Published in final edited form as:

Metabolomics. 2017 March ; 13(3): . doi:10.1007/s11306-016-1154-y.

Candidate serum metabolite biomarkers for differentiating gastroesophageal reflux disease, Barrett's esophagus, and high-grade dysplasia/esophageal adenocarcinoma

Matthew F. Buas^{1,2}, Haiwei Gu³, Danijel Djukovic³, Jiangjiang Zhu³, Lynn Onstad¹, Brian J. Reid^{1,4}, Daniel Raftery^{1,3,*}, and Thomas L. Vaughan^{1,5,*}

¹Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, 98109 USA

²Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY, 14263 USA

³Northwest Metabolomics Research Center, Department of Anesthesiology and Pain Medicine, University of Washington School of Medicine, Seattle, WA 98109 USA

⁴Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA 98109 USA

⁵Department of Epidemiology, University of Washington School of Public Health, Seattle, WA, 98109 USA

Abstract

Introduction/Objectives—Incidence of esophageal adenocarcinoma (EA), an often fatal cancer, has increased sharply over recent decades. Several important risk factors (reflux, obesity, smoking) have been identified for EA and its precursor, Barrett's esophagus (BE), but a key challenge remains identifying individuals at highest risk, since most with reflux do not develop BE, and most with BE do not progress to cancer. Metabolomics represents an emerging approach for identifying novel biomarkers associated with cancer development.

Methods—We used targeted liquid chromatography-mass spectrometry (LC-MS) to profile 57 metabolites in 322 serum specimens derived from individuals with gastroesophageal reflux disease

*Correspondence: T.L.V. (tvaughan@u.washington.edu, Phone: 206-667-5134, Fax: 206-667-4787; D.R. (draftery@uw.edu, Phone: 206-685-4753, Fax: 206-616-4819).

Supplemental Data

Supplemental Data include one figure and four tables.

Author contributions

Conception and design: M.F.B., T.L.V., D.R. Subject recruitment: T.L.V., B.J.R. Data acquisition: H.G., D.D., J.Z. Analysis and interpretation of data: M.F.B., H.G., D.D., L.O., D.R., T.L.V. Drafting of the manuscript: M.F.B., T.L.V., H.G., D.R. Study supervision: D.R., T.L.V. All authors critically revised the manuscript for intellectual content.

Conflict of Interest Statement:

D.R. holds equity and an executive role in Matrix-Bio, Inc. (IN, USA). The authors have no other potential conflicts of interest to disclose.

Compliance with Ethical Requirements:

All authors complied with ethical policies as described. Potential conflicts of interests were disclosed. This study included human participants and was approved by the Institutional Review Board at the Fred Hutchinson Cancer Research Center. Appropriate informed consent was obtained.

(GERD), BE, high-grade dysplasia (HGD), or EA, drawn from two well-annotated epidemiologic parent studies.

Results—Multiple metabolites differed significantly ($P<0.05$) between BE versus GERD ($n=9$), and between HGD/EA versus BE ($n=4$). Several top candidates (FDR $q \leq 0.15$), including urate, homocysteine, and 3-nitrotyrosine, are linked to inflammatory processes, which may contribute to BE/EA pathogenesis. Multivariate modeling achieved moderate discrimination between HGD/EA and BE (AUC=0.75), with less pronounced separation for BE versus GERD (AUC=0.64).

Conclusion—Serum metabolite differences can be detected between individuals with GERD versus BE, and between those with BE versus HGD/EA, and may help differentiate patients at different stages of progression to EA.

Keywords

Barrett's esophagus; esophageal adenocarcinoma; gastroesophageal reflux disease; metabolomics; serum biomarkers; risk stratification

Introduction

Incidence of esophageal adenocarcinoma (EA) has increased over 8-fold in recent decades, while median survival remains less than one year (1). Most EAs develop in a metaplastic epithelium termed Barrett's esophagus (BE), which can arise in the setting of chronic gastroesophageal reflux (GERD). Besides recurrent GERD symptoms, which occur in 20–30% of US adults, the major known risk factors are white race, male sex, abdominal obesity, and smoking (2–4). The prevailing clinical paradigm is to identify BE patients, enroll them in long-term surveillance, and aggressively treat them before they develop invasive cancer. However, 95% of BE patients will never progress to cancer, while only ~5% of EA cases are diagnosed as a result of screening and surveillance. Upper endoscopy with biopsy, the standard method of diagnosing and monitoring BE, is both invasive and expensive, and its frequent yet variable implementation has not appreciably increased early detection or decreased EA mortality (1). Development of new and non-invasive methods for stratification and assessment of patients at risk of EA would address important unmet clinical needs (1, 4).

Metabolomics has emerged as a promising domain for identifying molecular signatures associated with cancer (5–10). In contrast to genomics, transcriptomics, or proteomics, metabolomics describes the study of concentrations and fluxes of low-molecular-weight metabolites present in biofluids or tissue (11, 12). As metabolic fluctuations lie downstream of alterations at the DNA, RNA, and protein level, metabolomics offers a sensitive and comprehensive functional read-out of biological systems. Metabolite levels and metabolic pathways are significantly altered in a broad spectrum of human cancers (13–16). Adaptive changes such as increased rates of aerobic glycolysis (the Warburg effect), amino acid metabolism, and lipid turnover are thought to be important features of the neoplastic process that facilitate sustained cellular proliferation and tumor expansion (15, 16).

Several past metabolomics studies, generally of modest sample size, have focused on esophageal adenocarcinoma (17–25). In one of the earlier reports, ^1H magic angle spinning-

nuclear magnetic resonance (NMR) spectroscopy-based profiling of intact esophageal tissue samples identified a panel of metabolites that separated control tissues from tumor tissues and proximal histologically normal mucosa from individuals with EA (23). In a second study, Raftery and colleagues conducted a targeted mass-spectrometry (MS)-based analysis of eight nucleosides in serum samples from EA cases or healthy controls, and identified five as differentially abundant (19). Serum metabolites identified using NMR and LC-MS were combined to generate a signature capable of accurately discriminating between EA patients and healthy controls (24). Urinary metabolite panels identified by ¹H NMR spectroscopy have also facilitated separation of healthy controls from patients with BE or those with EA or esophageal squamous cell carcinoma (ESCC) (18). Most recently, serum levels of the amino acid L-proline, ketone body 3-hydroxybutyrate, and carbohydrate D-mannose were found to differ between EA cases versus healthy controls (21). Analyses in cell lines and isolated tissues specimens have described alterations in both energy metabolism and metabolically-relevant gene expression patterns across the disease continuum from BE to EA (26, 27). Collectively, while this overall body of work has suggested that metabolic perturbations are associated with BE/EA, very limited agreement between studies has been observed with respect to specific metabolite alterations, likely reflecting significant heterogeneity in study design, sample collection/processing, analytic platforms, and statistical approaches, as well as the confounding effects of modifying factors (age, race, smoking, etc).

In this discovery study, we used an optimized LC-MS targeted profiling platform to screen for differentially abundant serum metabolites among individuals with i) BE versus GERD, and ii) HGD/early-stage EA versus BE. We thus focused on clinically-relevant comparisons, rather than analyzing readily distinguishable groups such as cancer cases and healthy controls. A total of 322 serum samples (100 GERD, 122 BE, 100 HGD/EA) were selected from two prior epidemiologic studies of BE and EA and subjected to targeted metabolomic analysis of 57 metabolites, which collectively encompass several major metabolic pathways (glycolysis/TCA cycle, amino acid, nucleotide, and fatty acid metabolism). The potential utility of derived metabolite profiles to discriminate between these disease states was investigated through construction and evaluation of multivariate classification models.

Materials and Methods

Study design

Participants in this study were selected from individuals enrolled in one of two independent parent studies: i) a community clinic-based case-control study among patients with GERD (Study of Reflux Disease, SRD) (28, 29), and ii) a prospective longitudinal cohort study of BE patients (Seattle BE Project, SBEP) (30). Briefly, subjects from the SRD were residents of western Washington without a previous diagnosis of BE who underwent a baseline endoscopy for the investigation of chronic GERD symptoms. Participants with specialized intestinal metaplasia were classified as cases, and a subset with visible evidence of columnar epithelium were further classified as BE cases (29, 31). GERD controls consisted of a ~50% random sample of subjects who were biopsy-negative for specialized intestinal metaplasia. The SBEP includes >400 BE patients who underwent periodic endoscopic surveillance with

biopsy according to the Seattle Protocol (32). Diagnosis of dysplasia or EA, and/or confirmation of existing BE, was conducted by a single pathologist based on histologic assessment of biopsy specimens according to standard guidelines. Demographic and lifestyle characteristics were ascertained in both studies via similar structured interviews, either 1–2 months after endoscopy (SRD), or at baseline and follow-up endoscopy visits (SBEP).

The present study included a total of 322 participants: 162 individuals from the SRD (100 GERD cases and 62 BE cases), and 160 individuals from the SBEP (60 BE cases and 100 cases of high-grade dysplasia (HGD) or early-stage EA) (Figure S1). The BE case group from each study included an approximately equal number of long-segment BE (≥ 3 cm) and short-segment BE (<3 cm) cases. Participation was restricted to white subjects of age ≥ 20 years. Exclusions were made for GERD cases with visible evidence of columnar mucosa on endoscopy, and for EA cases with Stage II disease or higher. Cases were frequency balanced by age and sex across case type, within study, where feasible. This study was approved by the Institutional Review Board at the Fred Hutchinson Cancer Research Center.

Serum specimens

Blood samples from study participants selected for the present analyses were obtained within 0–120 days (median: 45.5 d) following endoscopy (SRD), or at one of the periodically-scheduled endoscopy visits during follow-up (SBEP). Among the SBEP BE subjects, blood samples were selected from the last available endoscopy visit at which the maximum histologic diagnosis was non-dysplastic BE. Among the combined HGD/EA case group, blood specimens were selected from an endoscopy visit at the time of EA diagnosis ($n=22$), shortly before a subsequent recorded EA diagnosis ($n=17$), or from the last available endoscopy date at which HGD was detected ($n=61$) (Table S1). Blood samples were allowed to clot for 45 min, centrifuged at 4°C for 10 min at $1300 \times g$, and the resulting sera were separated and aliquoted into separate tubes and frozen at -80°C . All specimens were processed and frozen for long-term storage within four hours post-draw. Each sample was assigned a unique laboratory identification number, which specified the order of processing and blinded laboratory personnel to sample identities. Samples of each case type were equally represented and randomly ordered within batch. Specimens were delivered on dry ice to the Northwest Metabolomics Research Center and stored at -80°C until use.

Laboratory methods

Frozen patient sera were thawed at room temperature for 30 min. Detailed protocols for sample preparation have been published previously (33). Two sample injections (2 μL and 10 μL) were used for LC-MS/MS analysis in positive and negative modes, respectively. ^{13}C -lactate and ^{13}C -tyrosine were added to each serum sample, after deproteinization, to monitor the system performance. Replicates of quality control serum samples were injected into the LC-MS/MS after every ~ 10 study samples, and derived profiles were used to calculate the coefficient of variation (CV) for each metabolite.

A robust targeted LC-MS/MS method has been developed and used in multiple studies in the Northwest Metabolomics Research Center (NW-MRC) (33–35). Briefly, LC-MS/MS profiling was performed on an Agilent 1260 LC (Agilent Technologies, Santa Clara, CA)

AB Sciex QTrap 5500 MS (AB Sciex, Toronto, Canada) system. MRM transitions were monitored in negative and positive mode, and extracted MRM peaks were integrated using MultiQuant 2.1 software (AB Sciex, Toronto, Canada).

Data processing and statistical analysis

Peak intensities were integrated using Analyst 1.5 software (AB SCIEX), and resulting integrals were used for further analyses. Profiled analytes were excluded if detectable signal was absent in >33% of all study samples, or if the CV across all QC samples exceeded 20%. Measurements for an individual metabolite were normalized to the mean signal within-batch for that metabolite and log₂-transformed. To evaluate relative differences in mean metabolite signals between case types, linear regression analyses were conducted in which metabolite values were regressed on case status (X, defined as 0/1), with adjustment for covariates (C_i): $M \sim \alpha_0 + [\alpha_1 C_1 + \dots + \alpha_n C_n] + \beta X$. Metabolites for which β differed significantly from zero (at P<0.05) were identified as preliminary candidates. Multiple comparisons were accounted for using the Benjamini-Hochberg false discovery rate (FDR) (36).

Two case type comparisons were examined: i) BE versus GERD (SRD samples), and ii) HGD/EA versus BE (SBEP samples). Inclusion of an individual covariate (C_i) in a given regression model was determined by assessment of the association between values of the metabolite and the covariate in samples of the reference case type (i: GERD and ii: BE): $M \sim \alpha + \gamma_i C_i | X=0$. Covariates for which γ_i differed significantly from zero (at P<0.05) were included in the primary regression model for the selected metabolite. Adjustment was considered for several variables considered risk factors for BE or EA—age (grouped-linear: 20–39, 40–49, 50–59, 60–69, 70–79, 80+ years), sex (male/female), smoking history (yes/no), and waist-to-hip-ratio (WHR, <0.9 or 0.9, males; <0.8 or 0.8, females)—and for duration of specimen storage at –80 °C (grouped-linear: 0–4, 5–9, 10–14, or 15–19 years). We adopted this approach to provide reasonable control for likely confounding factors while avoiding potentially biased inference from over-adjustment. Fold changes in geometric mean metabolite signals between case groups were calculated using covariate-adjusted metabolite values. For a given metabolite, if covariates C₁ and C₂ were included in the main regression model, adjusted metabolite values were obtained as follows: $M_{adj} = M_{unadj} - \delta_1 C_1 - \delta_2 C_2$, where δ_1 and δ_2 are coefficients derived from the regression $M \sim \alpha + \delta_1 C_1 + \delta_2 C_2 | X=0$.

For multivariate modeling, we first selected the top pool of candidates for each comparison based on a P-value threshold of 0.05 (n=9 metabolites for BE versus GERD, and n=4 metabolites for HGD/EA versus BE). Models for disease classification were constructed using regularized logistic regression with elastic net penalty (R package: glmnet) (37). Equal numbers of cases and controls were randomly allocated to a training set (75%), used for variable selection and model selection, or a testing set (25%), used for (preliminary) model validation. With the mixing parameter (α) set to 0.5, five-fold cross-validation was conducted within the training set only to select the optimal value of the penalty parameter (λ). Based on this value of λ , a model was generated using the complete training set data and used to predict class values for subjects in the testing set. Monte Carlo cross validation (MCCV) was conducted, such that the entire procedure described was repeated using 100 different training (and associated testing) sets randomly selected from the study sample.

Mean estimated area under the receiver-operating curve (AUC) was calculated across the 100 testing sets. Composite average ROC curves were constructed to summarize overall classification accuracy (R package: ROCR) (38). All univariate statistical analyses were conducted using Stata v13.1 (College Station, TX), while multivariate modeling was conducted using R v3.03.

Results

Study participant characteristics

Demographic and behavioral characteristics of included participants are shown in Table 1. Study participation was restricted to white individuals. Among individuals included from the SRD, BE cases were older and more likely to be male relative to GERD cases. The prevalence of central obesity (high WHR) was substantially higher for the BE versus GERD case group, while the percentage of ever-smokers was only slightly elevated. Among participants from the Seattle BE Project, HGD/EA cases were older and more likely to be male relative to BE cases. The vast majority of cases with either BE or HGD/EA exhibited central obesity, while ever-smoking prevalence was somewhat higher among individuals with HGD/EA.

Assessment of serum metabolite differences in cases with BE versus GERD, and HGD/EA versus BE

Detected metabolites represent several major metabolic pathways (glycolysis/TCA cycle, amino acid, nucleotide, and lipid metabolism) (Table S2). At the nominal significance threshold of $P < 0.05$, a number of metabolites were found to differ in mean signal intensity among individuals with BE versus GERD ($n=9$), and HGD/EA versus BE ($n=4$) (Tables 2 & S3). All nine of the metabolites identified in the BE versus GERD comparison were elevated in BE, while one of the four metabolites in the HGD/EA versus BE comparison (urate) was elevated in HGD/EA. Fold changes ranged from 1.07–1.29, and 0.90–1.11, respectively. When accounting for multiple comparisons, four metabolites differentiating between BE and GERD (creatine, homocysteine, 3-nitrotyrosine, and hydroxyproline/aminolevulinate) satisfied the FDR threshold of $q \leq 0.15$; the top differentiating metabolite between HGD/EA and BE (urate) had an FDR q value of 0.01. Box plots for selected top metabolites ($P < 0.01$) by case type are shown in Figure 1. In exploratory subgroup analyses, urate also remained highly significant ($P=0.001$, $q=0.058$) when the HGD/EA case group of 100 participants was restricted to the 22 included EA cases with blood specimens drawn at the time of cancer diagnosis (data not shown).

Multivariate disease classification models

To investigate the utility of combining multiple serum metabolites into a signature to discriminate between the indicated case types, we conducted multivariate modeling using regularized logistic regression with elastic net penalty. The mean AUC of models built using the top nine metabolites ($P < 0.05$) selected for BE versus GERD, and assessed across 100 rounds of MCCV, was 0.64 (SD=0.08); the mean AUC of models built using the top four metabolites ($P < 0.05$) selected for HGD/EA versus BE was 0.75 (SD=0.07) (Figure 2). Several individual metabolite variables were retained in 80% of all MCCV iterations for

the BE versus GERD classifier (n=5 of 9), and for the HGD/EA versus BE classifier (n=4 of 4) (Table S4). For each comparison, the metabolite retained in the highest percentage of MCCV iterations (models) was also the metabolite with the lowest P value in our primary linear regression analyses (creatine and urate, respectively).

Discussion

The prevention and control of EA has become an increasingly important public health challenge, as incidence of this rare but lethal cancer has risen 8-fold in recent decades (1). Despite identification of an epithelial precursor lesion and several strong epidemiologic risk factors, current screening and surveillance efforts, based largely on upper endoscopy with biopsy of patients with GERD or BE, have not significantly improved early detection or lowered EA mortality (1, 4). An important unmet need remains the development of robust non-invasive methods of assessing and stratifying at-risk patients according to their level of current, and risk of future, disease progression (39). Here, we adopted metabolomics profiling methods to identify serum metabolites correlated with distinct disease intermediates along the pathway to EA. Targeted interrogation of 57 serum metabolites encompassing several major metabolic pathways revealed multiple metabolites that differed in abundance when comparing patients with BE versus GERD, and HGD/EA versus BE. Multivariate models based on candidate metabolite markers yielded moderate discrimination between the HGD/EA and BE case groups, with lesser separation observed for BE versus GERD.

Despite a large number of cancer metabolomics studies in recent years, relatively few have focused on EA (or BE). These mostly small-scale investigations employed various NMR- or MS-based methods to conduct either targeted or global metabolite profiling in tissue, serum, or urine (17–25). Considerable variability has been observed among reported findings, likely reflecting significant differences not only in analytic platforms and types of biospecimens, but also in study design, subject selection, and sample preparation (40). Differentially abundant metabolites identified by NMR in EA tissue relative to normal tissue have included phosphocholine, glutamate, myo-inositol, adenosine-containing compounds, uracil, and inosine (23). Studies comparing serum from EA patients versus healthy controls, using either MS or NMR, have described differences in multiple nucleosides, amino acids, and other metabolites (19, 21, 24, 25). NMR profiling of urine from healthy controls or EA/ESCC patients has identified differences in sucrose, pantothenate, urea, methanol, pyroglutamate, hypoxanthine, fucose, 2-aminobutyrate, and cis-aconitate (18). In total, minimal consensus has emerged regarding consistent metabolite-based signatures of EA, and our biological understanding of the inter-relationships between metabolite levels in tissues, blood, and urine remains limited. Furthermore, no previous studies have included at-risk GERD patients as a comparison group.

The individual metabolite marker candidates identified in our analysis belong to several classes of compounds. The nine metabolites differentiating BE from GERD ($P<0.05$) included seven amino acids or amino acid derivatives (creatine, homocysteine, 3-nitrotyrosine, hydroxyproline/aminolevulinate, arginine, tyrosine, ornithine), one fatty acid (linoleic acid), and one sugar alcohol (sorbitol). The four metabolites differentiating

HGD/EA from BE ($P<0.05$) included one nucleotide derivative (urate), two amino acid derivatives (taurine and xanthurenate), and one sugar (erythrose). On a statistical level, the most significant signal observed was for urate ($P=0.0002$, $q=0.01$), in the HGD/EA versus BE comparison, and for creatine and homocysteine ($q<0.15$) in the BE versus GERD comparison. Urate also showed non-significant elevation in BE versus GERD ($P=0.12$), while homocysteine was similarly non-significantly elevated in HGD/EA versus BE ($P=0.17$).

It is of interest to note that several of the top identified metabolites (urate, homocysteine, and 3-nitrotyrosine) have been linked to inflammatory processes (41–43), as local and systemic inflammation are believed to play important roles in the pathogenesis of BE/EA (4). Urate is the monosodium form of uric acid, a derivative of purine nucleotides. High levels of circulating uric acid have been associated with a number of medical conditions, including gout, kidney stones, tumor lysis syndrome, metabolic syndrome, and diabetes (42), as well as with increased risk of several cancers (44, 45). Originally described as an antioxidant, uric acid has more recently been found to promote inflammation, both in its crystal and soluble forms, through activation of Toll receptors and stimulation of NF κ B, respectively (42). The amino acid homocysteine, a derivative of methionine, functions in one-carbon metabolism reactions, which are important for nucleotide synthesis and DNA methylation. Elevated circulating homocysteine has been associated with increased risk of head and neck cancers, colorectal cancer, and colorectal adenoma recurrence (46–48). Homocysteine induces inflammatory cytokine production in cultured human cells, and has been mechanistically linked to inflammatory bowel disease (IBD) and atherosclerosis (41, 49). Nitrotyrosine (NT) is derived from nitration of the amino acid tyrosine, mediated by reactive oxygen species such as peroxynitrite anion. NT is considered a marker of oxidative stress, inflammation, and protein damage, and has been detected in a number of disease states such as gastritis and IBD (43). Past studies have also described elevated NT immunoreactivity in esophageal tissue specimens from patients with BE (43, 50) and EA (43, 51), and in the esophagus of an animal model of EA (52).

Local and systemic inflammation represent plausible drivers of altered metabolite levels, as do increased rates of local cell proliferation associated with cancer precursors and early lesions. Metabolite alterations observed across case types may or may not reflect underlying metabolite changes within the esophageal epithelium itself. Interrogation of paired tissue biopsies from the same individuals and time points in which serum specimens were isolated will be required to address this uncertainty and facilitate accurate biological interpretation of the reported metabolite differences.

Non-invasive clinical assessment tools for patients at risk of EA could be of use in at least two clinical scenarios: i) initial evaluation of patients with severe or chronic GERD symptoms, to identify those most likely to have BE (or advanced pathology); and ii) additional monitoring of BE patients between surveillance endoscopies, to identify those most likely to have HGD or early-stage EA. Preliminary risk models for BE or EA based on classic epidemiologic risk factors such as age, sex, BMI/WHR, smoking status, and GERD symptoms (53–56) have provided a promising starting point for improved risk stratification tools, but the modest classification accuracy of most models requires improvement to afford

significant clinical utility. One emerging approach for enhancing non-invasive risk assessment is the Cytosponge, a cytological sampling device that yields esophageal epithelial cells for immunostaining or molecular analyses (57, 58). Our results suggest that the serum metabolome may represent another viable source of clinical information, and point to the potential utility of integrating novel candidate biomarkers with existing classical models to build more robust classifiers. Given that existing tissue-based variables such as dysplasia grade, DNA content flow cytometry, and somatic chromosome abnormalities provide strong discrimination among BE patients (59), serum-based assays will require substantial improvements to have direct clinical application. Future studies are also needed to probe whether serum metabolite signatures are specific to BE or HGD/EA, or may correlate with other conditions prevalent in a larger sample of at-risk individuals.

This discovery study had a number of key strengths. First, in contrast to past reports that primarily used healthy normal controls, our analysis focused on comparisons most relevant to clinical decision making and targeted risk assessment: BE versus GERD, and HGD/EA versus BE. Second, GERD, BE, and HGD/EA were all diagnosed according to standard definitions based on endoscopic and histopathologic evaluation, and serum specimens were processed according to uniform procedures and frozen at -80°C within four hours for long-term storage. Third, biospecimens were linked with extensive epidemiologic and clinical data from high-quality studies of BE and EA. Statistical analyses were adjusted as needed for a number of potential confounders, including age, sex, smoking, and waist-to-hip ratio. Fourth, to avoid capturing metabolite alterations associated exclusively with advanced/incurable disease, EA cases were restricted to Stage 0/1. Finally, our LC-MS-based profiling platform has been optimized for sensitivity and accuracy, and used successfully to identify and/or validate serum metabolite biomarker candidates for a number of diseases (33, 60).

Our study also had several limitations. First, while a total sample size of 322 participants was larger than that used in most previous metabolomics studies of BE/EA, it is nonetheless still relatively small when considering the modest magnitude of disease-associated serum metabolite fluctuations typically observed, coupled with the anticipated degree of natural biologic (and technical) variability in such measurements (61). Given that participants in the present analysis were drawn from two distinct parent studies (SRD & SBEP), we chose to retain two separate BE case groups and conduct only within-study comparisons (BE versus GERD, and HGD/EA versus BE), to eliminate the potential for confounding by systematic differences between participants/specimens by parent study (eg. SBEP blood samples were all collected under fasting conditions prior to endoscopy, while SRD samples were collected in the field under non-fasting circumstances). This decision effectively reduced study power by decreasing the size of our BE comparison groups, and precluded a comparison of HGD/EA vs. GERD patients. Second, while our targeted LC-MS analysis enabled assessment of 57 metabolites in multiple metabolic pathways, these compounds represent only a very small fraction of the total serum metabolome. The combination of disparate detection technologies, such as MS and NMR, and the pairing of targeted and global profiling approaches, may facilitate identification of a richer set of differentiating metabolites (24). Exploratory global MS profiling in the present study, however, did not yield strong additional signals (data not shown). Third, while our analyses made use of extensive covariate data on study participants for statistical adjustments, often not conducted

in metabolomics reports, the possibility remains for some degree of residual or unmeasured confounding, whereby observed metabolite differences between case types might reflect underlying subject/specimen characteristics. Medications and dietary patterns are two variables capable of influencing metabolite levels (62), which were not included in our analyses. However, given that we restricted EA cases to stage 0/1 (prior to surgical or chemoradiation treatment) and used refractory GERD cases as our primary “controls” (instead of healthy normals), substantial heterogeneity in medication usage between case groups appears relatively unlikely. It remains uncertain whether significant dietary differences exist between GERD, BE, and HGD/EA patients.

This study provides evidence that intermediate disease states preceding the development of EA correlate with detectable alterations in the serum metabolome. Our results expand upon the findings of a small number of past analyses by identifying metabolite differences across gradations of disease progression (BE versus GERD, and HGD/EA versus BE) that are most relevant to the clinical settings of screening and surveillance. Further studies in larger independent populations will be necessary to validate these results, and ultimately to evaluate the utility of serum metabolite profiles as a novel source of clinical information for improving non-invasive assessment and monitoring of patients at risk of EA.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Terri Watson, Tricia Christopherson, Paul Hansen, Jessica Arnaudo, David Cowan, and Carissa Sanchez for their contributions in project management, organization of biospecimens/data, or assistance with data retrieval. This work was principally supported by NIH grant R21CA178621 (T.L.V. and D.R.) and NIH training grant T32CA009168 (T.L.V.). Additional support was provided by NIH grants K05CA124911 (T.L.V.), P01CA091955 (B.J.R.), and 5P30CA015704 (D.R.).

References

1. Vaughan TL, Fitzgerald RC. Precision prevention of oesophageal adenocarcinoma. *Nat Rev Gastroenterol Hepatol*. 2015; 12:243–248. [PubMed: 25666644]
2. Cook MB, et al. Gastroesophageal reflux in relation to adenocarcinomas of the esophagus: A pooled analysis from the Barrett’s and Esophageal Adenocarcinoma Consortium (BEACON). *PLoS ONE*. 2014; 9:e103508. [PubMed: 25075959]
3. Kandulski A, Malfertheiner P. Gastroesophageal reflux disease—from reflux episodes to mucosal inflammation. *Nature Reviews Gastroenterology & Hepatology*. 2011; 9:15–22. [PubMed: 22105108]
4. Reid BJ, Li X, Galipeau PC, Vaughan TL. Barrett’s oesophagus and oesophageal adenocarcinoma: time for a new synthesis. *Nat Rev Cancer*. 2010; 10:87–101. [PubMed: 20094044]
5. Asiago VM, et al. Early detection of recurrent breast cancer using metabolite profiling. *Cancer Research*. 2010; 70:8309–8318. [PubMed: 20959483]
6. Denkert C, et al. Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. *Cancer Research*. 2006; 66:10795–10804. [PubMed: 17108116]
7. Mayers JR, et al. Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nature Medicine*. 2014; 20doi: 10.1038/nm.3686

8. Slupsky CM, et al. Urine metabolite analysis offers potential early diagnosis of ovarian and breast cancers. *Clinical Cancer Research*. 2010; 16:5835–5841. [PubMed: 20956617]
9. Sreekumar A, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature*. 2009; 457:910–914. [PubMed: 19212411]
10. Wikoff WR, et al. Diacetylspermine Is a Novel Prediagnostic Serum Biomarker for Non-Small-Cell Lung Cancer and Has Additive Performance With Pro-Surfactant Protein B. *J Clin Oncol*. 2015; doi: 10.1200/JCO.2015.61.7779
11. Gowda G, et al. Metabolomics-based methods for early disease diagnostics. *Expert Review of Molecular Diagnostics*. 2008; 8:617–633. [PubMed: 18785810]
12. Nicholson JK, Lindon JC, Holmes E. Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica; the fate of foreign compounds in biological systems*. 1999; 29:1181–1189. [PubMed: 10598751]
13. Gu H, Gowda G, Raftery D. Metabolic profiling: are we en route to better diagnostic tests for cancer? *Future Oncology*. 2012; 8:1207–1210. [PubMed: 23130920]
14. Liesenfeld DB, Habermann N, Owen RW, Scalbert A, Ulrich CM. Review of mass spectrometry-based metabolomics in cancer research. *Cancer Epidemiology Biomarkers and Prevention*. 2013; 22:2182–2201.
15. O'Connell TM. Recent advances in metabolomics in oncology. *Bioanalysis*. 2012; 4:431–451. [PubMed: 22394143]
16. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: A review. *Clinical Cancer Research*. 2009; 15:431–440. [PubMed: 19147747]
17. Dave U, et al. In vitro 1H-magnetic resonance spectroscopy of Barrett's esophageal mucosa using magic angle spinning techniques. *European Journal of Gastroenterology and Hepatology*. 2004; 16:1199–1205. [PubMed: 15489582]
18. Davis VW, Schiller DE, Eurich D, Sawyer MB. Urinary metabolomic signature of esophageal cancer and Barrett's esophagus. *World Journal of Surgical Oncology*. 2012; 10:271. [PubMed: 23241138]
19. Djukovic D, Baniyadi HR, Kc R, Hammoud Z, Raftery D. Targeted serum metabolite profiling of nucleosides in esophageal adenocarcinoma. *Rapid Communications in Mass Spectrometry*. 2010; 24:3057–3062. [PubMed: 20872639]
20. Doran ST, et al. Pathology of Barrett's esophagus by proton magnetic resonance spectroscopy and a statistical classification strategy. *American Journal of Surgery*. 2003; 185:232–238. [PubMed: 12620562]
21. Sanchez-Espiridon B, et al. Identification of Serum Markers of Esophageal Adenocarcinoma by Global and Targeted Metabolic Profiling. *Clin Gastroenterol Hepatol*. 2015; doi: 10.1016/j.cgh.2015.05.023
22. Suzuki M, Nishiumi S, Matsubara A, Azuma T, Yoshida M. Metabolome analysis for discovering biomarkers of gastroenterological cancer. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*. 2014; 966:59–69. [PubMed: 24636738]
23. Yakoub D, Keun HC, Goldin R, Hanna GB. Metabolic profiling detects field effects in nondysplastic tissue from esophageal cancer patients. *Cancer Research*. 2010; 70:9129–9136. [PubMed: 20884633]
24. Zhang J, et al. Esophageal cancer metabolite biomarkers detected by LC-MS and NMR methods. *PLoS ONE*. 2012; 7:e30181. [PubMed: 22291914]
25. Zhang J, et al. Metabolomics study of esophageal adenocarcinoma. *Journal of Thoracic and Cardiovascular Surgery*. 2011; 141:469–475. [PubMed: 20880550]
26. Phelan JJ, et al. Differential expression of mitochondrial energy metabolism profiles across the metaplasia-dysplasia-adenocarcinoma disease sequence in Barrett's oesophagus. *Cancer Letters*. 2014; 354:122–131. [PubMed: 25107643]
27. Suchorolski MT, Paulson TG, Sanchez Ca, Hockenbery D, Reid BJ. Warburg and Crabtree Effects in Premalignant Barrett's Esophagus Cell Lines with Active Mitochondria. *PLoS ONE*. 2013; 8:e56884. [PubMed: 23460817]

28. Edelstein ZR, Bronner MP, Rosen SN, Vaughan TL. Risk factors for Barrett's esophagus among patients with gastroesophageal reflux disease: a community clinic-based case-control study. *The American Journal of Gastroenterology*. 2009; 104:834–842. [PubMed: 19319131]
29. Edelstein ZR, Farrow DC, Bronner MP, Rosen SN, Vaughan TL. Central adiposity and risk of Barrett's esophagus. *Gastroenterology*. 2007; 133:403–411. [PubMed: 17681161]
30. Galipeau PC, et al. NSAIDs modulate CDKN2A, TP53, and DNA content risk for progression to esophageal adenocarcinoma. *PLoS Medicine*. 2007; 4:342–353.
31. Sampliner RE. Updated Guidelines for the Diagnosis, Surveillance, and Therapy of Barrett's Esophagus. *American Journal of Gastroenterology*. 2002; 97:1888–95. [PubMed: 12190150]
32. Reid BJ, Blount PL, Rabinovitch PS. Biomarkers in Barrett's esophagus. *Gastrointestinal Endoscopy Clinics of North America*. 2003; 13:369–97. [PubMed: 12916666]
33. Zhu J, et al. Colorectal Cancer Detection Using Targeted Serum Metabolic Profiling. *Journal of Proteome Research*. 2014; 13:4120–30. [PubMed: 25126899]
34. Carroll PA, et al. Deregulated Myc Requires MondoA/Mlx for Metabolic Reprogramming and Tumorigenesis. *Cancer Cell*. 2015; 27:271–285. [PubMed: 25640402]
35. Sperber H, et al. The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nat Cell Biol*. 2015; 17:1523–1535. [PubMed: 26571212]
36. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *Journal of Clinical Epidemiology*. 2014; 67:850–857. [PubMed: 24831050]
37. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*. 2010; 33:1–22. [PubMed: 20808728]
38. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: Visualizing classifier performance in R. *Bioinformatics*. 2005; 21:3940–3941. [PubMed: 16096348]
39. Vaughan TL. From genomics to diagnostics of esophageal adenocarcinoma. *Nature Genetics*. 2014; 46:806–807. [PubMed: 25070798]
40. Abbassi-Ghadi N, et al. Metabolomic profiling of oesophago-gastric cancer: A systematic review. *European Journal of Cancer*. 2013; 49:3625–3637. [PubMed: 23896378]
41. Danese S, et al. Homocysteine triggers mucosal microvascular activation in inflammatory bowel disease. *Am J Gastroenterol*. 2005; 100:886–895. [PubMed: 15784037]
42. Fini MA, Elias A, Johnson RJ, Wright RM. Contribution of uric acid to cancer risk, recurrence, and mortality. *Clin Transl Med*. 2012; 1:16. [PubMed: 23369448]
43. Vaninetti NM, et al. Inducible nitric oxide synthase, nitrotyrosine and p53 mutations in the molecular pathogenesis of Barrett's esophagus and esophageal adenocarcinoma. *Mol Carcinog*. 2008; 47:275–285. [PubMed: 17849424]
44. Kolonel LN, Yoshizawa C, Nomura AM, Stemmermann GN. Relationship of serum uric acid to cancer occurrence in a prospective male cohort. *Cancer Epidemiol Biomarkers Prev*. 1994; 3:225–228. [PubMed: 8019371]
45. Strasak AM, et al. Use of penalized splines in extended Cox-type additive hazard regression to flexibly estimate the effect of time-varying serum uric acid on risk of cancer incidence: a prospective, population-based study in 78,850 men. *Ann Epidemiol*. 2009; 19:15–24. [PubMed: 18835524]
46. Bobe G, et al. Serum adiponectin, leptin, C-peptide, homocysteine, and colorectal adenoma recurrence in the Polyp Prevention Trial. *Cancer Epidemiol Biomarkers Prev*. 2010; 19:1441–1452. [PubMed: 20501764]
47. Fanidi A, et al. A prospective study of one-carbon metabolism biomarkers and cancer of the head and neck and esophagus. *Int J Cancer*. 2015; 136:915–927. [PubMed: 24975698]
48. Miller JW, et al. Homocysteine, cysteine, and risk of incident colorectal cancer in the Women's Health Initiative observational cohort. *Am J Clin Nutr*. 2013; 97:827–834. [PubMed: 23426034]
49. Zhou J, Austin RC. Contributions of hyperhomocysteinemia to atherosclerosis: Causal relationship and potential mechanisms. *Biofactors*. 2009; 35:120–129. [PubMed: 19449439]
50. Jiménez P, et al. Free radicals and antioxidant systems in reflux esophagitis and Barrett's esophagus. *World J Gastroenterol*. 2005; 11:2697–2703. [PubMed: 15884106]

51. Allameh A, et al. Immunohistochemical analysis of selected molecular markers in esophagus precancerous, adenocarcinoma and squamous cell carcinoma in Iranian subjects. *Cancer Epidemiol.* 2009; 33:79–84. [PubMed: 19679052]
52. Goldstein SR, Yang GY, Chen X, Curtis SK, Yang CS. Studies of iron deposits, inducible nitric oxide synthase and nitrotyrosine in a rat model for esophageal adenocarcinoma. *Carcinogenesis.* 1998; 19:1445–1449. [PubMed: 9744541]
53. Rubenstein JH, et al. Prediction of Barrett's esophagus among men. *The American Journal of Gastroenterology.* 2013; 108:353–62. [PubMed: 23318485]
54. Thrift AP, Kendall BJ, Pandeya N, Vaughan TL, Whiteman DC. A Clinical Risk Prediction Model for Barrett Esophagus. *Cancer Prev Res (Phila).* 2012; 5:1115–1123. [PubMed: 22787114]
55. Thrift AP, Kendall BJ, Pandeya N, Whiteman DC. A Model to Determine Absolute Risk for Esophageal Adenocarcinoma. *Clinical Gastroenterology and Hepatology.* 2012; 11:138–44.e2. [PubMed: 23103823]
56. Xie S-H, Lagergren J. A Model for Predicting Individuals' Absolute Risk of Esophageal Adenocarcinoma: Moving towards Tailored Screening and Prevention. *Int J Cancer.* 2016; doi: 10.1002/ijc.29988
57. Ross-Innes CS, et al. Evaluation of a minimally invasive cell sampling device coupled with assessment of trefoil factor 3 expression for diagnosing Barrett's esophagus: a multi-center case-control study. *PLoS Med.* 2015; 12:e1001780. [PubMed: 25634542]
58. Weaver MJ, et al. Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nature genetics.* 2014; 46:837–843. [PubMed: 24952744]
59. Li X, et al. Assessment of esophageal adenocarcinoma risk using somatic chromosome alterations in longitudinal samples in Barrett's esophagus. *Cancer Prev Res (Phila).* 2015; doi: 10.1158/1940-6207.CAPR-15-0130
60. Baniasadi H, et al. Targeted metabolic profiling of hepatocellular carcinoma and hepatitis C using LC-MS/MS. *Electrophoresis.* 2013; 34:2910–2917. [PubMed: 23856972]
61. Sampson JN, et al. Metabolomics in epidemiology: Sources of variability in metabolite measurements and implications. *Cancer Epidemiology Biomarkers and Prevention.* 2013; 22:631–640.
62. Kim K, et al. Mealtime, temporal, and daily variability of the human urinary and plasma metabolomes in a tightly controlled environment. *PLoS ONE.* 2014; 9doi: 10.1371/journal.pone.0086223

Figure 1.

Box plots of top differentially abundant ($P < 0.01$) serum metabolites for BE versus GERD (A) or HGD/EA versus BE (B). Median values are indicated by horizontal lines within plotted boxes. Values for the 25th and 75th percentiles are indicated by the lower and upper box boundaries, respectively, with vertical lines (whiskers) marking the distribution of observed values, and outliers indicated by gray dots.

Figure 2. Estimated accuracy of multivariate classifiers based on Monte Carlo cross validation. Composite average receiver operating characteristic (ROC) curves summarizing classification performance across 100 iterations of MCCV for models constructed using elastic net regression. (A) BE versus GERD (mean area under the curve (AUC)=0.64, SD=0.08), (B) HGD/EA versus BE (mean AUC=0.75, SD=0.07).

Table 1

Participant characteristics.

	GERD ^a (n=100)		BE ^a (n=62)		BE ^b (n=60)		HGD+EA ^b (n=100)	
	n	%	n	%	n	%	n	%
Age								
20-39	18	18.0	8	12.9	1	1.7	0	0.0
40-49	24	24.0	12	19.4	4	6.7	5	5.0
50-59	33	33.0	23	37.1	14	23.3	14	14.0
60-69	15	15.0	9	14.5	22	36.7	29	29.0
70-79	10	10.0	9	14.5	15	25.0	35	35.0
80+	0	0.0	1	1.6	4	6.7	17	17.0
Mean (SD)	51.6 (12.5)		55.3 (12.5)		64.9 (10.9)		68.9 (10.4)	
Sex								
Female	42	42.0	14	22.6	16	26.7	8	8.0
Male	58	58.0	48	77.4	44	73.3	92	92.0
Race								
White	100	100.0	62	100.0	60	100.0	100	100
Smoking								
Never	46	46.0	25	40.3	24	40.0	27	27.0
Ever	54	54.0	37	59.7	36	60.0	73	73.0
WHR ^c								
Normal	29	29.0	9	14.5	4	6.7	6	6.0
Obese	71	71.0	53	85.5	56	93.3	94	94.0

^aStudy of Reflux Disease,

^bSeattle BE Project,

^cWaist-to-hip-ratio; normal WHR<0.9 (males) or <0.8 (females); High WHR 0.9 (males) or 0.8 (females)

Table 2

Top differentially abundant serum metabolites identified for BE versus GERD (A) or HGD/EA versus BE (B).

A.						
Metabolite	N	p ^a	q ^b	Cov ^c	FC ^d #	
1 Creatine	162	0.0035	0.11	2	1.29	
2 Homocysteine	162	0.0039	0.11		1.10	
3 3-Nitrotyrosine	162	0.010	0.15		1.10	
4 Hydroxyproline/Aminolevulinate	162	0.011	0.15		1.16	
5 Arginine	162	0.018	0.20		1.07	
6 Tyrosine	162	0.026	0.24		1.10	
7 Sorbitol	162	0.033	0.26		1.09	
8 Linoleic Acid	162	0.037	0.26		1.13	
9 Ornithine	162	0.046	0.29		1.10	
10 Urate	162	0.119	0.62	2	1.05	
B.						
Metabolite	N	p ^a	q ^b	Cov ^c	FC ^d #	
1 Urate	160	0.0002	0.01		1.11	
2 Taurine	160	0.016	0.31	5	0.90	
3 Erythrose	160	0.019	0.31		0.90	
4 Xanthurenate	160	0.021	0.31	2	0.94	
5 Aminoisobutyrate	160	0.054	0.61	1	1.04	
6 Sorbitol	160	0.116	0.81		1.06	
7 Valine	160	0.129	0.81	2	0.94	
8 Tryptophan	160	0.149	0.81	2	0.95	
9 Methionine	160	0.154	0.81	2	0.95	
10 Homocysteine	160	0.169	0.81		1.05	

^aP values were derived from linear regression of metabolite values on case status, with adjustment for indicated covariates^[c].

^bFalse discovery rate (Benjamini and Hochberg).

^cCovariates included in linear regression model: age (1), sex (2), smoking (3), WHR (4), specimen storage duration (5).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fold change of geometric mean signal,

Based on covariate-adjusted metabolite values. Metabolites above the horizontal lines reached nominal significance ($P < 0.05$).