Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2021 May ; 30(5): 1000–1008. doi:10.1158/1055-9965.EPI-20-1363.

Association of pre-diagnostic blood metabolomics with prostate cancer defined by ERG or PTEN molecular subtypes

Xiaoshuang Feng¹, Cindy Ke. Zhou¹, Clary B. Clish², Kathryn M. Wilson^{1,3}, Claire H. Pernar¹, Barbra A. Dickerman¹, Massimo Loda⁴, Stephen P. Finn⁵, Kathryn L. Penney^{1,3}, Daniel R. Schmidt^{6,7}, Matthew G. Vander Heiden^{4,6}, Edward L. Giovannucci^{1,3,8}, Ericka M. Ebot^{1,†}, Lorelei A. Mucci^{1,3,†}

¹Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA

²Broad Institute, Cambridge, MA, USA ³Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

⁴Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA ⁵Department of Histopathology Research, Trinity College, Dublin, Ireland ⁶David H. Koch Institute for Integrative Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA ⁷Department of Radiation Oncology, Beth Israel Deaconess Medical Center, Boston, MA, USA ⁸Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Abstract

Background: The *TMPRSS2: ERG* gene fusion and *PTEN* loss are two of the most common somatic molecular alterations in prostate cancer. Here we investigated the association of pre-diagnostic circulating metabolomics and prostate cancer defined by ERG or PTEN status to improve understanding of these etiologically distinct molecular prostate cancer subtypes.

Methods: The study was performed among 277 prostate cancer cases with ERG status, 211 with PTEN status, and 294 controls nested in the Health Professionals Follow-up Study (HPFS) and the Physicians' Health Study (PHS). We profiled 223 polar and non-polar metabolites using liquid chromatography-mass spectrometry in pre-diagnostic plasma specimens. We applied enrichment analysis and multinomial logistic regression models to identify biological metabolite classes and individual metabolites associated with prostate cancer defined by ERG or PTEN status.

Corresponding author: Xiaoshuang Feng, Department of Epidemiology, Harvard T. H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA, USA 02115; Tel: +86 18501284593; xfeng@hsph.harvard.edu.

[†]These authors share senior authorship equally.

Authors' Contributions

Conception and design: E. M. Ebot, L. A. Mucci

Development of methodology: X. Feng, C. H. Pernar, C. K. Zhou, E. M. Ebot

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C. B. Clish, M. F. Loda, S. Finn, D. R. Schmidt, M. G. Vander Heiden

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): X. Feng, C. K. Zhou, E. M. Ebot

Writing, review, and/or revision of the manuscript: X. Feng, C. K. Zhou, K. M. Wilson, B. A. Dickerman, M. F. Loda, K. L. Penney, S. Finn, D. R. Schmidt, M. G. Vander Heiden, E. L. Giovannucci, E. M. Ebot, L. A. Mucci

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C. B. Clish, C. K. Zhou

Study supervision: E. L. Giovannucci, L. A. Mucci

Disclosure of Potential Conflicts of Interest: No potential conflicts of interest were disclosed.

Results: Compared to non-cancer controls, sphingomyelins (P : 0.01), ceramides (P : 0.04), and phosphatidylethanolamines (P : 0.03) circulating levels were enriched among ERG-positive prostate cancer cases. Sphingomyelins (P : 0.02), ceramides (P : 0.005), and amino acids (P : 0.02) were enriched among tumors exhibiting PTEN-loss; unsaturated diacylglycerols (P : 0.003) were enriched among PTEN-intact cases; unsaturated triacylglycerols were enriched among both PTEN-loss (P : 0.001) and PTEN-intact (P : 0.0001) cases. While several individual metabolites identified in the above categories were nominally associated with ERG or PTEN defined prostate cancer, none remained significant after accounting for multiple testing.

Conclusions: The molecular process of prostate carcinogenesis may be distinct for men with different metabolomic profiles.

Impact: These novel findings provide insights into the metabolic environment for the development of prostate cancer.

Keywords

Metabolomics; Prostate cancer; ERG; PTEN; Sphingomyelin

Introduction

The *TMPRSS2:ERG* gene fusion leading to ERG overexpression is the most common somatic event in primary prostate cancer (1), which is found in ~25% of tumors in men of Asian and African descent and ~50% of tumors in men of European descent (2). *PTEN* is the most frequently deleted tumor suppressor gene in prostate cancer, with approximately 15% of primary prostate cancer cases showing homozygous deletions spanning the *PTEN* locus (3). Prior studies have suggested that the *PTEN* deletion are associated with worse outcomes, however, *TMPRSS2: ERG* fusion is not, but it does appear to be an etiologically distinct subtype (4,5). Moreover, it has been reported a close biological relationship between ERG overexpression and *PTEN* loss (6,7), and may delineate distinct prostate cancer subtypes with different prognosis if they co-occurred (8).

Accumulating evidence suggests that ERG and *PTEN* define etiologically distinct molecular subtypes. Observational studies have demonstrated an association between risk factors and prostate cancer defined by ERG and *PTEN*. For example, obesity (9), height (10), tomato sauce consumption (11), and vigorous physical activity (12) were found to be associated with ERG positive prostate cancer; statin use was found to be associated with lower risk of *PTEN*-null prostate cancer (13). Furthermore, ERG positive prostate cancers are characterized by alterations in insulin receptor (IR), insulin growth factor 1 receptor (IGF-1R), other tumor-specific metabolic alterations (14), and several prostate cancer genetic risk variants (15).

Given metabolites are the final downstream products of the genome as well as the intermediates or end products of multiple enzymatic reactions on external exposures, the metabolomic profile prior to disease onset might provide implications to refine existing risk factors and improve our understanding of etiologically distinct molecular subtypes of prostate cancer. We hypothesized that the systemic metabolic state could contribute to the

selective pressure in the prostate environment resulting in outgrowth of malignant clones with genetic alterations that confer increased fitness in a specific metabolic environment. To date, no studies have investigated circulating metabolomic profiles and prostate cancer molecular subtypes. To fill this gap, we leveraged metabolomics data from the prospective Health Professionals Follow-up Study (HPFS) and the Physicians' Health Study (PHS), to investigate the association of pre-diagnostic blood metabolites with prostate cancer defined by ERG or PTEN status.

Materials and methods

Study population

Data for this study was derived from two prospective US cohorts, HPFS and PHS. Details on the design and characteristics of the overall HPFS and PHS cohorts have been described elsewhere (16,17). In brief, the HPFS was initiated in 1986 among 51,529 male health professionals aged 40 to 75 years at baseline. Participants completed an initial questionnaire at baseline and have been followed up by mailed questionnaires every two years to update exposure information and ascertain incident diseases including prostate cancer, and the follow-up rates for each biennial cycle have consistently been >90% (18). During 1993 to 1995, blood samples were collected from 18,225 HPFS participants (19). The PHS was initiated in 1982 as a randomized, double-blind, placebo-controlled trial to test the effects of low-dose aspirin and beta-carotene in the primary prevention of cardiovascular diseases and cancer among 22,071 US male physicians, aged 40 to 84 at baseline. Baseline blood specimens were collected and frozen for later analyses from 14,916 participants during 1982–1984 (20). Self-reported blood fasting time were recorded in both cohorts.

For the current study, we included participants who had been previously selected for nested case-control studies of metabolomics in the HPFS and PHS. In HPFS, we included a case-control study of 213 advanced (T3b/T4/N1/M1/fatal) prostate cancer cases and 213 matched controls, matched on age at blood draw, PSA screening, as well as time, season and calendar year of blood draw. We additionally included a case study comprised of 294 non-advanced cases diagnosed between 1993–2014 for whom tumor tissue was available. We excluded 1 case for missing clinical TNM stage, 257 cases without ERG or PTEN information, and 13 controls for confirmed diagnosis of prostate cancer using medical records. This resulted in 200 controls, 242 prostate cancer cases with ERG data, and 194 cases with PTEN data from HPFS for this study.

In PHS, we first included a case-control study which comprised of 181 prostate cancer cases (100 cT1–3 and Gleason score ≤ 8 cases + 81 cT4/N1/M1 cases) and 100 controls free from prostate cancer at the time of diagnosis of matched cases, frequency matched on age at baseline (40–49/50–59/60–69/70+ years) and fasting status (last meal < 8 vs. ≥ 8 hours). We then included a case-only study among localized, low-grade prostate cancers (i.e., cT1–3 and Gleason score = 2–7). This study was comprised of 48 cases who died of prostate cancer within ten years and 85 cases who survived ten years and beyond. 46 men in PHS were excluded because of insufficient blood sample remaining to undertake the metabolomics assays. After excluding 239 cases without ERG or PTEN information, 94 controls, 35

prostate cancer cases with ERG data, and 17 cases having PTEN data were included from PHS.

Ethics statement

The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required.

Outcome assessment

Incident prostate cancers were initially self-reported on questionnaires, followed by cancer diagnosis adjudication by extracting clinical and treatment information from medical records and pathology reports (21). Archival prostate tumor tissue from approximately half of prostate cancer patients in HPFS was retrieved and underwent central histopathologic review by study pathologists for the standardized tumor grading.

ERG and PTEN immunohistochemistry

ERG and PTEN status was determined by immunohistochemistry (IHC) using previously constructed and validated tumor tissue microarrays (TMAs) (22). Tumors were classified as ERG-positive if the case had positive ERG staining (antibody: Clone EPR3864, Epitomics, Inc., Burlingame, CA) within prostate cancer epithelial cells on at least one TMA core (3 cores per subject). ERG IHC status is strongly associated with fusion status as assessed by fluorescence *in situ* hybridization (23). Tumors were classified as PTEN-loss if PTEN IHC expression (antibody: Clone D4.3 XP; Cell Signaling Technologies, Danvers, MA) was either markedly decreased or entirely lost across >10% of tumor cells compared with surrounding benign glands or stroma (8). PTEN IHC status is strongly associated with PTEN homozygous genetic deletion (8,24).

Assessment of metabolites

Plasma metabolites were profiled at the Broad Institute (Cambridge, MA) using untargeted liquid chromatography tandem mass spectrometry methods as described previously (25). Briefly, two methods were applied for the measurement of circulating metabolites: 1) amines and polar metabolites that ionize in the positive ion mode were measured using an liquid chromatography-mass spectrometry (LC-MS) platform comprised of an Open Accela 1250 U-HPLC coupled with a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific; Waltham, MA); 2) polar and non-polar lipids were measured using an LC-MS platform comprised of a Shimazu Nexera X2 U-HPLC coupled to an Exactive Plus orbitrap mass spectrometer (Thermo Fisher Scientific; Waltham, MA). We dropped unknown metabolites, metabolites with coefficient of variation (CV) higher than 25% or intraclass correlation coefficient (ICC) less than 0.4 as indicators of interassay reproducibility and within-person reproducibility, metabolites with undetectable levels in >10% of participants, and metabolites that could not be reasonably reproducible in samples with delayed processing in previous pilot tests (25,26). Metabolite values below the detection limit were assigned a half value of detection limit.

Statistical analyses

To normalize the distribution of metabolite data, the relative abundance of each metabolite was initially transformed using the natural logarithm, and then scaled to a mean value=0 and a standard deviation (SD)=1. To evaluate whether a subset of metabolites was associated with molecular subtype of prostate cancer, metabolites were grouped into categories based on their biological class. First, for all individual metabolites, multinomial logistic regression models were applied using the R package “multinom” to obtain *P*-values for the association between each individual metabolite and risk of prostate cancer by ERG and PTEN status. Second, enrichment analysis was performed using Mean-rank Gene Set Test (GSA) implemented with the R package “geneSetTest”. This analysis ranks individual metabolites by their associated *P*-value (from previous step) and identifies metabolites classes whose metabolites are highly ranked relative to other metabolites (27) in the ranked list. Metabolite classes with nominal *P*-values less than 0.05 were included for the further analyses. Multinomial logistic regression models were applied using the R package “multinom” to estimate odds ratios (ORs) and 95% confidence intervals (CI) for the associations of individual metabolites with prostate cancer by ERG and PTEN status. In both the group analyses and individual analyses, we adjusted for the following covariates, assessed immediately before or at the time of blood draw: age at blood draw, cohort, height, body mass index (BMI), physical activity, smoking status, fasting status, and season.

We conducted sensitivity analyses to evaluate the robustness of our results. First, to minimize the likelihood that latent disease influenced metabolite measures, we excluded cases whose time from blood collection to disease onset was less than 3 years (79 ERG defined cases, 52 PTEN defined cases). Second, to assess potential heterogeneity by diabetes, which has been reported to be inversely associated with risk of prostate cancer (28) and associated with diverse alterations in the metabolome (29), we excluded participants who had self-reported diabetes at the time of blood draw (14 control, 5 ERG-defined cases, 5 PTEN-defined cases). Third, to assess potential heterogeneity by race, we restricted the analysis to white men (excluded 27 control, 20 ERG defined cases, 10 PTEN defined cases).

All tests of statistical significance were two-sided. To account for multiple comparisons, the false discovery rate was controlled to 0.05 using the Benjamini & Hochberg approach. All analyses were conducted using SAS 9.4 software (SAS Institute, Cary, NC) and R 3.6.1 (cran.r-project.org). *P*-values were considered significant at values <0.05.

Results

Baseline characteristics

Blood-based metabolomics data were available for 277 cases with ERG status, 211 cases with PTEN status, and 294 controls. Compared with men without prostate cancer, men with prostate cancer were less likely to be diabetics and current smokers and included a higher proportion of men who provided fasting blood samples and men who delivered samples in winter (Table 1). As expected, the proportion of PSA level greater than 4 ng/ml at diagnosis and the proportion of advanced and high-grade cancers were higher for PTEN-loss *vs.* PTEN-intact cases, but similar by ERG status.

Enrichment analysis

We classified 223 known metabolites into 15 categories based on their chemical class, including sphingomyelin, ceramides, unsaturated triacylglycerols, saturated triacylglycerols, unsaturated diacylglycerols, saturated diacylglycerols, phosphatidylcholines, phosphatidylethanolamines, lysophosphatidylcholines, lysophosphatidylethanolamines, cholesterol esters, amino acids, amino acid derivatives, carnitines, as well as others (Supplementary table 1). Table 2 shows the nominal and FDR-corrected *P*-values for the associations between the metabolite categories and ERG and PTEN defined prostate cancer, compared with controls. We found an enrichment of sphingomyelins (*P* 0.01, FDR-*P* 0.15), ceramides (*P* 0.04, FDR-*P* 0.18), and phosphatidylethanolamines (*P* 0.03, FDR-*P* 0.18) in serum of men who developed ERG-positive prostate cancer compared to controls. In contrast, there were no nominally significant associations with ERG-negative disease.

For PTEN defined prostate cancer, unsaturated triacylglycerols were associated with both PTEN-loss (*P* 0.001, FDR-*P* 0.01) and PTEN-intact cases (*P* 0.0001, FDR-*P* 0.002) compared to non-cancer controls. Sphingomyelins (*P* 0.02, FDR-*P* 0.09), ceramides (*P* 0.005, FDR-*P* 0.04), and amino acids (*P* 0.02, FDR-*P* 0.09) were associated with PTEN-loss cases, while unsaturated diacylglycerols were associated with PTEN-intact cases (*P* 0.003, FDR-*P* 0.02).

Individual metabolites associated with ERG defined prostate cancer

Figure 1 shows associations of individual metabolites in the metabolite classes which nominal *P*-values less than 0.05 with risk of ERG defined prostate cancer. Among the ten sphingomyelins metabolites, eight were positively associated with ERG-positive prostate cancer. Specifically, higher risks were observed for men with higher concentrations of C24:0 SM (OR 1.29, 95% CI 1.03–1.62), C18:2 SM (OR 1.26, 95% CI 1.00–1.58), and C22:1 SM (OR 1.26, 95% CI 1.01–1.59). Additionally, all (n=4) ceramide metabolites were positively associated with ERG-positive cancer, while 19 of 21 phosphatidylethanolamine were inversely associated with ERG-positive cancer. Compared to ERG-positive prostate cancer cases, the associations of individual metabolites for the risk of ERG-negative prostate cancer in the sphingomyelins, ceramide, and phosphatidylethanolamine classes were weaker and conflicting. For example, except C18:0 SM (OR 0.81, 95% CI 0.65–1.00), there was no metabolite significantly associated with the risk of ERG-negative prostate cancer; higher risk of ERG-positive prostate cancer was observed for men with higher concentrations of C22:1 SM, but not for ERG-negative prostate cancer (OR 0.98, 95% CI 0.80–1.21). None of the metabolites remained statistically significantly associated with ERG defined prostate cancer after accounting for multiple comparisons (Supplementary table 2), and association of individual metabolites with ERG-defined prostate cancer in the non-significant metabolic classes were shown in the Supplementary table 3.

Individual metabolites associated with PTEN defined prostate cancer

Figure 2 shows associations of individual metabolites in the metabolite classes which nominal *P*-values less than 0.05 with PTEN defined prostate cancer. For the PTEN-loss prostate cancer, among the ten sphingomyelins metabolites, eight were showed positive association. Specifically, C18:2 SM (OR 1.64, 95% CI 1.11–2.42), C22:1 SM (OR 1.60,

95% CI 1.07–2.39), and C22:0 SM (OR 1.49, 95% CI 1.01–2.22) were positively associated with PTEN-loss prostate cancer diagnosis with nominal significance (Figure 2A). Meanwhile, all ceramide (n=4) and unsaturated diacylglycerol (n=11) metabolites, and 92% metabolites in the amino acids class were also positively associated with PTEN-loss cancer (Figure 2B–D). However, among unsaturated triacylglycerols, the direction of the association was diverse, for example, C52:3 TAG (OR 1.54, 95% CI 1.02–2.32) and C48:3 TAG (OR 1.52, 95% CI 1.01–2.29) were positively associated with PTEN-loss prostate cancer, while C58:8 TAG was inversely associated with PTEN-loss cancer (OR 0.65, 95% CI 0.45–0.95) (Figure 2E). It's interesting that individual metabolites which significantly associated with PTEN-loss prostate cancer showed much weaker associations with PTEN-intact prostate cancer. Such as C58:11 TAG (PTEN-loss: OR 0.69, 95% CI 0.50–0.97; PTEN-intact: OR 0.90, 95% CI 0.73–1.11). None of the metabolites remained statistically significantly associated with PTEN defined prostate cancer after accounting for multiple comparisons (Supplementary table 4), and the association of individual metabolites with PTEN-defined prostate cancer in the non-significant metabolic classes were shown in the Supplementary table 5.

Sensitivity analysis

We performed a number of sensitivity analyses, excluding subjects who had cancer diagnosed within 3 years after blood collection, excluding subjects with a history of diabetes, and restricting the analysis to self-reported white men. We did not observe qualitative differences in associations for the identified metabolites groups. In particular, the enrichment of sphingomyelins for ERG-positive prostate cancer, as well as ceramides and unsaturated triacylglycerols for PTEN-defined prostate cancers were robust in all three sensitivity analyses (Supplementary table 6 and 7).

Discussion

In this population-based case-control study, we investigated whether pre-diagnostic circulating metabolites were associated with the development of ERG or PTEN defined prostate cancer. Among the 15 predefined metabolite groups, sphingomyelins, ceramides, and phosphatidylethanolamines were enriched in ERG-positive prostate cancer cases compared to controls. Unsaturated triacylglycerols were enriched in both PTEN-loss and PTEN-intact prostate cancer, whereas sphingomyelins, ceramides, and amino acids were enriched in PTEN-loss prostate cancer and unsaturated diacylglycerols were enriched in PTEN-intact prostate cancer. Although no individual metabolites were statistically significantly associated with ERG or PTEN defined prostate cancer after accounting for multiple testing, the patterns of metabolites associated with these subtypes remained. These results suggest that alterations in overall pathways of metabolites, rather than individual metabolites, are better able to define the systemic metabolic state associated with ERG or PTEN prostate cancer.

Ten nested case-control studies have been conducted to prospectively examine the association between circulating metabolites and prostate cancer risk, including three studies in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (30–32),

three in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) cohort (33–35), and studies from the Janus Serum Bank Cohort in Norway (36), Prostate, Lung, Colorectal, and Ovarian Cancer Screening (PLCO) Trial (37), Cancer Prevention Study-II Nutrition Cohort (38), and Northern Sweden Health and Disease Study (39). However, none of the studies have examined the associations for molecular subtyped prostate cancers, and some findings could not be replicated in different cohorts. For example, among the two studies which applied GSA similar to our study to measure associations between chemical classes of metabolites and risk of prostate cancer, ATBC reported inverse associations between energy and lipid compounds and prostate cancer (34), however, PLCO found a very different metabolite-risk profile that featured primarily amino acids and peptides (37). The differential findings across studies may reflect differences in metabolomic assay tools, biospecimen processing, study population, and metabolomic approaches for dealing with metabolites and case identification (screening detected versus clinically detected). However, it is also possible that the molecular subtype of prostate cancer cases might contribute to the interpretation for the distinct patterns in different cohorts.

Large-scale cohort studies have found that high levels of pre-diagnostic circulating ligand insulin growth factor 1 (IGF-1) are associated with an increased risk of prostate cancer (40). In addition, our previous findings suggested that ERG-positive tumors are characterized by altered metabolic signaling pathways, such as higher expression of IR and IGF-1R, compared with ERG-negative tumors (14). Meanwhile, IGF-1 signaling downregulates the expression of PTEN as the PI3K/PTEN/Akt pathway is the downstream of IGF1/IGF-1R (41). Given that insulin sensitivity is strongly influenced by the presence of insulin receptor, IGFs, and IGF receptors (42), we hypothesized that metabolites correlated with insulin sensitivity or IGF signaling might be differently enriched in the pre-diagnosed samples for prostate cancer patients with ERG or PTEN subtypes.

Sphingomyelin is generated from ceramide by the transfer of phosphocholine from phosphatidylcholine with the generation of diacylglycerol through sphingomyelin synthase. Accumulation of metabolites in the sphingomyelin signaling pathway including ceramides, sphingomyelins and diacylglycerols were found to be involved in the development of insulin resistance (43,44), and several metabolites in sphingomyelins and ceramides classes were reported to be positively related to the risk of fatal prostate cancer (31,38), which might contribute to the interpretation for the positive associations of sphingomyelins and ceramides with risk of EGR-positive and PTEN-loss prostate cancer in our study. In addition, although the weak and diverse directions for the associations between metabolites in the unsaturated diacylglycerols class and the risk of fatal or advanced prostate cancer were reported by Wang Y et al. (38), we observed that unsaturated diacylglycerols, were non-significantly, positively associated with both PTEN-loss and -intact prostate cancer risk, though, owing to the limited case number, the association needs further validation.

The association between amino acids and prostate cancer risk has been inconsistent. Alanine, lysine, methionine, phenylalanine, arginine, and tryptophan were inversely associated with prostate cancer in ATBC or PLCO (34,37), while these results were not replicated in EPIC (31) or in our study. Moreover, the enrichment of amino acids in PTEN-loss prostate cancer was not replicated in our sensitivity analyses.

Phosphatidylethanolamines, an abundant membrane phospholipid that is essential for membrane integrity, cell division, and membrane protein topology, has been identified as one of the biomarkers for diagnosis of prostate cancer with a high sensitivity, specificity and accuracy (45). Dalmau et al found that when prostate cancer cells undergo epithelial to mesenchymal transition (EMT) induction, which plays a crucial role in cancer metastasis, the level of twelve unsaturated triacylglycerols were increased in cells (46). Among the eleven metabolites that overlapped in our study, C52:3 TAG was found to be positively associated with PTEN-loss prostate cancer risk with nominal $P < 0.05$, seven others were found to be positively but non-significantly associated both PTEN-loss and -intact prostate cancer risk.

There are several potential limitations in our study. First, the sample size was modest, particularly for the PTEN-loss prostate cancer cases. Although differential associations of metabolites were observed according to ERG and PTEN status, no associations reached statistical significance after accounting for multiple comparisons. Given that, it's more difficult to explore the metabolomic profiles for prostate cancer cases co-occurred ERG-positive and PTEN-loss. However, we conducted an exploratory analysis and divided cases into PTEN intact/ERG-negative (n=182) and PTEN-loss & ERG-positive (n=22) group. The results showed that for the PTEN-loss & ERG-positive group, the significant enrichment of ceramides ($P < 0.001$), but not Sphingomyelins ($P = 0.12$) were observed; for the PTEN intact/ERG-negative group, unsaturated triacylglycerols ($P < 0.001$) and unsaturated diacylglycerols ($P = 0.004$) were enriched. A larger epidemiological study with access to both pre-diagnostic bloods and tumor tissue materials for molecular subtyping will be needed to confirm what we observed in this study. Such effort may require the pooling of data across cohorts. Second, more than 90% of the men in our study are of European descent, which may limit the generalization of the results to other populations. This is particularly important given that the prevalence of *TMPRSS2: ERG* is lower in men of African and Asian ancestry. The strengths of our study include the combination of detailed longitudinal data on external exposures, pre-diagnostic metabolite profiling, and molecular defined outcomes together for the innovative analysis. The inclusion of well captured information on BMI, smoking status, fasting status, and the season of blood draw allowed us to evaluate associations independent of these factors.

In summary, the present study demonstrates the distinct metabolomic profiles associated with ERG and PTEN defined prostate cancer in men nested from two prospective cohorts with long-term follow-up. Metabolites in the sphingomyelin signaling pathway, especially sphingomyelins and ceramides, appear to be positively associated with ERG-positive and PTEN-loss prostate cancer, and might contribute to the development of prostate cancer in different molecular subtypes. Larger studies are needed to confirm our findings, and further our understanding of the environmental factors that contribute to the etiology of distinct molecular subtypes of prostate cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY; The authors assume full responsibility for analyses and interpretation of these data.

Financial support: The Health Professionals Follow-up Study is supported by U01 CA 167552 from the National Cancer Institute. The tissue microarrays were constructed by the Tissue Microarray Core Facility at the Dana-Farber/Harvard Cancer Center (P30 CA 006516). This project was supported in part by the Dana-Farber/Harvard Cancer Center SPORE in Prostate Cancer (P50 090381) and the US Army Prostate Cancer Research Program. X Feng was supported by the program of China Scholarships Council (No.201806210455). L.A. Mucci, K.L. Penney, S.P. Finn, and K.M. Wilson were Prostate Cancer Foundation Young Investigators. D.R. Schmidt is supported by the Harvard Catalyst/Harvard Clinical and Translational Science Center (NIH Award KL2 TR002542). B.A. Dickerman is supported by National Institutes of Health Grant (K99 CA248335).

Reference

1. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science (New York, NY)* 2005;310(5748):644–8.
2. Zhou CK, Young D, Yeboah ED, Coburn SB, Tettey Y, Biritwum RB, et al. TMPRSS2:ERG Gene Fusions in Prostate Cancer of West African Men and a Meta-Analysis of Racial Differences. *Am J Epidemiol* 2017;186(12):1352–61. [PubMed: 28633309]
3. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* 2015;163(4):1011–25. [PubMed: 26544944]
4. Wang Y, Dai B. PTEN genomic deletion defines favorable prognostic biomarkers in localized prostate cancer: a systematic review and meta-analysis. *International journal of clinical and experimental medicine* 2015;8(4):5430–7. [PubMed: 26131120]
5. Song C, Chen H. Predictive significance of TMRPSS2-ERG fusion in prostate cancer: a meta-analysis. *Cancer Cell Int* 2018;18(177):018–0672.
6. Squire JA. TMPRSS2-ERG and PTEN loss in prostate cancer. *Nature genetics* 2009;41(5):509–10. [PubMed: 19399032]
7. Yoshimoto M, Joshua AM, Cunha IW, Coudry RA, Fonseca FP, Ludkovski O, et al. Absence of TMPRSS2:ERG fusions and PTEN losses in prostate cancer is associated with a favorable outcome. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2008;21(12):1451–60.
8. Ahearn TU, Pettersson A, Ebot EM, Gerke T, Graff RE, Morais CL, et al. A Prospective Investigation of PTEN Loss and ERG Expression in Lethal Prostate Cancer. *JNCI: Journal of the National Cancer Institute* 2016;108(2):djv346. [PubMed: 26615022]
9. Egbers L, Luedeke M, Rinckleb A, Kolb S, Wright JL, Maier C, et al. Obesity and Prostate Cancer Risk According to Tumor TMPRSS2:ERG Gene Fusion Status. *Am J Epidemiol* 2015;181(9):706–13. [PubMed: 25852077]
10. Graff RE, Ahearn TU, Pettersson A, Ebot EM, Gerke T, Penney KL, et al. Height, Obesity, and the Risk of TMPRSS2:ERG-Defined Prostate Cancer. *Cancer Epidemiol Biomarkers Prev* 2018;27(2):193–200. [PubMed: 29167279]
11. Graff RE, Pettersson A, Lis RT, Ahearn TU, Markt SC, Wilson KM, et al. Dietary lycopene intake and risk of prostate cancer defined by ERG protein expression. *Am J Clin Nutr* 2016;103(3):851–60. [PubMed: 26817504]
12. Pernar CH, Ebot EM, Pettersson A, Graff RE, Giunchi F, Ahearn TU, et al. A Prospective Study of the Association between Physical Activity and Risk of Prostate Cancer Defined by Clinical Features and TMPRSS2:ERG. *Eur Urol* 2019; 76(1):33–40. [PubMed: 30301696]
13. Allott EH, Ebot EM, Stopsack KH, Gonzalez-Feliciano AG, Markt SC, Wilson KM, et al. Statin use is associated with lower risk of PTEN-null and lethal prostate cancer. *Clin Cancer Res* 2020; 26(5):1086–1093. [PubMed: 31754047]

14. Pettersson A, Lis RT, Meisner A, Flavin R, Stack EC, Fiorentino M, et al. Modification of the association between obesity and lethal prostate cancer by TMPRSS2:ERG. *J Natl Cancer Inst* 2013;105(24):1881–90. [PubMed: 24292212]
15. Penney KL, Pettersson A, Shui IM, Graff RE, Kraft P, Lis RT, et al. Association of Prostate Cancer Risk Variants with TMPRSS2:ERG Status: Evidence for Distinct Molecular Subtypes. *Cancer Epidemiol Biomarkers Prev* 2016;25(5):745–9. [PubMed: 26941365]
16. Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. A prospective cohort study of vasectomy and prostate cancer in US men. *Jama* 1993;269(7):873–7. [PubMed: 8426446]
17. Manson JE, Buring JE, Satterfield S, Hennekens CH. Baseline characteristics of participants in the Physicians' Health Study: a randomized trial of aspirin and beta-carotene in U.S. physicians. *Am J Prev Med* 1991;7(3):150–4. [PubMed: 1931143]
18. Rimm EB, Stampfer MJ, Colditz GA, Giovannucci E, Willett WC. Effectiveness of various mailing strategies among nonrespondents in a prospective cohort study. *Am J Epidemiol* 1990;131(6):1068–71. [PubMed: 2343859]
19. Platz EA, Pollak MN, Willett WC, Giovannucci E. Vertex balding, plasma insulin-like growth factor 1, and insulin-like growth factor binding protein 3. *J Am Acad Dermatol* 2000;42(6):1003–7. [PubMed: 10827403]
20. Ridker PM, Vaughan DE, Stampfer MJ, Sacks FM, Hennekens CH. A cross-sectional study of endogenous tissue plasminogen activator, total cholesterol, HDL cholesterol, and apolipoproteins A-I, A-II, and B-100. *Arterioscler Thromb* 1993;13(11):1587–92. [PubMed: 8218099]
21. Wilson KM, Kasperzyk JL, Rider JR, Kenfield S, van Dam RM, Stampfer MJ, et al. Coffee consumption and prostate cancer risk and progression in the Health Professionals Follow-up Study. *Journal of the National Cancer Institute* 2011;103(11):876–84. [PubMed: 21586702]
22. Pettersson A, Graff RE, Bauer SR, Pitt MJ, Lis RT, Stack EC, et al. The TMPRSS2:ERG rearrangement, ERG expression, and prostate cancer outcomes: a cohort study and meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2012;21(9):1497–509. [PubMed: 22736790]
23. Park K, Tomlins SA, Mudaliar KM, Chiu YL, Esgueva R, Mehra R, et al. Antibody-based detection of ERG rearrangement-positive prostate cancer. *Neoplasia* 2010;12(7):590–8. [PubMed: 20651988]
24. Lotan TL, Gurel B, Sutcliffe S, Esopi D, Liu W, Xu J, et al. PTEN protein loss by immunostaining: analytic validation and prognostic indicator for a high risk surgical cohort of prostate cancer patients. *Clin Cancer Res* 2011;17(20):6563–73. [PubMed: 21878536]
25. Townsend MK, Clish CB, Kraft P, Wu C, Souza AL, Deik AA, et al. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. *Clin Chem* 2013;59(11):1657–67. [PubMed: 23897902]
26. Mayers JR, Wu C, Clish CB, Kraft P, Torrence ME, Fiske BP, et al. Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nat Med* 2014;20(10):1193–8. [PubMed: 25261994]
27. Michaud J, Simpson KM, Escher R, Buchet-Poyau K, Beissbarth T, Carmichael C, et al. Integrative analysis of RUNX1 downstream pathways and target genes. *BMC Genomics* 2008;9(363):1471–2164.
28. Jian Gang P, Mo L, Lu Y, Runqi L, Xing Z. Diabetes mellitus and the risk of prostate cancer: an update and cumulative meta-analysis. *Endocr Res* 2015;40(1):54–61. [PubMed: 25105463]
29. Guasch-Ferre M, Hruby A, Toledo E, Clish CB, Martinez-Gonzalez MA, Salas-Salvado J, et al. Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. *Diabetes Care* 2016;39(5):833–46. [PubMed: 27208380]
30. Kuhn T, Floegel A, Sookthai D, Johnson T, Rolle-Kampczyk U, Otto W, et al. Higher plasma levels of lysophosphatidylcholine 18:0 are related to a lower risk of common cancers in a prospective metabolomics study. *BMC Med* 2016;14(13):016–0552.
31. Schmidt JA, Fensom GK, Rinaldi S, Scalbert A, Appleby PN, Achaintre D, et al. Pre-diagnostic metabolite concentrations and prostate cancer risk in 1077 cases and 1077 matched controls in the European Prospective Investigation into Cancer and Nutrition. *BMC Med* 2017;15(1):122. [PubMed: 28676103]

32. Schmidt JA, Fensom GK, Rinaldi S, Scalbert A, Appleby PN, Achaintre D, et al. Patterns in metabolite profile are associated with risk of more aggressive prostate cancer: A prospective study of 3,057 matched case-control sets from EPIC. *Int J Cancer* 2019;5(10):32314.
33. Mondul AM, Moore SC, Weinstein SJ, Mannisto S, Sampson JN, Albanes D. 1-stearoylglycerol is associated with risk of prostate cancer: results from serum metabolomic profiling. *Metabolomics* 2014;10(5):1036–41. [PubMed: 25254003]
34. Mondul AM, Moore SC, Weinstein SJ, Karoly ED, Sampson JN, Albanes D. Metabolomic analysis of prostate cancer risk in a prospective cohort: The alpha-tocopherol, beta-carotene cancer prevention (ATBC) study. *Int J Cancer* 2015;137(9):2124–32. [PubMed: 25904191]
35. Huang J, Mondul AM, Weinstein SJ, Derkach A, Moore SC, Sampson JN, et al. Prospective serum metabolomic profiling of lethal prostate cancer. *Int J Cancer* 2019;145(12):3231–43. [PubMed: 30779128]
36. de Vogel S, Ulvik A, Meyer K, Ueland PM, Nygard O, Vollset SE, et al. Sarcosine and other metabolites along the choline oxidation pathway in relation to prostate cancer--a large nested case-control study within the JANUS cohort in Norway. *Int J Cancer* 2014;134(1):197–206. [PubMed: 23797698]
37. Huang J, Mondul AM, Weinstein SJ, Koutros S, Derkach A, Karoly E, et al. Serum metabolomic profiling of prostate cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. *Br J Cancer* 2016;115(9):1087–95. [PubMed: 27673363]
38. Wang Y, Jacobs EJ, Carter BD, Gapstur SM, Stevens VL. Plasma Metabolomic Profiles and Risk of Advanced and Fatal Prostate Cancer. *European urology oncology* 2019; S2588–9311(19)30108–7.
39. Röhnisch HE, Kyrø C, Olsen A, Thysell E, Hallmans G, Moazzami AA. Identification of metabolites associated with prostate cancer risk: a nested case-control study with long follow-up in the Northern Sweden Health and Disease Study. *BMC Med* 2020;18(1):187. [PubMed: 32698845]
40. Travis RC, Appleby PN, Martin RM, Holly JMP, Albanes D, Black A, et al. A Meta-analysis of Individual Participant Data Reveals an Association between Circulating Levels of IGF-I and Prostate Cancer Risk. *Cancer research* 2016;76(8):2288–300. [PubMed: 26921328]
41. Zu K, Martin NE, Fiorentino M, Flavin R, Lis RT, Sinnott JA, et al. Protein expression of PTEN, insulin-like growth factor I receptor (IGF-IR), and lethal prostate cancer: a prospective study. *Cancer Epidemiol Biomarkers Prev* 2013;22(11):1984–93. [PubMed: 23983239]
42. Rajpathak SN, Gunter MJ, Wylie-Rosett J, Ho GY, Kaplan RC, Muzumdar R, et al. The role of insulin-like growth factor-I and its binding proteins in glucose homeostasis and type 2 diabetes. *Diabetes Metab Res Rev* 2009;25(1):3–12. [PubMed: 19145587]
43. Straczkowski M, Kowalska I, Nikolajuk A, Dzienis-Straczkowska S, Kinalska I, Baranowski M, et al. Relationship between insulin sensitivity and sphingomyelin signaling pathway in human skeletal muscle. *Diabetes* 2004;53(5):1215–21. [PubMed: 15111489]
44. Sokolowska E, Blachnio-Zabielska A. The Role of Ceramides in Insulin Resistance. *Front Endocrinol* 2019;10(577).
45. Zhou X, Mao J, Ai J, Deng Y, Roth MR, Pound C, et al. Identification of plasma lipid biomarkers for prostate cancer by lipidomics and bioinformatics. *PLoS One* 2012;7(11):e48889. [PubMed: 23152813]
46. Dalmau N, Jaumot J, Tauler R, Bedia C. Epithelial-to-mesenchymal transition involves triacylglycerol accumulation in DU145 prostate cancer cells. *Mol Biosyst* 2015;11(12):3397–406. [PubMed: 26474270]

A. Sphingomyelins

ERG-positive vs. Control

ERG-negative vs. Control

B. Ceramides

C. Phosphatidylethanolamines

OR per 1 SD

Figure 1. Lollipop plot of odds ratios (per 1 SD) for individual metabolites in nominally significant metabolic classes associated with risk of developing ERG-positive (n=128) or ERG-negative (n=149) prostate cancer. (A) Sphingomyelins. (B) Ceramides. (C) Phosphatidylethanolamines. Adjusted for age at blood draw, cohort (HPFS/ PHS), height (continuous), BMI (continuous), physical activity (MET-hr/week, continuous), cigarette smoking status (never, past, current, missing), fasting status (<8 hours, 8 hours, missing), season (winter, spring, summer, fall). *nominal $P < 0.05$.

Figure 2.

Lollipop plot of odds ratios (per 1 SD) for individual metabolites in nominally significant metabolic classes with risk of developing PTEN-loss (n=34) or PTEN-intact (n=177) prostate cancers. (A) Sphingomyelins. (B) Ceramides. (C) Unsaturated Diacylglycerols. (D) Amino acids. (E) Unsaturated Triacylglycerols.

Adjusted for age at blood draw, cohort (HPFS/ PHS), height (continuous), BMI (continuous), physical activity (MET-hr/week, continuous), cigarette smoking status (never, past, current, missing), fasting status (<8 hours, 8 hours, missing), season (winter, spring, summer, fall).

*nominal $P<0.05$.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Baseline characteristics of controls and prostate cancer cases by molecular subtype, Health Professionals Follow-up Study (HPFS) and Physicians' Health Study (PHS)

	Controls	Prostate cancer cases			
		ERG-positive	ERG-negative	PTEN-loss	PTEN-intact
N	294	128	149	34	177
Cohort, N (%)					
<i>HPFS</i>	200 (68.0)	115 (89.8)	127 (85.2)	28 (82.4)	166 (93.8)
<i>PHS</i>	94 (32.0)	13 (10.2)	22 (14.8)	6 (17.6)	11 (6.2)
Age at blood draw, mean (SD)	62 (9.5)	59 (7.5)	61 (6.7)	61 (6.8)	60 (7.0)
Height, centimeters, mean (SD)	177.3 (6.9)	178.5 (6.5)	178.5 (6.4)	176.5 (4.8)	179.2 (6.3)
BMI at blood draw, kg/m², mean (SD)	25.3 (3.3)	25.7 (3.2)	25.5 (3.0)	25.5 (2.5)	25.7 (3.2)
Physical activity, MET-hr/week^a, mean (SD)	28.7 (26.0)	32.5 (33.1)	35.5 (28.1)	41.1 (38.5)	31.6 (28.4)
White race, N (%)	267 (90.8)	119 (93.0)	138 (92.6)	33 (97.1)	168 (94.9)
History of diabetes, N (%)	14 (4.8)	2 (1.6)	3 (2.0)	1 (2.9)	4 (2.3)
Smoking status at blood draw, N (%)					
<i>never</i>	132 (44.9)	63 (49.2)	76 (51.0)	14 (41.2)	92 (52.0)
<i>past</i>	132 (44.9)	50 (39.0)	61 (40.9)	17 (50.0)	67 (37.8)
<i>current</i>	20 (6.8)	8 (6.3)	8 (5.4)	1 (2.9)	11 (6.2)
<i>Missing</i>	10 (3.4)	7 (5.5)	4 (2.7)	2 (5.9)	7 (4.0)
Season of blood draw, %					
<i>Winter</i>	42 (14.3)	33 (25.8)	23 (15.4)	10 (29.4)	33 (18.7)
<i>Spring</i>	41 (13.9)	12 (9.4)	28 (18.8)	6 (17.6)	25 (14.1)
<i>Summer</i>	62 (21.1)	28 (21.9)	40 (26.9)	7 (20.6)	42 (23.7)
<i>Fall</i>	149 (50.7)	55 (42.9)	58 (38.9)	11 (33.4)	77 (43.5)
Fasting status, hours, %					
<i><8</i>	127 (43.2)	53 (41.4)	54 (36.2)	12 (35.3)	64 (36.2)
<i>8</i>	148 (50.3)	68 (53.1)	80 (53.7)	20 (58.8)	103 (58.2)
<i>Missing</i>	19 (6.5)	7 (5.5)	15 (10.1)	2 (5.9)	10 (5.6)
Time from blood draw to cancer diagnosis, years, mean (SD)	NA	5.6 (4.0)	5.9 (4.4)	5.7 (4.4)	5.7 (3.6)
PSA level at diagnosis, ng/mL					
<i><4</i>	NA	18 (14.1)	17 (11.4)	1 (2.9)	24 (13.5)
<i>4–9.9</i>	NA	72 (56.2)	76 (51.0)	18 (52.9)	101 (57.1)
<i>10</i>	NA	29 (22.7)	42 (28.2)	14 (41.2)	20 (22.6)
<i>Missing</i>	NA	9 (7.0)	14 (9.4)	1 (2.9)	12 (6.8)
Advanced prostate cancer^b, %	NA	31 (24.2)	32 (21.5)	16 (47.1)	34 (19.2)
High-grade prostate cancer^c, %	NA	54 (42.2)	61 (40.9)	25 (73.5)	69 (39.0)

^aMET–hr/week: metabolic equivalent of task (MET) hours per week.

^bPathological or clinical T3b, T4, N1 or M1 or metastasis to other organs over follow-up or death.

^cGleason score $\geq 4+3$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Enrichment analysis for the associations of metabolomic classes with ERG and PTEN defined prostate cancer

Classes (n metabolites)	ERG-positive vs. Control*			ERG-negative vs. Control*			PTEN-loss vs. Control*			PTEN-intact vs. Control*		
	P	FDR-P	P	P	FDR-P	P	P	FDR-P	P	P	FDR-P	P
Lipids and lipid metabolites												
Sphingomyelins (10) ^{a, b}	0.01	0.15	0.07	0.07	0.85	0.02	0.09	0.09	0.52	0.97		
Ceramides (4) ^{a, b}	0.04	0.18	0.87	0.87	0.95	0.005	0.04	0.04	0.11	0.44		
Unsaturated Triacylglycerols (40) ^{b, c}	0.92	0.98	0.13	0.13	0.85	0.001	0.01	0.01	0.0001	0.002		
Saturated Triacylglycerols (5)	0.06	0.18	0.93	0.93	0.95	0.96	0.99	0.99	0.91	1.00		
Unsaturated Diacylglycerols (11) ^c	0.89	0.98	0.83	0.83	0.95	0.27	0.66	0.66	0.003	0.02		
Saturated Diacylglycerols (2)	0.13	0.34	0.84	0.84	0.95	0.92	0.99	0.99	0.85	1.00		
Phosphatidylcholines (36)	0.70	0.98	0.30	0.30	0.90	0.92	0.99	0.99	0.93	1.00		
Phosphatidylethanolamines (21) ^a	0.03	0.18	0.17	0.17	0.85	0.96	0.99	0.99	0.99	1.00		
Lysophosphatidylcholines (10)	0.97	0.98	0.95	0.95	0.95	0.94	0.99	0.99	0.97	1.00		
Lysophosphatidylethanolamines (6)	0.80	0.98	0.63	0.63	0.95	0.91	0.99	0.99	0.43	0.93		
Cholesterol esters (12)	0.17	0.36	0.83	0.83	0.95	0.80	0.99	0.99	0.78	1.00		
Amino acids (25) ^b	0.05	0.18	0.27	0.27	0.90	0.02	0.09	0.09	0.41	0.93		
Amino acid derivatives (6)	0.40	0.75	0.85	0.85	0.95	0.22	0.65	0.65	0.20	0.60		
Carnitines (23)	0.92	0.98	0.56	0.56	0.95	0.99	0.99	0.99	0.12	0.44		
Other (12)	0.98	0.98	0.37	0.37	0.93	0.40	0.86	0.86	1.00	1.00		

* Adjusted for age at blood draw, cohort (HPFS/ PHS), height (continuous), BMI (continuous), physical activity (MET-hr/week, continuous), cigarette smoking status (never, past, current, missing), fasting status (<8 hours, 8 hours, missing), season (winter, spring, summer, fall).

^a: nominal $P < 0.05$ for group of ERG+ VS. Control

^b: nominal $P < 0.05$ for group of PTEN loss VS. Control

^c: nominal $P < 0.05$ for group of PTEN intact VS. Control.

FDR, false discovery rate