Published OnlineFirst August 17, 2015; DOI: 10.1158/1055-9965.EPI-15-0427

Research Article

Cancer Epidemiology, Biomarkers & Prevention

Investigation of Metabolomic Blood Biomarkers for Detection of Adenocarcinoma Lung Cancer

Johannes F. Fahrmann[1], Kyoungmi Kim[2], Brian C. DeFelice[1], Sandra L. Taylor[2], David R. Gandara[3], Ken Y. Yoneda[4], David T. Cooke[5], Oliver Fiehn[1,6], Karen Kelly[3], and Suzanne Miyamoto[3]

Abstract

Background: Untargeted metabolomics was used in case–control studies of adenocarcinoma (ADC) lung cancer to develop and test metabolite classifiers in serum and plasma as potential biomarkers for diagnosing lung cancer.

Methods: Serum and plasma were collected and used in two independent case–control studies (ADC1 and ADC2). Controls were frequency matched for gender, age, and smoking history. There were 52 adenocarcinoma cases and 31 controls in ADC1 and 43 adenocarcinoma cases and 43 controls in ADC2. Metabolomics was conducted using gas chromatography time-of-flight mass spectrometry. Differential analysis was performed on ADC1 and the top candidates (FDR < 0.05) for serum and plasma used to develop individual and multiplex classifiers that were then tested on an independent set of serum and plasma samples (ADC2).

Results: Aspartate provided the best accuracy (81.4%) for an individual metabolite classifier in serum, whereas pyrophosphate had the best accuracy (77.9%) in plasma when independently tested. Multiplex classifiers of either 2 or 4 serum metabolites had an accuracy of 72.7% when independently tested. For plasma, a multimetabolite classifier consisting of 8 metabolites gave an accuracy of 77.3% when independently tested. Comparison of overall diagnostic performance between the two blood matrices yielded similar performances. However, serum is most ideal given higher sensitivity for low-abundant metabolites.

Conclusion: This study shows the potential of metabolite-based diagnostic tests for detection of lung adenocarcinoma. Further validation in a larger pool of samples is warranted.

Impact: These biomarkers could improve early detection and diagnosis of lung cancer. *Cancer Epidemiol Biomarkers Prev; 24(11); 1716–23. ©2015 AACR.*

Introduction

Lung cancer continues to be a leading cause of cancer mortality in both men and women in the United States (1, 2). Among the different lung cancer types, non–small cell lung cancer (NSCLC) accounts for approximately 85% of all lung cancer cases, with adenocarcinoma being the most common histologic type (3).

Recently, the National Lung Cancer Screen Trial (NLST) demonstrated that low-dose CT (LDCT) screening could reduce mortality due to lung cancer by 20%. However, LDCT screening is largely hindered by high false-positive rates (96%), particularly in high-risk populations (heavy smokers), due to the low prevalence rates (less than 2%) of malignant tumors and high incidence of benign lung nodules. Consequently, complementary biomarkers that can be used in conjunction with LDCT screening to improve diagnostic capacities and reduce false-positive rates are highly desirable (4, 5). Preferably, such complementary tools should be noninvasive and exhibit high sensitivity and specificity. The application of "-omic" sciences (genomics, transcriptomics, proteomics, and metabolomics) represents valuable tools for the discovery and validation of potential biomarkers that can be used for detection of NSCLC. Of these omic sciences, metabolomics has received considerable attention for its application in cancer (6). Metabolomics is the assessment of small molecules and biochemical intermediates (metabolites) using analytic instrumentation. Metabolites in blood are the product of all cellular processes, which are highly responsive to conditions of disease and environment, and represent the final output products of all organs forming a detailed systemic representation of an individual's current physiologic state (7).

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, 9). As such, the application of metabolomics in NSCLC adenocarcinoma represents a promising avenue of new research for the identification and validation of potential biomarkers associated with diagnosis and prognosis.

In this study, we used an untargeted metabolomics approach using gas chromatography time-of-flight mass spectrometry

[1]University of California, Davis Genome Center, Davis, California. [2]Division of Biostatistics, Department of Public Health Sciences, School of Medicine, University of California Davis, Davis, California. [3]Division of Hematology and Oncology, Department of Internal Medicine, School of Medicine, University of California, Davis Medical Center, Sacramento, California. [4]Division of Pulmonary Medicine, Department of Internal Medicine, UC Davis Medical Center, Sacramento, California. [5]Division of Thoracic Surgery, Department of Surgery, School of Medicine, University of California, Davis Medical Center, Sacramento, California. [6]Department of Biochemistry, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Suzanne Miyamoto, UC Davis, 4501 X Street, Suite 3016, Sacramento, CA 95817. Phone: 916-734-3769; Fax: 916-734-5356; E-mail: smiyamoto@ucdavis.edu

doi: 10.1158/1055-9965.EPI-15-0427

Published OnlineFirst August 17, 2015; DOI: 10.1158/1055-9965.EPI-15-0427

(GCTOFMS) to analyze the metabolome of serum and plasma samples both collected from the same patients that were organized into two independent case–control studies (ADC1 and ADC2). In both studies, only NSCLC adenocarcinoma was investigated. The overall objectives were to (i) determine whether individual or combinations of metabolites could be used as a diagnostic test to distinguish NSCLC adenocarcinoma from controls and (ii) to determine which, plasma or serum, provides more accurate classifiers for the detection of lung cancer. We developed individual and multimetabolite classifiers using a training test from the ADC1 study and evaluated the performance of the constructed classifiers, individually or in combination, in an independent test/validation study (ADC2).

Materials and Methods

Patient population and collection of patient samples

Subjects were recruited over a 4-year period (2010–2014) from the UC Davis Medical Center and Cancer Center Clinics. All subjects were diagnosed with NSCLC adenocarcinoma before specimen collection. Blood samples (serum and plasma) were collected from NSCLC adenocarcinoma and control subjects with patient consent using approved IRB protocols (LC001 for cancer cases and LC002 for control cases). The control population was heavily recruited from spouses and family members accompanying a lung cancer patient to their clinic to maintain as much of a similar environment and life styles, especially diet and smoking history, as possible. Cases were frequency matched with controls for gender, age, and smoking history. Only cases diagnosed with NSCLC adenocarcinoma were used in these studies. Fasting status was not controlled for as individuals were recruited upon their arrival to the clinic.

Patient characteristics are described in Table 1. Detailed information on blood sample collection protocols is provided in Supplementary Methods.

Metabolomic profiling

The MiniX database (10) was used as a Laboratory Information Management System (LIMS) and for sample randomization before all analytic procedures. Sample identifications were kept blinded during the entire metabolomics analysis to minimize potential bias. Serum and plasma samples were equally distributed for analysis so they could be compared directly.

Plasma and serum sample preparation. Detailed information on sample preparation, instrument parameters, and data acquisition is provided in Supplementary Methods.

Samples (30 μL; serum or plasma) were thawed, extracted, and derivatized as previously described (11). Mass spectrometry analysis and data acquisition were performed using an Agilent 7890A gas chromatograph coupled to a Leco Pegasus IV time-of-flight (TOF) spectrometer. Acquired spectra were further processed using the BinBase database (10, 12).

Data analysis

Before statistical analysis, metabolite intensity values were total quantity normalized and $\log_2$ transformed. Missing intensity values were imputed with one-half the minimum observed matrix and metabolite specific value. Differential analysis was implemented to identify significant metabolomic differences between cancer and control samples in serum and plasma separately for both the training (ADC1) and test (ADC2) set. For each matrix, intensity values were regressed on the covariates [age, gender, and smoking history (packs per year)] and the residuals used to calculate *t* statistics for the difference between cancer and control groups adjusting for the covariates. Significance between cancer and control groups was determined on the basis of a permutation null distribution consisting of 100,000 permutations. FDRs were calculated to account for multiple testing and FDR < 0.05 was considered as significant.

Development of classifiers was carried out on the training set (ADC1). Classifiers consisted of individual metabolites and as a multiplex panel, for classifying samples as cancer or control (13). Only metabolites with a significant FDR (<0.05) were used in constructing classifiers. Furthermore, in developing classifiers, we used residuals from adjusting for age, gender, and smoking history and scaled the residuals to a variance of 1 for comparability between datasets.

Development of classifiers. Classifiers were developed using a strategy based on the use of "voting classifiers" as previously

Table 1. Patient characteristics

	ADC1 (training set)[a] for development		ADC2 (test set)[a] for validation
Variable	**Plasma**	**Serum**	**Plasma/serum**
Total sample size, *N*	83	80	86
Healthy controls, *N* (%)	31 (37.3%)	31 (38.8%)	43 (50%)
Cancer cases, *N* (%)	52 (62.7%)	49 (61.2%)	43 (50%)
By stage, *N* (%)			
I	21 (40.38)	19 (36.54)	18 (41.86)
II	7 (13.46)	7 (13.46)	3 (6.98)
III	14 (26.92)	14 (26.92)	7 (16.28)
IV	10 (19.23)	9 (17.31)	15 (34.88)
Gender, *N* (males/females)			
Controls	11/20	11/20	21/22
Cancer cases	17/35	17/32	21/22
Age (y), mean ± SD			
Controls	64.1 ± 8.97	64.1 ± 8.97	65.9 ± 8.05
Cancer cases	65.9 ± 9.66	65.9 ± 9.87	67.3 ± 10.10
Packs per year, mean ± SD			
Controls	29.8 ± 19.54	29.8 ± 19.54	38.6 ± 26.46
Cancer cases	34.6 ± 19.33	33.9 ± 20.06	39.5 ± 27.23

[a]No statistical differences in variables between cases and controls within each set and between the training set and test set. Control group for serum ADC1 is same as for plasma ADC1.

Published OnlineFirst August 17, 2015; DOI: 10.1158/1055-9965.EPI-15-0427

described (13). Detailed information regarding classifier development is provided in the Supplementary Methods.

Results

Subject characteristics

Patient characteristics for the two independent studies are provided in Table 1. The first set (ADC1) used for biomarker development consisted of serum and plasma samples obtained from 52 stages I–IV NSCLC adenocarcinoma patients (52 plasma and 49 serum), and 31 healthy controls (31 pairs of serum and plasma) for a total of 163 samples. Thirty-one control patients were enrolled as individual control subjects matched multiple cancer samples. This set was regarded as the training set for biomarker discovery and classifier development. A second, independent case–control study (ADC2) consisting of serum and plasma samples collected from 43 stage I–IV NSCLC adenocarcinoma patients and 43 healthy controls (total 172 samples) was used as an independent test set for biomarker evaluation. Samples for ADC2 were collected and analyzed at a different time from the ADC1. There were no significant differences between the matching variables of age, gender, and smoking history (packs per year) for the cases and control cohorts in the two separate case–control studies (ADC1 and ADC2).

Identification of metabolites discriminatory of NSCLC adenocarcinoma in serum

Untargeted GCTOF–based metabolomics was conducted on each sample in the ADC1 and then ADC2 set. A total of 511 metabolites were detected in ADC1, of which 181 had known annotated structures, whereas in ADC2 413 metabolites, of which 152 were known (Supplementary Tables S1 and S2). Of all metabolites detected between the two studies, 296 were repeatedly measured in serum and plasma (Supplementary Tables S1 and S2). Notably, many of the metabolites that were unique to either ADC1 or ADC2 were unknowns. These unknowns may represent artifacts, low-abundant compounds or xenobiotic compounds that were removed during the data-filtering process (see Supplementary Methods). Pearson correlation coefficients illustrating the association between measured metabolites are provided in Supplementary Table S3. Differential analysis (cancer vs. control) on the ADC1 training set identified 80 differential metabolites in the serum with a raw P value of <0.05 (Supplementary Table S1). Only four metabolites (xylose, glutamate, aspartate, and Bin_225393) remained significant after FDR adjustment (FDR < 0.05) in ADC1 (Supplementary Table S4). Three of the four significant metabolites (glutamate, aspartate, and Bin_225393) were found to be elevated in cancer relative to control in ADC1, whereas xylose was decreased (box plots for each metabolite are provided in Supplementary Fig. S1). Using the independent test set (ADC2), we conducted a separate differential analysis to confirm whether the metabolites identified in the first study (ADC1) were significantly and consistently differential in a different cohort of samples. Out of the 80 metabolites with a raw P value of <0.05 in ADC1, 15 (18.8%) were also found to be significant (raw P value < 0.05) in ADC2 (Supplementary Table S4). More importantly, all 15 metabolites indicated a similar fold change (increased or decreased) in both studies (Supplementary Table S4). When comparing only those metabolites that were significantly different following FDR adjustment in ADC1, 3 of the 4 metabolites (aspartate, glutamate, and Bin_225393) were also found to be significantly (FDR-adjusted $P < 0.05$) elevated in adenocarcinoma in ADC2 (Supplementary Fig. S1).

Developing serum metabolite classifiers using the ADC1 training set

Serum classifiers were developed using the metabolites whose peak intensity were significantly differential in relation to the cancer presence in ADC1 after FDR adjustment. Determination of classification thresholds and rules, construction of classifiers and cross-validation of classifier performance for both individual metabolites and panel of metabolites was performed as described in Materials and Methods. Performances of the developed classifiers for each individual metabolite in ADC1 are provided in Table 2. Individually, Bin_225393 displayed the highest individual accuracy of 72.5% (AUC = 0.766; Table 2 and Fig. 1A). An ROC curve plus the confidence interval for Bin_225393 is provided in Fig. 1A and Supplementary Table S5, respectively. The mass spectrum for Bin_225393 is provided in Supplementary Fig. S2. We next evaluated whether combining multiple metabolites in ADC1 could yield improved classifications. The order that each metabolite entered the multiplex classifier is provided in Table 2. Overall, the highest accuracies achieved were 72.5% and 68.8% consisting of Bin_225393 alone and the first three metabolites (Bin_225393, aspartate, and xylose), respectively (Table 2). An ROC curve plus the confidence interval for the three metabolite classifiers (Bin_225393, aspartate, and xylose) is provided in Fig. 1B and Supplementary Table S5, respectively.

Table 2. Performances of developed individual- and multimetabolite serum classifiers in ADC1 (training) and ADC2 (test) sets

Individual metabolite classifiers

	ADC1 (training)	ADC2 (test)		
Metabolite	Accuracy[a]	Accuracy[b]	Sensitivity	Specificity
Xylose	53.8	50.0	76.7	23.3
Glutamate	61.3	74.4	65.1	83.7
Aspartate	63.8	81.4	67.4	95.4
Bin_225393	72.5	64.0	74.4	53.5

Multimetabolite classifiers

Metabolite	Accumulated accuracy[a]	Accumulated accuracy[b]	Sensitivity	Specificity
Bin_225393[1]	72.5	64.0	74.4	53.5
Asparate[2]	68.1	72.7	70.9	74.4
Xylose[3]	68.8	67.4	74.4	60.5
Glutamate[4]	66.9	72.7	70.9	74.4

NOTE: [1–4], denotes entry into classifier.

[a]ADC1 (training set).

[b]ADC2 (test set).

Published OnlineFirst August 17, 2015; DOI: 10.1158/1055-9965.EPI-15-0427

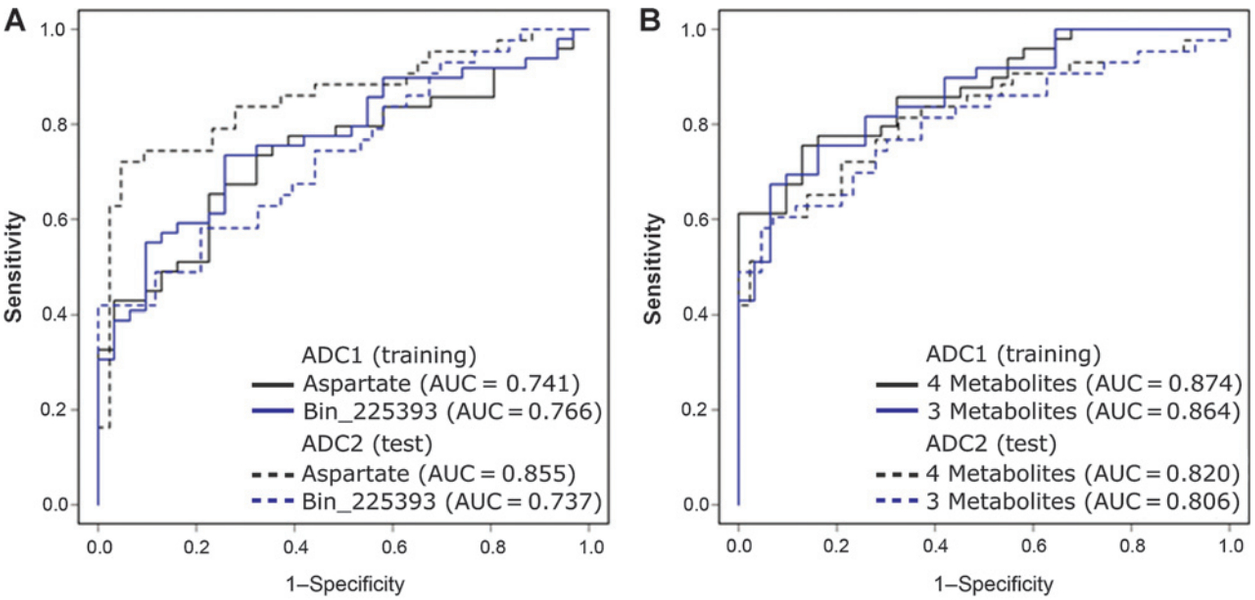

Figure 1.
ROC curves for individual- and multimetabolite classifiers in serum. A, ROC curves for aspartate and Bin_225393 in serum. B, ROC curves for two multimetabolite classifiers in serum. The four-metabolite classifier contains all the metabolites included in the classifier (Table 2). The three-metabolite classifier includes Bin_225393, aspartate and xylose. Confidence intervals for AUCs are provided in Supplementary Table S5.

Testing/validation of serum classifiers developed with ADC1 training set in an independent ADC2 test set

We next evaluated the performance of the serum metabolite classifiers developed using the ADC1 training set on the independent ADC2 test set. Individually, aspartate yielded the best performance with a classification accuracy of 81.4% (AUC = 0.855) when tested in ADC2 (Table 2). An ROC curve plus the confidence interval for aspartate is provided in Fig. 1A and Supplementary Table S5, respectively. Then we tested the multimetabolite classifiers consisting of up to 4 metabolites developed with the ADC training set in the test set for assessment of performance accuracy. The highest performance was achieved with all four metabolites in the classifier yielding an accuracy of 72.7% (Table 2). An ROC curve plus the confidence interval for the 4 metabolite classifiers (Bin_225393, aspartate, and xylose) is shown in Fig. 1B and Supplementary Table S5, respectively.

Identification of metabolites discriminatory of NSCLC adenocarcinoma in plasma

In addition to serum, we examined the performance of plasma-derived metabolite classifiers as potential biomarkers for NSCLC adenocarcinoma. Differential analysis identified 68 differential metabolites in plasma samples with a raw P value of <0.05 (Supplementary Table S2) in the ADC1 set. Only 14 (21%) of the 68 metabolites remained significant following FDR adjustment (Supplementary Table S2). Of these 68 metabolites, 18 (26.5%) were also found to be significantly different in the ADC2 test set (Supplementary Table S6). When comparing only those 14 metabolites that remained significant after FDR-adjustment in the ADC1 set, 6 (pyrophosphate, maltotriose, citrulline, adenosine-5-phosphate, Bin_226841, and Bin_36799) were also found to remain significant following FDR adjustment (FDR P value <0.05) in the ADC2 set (Supplementary Table S2). All 6 of these metabolites displayed the same direction of change in ADC2 (increased or decreased) as observed in ADC1 (Supplementary Fig. S1).

Developing plasma metabolite classifiers using the ADC1 training set

Plasma classifiers were developed from the 13 (Supplementary Table S2) discriminating metabolites that remained significant following FDR adjustment. Performances of the developed classifiers for each individual metabolite are provided in Table 3. Four metabolites (maltotriose, maltose, cellobiotol, and Bin_715929) had individual accuracy scores above 70% (Table 3). However, three metabolites (cellobiotol, Bin_715929, and Bin_299216) were not detected in the ADC2 test set and consequently excluded when developing classifiers to apply to the test set for performance evaluation. The nondetection of cellobiotol and unknown compounds (Bin_715929 and Bin_299216) is suspected to be the consequence of low spectral abundance or only being detected in few patients, thus resulting in removal of these compounds during the data-filtering processes (see Supplementary Methods). Overall, maltose performed the best with an accuracy of 72.3% (Table 3 and Fig. 2A). An ROC curve plus the confidence interval for maltose is provided in Fig. 2A and Supplementary Table S5, respectively. We subsequently evaluated whether combining multiple plasma metabolites could serve a better classification test. The order that each metabolite entered the classifier is provided in Table 3. Overall, the highest accuracy achieved for the plasma classifiers was 79.5% using a panel of five metabolites (Table 3 and Fig. 2B), suggesting that several metabolites in the classifier can improve classification relative to individual metabolite classifiers. An ROC curve plus confidence intervals for the 5-metabolite classifier is provided in Fig. 2B and Supplementary Table S5, respectively.

Published OnlineFirst August 17, 2015; DOI: 10.1158/1055-9965.EPI-15-0427

Table 3. Performances of developed individual- and multimetabolite plasma classifiers in ADC1 (training) and ADC2 (test) sets

Individual metabolite classifier				
	ADC1 (training)	ADC2 (test)		
Metabolite	Accuracy[a]	Accuracy[b]	Sensitivity	Specificity
Tryptophan	50.6	52.3	39.5	65.1
Pyrophophosphate	66.3	77.9	60.5	95.4
Maltotriose	71.1	64.0	79.1	48.8
Maltose	72.3	62.8	55.8	69.8
Cystine	66.3	68.6	58.1	79.1
Citrulline	67.5	66.3	55.8	76.7
Cellobiotol[d]	71.1	ND[c]	ND	ND
Adenosine-5-phosphate	68.7	72.1	67.4	76.7
3-Phosphoglycerate	69.9	51.2	30.2	72.1
Bin_226841	60.2	64.0	37.2	90.7
Bin_715929[d]	72.3	ND	ND	ND
Bin_367991	59.0	65.1	39.5	90.7
Bin_299216[d]	60.2	ND	ND	ND

Multimetabolite classifier				
Metabolite	Accumulated accuracy[a]	Accumulated accuracy[b]	Sensitivity	Specificity
Maltose[1]	72.3	62.8	55.8	69.8
Maltotriose[2]	71.7	63.4	67.4	59.3
Cystine[3]	75.9	69.8	69.8	69.8
3-Phosphoglycerate[4]	76.5	69.2	61.6	76.7
Citrulline[5]	79.5	73.3	65.1	81.4
Pyrophosphate[6]	77.1	75.6	64.0	87.2
Tryptophan[7]	77.1	76.7	60.5	93.0
Adenosine-5-Phosphate[8]	75.3	77.3	64.0	90.7
Bin_226841[9]	77.1	76.7	60.5	93.0
Bin_367991[10]	73.5	75.0	55.8	94.2

NOTE: Superscript numbers 1 to 10 denote entry into the classifier.
[a]ADC1 (training set).
[b]ADC2 (test set).
[c]Not detected.
[d]Metabolite not included in the multimetabolite classifier.

Testing/validation of plasma classifiers developed with ADC1 training set in an independent ADC2 test set

We next evaluated the performance of the plasma metabolite classifiers developed using the training set (ADC1) in the discovery phase on the independent ADC2 test set. Individually, plasma metabolites indicated modest performances in classifying ADC2 samples with pyrophosphate achieving the highest accuracy (77.9%) and specificity (95.4%) but low sensitivity (60.5%; Table 3). An ROC curve plus confidence intervals for pyrophosphate is shown in Fig. 2A and Supplementary Table S5,

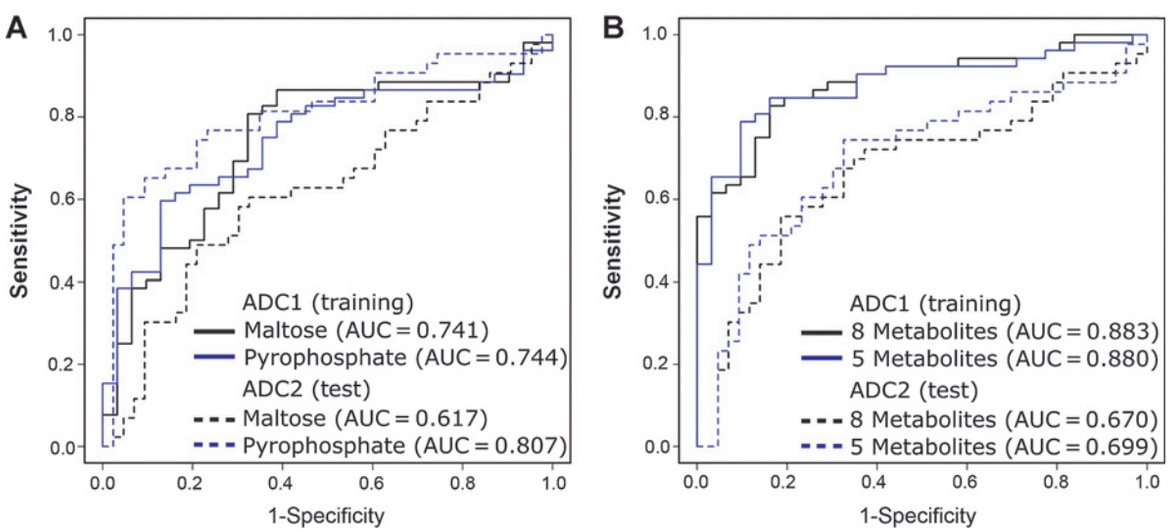

Figure 2.
ROC curves for individual- and multimetabolite classifiers in plasma. A, ROC curves for maltose and pyrophosphate in plasma. B, ROC curves for the best multimetabolite ADC1 classifier for plasma consisting of five metabolites (blue line) and the classifier consisting of eight (black line) when applied to ADC2 are shown (Table 3). Confidence intervals for AUCs are provided in Supplementary Table S5.

Published OnlineFirst August 17, 2015; DOI: 10.1158/1055-9965.EPI-15-0427

respectively. We afterwards evaluated the developed multiplex classifiers in the independent test set using the refined 10 metabolites classifiers developed in the training set. Collectively, the best performance was achieved when using a combination of 8 metabolites in the plasma classifier resulting in an accuracy of 77.3% (Table 3). An ROC curve plus confidence intervals for the 8-metabolite classifier is shown in Fig. 2A and Supplementary Table S5, respectively.

Comparison of classifier performances between plasma and serum samples

It is of particular interest for clinical utility to determine which type of blood specimen (serum or plasma) is best suited for obtaining optimal classifiers for the detection of lung cancer. For this reason, we collected both serum and plasma from the same individuals and developed the classifiers from these two biofluids independently. Collectively, three metabolites (maltotriose, glutamate, and Bin_223618) were found to be consistently differential between ADC and control (Supplementary Tables S1 and S2) in both serum and plasma.

Comparison of individual metabolite classifier performances in either serum or plasma yielded comparable results although classifier accuracies were slightly better in serum (ranging from 50%–81% accuracy) than plasma (ranging from 51%–78% accuracy; Tables 2 and 3). However, plasma provided slightly better performance metrics using a multimetabolite classifier compared with serum (77.3% versus 72.7% accuracy in ADC2, respectively; Tables 2 and 3), indicating that serum may be better suited for individual-metabolite classifiers; whereas plasma may be more suited for multimetabolite classifiers.

Discussion

In the present study, we identified multiple circulating metabolites (annotated and unknown) that are significantly elevated or reduced in patients with NSCLC adenocarcinoma compared with healthy controls. In the discovery phase of our experimental design, we identified and developed classifiers in a training set that were then applied to an independent test set for testing/validation. This approach is consistent with the guidelines set forth by the U.S. National Cancer Institute for evaluating potential diagnostic cancer biomarkers (14). Within each study, sample sets were matched by age, gender, and smoking history, and rigorous statistical evaluations were implemented to analyze the test performance of the metabolite compositions.

Overall, the individual-metabolite classifier aspartate showed the best accuracy (81.4%) in serum, whereas pyrophosphate provided the best accuracy (77.9%) in plasma when tested in the independent ADC2 test/validation study. A combination of either 2 or 4 metabolites (Table 2) in the serum classifier gave the best performance with an accuracy of 72.7% when applied to the ADC2 test set. For plasma, a multipanel classifier consisting of 8 metabolites (Table 3) provided the best performance with an accuracy of 77.3% in the ADC2 test set. Although the ideal (but unrealistic) situation is for 100% sensitivity and specificity, in this study we focused on building a multiplex test with high specificity and less emphasis on sensitivity. In this way, nearly all of the true negative and false positives could be correctly identified as cancer free.

The performance of the developed classifiers in our study for the identification of NSCLC adenocarcinoma is comparable with that of others (15–17). Patz and colleagues (15) illustrated that a panel of four serum proteins (CEA, retinol binding protein, a1-antitrypsin, and squamous cell carcinoma antigen) was found to have a sensitivity of 89.3% and a specificity of 84.7% in a case–control training set for lung cancer that in a validation study yielded a sensitivity of 77.8% and a specificity of 75.4%. Li and colleagues (16) developed a 13-protein classifier from a panel of 371 protein candidates, previously identified in 143 plasma samples obtained from patients with benign and malignant lung nodules. This 13-protein classifier was validated on an independent set of plasma samples ($n = 104$) yielding a negative predictive value (NPV) of 90% and a specificity of 44% $\pm$ 13%. Results from the analysis of a third independent study showed an NPV of 94% and a specificity of 56% (16). AUCs of 0.82, 0.60, and 0.74 were obtained for discovery, validation study 1, and validation study 2, respectively (16). As in our study, they also determined that their classifier score was independent of smoking history and age (16).

Recently, Sozzi and colleagues (17) demonstrated the effectiveness of a plasma miRNA signature classifier using samples collected from smokers within the randomized Multicenter Italian Lung Detection trial, both individual and in conjunction with LDCT. Using this miRNA classifier, a sensitivity of 87% and a specificity of 81% were achieved for all case–control samples, whereas 88% sensitivity and 80% specificity were achieved for cases from the low-dose CT arm (17).

Taken together, both methods (miRNA and metabolomics) or all methods (miRNAs, proteomics, and metabolomics) might eventually be combined to produce a test with even better performance than each individual test for early detection of all types of lung cancer. Despite the potential prospects of combined methodologies (such as miRNA and metabolomics), our results highlight the application of metabolomics in the discovery phase of potential biomarkers and yield candidate classifiers that will be expanded upon in future studies. Moreover, we have tested both plasma and serum from the same individuals to determine which type of blood specimen would be more suitable for metabolomics-derived classifiers. Specifically, we aimed to determine which biofluid would provide the most reliable classifiers with the potential for general utility as this is a clinically relevant question. Overall, there were no major advantages by using either serum or plasma as both blood-specimen types yielded comparable results in overall performances; although serum-based classifiers performed slightly better for individual-metabolite classifiers, plasma performed slightly better for multimetabolite classifiers. These findings are in agreement with those by Wedge and colleagues and Yu and colleagues (18, 19) who similarly stated that the two biofluids were comparable, although in both studies plasma provided slightly better results. However, despite their findings, Yu and colleagues (19) also found that metabolite concentrations were higher in serum, allowing for more sensitive results in biomarker detection. We suggest that serum is most ideal, given higher sensitivity for low-abundant metabolites as the low-abundant metabolites/compounds could potentially be of most biologically relevant.

Although the application of metabolomics in the identification of potential biomarkers is of immense value, it also provides useful information about the pathophysiology of the disease. Recently, we distinguished metabolic differences between matched malignant and nonmalignant lung tissue from subjects with early-stage (stage IA–B) NSCLC adenocarcinoma (20). Particularly, we identified that glutamate, malate, adenosine-5-phosphate, and xanthine were significantly increased, whereas glutamine was

Published OnlineFirst August 17, 2015; DOI: 10.1158/1055-9965.EPI-15-0427

reduced in malignant tissue compared with nonmalignant tissue (20). In the current investigation, we also found that glutamate, malate, adenosine-5-phosphate, and xanthine were consistently elevated in NSCLC, whereas glutamine was consistently reduced. However, only glutamate was found to be reliably significant. The elevation in glutamate and reduction in glutamine is particularly noteworthy, given the recently recognized importance of glutamate and glutamine in energy metabolism and macromolecule biosynthesis in lung cancer (21). The abundances of these metabolites and others mentioned above in serum or plasma may, therefore, be a reflection of tumorigenesis and cumulatively point toward alterations in nucleotide and energy metabolism.

Finally, it is important to recognize both the strengths and limitations of this study. A strength of this study is that our analysis examined only adenocarcinoma NSCLC. By evaluating only adenocarcinoma, we exclude any potential biases from mixed pathologies during the classifier construction allowing for the identification of adenocarcinoma-specific biomarkers. This is particularly important for diagnosis and treatment options. Unlike many biomarkers studies, candidate biomarkers were selected on the basis of their FDRs and not on raw *P* values, leading to less false-positive rates of developing biomarkers. A limitation of this study is the relatively small sample size for each cohort (52 cases, 31 controls for ADC1, and 43 cases and 43 controls for ADC2) because patient variability can be a big factor in smaller studies (22). However, this study is still part of the discovery phase and, as such, will require further validation in a larger population. In addition, values of detected metabolites are based on qualitative peak heights rather than absolute concentrations. The intent of this study was to obtain maximal coverage of the metabolome for the identification, generation, and evaluation of classifiers that could be used as potential diagnostic markers for NSCLC adenocarcinoma. However, it will be important to verify and quantitate absolute concentrations in future studies. The inclusion of both early- and late-stage adenocarcinoma may also lead to masking of potential biomarkers, given that heterogeneity exists among tumor stage. Finally, although collection of samples from a single institution can be a strength due to consistency in protocols, it also poses as a limitation as there is potential bias of site-specific confounding effects attributed to differences in sample collection and sample handling. Therefore, larger studies are warranted on only early-stage NSCLC adenocarcinoma and also include samples from multiple institutions to evaluate reproducibility and performance independent of the collection site.

In conclusion, our results highlight the application and validity of metabolomics in the discovery and validation of candidate biomarkers for diagnosing NSCLC adenocarcinoma. More specifically, we have identified individual metabolites and multimetabolite panels as part of the discovery phase, which will serve as the basis for classifiers used in future studies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: K. Kim, D.R. Gandara, K.Y. Yoneda, D.T. Cooke, O. Fiehn, K. Kelly, S. Miyamoto

Development of methodology: K. Kim, B.C. DeFelice, O. Fiehn, K. Kelly

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): B.C. DeFelice, D.R. Gandara, K.Y. Yoneda, O. Fiehn, K. Kelly, S. Miyamoto

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J.F. Fahrmann, K. Kim, S.L. Taylor, O. Fiehn, K. Kelly, S. Miyamoto

Writing, review, and/or revision of the manuscript: J.F. Fahrmann, K. Kim, D.R. Gandara, K.Y. Yoneda, O. Fiehn, K. Kelly, S. Miyamoto

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): K. Kim, D.R. Gandara, O. Fiehn, S. Miyamoto

Study supervision: K. Kim, D.T. Cooke, K. Kelly, S. Miyamoto

Acknowledgments

The authors acknowledge the assistance of Valerie Kuderer, Olesya Litovka, and Yahanna Sandoval Torres with collection of samples and managing the clinical specimen database. The authors also acknowledge the services and help with the storage of samples in the UC Davis Cancer Biorepository (Regina Gandour-Edwards and Irmgard Feldman).

Grant Support

This study is supported by Tobacco Related Disease Research Program 20PT0034 (to K. Kelly), LUNGevity Foundation and The Thomas G. Labrecque Foundation 201118739 (to S. Miyamoto), NIH West Coast Metabolomics Resource Core U24 DK097154 (to O. Fiehn), and W81XWH-10-1-0635 DOD LCRP (to D.R. Gandara).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 21, 2015; revised July 7, 2015; accepted July 27, 2015; published OnlineFirst August 17, 2015.

References

1. Prevention CfDCa. National Center for Health Statistics. CDC WONDER On-line Database, compiled from Compressed Mortality File 1999–2012. 2014;Series 20 No. 2R.
2. Society AC. Cancer Facts and Figures. 2014:1–60. Available form: http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2014/index.
3. Breathnach OS, Freidlin B, Conley B, Green MR, Johnson DH, Gandara DR, et al. Twenty-two years of phase III trials for patients with advanced non–small cell lung cancer: sobering results. J Clin Oncol 2001;19:1734–42.
4. Pass HI, Beer DG, Joseph S, Massion P. Biomarkers and molecular testing for early detection, diagnosis, and therapeutic prediction of lung cancer. Thorac Surg Clin 2013;23:211–24.
5. Hassanein M, Callison JC, Callaway-Lane C, Aldrich MC, Grogan EL, Massion PP. The state of molecular biomarkers for the early detection of lung cancer. Cancer Prev Res 2012;5:992–1006.
6. Serkova NJ, Glunde K. Metabolomics of cancer. Methods Mol Biol 2009;520:273–95.
7. Kwon H, Oh S, Jin X, An YJ, Park S. Cancer metabolomics in basic science perspective. Arch Pharm Res 2015;38:372–80.
8. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. Clin Cancer Res 2009;15:431–40.
9. Claudino WM, Goncalves PH, di Leo A, Philip PA, Sarkar FH. Metabolomics in cancer: a bench-to-bedside intersection. Crit Rev Oncol Hematol 2012;84:1–7.
10. Scholz M, Fiehn O. SetupX–a public study design database for metabolomic projects. Pac Symp Biocomput 2007;12:169–80.
11. Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee do Y, Lu Y, et al. Quality control for plant metabolomics: reporting MSI-compliant studies. Plant J 2008;53:691–704.
12. Fiehn O, Wohlgemuth G, Scholz M. Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. Data Integration in the Life Sciences, Proceedings 2005;3615:224–39.
13. Taylor SL, Kim K. A jackknife and voting classifier approach to feature selection and classification. Cancer Inform 2011;10:133–47.

14. McShane LM, Cavenagh MM, Lively TG, Eberhard DA, Bigbee WL, Williams PM, et al. Criteria for the use of omics-based predictors in clinical trials: explanation and elaboration. BMC Med 2013;11:220.
15. Patz EF Jr, Campa MJ, Gottlin EB, Kusmartseva I, Guan XR, Herndon JE II. Panel of serum biomarkers for the diagnosis of lung cancer. J Clin Oncol 2007;25:5578–83.
16. Li XJ, Hayward C, Fong PY, Dominguez M, Hunsucker SW, Lee LW, et al. A blood-based proteomic classifier for the molecular characterization of pulmonary nodules. Sci Transl Med 2013;5:207ra142.
17. Sozzi G, Boeri M, Rossi M, Verri C, Suatoni P, Bravi F, et al. Clinical utility of a plasma-based miRNA signature classifier within computed tomography lung cancer screening: a correlative MILD trial study. J Clin Oncol 2014;32:768–73.
18. Wedge DC, Allwood JW, Dunn W, Vaughan AA, Simpson K, Brown M, et al. Is serum or plasma more appropriate for intersubject comparisons in metabolomic studies? An assessment in patients with small-cell lung cancer. Anal Chem 2011;83:6689–97.
19. Yu Z, Kastenmuller G, He Y, Belcredi P, Moller G, Prehn C, et al. Differences between human plasma and serum metabolite profiles. PLoS ONE 2011;6:e21230.
20. Wikoff W, Grapov D, Fahrmann J, DeFelice B, Rom W, Pass H, et al. Metabolomic markers of altered nucleotide metabolism in early stage adenocarcinoma. Cancer Prev Res 2015;8:410–8.
21. Mohamed A, Deng X, Khuri FR, Owonikoko TK. Altered glutamine metabolism and therapeutic opportunities for lung cancer. Clin Lung Cancer 2014;15:7–15.
22. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 2001;93:1054–61.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Investigation of Metabolomic Blood Biomarkers for Detection of Adenocarcinoma Lung Cancer

Johannes F. Fahrmann, Kyoungmi Kim, Brian C. DeFelice, et al.

Cancer Epidemiol Biomarkers Prev 2015;24:1716-1723. Published OnlineFirst August 17, 2015.