS/MS: liquid chromatography coupled with tandem mass spectrometry; LOOCV: leave-one-out cross-validation; MTHFR: methylenetetrahydrofolate reductase; NPV: negative predictive value; NSCLC: non-small cell lung cancer; NSE: neuron-specific enolase; PE: phosphatidylethanolamine; PEMT: phosphatidylethanolamine N-methyltransferase; PI: Phosphatidylinositol; PLS-DA: partial least square discriminant analysis; PPV: positive predictive value; ROC: receiver operator characteristic; SCLC: small cell lung cancer; SD: standard deviation; sFas: soluble Fas; SM: Sphingomyelin; sMRM: scheduled multiple reaction monitoring; VIP: variance in projection.

AUTHOR CONTRIBUTIONS

Conceptualization: LY and XZ; methodology: LY; software: KL and XL; validation: LY and XL; formal analysis: LY and XL; investigation: CG and TS; resources: XZ; data curation: KL, XL, CG and TS; writing-original draft preparation: LY; writing-review & editing: LY, KL, XL, CG, TS and XZ; visualization: KL and XL; supervision: XZ; project administration: XZ; funding acquisition: XZ.

.

CONFLICTS OF INTEREST

The authors of this manuscript declare no conflicts of interest.

FUNDING

This project was funded by the "Plan of Cultivating Future Talents of Shengjing Hospital, China". The funder has no roles in the design and implement of the study.

REFERENCES

1. Vineis P, Wild CP. Global cancer patterns: causes and prevention. Lancet. 2014; 383:549–57. https://doi.org/10.1016/S0140-6736(13)62224-2 PMID:24351322

2. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ Jr, Wu YL, Paz-Ares L. Lung cancer: current therapies and new targeted treatments. Lancet. 2017; 389:299–311. https://doi.org/10.1016/S0140-6736(16)30958-8 PMID:27574741

3. Thunnissen E, Borczuk AC, Flieder DB, Witte B, Beasley MB, Chung JH, Dacic S, Lantuejoul S, Russell PA, den Bakker M, Botling J, Brambilla E, de Cuba E, et al. The Use of Immunohistochemistry Improves the Diagnosis of Small Cell Lung Cancer and Its Differential Diagnosis. An International Reproducibility Study in a Demanding Set of Cases. J Thorac Oncol. 2017; 12:334–46. https://doi.org/10.1016/j.jtho.2016.12.004 PMID:27998793

4. den Bakker MA, Willemsen S, Grünberg K, Noorduijn LA, van Oosterhout MF, van Suylen RJ, Timens W, Vrugt B, Wiersma-van Tilburg A, Thunnissen FB. Small cell carcinoma of the lung and large cell neuroendocrine carcinoma interobserver variability. Histopathology. 2010; 56:356–63. https://doi.org/10.1111/j.1365-2559.2010.03486.x PMID:20459535

5. Killock D. Lung cancer: liquid biopsy of SCLC chemosensitivity. Nat Rev Clin Oncol. 2017; 14:2. https://doi.org/10.1038/nrclinonc.2016.201 PMID:27922046

6. Isgrò MA, Bottoni P, Scatena R. Neuron-Specific Enolase as a Biomarker: Biochemical and Clinical Aspects. Adv Exp Med Biol. 2015; 867:125–43. https://doi.org/10.1007/978-94-017-7215-0_9 PMID:26530364

7. Korse CM, Holdenrieder S, Zhi XY, Zhang X, Qiu L, Geistanger A, Lisy MR, Wehnl B, van den Broek D, Escudero JM, Standop J, Hu M, Molina R. Multicenter evaluation of a new progastrin-releasing peptide (ProGRP) immunoassay across Europe and China. Clin Chim Acta. 2015; 438:388–95. https://doi.org/10.1016/j.cca.2014.09.015 PMID:25262909

8. Cui S, Li K, Ang L, Liu J, Cui L, Song X, Lv S, Mahmud E. Plasma Phospholipids and Sphingolipids Identify Stent Restenosis After Percutaneous Coronary Intervention. JACC Cardiovasc Interv. 2017; 10:1307–16. https://doi.org/10.1016/j.jcin.2017.04.007 PMID:28624380

9. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol. 2016; 17:451–59. https://doi.org/10.1038/nrm.2016.25 PMID:26979502

10. Yu L, Li K, Zhang X. Next-generation metabolomics in lung cancer diagnosis, treatment and precision medicine: mini review. Oncotarget. 2017; 8:115774–86. https://doi.org/10.18632/oncotarget.22404 PMID:29383200

11. Han J, Xia Y, Lin L, Zhang Z, Tian H, Li K. Next-generation Metabolomics in the Development of New Antidepressants: Using Albiflorin as an Example. Curr Pharm Des. 2018; 24:2530–40. https://doi.org/10.2174/1381612824666180727114134 PMID:30051781

12. Yu L, Li K, Xu Z, Cui G, Zhang X. Integrated omics and gene expression analysis identifies the loss of metabolite-metabolite correlations in small cell lung cancer. Oncotargets Ther. 2018; 11:3919–29. https://doi.org/10.2147/OTT.S166149 PMID:30013371

13. Freeman JD, Rosman LM, Ratcliff JD, Strickland PT, Graham DR, Silbergeld EK. State of the Science in Dried Blood Spots. Clin Chem. 2018; 64:656–79. https://doi.org/10.1373/clinchem.2017.275966 PMID:29187355

14. Wang Q, Sun T, Cao Y, Gao P, Dong J, Fang Y, Fang Z, Sun X, Zhu Z. A dried blood spot mass spectrometry metabolomic approach for rapid breast cancer detection. Onco Targets Ther. 2016; 9:1389–98. https://doi.org/10.2147/ott.s95862 PMID:27042107

15. Jing Y, Wu X, Gao P, Fang Z, Wu J, Wang Q, Li C, Zhu Z, Cao Y. Rapid differentiating colorectal cancer and colorectal polyp using dried blood spot mass spectrometry metabolomic approach. IUBMB Life. 2017; 69:347–54. https://doi.org/10.1002/iub.1617 PMID:28322027

16. Ouellette D, Desbiens G, Emond C, Beauchamp G. Lung cancer in women compared with men: stage, treatment, and survival. Ann Thorac Surg. 1998; 66:1140–43. https://doi.org/10.1016/S0003-4975(98)00557-8 PMID:9800795

17. Tseng JE, Rodriguez M, Ro J, Liu D, Hong WK, Mao L. Gender differences in p53 mutational status in small cell lung cancer. Cancer Res. 1999; 59:5666–70. PMID:10582680

18. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics. 2013; 9:280–99. https://doi.org/10.1007/s11306-012-0482-9 PMID:23543913

19. Zamay TN, Zamay GS, Kolovskaya OS, Zukov RA, Petrova MM, Gargaun A, Berezovski MV, Kichkailo AS. Current and Prospective Protein Biomarkers of Lung Cancer. Cancers (Basel). 2017; 9:E155. https://doi.org/10.3390/cancers9110155 PMID:29137182

20. Ivanisevic J, Elias D, Deguchi H, Averell PM, Kurczy M, Johnson CH, Tautenhahn R, Zhu Z, Watrous J, Jain M, Griffin J, Patti GJ, Siuzdak G. Arteriovenous Blood Metabolomics: A Readout of Intra-Tissue Metabostasis. Sci Rep. 2015; 5:12757. https://doi.org/10.1038/srep12757 PMID:26244428

21. Lew J, Sanghavi M, Ayers CR, McGuire DK, Omland T, Atzler D, Gore MO, Neeland I, Berry JD, Khera A, Rohatgi A, de Lemos JA. Sex-Based Differences in Cardiometabolic Biomarkers. Circulation. 2017; 135:544–55. https://doi.org/10.1161/CIRCULATIONAHA.116.023005 PMID:28153991

22. de Torres JP, Casanova C, Pinto-Plata V, Varo N, Restituto P, Cordoba-Lanus E, Baz-Dávila R, Aguirre-Jaime A, Celli BR. Gender differences in plasma biomarker levels in a cohort of COPD patients: a pilot study. PLoS One. 2011; 6:e16021. https://doi.org/10.1371/journal.pone.0016021 PMID:21267454

23. Izbicka E, Streeper RT, Michalek JE, Louden CL, Diaz A 3rd, Campos DR. Plasma biomarkers distinguish non-small cell lung cancer from asthma and differ in men and women. Cancer Genomics Proteomics. 2012; 9:27–35. https://doi.org/10.1158/1535-7163.targ-09-c15 PMID:22210046

24. Resseguie M, Song J, Niculescu MD, da Costa KA, Randall TA, Zeisel SH. Phosphatidylethanolamine N-methyltransferase (PEMT) gene expression is induced by estrogen in human and mouse primary hepatocytes. FASEB J. 2007; 21:2622–32. https://doi.org/10.1096/fj.07-8227com PMID:17456783

25. Chen JQ, Brown TR, Russo J. Regulation of energy metabolism pathways by estrogens and estrogenic chemicals and potential implications in obesity associated with increased exposure to endocrine disruptors. Biochim Biophys Acta. 2009; 1793:1128–43. https://doi.org/10.1016/j.bbamcr.2009.03.009 PMID:19348861

26. Lucki NC, Sewer MB. Multiple roles for sphingolipids in steroid hormone biosynthesis. Subcell Biochem. 2008; 49:387–412. https://doi.org/10.1007/978-1-4020-8831-5_15 PMID:18751920

27. Clegg DJ, Mauvais-Jarvis F. An integrated view of sex differences in metabolic physiology and disease. Mol Metab. 2018; 15:1–2. https://doi.org/10.1016/j.molmet.2018.06.011 PMID:30032908

28. Liu P, Cheng H, Roberts TM, Zhao JJ. Targeting the phosphoinositide 3-kinase pathway in cancer. Nat Rev Drug Discov. 2009; 8:627–44. https://doi.org/10.1038/nrd2926 PMID:19644473

29. Eberlin LS, Gabay M, Fan AC, Gouw AM, Tibshirani RJ, Felsher DW, Zare RN. Alteration of the lipid profile in lymphomas induced by MYC overexpression. Proc Natl Acad Sci USA. 2014; 111:10450–55. https://doi.org/10.1073/pnas.1409778111 PMID:24994904

30. Marien E, Meister M, Muley T, Fieuws S, Bordel S, Derua R, Spraggins J, Van de Plas R, Dehairs J, Wouters J, Bagadi M, Dienemann H, Thomas M, et al. Non-small cell lung cancer is characterized by dramatic changes in phospholipid profiles. Int J Cancer. 2015; 137:1539–48. https://doi.org/10.1002/ijc.29517 PMID:25784292

31. Ogretmen B. Sphingolipid metabolism in cancer signalling and therapy. Nat Rev Cancer. 2018; 18:33–50. https://doi.org/10.1038/nrc.2017.96 PMID:29147025

32. Liu X, Chen B, You W, Xue S, Qin H, Jiang H. The membrane bile acid receptor TGR5 drives cell growth and migration via activation of the JAK2/STAT3 signaling pathway in non-small cell lung cancer. Cancer Lett. 2018; 412:194–207. https://doi.org/10.1016/j.canlet.2017.10.017 PMID:29074425

33. Pyszniak M, Tabarkiewicz J, Łuszczki JJ. Endocannabinoid system as a regulator of tumor cell malignancy - biological pathways and clinical significance. Oncotargets Ther. 2016; 9:4323–36. https://doi.org/10.2147/OTT.S106944 PMID:27486335

34. Sailler S, Schmitz K, Jäger E, Ferreiros N, Wicker S, Zschiebsch K, Pickert G, Geisslinger G, Walter C, Tegeder I, Lötsch J. Regulation of circulating endocannabinoids associated with cancer and metastases in mice and humans. Oncoscience. 2014; 1:272–82. https://doi.org/10.18632/oncoscience.33 PMID:25594019

35. Hermanson DJ, Marnett LJ. Cannabinoids, endocannabinoids, and cancer. Cancer Metastasis Rev. 2011; 30:599–612. https://doi.org/10.1007/s10555-011-9318-8 PMID:22038019

36. Tong W, Tong G, Jin D, Lv Q. MTHFR C677T and A1298C polymorphisms and lung cancer risk in a female Chinese population. Cancer Manag Res. 2018; 10:4155–61. https://doi.org/10.2147/CMAR.S176263 PMID:30323671

37. Shi Q, Zhang Z, Li G, Pillow PC, Hernandez LM, Spitz MR, Wei Q. Sex differences in risk of lung cancer

associated with methylene-tetrahydrofolate reductase polymorphisms. Cancer Epidemiol Biomarkers Prev. 2005; 14:1477–84.
https://doi.org/10.1158/1055-9965.EPI-04-0905
PMID:15941959

38. Alexanian A, Sorokin A. Targeting 20-HETE producing enzymes in cancer - rationale, pharmacology, and clinical potential. Onco Targets Ther. 2013; 6:243–55.
https://doi.org/10.2147/ott.s31586 PMID:23569388

39. de Carvalho LP, Tan SH, Ow GS, Tang Z, Ching J, Kovalik JP, Poh SC, Chin CT, Richards AM, Martinez EC, Troughton RW, Fong AY, Yan BP, et al. Plasma Ceramides as Prognostic Biomarkers and Their Arterial and Myocardial Tissue Correlates in Acute Myocardial Infarction. JACC Basic Transl Sci. 2018; 3:163–75.
https://doi.org/10.1016/j.jacbts.2017.12.005
PMID:30062203

40. Drobnik W, Liebisch G, Audebert FX, Frohlich D, Gluck T, Vogel P, Rothe G, Schmitz G. Plasma ceramide and lysophosphatidylcholine inversely correlate with mortality in sepsis patients. J Lipid Res. 2003; 44:754–61.
https://doi.org/10.1194/jlr.M200401-JLR200
PMID:12562829

41. McEligot AJ, Ziogas A, Pfeiffer CM, Fazili Z, Anton-Culver H. The association between circulating total folate and folate vitamers with overall survival after postmenopausal breast cancer diagnosis. Nutr Cancer. 2015; 67:442–48.
https://doi.org/10.1080/01635581.2015.1002623
PMID:25647689

42. Hung CY, Chao A, Wang CC, Wu RC, Lu KY, Lu HY, Lai CH, Lin G. Glycerophospholipids and sphingolipids correlate with poor prognostic genotypes of human papillomavirus in cervical cancer: global lipidomics analysis. Anal Methods. 2018; 10:4970–77.
https://doi.org/10.1039/C8AY01691G

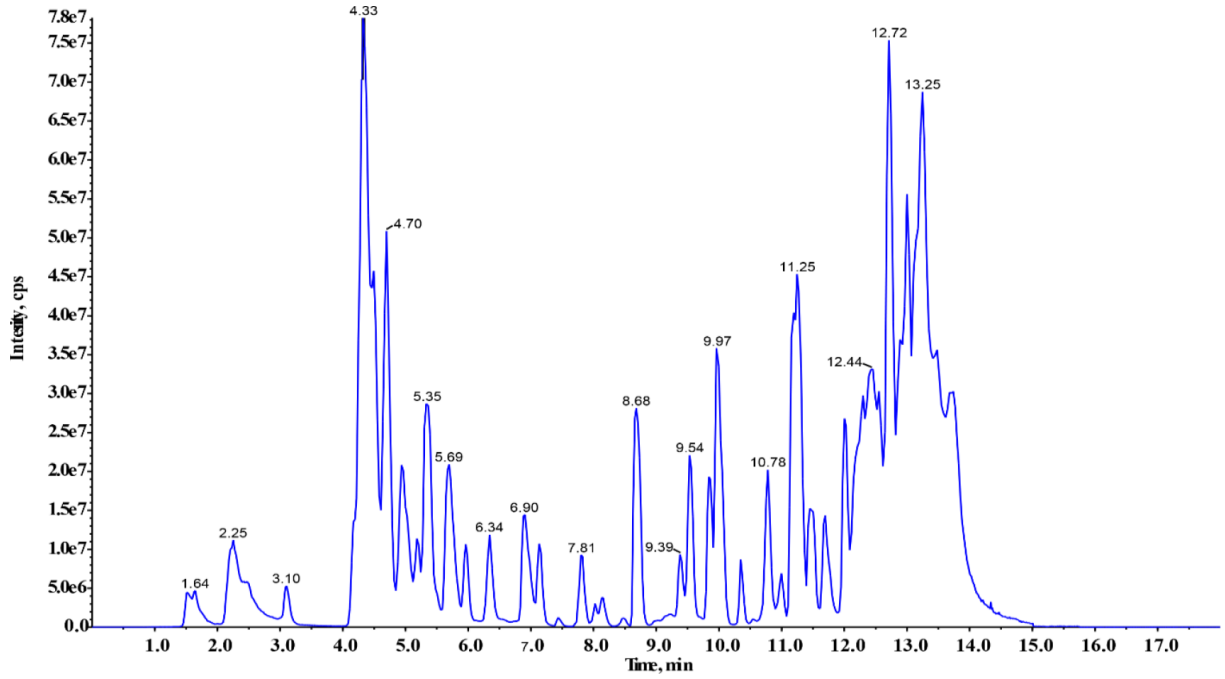
43. Drolet J, Tolstikov V, Williams BA, Greenwood BP, Hill C, Vishnudas VK, Sarangarajan R, Narain NR, Kiebish MA. Integrated Metabolomics Assessment of Human Dried Blood Spots and Urine Strips. Metabolites. 2017; 7:E35.
https://doi.org/10.3390/metabo7030035
PMID:28714878

44. Li K, Wang X, Pidatala VR, Chang CP, Cao X. Novel quantitative metabolomic approach for the study of stress responses of plant root metabolism. J Proteome Res. 2014; 13:5879–87.
https://doi.org/10.1021/pr5007813
PMID:25327737

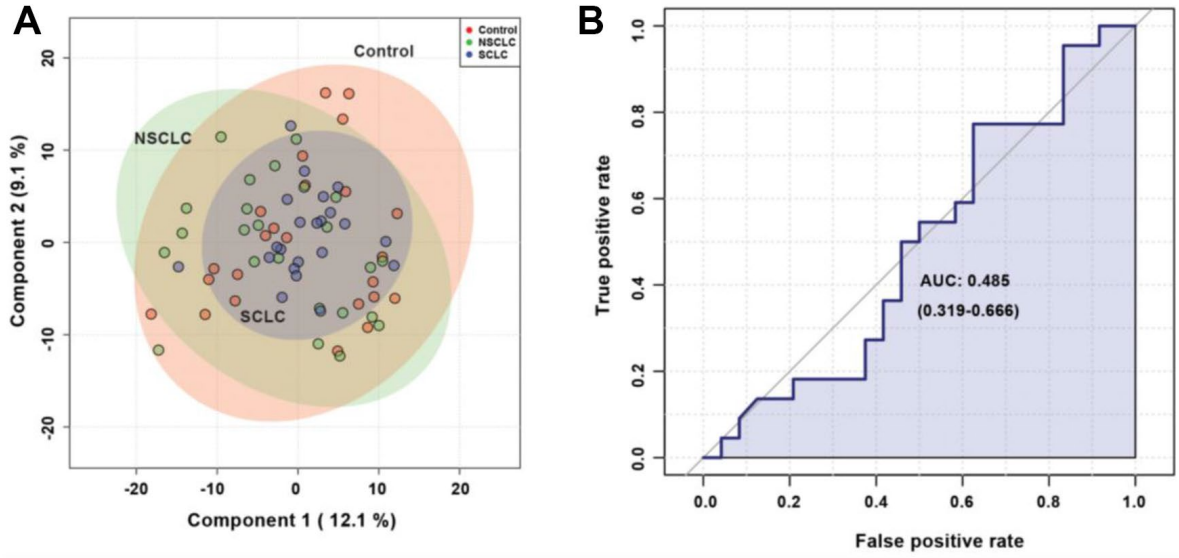
45. Yuan M, Breitkopf SB, Yang X, Asara JM. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. Nat Protoc. 2012; 7:872–81.
https://doi.org/10.1038/nprot.2012.024
PMID:22498707

SUPPLEMENTARY MATERIALS

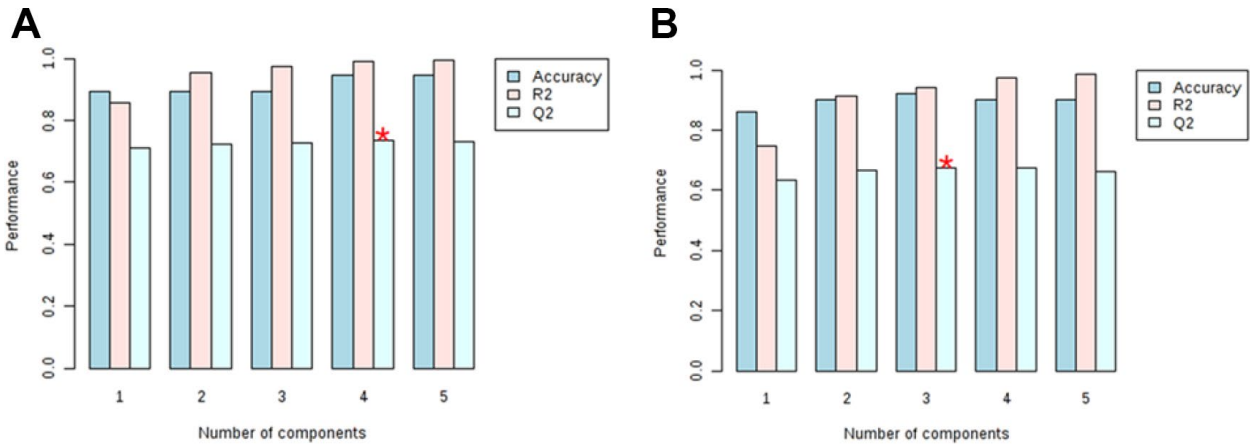
Supplementary Figures



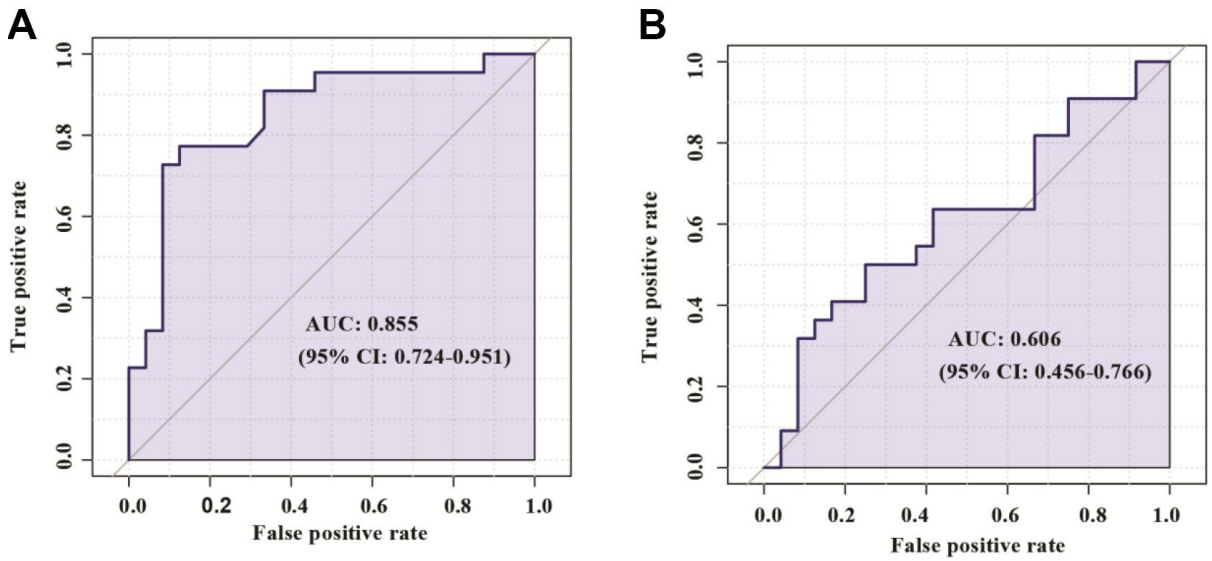
Supplementary Figure 1. The representative chromatogram of the metabolomic profiling of a dried blood spot (DBS) specimen.



Supplementary Figure 2. No unique metabolic features were discovered to distinguish SCLC from controls and NSCLC when male and female samples were combined for data analysis. (**A**) The metabolomic profile of SCLC was not separated from the controls and NSCLC by PLS-DA when all the samples were analyzed together. (**B**) ROC curve analysis showed low diagnostic accuracy for SCLC when males and females were combined. Abbreviations: NSCLC = non-small cell lung cancer; PLS-DA = partial least square discriminant analysis; ROC = receiver operator characteristic; SCLC = small cell lung cancer.



Supplementary Figure 3. LOOCV confirmed that the PLS-DA models were able to effectively and accurately discriminate SCLC patients from NSCLC and controls in male and female subjects, respectively. (**A**) male subjects; (**B**) female subjects. Abbreviations: LOOCV = leave-one-out cross-validation; NSCLC = non-small cell lung cancer; PLS-DA = partial least square discriminant analysis; SCLC = small cell lung cancer.



Supplementary Figure 4. The classification performance of the univariate model for the diagnosis of SCLC in male subjects. (**A**) The discovery set; (**B**) Validation set. The single analyte used in the figure was 2-arachidonylglycerol. Abbreviations: SCLC = small cell lung cancer.

Supplementary Tables

Supplementary Table 1. The characteristics of the subjects in the validation cohort.

Variables	Control	NSCLC	SCLC	P value
N	27	20	31	NA
Age	61.5 ± 5.4	62.1 ± 4.3	62.9 ± 6.1	0.62
Male (%)	13 (48.2%)	12 (60%)	15 (48.4%)	0.61
Smoking Status (%)				
Never	13 (48.1%)	9 (45.0%)	14 (45.2%)	0.99
Former	2 (7.4%)	3 (15.1%)	4 (10.1%)	0.72
Current	12 (44.4%)	8 (40.0%)	13 (41.9%)	0.99
Alcohol drinking status				
Never	15 (55.6%)	12 (60.0%)	16 (51.6%)	0.99
Former	3 (11.1%)	2 (10.0%)	4 (12.3%)	0.99
Current	9 (33.3%)	6 (30.0%)	11 (35.5%)	0.99

Abbreviations: SCLC = small cell lung cancer; NSCLC = non-small cell lung cancer.

Supplementary Table 2. The biochemical pathways, VIP scores, k-mean cluster and LASSO frequency of the selected 5 diagnostic metabolites for SCLC in males.

Metabolites	Biochemical pathways	VIP scores in the PLS-DA model	K-means clusters	Frequencies using LASSO modeling
PI(18:0/18:0)	Phospholipids metabolism	3.13	2	70%
Cer(d18:1/22:0 OH)	Sphingolipids metabolism	2.81	5	70%
Cholic acid	Bile acids metabolism	2.59	3	70%
2-arachidonylglycerol	Endocannabinoid metabolism	2.57	4	70%
IMP	Purine metabolism	2.11	2	70%

Abbreviations: IMP = inosine monophosphate; Cer = Ceramide; PI = Phosphatidylinositol; LASSO = least absolute shrinkage and selection operator; PLS-DA = partial least square discriminant analysis; SCLC = small cell lung cancer; VIP = variance in projection.

Supplementary Table 3. The biochemical pathways, VIP scores, k-mean cluster and LASSO frequency of the selected 5 diagnostic metabolites for SCLC in females.

Metabolites	Biochemical pathways	VIP scores in the PLS-DA model	K-means clusters	Frequencies using LASSO modeling
PE(18:1/20:4)	Phospholipids metabolism	2.87	5	64%
5-Methyltetrahydrofolic acid	One-carbon metabolism	2.85	3	64%
Desmosterol	Steroid metabolism	2.36	2	60%
4,5-Dihydroorotic acid	Pyrimidine metabolism	2.12	1	60%
9-HETE	Eicosanoid metabolism	2.64	4	60%

Abbreviations: 9-HETE = 9-hydroxyeicosatetraenoic acid; LASSO = least absolute shrinkage and selection operator; PE = phosphatidylethanolamine; PLS-DA = partial least square discriminant analysis; SCLC = small cell lung cancer; VIP = variance in projection.

Supplementary Table 4. The multinomial logistic regression analysis of the 5 diagnostic metabolites for SCLC in males.

Metabolites	Multinomial logistic regression coefficients	P value	Odds ratio for SCLC (95% CI)
2-AG	0.18	0.041	1.19 (1.01 - 1.37)
Cholic acid	0.91	0.006	2.48 (2.19 – 2.76)
PI (18:0/18:0)	0.51	0.01	1.61 (1.13 - 1.85)
IMP	0.23	0.034	1.26 (1.06 - 1.89)
Cer (d18:1/22:0 OH)	0.47	0.025	1.59 (1.41 - 1.78)

Abbreviations: 2-AG = 2-arachidonylglycerol; PI = Phosphatidylinositol; IMP = inosine monophosphate; Cer = Ceramide; SCLC = small cell lung cancer; CI = confidence interval.

Supplementary Table 5. The multinomial logistic regression analysis of the 5 diagnostic metabolites for SCLC in females.

Metabolites	Multinomial logistic regression coefficients	P value	Odds ratio for SCLC (95% CI)
PE (18:1/20:4)	0.85	0.017	2.34 (2.19 - 2.48)
5-Methyltetrahydrofolic acid	0.33	0.035	1.39 (1.11 - 1.67)
Desmosterol	1.21	0.002	3.35 (3.06 - 3.63)
4, 5-Dihydroorotic acid	0.95	0.005	2.59 (2.31 - 2.87)
9-HETE	0.45	0.031	1.56 (1.27 - 1.84)

Abbreviations: CI = confidence interval; SCLC = small cell lung cancer; PE = phosphatidylethanolamine; 9-HETE = 9-Hydroxyeicosatetraenoic acid.

Supplementary Table 6. The validation of the developed discriminant models using another independent cohort.

Gender	Discriminant models	Accuracy (%) (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
Male	PI(18:0/18:0), Cer(d18:1/22:0 OH), 2-Arachidonylglycerol, IMP and Cholic acid	91.3 (74.1-97.1)	86.6 (59.5-98.3)	84.1 (63.9-95.5)	76.5 (56.4-89.1)	91.3 (74.4-97.5)
Female	PE(18:1/20:4), 5-Methyltetrahydrofolic acid, Desmosterol, 4,5-Dihydroorotic acid and 9-HETE	81.3 (54.4-95.9)	86.4 (65.1-97.1)	84.2 (68.8-93.4)	81.3 (59.6-92.7)	86.4 (69.3-94.7)

Abbreviations: 9-HETE = 9-hydroxyeicosatetraenoic acid; Cer = Ceramide; PI = Phosphatidylinositol; CI = confidence interval; IMP = inosine monophosphate; PE = phosphatidylethanolamine.