AJCN. First published ahead of print June 28, 2017 as doi: 10.3945/ajcn.116.150912.

Nutritional metabolomics and breast cancer risk in a prospective study

Mary C Playdon,[1,2] *Regina G Ziegler,*[2] *Joshua N Sampson,*[2] *Rachael Stolzenberg-Solomon,*[2] *Henry J Thompson,*[3] *Melinda L Irwin,*[1,4] *Susan T Mayne,*[1,5] *Robert N Hoover,*[2] *and Steven C Moore*[2]

[1]Yale School of Public Health, Yale University, New Haven, CT; [2]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD; [3]Cancer Prevention Laboratory, Colorado State University, Fort Collins, CO; [4]Yale Cancer Center, New Haven, CT; and [5]US Food and Drug Administration, College Park, MD

ABSTRACT

Background: The epidemiologic evidence for associations between dietary factors and breast cancer is weak and etiologic mechanisms are often unclear. Exploring the role of dietary biomarkers with metabolomics can potentially facilitate objective dietary characterization, mitigate errors related to self-reported diet, agnostically test metabolic pathways, and identify mechanistic mediators.

Objective: The aim of this study was to evaluate associations of diet-related metabolites with the risk of breast cancer in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial.

Design: We examined prediagnostic serum concentrations of diet-related metabolites in a nested case-control study in 621 postmenopausal invasive breast cancer cases and 621 matched controls in the multicenter PLCO cohort. We calculated partial Pearson correlations between 617 metabolites and 55 foods, food groups, and vitamin supplements on the basis of the 2015 Dietary Guidelines for Americans and derived from a 137-item self-administered food-frequency questionnaire. Diet-related metabolites (P-correlation $< 1.47 \times 10^{-6}$) were evaluated in breast cancer analyses. ORs for the 90th compared with the 10th percentile were calculated by using conditional logistic regression, with body mass index, physical inactivity, other breast cancer risk factors, and caloric intake controlled for (false discovery rate <0.2).

Results: Of 113 diet-related metabolites, 3 were associated with overall breast cancer risk (621 cases): caprate (10:0), a saturated fatty acid (OR: 1.77; 95% CI = 1.28, 2.43); γ-carboxyethyl hydrochroman (γ-CEHC), a vitamin E (γ-tocopherol) derivative (OR: 1.64; 95% CI: 1.18, 2.28); and 4-androsten-3β,17β-diol-monosulfate (1), an androgen (OR: 1.61; 95% CI: 1.20, 2.16). Nineteen metabolites were significantly associated with estrogen receptor (ER)–positive (ER^+) breast cancer (418 cases): 12 alcohol-associated metabolites, including 7 androgens and α-hydroxyisovalerate (OR: 2.23; 95% CI: 1.50, 3.32); 3 vitamin E (tocopherol) derivatives (e.g., γ-CEHC; OR: 1.80; 95% CI: 1.20, 2.70); butter-associated caprate (10:0) (OR: 1.81; 95% CI: 1.23, 2.67); and fried food–associated 2-hydroxyoctanoate (OR: 1.46; 95% CI: 1.03, 2.07). No metabolites were significantly associated with ER-negative breast cancer (144 cases).

Conclusions: Prediagnostic serum concentrations of metabolites related to alcohol, vitamin E, and animal fats were moderately strongly associated with ER^+ breast cancer risk. Our findings show how nutritional metabolomics might identify diet-related exposures that modulate cancer risk. This trial was registered at clinicaltrials.gov as NCT00339495. *Am J Clin Nutr* doi: https://doi.org/10.3945/ajcn.116.150912.

Keywords: breast cancer, diet, nutrition, metabolomics, biomarker, tocopherol, alcohol, androgen, fat

INTRODUCTION

Diet has long been thought to play a role in breast cancer etiology on the basis of ecological and migrant studies, with support from animal experiments. Epidemiologic studies and randomized clinical trials, however, have been inconclusive for most dietary exposures (1, 2), possibly due to imprecision and inaccuracy in the self-reported dietary assessments used in population-based studies (3). The role of dietary fat has been particularly controversial. Large randomized controlled trials of low-fat or modified-fat (e.g., Mediterranean) diets have generated mixed findings, including inverse (4) or suggestive inverse associations with breast cancer incidence (5) and survival (6) and null associations (7–9). However, these interventions had diet adherence issues, leading to a loss of study power. Furthermore, each intervention tested a specific defined diet, rather than agnostically evaluating a full range of dietary patterns. Heterogeneity in diet–breast cancer associations was also observed by baseline fat intake and hormone receptor subtype (5); failure to

The study and analyses were supported by intramural funding provided to the Division of Cancer Epidemiology and Genetics of the National Cancer Institute and in part by the Breast Cancer Research Stamp Fund, as well as by the Yale–National Cancer Institute predoctoral training grant T32 CA105666 (to STM). None of the funding bodies had any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The opinions and conclusions expressed in this article are solely the views of the authors and do not necessarily reflect those of the Food and Drug Administration.

Supplemental Figures 1–3, Supplemental Tables 1–10, and Supplemental Methods are available from the "Online Supporting Material" link in the online posting of the article and from the same link in the online table of contents at http://ajcn.nutrition.org.

Address correspondence to MCP (e-mail: mary.playdon@nih.gov).

Abbreviations used: DQx, dietary questionnaire; ER, estrogen receptor; ER^+, estrogen receptor–positive; ER^-, estrogen receptor–negative; GGM, Gaussian graphical model; HEI-2010, Healthy Eating Index–2010; LOD, limit of detection; MHT, menopausal hormone therapy; PCA, principal components analysis; PLCO, Prostate, Lung, Colorectal and Ovarian; γ-CEHC, γ-carboxyethyl hydrochroman; 90vs10, 90th versus 10th percentile.

Received December 15, 2016. Accepted for publication May 30, 2017.
doi: https://doi.org/10.3945/ajcn.116.150912.

Copyright (C) 2017 by the American Society for Nutrition

consider heterogeneity may attenuate overall associations. Among dietary exposures, the only one for which there is consistent and convincing evidence for an association with breast cancer is alcohol consumption (1, 10). However, the mechanisms linking alcohol and breast cancer risk are incompletely understood.

Nutritional metabolomics—the study of diet-related metabolites toward the end of elucidating diet and health outcome relations (11–13)—may provide a means to more accurately capture exogenous dietary exposures, evaluate endogenous biomarkers that mediate diet-cancer relations, or both. Nutritional metabolomics allows for simultaneous evaluation of several diet-related exposures and the metabolic processes they influence. Recently, we used a nutritional metabolomics approach in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial cohort (14) and a colon adenoma case-control study (15) to identify serum metabolites that were significantly correlated with specific dietary exposures, including consumption of fruit, vegetables, meats, fish, nuts, fats, alcohol, coffee, and multivitamins.

No published prospective studies, to our knowledge, have used a nutritional metabolomics approach to examine diet-related metabolite associations with subsequent breast cancer risk. We therefore conducted a nested case-control study of prediagnostic serum concentrations of diet-related metabolites and postmenopausal breast cancer in women in the PLCO cohort. Our aims were as follows: *1*) to replicate and expand the diet-related metabolites identified in previous studies by using a more sensitive and comprehensive metabolomics platform; *2*) to evaluate relations between diet-related metabolites and breast cancer risk, overall and stratified by estrogen receptor (ER) status; and *3*) interpret these relations in terms of the underlying metabolic processes.

METHODS

Study design and population

The PLCO trial is a randomized multicenter trial that assigned 78,000 men and 77,000 women to either a cancer screening intervention or standard care between 1993 and 2001 (**Supplemental Figure 1**) (16). Participants were followed over time for multiple cancer outcomes. Eligible participants were aged 55–74 y at baseline, with no previous history of prostate, lung, colorectal, or ovarian cancer. The PLCO trial was registered at clinicaltrials.gov as NCT00339495. All of the participants provided informed consent and completed self-administered questionnaires at baseline about personal characteristics, health-related behaviors, medical history, and cancer risk factors. In the screening arm, participants completed a food-frequency questionnaire [PLCO dietary questionnaire (DQx)] (17) and provided blood samples at defined times during the study according to a standardized protocol (18). The study was approved by institutional review boards at the US National Cancer Institute and 10 participating screening centers. We conducted a nested case-control study of postmenopausal breast cancer in women who completed the baseline risk factor questionnaire and DQx, donated a blood sample 1 y after baseline, and had no history of breast or rare cancer before blood collection. All first primary invasive breast cancer cases (International Classification of Diseases, Ninth Revision, codes 174.0–174.9) among women not using menopausal hormone therapy (MHT) at baseline were included (n = 468 cases). To increase ER-negative (ER^-) cases, progesterone receptor (PR)–negative cases, or both, ER^- or PR-negative cases who used MHT at baseline were included (n = 153). With the use of incidence density sampling, 621 controls were matched to 621 cases (n = 418 ER^+, 144 ER^-, and 59 equivocal) on study baseline age (±2 y), date of blood collection (±3 mo), year of diagnosis, cohort entry, and MHT use at baseline.

Outcome ascertainment

Confirmed incident breast cancer cases were ascertained by annual questionnaire, state cancer registries, the National Death Index, and physician and next-of-kin reporting (18). We restricted analyses to cases confirmed by hospital records during follow-up to 19 November 2013. Tumors were considered ER^+ if ≥10% of cells were immunohistochemically stained positive. Equivocal results (1–9% positively stained) or indeterminate status were excluded from the hormone receptor–specific analyses.

Dietary assessment

The DQx measured the typical frequency of intake during the 12 mo before baseline for 137 foods, including alcohol, and typical portion sizes for 77 items (19–21). The frequency of intake and serving size of each food item were converted into grams per day of the food and categorized into 55 a priori–defined food groups on the basis of the 2015 Dietary Guidelines for Americans (**Supplemental Table 1**) (22). The DQx ascertained vitamin C, D, and E and calcium supplement intakes at baseline (continuous) and recent multivitamin use (now and 2 y ago, yes or no). We also calculated the Healthy Eating Index–2010 (HEI-2010) score (23). Participants missing quantitative responses to ≥8 food-frequency questions or who were in the highest or lowest 1% for energy intake were excluded from the diet-metabolite analysis (n = 115). For participants missing quantitative responses for <8 food-frequency questions, "0" was assigned to the missing items.

Metabolite assessment

Nonfasting serum samples were prospectively collected at the first follow-up, ∼1 y after the baseline visit. Blood samples were stored at −70°C and analyzed in 2014 (24). Parent compounds and their downstream metabolites with molecular weights <1 kDa, herein collectively referred to as "metabolites," were measured by using the Metabolon Inc. platform. These included amino acids, monosaccharides, small lipids, cofactors and vitamins, intermediates of energy metabolism, and exogenous chemical and food and plant xenobiotics. Liquid chromatography–tandem mass spectrometry and gas chromatography–mass spectroscopy measured spectral peaks (25) (detailed in **Supplemental Methods**), which were identified by comparing against a library of chemical standards (>4500 known metabolites and 9000 unidentified metabolites).

In the PLCO study, we measured 1057 serum metabolites (672 identifiable). Values below the limit of detection (LOD) were assigned the minimum of observed values for that metabolite (**Supplemental Figure 2**). Metabolites for which ≥90% of samples had values less than the LOD were excluded (n = 55), which left 617 metabolites of known identity for analysis. This exclusion criterion was chosen to allow for the assessment of

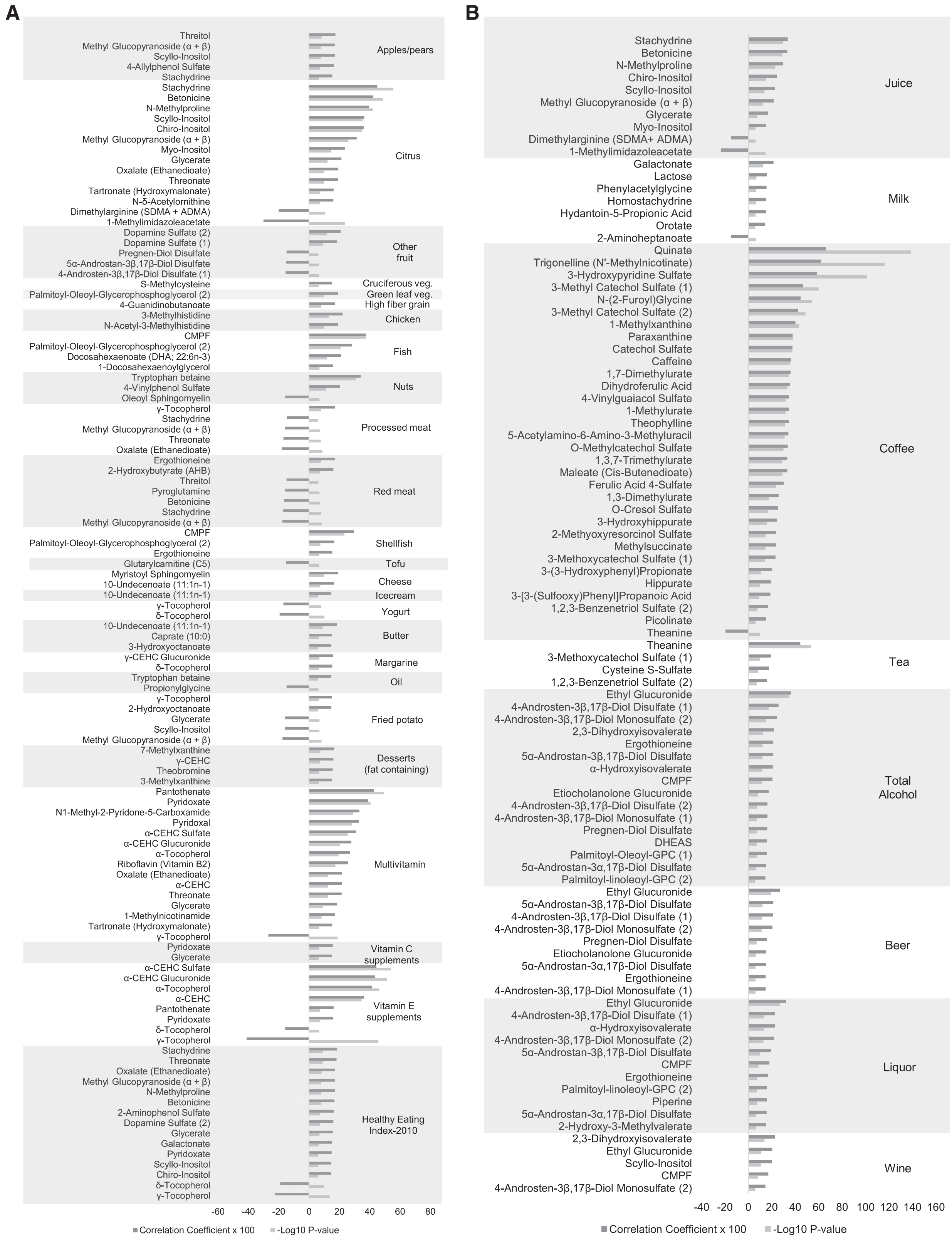

FIGURE 1 Serum nonfasting metabolites associated with food, diet quality, dietary supplements, and beverages (Bonferroni correction) in a nested case-control study within the PLCO cancer screening trial cohort (cases and controls: $n = 1127$). Partial Pearson correlations adjusted for case status (case, control), age at blood draw (years), smoking history (never or missing, former, or current), diabetes history (no or yes), BMI (in kg/m^2; <25, 25 to <30 or missing, or ≥30), education (≤high school, post–high school training other than college, some college, college graduate, or postgraduate), vigorous physical activity (in hours per week; none, <1, 1, 2, 3, ≥4, or missing), and daily caloric intake (quintiles). Analyses were adjusted for multiple comparisons by using Bonferroni correction [$P < 0.05/(617 \times 55) = 1.5 \times 10^{-6}$]. Alcohol use was adjusted for coffee intake and vice versa. The Healthy Eating Index–2010 was adjusted for

(*continued on next page*)

metabolites associated with diet items consumed only episodically or by a small proportion of participants (e.g., alcohol, dietary supplements). Metabolite peak intensity was normalized by using the median peak intensity for each run-day per metabolite. Reliability was assessed by using 86 masked replicates from 7 study participants and 1 quality-control pool spread at approximately even intervals throughout the sample set. Case-control pairs were analyzed in the same batch. Medians (IQRs) for intraclass correlation coefficients (calculated by using 86 replicates) and CVs were 0.94 (0.88–0.98) and 0.21 (0.15–0.30), respectively.

Statistical analysis

Our analysis was conducted in 2 stages. First, we estimated correlations between 617 named metabolites and 55 individual foods, food groups, beverages, vitamin supplements, and the HEI-2010 by using partial Pearson correlation, adjusted for case-control status, age at blood draw, smoking history, diabetes history, BMI, education, weekly hours of vigorous physical activity, and daily caloric intake (categories shown in **Figure 1**). Metabolite and dietary exposure values were log-transformed to account for non-normal distributions. Alcohol was controlled for coffee intake and vice versa due to evidence of confounding, because alcohol was associated with known coffee phytochemicals. Individual dietary supplement intake was adjusted for multivitamin supplement use. The HEI-2010 was adjusted for any dietary supplement use (yes or no). Metabolites associated with dietary exposures at a Bonferroni-corrected α level $<1.47 \times 10^{-6}$ [$P < 0.05/(617 \times 55)$] were selected for further analysis.

In the second stage, we calculated ORs and 95% CIs for the associations between diet-related metabolites and breast cancer risk by using conditional logistic regression. We compared risks at the 90th and the 10th percentiles of the distribution of log-relative metabolite intensity [i.e. OR: $e^{\beta(X90-X10)}$], where β is the coefficient from the conditional logistic regression of case-control status and (X_{10}, X_{90}) the 10th and 90th percentiles of the log-metabolite values. Case-control matching was retained for subtype analyses. Models were adjusted for age at blood draw, race/ethnicity, putative breast cancer risk factors, and daily caloric intake (categorization detailed in Results tables). We present analyses both with and without adjustment for BMI, because energy balance may mediate diet–breast cancer associations (26). For the breast cancer analyses, we controlled for multiple comparisons by using a false discovery rate <0.2 (Benjamini-Hochberg procedure) (27, 28). All of the statistical tests were 2-sided.

In sensitivity analyses, we evaluated multiplicative effect-modification of metabolite–breast cancer associations by follow-up time (less than or equal to or greater than the median of 6.7 y) by using the Wald test for homogeneity with a Bonferroni-corrected $\alpha < 4.42 \times 10^{-4}$. Because tocopherols are transported in lipoproteins and lipoprotein concentrations can confound risk associations with serum tocopherols (29), we adjusted tocopherol metabolites for total circulating cholesterol, and assessed mutual adjustment of tocopherol metabolites. We also adjusted steroid alcohol–related metabolites for non-steroid alcohol–related metabolites and vice versa. The linearity of select metabolite–ER^+ breast cancer associations was assessed by using cubic splines with likelihood ratio tests (30).

We identified diet-related multiple-metabolite profiles by using principal components analysis (PCA) with the principal axis method and varimax (orthogonal) rotation to group the diet-related metabolites identified in the first analysis stage. We then measured principal component associations with breast cancer by using conditional logistic regression. All of the sensitivity analyses and the PCA used fully adjusted models, including BMI.

With the use of Gaussian graphical modeling (GGM), we modeled the bivariate relations between the Bonferroni-significant diet-related metabolites to summarize how they group together as an interconnected metabolic network (19, 31). Direct relations between metabolites were measured by using conditional correlation, where each correlation is conditioned on all other metabolites in the analysis, including any indirectly correlated metabolites. Conditional correlations between metabolites (absolute value of $r > 0.2$) were visualized by linking them together by a line to represent direct relations. Partial Pearson correlation, logistic regression, and PCA analyses were performed in SAS (version 9.3; SAS Institute); GGM was conducted in R (R Project, version 3.1.2) with visualization using Cytoscape (32).

RESULTS

Study population characteristics

Women's mean ± SD age was 64 ± 5.3 y, on average, at the time of blood collection. The median (interdecile range) follow-up time from baseline was 6.7 y (2.1–11.6 y). Compared with controls, cases were more likely to have an older age at natural menopause, a history of benign breast disease, and a family history of breast cancer and were less likely to have never smoked (**Table 1**). Self-reported dietary intakes for the 12 mo before the study baseline for all women are presented in Supplemental Table 1.

Dietary exposure–metabolite associations

Of the 617 identifiable serum metabolites we evaluated, 113 were significantly associated with ≥1 dietary exposures, with Bonferroni correction for multiple comparisons (Figure 1; additional details in **Supplemental Table 2**, including, for each metabolite, the percentage of participants less than the LOD and intraclass correlation coefficient). Of the 113 diet-related metabolites, the median (IQR) for the percentage of participants with metabolite values less than the LOD was 0% (0–9%). In addition, 87% of metabolites had metabolite intensities above the LOD for >80% of participants. These diet-related metabolites represent both exogenous food breakdown products and

Figure 1. (continued) any supplement use. Individual vitamin supplement use was adjusted for multivitamin use. Metabolites associated with oranges and grapefruit, separately, can be found in Supplemental Table 2. Information on mass-to-charge ratio, retention index, number of metabolites below the limit of detection, metabolomics platform, metabolite identification number [Human Metabolome Database (http://www.hmdb.ca/metabolites/) or PubChem (https://pubchem.ncbi.nlm.nih.gov/)], and metabolic pathways is available in Supplemental Table 2. ADMA, asymmetric dimethylarginine; AHB, α-hydroxybutyrate; CEHC, carboxyethyl hydrochroman; CMPF, 3-carboxy-4-methyl-5-propyl-2-furanpropionic acid; DHEAS, dehydroepiandrosterone sulfate; GPC, glycerophosphocholine; PLCO, Prostate, Lung, Colorectal and Ovarian; SDMA, symmetric dimethylarginine; veg., vegetables.

TABLE 1
Baseline characteristics of participants in a nested case-control study within the PLCO cancer screening trial[1]

Characteristic	Cases (n = 621), %	Controls (n = 621), %	P^2
Age at blood draw,[3] y			0.14
<60	23	23	
60 to <65	34	34	
65 to <70	24	23	
≥70	19	20	
Age at menarche, y			0.18
<12	22	18	
12–13	52	57	
≥14	26	25	
Age at first child, number of live births			0.07
Nulliparous	9	9	
<20 y, ≥1	14	14	
20–24 y, 1–2	14	10	
20–24 y, ≥3	32	39	
25–29 y, 1–2	10	8	
25–29 y, ≥3	12	12	
≥30 y, ≥1	10	7	
Type of menopause, age at menopause			0.04
Natural, <45 y	8	6	
Natural, 45–49 y	17	19	
Natural, 50–54 y	35	35	
Natural, ≥55 y	10	7	
Bilateral oopherectomy	8	11	
Drug therapy or radiation	4	2	
Hysterectomy	19	20	
MHT use[4]			0.02
Never	70	64	
Former	18	24	
Current	12	12	
History of benign breast disease			0.001
No	69	77	
Yes	31	23	
First-degree female family history			0.02
No	80	85	
Yes	20	15	
Race/ethnicity			0.29
Non-Hispanic white	90	91	
Other	10	9	
Education			0.56
≤12 y	35	34	
Post–high school training other than college	12	13	
Some college	21	22	
College	18	15	
Postgraduate	14	16	
Smoking history			0.001
Never	54	65	
Former	39	29	
Current	8	6	
Drinking history			0.56
Never	9	9	
Former	13	14	
Current	70	71	
Missing	8	6	
Diabetes history			0.34
No	93	95	
Yes	7	5	
Vigorous physical activity, h/wk			0.10
None	14	12	
<1	18	18	
1	12	9	
2	15	19	
3	16	16	
≥4	17	20	
Missing	8	7	
BMI, kg/m^2			0.08
<25	32	37	
25 to <30	40	39	
≥30	29	24	

[1] Values may not sum to 100% due to rounding. MHT, menopausal hormone therapy; PLCO, Prostate, Lung, Colorectal, and Ovarian.

[2] P for difference by case-control status is based on a univariate conditional logistic regression global likelihood ratio test.

[3] Cases and controls were matched on study baseline age (±2 y) and month per year of blood collection (±3 mo) but not on age at blood draw.

[4] MHT users were excluded at cohort entry, except for estrogen receptor–negative or progesterone receptor–negative cases, or both, and their matched controls.

endogenous metabolic factors correlated with the consumption of specific foods. Of the 55 dietary exposures we defined, on the basis of the 2015 Dietary Guidelines for Americans, 34 were significantly related to ≥1 serum metabolite. A total of 222 dietary exposure–metabolite associations (all $r \geq \pm 0.14$), reached significance. Forty-two of these diet-metabolite correlations were moderately strong (absolute value of $r > 0.3$), including associations with citrus, coffee, juice, alcohol, liquor, nuts, tea, fish, shellfish, multivitamins, and vitamin E supplements. Several of these associations were potentially novel (i.e., no published epidemiologic studies in free-living participants, to our knowledge), including theanine (tea; r = 0.43), 3-hydroxypyridine-sulfate (coffee; r = 0.58), and methyl-glucopyranoside-(α/β) (citrus; r = 0.31). There were 42 metabolites that were correlated with >1 dietary exposure (**Supplemental Table 3**).

Dietary metabolite–breast cancer associations

Twelve diet-related metabolites were significantly associated with breast cancer risk, after adjustment for multiple comparisons (false discovery rate <0.2), including metabolites related to the consumption of butter, margarine, alcohol, and fat-containing desserts (**Table 2**). With further adjustment for BMI, the associations with butter-related caprate (10:0) [$OR_{90th\ versus\ 10th\ percentile}$ (OR_{90vs10}): 1.77; 95% CI: 1.28, 2.43], dessert-related γ-carboxyethyl hydrochroman (γ-CEHC) (OR_{90vs10}: 1.64; 95% CI: 1.18, 2.28), and alcohol-related 4-androsten-3β,17β-diol-monosulfate (1) (OR_{90vs10}: 1.61; 95% CI: 1.20, 2.16) remained significant; the other associations were, in general, modestly attenuated and became nonsignificant (change in ORs: +2% to −6%).

Nineteen metabolites were significantly associated with ER$^+$ breast cancer; all associations remained after adjustment for BMI (**Table 3**). Two were butter-related: caprate (10:0) (OR: 1.81; 95%

TABLE 2

Top 20 diet-related serum metabolites associated with breast cancer in a nested case-control study within the PLCO cancer screening trial[1]

Metabolite	Dietary exposure	Correlation coefficient[2]	90th vs. 10th percentile[3] Multivariate-adjusted OR (95% CI)	P	90th vs. 10th percentile[4] BMI and multivariate-adjusted OR (95% CI)	P
Caprate-(10:0)	Butter	+0.15	1.76 (1.28, 2.42)	5.1×10^{-4}*	1.77 (1.28, 2.43)	4.8×10^{-4}*
γ-CEHC	Desserts	+0.16	1.71 (1.24, 2.38)	1.3×10^{-3}*	1.64 (1.18, 2.28)	3.6×10^{-3}*
4-Androsten-3β,17β-diol monosulfate-(1)	Alcohol (beer)	+0.16	1.67 (1.24, 2.23)	6.8×10^{-4}*	1.61 (1.20, 2.16)	1.8×10^{-3}*
DHEAS	Alcohol	+0.16	1.52 (1.11, 2.07)	8.2×10^{-3}*	1.49 (1.09, 2.04)	1.2×10^{-2}
α-Hydroxyisovalerate	Alcohol (liquor)	+0.23	1.53 (1.13, 2.06)	5.6×10^{-3}*	1.48 (1.09, 2.00)	1.1×10^{-2}
4-Androsten-3β,17β-diol disulfate-(1)	Alcohol (beer, liquor)	+0.26	1.55 (1.14, 2.11)	4.8×10^{-3}*	1.48 (1.09, 2.03)	1.3×10^{-2}
Pregnen-diol disulfate	Alcohol (beer)	+0.16	1.52 (1.11, 2.09)	9.8×10^{-3}*	1.48 (1.08, 2.04)	1.6×10^{-2}
δ-Tocopherol	Margarine	+0.15	1.56 (1.09, 2.22)	1.4×10^{-2}*	1.47 (1.03, 2.11)	3.6×10^{-2}
4-Androsten-3β,17β-diol disulfate-(2)	Alcohol	+0.16	1.50 (1.11, 2.02)	9.0×10^{-3}*	1.44 (1.06, 1.95)	2.0×10^{-2}
γ-CEHC glucuronide	Margarine	+0.16	1.53 (1.08, 2.15)	1.6×10^{-2}*	1.44 (1.01, 2.04)	4.2×10^{-2}
4-Androsten-3β,17β-diol monosulfate-(2)	Alcohol (beer, liquor, wine)	+0.24	1.48 (1.12, 1.95)	6.3×10^{-3}*	1.43 (1.08, 1.90)	1.3×10^{-2}
10-Undecenoate-(11:1n–1)	Butter, cheese, ice cream	+0.18	1.44 (1.05, 1.98)	2.6×10^{-2}	1.43 (1.04, 1.97)	2.9×10^{-2}
5α-Androstan-3α,17β-diol disulfate	Alcohol (beer, liquor)	+0.15	1.42 (1.06, 1.91)	2.1×10^{-2}*	1.39 (1.03, 1.88)	2.9×10^{-2}
Tryptophan betaine	Nuts, oil	+0.34	1.38 (1.03, 1.84)	2.9×10^{-2}	1.38 (1.03, 1.85)	3.0×10^{-2}
2-Hydroxyoctanoate	Fried potatoes	+0.15	1.39 (1.04, 1.86)	2.8×10^{-2}	1.38 (1.03, 1.85)	3.3×10^{-2}
Ethyl glucuronide	Alcohol (beer, liquor, wine)	+0.36	1.31 (0.99, 1.72)	5.6×10^{-2}	1.33 (1.01, 1.75)	4.2×10^{-2}
Piperine	Liquor	+0.16	0.75 (0.56, 1.00)	5.0×10^{-2}	0.74 (0.56, 0.99)	4.5×10^{-2}
Ergothioneine	Alcohol (beer, liquor), shellfish	+0.21	1.28 (0.99, 1.65)	6.2×10^{-2}	1.29 (1.00, 1.67)	5.3×10^{-2}
α-Tocopherol	Multivitamin, vitamin E supplements	+0.41	0.71 (0.51, 0.99)	4.2×10^{-2}	0.73 (0.52, 1.03)	7.0×10^{-2}
Propionylglycine	Oil	−0.15	0.78 (0.62, 0.99)	3.99×10^{-2}	0.81 (0.64, 1.03)	8.1×10^{-2}

[1] n = 621 incident cases and 621 matched controls. The top 20 metabolites are presented to highlight all significant associations with a consistent number of results for the overall breast cancer and breast cancer subtype analyses. For dietary exposure, "Alcohol" refers to all alcohol types combined. Specific alcohol type associations are shown in parentheses. *Significant after controlling for multiple comparisons by using a false discovery rate <0.2. DHEAS, dehydroisoandrosterone sulfate; PLCO, Prostate, Lung, Colorectal, and Ovarian; γ-CEHC, γ-carboxyethyl hydrochroman.

[2] Direction and magnitude of correlation between dietary exposure and metabolite; the highest correlation is presented for metabolites correlated with multiple foods.

[3] Risks at the 90th compared with the 10th percentile of the distribution of log-relative metabolite intensity. Conditional logistic regression, controlling for age at blood draw (years), race/ethnicity (non-Hispanic white or other), age at menarche (<12 y, 12–13 y or missing, or ≥14 y), age at first child and number of live births (nulliparous, <20 y and ≥1 live birth, 20–29 y and 1–2 live births, 20–29 y and ≥3 live births or missing, or ≥30 y and ≥1 live births), type of menopause and age at menopause (natural and <45 y, natural and 45–49 y, natural and 50–54 y, natural and ≥55 y, bilateral oophorectomy or drug therapy or radiation, or hysterectomy or missing), menopausal hormone therapy use (never, former, or current), history of benign breast disease (no or missing, or yes), first-degree female family history of breast cancer (no or missing, or yes), education (≤12 y, post–high school training besides college, some college or missing, college, or postgraduate), smoking history (never or missing, former, or current), drinking history (never or missing, former, or current), diabetes history (no or yes), vigorous physical activity (in hours per week; none, <1, 1, 2, 3, ≥4, or missing), and daily caloric intake (quintiles). Alcohol-related metabolites were not adjusted for self-reported alcohol use.

[4] Risks at the 90th compared with the 10th percentile of the distribution of log-relative metabolite intensity. Conditional logistic regression with control for the same covariates as in footnote 3, with additional control for BMI (in kg/m^2; <25, 25 to <30, ≥30, or missing).

CI: 1.23, 2.67) and 10-undecenoate (11:1n–1) (OR_{90vs10}: 1.59; 95% CI: 1.07, 2.36); 1 was fried food–related: 2-hydroxyoctanoate (OR_{90vs10}: 1.46; 95% CI: 1.03, 2.07). Twelve were alcohol-related, including the following: α-hydroxyisovalerate (OR_{90vs10}: 2.23; 95% CI: 1.50, 3.32), 4-androsten-3β,17β-diol-disulfate (2) (OR_{90vs10}: 2.08; 95% CI: 1.41, 3.07), and 6 additional androgens. Three were vitamin E (tocopherol)–related, with positive or inverse associations with ER^+ breast cancer depending on the tocopherol isomer: γ-CEHC (positively associated with fat-containing desserts; OR_{90vs10}: 1.80; 95% CI: 1.20, 2.70), δ-tocopherol (positively associated with margarine; OR_{90vs10}: 1.74; 95% CI: 1.12, 2.71), and α-tocopherol (positively associated with multivitamins and vitamin E supplements; OR_{90vs10}: 0.56; 95% CI: 0.37, 0.85). Metabolites associated with margarine (γ-CEHC glucuronide; OR_{90vs10}: 4.2; 95% CI: 1.43, 12.40), desserts (γ-CEHC; OR: 4.25; 95% CI: 1.36, 13.28), and butter [caprate (10:0); OR: 3.00; 95% CI: 1.21, 7.46] were positively, although not significantly, associated with ER^- breast cancer (**Supplemental Table 4**). None of the metabolite–breast cancer associations described above was modified by follow-up time (≤6.7 compared with >6.7 y).

Adjustment for serum cholesterol did not materially change associations between tocopherol metabolites and overall and ER^+ breast cancer (change in ORs: <5%), which remained significant. We further adjusted steroid alcohol–related metabolites for non-steroid alcohol–related metabolites and vice versa. Steroid metabolite–ER^+ breast cancer associations were attenuated (change in ORs: −7% to −20%), but all associations remained significant.

The non-steroid alcohol–related metabolite associations were not modified. The dose-response relations of circulating caprate-(10:0), γ-CEHC, α-hydroxyisovalerate, 4-androsten-3β,17β-diol-disulfate (2), and α-tocopherol with ER^+ breast cancer did not depart significantly from log linearity (**Figure 2**).

For the PCA of the 113 diet-related metabolites, 3 principal components each had eigenvalues >1, which explained ≥5% of total variance and met criteria for component loading and simple structure. The PCA created 3 sets of linearly uncorrelated variables, as follows: "coffee" (31 metabolites), "HEI-2010/multivitamin" (23 metabolites), and "alcohol" (9 metabolites). Of these, the "alcohol" component was significantly associated with overall breast cancer (OR_{90vs10}: 1.53; 95% CI: 1.12, 2.10) and ER^+ breast cancer (OR_{90vs10}: 2.21; 95% CI: 1.48, 3.32) (**Supplemental Tables 5** and **6**).

The association networks for pairs of diet-related metabolites with conditional correlations ≥0.2 (GGM) consisted of 92 direct (paired) relations, represented by connected lines. **Figure 3** presents results for diet exposures for which ≥1 correlated metabolite was associated with ER^+ breast cancer; the network for all 113 diet-related metabolites is shown in **Supplemental Figure 3**. Metabolites associated with ER^+ breast cancer fell into 3 distinct metabolic networks: metabolites related to alcohol, vitamin E, and fats and oils.

Finally, we qualitatively compared the associations between dietary exposures and breast cancer measured by potential biomarkers compared with self-report. Self-reported butter, fried potatoes, margarine, desserts, and alcohol were not associated with breast cancer risk. Vitamin E supplement intake was inversely associated with ER^+ breast cancer (OR_{90vs10}: 0.59; 95% CI: 0.40, 0.89) (**Supplemental Table 7**). We present a metabolite by metabolite correlation matrix and a food by metabolite correlation matrix in **Supplemental Tables 8** and **9**.

DISCUSSION

Our study provides a conceptual framework for utilizing metabolomics to study dietary associations with cancer. We identified 113 metabolites that were associated with ≥1 food, beverage, or dietary supplement, thus providing possible biomarkers for these dietary exposures should they be replicated in future studies, conform to a classical measurement model, or be further identified as a metabolite constituent of the associated food (13, 33). In our breast cancer analysis—which was, to our knowledge, the first to use metabolomics to agnostically evaluate prediagnostic circulating diet-related metabolites in relation to breast cancer risk—we found that metabolites related to alcohol, tocopherols (vitamin E), and animal fats were moderately associated with overall breast cancer risk, primarily driven by results for ER^+ cases. These associations were stronger than those observed for the same exposure measures based on self-reported diet, both within our study population and comparing against historical effect estimates, such as for alcohol (10).

We were able to replicate previously published diet-metabolite associations on the basis of both the Metabolon, Inc., and other platforms, which is a promising result that suggests that nutritional metabolomics can be robust and informative. In the 1242 female participants in our study, we replicated approximately three-quarters of the 40 Bonferroni-significant candidate dietary biomarkers previously uncovered in the PLCO cohort by using the Metabolon platform (14). In addition, for 15 dietary exposures, we replicated diet-metabolite associations reported by nutritional metabolomics studies with the use of other study populations, dietary assessment tools, or metabolomics platforms (14, 15, 34–46) (**Supplemental Table 10**). Because our current study used a newer, more sensitive platform than the earlier PLCO study (14), we measured 84 more identifiable metabolites that could serve as candidate dietary biomarkers. For example, theanine, an amino acid extracted from tea leaves and studied in vitro (47, 48), was correlated with tea intake; methyl-glucopyranoside-(α/β) was associated with fruit and citrus; and 3-hydroxypyridine-sulfate was associated with coffee. These are phytochemical constituents of the associated foods, which explains why they may function as dietary biomarkers.

Metabolites that were associated with multiple dietary exposures included those associated with different varieties of specific food groups (e.g., alcohol, fruit, tea and coffee, dairy products, and dietary supplements). 3-Carboxy-4-methyl-5-propyl-2-furanpropionic acid and ergothioneine were each correlated with both fish and shellfish and alcohol intake, which could reflect correlated intakes of fish and alcohol in the study population (alcohol: $r = 0.16$ with fish, $r = 0.26$ with shellfish; $P < 0.001$). In addition, in univariate analyses, piperine, a metabolite derived from pepper, was associated with liquor intake. However, in fully conditional models of metabolite associations with other metabolites, piperine was associated with γ-tocopherol, a metabolite we found to be linked to intakes of specific fats. This may reflect a complex pattern of confounding, such as if piperine were an ingredient in a food often consumed concurrently with liquor.

In the breast cancer analysis, metabolites positively associated with ER^+ breast cancer included the following: butter-related caprate (10:0), a medium-chain SFA; dairy fat–related 10-undecenoate (11:1n–1), an odd-carbon MUFA; fried food–related 2-hydroxyoctanoate, a hydroxy fatty acid; and dessert-related γ-CEHC. The association of dietary fat with postmenopausal breast cancer has been controversial due to conflicting findings in the literature. Such inconsistencies may reflect varying degrees of measurement error in assessing fat intake (49–53), etiologic associations that vary according to type of dietary fat consumed, or both, with potentially stronger positive associations for SFAs than for other fats (54). Nutritional metabolomics may partly circumvent these problems, by providing objectively quantifiable biomarkers that reflect a full range of the types of dietary fat typically consumed. Suggested mechanisms for the effect of dietary fat on breast carcinogenesis include increased estrogen synthesis in adipose tissue, inflammation, and alterations in several physiologic processes (e.g., breast cell response to growth factors, immune function, and tumor suppression) (54). Few studies, to our knowledge, have reported on the fatty acids implicated in our analysis, although a urine analysis found that caprate (10:0) concentrations were higher in mice with breast tumors than in controls (55).

We found 7 endogenous androgen pathway metabolites related to alcohol consumption and ER^+ breast cancer; estrogen pathway metabolites could not be evaluated because none were measured by the metabolomics platform. Alcohol intake is the only well-established dietary risk factor for breast cancer (1), with an attributable risk of ~5% (56). Several mechanisms have been proposed. A randomized alcohol feeding study in postmenopausal women reported that alcohol increased serum dehydroepiandrosterone

TABLE 3

Top 20 diet-related serum metabolites associated with ER^+ breast cancer in a nested case-control study within the PLCO cancer screening trial[1]

			90th vs. 10th percentile[3]		90th vs. 10th percentile[4]	
Metabolite	Food	Correlation coefficient[2]	Multivariate-adjusted OR (95% CI)	P	BMI and multivariate-adjusted OR (95% CI)	P
α-Hydroxyisovalerate	Alcohol (liquor)	+0.23	2.22 (1.51, 3.29)	6.0×10^{-5}*	2.23 (1.50, 3.32)	7.4×10^{-5}*
4-Androsten-3β,17β-diol disulfate-(2)	Alcohol	+0.16	2.13 (1.45, 3.13)	1.2×10^{-4}*	2.08 (1.41, 3.07)	2.4×10^{-4}*
4-Androsten-3β,17β-diol monosulfate-(1)	Alcohol (beer)	+0.16	2.11 (1.43, 3.11)	1.7×10^{-4}*	2.06 (1.39, 3.05)	3.0×10^{-4}*
DHEAS	Alcohol	+0.16	1.96 (1.32, 2.91)	9.0×10^{-4}*	1.93 (1.30, 2.88)	1.2×10^{-3}*
5α-Androstan-3α,17β-diol disulfate	Alcohol (beer, liquor)	+0.15	1.92 (1.31, 2.80)	8.1×10^{-4}*	1.89 (1.29, 2.80)	1.1×10^{-3}*
Pregnen-diol disulfate	Alcohol (beer)	+0.16	1.92 (1.29, 2.84)	1.2×10^{-3}*	1.88 (1.27, 2.79)	1.8×10^{-3}*
4-Androsten-3β,17β-diol monosulfate-(2)	Alcohol (beer, liquor, wine)	+0.24	1.92 (1.34, 2.76)	4.2×10^{-4}*	1.87 (1.30, 2.70)	7.6×10^{-4}*
4-Androsten-3β,17β-diol disulfate-(1)	Alcohol (beer, liquor)	+0.26	1.92 (1.30, 2.84)	1.1×10^{-3}*	1.87 (1.25, 2.77)	2.1×10^{-3}*
Caprate-(10:0)	Butter	+0.15	1.82 (1.24, 2.67)	2.3×10^{-3}*	1.81 (1.23, 2.67)	2.7×10^{-3}*
γ-CEHC	Desserts	+0.16	1.85 (1.24, 2.77)	2.7×10^{-3}*	1.80 (1.20, 2.70)	4.7×10^{-3}*
Ethyl glucuronide	Alcohol (beer, liquor, wine)	+0.36	1.73 (1.23, 2.41)	1.5×10^{-3}*	1.76 (1.26, 2.47)	1.1×10^{-3}*
α-Tocopherol	Multivitamin, vitamin E supplements	+0.41	0.56 (0.37, 0.84)	5.3×10^{-3}*	0.56 (0.37, 0.85)	6.8×10^{-3}*
Etiocholanolone glucuronide	Alcohol (beer)	+0.18	1.76 (1.16, 2.67)	7.7×10^{-3}*	1.74 (1.15, 2.65)	9.2×10^{-3}*
δ-Tocopherol	Margarine	+0.15	1.80 (1.17, 2.79)	8.0×10^{-3}*	1.74 (1.12, 2.71)	1.5×10^{-2}*
2-Hydroxy-3-methylvalerate	Liquor	+0.15	1.74 (1.15, 2.65)	9.0×10^{-3}*	1.72 (1.12, 2.62)	1.2×10^{-2}*
Pyroglutamine	Red meat	−0.16	1.60 (1.08, 2.38)	2.1×10^{-2}*	1.59 (1.07, 2.37)	2.3×10^{-2}*
10-Undecenoate-(11:1n–1)	Butter, cheese, ice cream	+0.18	1.59 (1.07, 2.36)	2.2×10^{-2}*	1.59 (1.07, 2.36)	2.3×10^{-2}*
5α-Androstan-3β,17β-diol disulfate	Alcohol (beer, liquor)	+0.15	1.53 (1.08, 2.17)	1.6×10^{-2}*	1.50 (1.06, 2.13)	2.3×10^{-2}*
2-Hydroxyoctanoate	Fried potatoes	+0.15	1.47 (1.04, 2.08)	3.1×10^{-2}*	1.46 (1.03, 2.07)	3.2×10^{-2}*
Ergothioneine	Alcohol (beer, liquor), shellfish	+0.21	1.41 (1.01, 1.95)	4.3×10^{-2}	1.42 (1.02, 1.97)	3.8×10^{-2}

[1] n = 418 incident cases and 437 matched controls. The top 20 metabolites are presented to highlight all significant associations with a consistent number of results for the overall breast cancer and breast cancer subtype analyses. For dietary exposure, "Alcohol" refers to all alcohol types combined. Specific alcohol type associations are shown in parentheses. *Significant after controlling for multiple comparisons by using a false discovery rate <0.2. DHEAS, dehydroisoandrosterone sulfate; ER^+, estrogen receptor–positive; PLCO, Prostate, Lung, Colorectal, and Ovarian; γ-CEHC, γ-carboxyethyl hydrochroman.

[2] Direction and magnitude of correlation between dietary exposure and metabolite; the highest correlation is presented for metabolites correlated with multiple foods.

[3] Risks at the 90th compared with the 10th percentile of the distribution of log-relative metabolite intensity. Conditional logistic regression, controlling for age at blood draw (years), race/ethnicity (non-Hispanic white or other), age at menarche (<12 y, 12–13 y or missing, or ≥14 y), age at first child and number of live births (nulliparous, <20 y and ≥1 live birth, 20–29 y and 1–2 live births, 20–29 y and ≥3 live births or missing, or ≥30 y and ≥1 live birth), type of menopause and age at menopause (natural and <45 y, natural and 45–49 y, natural and 50–54 y, natural and ≥55 y, bilateral oophorectomy or drug therapy or radiation, or hysterectomy or missing), menopausal hormone therapy use (never, former or current), history of benign breast disease (no or missing or yes), first-degree female family history of breast cancer (no or missing, or yes), education (≤12 y, post–high school training besides college, some college or missing, college, or postgraduate), smoking history (never or missing, former, or current), drinking history (never or missing, former, or current), diabetes history (no or yes), vigorous physical activity (in hours per week; none, <1, 1, 2, 3, ≥4, or missing), and daily caloric intake (quintiles). Alcohol-related metabolites were not adjusted for self-reported alcohol use.

[4] Risks at the 90th compared with the 10th percentile of the distribution of log-relative metabolite intensity. Conditional logistic regression with control for the same covariates as in footnote 3, with additional control for BMI (in kg/m^2; <25, 25 to <30, ≥30, or missing).

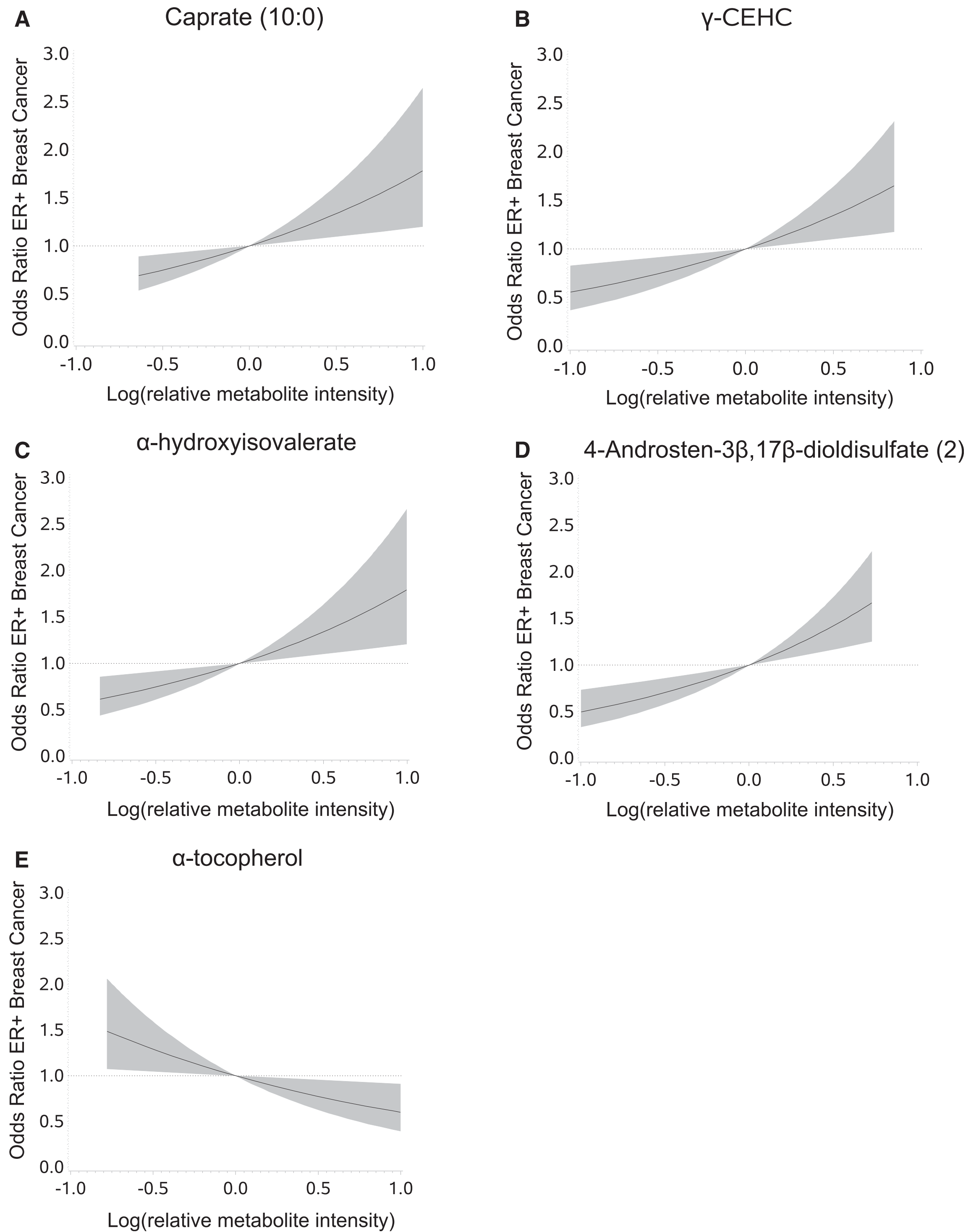

FIGURE 2 ORs for the association between diet-related metabolites and ER^+ breast cancer in the PLCO cancer screening trial cohort. (A) Caprate (10:0) (butter-related; *P*-linearity = 0.003, *P*-nonlinearity = 0.99); (B) γ-CEHC (dessert-related; *P*-linearity = 0.003, *P*-nonlinearity = 0.66); (C) α-hydroxyisovalerate (alcohol-related; *P*-linearity = 0.003, *P*-nonlinearity = 0.57); (D) 4-androsten-3β,17β-diol disulfate (2) (alcohol-related; *P*-linearity = 0.0003, *P*-nonlinearity = 0.75); (E) α-tocopherol (multivitamin- and vitamin E supplement–related; *P*-linearity = 0.02, *P*-nonlinearity = 0.56). Estimates were obtained from restricted cubic spline logistic regression models. All models were significant for a linear relation; thus, the linear relations are presented. Models were adjusted for age at blood draw (years), age at menarche (<12 y, 12–13 y or missing, or ≥14 y), age at first child and number of live births (nulliparous, <20 y and ≥1 live birth, 20–29 y and 1–2 live births, 20–29 y and ≥3 live births or missing, or ≥30 y and ≥1 live birth), type of menopause and age at menopause (natural and <45 y, natural and 45–49 y, natural and 50–54 y, natural and ≥55 y, bilateral oophorectomy or drug therapy or radiation, or hysterectomy or missing), menopausal hormone therapy use (never, former, or current), history of benign breast disease (no or missing, or yes), first-degree female family history of breast cancer (no or missing, or yes), education (≤12 y, post–high school training besides college, some college or missing, college, or postgraduate), smoking history (never or missing, former, or current), drinking history (never or missing, former, or current), diabetes history (no or yes), vigorous physical activity (in hours per week; none, <1, 1, 2, 3, ≥4, or missing), and daily caloric intake (quintiles). Alcohol-related metabolites were not adjusted for self-reported alcohol use. ER^+, estrogen receptor–positive; PLCO, Prostate, Lung, Colorectal, and Ovarian; γ-CEHC, γ-carboxyethyl hydrochroman.

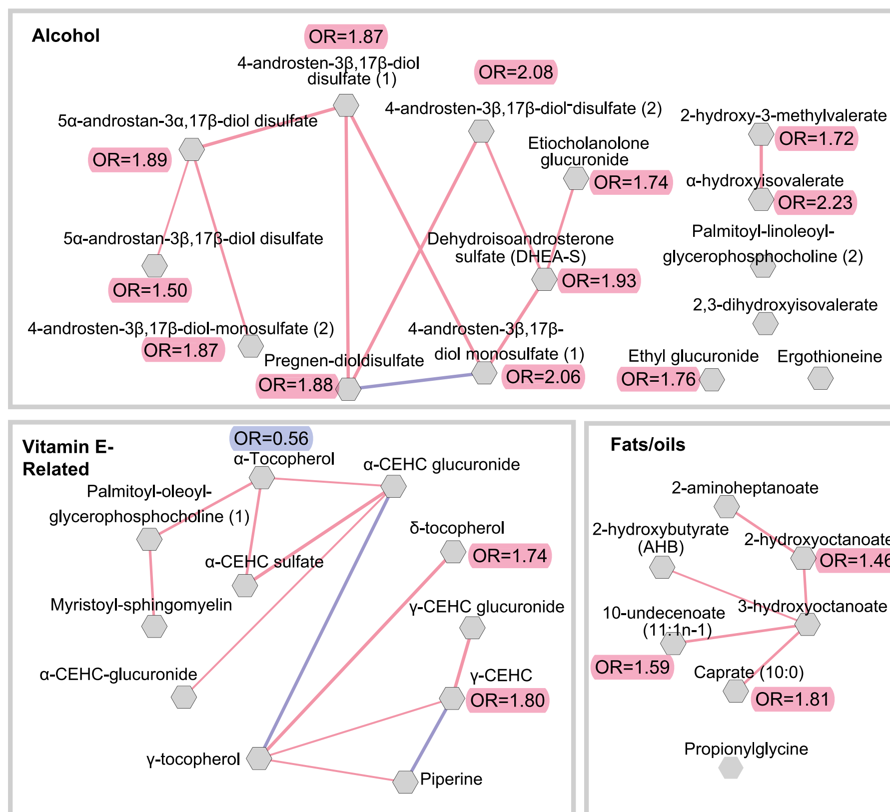

FIGURE 3 Gaussian graphical model of 113 diet-related metabolites measured in a nested case-control study within the PLCO cancer screening trial cohort. Metabolites are pictured as hexagons, and those pairs with an absolute value of conditional correlation >0.2 are connected by a line. Pink lines represent positive conditional correlations. Purple lines represent inverse conditional correlations. Hexagons are color-coded by their association with ER$^+$ breast cancer. Metabolites highlighted pink are positively associated with ER$^+$ breast cancer. Metabolites highlighted purple are inversely associated with ER$^+$ breast cancer. CEHC, carboxyethyl hydrochroman; ER$^+$, estrogen receptor–positive; PLCO, Prostate, Lung, Colorectal, and Ovarian.

sulfate (57), a precursor of circulating androgens and estrogens (58, 59). Moreover, pooled analyses of prospective epidemiologic studies found that dehydroepiandrosterone sulfate, other androgens, and estrogens were strongly positively associated with both alcohol consumption (60) and postmenopausal breast cancer (61, 62), with effect sizes similar to those in our study (i.e., ORs: ∼1.5–2 comparing extreme quintiles). Thus, alcohol may influence postmenopausal breast cancer risk through increased production of androgens, estrogens, and their metabolites. Alcohol-related metabolites unrelated to androgen metabolism were also associated with ER$^+$ breast cancer, including ethyl-glucuronide (a metabolite of ethanol breakdown) and α-hydroxyisovalerate (a hydroxy-fatty acid derivative) (63). Their associations were unchanged with adjustment for androgen-pathway metabolites, suggesting that alcohol may operate through androgen-independent mechanisms to increase breast cancer risk.

We also found breast cancer associations with vitamin E tocopherols and their metabolites. Specifically, α-tocopherol was inversely associated, whereas δ-tocopherol and γ-CEHC (conjugated γ-tocopherol) were positively associated with ER$^+$ breast cancer. α-Tocopherol, the most abundant circulating tocopherol isoform, which protects polyunsaturated fats from lipid peroxidation, is found in vitamin E supplements, olive and sunflower oils, wheat germ, and nuts; γ-tocopherol is found in vegetable oil and margarine (64); and δ-tocopherol is found in soy and castor oils and margarine (65). Supplementation with α-tocopherol has been shown to increase circulating α-tocopherol and suppress γ-tocopherol (64) and vice versa (66). Experimental research indicates that α- and γ-tocopherols may reduce cell proliferation and increase apoptosis, 2 hallmarks of cancer (67–69). However, evidence from large prospective studies has been conflicting, including no breast cancer association with circulating α- or γ-tocopherol (70–76), a positive association with γ-tocopherol (77), and an inverse association with α-tocopherol (78). The inverse association of α-tocopherol with breast cancer may reflect a healthy lifestyle, as evidenced by its correlation with dietary supplement use and the HEI-2010. The δ- and γ-tocopherol findings could indicate a direct association of vitamin E with breast cancer or could reflect the influence of specific sources of dietary fat.

To our knowledge, ours is the first study to use metabolomics to agnostically evaluate prediagnostic circulating diet-related metabolites in relation to breast cancer risk. Other studies have been retrospective, and these metabolite profiles may reflect physiologic changes driven by the occurrence and treatment of breast cancer (79). Strengths of our study include prediagnostically collected serum, comparable handling of samples between cases and controls, large study size, large number of identified metabolites, and metabolomics platform reliability. However, there are several limitations. Serum samples were obtained 1 y after the food-frequency questionnaire; major diet changes over this period could attenuate observed diet-metabolite correlations. Participants were not required to fast. We did not have data on hours since last meal and were unable to adjust for fasting status.

Previous analyses have identified the minimal contribution of fasting status to overall metabolite variability (80). Our analysis was restricted to identified metabolites measured by the metabolomics platform. We did not evaluate parent estrogens and metabolites, the inclusion of which would complement our findings. The study population was predominantly white postmenopausal women; thus, generalizability to premenopausal women and diverse race/ethnicity is unknown. The nutritional biomarkers identified in this study were not selected a priori. This approach could risk a degree of overfitting; however, these metabolites largely reproduce those previously observed and, moreover, are directly relevant for foods consumed within the PLCO cohort. Furthermore, we were underpowered to detect ER^- breast cancer associations, which possibly explained the null findings. Although we adjusted analyses for potential confounders, there is the possibility of residual or unknown confounding. Finally, we cannot rule out that some results were due to chance, although we minimized the potential for false positives through stringent control for multiple comparisons. Replication with other metabolomics platforms and in other cohorts will be necessary.

In summary, our findings generate hypotheses about the role of diet in postmenopausal breast cancer etiology, particularly for ER^+ breast cancer. To date, findings from previous studies of self-reported diet have convincingly implicated alcohol consumption as a breast cancer risk factor. Our results, based on candidate dietary biomarkers, suggest that androgen-dependent as well as androgen-independent mechanisms may mediate alcohol-induced breast carcinogenesis. Tocopherols and animal fat may also play a role in breast cancer etiology. Applying a nutritional metabolomics approach to understanding breast cancer etiology could generate new hypotheses about the role of diet and strengthen dietary guidance for cancer prevention.

The authors' responsibilities were as follows—MCP and SCM: designed the research, wrote the manuscript, and had primary responsibility for the final content; RNH, JNS, and RGZ: provided the essential materials; MCP, JNS, and SCM: analyzed the data; and all authors: provided critical intellectual content to review and edit the manuscript, and read and approved the final manuscript. None of the authors reported a conflict of interest.

REFERENCES

1. World Cancer Research Fund International. Continuous Update Project report. Food, nutrition, physical activity, and the prevention of breast cancer [Internet]. 2010. [cited 2015 Jun 1]. Available from: http://www.dietandcancerreport.org/cancer_resource_center/downloads/cu/Breast-Cancer-2010-Report.pdf.
2. Willett W. The search for the causes of breast and colon cancer. Nature 1989;338:389–94.
3. Schatzkin A, Greenwald P, Byar DP, Clifford CK. The dietary fat–breast cancer hypothesis is alive. JAMA 1989;261:3284–7.
4. Toledo E, Salas-Salvado J, Donat-Vargas C, Buil-Cosiales P, Estruch R, Ros E, Corella D, Fito M, Hu FB, Aros F, et al. Mediterranean diet and invasive breast cancer risk among women at high cardiovascular risk in the PREDIMED trial: a randomized clinical trial. JAMA Intern Med 2015;175:1752–60.
5. Prentice RL, Caan B, Chlebowski RT, Patterson R, Kuller LH, Ockene JK, Margolis KL, Limacher MC, Manson JE, Parker LM, et al. Low-fat dietary pattern and risk of invasive breast cancer: the Women's Health Initiative randomized controlled dietary modification trial. JAMA 2006;295:629–42.
6. Chlebowski RT, Blackburn GL, Thomson CA, Nixon DW, Shapiro A, Hoy MK, Goodman MT, Giuliano AE, Karanja N, McAndrew P, et al. Dietary fat reduction and breast cancer outcome: interim efficacy results from the Women's Intervention Nutrition Study. J Natl Cancer Inst 2006;98:1767–76.
7. Pierce JP, Faerber S, Wright FA, Rock CL, Newman V, Flatt SW, Kealey S, Jones VE, Caan BJ, Gold EB, et al. A randomized trial of the effect of a plant-based dietary pattern on additional breast cancer events and survival: the Women's Healthy Eating and Living (WHEL) Study. Control Clin Trials 2002;23:728–56.
8. Pierce JP, Natarajan L, Caan BJ, Parker BA, Greenberg ER, Flatt SW, Rock CL, Kealey S, Al-Delaimy WK, Bardwell WA, et al. Influence of a diet very high in vegetables, fruit, and fiber and low in fat on prognosis following treatment for breast cancer: the Women's Healthy Eating and Living (WHEL) randomized trial. JAMA 2007;298:289–98.
9. Martin LJ, Li Q, Melnichouk O, Greenberg C, Minkin S, Hislop G, Boyd NF. A randomized trial of dietary intervention for breast cancer prevention. Cancer Res 2011;71:123–33.
10. Key J, Hodgson S, Omar RZ, Jensen TK, Thompson SG, Boobis AR, Davies DS, Elliott P. Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues. Cancer Causes Control 2006;17:759–70.
11. Bictash M, Ebbels TM, Chan Q, Loo RL, Yap IK, Brown IJ, de Iorio M, Daviglus ML, Holmes E, Stamler J, et al. Opening up the "black box": metabolic phenotyping and metabolome-wide association studies in epidemiology. J Clin Epidemiol 2010;63:970–9.
12. O'Sullivan A, Gibney MJ, Brennan L. Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. Am J Clin Nutr 2011;93:314–21.
13. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hooft JJ, Wishart DS. The food metabolome: a window over dietary exposure. Am J Clin Nutr 2014;99:1286–308.
14. Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr 2014;100:208–17.
15. Playdon MC, Sampson JN, Cross AJ, Sinha R, Guertin KA, Moy KA, Rothman N, Irwin ML, Mayne ST, Stolzenberg-Solomon R, et al. Comparing metabolite profiles of habitual diet in serum and urine. Am J Clin Nutr 2016;104:776–89.
16. Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, Fogel R, Gelmann EP, Gilbert F, Hasson MA, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Control Clin Trials 2000;21:273S–309S.
17. Subar AF, Ziegler RG, Thompson FE, Johnson CC, Weissfeld JL, Reding D, Kavounis KH, Hayes RB. Prostate, Lung, Colorectal Ovarian cancer screening trial I. Is shorter always better? Relative importance of questionnaire length and cognitive ease on response rates and data quality for two dietary questionnaires. Am J Epidemiol 2001;153:404–9.
18. Hayes RB, Sigurdson A, Moore L, Peters U, Huang WY, Pinsky P, Reding D, Gelmann EP, Rothman N, Pfeiffer RM, et al. Methods for etiologic and early marker investigations in the PLCO trial. Mutat Res 2005;592:147–54.
19. Bartel J, Krumsiek J, Theis FJ. Statistical methods for the analysis of high-throughput metabolomics data. Comput Struct Biotechnol J 2013;4:e201301009.
20. Subar AF, Thompson FE, Smith AF, Jobe JB, Ziegler RG, Potischman N, Schatzkin A, Hartman A, Swanson C, Kruse L, et al. Improving food frequency questionnaires—a qualitative approach using cognitive interviewing. J Am Diet Assoc 1995;95:781–8.
21. Thompson FE, Subar AF, Brown CC, Smith AF, Sharbaugh CO, Jobe JB, Mittl B, Gibson JT, Ziegler RG. Cognitive research enhances accuracy of food frequency questionnaire reports: results of an experimental validation study. J Am Diet Assoc 2002;102:212–25.
22. USDA. MyPlate [Internet] [cited 2014 Oct 23]. Available from: http://www.choosemyplate.gov.
23. Guenther PM, Casavale KO, Reedy J, Kirkpatrick SI, Hiza HA, Kuczynski KJ, Kahle LL, Krebs-Smith SM. Update of the Healthy Eating Index: HEI-2010. J Acad Nutr Diet 2013;113:569–80.
24. Carrick DM, Black A, Gohagan JK, Khan A, Pettit K, Williams C, Yu K, Yurgalevitch S, Huang WY, Zhu C. The PLCO biorepository: creating, maintaining, and administering a unique biospecimen resource. Rev Recent Clin Trials 2015;10:212–22.
25. Evans AM, Bridgewater BR, Liu Q, Mitchell MW, Robinson RJ, Dai H, Stewart SJ, DeHaven CD, Miller LAD. High resolution mass spectrometry improves data quantity and quality as compared to unit mass resolution mass spectrometry in high-throughput profiling metabolomics. Metabolomics 2014;4:132.

26. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL, Gaudet M, Schmidt MK, Broeks A, Cox A, et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. J Natl Cancer Inst 2011;103:250–63.
27. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc B 1995;57:289–300.
28. Ganna A, Lee D, Ingelsson E, Pawitan Y. Rediscovery rate estimation for assessing the validation of significant findings in high-throughput studies. Brief Bioinform 2015;16:563–75.
29. Mondul AM, Moore SC, Weinstein SJ, Karoly ED, Sampson JN, Albanes D. Metabolomic analysis of prostate cancer risk in a prospective cohort: the Alpha-Tocolpherol, Beta-Carotene Cancer Prevention (ATBC) Study. Int J Cancer 2015;137:2124–32.
30. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med 1989;8:551–61.
31. Krumsiek J, Suhre K, Illig T, Adamski J, Theis FJ. Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data. BMC Syst Biol 2011;5:21.
32. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
33. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. J Natl Cancer Inst 2011;103:1086–92.
34. Pallister T, Jennings A, Mohney RP, Yarand D, Mangino M, Cassidy A, MacGregor A, Spector TD, Menni C. Characterizing blood metabolomics profiles associated with self-reported food intakes in female twins. PLoS One 2016;11:e0158568.
35. Pujos-Guillot E, Hubert J, Martin JF, Lyan B, Quintana M, Claude S, Chabanas B, Rothwell JA, Bennetau-Pelissero C, Scalbert A, et al. Mass spectrometry-based metabolomics for the discovery of biomarkers of fruit and vegetable intake: citrus fruit as a case study. J Proteome Res 2013;12:1645–59.
36. Lloyd AJ, Beckmann M, Fave G, Mathers JC, Draper J. Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. Br J Nutr 2011;106: 812–24.
37. Andersen MB, Kristensen M, Manach C, Pujos-Guillot E, Poulsen SK, Larsen TM, Astrup A, Dragsted L. Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics. Anal Bioanal Chem 2014;406:1829–44.
38. Andersen MB, Rinnan A, Manach C, Poulsen SK, Pujos-Guillot E, Larsen TM, Astrup A, Dragsted LO. Untargeted metabolomics as a screening tool for estimating compliance to a dietary pattern. J Proteome Res 2014;13:1405–18.
39. Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas ME, Kochhar S, Stamler J, Holmes E, Elliott P, Nicholson JK. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. Am J Clin Nutr 2010;92:436–43.
40. Zheng Y, Yu B, Alexander D, Steffen LM, Boerwinkle E. Human metabolome associates with dietary intake habits among African Americans in the atherosclerosis risk in communities study. Am J Epidemiol 2014;179:1424–33.
41. Edmands WM, Beckonert OP, Stella C, Campbell A, Lake BG, Lindon JC, Holmes E, Gooderham NJ. Identification of human urinary biomarkers of cruciferous vegetable consumption by metabonomic profiling. J Proteome Res 2011;10:4513–21.
42. Hanhineva K, Lankinen MA, Pedret A, Schwab U, Kolehmainen M, Paananen J, de Mello V, Sola R, Lehtonen M, Poutanen K, et al. Nontargeted metabolite profiling discriminates diet-specific biomarkers for consumption of whole grains, fatty fish, and bilberries in a randomized controlled trial. J Nutr 2015;145:7–17.
43. Guertin KA, Loftfield E, Boca SM, Sampson JN, Moore SC, Xiao Q, Huang WY, Xiong X, Freedman ND, Cross AJ, et al. Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer. Am J Clin Nutr 2015;101:1000–11.
44. Rothwell JA, Fillatre Y, Martin JF, Lyan B, Pujos-Guillot E, Fezeu L, Hercberg S, Comte B, Galan P, Touvier M, et al. New biomarkers of coffee consumption identified by the non-targeted metabolomic profiling of cohort study subjects. PLoS One 2014;9:e93474.
45. Edmands WM, Ferrari P, Rothwell JA, Rinaldi S, Slimani N, Barupal DK, Biessy C, Jenab M, Clavel-Chapelon F, Fagherazzi G, et al. Polyphenol metabolome in human urine and its association with intake of polyphenol-rich foods across European countries. Am J Clin Nutr 2015;102:905–13.
46. Lang R, Wahl A, Stark T, Hofmann T. Urinary N-methylpyridinium and trigonelline as candidate dietary biomarkers of coffee consumption. Mol Nutr Food Res 2011;55:1613–23.
47. Lee JE, Lee BJ, Chung JO, Kim HN, Kim EH, Jung S, Lee H, Lee SJ, Hong YS. Metabolomic unveiling of a diverse range of green tea (Camellia sinensis) metabolites dependent on geography. Food Chem 2015;174:452–9.
48. Vuong QV. Epidemiological evidence linking tea consumption to human health: a review. Crit Rev Food Sci Nutr 2014;54:523–36.
49. Murtaugh MA, Herrick J, Sweeney C, Guiliano A, Baumgartner K, Byers T, Slattery M. Macronutrient composition influence on breast cancer risk in Hispanic and non-Hispanic white women: the 4-Corners Breast Cancer Study. Nutr Cancer 2011;63:185–95.
50. Bingham SA, Day N. Commentary: fat and breast cancer: time to re-evaluate both methods and results? Int J Epidemiol 2006;35:1022–4.
51. Freedman LS, Potischman N, Kipnis V, Midthune D, Schatzkin A, Thompson FE, Troiano RP, Prentice R, Patterson R, Carroll R, et al. A comparison of two dietary instruments for evaluating the fat-breast cancer relationship. Int J Epidemiol 2006;35:1011–21.
52. Key TJ, Appleby PN, Cairns BJ, Luben R, Dahm CC, Akbaraly T, Brunner EJ, Burley V, Cade JE, Greenwood DC, et al. Dietary fat and breast cancer: comparison of results from food diaries and food-frequency questionnaires in the UK Dietary Cohort Consortium. Am J Clin Nutr 2011;94:1043–52.
53. Cao Y, Hou L, Wang W. Dietary total fat and fatty acids intake, serum fatty acids and risk of breast cancer: a meta-analysis of prospective cohort studies. Int J Cancer 2016;138:1894–904.
54. Sczaniecka AK, Brasky TM, Lampe JW, Patterson RE, White E. Dietary intake of specific fatty acids and breast cancer risk among postmenopausal women in the VITAL cohort. Nutr Cancer 2012;64:1131–42.
55. Johnson CH, Manna SK, Krausz KW, Bonzo JA, Divelbiss RD, Hollingshead MG, Gonzalez FJ. Global metabolomics reveals urinary biomarkers of breast cancer in a Mcf-7 xenograft mouse model. Metabolites 2013;3:658–72.
56. Schütze M, Boeing H, Pischon T, Rehm J, Kehoe T, Gmel G, Olsen A, Tjønneland AM, Dahm CC, Overvad K, et al. Alcohol attributable burden of incidence of cancer in eight European countries based on results from prospective cohort study. BMJ 2011;342:d1584.
57. Dorgan JF, Baer DJ, Albert PS, Judd JT, Brown ED, Corle DK, Campbell WS, Hartman TJ, Tejpar AA, Clevidence BA, et al. Serum hormones and the alcohol-breast cancer association in postmenopausal women. J Natl Cancer Inst 2001;93:710–5.
58. Kalogera E, Pistos C, Provatopoulou X, Athanaselis S, Spiliopoulou C, Gounaris A. Androgen glucuronides analysis by liquid chromatography tandem-mass spectrometry: could it raise new perspectives in the diagnostic field of hormone-dependent malignancies? J Chromatogr B Analyt Technol Biomed Life Sci 2013;940:24–34.
59. Danforth KN, Eliassen AH, Tworoger SS, Missmer SA, Barbieri RL, Rosner BA, Colditz GA, Hankinson SE. The association of plasma androgen levels with breast, ovarian and endometrial cancer risk factors among postmenopausal women. Int J Cancer 2010;126:199–207.
60. Key TJ, Appleby PN, Reeves GK, Roddam AW, Helzlsouer KJ, Alberg AJ, Rollison DE, Dorgan JF, Brinton LA, Overvad K, et al.; Endogenous Hormones and Breast Cancer Collaborative Group. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. Br J Cancer 2011;105:709–22.
61. Key T, Appleby P, Barnes I, Reeves G; Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 2002;94:606–16.
62. Key TJ, Appleby PN, Reeves GK, Travis RC, Brinton LA, Helzlsouer KJ, Dorgan JF, Gapstur SM, Gaudet MM, Kaaks R, et al. Steroid hormone measurements from different types of assays in relation to body mass index and breast cancer risk in postmenopausal women: reanalysis of eighteen prospective studies. Steroids 2015;99:49–55.
63. Troche JR, Mayne ST, Freedman ND, Shebl FM, Guertin KA, Cross AJ, Abnet CC. Alcohol consumption-related metabolites in relation to colorectal cancer and adenoma: two case-control studies using serum biomarkers. PLoS One 2016;11:e0150962.

64. Jiang Q, Christen S, Shigenaga MK, Ames BN. γ-Tocopherol, the major form of vitamin E in the US diet, deserves more attention. Am J Clin Nutr 2001;74:714–22.
65. Smolarek AK, Suh N. Chemopreventive activity of vitamin E in breast cancer: a focus on gamma- and delta-tocopherol. Nutrients 2011;3: 962–86.
66. Yoshikawa S, Morinobu T, Hamamura K, Hirahara F, Iwamoto T, Tamai H. The effect of gamma-tocopherol administration on alpha-tocopherol levels and metabolism in humans. Eur J Clin Nutr 2005; 59:900–5.
67. Kline K, Yu W, Sanders BG. Vitamin E and breast cancer. J Nutr 2004; 134:3458S–62S.
68. Fortmann SP, Burda BU, Senger CA, Lin JS, Whitlock EP. Vitamin and mineral supplements in the primary prevention of cardiovascular disease and cancer: an updated systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med 2013;159:824–34.
69. Moyer VA; US Preventive Services Task Force. Vitamin, mineral, and multivitamin supplements for the primary prevention of cardiovascular disease and cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2014;160:558–64.
70. Dorjgochoo T, Gao YT, Chow WH, Shu XO, Li H, Yang G, Cai Q, Rothman N, Cai H, Franke AA, et al. Plasma carotenoids, tocopherols, retinol and breast cancer risk: results from the Shanghai Women Health Study (SWHS). Breast Cancer Res Treat 2009;117:381–9.
71. Bakker MF, Peeters PH, Klaasen VM, Bueno-de-Mesquita HB, Jansen EH, Ros MM, Travier N, Olsen A, Tjonneland A, Overvad K, et al. Plasma carotenoids, vitamin C, tocopherols, and retinol and the risk of breast cancer in the European Prospective Investigation into Cancer and Nutrition cohort. Am J Clin Nutr 2016;103:454–64.
72. Epplein M, Shvetsov YB, Wilkens LR, Franke AA, Cooney RV, Le Marchand L, Henderson BE, Kolonel LN, Goodman MT. Plasma carotenoids, retinol, and tocopherols and postmenopausal breast cancer risk in the Multiethnic Cohort Study: a nested case-control study. Breast Cancer Res 2009;11:R49.
73. Maillard V, Kuriki K, Lefebvre B, Boutron-Ruault MC, Lenoir GM, Joulin V, Clavel-Chapelon F, Chajes V. Serum carotenoid, tocopherol and retinol concentrations and breast cancer risk in the E3N-EPIC study. Int J Cancer 2010;127:1188–96.
74. Dorgan JF, Sowell A, Swanson CA, Potischman N, Miller R, Schussler N, Stephenson HE Jr. Relationships of serum carotenoids, retinol, alpha-tocopherol, and selenium with breast cancer risk: results from a prospective study in Columbia, Missouri (United States). Cancer Causes Control 1998;9:89–97.
75. Sato R, Helzlsouer KJ, Alberg AJ, Hoffman SC, Norkus EP, Comstock GW. Prospective study of carotenoids, tocopherols, and retinoid concentrations and the risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2002;11:451–7.
76. Tamimi RM, Hankinson SE, Campos H, Spiegelman D, Zhang S, Colditz GA, Willett WC, Hunter DJ. Plasma carotenoids, retinol, and tocopherols and risk of breast cancer. Am J Epidemiol 2005;161:153–60.
77. Kabat GC, Kim M, Adams-Campbell LL, Caan BJ, Chlebowski RT, Neuhouser ML, Shikany JM, Rohan TE, Investigators WHI. Longitudinal study of serum carotenoid, retinol, and tocopherol concentrations in relation to breast cancer risk among postmenopausal women. Am J Clin Nutr 2009;90:162–9.
78. Hu F, Wu Z, Li G, Teng C, Liu Y, Wang F, Zhao Y, Pang D. The plasma level of retinol, vitamins A, C and alpha-tocopherol could reduce breast cancer risk? A meta-analysis and meta-regression. J Cancer Res Clin Oncol 2015;141:601–14.
79. Liesenfeld DB, Habermann N, Owen RW, Scalbert A, Ulrich CM. Review of mass spectrometry-based metabolomics in cancer research. Cancer Epidemiol Biomarkers Prev 2013;22:2182–201.
80. Fages A, Ferrari P, Monni S, Dossus L, Floegel A, Mode N, Johansson M, Travis RC, Bamia C, Sanchez-Perez MJ, et al. Investigating sources of variability in metabolomic data in the EPIC study: the Principal Component Partial R-square (PC-PR2) method. Metabolomics 2014;10:1074–83.