Published in final edited form as:

Lung Cancer. 2017 December ; 114: 70–78. doi:10.1016/j.lungcan.2017.10.018.

Circulating metabolite profiles to predict overall survival in advanced non-small cell lung cancer patients receiving first-line chemotherapy

Jie Shen¹, Yuanqing Ye¹, David W. Chang¹, Maosheng Huang¹, John V. Heymach², Jack A. Roth³, Xifeng Wu^{1,*}, and Hua Zhao¹

¹Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA 77030

²Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA 77030

³Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA 77030

Abstract

Objectives—The prognosis for advanced-stage non-small cell lung cancer (NSCLC) is usually poor. However, survival may be variable and difficult to predict. In the current study, we aimed to identify circulating metabolites as potential predictive biomarkers for overall survival of advanced-stage (III/IV) NSCLC patients treated with first-line platinum-based chemotherapy.

Materials and methods—Using two-stage study design, we performed global metabolomic profiling in blood of 220 advanced-stage NSCLC patients, including 110 with poor survival and 110 with good survival. Metabolomic profiling was conducted using Metabolon platform. The association of each metabolite with survival was assessed by Cox proportional hazard regression model with adjustment for covariates.

Results and conclusion—We found levels of 4 metabolites, caffeine, paraxanthine, stachydrine, and methyl glucopyranoside (alpha + beta), differed significantly between NSCLC patients with poor and good survival in both discovery and validation phases ($P < 0.05$). Interestingly, majority of the identified metabolites are involved in caffeine metabolism, and 2 metabolites are related to coffee intake. In fact, caffeine metabolism pathway was the only significant pathway identified which significantly differed between NSCLC patients with poor and good survival ($P = 1.48 \times 10^{-7}$) in the pathway analysis. We also found 4 metabolites whose levels were significantly associated with good survival in both discovery and validation phases. Strong cumulative effects on overall survival were observed for these 4 metabolites. In conclusion, we

*Corresponding Author: Xifeng Wu, MD, PhD, Department of Epidemiology, Unit 1340, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030. Tel: 713-745-2485; xwu@mdanderson.org.

Disclosure of potential conflicts of interest: No potential conflicts of interest were disclosed.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

identified a panel of metabolites in caffeine metabolism pathway that may predict survival outcome in advanced-stage NSCLC patients. The identified small metabolites may be useful biomarker candidates to help identify patients who may benefit from platinum-based chemotherapy.

Keywords

circulating metabolite; metabolomics; non-small cell lung cancer; survival; platinum-based chemotherapy

1. Introduction

Lung cancer is the one of the most common cancers, and the most common cause of cancer-related death in men and second most common in women (1). The high mortality of lung cancer is mainly due to late diagnosis, resulting in poor response to clinical intervention (2). Platinum-based chemotherapy is the standard treatment for advanced-stage (III/IV) NSCLC (3). Although the prognosis is dismal, significant variations still exist. Unfortunately, current prognostic prediction tools, mainly built upon clinicopathological characteristics and/or proliferation markers, are not sufficiently accurate to predict which patients have virtually no benefit from the standard treatment and would, therefore, be candidates for alternative therapy, and which patients have high response to standard treatment such that more intensive therapy would not be required (4,5). For example, downregulated survivin has been found to be significantly correlated with improved overall survival in NSCLC patients who received platinum-based therapy. However, the potential of survivin as a prognostic marker is variably reported depending on the lung cancer subtype and stage (4, 6). Large scale molecular characterization in tumor tissues may help, but somatic information for gene mutation and expression profiles are still rarely available in clinical practice. Therefore, there is an urgent need for useful, noninvasive prognostic markers including metabolic biomarkers for advanced-stage lung cancer.

Altered metabolic homeostasis plays important roles in the development of tumors (7). Metabolites are sensitive to physiologic and pathologic stimuli, environmental exposures, pharmaceutical interventions, and have potential to serve as indicators of the overall physiological status, such as cancer. Metabolites are end products of intercellular pathways, and are much more stable compared to the upstream products, such as mRNAs and proteins (8,9). Therefore, metabolomic profiling, which systematically quantify small-molecule metabolites and record unique chemical fingerprints, has emerged as an important tool to identify diagnostic and prognostic biomarkers in various biospecimens including serum, plasma, and urine. An increasing number of studies have utilized metabolomic profiling to identify biomarkers associated with early diagnostic, tumor characterization, prognosis, treatment response, and toxicity in various types of cancers (2)(10-24). In lung cancer, metabolomic studies have been performed in blood and urine samples to identify metabolite biomarker for lung cancer diagnosis (2)(25-30), tumor characterization (31-2) and progression (2,27). In our previous lung cancer study, we found that low levels of serum bilirubin were associated with increased risk of lung cancer incidence and mortality in male smokers (25).

To our knowledge, there has been no study to date examining the relation between serum metabolomic profiles and clinical outcome in advanced-stage lung cancer patients. To fill the gap, we conducted a two-stage study to examine the role of blood metabolites in predicting overall survival among advanced-stage NSCLC patients who received platinum-based chemotherapy.

2. Material and methods

2.1. Study subjects

In this study, a total of 220 Caucasian advanced-stage (III/IV) NSCLC patients who received first-line platinum-based chemotherapy with or without radiotherapy were included. All study subjects are participants in an ongoing lung cancer case-control study at the University of Texas MD Anderson Cancer Center (Houston, TX). The study subjects were accrued from 1997 to 2009, and their median follow-up time are 14.0 months. Details of subject recruitment methods have been reported previously (25). All study subjects completed an in-person interview using a structured questionnaire and donated 40 mL blood sample for molecular analysis. The blood samples were collected prior to any cancer treatment. Demographic characteristics, smoking history, family history of cancer, and exposure data were collected. Among these 220 patients, 110 had good survival (defined as survival >18 months) and 110 had poor survival (defined as survival <12 months). About 120 patients received chemotherapy only and 100 received chemoradiation. The poor and good survival groups were well-matched on age (+/-5 years), gender and mean blood sample collection time (+/-1 year). In current study, the study subjects were divided into discovery cohort (n= 110, 55 with good survival and 55 with poor survival), and validation cohort (n= 110, 55 with good survival and 55 with poor survival). Serum and plasma samples were used in discovery and validation phase, respectively. The clinical data for the cases were obtained from medical record review. The study protocol was approved by the Institutional Review Board of The University of Texas M. D. Anderson Cancer Center.

2.2. Metabolite profiling

Global metabolomic profiling was conducted with 200ul serum samples for discovery cohort and 200ul plasma samples for validation cohort at Metabolon Inc. (Durham, NC) as described previously (16). Internal controls were used to control for experimental variability, which included controls for injection, processing, and alignment standards for quality assurance/quality control (QA/QC) procedures. All samples were kept at -80°C until analyses were performed.

2.3. Pathway analysis

MetaboAnalyst 3.0 was used to identify the significant pathway of altered metabolites. This pathway analysis uses the high-quality KEGG metabolic pathways as the backend knowledge base, and integrates powerful pathway enrichment analysis and pathway topology analysis. First, by using the compound identifications (common names, HMDB IDs or KEGG IDs), 11 top differentially expressed metabolites in both discovery and validation phases were matched with the compounds contained in the pathway analysis. The metabolites which could not be matched were excluded from subsequent pathway analysis.

2.4. Statistical analysis

Stata (version 14.0; StataCorp, College Station, TX) was used for the statistical analysis. Differences in distribution of the host characteristics between poor survival and good survival were evaluated by Pearson χ^2 test for categorical variables and Student t test or Wilcoxon ranksum test for continuous variables. Only metabolites with detectable expression in at least 80% of samples were kept for analysis. There are 623 metabolites in the discovery data and 513 metabolites in the validation data which satisfied the criteria. A total of 454 metabolites overlap between the discovery and the validation phase were included in the final analysis. The difference between the poor survival and good survival were assessed by Wilcoxon rank sum test in both the discovery and validation phases and in the subgroup analysis for adenocarcinoma. In comparing the significant metabolites identified from the discovery and validation cohorts, we first chose 0.05 as P value threshold to identify significant metabolites. Thereafter, to further identify more candidate metabolites, we increased P value threshold for significance to 0.10 to find more significant metabolites and include them in the pathway analysis. The association of each metabolite with survival was also assessed by Cox proportional hazard regression model while adjusting by smoking status, stage, pathology, grade, performance status and treatment regimen (chemotherapy and/or radiotherapy). The risk score was defined as the weighted sum of metabolite where the weight is the beta coefficient. Additionally, for metabolites and risk score, we applied Kaplan-Meier analysis and the log-rank tests to evaluate the differences in survival on categorized metabolites/risk score subgroups.

3. Results

Characteristics of the study populations

Demographic and pathological characteristics of the patient cohort were shown in Table 1. Briefly, a total of 220 study subjects (discovery cohort: 110, validation cohort: 110) were included in the analysis. Among them, the discovery cohort included 55 patients with good survival and 55 with poor survival, and validation cohort included 55 patients with good survival and 55 with poor survival. In the discovery cohort, the poor and good survival groups were well-matched on age, gender, mean blood sample collection time, tumor grade, and tumor type. The poor survival group had more current smokers and fewer never smokers than the good survival group ($P=0.015$). In addition, the poor survival group had more patients with suboptimal performance status (2-3) than the good survival group ($P=0.001$). In the validation cohort, the poor and good survival groups were well-matched on all selected characteristics except tumor stage. The poor survival group had greater percentage of patients with grade IV NSCLC than the good survival group (64% vs. 45%, respectively; $P=0.056$), although the difference is only borderline significant.

Global metabolomic profiling of lung cancer

The basic study scheme was shown in Figure 1. We performed the global metabolomic profiling for discovery cohort first, then for validation cohort, and identified top significant metabolites that differed between NSCLC patients with poor and good survival in both phases. A total of 807 metabolites were detected in the serum samples of NSCLC patients. After quality control analyses, we observed 623 and 513 serum metabolites with less than

20% missing values in the discovery and validation cohorts, respectively. Among them, 454 serum metabolites were observed in both cohorts. In the discovery cohort, 78 serum metabolites whose levels significantly differed between the poor and good survival groups were identified ($P<0.05$) (Supplemental Table 1). Among them, 39 metabolites were increased in the poor survival group, and 39 were decreased. In addition, 20 serum metabolites whose levels differed significantly between the poor and good survival groups were identified ($P<0.01$). Among them, 10 metabolites were increased in the poor survival group, and 10 were decreased. In the validation cohort, 49 serum metabolites whose levels differed significantly between the poor and good survival groups were identified ($P<0.05$) (Supplemental Table 2). Among them, 22 metabolites were increased in the poor survival group, and 27 were decreased. In addition, 8 serum metabolites whose levels differed significantly between the poor and good survival groups were identified ($P<0.01$). Among them, 4 metabolites were increased in the poor survival group, and 4 were decreased.

In comparing the significant metabolites identified from the discovery and validation cohorts, 4 significant metabolites ($P<0.05$) identified in the discovery cohort were also significant in the validation cohort (Table 2). They were caffeine, paraxanthine, stachydrine, methyl glucopyranoside (alpha + beta). Among them, mean levels of caffeine and paraxanthine were significantly higher in the poor than good survival groups. On the other hand, mean levels of stachydrine and methyl glucopyranoside (alpha + beta) were significantly lower in the poor than good survival groups. Interestingly, both caffeine and paraxanthine belong to caffeine metabolism pathway, and both stachydrine, methyl glucopyranoside (alpha + beta) belong to xenobiotics pathway. To further identify more candidate metabolites, we increased P value threshold for significance to 0.10. We found 7 additional metabolites identified in the discovery cohort were also significant in the validation cohort ($P<0.10$) (Table 2). They were 1,3,7-trimethylurate, 1,7-dimethylurate, 1-methylurate, theophylline, 1,3-dimethylurate, quinate, and theophylline. Among them, quinate belongs to xenobiotics pathway, and trigonelline (N'-methylnicotinate) belongs to carbohydrates pathway. Both quinate and trigonelline (N'-methylnicotinate) are serum dietary biomarkers that strongly associated with coffee intake (33). Interestingly, 5 of the metabolites, 1,3,7-trimethylurate, 1,7-dimethylurate, 1-methylurate, theophylline, 1,3-dimethylurate, also belong to caffeine metabolism. In the current study, as much as 7 significant metabolites in caffeine metabolism pathway showed significantly higher levels in NSCLC patients with poor survival than those with good survival, suggesting the increased activity of caffeine metabolism pathway in NSCLC patients with poor survival compared to ones with good survival (Table 2, Fig 2A, 2B).

Pathway analysis of significant metabolites

We conducted pathway analysis for the 11 significant metabolites identified in the discovery and validation cohorts by using MetaboAnalyst 3.0. This pathway analysis tool uses the high-quality KEGG metabolic pathways as the backend knowledge base and integrates powerful pathway enrichment analysis and pathway topology analysis. Specifically, pathway topology analysis is utilized to analyze the impact of one metabolite or a group of metabolites in a specific pathway. First, we matched these 11 metabolites with the compounds contained in the pathway analysis. Because methyl glucopyranoside (alpha +

beta) did not find any match, it was excluded from subsequent pathway analysis. All other 10 metabolites found exact match (data not shown). Stachydrine matched proline betaine in the analysis. A total of 3 pathways were identified: caffeine metabolism pathway, phenylalanine, tyrosine and tryptophan biosynthesis pathway, and nicotinate and nicotinamide metabolism pathway (Table 3). However, caffeine metabolism pathway is the only significant pathway ($P=1.48E-07$) with high impact factor of 0.36. The other 2 pathways, phenylalanine, tyrosine and tryptophan biosynthesis pathway and nicotinate and nicotinamide metabolism pathway either showed borderline significant difference or did not reach significance level ($P = 0.076$ and 0.121 , respectively). The highlighted metabolite in phenylalanine, tyrosine and tryptophan biosynthesis pathway is quinate, which belongs to Xenobiotics superpathway/Food Component/Plant subpathway in metabolomic profiling. Even though as much as 3 metabolites in Xenobiotics superpathway/Food Component/Plant subpathway were identified in metabolomic profiling, this pathway was not identified in MetaboAnalyst pathway analysis. It is likely that MetaboAnalyst and Metabolon use different bioinformatics tools to define metabolic pathways. Next, we explored whether the relationships between serum metabolites and the survival outcome differed by tumor type. We found that levels of quinate were significantly higher in lung adenocarcinoma patients with poor survival than those with good survival in both discovery and validation phases ($P = 0.038$ and 0.022 , respectively) (Supplemental Table 3).

Association of 4 metabolites with good overall survival

Next, we conducted Cox proportional hazard regression model to assess the association of each metabolite with survival. We found 4 metabolites (1-oleoylglycero phosphoethanolamine, o-cresol sulfate, palmitoyl ethanolamide, 1- palmitoylglycero phosphoethanolamine) whose levels were significantly associated with good survival in both discovery and validation phases (HR: 0.47, 95% CI: 0.31 – 0.72, $P = 0.0005$; HR: 0.54, 95% CI: 0.34 – 0.84, $P = 0.007$; HR: 0.63, 95% CI: 0.41 -0.97, $P = 0.034$; HR: 0.68, 95% CI: 0.43 – 1.05, $P = 0.08$ for discovery phase, respectively, and HR: 0.50, 95% CI: 0.32 – 0.77, $P = 0.002$; HR: 0.64, 95% CI: 0.41 -0.99, $P = 0.047$; HR: 0.58, 95% CI: 0.37 – 0.90, $P = 0.016$; HR: 0.31, 95% CI: 0.18 – 0.52, $P = 1.21E-05$ for validation phase, respectively) (Supplemental Table 4). Among them, 1-oleoylglycero phosphoethanolamine and 1- palmitoylglycero phosphoethanolamine belong to Lysolipid subpathway, o-cresol sulfate belongs to Phenylalanine and Tyrosine Metabolism subpathway, and palmitoyl ethanolamide belongs to Endocannabinoid subpathway. Finally, Kaplan Meier survival analysis was conducted to assess the effects of these 4 metabolites on overall survival, either singly (Fig. 3A) or jointly using a weighted risk score (Fig. 3B), in both discovery and validation phases. Fig. 3A shows that patients with low levels of 1-palmitoylglycerophosphoethanolamine and 1-oleoylglycero phosphoethanolamine had significantly decreased overall survival compared to patients with high levels (low expression group: MST = 8 months and 7 months, respectively, vs. high expression group: MST = 17 months and 18 months, respectively, for discovery phase; and low expression group: MST = 7 months and 11 months, respectively, vs. high expression group: MST = 17 months and 27 months, respectively, for validation phase; all log rank P s < 0.05). We also combined all 4 prognostic metabolites to develop a weighted risk score for survival. As indicated by Fig. 3B, subjects with high risk score displayed significantly worse overall survival compared to those with low risk score in both

discovery and validation phases (low risk score: MST = 7 and 11 months, respectively, vs. high risk score: MST = 17 and 28 months, respectively; log rank P s < 0.05). Our study suggested patients with higher levels of these 4 metabolites may benefit from the first-line platinum-based chemotherapy, and have better survival.

Cumulative effects

Interestingly, we observed significant cumulative effects of these 4 significant metabolites (1-oleoylglycero phosphoethanolamine, o-cresol sulfate, palmitoyl ethanolamide, 1-palmitoylglycero phosphoethanolamine) on overall survival. Patients with greater number of these metabolites expressed at lower levels (low level metabolites [LLMs]) (< median cutoff in controls) had much worse overall survival compared to patients with fewer or no LLM in both discovery and validation phases (Supplemental Figure 1). Compared with subjects without any LLM, patients with 1 or 2 LMs had a 3.4-fold and 3.1-fold increased risk of death, respectively, for discovery and validation cohorts. Patients with 3 or 4 LLMs had a combined 4.8-fold and 5.5-fold increased risk of death, for discovery and validation cohort, respectively (discovery cohort: HR:3.36, 95% CI: 1.62–6.95, P = 0.001, for patients with 1 or 2 LLMs; HR: 4.77, 95% CI: 2.24–10.15, P = 5.09E-05, for patients with 3 or 4 LLMs; validation cohort: HR: 3.06, 95% CI: 1.52–6.19, P = 0.002, for patients with 1 or 2 LLMs; HR: 5.46, 95% CI: 2.61–11.42, P = 6.57E-06, for patients with 3 or 4 LLMs) (Supplemental Table 5).

4. Discussion

In our current study with 220 advanced-stage lung cancer patients who were treated with platinum-based chemotherapy, we identified a panel of metabolites whose levels differed significantly between NSCLC patients with poor and good survival, and 4 metabolites which were significantly associated with good survival, in both discovery and validation phases. In the pathway analysis, caffeine metabolism pathway was identified as the only significant metabolic pathway in predicting the survival outcome among NSCLC patients.

We found the levels of 7 metabolites of caffeine metabolism pathway differed significantly in NSCLC patients with poor and good survival. Caffeine has been reported to have an antineoplastic effect in experimental animals, induces cell differentiation and inhibits mitosis in tissue cultures. For example, 3 significant metabolites of caffeine metabolism pathway identified in our study, theophylline, caffeine and paraxanthine, have been reported to be inversely associated with colorectal cancer (34). In addition, in women with breast cancer, caffeine and/or coffee intake significantly associated with tumor differentiation which may slow tumor growth (35). There are very few reports about caffeine and lung cancer. Caffeine may mediate DNA-repair inhibition, and women with rapid CYP1A2 activity and slow NAT2 were at higher risk for lung adenocarcinoma (36-7). Even though caffeine has been reported to prevent the p53-dependent apoptosis pathway via inhibition of ATM kinase in leukemia, caffeine induces p53-mediated apoptosis through Bax protein in NSCLC cells, suggesting the varied roles caffeine plays in different cancers (38). How caffeine and the metabolites of caffeine metabolism pathway identified in our study are related to survival

outcome of NSCLC patients in current study are unclear and the underlying mechanism remains to be elucidated.

Interestingly, caffeine exposure has been reported to sensitize tumor cells to ionizing radiation, which are associated with the disruption of DNA damage-responsive cell cycle checkpoints including activation of p34cdc2 kinase (39-40). Identified metabolites of caffeine metabolism pathway in the current study may be related to deregulated cell cycle checkpoints and phosphorylation of p53 in NSCLC patients who received chemoradiation therapy, thereby influencing their survival. How caffeine influences the survival of NSCLC patients treated with chemotherapy requires further study.

Intriguingly, in addition to metabolites of caffeine metabolism pathway, we also found 2 metabolites which strongly associated with coffee intake, quinate and trigonelline(N'-methylnicotinate), which differed significantly in NSCLC patients with poor and good survival (33). Because coffee contained caffeine, metabolites of caffeine metabolism pathway and metabolites related to caffeine intake may share common pathway and indirectly related to each other. There are very few reports about coffee intake and cancer prognosis. Coffee has been reported to favorably affect inflammation, lung function and insulin sensitivity, and subsequently reduce mortality risk (41). There are some reports about coffee intake and cancer risk, and the results are variable (34)(42-43). Especially, it has been reported that coffee intake may be associated with increased risk of lung cancer. How coffee intake and the metabolites associated with coffee intake identified in our study are related to survival outcome of NSCLC patients, and how quinate and trigonelline(N'-methylnicotinate) are related to caffeine metabolism pathway are unclear and warrant further investigation. Our study may provide a clue for the association of caffeine metabolism pathway and coffee intake with the prognosis of NSCLC patients.

In the current study, we also found that the levels of stachydrine and methyl glucopyranoside (alpha + beta), which belong to xenobiotics superpathway/food component/plant subpathway, are significantly lower in NSCLC patients with poor survival compared to ones with good survival. Stachydrin was reported to be a serum dietary biomarker that strongly associated with citrus intake, and methyl glucopyranoside (alpha + beta) is a monosaccharide derived from glucose (44). Stachydrine is reported to have *in vitro* anticancer activity in prostate cancer cell lines, remarkably inhibited the invasive capacity of malignant cancer cells, and is a potent anti-metastatic agent (45). There are no reports of methyl glucopyranoside (alpha + beta)'s role in cancer. How low levels of these two metabolites contributed to the poor survival of NSCLC patients in the current study remains to be determined.

In addition, our study found 4 metabolites (1-oleoylglycero phosphoethanolamine, o-cresol sulfate, palmitoyl ethanolamide, 1-palmitoylglycero phosphoethanolamine) whose levels were significantly associated with good survival in both discovery and validation phases. No published reports have examined the relation of these 4 metabolites with lung cancer risk or outcome, and their roles in lung carcinogenesis are unknown. Nevertheless, one of the compounds, palmitoyl ethanolamide, has reported anti-inflammatory properties suggesting one of the possible protective biological mechanisms (46).

5. Conclusions

To date, there are no prognostic biomarkers that are sensitive and accurate enough to be used in the clinic. We have identified a panel of serum metabolites that are novel, noninvasive potential prognostic biomarkers for advanced-stage NSCLC patients treated with platinum-based chemotherapy. These metabolites may predict overall survival in NSCLC patients and help to identify NSCLC patients who may benefit from platinum-based chemotherapy. In the current study, we used a two-stage study design with global metabolomic profiling in both discovery and validation phases to comprehensively identify prognostic metabolic markers and minimize chance findings. In addition, we restricted our analysis to advanced-stage patients and well matched our patients groups of good and poor survival by age, gender, and blood sample collection time to lessen biases from covariate factors. However, there are also some limitations in our study. For example, potential confounding from dietary factors such as coffee consumption was not taken into consideration in the analysis. In addition, we collected non-fasting blood and used serum samples for discovery phase and plasma samples for the validation phase, which might cause experimental variabilities due to prior diet and variant sample source. However, it has been reported that there is close correlation of metabolic profiling results from serum and plasma samples (47). Furthermore, as expected, different sets of significant metabolites were identified when we performed different analysis (Wilcoxon Rank sum-test vs. Cox proportional hazard regression model). Further validation of our results in large-scale and functional studies are warranted to not only corroborate these findings but also explore their biological mechanisms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Cancer Institute grants P50 CA070907, R01 CA176568, CPRIT RP130502, and MD Anderson institutional support for Center for Translational and Public Health Genomics and Duncan Family Institute for Cancer Prevention and Risk Assessment.

References

1. Derman BA, Mileham KF, Bonomi PD, Batus M, Fidler MJ. Treatment of advanced squamous cell carcinoma of the lung: a review. *Transl Lung Cancer Res.* 2015; 4:524–32. [PubMed: 26629421]
2. Puchades-Carrasco L, Jantus-Lewintre E, Pérez-Rambla C, García-García F, Lucas R, Calabuig S. Serum metabolomic profiling facilitates the non-invasive identification of metabolic biomarkers associated with the onset and progression of non-small cell lung cancer. *Oncotarget.* 2016; 7:12904–16. [PubMed: 26883203]
3. Scheff RJ, Schneider BJ. Non-small-cell lung cancer: treatment of late stage disease: chemotherapeutics and new frontiers. *Semin Intervent Radiol.* 2013; 30:191–8. [PubMed: 24436536]
4. Cho HJ, Kim HR, Park YS, Kim YH, Kim DK, Park SI. Prognostic value of survivin expression in stage III non-small cell lung cancer patients treated with platinum-based therapy. *Surg Oncol.* 2015; 24:329–34. [PubMed: 26690822]
5. Lindquist D, Kvarnbrink S, Henriksson R, Hedman H. LRIG and cancer prognosis. *Acta Oncol.* 2014; 53:1135–42. [PubMed: 25180912]

6. Lu B, Gonzalez A, Massion PP, Shyr Y, Shaktour B, Carbone DP. Nuclear survivin as a biomarker for non-small cell lung cancer. *Br J Cancer*. 2004; 91:537–40. [PubMed: 15266313]
7. Claudino WM, Goncalves PH, di Leo A, Philip PA. Metabolomics in cancer: a bench-to-bedside intersection. *Crit Rev Oncol Hematol*. 2012; 84:1–7. [PubMed: 22429650]
8. Go VL, Nguyen CT, Harris DM, Lee WN. Nutrient-gene interaction: metabolic genotype-phenotype relationship. *J Nutr*. 2005; 135:3016S–3020S. [PubMed: 16317163]
9. Bictash M, Ebbels TM, Chan Q, Loo RL, Yap IK, Brown IJ. Opening up the “Black Box”: metabolic phenotyping and metabolome-wide association studies in epidemiology. *J Clin Epidemiol*. 2010; 63:970–9. [PubMed: 20056386]
10. Asiago VM, Alvarado LZ, Shanaiah N, Gowda GA, Owusu-Sarfo K, Ballas RA. Early detection of recurrent breast cancer using metabolite profiling. *Cancer Res*. 2010; 70:8309–18. [PubMed: 20959483]
11. Bertini I, Cacciatore S, Jensen BV, Schou JV, Johansen JS, Kruhøffer M, et al. Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer. *Cancer Res*. 2012; 72:356–64. [PubMed: 22080567]
12. Davis VW, Schiller DE, Eurich D, Sawyer MB. Urinary metabolomics signature of esophageal cancer and barrett's esophagus. *World J Surg Oncol*. 2012; 10:271. [PubMed: 23241138]
13. Shen J, Yan L, Liu S, Ambrosone CB, Zhao H. Plasma metabolomic profiles in breast cancer patients and healthy controls: by race and tumor receptor subtypes. *Transl Oncol*. 2013; 6:757–65. [PubMed: 24466379]
14. Sitter B, Bathen TF, Singstad TE, Fjøsne HE, Lundgren S, Halgunset J. Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy. *NMR Biomed*. 2010; 23:424–31. [PubMed: 20101607]
15. Slupsky CM, Steed H, Wells TH, Dabbs K, Schepansky A, Capstick V. Urine metabolite analysis offers potential early diagnosis of ovarian and breast cancers. *Clin Cancer Res*. 2010; 16:5835–41. [PubMed: 20956617]
16. Sanchez-Espiridion B, Liang D, Ajani JA, Liang S, Ye Y, Hildebrandt MA. Identification of Serum Markers of Esophageal Adenocarcinoma by Global and Targeted Metabolic Profiling. *Clin Gastroenterol Hepatol*. 2015; 13:1730–1737. [PubMed: 25998788]
17. Bathe OF, Shaykhutdinov R, Kopciuk K, Weljie AM, McKay A, Sutherland FR. Feasibility of identifying pancreatic cancer based on serum metabolomics. *Cancer Epidemiol Biomarkers Prev*. 2011; 20:140–7. [PubMed: 21098649]
18. MacIntyre DA, Jimenez B, Lewintre EJ, Martin CR, Schafer H, Ballesteros CG. Serum metabolome analysis by ¹H-NMR reveals differences between chronic lymphocytic leukaemia molecular subgroups. *Leukemia*. 2010; 24:788–797. [PubMed: 20090781]
19. Tiziani S, Lopes V, Gunther UL. Early stage diagnosis of oral cancer using ¹H NMR-based metabolomics. *Neoplasia*. 2009; 11:269–276. [PubMed: 19242608]
20. Urayama S, Zou W, Brooks K, Tolstikov V. Comprehensive mass spectrometry based metabolic profiling of blood plasma reveals potent discriminatory classifiers of pancreatic cancer. *Rapid communications in mass spectrometry : RCM*. 2010; 24:613–620. [PubMed: 20143319]
21. Guan W, Zhou M, Hampton CY, Benigno BB, Walker LD, Gray A. Ovarian cancer detection from metabolomic liquid chromatography/ mass spectrometry data by support vector machines. *BMC bioinformatics*. 2009; 10:259. [PubMed: 19698113]
22. Zhu J, Djukovic D, Deng L, Gu H, Himmati F, Chiorean EG. Colorectal cancer detection using targeted serum metabolic profiling. *Journal of proteome research*. 2014; 13:4120–30. [PubMed: 25126899]
23. Kim K, Aronov P, Zakharkin SO, Anderson D, Perroud B, Thompson IM. Urine metabolomics analysis for. 2009; 8:558–70.
24. Zhao H, Heimberger AB, Lu Z, Wu X, Song R, Shen J. Metabolomics profiling in plasma samples from glioma patients correlates with tumor phenotypes. *Oncotarget*. 2016; 7:20486–95. [PubMed: 26967252]
25. Wen CP, Zhang F, Liang D, Wen C, Gu J, Ye Y, et al. The ability of bilirubin in identifying smokers with higher risk of lung cancer: a large cohort study in conjunction with global metabolomic profiling. *Clin Cancer Res*. 2015; 21:193–200. [PubMed: 25336700]

26. Maeda J, Higashiyama M, Imaizumi A, Nakayama T, Yamamoto H, Daimon T. Possibility of multivariate function composed of plasma amino acid profiles as a novel screening index for non-small cell lung cancer: a case control study. *BMC Cancer*. 2010; 10:690. [PubMed: 21176209]
27. Hori S, Nishiumi S, Kobayashi K, Shinohara M, Hatakeyama Y, Kotani Y. A metabolomic approach to lung cancer. *Lung Cancer*. 2011; 74:284–92. [PubMed: 21411176]
28. Deja S, Porebska I, Kowal A, Zabek A, Barg W, Pawelczyk K. Metabolomics provide new insights on lung cancer staging and discrimination from chronic obstructive pulmonary disease. *Journal of pharmaceutical and biomedical analysis*. 2014; 100:369–380. [PubMed: 25213261]
29. Chen Y, Ma Z, Min L, Li H, Wang B, Zhong J. Biomarker identification and pathway analysis by serum metabolomics of lung cancer. *Biomed Res Int*. 2015; 2015:183624.
30. Chen Y, Ma Z, Li A, Li H, Wang B, Zhong J. Metabolomic profiling of human serum in lung cancer patients using liquid chromatography/hybrid quadrupole time-of-flight mass spectrometry and gas chromatography/mass spectrometry. *J Cancer Res Clin Oncol*. 2015; 141:705–18. [PubMed: 25293627]
31. Jordan KW, Adkins CB, Su L, Halpern EF, Mark EJ, Christiani DC. Comparison of squamous cell carcinoma and adenocarcinoma of the lung by metabolomic analysis of tissue-serum pairs. *Lung Cancer*. 2010; 68:44–50. [PubMed: 19559498]
32. Lokhov PG, Balashova EE, Voskresenskaya AA, Trifonova OP, Maslov DL, Archakov AI. Mass spectrometric signatures of the blood plasma metabolome for disease diagnostics. *Biomed Rep*. 2016; 4:122–6. [PubMed: 26870348]
33. Guertin KA, Loftfield E, Boca SM, Sampson JN, Moore SC, Xiao Q. Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer. *Am J Clin Nutr*. 2015; 101:1000–11. [PubMed: 25762808]
34. Guertin KA, Loftfield E, Boca SM, Sampson JN, Moore SC, Xiao Q. Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer. *Am J Clin Nutr*. 2015; 101:1000–11. [PubMed: 25762808]
35. Pozner J, Papatestas AE, Fagerstrom R, Schwartz I, Saevitz J, Feinberg M. Association of tumor differentiation with caffeine and coffee intake in women with breast cancer. *Surgery*. 1986; 100:482–8. [PubMed: 3738767]
36. Ogasawara H, Nishio K, Ishida T, Arioka H, Fukuoka K, Saijo N. In vitro enhancement of antitumor activity of a water-soluble duocarmycin derivative, KW-2189, by caffeine-mediated DNA-repair inhibition in human lung cancer cells. *Jpn J Cancer Res*. 1997; 88:1033–7. [PubMed: 9439677]
37. Seow A, Zhao B, Lee EJ, Poh WT, Teh M, Eng P. Cytochrome P4501A2 (CYP1A2) activity and lung cancer risk: a preliminary study among Chinese women in Singapore. *Carcinogenesis*. 2001; 22:673–7. [PubMed: 11285205]
38. Ito K, Nakazato T, Miyakawa Y, Yamato K, Ikeda Y, Kizaki M. Caffeine induces G2/M arrest and apoptosis via a novel p53-dependent pathway in NB4 promyelocytic leukemia cells. *J Cell Physiol*. 2003; 196:276–83. [PubMed: 12811820]
39. Sarkaria JN, Busby EC, Tibbetts RS, Roos P, Taya Y, Karnits LM. Inhibition of ATM and ATR kinase activities by the radiosensitizing agent, caffeine. *Cancer Res*. 1999; 59:4375–82. [PubMed: 10485486]
40. Yao SL, Akhtar AJ, McKenna KA, Bedi GC, Sidransky D, Mabry M. Selective radiosensitization of p53-deficient cells by caffeine mediated activation of p34cdc2 kinase. *Nature Med*. 1996; 2:1140–3. [PubMed: 8837615]
41. Loftfield E, Freedman ND, Graubard BI, Guertin KA, Black A, Huang WY. Association of Coffee Consumption With Overall and Cause-Specific Mortality in a Large US Prospective Cohort Study. *Am J Epidemiol*. 2015; 182:1010–22. [PubMed: 26614599]
42. Hashibe M, Galeone C, Buys SS, Gren L, Boffetta P, Zhang ZF. Coffee, tea, caffeine intake, and the risk of cancer in the PLCO cohort. *Br J Cancer*. 2015; 113:809–16. [PubMed: 26291054]

43. Guertin KA, Freedman ND, Loftfield E, Graubard BI, Caporaso NE, Sinha R. Coffee consumption and incidence of lung cancer in the NIH-AARP Diet and Health Study. *Int J Epidemiol.* 2016; 45:929–39. [PubMed: 26082405]
44. Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. *Am J Clin Nutr.* 2014; 100:208–17. [PubMed: 24740205]
45. Rathee P, Rathee D, Rathee D, Rathee S. In vitro anticancer activity of stachydrine isolated from *Capparis decidua* on prostate cancer cell lines. *Nat Prod Res.* 2012; 26:1737–40. [PubMed: 21988653]
46. Impellizzeri D, Ahmad A, Bruschetta G, Di Paola R, Crupi R, Paterniti I. The anti-inflammatory effects of palmitoylethanolamide (PEA) on endotoxin-induced uveitis in rats. *Eur J Pharmacol.* 2015; 761:28–35. [PubMed: 25934566]
47. Yu Z, Kastenmüller G, He Y, Belcredi P, Möller G, Prehn C. Differences between human plasma and serum metabolite profiles. *PLoS One.* 2011; 6:e21230. [PubMed: 21760889]

Highlights

- A two-stage, global metabolite profiling was conducted in lung cancer patients
- Four metabolites, including caffeine, were associated with overall survival.
- Majority of significant metabolites were involved in caffeine metabolism pathway.

Figure 1.

Overall study design. *metabolites observed in both discovery and validation cohorts after quality control analyses. #metabolites whose levels significantly differed between the poor and good survival groups ($P < 0.05$) identified in discovery and validation cohort.

&metabolites whose levels differed significantly between the poor and good survival groups ($P < 0.01$) identified in discovery and validation cohort. \$metabolites whose levels differed significantly between the poor and good survival groups in both cohorts.

Figure 2.

The circulating levels of metabolites of caffeine metabolism pathway were significantly higher in NSCLC patients with poor survival compared to ones with good survival. **A)** Caffeine metabolism pathway. The level of metabolites in red is significantly higher in NSCLC patients with poor survival compared to good survival group in both discovery and validation stages ($P < 0.05$). The level of metabolite in blue is significantly higher ($P < 0.05$) in poor survival group compared to good survival in one stage but borderline significantly higher ($P < 0.10$) in another stage. **B)** Relative circulating levels of seven metabolites of caffeine metabolism pathway which show significantly higher levels (caffeine and paraxanthine) ($P < 0.05$) and borderline significantly higher levels (1,3,7-trimethylurate, 1,7-dimethylurate, 1-methylurate, theophylline, 1,3-dimethylurate) ($P < 0.10$) in NSCLC patients with poor survival (1) compared to good survival group (0). NSCLC, non-small cell lung cancer.

Figure 3. Kaplan–Meier survival estimates of 2 significant metabolites and overall survival in advanced NSCLC patients treated with platinum-based chemotherapy: **A)** 1-palmitoylglycerophosphoethanolamine; **B)** 1-oleoylglycerophosphoethanolamine. Low and High groups were stratified by median value. **C)** Kaplan–Meier survival estimates of patients stratified by risk score (RS) derived from 4 metabolites (i.e. 1-oleoylglycerophosphoethanolamine, o-cresol sulfate, palmitoyl ethanolamide, 1-palmitoylglycerophosphoethanolamine) and overall survival in discovery and validation phases. Low- and

High-RS groups were stratified by median value. MST, median survival time; NSCLC, non-small cell lung cancer.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1
Demographic and clinical characteristics of the patient cohort

Variables	Discovery (n=110)			Validation (n=110)		
	Poor survival, n(%)	Good survival, n(%)	P value	Poor survival, n(%)	Good survival, n(%)	P value
Age, mean(SD)	61.35(10.36)	60.25(6.85)	0.516	60.24(9.41)	60.05(8.92)	0.917
Mean blood setup time, years (SD)	8.74(1.97)	9.04(1.77)	0.403	11.11(3.25)	11.09(3.22)	0.977
Sex						
Male	30(54.55)	30(54.55)		32(58.18)	32(58.18)	
Female	25(45.45)	25(45.45)	1.000	23(41.82)	23(41.82)	1.000
Smoking Status						
Never	5(9.09)	17(30.91)		5(9.09)	9(16.36)	
Former	24(43.64)	20(36.36)		22(40.00)	22(40.00)	
Current	26(47.27)	18(32.73)	0.015	28(50.91)	24(43.64)	0.484
Type						
Adenocarcinoma	37(67.27)	39(70.91)		25(45.45)	29(52.73)	
Squamous cell carcinoma	7(12.73)	7(12.73)		9(16.36)	10(18.18)	
Other	11(20.00)	9(16.36)	0.881	21(38.18)	16(29.09)	0.599
Grade						
Well-differentiated	2(3.64)	3(5.45)		1(1.82)	3(5.45)	
Moderately-differentiated	2(3.64)	5(9.09)		4(7.27)	8(14.55)	
poorly-differentiated/undifferentiated	21(38.18)	21(38.18)		31(56.36)	23(41.82)	
Unknown	30(54.55)	26(47.27)	0.621	19(34.55)	21(38.18)	0.306
Performance Status						
0	5(9.09)	16(29.09)		11(20.00)	18(32.73)	
1	34(61.82)	36(65.45)		30(54.55)	27(49.09)	
2-3	13(23.64)	1(1.82)		10(18.18)	4(7.27)	
Unknown	3(5.45)	2(3.64)	0.001	4(7.27)	6(10.91)	0.186
Stage						
III				20(36.36)	30(54.55)	
IV	55(100)	55(100)	NA	35(63.64)	25(45.45)	0.056
Treatment						

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Variables	Discovery (n=110)		Validation (n=110)		P value
	Poor survival, n(%)	Good survival, n(%)	Poor survival, n(%)	Good survival, n(%)	
Chemotherapy only (platinum based)	48(87.27)	55(100.00)	35(63.64)	24(43.64)	
Chemoradiation	7(12.73)	0	20(36.36)	31(56.36)	0.035

Table 2
List of top 11 significant metabolites that differed between NSCLC patients with good and poor survival

Metabolite	Discovery			Validation			P value	Superpathway/Subpathway	HMDB	KEGG
	Poor survival (N=55)	Good survival (N=55)	P	Poor survival (N=55)	Good survival (N=55)	P				
P<0.05 in both discovery and validation phases										
Caffeine	28.66(1.90)	27.89(2.02)	0.010	28.86(1.36)	28.06(2.17)	0.037	Caffeine metabolism/Xanthine Metabolism	HMDB01847	C07481	
Paraxanthine	26.41(1.86)	25.73(1.69)	0.015	26.60(1.55)	25.79(2.11)	0.023	Caffeine metabolism/Xanthine Metabolism	HMDB01860	C13747	
Stachydine	27.10(2.31)	28.06(2.22)	0.023	27.83(2.31)	28.91(1.90)	0.014	Xenobiotics/Food Component/Plant	HMDB04827	C10172	
Methyl glucopyranoside (alpha + beta)	18.65(2.17)	19.42(1.89)	0.049	19.85(1.61)	20.49(1.69)	0.034	Xenobiotics/ Food Component/Plant	-	-	
P<0.10 in both discovery and validation phases										
1,3,7-trimethylurate	19.07(1.67)	18.22(1.85)	0.007	21.87(1.49)	21.30(1.82)	0.080	Caffeine metabolism/Xanthine Metabolism	HMDB02123	C16361	
Quinate	24.43(2.33)	23.40(2.34)	0.013	25.17(1.85)	24.25(2.55)	0.076	Xenobiotics/Food Component/Plant	METPA0024	C00296	
Trigonelline (N' methylnicotinate)	25.66(1.85)	24.85(1.99)	0.022	25.91(1.33)	25.26(1.86)	0.082	Carbohydrates/Nicotinate and Nicotinamide Metabolism	HMDB00875	C01004	
1,7-dimethylurate	22.21(1.37)	21.65(1.53)	0.051	22.49(1.50)	21.89(1.73)	0.015	Caffeine metabolism/Xanthine Metabolism	HMDB11103	C16356	
1-methylurate	21.82(1.26)	21.24(1.36)	0.057	21.04(1.46)	20.49(1.61)	0.047	Caffeine metabolism/Xanthine Metabolism	HMDB03099	C16359	
Theophylline	24.90(1.45)	24.46(1.42)	0.06	24.72(1.76)	24.03(2.09)	0.007	Caffeine metabolism/Xanthine Metabolism	HMDB01889	C07130	
1,3-dimethylurate	19.42(1.38)	18.74(1.76)	0.076	18.77(1.59)	18.14(1.94)	0.027	Caffeine metabolism/Xanthine Metabolism	HMDB01857	-	

Table 3
Top metabolic pathway associated with poor survival among NSCLC patients

Pathway Name	Total	Hits	P value	-log(p)	Holm P	FDR	Impact
Caffeine metabolism	21	4	1.48×10^{-7}	15.729	1.18×10^{-5}	1.18×10^{-5}	0.3628
Phenylalanine, tyrosine and tryptophan biosynthesis	27	1	0.0760	2.5768	1	1	0
Nicotinate and nicotinamide metabolism	44	1	0.1213	2.1095	1	1	0