Title

Serum phosphatidylethanolamine levels distinguish benign from malignant solitary pulmonary nodules and represent a potential diagnostic biomarker for lung cancer.

Journal

Cancer biomarkers : section A of Disease markers, 16(4)

ISSN

1574-0153

Authors

Fahrmann, Johannes F
Grapov, Dmitry
DeFelice, Brian C
et al.

Publication Date

2016-03-01

DOI

10.3233/cbm-160602

Peer reviewed

Cancer Biomarkers 16 (2016) 609–617
DOI 10.3233/CBM-160602
IOS Press

Serum phosphatidylethanolamine levels distinguish benign from malignant solitary pulmonary nodules and represent a potential diagnostic biomarker for lung cancer

Johannes F. Fahrmann[a,1], Dmitry Grapov[b,1], Brian C. DeFelice[a], Sandra Taylor[c], Kyoungmi Kim[c], Karen Kelly[d], William R. Wikoff[a], Harvey Pass[e], William N. Rom[f], Oliver Fiehn[a,g] and Suzanne Miyamoto[d,*]

[a]*University of California, Davis Genome Center, Davis, California, CA, USA*
[b]*CDS Creative Data Solutions, Ballwin, MO, USA*
[c]*Division of Biostatistics, Department of Public Health Sciences, School of Medicine, University of California, Davis, California, CA, USA*
[d]*Division of Hematology and Oncology, Department of Internal Medicine, School of Medicine, University of California , Davis Medical Center, Sacramento, CA, USA*
[e]*Division of Thoracic Surgery, Department of Cardiothoracic Surgery, Langone Medical Center, New York University, New York, NY, USA*
[f]*Division of Pulmonary, Critical Care, and Sleep, NYU School of Medicine, New York, NY, USA*
[g]*Department of Biochemistry, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi-Arabia*

Abstract.
BACKGROUND: Recent computed tomography (CT) screening trials showed that it is effective for early detection of lung cancer, but were plagued by high false positive rates. Additional blood biomarker tests designed to complement CT screening and reduce false positive rates are highly desirable.
OBJECTIVE: Identify blood-based metabolite biomarkers for diagnosing lung cancer.
MEHTODS: Serum samples from subjects participating in a CT screening trial were analyzed using untargeted GC-TOFMS and HILIC-qTOFMS-based metabolomics. Samples were acquired prior to diagnosis (pre-diagnostic, $n = 17$), at-diagnosis ($n = 25$) and post-diagnosis ($n = 19$) of lung cancer and from subjects with benign nodules ($n = 29$).
RESULTS: Univariate analysis identified 40, 102 and 30 features which were significantly different between subjects with malignant (pre-, at- and post-diagnosis) solitary pulmonary nodules (SPNs) and benign SPNs, respectively. Ten metabolites were consistently different between subjects presenting malignant (pre- and at-diagnosis) or benign SPNs. Three of these 10 metabolites were phosphatidylethanolamines (PE) suggesting alterations in lipid metabolism. Accuracies of 77%, 83% and 78% in the pre-diagnosis group and 69%, 71% and 67% in the at-diagnosis group were determined for PE(34:2), PE(36:2) and PE(38:4), respectively.
CONCLUSIONS: This study demonstrates evidence of early metabolic alterations that can possibly distinguish malignant from benign SPNs. Further studies in larger pools of samples are warranted.

Keywords: Lung cancer, metabolomics, phospholipids, biomarkers, solitary pulmonary nodules

[1]These authors contributed equally to this work.

*Corresponding author: Suzanne Miyamoto, Division of Hematology/Oncology, UC Davis Cancer Center, 4501 X Street, Room 3016, Sacramento, CA 95817, USA. Tel.: +1 916 734 3769; Fax: +1 916 734 5415; E-mail: smiyamoto@ucdavis.edu.

ISSN 1574-0153/16/$35.00

1. Introduction

Lung cancer continues to be a leading cause of cancer mortality in both men and women in the United States with 221,200 new cases and 158,040 deaths estimated in 2015 [1]. Since the relative 5-year survival rate is considerably better when lung cancer is diagnosed in the early stage (16%–49% survival) compared with metastatic late stage disease (∼2% survival) [2], early diagnosis is extremely important to pursue. To-date, the most successful effort for early detection of lung cancer is the National Lung Cancer Screening Trial (NLST) study that demonstrated value of low-dose computed tomography (LDCT) scan screening in reducing mortality by 20% compared with chest radiograph screening [3]. An NYU-conducted LDCT scan screening study found that increased age and emphysema are risk factors for the presence of non-calcified nodules (NCNs) on CT-scans. Having additional risk factors of decreased FEV1 and FVC, smoking history, and the presence of multiple sub-solid nodules were the strongest predictors that an individual with NCNs could be diagnosed with lung cancer [4]. Despite the success of LDCT screening in the early diagnosis of lung cancer, a major concern was the high prevalence (96%) of false positives [3]. Contributing to the problem is that there is no consensus on how to manage these solitary pulmonary nodules (SPNs). The concern with screening, then, is over-treatment, anxiety induction and excessive use of invasive procedures. There is a critical need for additional tests that can better stratify the SPNs found with LDCT into malignant and non-malignant. Such a complementary test should preferably be non-invasive and exhibit high sensitivity and specificity.

The use of analytical methods targeting blood analysis has generated considerable interest in biomarker discover for lung and other cancers, especially comprehensive molecular analyses (RNAseq, proteomic, glycomic, metabolomic) because of their precise, quantitative measurements [5]. Metabolomic analysis measures the end products of all cellular, tissue and organ activities in the body [6]. Metabolomics has been applied to gain new insights into the pathology of cancer, develop methods predictive of disease onset and reveal new biomarkers associated with diagnosis and prognosis [6–8]. Using a metabolomics approach, we previously demonstrated the value of metabolite-derived classifiers for the early detection of non-small cell lung adenocarcinoma [9,10].

In the current study, we expand upon our initial experimental findings as part of the discovery phase by evaluating metabolites in serum from subjects with benign or malignant SPNs using a combined approach of gas chromatography time-of-flight mass spectrometry (GC-TOFMS) and hydrophilic liquid chromatography accurate mass quadrupole time-of-flight mass spectrometry (HILIC-qTOFMS). Furthermore, we evaluated serum collected pre-diagnosis and at-diagnosis of lung cancer in addition to samples obtained post-surgical intervention from subjects with malignant SPNs (post-diagnosis). We hypothesize that our systems biology approach to identify candidate metabolomics biomarkers will ultimately lead to improved early detection of lung cancer and can be used in as a companion blood test to LDCT screening.

2. Methods

2.1. Patient population and collection of patient samples

The study subjects were recruited from the NYU Lung Cancer Biomarker Center and all gave informed consent for the IRB approved NYU Lung Cancer Biomarker Center Protocol #8896. The NYU Lung Cancer Biomarker Center performs low-dose CT-scan screening for high-risk smokers as part of the National Cancer Institute's Early Detection Research Network Program (EDRN). Lung cancer cases for this study were confirmed by pathology (Table 1). Patients with benign nodules were those with stable nodules or ground glass opacity (GGO) over at least two year period with annual CT scans performed. Blood samples were collected from all patients enrolled in the protocol before diagnosis (>6 months prior to surgery, pre-diagnostic). Additional blood samples were collected from some of the lung cancer patients, at-diagnosis (at-diagnosis) and post-treatment (after surgery). All but one patient had their "at-diagnosis" sample collected before surgery, usually within days prior to surgery, likely on their pre-operation visit. Surgery was usually performed within 1–3 months of diagnosis. Diagnosis was made by biopsy and/or surgery with all tissue diagnoses confirmed by pathology. Post-diagnosis samples were collected at least one month post-surgery. Selection of cases was performed by NYU and the blinded serum samples were sent to University of California, Davis Medical Center (UCDMC) for analysis. None of the study subjects had previous cancer or chemotherapy. All subjects had blood drawn by EDRN protocol, performed spirometry according to ATS guide-

Table 1
Patient characteristics

	Benign	Cancer (Pre-diagnosis)	Cancer (At-diagnosis)	Cancer (Post-diagnosis)
Subjects, N	29	17	25	19
By histology type				
Adenocarcinoma, N (%)	–	14* (82)	21* (84)	17 (89)
By Stage, N (%)				
Stage I	–	12 (86)	19 (90)	16 (94)
Stage II	–	1 (7)	1 (5)	1 (6)
Stage IV	–	1 (7)	1 (5)	0 (0)
Squamous Cell Carcinoma, N (%)	–	1 (6)	2 (8)	2 (11)
By Stage, N (%)				
Stage I	–	1 (100)	1 (50)	2 (100)
Stage III		0 (0)	1 (50)	0 (0)
Small Cell, N (%)	–	3* (18)	3* (12)	0 (0)
By Stage, N (%)				
Stage I	–	2 (67)	2 (67)	0 (0)
Stage IV	–	1 (33)	1 (33)	0 (0)
Gender, N (Males/Females)	10/19	7/10	8/17	6/13
Age, Mean ± SD	67 ± 6	68 ± 7	67 ± 7	64 ± 5
BMI, Mean ± SD	28 ± 5	25 ± 3**	28 ± 5	28 ± 6
Packs per year, Mean ± SD	53 ± 20	57 ± 18	54 ± 17	56 ± 23
Emphysema/COPD, N (%)	21 (72)	11 (65)	14 (56)	13 (68)

*One subject has both non-small cell adenocarcinoma and small cell lung cancer; **Significantly different (p: < 0.05) from subjects with benign solitary pulmonary nodules.

lines, and answered questionnaires with smoking and occupational history. Blood samples were processed within two hours and then stored at −80°C. Matching of subjects with malignant SPNs with those having benign SPNs was based on age, gender, smoking in pack-years, and date of sample collection. Samples were thawed twice for aliquoting (once at the NYU and once at UCDMC) prior to metabolomic analysis.

2.2. Metabolomic profiling

The MiniX database [11] was used as a Laboratory Information Management System (LIMS) and for sample randomization prior to all analytical procedures. Sample identifications were kept blinded during the entire metabolomics analysis to minimize potential bias.

Detailed information on sample preparation, instrument parameters and data acquisition are provided in *Supplemental Methods*.

2.2.1. GC-TOFMS analysis

Serum samples (30 μL) were thawed, extracted and derivatized as previously described [9]. Mass spectrometry analysis and data acquisition was performed using an Agilent 7890A gas chromatograph coupled to a Leco Pegasus IV time-of-flight (TOF) spectrometer. This method is specifically useful for primary metabolites including sugars, amino acids, and hydroxyl acids. Acquired spectra were further processed using the BinBase database [11,12].

2.2.2. Hydrophilic interaction (HILIC)-qTOFMSMS analysis

Serum samples (30 μL) were thawed on ice and metabolites extracted using ice cold 3:1 methanol: water. Supernatant containing extracted metabolites were dried to completeness using a Labconco Centrivap, resuspended in 30 μL of 4:1 acetonitrile:water and submitted to analysis by Hydrophilic Interaction Liquid Chromatography Accurate Mass Quadrupole Time-Of-Flight Mass Spectrometry (HILIC-qTOFMS) which is specifically useful for detecting biogenic amines but will also detect and separate complex lipids.

Liquid chromatography-mass spectrometry (LC-MS) was performed using an Agilent 6530 Accurate Mass Quadrupole Time-of-Flight (QTOF) with a Jet-Stream ion source in positive ionization mode was coupled with an Agilent 1290 Series UHPLC. Data were collected at a rate of 2 spectra per second. Chromatographic separation was performed using a Waters Acquity BEH HILIC 2.1 × 150 mm column, particle size 1.7 μm. Mobile phase A was composed of water with added modifiers; 4 mM acetic acid and 6 mM ammonium acetate; mobile phase B was 9:1 acetonitrile: water using the same modifiers as mobile phase A. The mobile phase gradient started from 0 min 100% with a flow-rate of 0.4 mL/min (B), 0–14 min 70% (B), 14–14.2 min 45% (B), and 14.2–17.1 min 45% (B) with the flow-rate adjusted to 0.45 mL/min, 17.1–20 min

100% (B) and 20–25 min 100% (B) with the flow-rate returned to 0.4 mL/min. The injection volume was 5 μL with column temperature held constant at 45°C.

Instrument data was converted to MZdata format using Agilent's MassHunter software and processed using MZmine 2 [13]. Identification of features were conducted using NISTMS.exe comparing acquired MSMS spectra collected from pooled serum samples against MSMS spectral libraries including Lipidblast [14], Metlin, Massbank, HMDB and Lipidmaps.

2.3. Data analysis

2.3.1. Statistical analyses

Statistical analyses were carried out on gender + age + BMI + (smoking) packs per year covariate-adjusted metabolite values. A linear model was used to describe differences in serum metabolite abundances due to gender + age + BMI + packs per year; the residuals from which were tested for differences between benign and malignant (pre-diagnosis, at-diagnosis and post-diagnosis) SNPs. Covariate-adjustment was conducted in R [15]. Covariate-adjusted data was $\log_{10}$ transformed and significance was determined using a Mann-Whitney U test. The significance levels (i.e. p-values) were adjusted for multiple hypothesis testing according to Benjamini and Hochberg [16] at a false discovery rate (FDR) of < 0.05. All statistical analyses were conducted using DeviumWeb [17].

Classical Receiver Operating Characteristic (ROC) Curves were generated using Metaboanalyst v3.0 [18]. pROC::roc was used to calculate the ROC curve [19]. Subsequently, the optimal threshold and statistics were extracted based on the Younden method pROC:: cords [19].

Figures and metabolic trajectories were generated in GraphPad Prism v5.0.

3. Results

3.1. Subject characteristics

Patient characteristics for the respective cohorts are provided in Table 1. The cohorts consisted of serum samples collected from 29 subjects who presented SPNs but were found to be benign upon follow-up for 2 years; serum samples acquired from 17 subjects prior to diagnosis of lung cancer (>6 months, pre-diagnosis), serum samples acquired from 25 subjects when they were diagnosed with lung cancer (at-diagnosis) and serum samples from 19 subjects with lung cancer following surgical removal of malignant nodules (>1 month, post-diagnosis). For 10 patients, samples were collected at three time-points (pre-diagnostic, at-diagnosis and post-diagnosis), while for 13 patients, samples were only available at two time-points and for 8 subjects, only one time-point serum was collected. Although these samples were collected within a larger prospective cohort study, data could not be analyzed in a paired- or longitudinal-manner because not all cases had all three samples collected (pre-diagnosic, at-diagnosis and post-diagnosis). Instead, we used these samples for cross-sectional comparisons. Non-small cell lung adenocarcinoma was the most common histological type (82%, 84%, and 89% for pre-diagnosis, at-diagnosis and post-diagnosis, respectively) in our study (Table 1). The cancer population was predominately composed of those subjects with early stage (Stage IA/1B) cancer (Table 1). There were no significant differences in the matching variables of age, gender, and smoking packs per year between the respective groups; however BMI was significantly lower in the pre-diagnosis group relative to the benign cohort (Table 1).

3.2. Identification of metabolites: Distinguishing malignant SNPs from benign SNPs in serum

Untargeted GC-TOFMS and UPHLC-qTOFMS based-metabolomics was conducted on each sample. A total of 1332 features were detected between the two platform-types of which 186 had known annotated structures. Univariate analysis identified 40, 102 and 30 serum metabolites which were significantly (raw p-value < 0.05) different between subjects with benign SPNs and the pre-diagnosis, at-diagnosis or post-diagnosis cohort, respectively (Supplemental Table S1 and Supplemental Fig. S2). No metabolites were found to be significantly different following FDR-adjustment. Of the significantly (raw p-value < 0.05) different metabolites, 10 were found to be consistently increased or decreased in the pre-diagnosis and at-diagnosis group relative to the benign cohort (Fig. 1). Three of these 10 features were identified as phosphatidylethanolamines (PE34:2, PE36:2 and PE38:4). All three PEs were elevated in the cancer group relative to patients with benign tissue diagnoses (Fig. 1). In general, measured PEs tended to be elevated in subjects with cancer relative to subjects with benign SPNs (Supplemental Table S1). The other 7 peaks (repre-

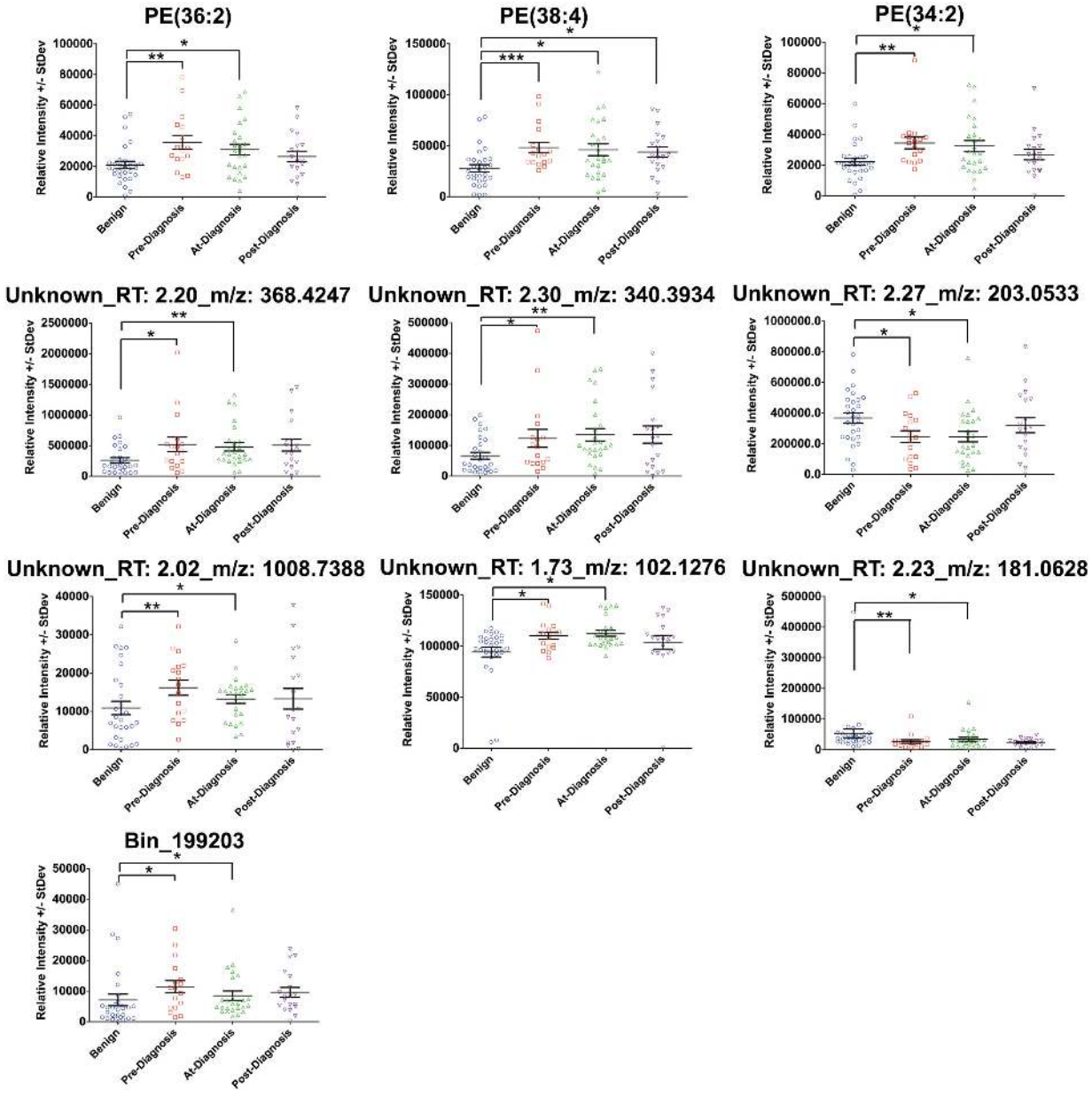

Fig. 1. Distribution of the 10 Serum Metabolites Significantly Different between Subjects with Malignant (Pre- and At-Diagnosis) or Benign Solitary Pulmonary Nodules. Relative Intensities $\pm$ StDev for the 10 circulating metabolites that were significantly increased or decreased between subjects with malignant (pre- and at-diagnosis) or benign solitary pulmonary nodules are shown for benign, pre-diagnosis, at-diagnosis and post–diagnosis. *: p-value < 0.05; **: p-value < 0.01. (Colours are visible in the online version of the article; http://dx.doi. org/10.3233/CBM-160602)

senting different compounds) were composed of one unknown compound from the GC-TOFMS analysis and 6 unknown compounds from the HILIC-qTOFMS analysis. The mass spectrum for Bin#199203 is provided in Supplemental Fig. S3. Based on characteristic masses and its retention time, this compound belongs to the group of carbohydrate derivatives. It matches a spectrum in the NIST14 mass spectral library, 1-methylgalactose, as well as its predicted retention time. Many of these metabolites were no longer significantly different in the post-diagnosis group compared to the benign group (Fig. 1). Individual metabolite trajectories for those subjects that had all three time periods collected ($n = 9$) are shown in Supplemental Fig. S1.

Table 2
Classification performances of serum metabolites significantly different between subjects with malignant (pre-diagnosis and at-diagnosis) or benign solitary pulmonary nodules

	Pre-diagnosis				At-diagnosis			
Variable	AUC	Specificity	Sensitivity	Accuracy	AUC	Specificity	Sensitivity	Accuracy
PE(34:2)	0.80	66	88	77	0.67	83	56	69
PE(36:2)	0.77	83	82	83	0.67	83	60	71
PE(38:4)	0.79	62	94	78	0.71	83	52	67
Bin_199203	0.72	76	71	73	0.67	45	92	68
Unknown_RT: 1.73_m/z: 102.1276	0.71	86	53	70	0.75	52	96	74
Unknown_RT: 2.02_m/z: 1008.7388	0.69	48	88	68	0.64	66	68	67
Unknown_RT: 2.2_m/z: 368.4247	0.70	66	77	71	0.74	66	88	77
Unknown_RT: 2.23_m/z: 181.0628	0.77	66	82	74	0.66	62	72	67
Unknown_RT: 2.27_m/z: 203.0533	0.68	45	88	67	0.71	86	52	69
Unknown_RT: 2.3_m/z: 340.3934	0.69	55	88	72	0.73	66	84	75

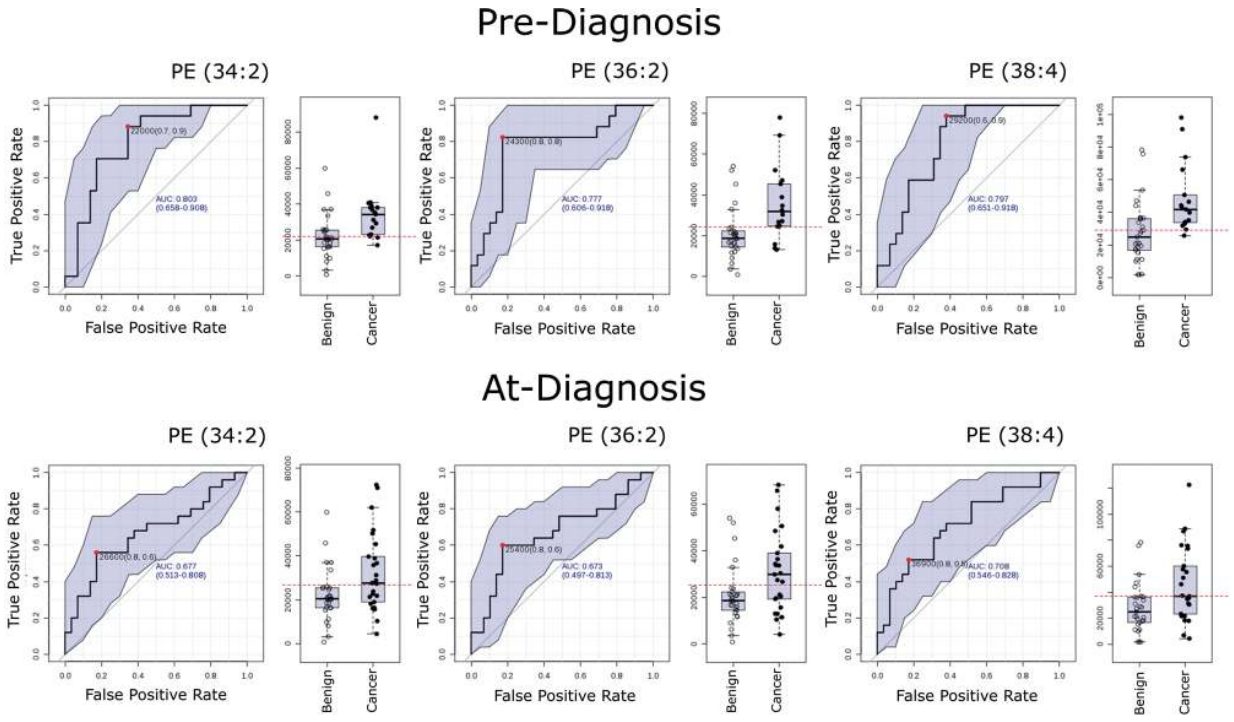

Fig. 2. ROC Curves for PE(34:2), PE(36:2) and PE(38:4) in the Pre- and At-Diagnosis Cohorts. Receiver Operating Characteristic Curves (ROC) plus the 95% Confidence Intervals are shown for PE(34:2), PE(36:2) and PE(38:4) in the pre-diagnosis and at-diagnosis cohorts, respectively. (Colours are visible in the online version of the article; http://dx.doi.org/10.3233/CBM-160602)

3.3. Evaluation of diagnostic capacity of conserved metabolic alterations

To explore the diagnostic capacity of the 10 conserved metabolites that distinguished subjects with malignant SNPs (pre-diagnosis and at-diagnosis) from subjects with benign SNPs, we evaluated their classification performance as single-metabolite classifiers. We acknowledge the small population in this study and, as such, we evaluated the entire set of subjects (i.e. training set) per cancer group (pre-diagnosis and at-diagnosis) relative to the benign group. Performance metrics for the individual metabolite-classifiers are provided in Table 2. Individually, PE(34:2), PE(36:2) and PE(38:4) had the best accuracies in the pre-diagnostic group (77%, 83% and 78%, respectively); however, these lipid species showed modest performance in the at-diagnosis group (accuracy = 69%, 71% and 67%, respectively) (Table 2). The accuracy of the unknown compounds in the pre-diagnostic samples ranged from 67% to 74% (Table 2). For the at-diagnosis samples, the unknown compounds had accu-

racies that ranged from 67 to 75% (Table 2).

Individual ROC curves with 95% C.I. for the three known compounds, PE(34:2), PE(36:2) and PE(38:4), which are the top performing candidates in this study, are provided in Fig. 2. Despite the small sample size these findings provide preliminary evidence that circulating metabolites, particularly PEs, have diagnostic value.

4. Discussion

Lung cancer continues to be a leading cause of cancer mortality in the United States. Although LDCT screening has drastically aided in early detection of lung cancer, it is hindered by high false positive rates. Consequently, determining whether a SNP is malignant upon presentation and deciding the appropriate course of management remains challenging and difficult.

There is considerable clinical need to find complementary screening methodologies that can distinguish between a benign and malignant SNP. In the current study, as part of the discovery phase, a multiplatform metabolomics approach was utilized to discover metabolic perturbations in serum that distinguish subjects with benign SNPs versus those with malignant SNPs. Metabolites are sensitive to physical characteristics such as gender, age and BMI [20,21] and environmental stimuli (smoking) [22]. Consequently, we evaluated metabolite abundances adjusted for the mentioned above covariates to minimize potential bias. Despite the small sample size, these findings provide preliminary evidence of serum metabolic features with diagnostic potential that will be expanded upon in future studies.

Compared to subjects with benign SPNs, subjects that presented malignant SPNs indicated elevated levels of serum phosphatidylethanolamines, particularly, PE34:2, PE36:2 and PE38:4. Moreover, elevation in these specific PEs were maintained pre-diagnosis and at-diagnosis of lung cancer and reasonably classified SPNs as malignant with accuracies 67–83%. Notably, classification performances of the mentioned above PEs were better in the pre-diagnosis group compared to the at-diagnosis group. While it cannot be ruled out that this may be due to differences in sample sizes, circulating abundances of phospholipids have been shown to decrease with tumor progression [23]. This suggests that early aberrations in lipid metabolism accompany tumorigenesis and that these behaviors may alter post transformation. Indeed, NSCLC tumors exhibit drastic changes in lipid profiles, including elevations in PEs, compared to matched control tissue [24,25]. Consistent with our findings, Gao et al. also found that PEs, including PE(38:4), tended to be elevated in plasma from subjects with malignant nodules compared to those with benign nodules [26]. Huang et al. previously illustrated that A549 lung adenocarcinoma cells increase secretion of phosphatidylethanolamine binding protein (PEBP) [27]. PEBP is overexpressed in lung cancer and has been shown to modulate development, invasion, metastatic potential of tumors [28,29]. Thus, the elevation in PEs may, in part, act as agonists of PEBP-mediated signaling transduction. Interestingly, PE38:4, provisionally assigned as PE (18:0_20:4), contains an arachidonic acid side chain. Free fatty acids arachidonic acid and linoleic acid and their derived-eicosanoids have been shown to be elevated in serum from subjects with NSCLC adenocarcinoma relative to control [30]. Arachidonic acid serves as a central precursor for a variety of pro-inflammatory and immune modulatory lipid signaling mediators [31]. We suspect that increased PEs are likely related to an inflammatory response rather than a specific marker of lung cancer. It is interesting to note that, with the exception of PE(38:4), levels of PE(36:2) and PE(34:2) declined following surgical removal of malignant nodules. Moreover, when evaluating those subjects which had all three time-points (pre-diagnosis, at-diagnosis and post-diagnosis), this general trend was conserved. Serum levels of PEs present promising diagnostic markers for distinguishing between malignant and benign SPNs that will be further explored in future studies.

The other seven top candidates were unknown compounds. One was detected by GC-TOFMS and the other 6 were found by HILIC-qTOF MS. Due to sample limitations and noisy MSMS spectra, identification of these unknown compounds was hampered. While additional analyses will be needed to identify these unknown compounds, it is interesting to note that the unknown primary metabolism compound best matched the mass spectrum of a methylated hexose. Both hexose metabolism and methylation (one-carbon) metabolism is known to be intricately involved in the etiology of tumors. If these compounds are identified and validated in further lung nodule studies, the identity of these compounds can be investigated through additional analysis.

Despite changes in the mentioned above metabolites, we did not observe differences in metabolites

which we had previously identified as metabolite (single or multiplex)-classifiers for diagnosis of non-small cell lung adenocarcinoma, including diacetylspermine [9,10]. However, it should be mentioned that the population size in the current study is small with repeated measurements for select subjects and contained mixed pathologies, whereas our former studies were exclusively focused on NSCLC adenocarcinoma.

A review on imaging of solitary pulmonary nodule by Sim and Poon [32] describes certain limitations of using follow-up imaging for SPNs, since there are still wide differential diagnoses of benign and malignant SPNs by current conventional imaging methods. Although ^{18}F FDG PET imaging combined with CT improves the diagnostic accuracy of imaging, and new CT imaging methods (dual-energy CT, perfusion CT, magnetic resonance (MR) imaging, dynamic contrast enhancement or diffusion-weighted imaging) are being developed, Sim and Poon stress that tissue diagnosis by resection or percutaneous biopsy of SPN still needs to be performed in patients determined by clinical stratification to be at moderate or high risk of malignancy.

In conclusion, our preliminary findings highlight the application of a metabolomics to identify serum-based metabolites that can distinguish malignant and benign SNPs as part of the discovery phase. Specifically, we hypothesize that elevations in serum phosphatidylethanolamines may serve as potential diagnostic markers that will be explored in conjunction with previously identified metabolite-classifiers in larger future studies.

Acknowledgments

We acknowledge the efforts of Ting-An Yie MS, who processed all of the samples, Ellen Eylers RN, who was the study nurse, Katrina Bantis and other members of the Rom laboratory for assistance with selection of patient samples, patient data information, aliquoting and transportation of samples to UCDMC and follow-up information and Carol Stroble (Miyamoto lab) for assistance with the sample database and handling and distribution of the clinical samples.

Financial support

We acknowledge the following grants: NCI UO1 CA 086137 NYU Lung Cancer Biomarker Center (Rom), LUNGevity Foundation and The Thomas G. Labrecque Foundation 201118739 (Miyamoto), Tobacco Related Disease Research Program 20PT0034 (Kelly) and NIH West Coast Metabolomics Resource Core U24 DK097154 (Fiehn).

Supplementary data

The supplementary files are available to download from http://dx.doi.org/10.3233/CBM-160602.

References

[1] R.L. Siegel, K.D. Miller and A. Jemal, Cancer statistics, 2015, *CA Cancer J Clin* **65** (2015), 5-29.
[2] L.A. Ries, Influence of extent of disease, histology, and demographic factors on lung cancer survival in the SEER population-based data, *Semin Surg Oncol* **10** (1994), 21-30.
[3] B.S. Kramer, C.D. Berg, D.R. Aberle and P.C. Prorok, Lung cancer screening with low-dose helical CT: results from the National Lung Screening Trial (NLST), *J Med Screen* **18** (2011), 109-11.
[4] A.K. Greenberg, F. Lu, J.D. Goldberg, E. Eylers, J.C. Tsay, T.A. Yie, D. Naidich, G. McGuinness, H. Pass, K.M. Tchou-Wong, D. Addrizzo-Harris, A. Chachoua, B. Crawford and W.N. Rom, CT scan screening for lung cancer: risk factors for nodules and malignancy in a high-risk urban cohort, *PLoS One* **7** (2012), e39403.
[5] D. Ghosh and L.M. Poisson, "Omics" data and levels of evidence for biomarker discovery, *Genomics* **93** (2009), 13-6.
[6] W.M. Claudino, P.H. Goncalves, A. di Leo, P.A. Philip and F.H. Sarkar, Metabolomics in cancer: a bench-to-bedside intersection, *Crit Rev Oncol Hematol* **84** (2012), 1-7.
[7] N.J. Serkova and K. Glunde, Metabolomics of cancer, *Methods Mol Biol* **520** (2009), 273-95.
[8] J.L. Spratlin, N.J. Serkova and S.G. Eckhardt, Clinical applications of metabolomics in oncology: a review, *Clin Cancer Res* **15** (2009), 431-40.
[9] J.F. Fahrmann, K. Kim, B.C. DeFelice, S.L. Taylor, D.R. Gandara, K.Y. Yoneda, D.T. Cooke, O. Fiehn, K. Kelly and S. Miyamoto, Investigation of Metabolomic Blood Biomarkers for Detection of Adenocarcinoma Lung Cancer, *Cancer Epidemiol Biomarkers Prev* (2015).
[10] W.R. Wikoff, S. Hanash, B. DeFelice, S. Miyamoto, M. Barnett, Y. Zhao, G. Goodman, Z. Feng, D. Gandara, O. Fiehn and A. Taguchi, Diacetylspermine is a Novel Prediagnostic Serum Biomarker for Non-Small-Cell Lung Cancer and Has Additive Performance With Pro-Surfactant Protein B, *J Clin Oncol* (2015).
[11] M. Scholz and O. Fiehn, SetupX–a public study design database for metabolomic projects, *Pac Symp Biocomput* (2007), 169-80.
[12] O. Fiehn, G. Wohlgemuth and M. Scholz, Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata, *Data Integration in the Life Sciences, Proceedings* **3615** (2005), 224-239.
[13] T. Pluskal, S. Castillo, A. Villar-Briones and M. Oresic, MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinformatics* **11** (2010), 395.

[14] T. Kind, K.H. Liu, Y. Lee do, B. DeFelice, J.K. Meissen and O. Fiehn, LipidBlast in silico tandem mass spectrometry database for lipid identification, *Nat Methods* **10** (2013), 755-8.
[15] R Development Core Team. R: A language and environment for statistical computing, *R Foundation for Statistical Computing* (2011), ISBN 3-900051-900007-900050.
[16] Y. Benjamini and Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society. Series B (Methodological)* (1995), 289-300.
[17] D. Grapov, DeviumWeb: Dynamic Multivariate Data Analysis and Visualization Platform, 2014.
[18] J. Xia, I.V. Sinelnikov, B. Han and D.S. Wishart, MetaboAnalyst 3.0-making metabolomics more meaningful, *Nucleic Acids Res* (2015).
[19] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.C. Sanchez and M. Muller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, *BMC Bioinformatics* **12** (2011), 77.
[20] E. Blaak, Gender differences in fat metabolism, *Curr Opin Clin Nutr Metab Care* **4** (2001), 499-502.
[21] N. Barzilai, D.M. Huffman, R.H. Muzumdar and A. Bartke, The critical role of metabolic pathways in aging, *Diabetes* **61** (2012), 1315-22.
[22] A. Chiolero, D. Faeh, F. Paccaud and J. Cornuz, Consequences of smoking for body weight, body fat distribution, and insulin resistance, *Am J Clin Nutr* **87** (2008), 801-9.
[23] R.A. Murphy, T.F. Bureyko, M. Mourtzakis, Q.S. Chu, M.T. Clandinin, T. Reiman and V.C. Mazurak, Aberrations in plasma phospholipid fatty acids in lung cancer patients, *Lipids* **47** (2012), 363-9.
[24] E. Marien, M. Meister, T. Muley, S. Fieuws, S. Bordel, R. Derua, J. Spraggins, R. Van de Plas, J. Dehairs, J. Wouters, M. Bagadi, H. Dienemann, M. Thomas, P.A. Schnabel, R.M. Caprioli, E. Waelkens and J.V. Swinnen, Non-small cell lung cancer is characterized by dramatic changes in phospholipid profiles, *Int J Cancer* **137** (2015), 1539-48.
[25] D. Zinrajh, G. Horl, G. Jurgens, J. Marc, M. Sok and D. Cerne, Increased phosphatidylethanolamine N-methyltransferase gene expression in non-small-cell lung cancer tissue predicts shorter patient survival, *Oncol Lett* **7** (2014), 2175-2179.
[26] L. Gao, Z. Wen, C. Wu, T. Wen and C.N. Ong, Metabolic profiling of plasma from benign and malignant pulmonary nodules patients using mass spectrometry-based metabolomics, *Metabolites* **3** (2013), 539-51.
[27] L.J. Huang, S.X. Chen, W.J. Luo, H.H. Jiang, P.F. Zhang and H. Yi, Proteomic analysis of secreted proteins of non-small cell lung cancer, *Ai Zheng* **25** (2006), 1361-7.
[28] G.P. Yu, G.Q. Chen, S. Wu, K. Shen and Y. Ji, The expression of PEBP4 protein in lung squamous cell carcinoma, *Tumour Biol* **32** (2011), 1257-63.
[29] X. Wang, N. Li, B. Liu, H. Sun, T. Chen, H. Li, J. Qiu, L. Zhang, T. Wan and X. Cao, A novel human phosphatidylethanolamine-binding protein resists tumor necrosis factor alpha-induced apoptosis by inhibiting mitogen-activated protein kinase pathway activation and phosphatidylethanolamine externalization, *J Biol Chem* **279** (2004), 45855-64.
[30] J. Liu, P.J. Mazzone, J.P. Cata, A. Kurz, M. Bauer, E.J. Mascha and D.I. Sessler, Serum free fatty acid biomarkers of lung cancer, *Chest* **146** (2014), 670-9.
[31] J.Z. Haeggstrom, A. Rinaldo-Matthis, C.E. Wheelock and A. Wetterholm, Advances in eicosanoid research, novel therapeutic implications, *Biochem Biophys Res Commun* **396** (2010), 135-9.
[32] Y.T. Sim and F.W. Poon, Imaging of solitary pulmonary nodule-a clinical review, *Quant Imaging Med Surg* **3** (2013), 316-26.