Published in final edited form as:

Clin Gastroenterol Hepatol. 2015 October ; 13(10): 1730–1737.e9. doi:10.1016/j.cgh.2015.05.023.

Identification of Serum Markers of Esophageal Adenocarcinoma by Global and Targeted Metabolic Profiling

Beatriz Sanchez-Espiridion^{1,*}, Dong Liang^{2,*}, Jaffer A. Ajani³, Su Liang², Yuanqing Ye¹, Michelle TA Hildebrandt¹, Jian Gu¹, and Xifeng Wu¹

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

²Department of Pharmaceutical Sciences, Texas Southern University, Houston, TX

³Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX

Abstract

BACKGROUND & AIMS—We aimed to identify new serum biomarkers of esophageal adenocarcinoma (EAC).

METHODS—We performed metabolomic analyses of serum samples from 30 patients with histologically confirmed EAC (cases) from The University of Texas MD Anderson Cancer Center and 30 patients without EAC (controls). We identified metabolites whose levels differed significantly between cases and controls and validated those with the greatest difference in an analysis of 321 EAC cases and 331 controls. We generated a metabolite risk score (MRS) for the metabolites.

RESULTS—The levels of 64 metabolites differed significantly between EAC cases and controls. The metabolites with the greatest difference were: amino acid L-proline (LP), ketone body 3-hydroxybutyrate (BHBA), and carbohydrate D-mannose (DM) different; these differences were confirmed in the validation set. Cases had lower mean levels of LP than controls (22.78±6.79 ug/ml vs 28.24±8.64 ug/ml; $P<.001$) and higher levels of BHBA (18.06±17.84 ug/ml vs 7.73±9.92 ug/ml; $P<.001$) and DM (9.87±4.28 ug/ml vs 6.28 ±3.61 ug/ml; $P<.001$). Levels of DM were significantly higher in patients with late-stage EAC than early-stage EAC (10.61±4.79 ug/ml vs 8.97±3.36 ug/ml; $P=.005$). Higher levels of LP were associated with a significant decreased in

Correspondence: Xifeng Wu, Department of Epidemiology, Unit 1340, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-745-2485; Fax: 713-792-4657; xwu@mdanderson.org.

*Authors contributed equally to this work

Disclosures: The authors declare no potential conflicts.

Writing assistance: There was not writing assistance from outside source.

Author contribution: Study concept and design: Wu X., Gu J. Sanchez-Espiridion B; Acquisition of data: Liang D, Sanchez-Espiridion B.; Analysis and interpretation of data: Wu X, Ye Y, Liang D, Sanchez-Espiridion B ; Drafting of the manuscript: Wu X, Gu J, Sanchez-Espiridion B; Critical revision of the manuscript for important intellectual content; Wu X, Gu J, Hildebrandt TA M, Ajani A J ; Statistical analysis: Ye Y; Study supervision: Wu X, Gu J.

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.

risk of EAC (odds ratio [OR] =0.26; 95% confidence interval [CI], 0.18–0.38). A significant increase in risk of EAC was associated with higher levels of BHBA (OR=4.05; 95% CI, 2.84–5.78) and DM (OR=7.04; 95% CI, 4.79–10.34). Levels of all 3 metabolites associated with EAC risk in quartile analyses; the level of risk conferred by the metabolites increased with smoking status and body mass index. Individuals with a high MRS had a significant (7.76-fold) increase in risk of EAC vs those with low a MRS. Smokers with a high MRS had the greatest risk of EAC (OR=20.26; 95% CI, 11.19–36.68), compared with never smokers with a low MRS.

CONCLUSIONS—Based on a case vs control metabolic profile analysis, levels of LP, BHBA and DM are associated with risk of EAC. These markers might be used as prognostic factors for patients with EAC.

Keywords

Metabolomics; Esophageal adenocarcinoma; Esophageal cancer; Serum Biomarkers

Introduction

Esophageal cancer (EC) is the eighth most common cancer in the world, with approximately 300,000 new diagnosed cases each year. In United States, 18,170 new cases and 15,450 deaths were expected in 2014¹. There are two major subtypes, esophageal adenocarcinoma (EAC) and squamous cell carcinoma (SCC). In the Western world, EAC has the highest increasing incidence among solid tumors, accounting for more than 80% of the newly diagnosed cases in United States. Despite improvement in diagnosis, the prognosis and survival of patients remains poor (5-year survival rates remain around 15%). Development of new risk prediction and early detection tools, especially biomarkers with high sensitivity and specificity, are in demand.

Metabolomics is the systematic study of the unique chemical fingerprints generated by metabolic processes of an organism. This technology, based on combining metabolic profiling techniques and multivariate statistical approaches, has been shown to accurately quantify global changes in metabolic profiles of individuals in response to disease or treatment using biofluids (serum, plasma, and urine)². Novel early diagnostic biomarkers using metabolomics have been found in several cancer types and this approach has also been used for identifying biomarkers associated with prognosis, recurrence³, treatment response, and toxicity⁴, supporting the potential of metabolomics as an early diagnostic and prognostic tool for cancer.

To date, metabolomic studies on EC included profiling serum and urine samples to identify differential metabolite markers between patients and controls^{5,6} and potentially useful metabolic profiles for diagnosis of either SCC⁷ and EAC⁸. However, no external validation was included and no insight into biochemical processes altered in EC was provided. We hypothesized that metabolomic profiling may significantly improve the chances for the discovery of cancer-related biomarkers. Therefore, this work aimed to conduct a comprehensive metabolomics study in a large case-control dataset, using a two-stage approach to search for EAC-specific metabolomic signatures, and to identify potential novel

biomarkers for EAC risk stratification. We also assessed whether the identified metabolites may interact with smoking and body mass index (BMI), to modify EAC risk.

Methods

Study design

This is a retrospective study in which two sets of case-control pairs, each consisting of 30 cases and 30 controls, were included in discovery phase for global metabolomic profiling and screening. Validation phase comprised 321 EAC cases and 331 controls. Cases and controls in both phases were frequency-matched on age and gender and were all Caucasians. Cases with histologically confirmed EAC and no prior chemotherapy or radiotherapy were recruited between January 2004 and December 2014 from The University of Texas MD Anderson Cancer Center through a daily review of computerized appointment schedules. There was no age, gender, or clinical stage restrictions. Controls were selected from general population, who had no prior history of cancer except non melanoma skin cancer and matched to cases by age, gender, ethnicity, and residence. The controls were identified and recruited using the random digit dialing method. The same questionnaire data were collected from the cases and controls, which included demographic and epidemiologic risk factor data for EAC, including smoking, obesity, and medical history of GERD and Barrett's esophagus. BMI (kg/m²) was categorized according to the standard classification of the World Health Organization (non-obese or normal/overweight<30; obese ≥30). An individual who never smoked or smoked no more than 100 cigarettes in his or her lifetime was considered a never smoker. An individual who smoked at least 100 cigarettes in his or her life time was defined as an ever smoker. Individuals with stage I, IIa and IIb were classified as early-stages whereas individuals with stage III, IVa and IVb were categorized as late-stages. All study participants signed a written informed consent, and the study was approved by the Institutional Review Boards of The University of Texas MD Anderson Cancer Center (Houston, TX).

Global metabolic profiling and individual targeted metabolomic validation

Global metabolomic profiling analyses were conducted with 200 µl of serum samples at Metabolon Inc. (Durham, NC) as previously described⁹. Individual targeted metabolomic validation was carried out for the three individual selected metabolites in validation dataset using 100 µl of serum for each metabolite by using validated LC-MS/MS methods at Texas Southern University. The LC-MS/MS assays were validated according to FDA's requirements described in its 'Guidance for Industry: Bioanalytical Method Validation'¹⁰. Our LC-MS/MS assay met FDA's validation requirements. In addition, other criteria such as selectivity, recovery, matrix effect, stability, and reproducibility of the assay were also well demonstrated¹¹.

Pathway analyses and individual metabolite selection

For pathway identification of the altered metabolites, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database¹² was used. EAC-associated metabolites were mapped to metabolic maps using their compound IDs, followed by identification of the anabolic and catabolic enzymes in the mapped pathways. To delineate the overall biological processes

altered in EAC, we examined whether EAC-associated metabolites and their corresponding mapped genes by using the Enrich analysis program¹³. We then used a pathway-based selection approach for identifying top metabolic markers. Three of the top differentially expressed metabolites in both discovery datasets were selected and analyzed by targeted metabolomic profiling in the validation phase based on their biological relevance, differential expression levels, and validated assay availability.

Statistical Methods

Statistical analyses were performed using Stata (version 10.1; StataCorp). Differences in distribution of the host characteristics between cases and controls were evaluated by Pearson χ^2 test for categorical variables whereas Student *t*-test was used to test differences for continuous variables.

For global metabolomic profiling analysis of the case-control series in discovery phase, data were log2 transformed. Only metabolites with detectable expression in at least 80% of samples were kept for analysis. A total number of 330 metabolites from Discovery I dataset and 341 from Discovery II dataset were included for analyses. Wilcoxon-rank sum test was conducted to test the difference between cases and control as overall mean values and in stratified analyses by sex, age (<63 vs ≥63), smoking status (Never vs. Ever), BMI(<30 vs ≥30), and to test the difference between late and early stage cases. Association between LP, BHBA and DM with EAC risk was estimated by logistic regression analyses in validation dataset to determine the odds ratio (OR) and 95% confidence interval (95% CI) adjusting for age, gender, smoking status, BMI and batch effect.

Receiver operating characteristic (ROC) curves were used to determine the value with the highest sensitivity and specificity of individual metabolites to be used as cutoff point for dichotomizing patients. Individual metabolites were also analyzed as categorical variables by setting cutoff points at quartile values in control group. In addition, stratified analyses according to age, gender, smoking status and BMI were performed. We then added an interaction term to the logistic regression models using median as cutoff point for a better data distribution to test the interaction between the metabolites and smoking status and BMI in modulating EAC risk. The combined metabolite risk score (MRS) for each patient was derived by linear combination of the product of reference-normalized level of each of the three metabolites by its logistic regression corresponding coefficient. All cases were dichotomized by the using the ROC method into high or low risk groups, respectively. We also added an interaction term to the logistic regression models to test the interaction between the MRS and BMI and smoking in modulating EAC risk. All statistical tests were two sided, with statistical significance set at $P < 0.005$.

Results

Patient characteristics

We included 30 cases and 30 controls in each of the two discovery phases for screening and identification of promising metabolites. The validation phase consisted of 321 cases and 331 controls who were frequency-matched on age and gender (Table 1). Mean age was

62.99±10.69 years in cases and 62.88±10.33 in controls (P=0.890). Over 90% of cases were males, consistent with the EAC male dominance. As expected, cases had significantly higher proportion of ever smokers than controls (73.89% vs. 51.96%, P<0.001). Among ever smokers, self-reported median pack-years were significantly higher in cases than in controls (6.14% vs. 36.91%, P<0.001). There was no significant difference in neither BMI nor diabetes status between cases and controls. Among cases in the validation set, 146 subjects were early-stage patients while 175 were late-stage.

Differentially expressed metabolites in the discovery phase

A total of 437 metabolites were detected in the discovery phase. After quality control analyses, 331 metabolites (Discovery I) and 341 metabolites (Discovery II) were included for analyses (Supplementary Figure 1). Wilcoxon rank-sum test identified 131 metabolites in Discovery I and 123 metabolites in Discovery II as showing significant differential levels between cases and controls (p<0.005). Taken together, 64 metabolites had consistent differential levels between cases and controls in both discovery batches (P<0.005). Out of these 64 metabolites, 22 had significantly higher levels in cases compared to controls whereas 42 exhibited lower levels (Supplementary Table 1).

Pathway enrichment analyses of EAC-associated metabolites

To explore the relationships between these metabolic biomarkers, their corresponding metabolic pathways were analyzed. These metabolites were mapped to 5 super-pathways (amino acids, carbohydrates, lipids, peptides and xenobiotics) and 27 sub-pathways revealing an increase in amino acid metabolism and changes in carbohydrates, lipids and gamma-glutamyl aminoacids as potential pathways involved in EAC (Supplementary Table 2). A similar enrichment network was identified by the bioinformatics tool Enrich¹³ pointing to alterations in protein biosynthesis, fatty acids, energy and bile acid metabolism as key metabolic EAC-pathways, with Arginine and Proline metabolism identified as the main significantly overrepresented pathway (P<0.001; FDR<0.01) (Supplementary Figure 2).

Individual targeted validation of selected metabolites

Metabolites that exhibited a significant trend in altered levels from normal individuals to early- to late-stage EAC cases are potential early detection biomarkers of EAC. Out of the 64 significant metabolites, we selected three metabolites, an amino acid LP, a ketone body BHBA, and a carbohydrate DM based on their differential expression levels, biological significance, and previous literature reports. We developed validated standard LC/MS-MS assays and measured their levels in 321 EAC cases and 331 matched controls for validation. Consistent with discovery phase, cases had lower levels of LP than controls (mean±SD: 22.78±6.79 vs. 28.24±8.64 ug/ml; P<0.001) and higher levels of BHBA (18.06±17.84 vs. 7.73±9.92 ug/ml; P<0.001) and DM (9.87±4.28 vs. 6.28±3.61 ug/ml; P<0.001) (Table 2). Results of stratified analyses by gender, age, smoking status and BMI were consistent with overall results (Supplementary Table 3).

Association of serum metabolite levels with EAC stage

Since clinical stage is an established EAC prognostic factor EAC, we compared the levels of these 3 metabolites between controls and patients with early-stage and late-stage in validation set. Controls had higher levels of LP and lower levels of BHBA and DM when compared to early-stage or late-stage cases separately ($P<0.001$ for all comparisons).

When comparing early- to late-stage patients, no significantly different levels were found for LP. A borderline significance was found for BHBA with higher levels in late-stage cases when compared to early-stage (mean $\pm$ SD, 19.89 ± 18.97 vs 15.89 ± 16.19 ug/ml, $P=0.071$) and significantly higher levels when compared to controls (mean $\pm$ SD, 19.89 ± 18.97 vs 7.73 ± 9.92 ug/ml, $P<0.001$). Interestingly, levels of DM were significantly higher in late-stage patients when compared to early-stage patients (mean $\pm$ SD, 10.61 ± 4.79 vs 8.97 ± 3.36 ug/ml; $P=0.005$) and controls (10.61 ± 4.79 vs 6.28 ± 3.61 ug/ml; $P<0.001$) (Table 3).

Associations between LP, BHBA and DM and EAC risk

We next analyzed the association between the 3 validated metabolites with EAC risk by unconditional multivariate logistic regression analyses in validation set. Using the ROC method to determine the cutoff point of each metabolite, higher levels of LP were associated with a significant decreased EAC risk (OR=0.26; 95% confidence interval (CI), [0.18–0.38]; $P<0.001$). However, a significantly increased EAC risk was associated with higher levels of either BHBA (OR =4.05, 95% CI, 2.84–5.78) or DM (OR =7.04, 95% CI, 4.79–10.34).

Quartile analyses showed a dose-response relationship. Using the lowest quartile as reference, the ORs for the 2nd, 3rd and 4th quartiles were 0.41 (95% CI, 0.26–0.64), 0.28 (95% CI, 0.18–0.45) and 0.14 (95% CI, 0.08–0.25), respectively (P for trend <0.001) for LP. In contrast, using the lowest quartile as reference, the ORs for the 2nd, 3rd and 4th quartiles were 2.97 (95% CI, 1.55–5.70), 4.06 (95% CI, 2.15–7.70) and 9.45 (95% CI, 5.15–17.34), respectively, (P for trend <0.001) for BHBA, and 2.54 (95% CI, 1.21–5.31), 4.26 (95% CI, 2.11–8.60) and 16.97 (95% CI, 8.63–33.37), respectively, (P for trend <0.001) for DM (Figure 1). Stratified analyses by sex, age, smoking status and BMI showed consistent results with overall analyses (Supplementary Table 4).

We also evaluated the joint effects of these metabolites with smoking status and BMI, two known EAC risk factors. A joint effect between low LP levels and smoking was seen. On the other hand, a joint effect was observed between either both high BHBA and DM levels and smoking. In terms of BMI, a joint effect between low LP levels and BMI (≥ 30) was also seen. A similar joint effect was seen for both high BHBA and DM levels and BMI (Table 4).

Development of a metabolite risk score

The combined effects of the three validated metabolites were investigated by calculating a metabolite risk score (MRS). When using ROC curve to determine the cutoff point, patients with high MRS exhibited a significantly increased EAC risk compared with those with low MRS (OR=7.76, 95% CI, 4.80–12.53; $p<0.001$). A joint effect between high MRS and smoking status was seen, with OR of 11.06 (95% CI, 5.05–24.19) and 4.14 (95% CI, 1.83–9.37) for individuals with one risk factor and OR of 23.40 (95% CI, 10.95–50.00) for

individuals with both risk factors (smoking and high MRS) with a significant interaction (P interaction <0.001) (Figure 2).

Discussion

Diagnostic and therapeutic biomarkers useful for EAC have the potential to facilitate early detection and increase long-term survival of cancer patients. Through a multistage study design, we found significant serum metabolites and pathways associated with EAC. In the discovery phase, 64 metabolites were significantly altered in the serum samples of EAC patients compared to controls, and these 64 metabolites belonged to 5 main metabolic superpathways (amino acids, carbohydrates, lipids, peptides and xenobiotics). Further pathway analyses showed a majority of these metabolites involved in protein biosynthesis, fatty acids, energy and bile acids metabolism.

It has been shown that tumors have the ability to induce metabolic alterations in the host organism involving aspects of the intermediate metabolism, the intracellular process by which nutritive material is converted into cellular materials, including energy, carbohydrate, fat and protein metabolism. Altered metabolic pathways including a number of amino acid pathways and energy metabolism have been already described in EAC¹⁴, in agreement with our results. In addition, differences in circulating bile acids may reflect altered bile acid biosynthesis. Bile acids can regulate fatty acid metabolism; therefore they might play a role in the observed fatty acid changes. As mentioned, an abnormal metabolism has been considered an important common characteristic of tumors and cancer cells¹⁵ and an increased in glycolysis known as the “Warburg effect” has been well characterized. This deregulated energy metabolism of cancer cells may also modify many related metabolic pathways that influence various biological processes, such as cell proliferation and apoptosis.

To validate the potential metabolic biomarkers, an independent validation was performed. Consistent with discovery phase, cases had lower levels of LP and higher levels of BHBA and DM than controls. The non-essential amino acid LP and its role in cancer metabolism has been recently highlighted^{16,17} whereas alterations in ketone bodies (BHBA) have been shown to promote tumor growth and metastasis¹⁸ and altered carbohydrate (DM) levels have been related to cancer development and progression^{19,20}. Similar to our findings, altered LP levels have been found in other cancer types such as renal cell carcinoma²¹, and changes in LP levels have been associated with early detection of recurrence in breast cancer³. Studies done in plasma have reported decreased amino acids levels of patients with malignant tumors²². Decreased amino acids levels have been interpreted as the result of high-demand tumor metabolism associated with early signs of malnutrition in a tumor-bearing state. On the other hand, the ketone BHBA was found to be higher in cases when compared to controls in agreement with previous studies of other cancer types such as hepatocellular carcinoma¹⁹, ovarian cancer²³ and other cancer types including EAC⁵. The significant increase of serum concentrations of BHBA may arise from altered lipolysis linked to high energy demands in the cells. Ketone bodies, including BHBA, are able to replace glucose as an alternative energy source when glucose levels are low. Therefore, increased amounts of BHBA could be associated with lower glucose levels as supported by many cancer studies²⁴.

Mannose, which plays an important role in protein glycosylation, has not been well-documented in cancer. A recent study has shown an elevation of high-mannose glycan's in serum during breast cancer progression²⁰, suggesting an incomplete glycosylation process. In addition, DM has been shown to be elevated in other cancer types such as breast and ovarian cancer and associated with breast cancer metastasis and invasion²⁵, consistent with our observation of higher DM levels in late EAC stages.

We found joint effect of metabolites and smoking on elevating EAC risk. Previous studies have also showed smoking-related changes in human serum metabolize that included mostly amino acids and lipids²⁶. A similar significant joint effect was found between LP, BHBA and DM levels and BMI on EAC risk. Specific metabolites correlated with BMI and changes in the human serum metabolize associated with BMI have been reported²⁷. Furthermore, we evaluated the predictive value of the 3 individually selected metabolites by developing a MRS. EAC cases with high MRS showed a 5.79-fold increased risk when compared to the low risk group. Smokers who had high MRS exhibited the highest EAC risk with a 23.40-fold increased risk when compared to never smokers with low MRS. These data highlight the value of combining biomarkers with environment exposures in predicting the risk of cancer development.

The biological reproducibility might be a concern in metabolomics studies. In our study, we couldn't evaluate the reproducibility of the measurements across time and it should be mentioned as a limitation. However, recent studies have suggested that metabolomic profiles may be relatively stable^{28,29}.

To the best of our knowledge, this is the largest serum metabolomics study of EAC risk and the first one to perform a discovery phase global profiling followed by targeted validation in a large population using different assay platforms in the two phases. A previous metabolomics study of EAC serum specimens using a case-control design also included cases with Barrett's esophagus^{7, 8}. However, that study did not include an external validation and the sample sizes were small. Most studies have been focused on metabolic profiling of SCC cases in Chinese populations.

In summary, we found that serum levels of LP, BHBA and DM were significantly different between EAC cases and healthy controls and these metabolites may become novel biomarkers for the early detection of EAC. Our data support that serum metabolic profiling has the potential to identify novel biomarkers for EAC diagnosis and susceptibility and can also enhance our understanding of EAC development. Our results suggest that LP, BHBA and DM could become novel biomarkers for EAC risk stratification and DM could be utilized for therapy monitoring. Nevertheless, future prospective validation in independent cohorts is warranted to confirm our observations and further investigation is needed to determine the potential clinical utility of these initial findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support. This study was supported by a grant CA111922 from National Cancer Institute, and MD Anderson Cancer Center institutional support for the Center for Translational and Public Health Genomics.

Abbreviations used in this paper

EAC	esophageal adenocarcinoma
LP	L-Proline
BHBA	3-hydroxybutyrate
DM	D-mannose
SD	standard deviation
OR	odds ratio
CI	confidence interval
MRS	metabolite risk score
EC	esophageal cancer
SCC	squamous cell carcinoma
BMI	body mass index
LC-MS/MS	liquid chromatography–mass spectrometry/mass spectrometry
KEGG	Kyoto Encyclopedia of Genes and Genomes
Compound IDs	compound identifications

References

1. Siegel R, Ma J, Zou Z, et al. Cancer statistics, 2014. *CA Cancer J Clin.* 2014; 64:9–29. [PubMed: 24399786]
2. Gowda GA, Zhang S, Gu H, et al. Metabolomics-based methods for early disease diagnostics. *Expert Rev Mol Diagn.* 2008; 8:617–33. [PubMed: 18785810]
3. Asiago VM, Alvarado LZ, Shanaiah N, et al. Early detection of recurrent breast cancer using metabolite profiling. *Cancer Res.* 2010; 70:8309–18. [PubMed: 20959483]
4. Bertini I, Cacciatore S, Jensen BV, et al. Metabolomic nmr fingerprinting to identify and predict survival of patients with metastatic colorectal cancer. *Cancer Res.* 2012; 72:356–64. [PubMed: 22080567]
5. Zhang X, Xu L, Shen J, et al. Metabolic signatures of esophageal cancer: Nmr-based metabolomics and UHPLC-based focused metabolomics of blood serum. *Biochim Biophys Acta.* 2013; 1832:1207–16. [PubMed: 23524237]
6. Davis VW, Schiller DE, Eurich D, et al. Urinary metabolomic signature of esophageal cancer and barretts esophagus. *World J Surg Oncol.* 2012; 10:271. [PubMed: 23241138]
7. Liu R, Peng Y, Li X, et al. Identification of plasma metabolomic profiling for diagnosis of esophageal squamous-cell carcinoma using an UPLC/TOF/MS platform. *Int J Mol Sci.* 2013; 14:8899–911. [PubMed: 23615477]
8. Zhang J, Liu L, Wei S, et al. Metabolomics study of esophageal adenocarcinoma. *J Thorac Cardiovas Surg.* 2011; 141:469–75. 75 e1–4.
9. Lawton KA, Berger A, Mitchell M, et al. Analysis of the adult human plasma metabolome. *Pharmacogenomics.* 2008; 9:383–97. [PubMed: 18384253]

10. FDA. Guidance for Industry: Bioanalytical Method Validation. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; Rockville MD: 2001.
11. Liang S, Sanchez-Espiridion B, Xie H, et al. Determination of proline in human serum by a robust LC-MS/MS method: Application to identification of human metabolites as candidate biomarkers for esophageal cancer early detection and risk stratification. *Biomed Chromatog*. 2014 Aug 28. Epub ahead of print. 10.1002/bmc.3315
12. Kanehisa M. The kegg database. *Novartis Found Symp*. 2002; 247:91–101. discussion -3, 19–28, 244–52. [PubMed: 12539951]
13. Chen EY, Tan CM, Kou Y, et al. Enrichr: Interactive and collaborative html5 gene list enrichment analysis tool. *BMC bioinformatics*. 2013; 14:128. [PubMed: 23586463]
14. Zhang J, Bowers J, Liu L, et al. Esophageal cancer metabolite biomarkers detected by LC-MS and NMR methods. *PloS one*. 2012; 7:e30181. [PubMed: 22291914]
15. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*. 2011; 144:646–74. [PubMed: 21376230]
16. Phang JM, Liu W. Proline metabolism and cancer. *Front Biosci (Landmark Ed)*. 2012; 17:1835–45. [PubMed: 22201839]
17. Phang JM, Liu W, Hancock C, et al. The proline regulatory axis and cancer. *Front Oncol*. 2012; 2:60. [PubMed: 22737668]
18. Martinez-Outschoorn UE, Lin Z, Whitaker-Menezes D, et al. Ketone body utilization drives tumor growth and metastasis. *Cell Cycle*. 2012; 11:3964–71. [PubMed: 23082722]
19. Liu Y, Hong Z, Tan G, Dong X, Yang G, Zhao L, et al. NMR and LC/MS-based global metabolomics to identify serum biomarkers differentiating hepatocellular carcinoma from liver cirrhosis. *Int J Cancer*. 2014; 135(3):658–68. [PubMed: 24382646]
20. de Leoz ML, Young LJ, An HJ, et al. High-mannose glycans are elevated during breast cancer progression. *Mol Cell Proteomics*. 2011; 10:M110 002717. [PubMed: 21097542]
21. Mustafa A, Gupta S, Hudes GR, et al. Serum amino acid levels as a biomarker for renal cell carcinoma. *J Urol*. 2011; 186:1206–12. [PubMed: 21849193]
22. Shen J, Yan L, Liu S, et al. Plasma metabolomic profiles in breast cancer patients and healthy controls: By race and tumor receptor subtypes. *Transl Oncol*. 2013; 6:757–65. [PubMed: 24466379]
23. Fong MY, McDunn J, Kakar SS. Identification of metabolites in the normal ovary and their transformation in primary and metastatic ovarian cancer. *PloS one*. 2011; 6:e19963. [PubMed: 21625518]
24. Rocha CM, Carrola J, Barros AS, et al. Metabolic signatures of lung cancer in biofluids: Nmr-based metabolomics of blood plasma. *J Proteome Res*. 2011; 10:4314–24. [PubMed: 21744875]
25. Jobard E, Pontoizeau C, Blaise BJ, et al. A serum nuclear magnetic resonance-based metabolomic signature of advanced metastatic human breast cancer. *Cancer Lett*. 2014; 343:33–41. [PubMed: 24041867]
26. Xu T, Holzapfel C, Dong X, et al. Effects of smoking and smoking cessation on human serum metabolite profile: Results from the KORA cohort study. *BMC Med*. 2013; 11:60. [PubMed: 23497222]
27. Moore S, Matthews C, Sampson J, et al. Human metabolic correlates of body mass index. *Metabolomics*. 2014; 10:259–69. [PubMed: 25254000]
28. Townsend MK, Clish CB, Kraft P, et al. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. *Clin Chem*. 2013 Nov; 59(11):1657–67. [PubMed: 23897902]
29. Sampson JN, Boca SM, Shu XO, et al. Metabolomics in epidemiology: sources of variability in metabolite measurements and implications. *Cancer Epidemiol Biomarkers Prev*. 2013 Apr; 22(4): 631–40. [PubMed: 23396963]

Figure 1. Adjusted Odds ratio (95% CI) of EAC risk stratified by quartiles of L-Proline levels (A), 3-Hydroxybutyrate (B) and D-Mannose (C).

Figure 2.
Joint effects of Metabolic risk score and smoking status on EAC risk.

Table 1

Host Characteristics in Validation phase

Variables	n (%)		P value
	Cases (n=321)	Controls (n=331)	
Mean age, years (SD)	62.99(10.69)	62.88(10.33)	.890
Pack years. mean (SD) ^a	61.14(63.84)	36.91(44.20)	<.001
BMI, mean (SD)	28.47(5.33)	28.67(5.35)	.640
Sex, n (%)			
Male	293(91.28)	301(90.94)	
Female	28(8.72)	30(9.06)	.879
Smoked status, n (%)			
Never-smoker	82(26.11)	159(48.04)	
Ever-smoker	232(73.89)	172(51.96)	<.001
Pack years. Mean (SD) ^a			
<30	81(37.33)	92(55.09)	
>=30	136(62.67)	75(44.91)	.001
BMI			
<30	192(41.55)	222 (69.91)	
>=30	93(32.75)	96(30.09)	.234
Diabetes			
Yes	37(13.03)	45(13.68)	
No	247(86.97)	284(86.32)	.814

^aEver smokers only; SD: standard deviation; BMI: Body Mass Index

Significant P values in bold font

Table 2

Levels of individual selected metabolites in Discovery and Validation phase.

Metabolite	Cases			Controls		
	N	Mean ^a (SD)		N	Mean ^a (SD)	P value ^c
Discovery I^a						
L-Proline	30	2.21(0.47)		30	3.06(0.86)	<.001
3-Hydroxybutyrate	30	22.91(21.50)		30	7.95(8.18)	<.001
D-Mannose	30	8.06(2.96)		30	5.27(5.29)	<.001
Discovery II^a						
L-Proline	30	4.34(1.02)		30	5.41 (1.18)	.001
3-Hydroxybutyrate	30	41.06(50.74)		30	19.50 (37.99)	.002
D-Mannose	30	20.66(7.02)		30	13.88(5.49)	<.001
Validation^b						
L-Proline	316	22.78 (6.79)		331	28.24 (8.64)	<.001
3-Hydroxybutyrate	309	18.06(17.84)		329	7.73 (9.92)	<.001
D-Mannose	320	9.87 (4.28)		323	6.28 (3.61)	.001

^a Values are relative expression values (x 10⁹) (Metabolon platform)^b Values are absolute mean values in ug/ml^c Wilcoxon rank-sum test; SD: standard deviation; ug/ml, microgram/ milliliter

Significant P values in bold

Table 3

Levels of selected metabolites in controls, early-stage and late -stage EAC cases in validation set

Metabolite	Controls			Cases						P value ^b (Late stage vs Controls)	P value ^b (Early stage vs late stage)
				Early stage			Late stage				
	N	Mean	a(SD)	N	Mean	a(SD)	N	Mean	a(SD)		
L-Proline	331	28.24	(8.64)	145	22.35	(5.89)	171	23.14	(7.48)	<.001	.682
3-Hydroxybutyrate	329	7.73	(9.92)	141	15.89	(16.19)	168	19.89	(18.97)	<.001	.071
D-Mannose	323	6.28	(3.61)	145	8.97	(3.36)	175	10.61	(4.79)	<.001	.005

^a Absolute mean values in ug/ml,

^b Wilcoxon-rank sum test

SD: standard deviation; ug/ml, microgram/milliliter; Significant P values in bold

Table 4
Joint effects of L-Proline, BHBA and D-Mannose with smoking and BMI on EAC risk

Metabolite	Smoking Status	Cases, n(%)	Controls n(%)	OR ^a (95% CI)	P value	BMI	Cases, n (%)	Controls n (%)	OR ^a (95% CI)	P value
L-Proline^b										
>26.40	Never	15(4.82)	88(26.59)	1(reference)		<30	32(11.31)	113(35.42)	1(reference)	
>26.40	Ever	56(18.01)	77(23.26)	4.39 (2.24–8.62)	<.001	<30	33(11.66)	48(15.05)	3.34 (1.58–7.03)	.002
<=26.40	Never	66(21.22)	71(21.45)	5.75 (2.96–11.17)	<.001	>=30	159(56.18)	110(34.48)	4.76 (2.97–7.63)	<.001
<=26.40	Ever	174(55.95)	95(28.70)	11.19(5.96–21.02)	<.001	>=30	59(20.85)	48(15.05)	6.50(3.10–13.61)	<.001
P for trend					<.001					<.001
P interaction					.045					.022
3-Hydroxybutyrate										
<=4.19	Never	17(5.59)	79(24.01)	1(reference)		<30	44(15.94)	115(36.16)	1(reference)	
<=4.19	Ever	51(16.78)	86(26.14)	2.67(1.40–5.08)	.003	<30	20(7.25)	45(14.15)	1.52(0.70–3.30)	.284
>4.19	Never	63(20.72)	80(24.32)	3.51(1.87–6.60)	<.001	>=30	142(51.45)	107(33.65)	3.36(2.17–5.22)	<.001
>4.19	Ever	173(56.91)	84(25.53)	9.28(5.07–16.97)	<.001	>=30	70(25.36)	51(16.04)	5.09(2.55–10.16)	<.001
P for trend					<.001					<.001
P interaction					.978					.985
D-Mannose										
<=5.93	Never	10(3.23)	87(26.93)	1(reference)		<30	32(11.23)	113(36.33)	1(reference)	
<=5.93	Ever	35(11.29)	74(22.91)	4.53 (2.06–9.96)	<.001	<30	13(4.56)	43(13.83)	1.50(0.63–3.57)	.360
>5.93	Never	71(22.90)	69(21.36)	9.90(4.64–21.16)	<.001	>=30	160(56.14)	104(33.44)	5.01(3.12–8.05)	<.001
>5.93	Ever	194(62.58)	93(28.79)	20.47(9.84–42.60)	<.001	>=30	80(28.07)	51(16.40)	9.37(4.45–19.72)	<.001
P for trend					<.001					<.001
P interaction					.085					.626

^a Adjusted by age, gender, smoking status and BMI, SD : standard deviation, BMI: Body Mass index

^b Median values of individual metabolites (ug/mL) determined by the distribution of metabolite levels in controls Significant P values in bold