

Plasma lipid biomarker signatures in squamous carcinoma and adenocarcinoma lung cancer patients

Srinivasarao Ravipati, David R Baldwin⁺, Helen L Barr⁺⁺,
Andrew W. Fogarty⁺⁺⁺, David A. Barrett

Centre for Analytical Bioscience, School of Pharmacy, University of Nottingham, University Park, NG7 2RD, UK

⁺ Department of Respiratory Medicine, David Evans Centre, City Hospital Campus, Nottingham, NG5 1PB, UK.

⁺⁺ Division of Respiratory Medicine, Clinical Sciences Building, University of Nottingham, City Hospital Campus, Nottingham, NG5 1PB, UK.

⁺⁺⁺ Division of Epidemiology and Public Health, Clinical Sciences Building, University of Nottingham, City Hospital Campus, Nottingham, NG5 1PB, UK.

Address for Correspondence:

D.A. Barrett, Centre for Analytical Bioscience, School of Pharmacy, University of Nottingham, University Park, NG7 2RD, UK

Email: david.barrett@nottingham.ac.uk

Tel.: +44(0)1159515062. Fax: +44(0)1159515102

Running title: lipidomic biomarkers for lung cancer

Abstract

There is a clinical need for reliable biomarkers for lung cancer that permit early diagnosis of the disease and provide prediction of histological phenotype. A prospective study design was used with a study population of patients with suspected lung cancer. Blood samples were collected from 17 patients with histologically confirmed squamous cell lung carcinoma, 17 individuals with adenocarcinoma, and 17 control individuals who did not subsequently have a diagnosis of lung cancer or any other cancer. Blood plasma samples were analysed for their lipid profiles using liquid chromatography coupled with high resolution mass spectrometry. Data were analysed using multivariate statistical methods. There was good separation between histological subtypes and control groups and also between individuals with a subsequent diagnosis of adenocarcinoma and squamous cell carcinoma (sensitivity 80%, specificity 83%, $Q^2=0.70$). Alterations in the levels of different classes of lipids including triglycerides (TGs), phosphatidylinositols (PIs), phosphatidylcholines (PCs), phosphatidylethanolamines (PEs), free fatty acids, lysophospholipids and sphingolipids were observed in squamous carcinoma and adenocarcinoma lung cancer patients when compared with control patients. In conclusion, this study has identified candidate lipid biomarkers of non-small cell lung cancer patients which may be helpful to indicate the tumour subtype and to differentiate them from patients who do not have lung cancer. Measuring these biomarkers has the potential to improve diagnosis in patients with suspected lung cancer and risk stratification in screening.

Keywords: lung cancer, squamous carcinoma, adenocarcinoma, lipidomics

Introduction

Worldwide, more than 1.8 million people were newly diagnosed with lung cancer in 2012, and it now represents the leading global cause of death from cancer, accounting for 19.4 % of all cancer deaths in recent years (GLOBOCAN 2012). 75% of patients with lung cancer present with advanced, incurable disease (Oak et al. 2012), which is associated with five year survival rates of approximately 9% (De Angelis et al. 2014). However, survival rates are significantly better in those individuals who present earlier, with five year survival rates of 60-75% (Scott et al. 2007). Lung cancers are generally divided into two main categories: small cell lung cancer and non-small cell lung cancer (NSCLC). NSCLC accounts for approximately 90% of all lung cancers. NSCLC is divided further into adenocarcinoma, squamous cell carcinoma (SCC), and large cell carcinoma histologies (Cheng et al. 2012). Lung cancer is often insidious, and it may produce no symptoms until the disease is well advanced. Approximately 7-10% of patients with lung cancer are asymptomatic, and their cancers are diagnosed incidentally after a chest radiograph performed for other reasons (Maghfoor I, 2014). While the management options for lung cancer have increased substantially in the past decade, rapid diagnosis remains a challenge and new diagnostic tests would have the potential to benefit a large number of patients. In addition, new diagnostic tests for lung cancer would (together with existing investigations) have potential as prognostic and responsiveness to treatment indicators.

There is thus a pressing need for a new biofluid-based screening test that can identify high-risk individuals who can then be investigated using more invasive tissue biopsy methodology (Hassanein et al. 2012). The crucial role of lipids in cell, tissue and organ physiology is demonstrated by a large number of genetic studies and by many human diseases that involve the disruption of lipid metabolic enzymes and pathways (Wenk, 2005). A number of LC-MS methods have been applied to discover serum biomarkers of tumours, including those with high mortality, such as ovary, lung, liver, and pancreatic cancer. To date the large majority of these biomarkers have been identified as peptides or proteins. Proteomic analytical techniques (Yanagisawa et al. 2003, Hassanein et al. 2011) have been used in the study of lung cancer,

and this has yielded protein biomarkers that have been demonstrated to have high specificity in a study of 54 individuals diagnosed with NSCLC (Zeng et al. 2011). Serologic biomarkers of lung cancer have emerged recently: these include carcinoembryonic antigen, the cytokeratin 19 fragment CYFRA21-1, cancer antigen CA-125 (Cedres et al. 2011), plasma kallikrein (Chee et al. 2008), progastrin-releasing peptide (ProGRP), and neuron-specific enolase (NSE) (Wojcik et al. 2008). However, there are relatively few studies of small molecule serological biomarkers of lung cancer. Lipids have a prominent biology in lungs as surfactants (Serrano et al., 2006) and prostanoids (Keith et al. 2006), and in phospholipid signalling-related cancer biology (e.g. PI3-kinase ((Kong et al. 2008)).

Most of the studies of the role of lipids in lung cancer involve the targeted analysis of specific lipids. There is evidence for the involvement of oxylipins (Poczobutt et al. 2006) and phospholipid species (Tyurina et al. 2011) in tumour development from animal models of lung cancer. This is supported by clinical studies showing altered profiles of various lipids in plasma of lung cancer patients, including lysophospholipids (Dong et al. 2010; Guo et al. 2012), free fatty acids (Liu et al. 2014), sphingomyelins, plasmalogens, phosphoglycerides and related lipids (Smith et al. 2008; De Castro et al. 2008; Ma et al. 2011). Lipid profiling or lipidomics is an emerging field within metabolomics which focuses on monitoring changes in biofluid profiles of lipids and factors that interact with lipids. The chemical complexity of lipids, with many classes and sub-classes, makes lipid profiling of biofluids a challenging task (Yang et al. 2011). However, recent improvements in analytical approaches using ultra high performance liquid chromatography (UHPLC) and high resolution mass spectrometry (HRMS) make this field a promising new area for biomarker research (Yang et al. 2011, Gao et al. 2012, Haag et al. 2012). Here we apply UHPLC-HRMS lipidomics analysis to blood samples from patients with confirmed squamous cell or adenocarcinoma to test the hypotheses that blood small lipid molecules are biomarkers for lung cancer and can provide data to distinguish between squamous and adenocarcinoma histological sub-types of lung cancer.

Methods

Study population

The study population was recruited from individuals referred with suspected lung cancer to the lung cancer clinic at Nottingham University Hospitals. After providing consent, blood samples were collected in pre-chilled lithium heparin tubes and centrifuged within 30 min at 2000×g and 4°C for 10 min. Plasma was separated and immediately stored at -80°C until analysis. After the patients had received a diagnosis of squamous carcinoma or adenocarcinoma or confirmation that there was no evidence of lung cancer, the archived plasma samples were analysed for their lipid profiles. The study was approved by the Nottingham Research Ethics Committee (09/H0403/68).

Reagents and materials

A Milli-Q water purification system (Millipore, MA, USA) was used in the preparation of deionized water (18.2 mΩ). Acetonitrile, chloroform, ethyl acetate and hexane were HPLC grade purchased from Fischer scientific (Loughborough, UK). Methanol (LC-MS grade) and ammonium acetate were purchased from Sigma-Aldrich (Darmstadt, Germany). Isopropanol (LC-MS grade) and ethanol AR grade were obtained from Fischer Scientific (Loughborough, UK). 3,5 Di-tert-butyl-4-hydroxytoluene (BHT) was purchased from Supelco (PA, USA).

Sample preparation

Lipids were extracted from plasma samples (50 µL) by adding 0.5 mL of ice-cold (-20°C) chloroform/methanol (1:2), the frozen plasma sample being allowed to thaw in the presence of the extraction solvent. After brief vortex-mixing (20 s), 0.5 mL of water was added to the tube contents and mixed again for 10 min, centrifuged at 1000 g for 10 min at 4°C. An aliquot of the lower lipophilic phase (100 µL) was removed and mixed with an equal volume of isopropanol prior to injection.

Liquid chromatography-mass spectrometry (LC-MS) lipidomic analysis

LC-MS lipidomics analysis with high resolution mass spectrometry detection used an Accela high speed LC coupled to an Exactive MS (Thermo Fisher Scientific, USA). Data was acquired simultaneously in full scan ion mode (m/z 100-1200, resolution 25,000) in both positive and negative electrospray ionisation modes. The capillary temperature and heater temperature were maintained at 350° C and 300°C respectively in both positive and negative modes. The maximum scan inject time and micro scans were 100 ms and 1 respectively. LC-MS analysis was performed on an ACE C4, 300Å column (100x2.1 mm, 3 µm particle size; Phenomenex, Aberdeen, UK) maintained at a temperature of 40° C and a flow rate of 300 µL/min. The mobile phase consisted of (A) 5 mM ammonium acetate in water, (B) 5 mM ammonium acetate in methanol and (C) 5 mM ammonium acetate in acetonitrile. A binary gradient (30-100% B) from channels A, B was used for 0-16 min, 100% C for 16-18 min and 100%B from 19-20 min with a total run time of 20 min. The injection volume was 10 µL. All study samples and QC samples were analysed in a single continuous analytical run.

Data Analysis and Metabolite Identification

LC-MS raw data from the analysis of extracts of study and control samples were acquired using Xcalibur v2.1 software (Thermo Scientific, Hemel Hempstead UK). Each sample analysis produced a three dimensional data set consisting of m/z , retention time and ion signal intensity. The complete datasets from the adenocarcinoma, squamous patients, unknown histology patients and control groups were imported and processed by SIEVE v1.2 software (Thermo Scientific, Hemel Hempstead UK). Analytical method performance was validated by examining a representative set of 16 plasma lipids (details in Supplementary Information) in a series of QC samples prepared from a single separate source of control human plasma used in our laboratory ('plasma QC'). for retention timeshifts, relative standard deviations (RSD%) of peak areas and mass accuracy. SIMCA-P 13.0.2 version (Umetrics AB, Umea, Sweden) was used to carry out multivariate analysis including principal component analysis (PCA) and orthogonal partial least square-discriminant analysis (OPLS-DA). OPLS-DA was used to investigate the differences in lipid profiles between the two lung cancer histological groups and the control group data sets.

Initial models based on the entire datasets (n=17 in each group) were internally cross-validated. Further prediction models were based on randomly selected training and test sets for each group with sensitivity and specificity calculations reported. Potential candidates for discriminant markers were selected using S-plots by setting the cutoff values for both covariance $p(1)$ on x-axis and the correlation value, $p[\text{corr}]$ on y-axis. The two tails of S-plots represents the most contributing components to differentiate the two groups with more confidence and they are considered as potential candidates for discriminant lipid biomarkers. Candidates for discriminant biomarkers were also selected by variable importance in projection values (VIP) where VIP values greater than one were considered further as potential biomarkers. Tentative identification of key lipid biomarkers was achieved by using accurate mass determinations within a narrow m/z range (1 mDa in positive mode and 2 mDa in negative mode) to search appropriate metabolite databases including Lipid Maps (<http://www.lipidmaps.org/>) and the Human Metabolome database (<http://www.hmdb.ca/>).

Results

Demographics and clinical features of subjects

Table 1 shows the baseline demographics of the subjects. The groups were similar in terms of age, sex, current or ex-smoking habit and co-morbidity. Although there were more subjects reporting weight loss in the squamous cancer and adenocarcinoma groups, a quarter of the controls also reported weight loss. This reflected the nature of the controls, being patients rather than healthy volunteers.

Validation of LC-MS lipidomic method performance

The performance of the analytical LC-MS lipidomics method was evaluated using the plasma QC samples. All sample and plasma QC extracts were analysed in a single run where QC samples were interspaced with study samples. The plasma QC samples were closely clustered by PCA analysis and well separated from the study samples (Figure 1). In the QC datasets the %RSD values of peak areas of the representative plasma lipids were in the range of 5.6 to 10.2%,

retention time shifts were less than 0.1 min, the mass accuracy deviation ranged from 0.27 to 0.80 mDa in positive ion mode and 0.17 to 1.21 mDa in negative ion mode (Full QC information is shown in Supplementary Table S6). These QC results validate the LC-MS lipidomics analytical performance during the analysis of the study samples.

Plasma lipidomic analysis of non-small cell lung cancer patients and healthy volunteers

Principal Component Analysis (PCA, unsupervised) was used to give an overview of the data sets but did not reveal any clear separation between the cancer and control groups (Figure 1). Supervised orthogonal partial least squares discriminate analysis (OPLS-DA) was used to analyse the data sets to generate models to differentiate between squamous carcinoma (n=17), adenocarcinoma (n=17) and control (n=17) subjects (Figure 2), and to identify discriminating lipid ions that contributed to the separation of cancer and control groups. OPLS-DA shows a reasonable separation between control samples and either squamous carcinoma or adenocarcinoma samples, but does not distinguish clearly between the squamous and adenocarcinoma patients when all three groups are included in a single analysis. Therefore, further analysis was undertaken to examine controls versus squamous carcinoma and controls versus adenocarcinoma datasets separately.

Analysis of squamous carcinoma vs. control lipidomics data sets

Using OPLS-DA analysis a clear separation was observed between squamous carcinoma samples and control samples (Figure 3). The S-plot (Supplementary Figure S8A) shows the contribution of the measured variables in both squamous carcinoma patients and control groups and it was used to guide the identification of potential lipid biomarkers in addition to VIP scores. To further validate the differences observed between the control and squamous carcinoma samples, the OPLS-DA model was validated using a training set containing 12 squamous carcinoma patients and 11 control patients and a test set containing 5 squamous carcinoma patients and 6 control patients. The model build using the training set was used to predict the status of test set subjects (Supplementary Figure S9). Despite the relatively small size of the

study, Table 5 shows that the calculated sensitivity and specificity values with 4 out of 5 squamous carcinoma samples predicted correctly and 4 out of 6 control subject samples predicted correctly. The model was further validated by calculating area under receiver operating characteristic (ROC) curve (Eng 2007) (Figure 6A).

Analysis of adenocarcinoma vs. control lipidomic data sets

OPLS-DA analysis of the data sets generated from adenocarcinoma lung cancer patients (n=17) and healthy volunteers (n=17) gave separation between the two groups (Figure 4). The S-plot (Supplementary Figure S8B) shows the contribution of the measured variables in both adenocarcinoma patients and control groups and it was used to guide the identification of potential lipid biomarkers. As described previously, an OPLS-DA prediction model was built using subjects in the training set (adenocarcinoma patients (n=12) and control patients (n=11) to predict adenocarcinoma patients (adenocarcinoma patients (n=5) and control patients (n=6)). Using the prediction model 4 out of 5 adenocarcinoma samples were predicted correctly and 5 out of 6 control subject samples were predicted correctly (Supplementary Figure S10) and sensitivity and specificity values are reported in Table 5. The model was validated by calculating area under receiver operating characteristic (ROC) curve (Figure 6B).

Analysis of Adenocarcinoma Vs. Squamous datasets

Finally, a supervised OPLS-DA model was generated which showed a clear separation between adenocarcinoma and squamous carcinoma patients (Figure 5). The potential biomarkers were identified using S-plot (Supplementary Figure S8C). A training set containing 12 adenocarcinoma patients and 11 squamous carcinoma patients and a test set containing 5 adenocarcinoma patients and 6 squamous carcinoma patients were used. The obtained R²X and Q² values from training set were 0.61 and 0.70 respectively demonstrating an acceptable model. The model generated from training set was used to predict the classification of the remaining samples in test set. Finally, 4 out of 5 adenocarcinoma samples were predicted correctly and 5 out of 6 squamous carcinoma samples were predicted correctly (Supplementary Figure S11) and sensitivity and

specificity values of 80% and 83% respectively obtained (Table 5). The model was further validated by calculating area under receiver operating characteristic (ROC) curve (Figure 6C).

Plasma lipid biomarkers of squamous and adenocarcinoma

Potential markers contributing strongly to the separation between the plasma lipid profiles of squamous carcinoma adenocarcinoma and control subjects were identified using S plots (Figure S7). The exact masses of these biomarkers were then used to interrogate metabolite/lipid databases. The biomarkers that contributed most to the squamous carcinoma signature were membrane lipids including triglycerides (TG), ceramides (CE), sphingomyelins (SM), phosphatidylcholines (PC), lysophosphatidylcholines (LysoPC) and phosphatidylethanolamines (PE) (Table 2) while the biomarkers that contributed to adenocarcinoma were triglycerides (TG), diglycerides (DG), phosphatidylserines (PS), fatty acids, plasmalogens, lysophosphatidylethanolamines (LysoPE) and phosphatidylethanolamines (PE) (Table 3). The biomarkers mostly contributed to the differentiation between adenocarcinoma and squamous carcinoma histological subtypes were phosphatidylinositol (PI), triglycerides (TG), ceramides (CE), sphingomyelins (SM), phosphatidylcholines (PC), phosphatidylethanolamines (PE) and plasmalogens (Table 4). Figure 7 shows elevated levels of PS(36:1) observed in adenocarcinoma patients when compared to control and elevated levels of SM(d18:1/16:0) found in plasma of adenocarcinoma patients when compared with squamous carcinoma patients.

Discussion

We have used lipidomic analysis of blood plasma to classify those with and without lung cancer, and in those with lung cancer to distinguish between those having adenocarcinoma and squamous cell histological phenotypes. Our pilot data demonstrate the potential for this analytical methodology in the context of screening programmes for lung cancer and clinical management of patients with suspected lung cancer. In our small-scale study the obtained sensitivity and specificity values for the identification of adenocarcinoma were 80.0%, and 83.3% respectively and 80.0% and 66.7% for squamous carcinoma. These values are comparable with previous

protein-based biomarker studies for lung cancer which achieved sensitivity and specificity of 86.9% and 80% respectively (Yang et al. 2005) and 87.3% and 81.9% (Sreseli et al. 2010) and >70% (Liu et al. 2014).

The major strengths of our study are that the samples and controls were taken prospectively from patients attending the same cancer clinic with similar population demographics and that the analysis was performed by an investigator who was blinded to the cancer status. The use of patients with suspected lung cancer who provided prospective blood samples prior to learning if they had lung cancer or not is also important, as the psychological stress experienced by both those with and those without cancer will be similar, and it is possible that this could manifest itself in the lipidomic signature if we used a 'healthy' population as our control group. The plasma samples were rigorously collected using a standardised protocol and the samples analyses were conducted with no awareness of the cancer status of the samples.

Broadly, both squamous and adenocarcinomas were characterised by an increase in plasma di- and tri-glycerides and a decrease in phospholipids and lysophospholipids, although there were some exceptions to this (Tables 2&3). Cancer cells contain as high as 6.8% triglyceride fraction of total plasma membrane lipids (May et al. 1986). Changes in triglyceride levels in plasma have been associated with cancers (Ma et al. 2011, Fiorenza et al. 2000, Ulmer et al. 2009). Previous studies have noted a decrease in plasma phospholipids and lysophospholipids in advanced cancer, attributed to a loss of essential fatty acids in the periphery (Murphy et al. 2012, 2010, 2007). Hence our observations are consistent with previous studies. The variations in the levels of phospholipid fatty acids between stages of disease may be based on many factors such as greater demand for phospholipid in the context of lung cancer because of higher levels of cell replication and modified phospholipid metabolism (Murphy et al. 2012). The involvement of lysophosphatidylcholine, a precursor of lysophosphatidic acid has been observed in the progression of ovarian and other cancers and abnormal levels of lysophospholipids in plasma may be indicators of the presence of cancer (Murphy et al. 2007).

The lipid biomarkers which contributed most to differentiate between adenocarcinoma and squamous carcinoma were phospholipids, triglycerides and sphingomyelins (Table 4). Phospholipids such as phosphatidylinositols are precursors for secondary messengers and involve in controlling of various cellular mechanisms like cell growth, motility and proliferation through definitive interactions of proteins which bind to their phosphorylated head groups. These play an important role in cancer biology. The 3-phosphorylated phosphoinositides are produced by phosphatidylinositol-3- kinases (PI3Ks) which are most common drug targets in cancer treatment. The mutation of PI3K enzyme in some cancers leads to increased levels of 3-phosphorylated phosphoinositides which initiate growth factor stimulation and activation of protein kinase B and phosphoinositides dependent kinases (Fernandis et al. 2009, Wakelam et al. 2007). Plasmalogens contributed as suitable biomarkers in differentiating adenocarcinoma patients and healthy controls and also involved in differentiation of squamous and adenocarcinoma patients. Phospholipids that contain a vinyl ether-linkage are commonly known as plasmalogens. They involve in controlling of membrane fluidity and acts as mediators of membrane dynamics. They are also control the damage of nervous tissue associated with the accumulation of very long chain fatty acids. They are also have antioxidant capability and impaired and/or reduced levels of plasmalogens lead to human pathological disorders such as Alzheimer's disease and metastatic cancer (Brites et al. 2004, Smith et al. 2008).

The limitations of our data include the small size of the study population that is inevitable for a pilot study of a new technique. This may result in type II statistical error as we were unable to identify all small lipid molecules that contribute to the lipidomic signatures associated with non-small cell lung cancers. We were also unable to adjust for other demographic and lifestyle exposures that may modify the lipidomic signatures associated with non-small cell lung cancer. However, the fact that we were able to see differences despite these limitations suggests that these observations need further study initially to clarify if they are consistent and replicatable. It is also recognised that the distinction between adenocarcinoma and squamous subtype based on small biopsies and histology can be a simplification of what is found when the whole tumour

is examined. A small proportion of whole tumours may contain both subtypes. This might explain some overlap between the two groups.

The field of screening for non-small cell lung cancer is a challenging one to work in as definitive studies require high-risk populations of thousands of individuals to permit the detection of a clinically worthwhile difference in survival. Hence, the use of lipidomic methodology to permit the early assessment of individuals with suspected lung cancer is probably the field where these observations have the greatest potential to improve clinical care. In addition, we did not include individuals with small-cell lung cancer as this has a lower incidence and hence we would not have had sufficient power to permit identification of differences from those with either non-small cell lung cancer or no cancer, but clearly in larger studies this is an important group clinically. Although the effect of lung cancer on the plasma lipid signature may be responsible for our observed changes in the plasma lipid profiles, there may be other factors that we have not specifically considered such as diet, medication and nutrition which may also influence the lipid profile.

In conclusion, this small pilot study of the use of lipidomics has demonstrated that a lipidomics approach can identify individuals with non-small cell lung cancer from those who do not have lung cancer, and also has the potential to distinguish individuals with adenocarcinoma lung cancer from those with squamous cell lung cancer, prior to diagnosis. The next step is for larger studies that also include patients with small cell lung cancer to assess the potential of lipidomic signatures in the management of patients in the lung cancer clinic.

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Funding: this study was supported by funds provided by the Nottingham University Hospitals' Charity

References

Bird, S.S., Marur, R.V., Sniatynski, J.M., Greenberg, K.H. & Kristal, S.B. (2011). Serum lipidomics profiling using LC-MS and high energy collisional dissociation fragmentation: Focus on triglyceride detection and characterization. *Analytical chemistry*. 83, 6648-6657.

Brites, P., Waterham, H. R., Wanders, R. J. A. (2004). Functions and biosynthesis of plasmalogens in health and disease. *Biochimica et Biophysica Acta*. 1636, 219-231.

Cedres, S., Nunez, I., Longo, M., Martinez, P., Checa, E., Torrejon, D. & Felip, E. (2011). Serum tumor markers CEA, CYFRA21-1, and CA-125 are associated with worse prognosis in advanced non-small-cell lung cancer (NSCLS). *Clinical Lung Cancer*. 12, 172-9.

Chee, J., Naran, A., Misso, N. L., Thompson, P. J & Bhoola, K. D. (2008). Expression of tissue and plasma kallikreins and kinin B1 and B2 receptors in lung cancer. *The Journal of Biological Chemistry*. 389, 1225-1233.

Cheng, L., Alexander, R.E., MacLennan, G.T., Cummings, O.W., Montironi, R., Lopez-Beltran, A., Cramer, H.M., Davidson, D.D. & Zhang, S. (2012). Molecular pathology of lung cancer: key to personalized medicine. *Modern Pathology*. 25, 347-369.

de Angelis, R., Sant, M., Coleman, M. P., Francisci, S., Balli, P. et al. (2014). Cancer survival in Europe 1999-2007 by country and age: results of EURO CARE-5- a population-based study. *Lancet Oncology*. 15, 23-34.

de Castro, J., Rodriguez, M.C., Martinez-Zorzano, V. S., et al. (2008). Erythrocyte and platelet phospholipid fatty acids as markers of advanced nonsmall cell lung cancer: comparison with serum levels of sialic acid, TPS and Cyfra 21-1. *Cancer Investigation*, 26, 407-418.

Dong, J., Cai, X., Zhao, L., Xue, X., Zou, L., Zhang, X., Liang, X. (2010). Lysophosphatidylcholine profiling of plasma: discrimination of isomers and discovery of lung cancer biomarkers. *Metabolomics*. 6, 478-488.

Eng, J. (2007). ROC analysis: Web-based calculator for ROC curves. <http://www.jrocf.it.org>. Accessed Dec 2014.

Fernandis, Z. A. & Wenk, R.M. (2009). Lipid based biomarkers for cancer. *Journal of chromatography B*. 877, 2830-2835.

Fiorenza, A. M., Branchi, A., Sommariva, D. (2000). Serum lipoprotein profile in patients with cancer. A comparison with non-cancer subjects. *International Journal of Clinical and Laboratory Research*. 30(3), 141-145.

Fuchs, B. & Schiller, J. (2009). Application of MALDI-TOF mass spectrometry in lipidomics. *European Journal of Lipid Science and Technology*. 111,83-98.

Gao, X., Zhang, Q., Meng, D., Zhao, R., Fillmore, T.L., Chu, R.K., Zhou, J., Tang, K., Hu, Z., Moore, R.J., Smith, R.D., Katze, M.G. & Metz, T.O. (2012). A reversed-phase capillary ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) method for comprehensive top-down/bottom-up lipid profiling. *Analytical and Bioanalytical Chemistry*. 402, 2923-2933.

GLOBOCAN 2012: Estimated cancer incidence, mortality and prevalence worldwide in 2012. Revised 2015 consensus on section of cancer surveillance. http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx.

Guo, Y., Wang, X., Qiu, L., et al. (2012). Probing gender-specific lipid metabolites diagnostic biomarkers for lung cancer using Fourier transform ion cyclotron resonance mass spectrometry. *Clinica Chimica Acta*. 141, 135-141.

Haag, M., Schmidt, A., Sachsenheimer, T. & Brugger, B. (2012). Quantification of signaling lipids by nano-electrospray ionization tandem mass spectrometry (Nano-ESI MS/MS). *Metabolites*, 2, 57-76.

Hassanein, M., Callison, J. C., Callaway-Lane, C., Aldrich, M. C., Grogan, E. L. & Massion, P. P. (2012). The state of molecular biomarkers for the early detection of lung cancer. *Cancer Prevention Research*. 5, 992-1006.

Hassanein, M., Rahman, J. S. M., Chaurand, P. & Massion, P. P. (2011). Advances in proteomic strategies towards the early detection of lung cancer. *Proceedings of the American Thoracic Society*. 8, 183-188.

Ikeda, K., Oike, Y., Shimizu, T. & Taguchi, R. (2009). Global analysis of triacylglycerols including oxidized molecular species by reverse-phase high resolution LC/ESI-QTOF MS/MS. *Journal of Chromatography B*. 877, 2639-2647.

Imbert, L., Gaudin, M., Libong, D., Touboul, D., Abreu, S., Loiseau, P. M., Laprevote, O. & Chaminade, P. (2012). Comparison of electrospray ionization, atmospheric pressure chemical ionization and atmospheric pressure photoionization for a lipidomic analysis of *Leishmania donovani*. *Journal of Chromatography A*. 1242, 75-83.

Keith, R. L., Geraci, M. W. (2006). Prostacyclin in lung cancer. *Journal of Thoracic Oncology*. 1, 503–505.

Kong, D., Yamori, T. (2008). Phosphatidylinositol 3-kinase inhibitors: promising drug candidates for cancer therapy. *Cancer Science*. 99, 1734–1740.

Liu, J., Mazzone, P. J., Cata, J. P., Kurz, A., Bauer, M., Mascha, E. J., Sessler, D. I. (2014). Serum free fatty acid biomarkers of lung cancer. *Chest*, 146, 670-679.

Liu, Ying., Chen, Y., Shaner, R., Wang, E., Bowen, J.N., Matyunina, V.L., Walker, D.L., McDonald, F.J., Sullards, C.M. & Merrill Jr, H.A. (2010). Elevation of sulfatides in ovarian cancer: An integrated transcriptomic and lipidomic analysis including tissue-imaging mass spectrometry. *Journal of Molecular Cancer*. 9:186.

Ma, X. & Yang, J. (2011). Lipidomics in cancer biomarker discovery. *Omics technologies in cancer biomarker discovery edited by Zhang, X., Landis Bioscience, Texas USA*. 101-115.

Maghfoor I Lung Cancer, Non-Small Cell. <http://www.emedicine.com/med/topic1333.htm> .

Accessed December, 2014.

Masood, A.M., Rao, P.R., Acharya, K.J. & Blonder, J. (2012). Quantification of multiple sphingolipid classes using normal and reversed-phase LC-ESI-MS/MS: Comparative profiling of two cell lines. *Lipids*. 47,209-226.

May, G. L., Wright, L. C., Holmes, K.T., Williams, P. G., Smith, I.C.P., et al. (1986). Assignment of Methylene Proton Resonances in Nmr-Spectra of Embryonic and Transformed-Cells to Plasma-Membrane Triglyceride. *Journal of Biological Chemistry*. 261, 3048–3053.

Milne, S., Ivanova, P., Forrester, J. & Brown, H. A. (2006). Lipidomics: an analysis of cellular lipids by ESI-MS, *Methods*. 39, 92-103.

Murphy, M., Tanaka, T., Pang, J., Felix, E., Liu, S., Trost, R., Godwin, K.A., Newman, R. & Mills, G. (2007). Liquid chromatography mass spectrometry for quantifying plasma lysophospholipids: Potential biomarkers for cancer diagnosis. *Methods in enzymology*. 433, 1-25.

Murphy, M., Tanaka, T., Pang, J., Felix, E., Liu, S., Trost, R., Godwin, K.A., Newman, R. & Mills, G. (2007). Liquid chromatography mass spectrometry for quantifying plasma lysophospholipids: Potential biomarkers for cancer diagnosis. *Methods in enzymology*. 433, 1-25.

Murphy, R. A., Wilke, M. S., Perrine, M., Pawlowicz, M., Mourtzakis, M., Lieffers, J. R., Maneshgar, M., Bruera, E., Clandinin, M.T., Baracos, V. E., Mazurak, V.C. (2010). Loss of adipose tissue and plasma phospholipids: Relationship to survival in advanced cancer patients. *Clinical Nutrition*, 29(4), 482–487.

Murphy, R. A., Bureyko, T. F., Mourtzakis, M., Quincy, S., Chu, M., Clandinin, T.M., Reiman, T. & Mazurak, V. C. (2012). Aberrations in plasma phospholipid fatty acids in lung cancer patients. *Lipids*. 47, 363-369.

Oak, C.H., Wilson, D., Lee, H. J., Lim, H-J. & Park, E-K. (2012). Potential molecular approaches for the early diagnosis of lung cancer. *Molecular Medicine Reports*. 6, 931-936.

Poczobutt, J. M., Gijon, M., Amin, J., Hanson, Dwight., Li, H., Walker, D., Weiser-Evans, M., Lu, X., Murphy, R. C., Nemenoff, R. A. (2006). Eicosanoid profiling in an orthotopic model of lung cancer progression by mass spectrometry demonstrates selective production of leukotrienes by inflammatory cell of the microenvironment. *British Journal of cancer*. 95, 1131-1135.

Scott, W. J., Howington, J., Feigenberg, S., Movsas, B. & Pisters, K. (2007). American College of Chest Physicians: Treatment of non-small cell lung cancer stage I and stage II: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 132, 234S-242S.

Serrano, A. G., Perez-Gil, J. (2006). Protein-lipid interactions and surface activity in the pulmonary surfactant system. *Chemistry and Physics of Lipids*. 141, 105–118.

Smith, R. E., Lepsi, P., Di Luca, M., Bustos, C., Marra, F., A., de Alaniz, M. J., Marra, C. A. (2008). A reliable biomarker derived from plasmalogens to evaluate malignancy and metastatic capacity of human cancers. *Lipids*. 43, 79-89.

Sreseli, R.T., Binder, H., Kuhn, M., Digel, W., Veelken, H., Siemel, W., Passlick, B, Schumacher, M., Martens, U.M., Zimmermann, S. (2010). Identification of a 17-protein signature in the serum of lung cancer patients. *Oncology Reports*. 24, 263-270.

Tyurina, Y. Y., Tyurin, V. A., Kapralova, V. I., Wasserloos, K., Mosher, M., Epperly, M. W., Greenberger, J. S., Pitt, B. R. & Kagan, V. E. (2011). Oxidative lipidomics of γ -radiation-induced lung injury: Mass spectrometric characterization of cardiolipin and phosphatidylserine peroxidation. *Radiation research*. 175(5), 610-621.

Ulmer, H., Borena, W., Rapp, K., Klenk, J., Strasak, A., Diem, G., Concin, H., Nagel, G. (2009). Serum triglyceride concentrations and cancer risk in a large cohort study in Austria. *British Journal of Cancer*. 101(7), 1202-1206.

Wakelam, M. J., Pettitt, T. R., & Postle, A. D. (2007). Lipidomic analysis of signaling pathways. *Methods in enzymology*. 433, 233-245.

Wenk, M.R. (2005). The emerging field of lipidomics. *Nature reviews*. 4, 594-610.

Wojcik, E., Kulpa, J.K., Sas-Korczynska, B., Korzeniowski, S., Jakubowicz, J. (2008). ProGRP and NSE in Therapy Monitoring in Patients with Small Cell Lung Cancer. *Anticancer Research*. 28, 3027-3033.

Yanagisawa, K., Shyr, Y., Xu, B. J., Massion, P. P., Larsen, P. H., White, B.C., Roberts, J. R., Edgerton, M., Gonzalez, A., Nadaf, S., Moore, J. H., Caprioli, R M. & Carbone, D. P. (2003). Proteomic patterns of tumour subsets in non-small-cell lung cancer. *The Lancet*. 362, 433-39.

Yang, H., Zhang, W., Yang, H., Hu, C. (2005). Cathepsin B expression and its significance in non-small cell lung cancer tissues. *Journal of Central South University. Medical Sciences*. 30,729-30.

Yang, K. & Han, X. (2011). Accurate quantification of lipid species by electrospray ionization mass spectrometry- meets a key challenge in lipidomics. *Metabolites*. 1, 21-40.

Zeng, X., Hood, B. L., Zhao, T., Conrads, T. P., Sun, M., Goplakrishnan, V., Grover, H., Day, R. S., Weissfeld, J. L., Wilson, D. O., Siegfried, M., Bigbee, W. L. (2011). Lung cancer serum biomarker discovery using label free LC-MS/MS. *Journal of Thoracic Oncology*. 6, 725-734.

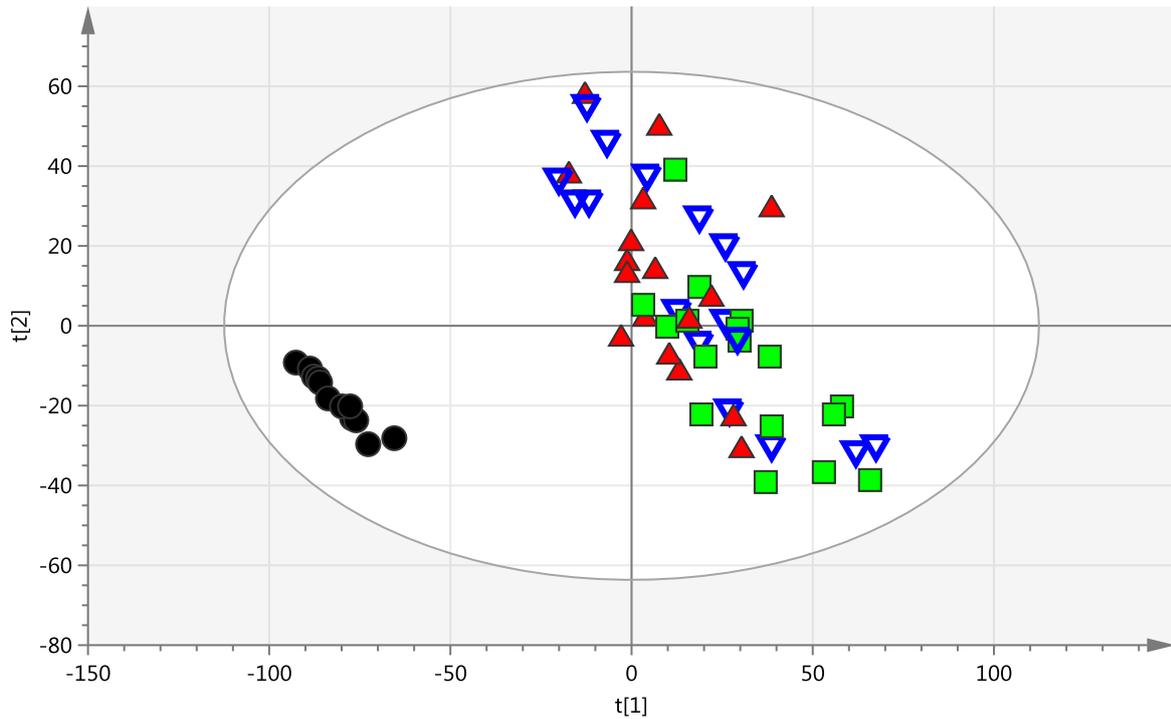


Figure 1: Principal component analysis (PCA) scores plot obtained from all lung cancer study samples (squamous carcinoma – triangles \blacktriangle (n=17), adenocarcinoma - inverted open triangles \blacktriangledown (n=17) and healthy individuals – squares \blacksquare (n=17)), and QCs –circles \bullet (n=12) samples ($R^2X=0.710$, $Q^2=0.390$, $A=11$, $N=63$).

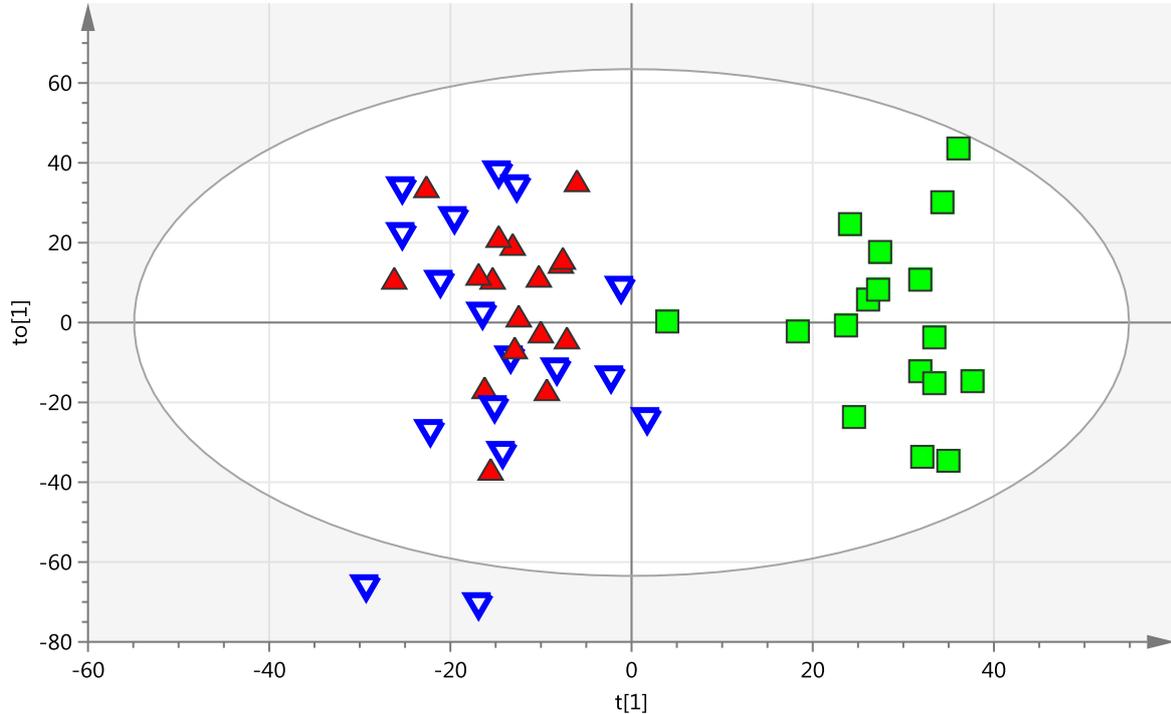


Figure 2: Orthogonal partial least squares discriminant analysis (OPLS-DA) scores plot obtained from squamous carcinoma samples (triangles \blacktriangle), adenocarcinoma samples (open inverted triangles \blacktriangledown) and control samples (squares \blacksquare) ($R^2X=0.260$, $R^2Y=0.879$, $Q^2=0.452$, $A=1+2+0$, $N=51$).

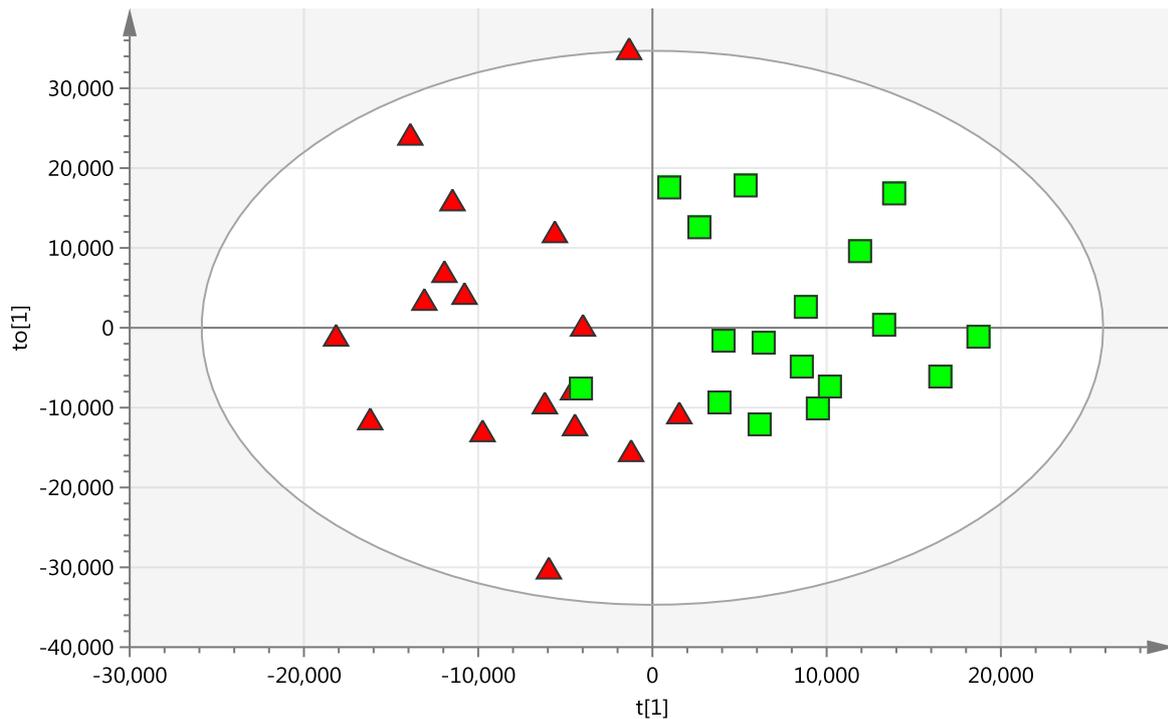


Figure 3: OPLS-DA scores plot obtained from squamous carcinoma (triangles ▲) and control (squares ■) samples. ($R^2X=0.262$, $R^2Y=0.668$, $Q^2=0.181$, $A=1+1+0$, $N=34$)

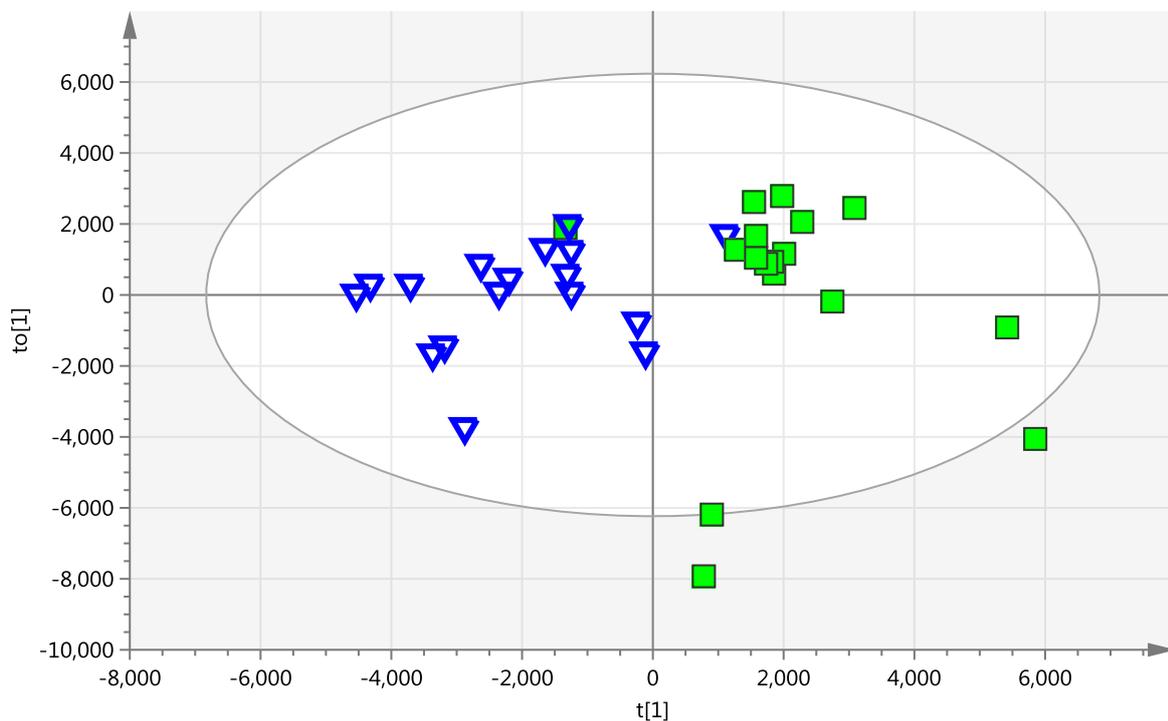


Figure 4: OPLS-DA scores plot obtained from adenocarcinoma (open inverted triangles ▼) and control (squares ■) samples. ($R^2X=0.302$, $R^2Y=0.642$, $Q^2=0.368$, $A=1+1+0$, $N=34$)

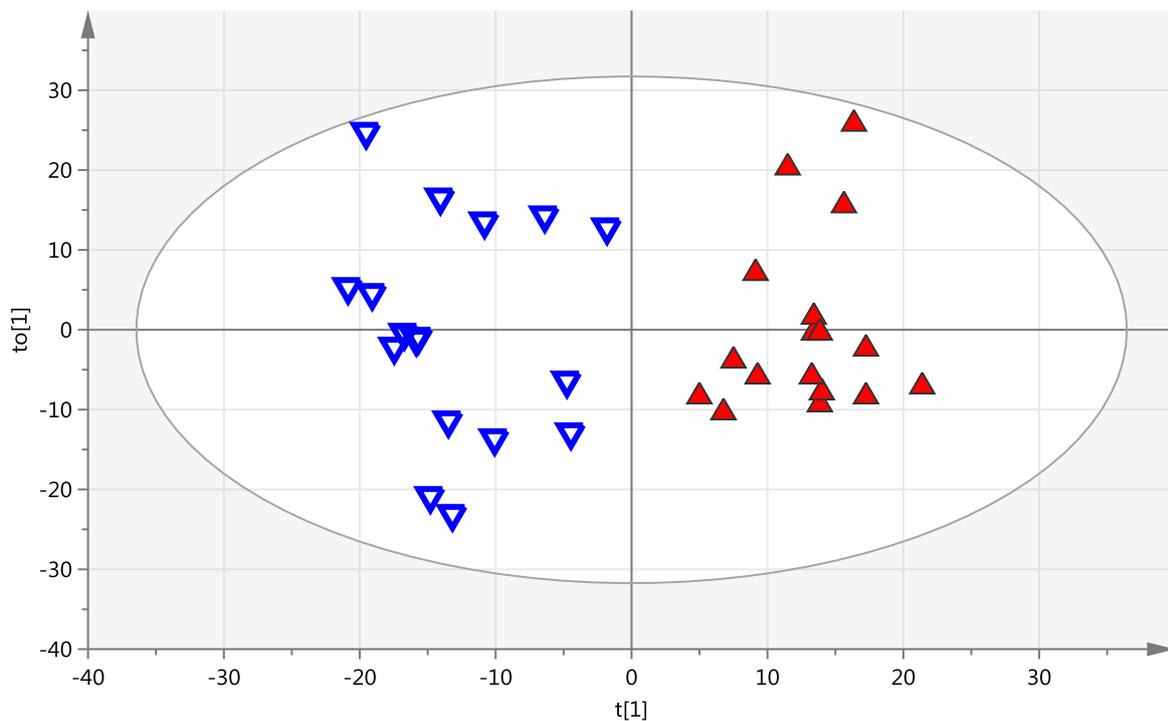


Figure 5: OPLS-DA scores plot obtained from adenocarcinoma (inverted open triangles ▼) and squamous carcinoma (triangles ▲) samples. ($R^2X=0.439$, $R^2Y= 0.874$, $Q^2=0.635$, $A=1+2+0$, $N=34$)

Table 1: Baseline demographics of the participants recruited to the study

	Squamous cell carcinoma N= 17	Adenocarcinoma N= 17	Controls N=17
Age in years, median (range)	69 (41 to 89)	75 (57 to 88)	70 (54 to 90)
Gender, males N (%)	13 (76)	10 (60)	9 (53)
Smoker or ex-smoker N (%)	16 (94)	17 (100)	13 (76)
Weight loss N (%)	11 (65)	6 (35)	4 (24)
Co-morbidity N (%)*	17 (100)	16 (94)	13 (76)
Number of medications, median (range)	4 (0-11)	4 (0-8)	2 (0-9)
Diagnosis of COPD N (%)	2 (12)	2 (12)	4 (24)
Diagnosis of hypertension N (%)	3 (18)	4 (24)	2 (12)
Diagnosis of diabetes mellitus N (%)	1 (6)	1 (6)	2 (12)

N= number of participants

% = Percentage of participants

Table 2: Biomarkers showing differences between squamous carcinoma patients and control subjects

Biomarker MW (Da)	Formula	Difference (mDa)	Lipid tentative identification	Adduct	Squamous carcinoma ↓↑
463.3491	C ₂₂ H ₄₂ NO ₇ P	0.67	Unknown	M+H	↑
878.7333	C ₅₇ H ₉₈ O ₆	0.25	TG(54:6)	M+H	↑
880.749	C ₅₇ H ₁₀₀ O ₆	0.24	TG(54:5)	M+H	↑
864.7206	C ₅₆ H ₉₆ O ₆	0.38	Unknown	M+NH ₄	↑
882.7634	C ₅₇ H ₁₀₂ O ₆	0.37	TG(54:4)	M+H	↑
326.0003	C ₁₇ H ₁₀ O ₇	0.53	Unknown	M+H	↑
884.7832	C ₅₇ H ₁₀₄ O ₆	0.33	TG(54:3)	M+Na	↑
327.9965	C ₁₀ H ₇ O ₈ N ₃ P	0.28	Unknown	M+H	↑
882.7523	-----	-----	Unknown	M+H	↑
464.3536	-----	-----	Unknown	M+H	↑
919.7595	C ₅₉ H ₉₈ O ₆	0.29	TG(58:8)	M+H	↑
757.5612	C ₄₂ H ₈₀ NO ₈ P	0.43	PC(34:2)	M+H	↓
785.5902	C ₄₄ H ₈₄ NO ₈ P	0.27	PC(36:2)	M+H	↓
523.3632	C ₂₆ H ₅₄ NO ₇ P	0.07	LysoPC(18:0)	M+H	↓
648.5845	C ₄₅ H ₇₆ O ₂	0.23	CE(18:2)	M+NH ₄	↓
814.6927	C ₄₇ H ₉₅ N ₂ O ₆ P	0.01	SM(d18:0/24:1)	M+H	↓
769.5621	C ₄₃ H ₈₀ NO ₈ P	0.40	PE(38:3)	M+NH ₄	↓
495.3312	C ₂₄ H ₅₀ NO ₇ P	0.08	LysoPC(16:0)	M+H	↓

Mass Difference (mDa) is the difference in exact mass between the measured value and the theoretical value

↑↓ - increase or decrease of biomarker compared with control group

PC phosphatidylcholine; DG diacylglycerol; TG triglyceride; SM sphingomyelin; PE phosphatidylethanolamine; LysoPC lysophosphatidylcholine; CE ceramide;

Table 3: Biomarkers showing differences between adenocarcinoma patients and control subjects

Biomarker MW (Da)	Formula	Difference (mDa)	Lipid tentative identification	Adduct	Adenocarcinoma ↓↑
758.5742	-----	-----	Unknown	M+H	↑
784.5892	-----	-----	Unknown	M+H	↑
789.5519	C ₄₂ H ₈₀ NO ₁₀ P	0.48	PS(36:1)	M+NH ₄	↑
860.6893	C ₅₆ H ₉₂ O ₆	0.49	Unknown	M+NH ₄	↑
791.5468	C ₄₅ H ₇₈ NO ₈ P	0.03	PE(40:6)	M-H	↑
310.2878	C ₂₀ H ₃₈ O ₂	0.10	Eicosenoic acid	M-H	↑
926.7442	-----	-----	Unknown	M+H	↑
338.3188	C ₂₂ H ₄₂ O ₂	0.205	Erucic acid	M-H	↑
976.7519	C ₆₅ H ₁₀₀ O ₆	0.18	TG(62:13)	M+NH ₄	↑
664.5066	C ₄₃ H ₆₈ O ₅	0.34	DG(40:8)	M+NH ₄	↑
960.8145	C ₆₃ H ₁₀₈ O ₆	0.14	TG(60:7)	M+NH ₄	↑
280.2408	C ₁₈ H ₃₂ O ₂	0.046	Linoelaidic acid	M-H	↑
793.5621	C ₄₅ H ₈₀ NO ₈ P	0.50	PE(40:5)	M+NH ₄	↓
751.5528	C ₄₃ H ₇₈ NO ₇ P	0.686	Plasmalogen(38:4)	M-H	↓
727.5528	C ₄₁ H ₇₈ NO ₇ P	0.686	Plasmalogen(36:2)	M-H	↓
692.5379	C ₄₅ H ₇₂ O ₅	0.38	DG(42:8)	M+Hac-H	↓
668.5379	C ₄₃ H ₇₂ O ₅	0.28	DG(40:6)	M+Hac-H	↓
765.5308	C ₄₃ H ₇₆ NO ₈ P	0.46	PE(38:5)	M+NH ₄	↓
749.5368	C ₄₃ H ₇₆ NO ₇ P	0.34	Plasmalogen(38:5)	M-H	↓
729.5688	C ₄₁ H ₈₀ NO ₇ P	0.11	Plasmalogen(36:1)	M-H	↓
481.3178	C ₂₃ H ₄₈ NO ₇ P	0.04	LysoPE(18:0/0:0)	M-H	↓

Mass Difference (mDa) is the difference in exact mass between the measured value and the theoretical value

↑↓ - increase or decrease of biomarker compared with control group

PS phosphatidylserine; DG diacylglycerol; TG triglyceride; PE

phosphatidylethanolamine; LysoPE lysophosphatidylethanolamine

Table 4: Biomarkers showing differences between adenocarcinoma and squamous patients

Biomarker MW (Da)	Formula	Difference (mDa)	Lipid tentative identification	Adduct	Adenocarcinoma	Squamous carcinoma
603.5292	-----	-----	Unknown	M+H	↓↑ ↑	↓↑ ↓
862.5571	C ₄₅ H ₈₃ O ₁₃ P	0.36	PI(36:2)	M+H	↑	↓
757.5622	C ₄₂ H ₈₀ NO ₈ P	1.45	PC(34:2)	M+Hac-H	↑	↓
743.5829	C ₄₂ H ₈₂ NO ₇ P	0.15	Plasmalogen(34:1)	M-H20-H	↑	↓
219.3668	-----	-----	Unknown	M-H	↑	↓
765.5978	-----	-----	Unknown	M-H	↑	↓
792.6267	C ₅₁ H ₈₄ O ₆	0.49	TG(48:7)	M-H	↑	↓
731.6067	C ₄₁ H ₈₄ N ₂ O ₆ P	0.32	SM(d18:1/18:0)	M+Hac-H	↑	↓
874.705	C ₅₇ H ₉₄ O ₆	0.024	TG(54:8)	M-H	↑	↓
717.5672	C ₄₀ H ₈₀ NO ₇ P	0.015	Plasmalogen(32:0)	M-H20-H	↑	↓
771.6142	C ₄₄ H ₈₆ NO ₇ P	1.22	Plasmalogen(36:1)	M-H20-H	↑	↓
703.5754	C ₃₉ H ₈₀ N ₂ O ₆ P	0.29	SM(d18:1/16:0)	M-H20-H	↓	↑
587.3678	-----	-----	Unknown	M-H	↓	↑
719.5465	C ₃₉ H ₇₈ NO ₈ P	0.27	PE(34:0)	M+H	↓	↑
819.5718	-----	-----	Unknown	M-H	↓	↑
444.3603	C ₂₉ H ₄₈ O ₃	0.21	4alpha-Carboxy-4beta-methyl-5alpha-cholesta-8-en-3beta-ol	M+Hac-H	↓	↑
666.5618	-----	-----	Unknown	M-H	↓	↑
809.6745	C ₄₈ H ₉₁ NO ₈	0.078	Glucosylceramide (d18:1/24:1)	M+NH ₄	↓	↑

Mass Difference (mDa) is the difference in exact mass between the measured value and the theoretical value

PI phosphatidylinositol; PC phosphatidylcholine; TG triglyceride; SM sphingomyelin; PE phosphatidylethanolamine;

Table 5: Prediction of lung cancer patients based on OPLS-DA multivariate models

PLS-DA model	Training set			Test set			Sensitivity (%)	Specificity (%)
	Squamous carcinoma	Adeno-carcinoma	Healthy volunteers	Squamous carcinoma	Adeno-carcinoma	Healthy volunteers		
Squamous carcinoma Vs. control	12	-----	11	5	-----	6	80	66.7
Adenocarcinoma Vs. Control	-----	12	11	-----	5	6	80	83.3
Adenocarcinoma Vs. Squamous carcinoma	12	11	-----	5	6	-----	80	83.3

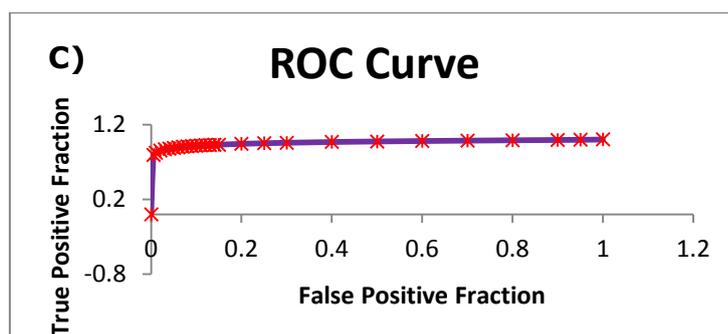
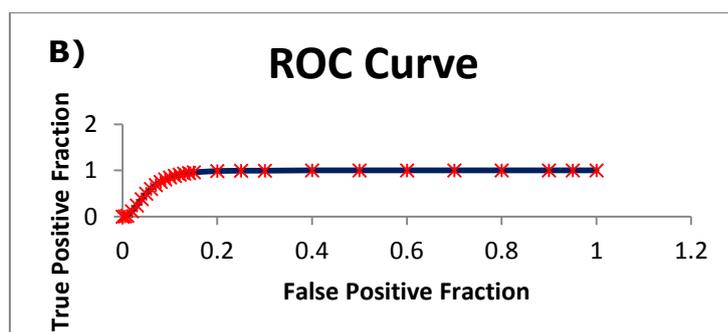
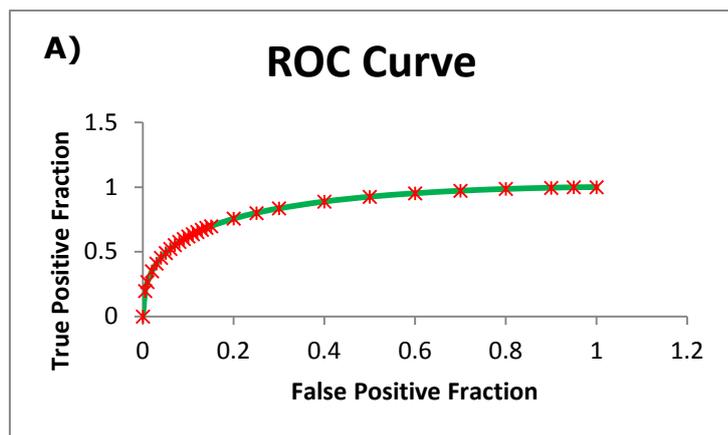


Figure 6: ROC curve is defined as true positive fraction versus false positive fraction. To affirm the validity of prediction OPLS-DA models of **(A)** squamous carcinoma vs control; **(B)** adenocarcinoma vs control; and **(C)** adenocarcinoma vs squamous carcinoma. Area under receiver operating characteristic (ROC) curves were calculated for each comparison. The area under the curve values were 0.86, 0.94 and 0.96 for squamous carcinoma vs control, adenocarcinoma vs control and adenocarcinoma vs squamous carcinoma (an ideal model would have an AUC of 1) which clearly states that the prediction models were robust.

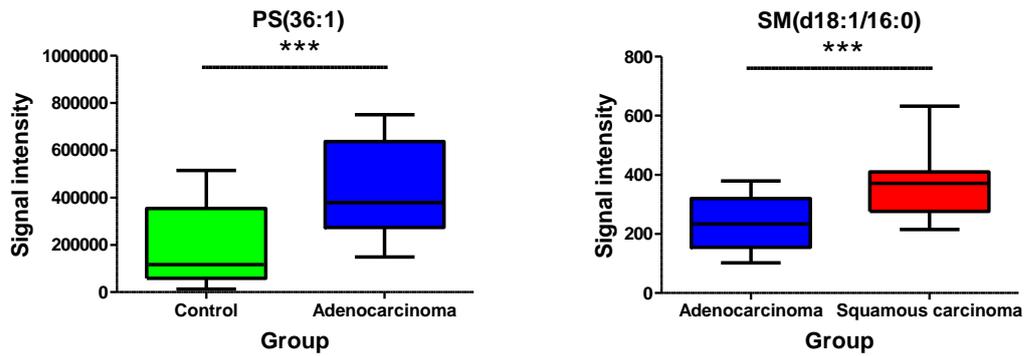


Figure 7: Box and whisker plots demonstrated altered levels of PS(36:1), plasmalogen (38:5) (adenocarcinoma vs control) and SM(d18:1/16:0) (adenocarcinoma vs squamous carcinoma) which were significant when compared groupings between control (n=17), squamous carcinoma (n=17) and adenocarcinoma (n=17). Statistical analysis performed using unpaired t-test, ***p<0.050 with Bonferroni correction..

SUPPLEMENTARY INFORMATION

Plasma lipid biomarker signatures in squamous carcinoma and adenocarcinoma lung cancer patients

Srinivasarao Ravipati, David A. Barrett, David R Baldwin⁺, Helen L Barr⁺⁺,
Andrew W. Fogarty⁺⁺⁺

Centre for Analytical Bioscience, School of Pharmacy, University of Nottingham, University Park, NG7 2RD

⁺ Department of Respiratory Medicine, David Evans Centre, City Hospital Campus, Nottingham, NG5 1PB.

⁺⁺ Division of Respiratory Medicine, Clinical Sciences Building, University of Nottingham, City Hospital Campus, Nottingham, NG5 1PB.

⁺⁺⁺ Division of Epidemiology and Public Health, Clinical Sciences Building, University of Nottingham, City Hospital Campus, Nottingham, NG5 1PB.

Correspondence and requests for reprints to Dr Andrew Fogarty:

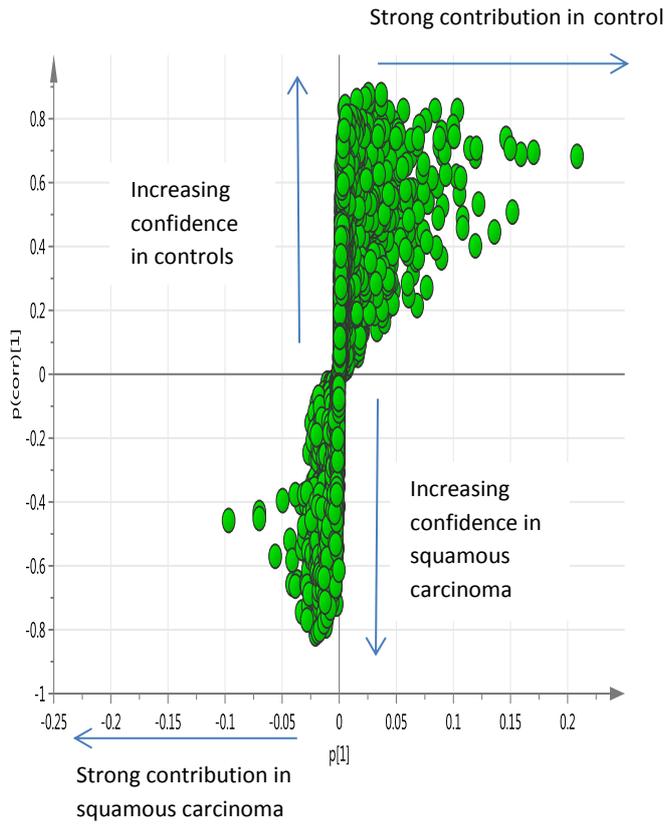
andrew.fogarty@nottingham.ac.uk

Division of Epidemiology and Public Health, University of Nottingham, Clinical Sciences Building, City Hospital, Nottingham NG5 1PB, UK

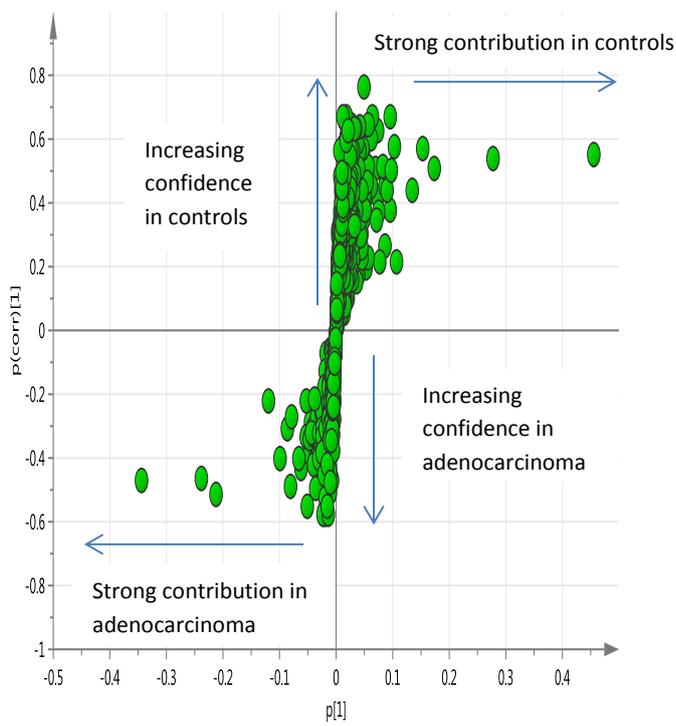
Tel 0044 115 8231715

Fax: 0044 115 8231946

(A)



(B)



(C)

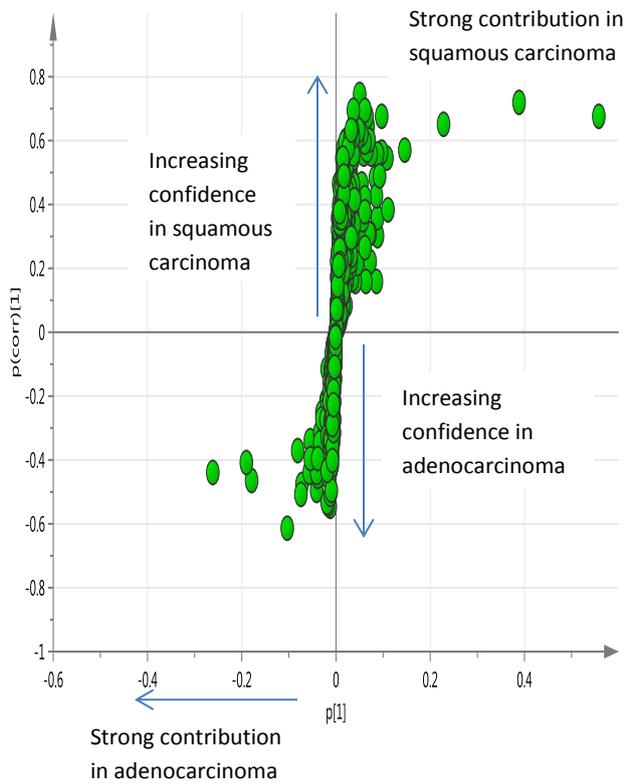


Figure S8: (A) S plot obtained from squamous carcinoma (n=17) and control (n=17) samples. ($R^2X=0.262$, $R^2Y=0.668$, $Q^2=0.181$, $A=1+1+0$, $N=34$) Figure S9: **(B)** S plot obtained from adenocarcinoma (n=17) and control (n=17) samples. ($R^2X=0.302$, $R^2Y=0.642$, $Q^2=0.368$, $A=1+1+0$, $N=34$); **(C)** S plot obtained from adenocarcinoma (n=17) and squamous carcinoma (n=17) samples. ($R^2X=0.670$, $R^2Y=0.419$, $Q^2=0.302$, $A=1+1+0$, $N=34$)

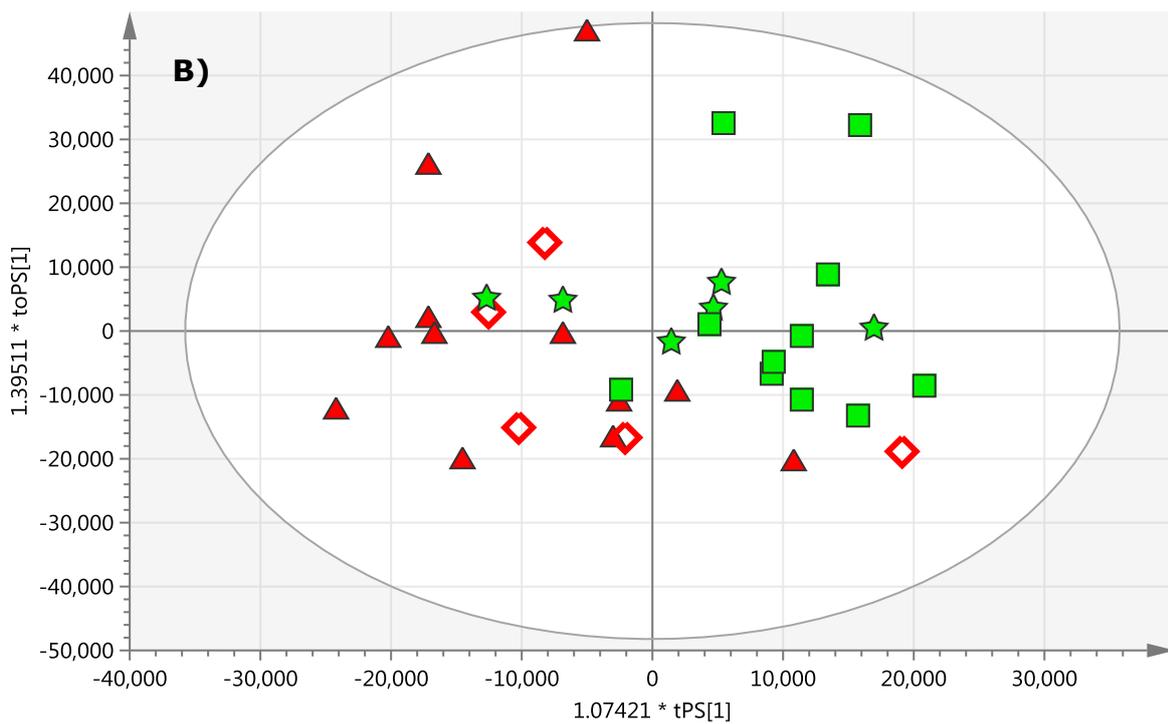
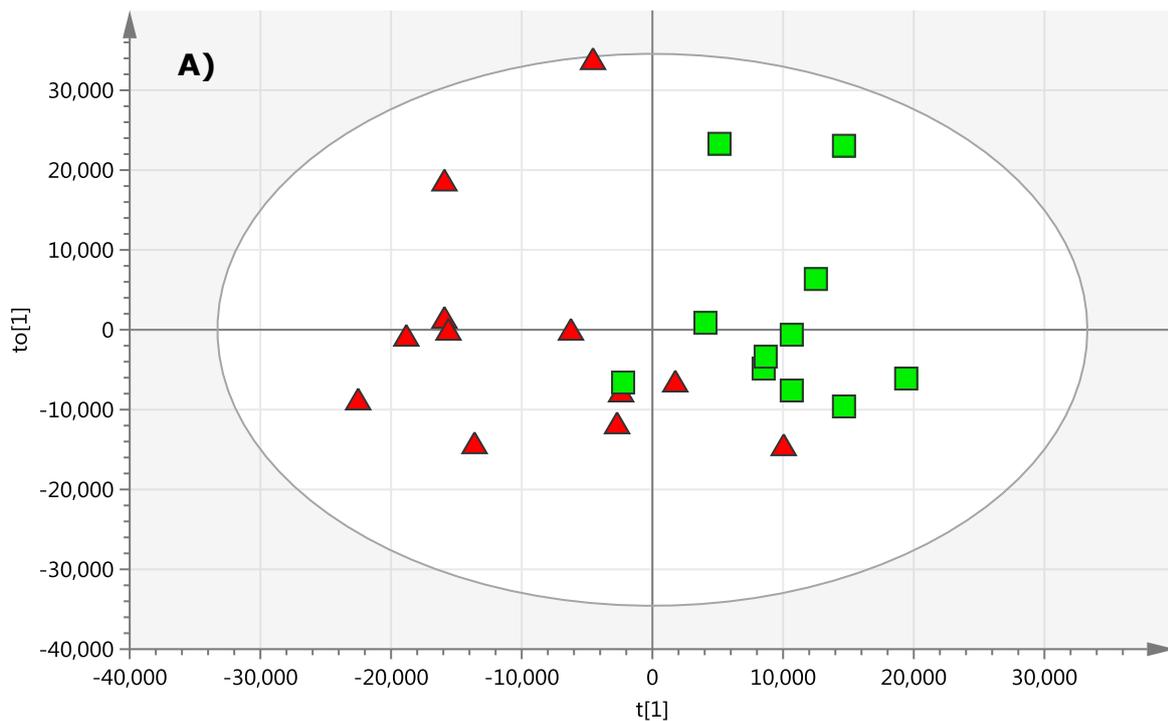


Figure S9: A) OPLS-DA scores plot (training set) obtained from squamous carcinoma (triangles \blacktriangle) and control (squares \blacksquare) samples. ($R^2X=0.342$, $R^2Y=0.705$, $Q^2=0.293$, $A=1+1+0$, $N=23$) **B)** OPLS-DA prediction of 11 additional subjects 5 squamous carcinoma patients (open diamonds \diamond) and 6 control samples (five point stars \star).

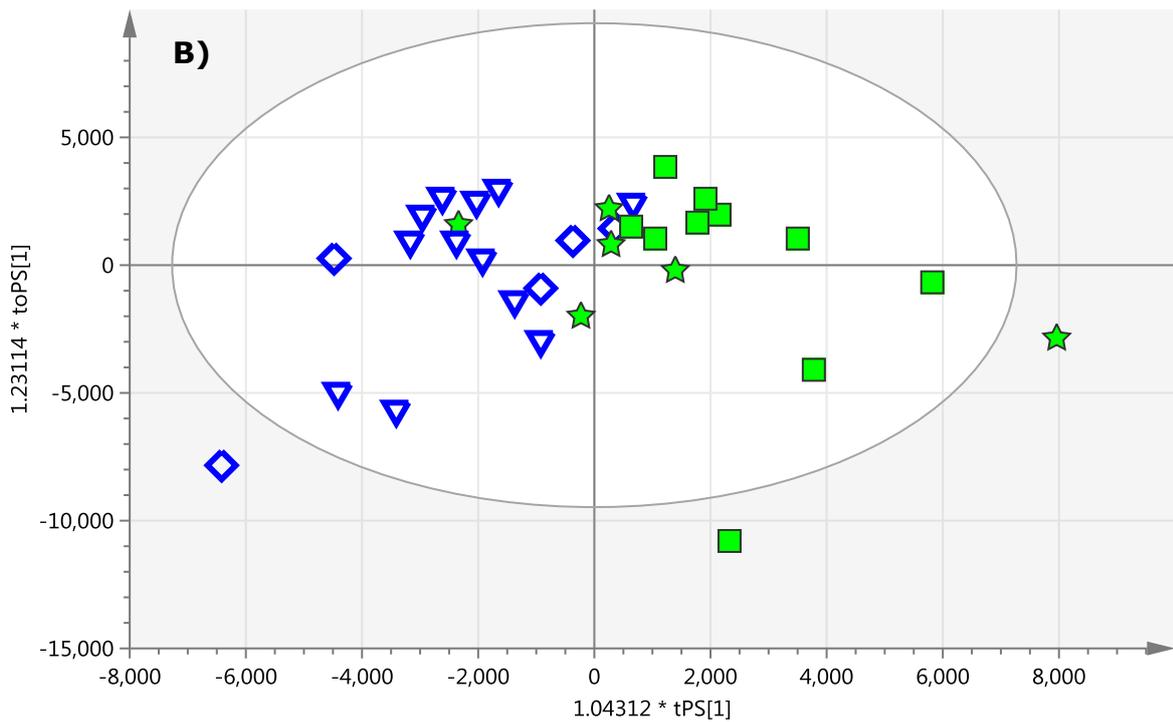
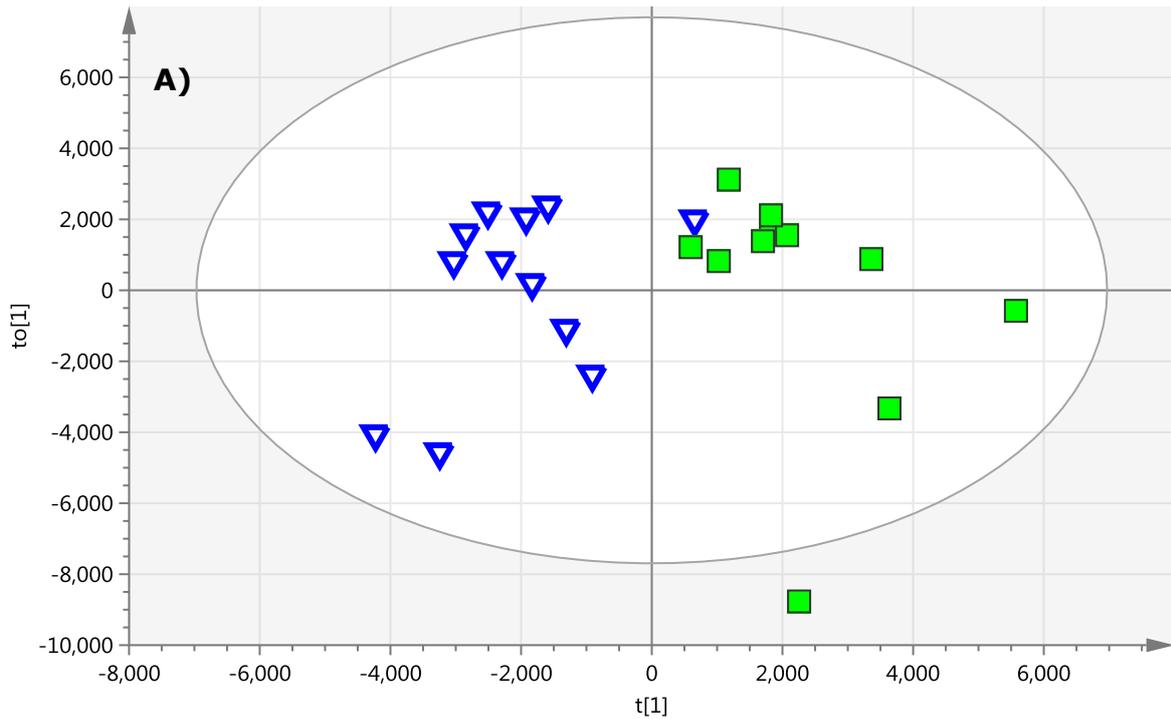


Figure S10: A) OPLS-DA scores plot (training set) obtained from adenocarcinoma (open inverted triangles ∇) and Control (squares \blacksquare) samples. ($R^2X=0.383$, $R^2Y= 0.743$, $Q^2=0.519$, $A=1+1+0$, $N=23$) **B)** OPLS-DA prediction of 11 additional subjects 5 adenocarcinoma (open diamond's \diamond) and 6 control samples (five point stars \star).

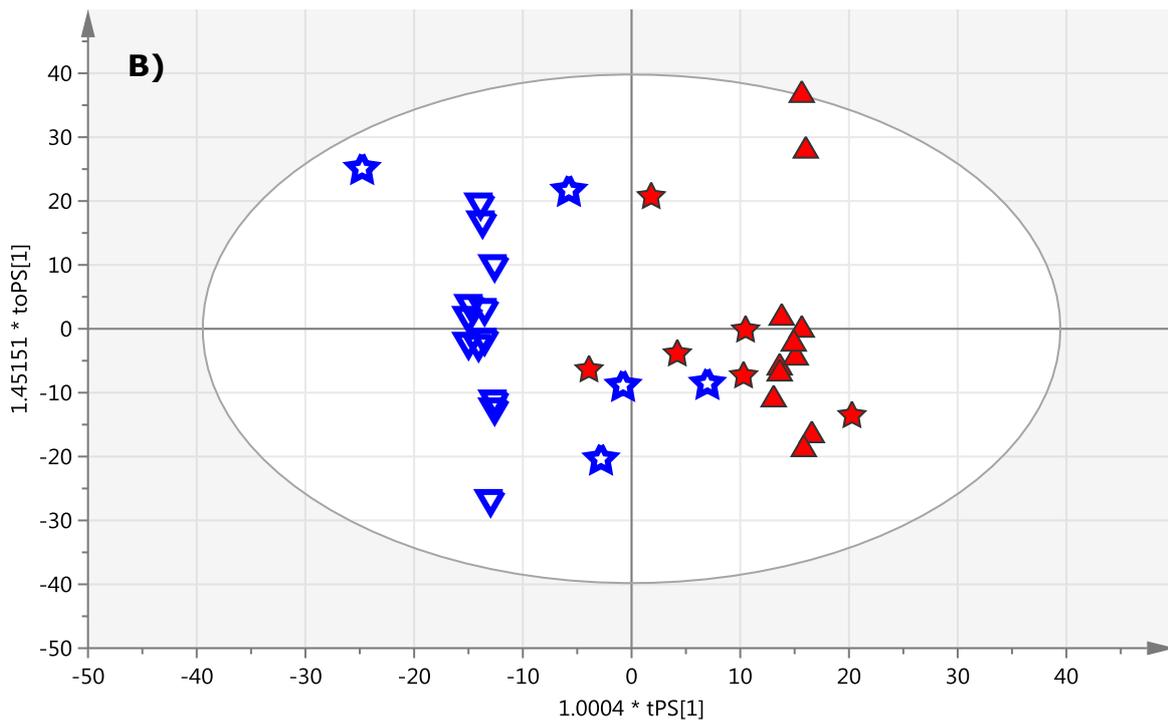
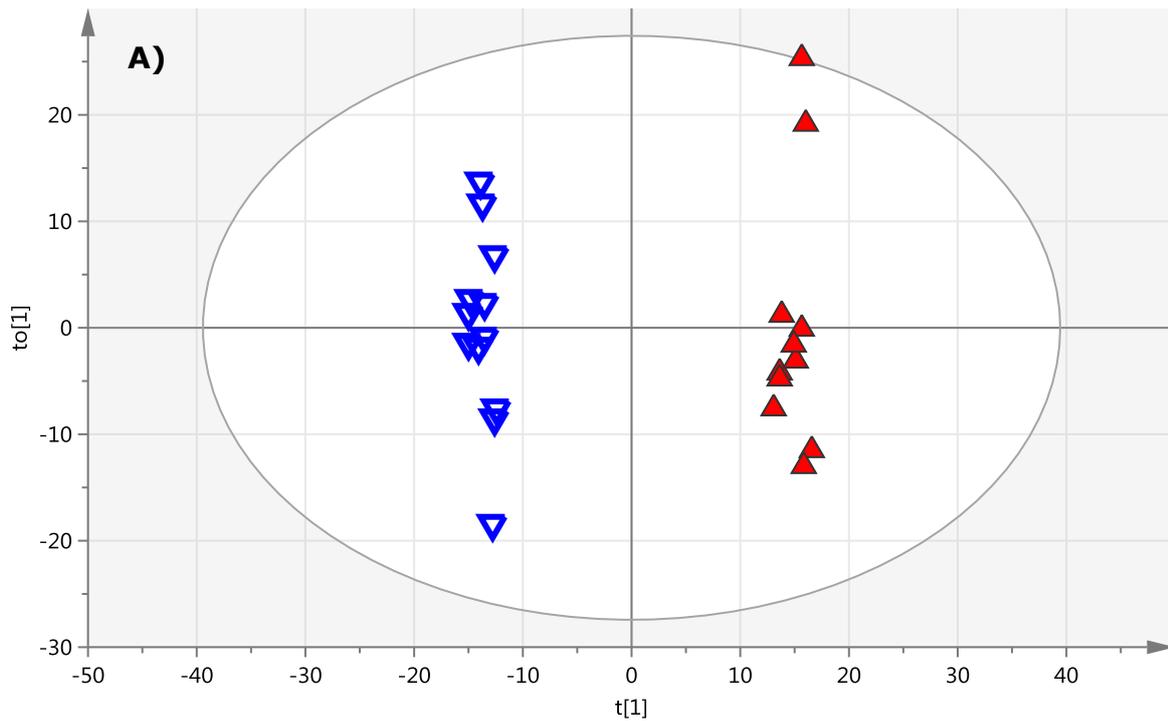


Figure S11: A) OPLS-DA scores plot (training set) obtained from randomly selected 33% adenocarcinoma (open inverted triangles ∇) and squamous carcinoma (\blacktriangle) patients. ($R^2X=0.612$, $R^2Y= 0.995$, $Q^2=0.699$, $A=1+4+0$, $N=23$) **B)** OPLS-DA prediction of 11 additional subjects 5 adenocarcinoma (open five point stars \star) and 6 squamous carcinoma samples (five point stars \star).

Table S6: Reproducible peak areas and retention times of different lipids observed in plasma QC samples which were analysed during the lung cancer biomarker study

Lipids	%RSD peak area	%RSD retention time	Mass deviation mDa
LysoPC(20:4)	8.95	0.47	0.43
α or γ Linolenic acid	9.61	0.30	0.85
Docohexaenoic acid	7.25	0.27	1.02
LysoPC(15:0)	10.23	0.32	0.38
LysoPC(16:0)	6.67	0.71	0.27
Arachidonic acid	7.59	0.12	0.63
LysoPC(18:1)	9.55	0.76	0.48
Palmitoleic acid	6.32	0.20	0.96
Linoleic acid	5.58	0.26	0.28
LysoPC(18:0)	9.23	0.21	0.63
LysoPE(22:6)	8.33	0.62	0.80
Oleic acid	6.62	0.17	0.36
13,14 Dihydro PGF1 α	8.82	0.23	0.57
Stearic acid	6.63	0.13	0.17
Eicosanoic acid	6.66	0.29	0.95
Ceramide (d18:1/16:0)	7.17	0.22	0.66
Heptadecanoic acid	8.67	0.19	0.83
Palmitic acid	6.82	0.31	0.99
Palmitoyl glycerol	7.89	0.16	1.21
HETEs	10.12	0.43	1.16
DHETs	8.13	0.29	1.05

LysoPC lysophosphatidylcholine; LysoPE lysophosphatidylethanolamine; PGF1 α prostaglandin F1 α ; PGF2 α prostaglandin F2 α ; HETE hydroxyeicosatetraenoic acid; DHET dihydroxyeicosatrienoic acid;