

# Plasma Metabolomic Signatures Associated with Long-term Breast Cancer Risk in the SU.VI.MAX Prospective Cohort



Lucie Lécuyer<sup>1</sup>, Céline Dalle<sup>2</sup>, Bernard Lyan<sup>2</sup>, Aicha Demidem<sup>3</sup>, Adrien Rossary<sup>3</sup>, Marie-Paule Vasson<sup>3,4</sup>, Mélanie Petera<sup>2</sup>, Marie Lagree<sup>5</sup>, Thomas Ferreira<sup>3</sup>, Delphine Centeno<sup>2</sup>, Pilar Galan<sup>1</sup>, Serge Hercberg<sup>1,6</sup>, Mélanie Deschasaux<sup>1</sup>, Valentin Partula<sup>1</sup>, Bernard Srour<sup>1</sup>, Paule Latino-Martel<sup>1</sup>, Emmanuelle Kesse-Guyot<sup>1</sup>, Nathalie Druésne-Pecollo<sup>1</sup>, Stéphanie Durand<sup>2</sup>, Estelle Pujos-Guillot<sup>2</sup>, and Mathilde Touvier<sup>1</sup>

## Abstract

**Background:** Breast cancer is a major cause of death in occidental women. The role of metabolism in breast cancer etiology remains unclear. Metabolomics may help to elucidate novel biological pathways and identify new biomarkers to predict breast cancer long before symptoms appear. The aim of this study was to investigate whether untargeted metabolomic signatures from blood draws of healthy women could contribute to better understand and predict the long-term risk of developing breast cancer.

**Methods:** A nested case-control study was conducted within the SU.VI.MAX prospective cohort (13 years of follow-up) to analyze baseline plasma samples of 211 incident breast cancer cases and 211 matched controls by LC/MS. Multivariable conditional logistic regression models were computed.

**Results:** A total of 3,565 ions were detected and 1,221 were retained for statistical analysis. A total of 73 ions were associated with breast cancer risk ( $P < 0.01$ ;  $FDR \leq 0.2$ ). Notably,

we observed that a lower plasma level of O-succinyl-homoserine ( $OR=0.70$ ,  $95\%CI=[0.55-0.89]$ ) and higher plasma levels of valine/norvaline [1.45 (1.15–1.83)], glutamine/isoglutamine [1.33 (1.07–1.66)], 5-aminovaleric acid [1.46 (1.14–1.87)], phenylalanine [1.43 (1.14–1.78)], tryptophan [1.40 (1.10–1.79)],  $\gamma$ -glutamyl-threonine [1.39 (1.09–1.77)], ATBC [1.41 (1.10–1.79)], and pregnene-triol sulfate [1.38 (1.08–1.77)] were associated with an increased risk of developing breast cancer during follow-up.

**Conclusion:** Several prediagnostic plasmatic metabolites were associated with long-term breast cancer risk and suggested a role of microbiota metabolism and environmental exposure.

**Impact:** After confirmation in other independent cohort studies, these results could help to identify healthy women at higher risk of developing breast cancer in the subsequent decade and to propose a better understanding of the complex mechanisms involved in its etiology.

<sup>1</sup>Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), French National Institute of Health and Medical Research (Inserm) U1153, French National Institute for Agricultural Research (Inra) U1125, French National Conservatory of Arts and Crafts (Cnam), Paris 13 University, Nutritional Epidemiology Research Team (EREN), Bobigny, France. <sup>2</sup>Clermont Auvergne University, INRA, UNH, Plateforme d'Exploration du Métabolisme, MetaboHUB Clermont, Clermont-Ferrand, France. <sup>3</sup>Clermont Auvergne University, INRA, UMR 1019, Human Nutrition Unit (UNH), CRNH Auvergne, Cellular Micro-Environment, Immunomodulation and Nutrition (ECREIN), Clermont-Ferrand, France. <sup>4</sup>Anti-cancer Center Jean-Perrin, CHU Clermont-Ferrand, France. <sup>5</sup>Clermont Auvergne University, Institut de Chimie de Clermont-Ferrand, Plateforme d'Exploration du Métabolisme, MetaboHUB-Clermont, BP 80026, Aubière, France. <sup>6</sup>Public Health Department, Avicenne Hospital, Bobigny, France.

**Note:** Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

Clinical Trials.gov registration ID: SU.VI.MAX, NCT00272428. Registered on January 3, 2006.

**Corresponding Author:** Lucie Lécuyer, Inserm - U1153 team 3 (Equipe de Recherche en Epidemiologie Nutritionnelle)/INRA U1125/Paris 13 University, 74 rue Marcel Cachin, Bobigny Cedex 93000, France. Phone: 331-4838-7638; Fax: 331-4838-8931; E-mail: llecuyer@eren.smbh.univ-paris13.fr

Cancer Epidemiol Biomarkers Prev 2019;28:1300–7

doi: 10.1158/1055-9965.EPI-19-0154

©2019 American Association for Cancer Research.

## Introduction

Breast cancer is the most common female cancer in the world ([www.dietandcancerreport.org](http://www.dietandcancerreport.org)) and ranks as the fifth cause of death from cancer overall (626,679 deaths worldwide in 2018; ref. 1). Despite some well-known risk factors such as adult height, reproductive life, and hormonal factors, and, for postmenopausal women, alcoholic drinks, body fatness, and adult weight gain ([www.wcrf.org/breast-cancer-2017](http://www.wcrf.org/breast-cancer-2017)), the role of metabolism in breast cancer etiology remains unclear and hypothesis-driven research has not yet allowed to entirely elucidate it.

Recent advances in modern explorative technologies, such as metabolomics, increase the capacity to identify additional risk factors and better understand mechanisms involved (2). Indeed, mass spectrometry (MS) and NMR allows detecting a wide range of low molecular weight biochemicals in biofluids, cells, or tissues (3) and thus can provide a signature of both exogenous exposures and endogenous metabolism. Therefore, metabolomics allows identifying biomarkers relevant to cancer etiology and elucidating novel biological pathways (4–6). Several cross-sectional epidemiologic studies highlighted differences between metabolomic profiles and patients with breast cancer and

controls (7, 8). An NMR-based prospective study (5 years of follow-up) recently showed the benefit of metabolomic signatures of breast cancer risk (9). However metabolomic signatures may be altered by the presence of a breast tumor itself (10), arguing for similar investigations with a long follow-up. Improving the ability to discriminate profiles at higher risk of developing breast cancer in the subsequent decades through the identification of prediagnostic biomarkers, to establish preventive approaches, remains an important focus in breast cancer research.

Two recent semiuntargeted MS studies investigated the associations between diet- (11) or body mass index (BMI)-related metabolites (12) and long-term breast cancer risk, and found promising results supporting the role of diet or BMI in postmenopausal breast cancer etiology. We previously set-up a prospective nested case-control study within the SU.VI.MAX cohort to investigate whether baseline metabolomic profiles from plasma of apparently healthy women could contribute to identify women who would develop breast cancer during the subsequent years, and better understand the etiology of this complex disease. Using an NMR metabolomics analysis, higher fasting plasma levels of valine, lysine, glutamine, and glucose and lower plasma levels of lipoproteins and glycerol-derived compounds were notably found to be associated with a higher risk of developing breast cancer (13). To increase the number of metabolites detected both in terms of sensitivity and metabolome coverage, this study applied for the first time a fully untargeted and nonoriented MS analysis, with the objective to highlight new endogenous or exogenous metabolites associated with long-term risk of breast cancer.

## Materials and Methods

### Population study

This analysis was based on a case-control study nested within the SU.VI.MAX prospective cohort (clinicaltrials.gov; NCT00272428). The latter was initially designed to investigate the influence of a daily supplementation with nutritional doses of antioxidants on the incidence of cardiovascular diseases and cancers. The study design and methods have been previously detailed (14, 15). Briefly, this study is a population-based, double-blind, placebo-controlled, randomized trial enrolling a total of 13,017 participants recruited in 1994–1995. The intervention study lasted 8 years, and observational follow-up of health events was maintained until September 2007. Written informed consent was obtained from all participants. The trial was approved by the Ethics Committee for Studies with Human Subjects of Paris-Cochin Hospital (CCPPRB 706/2364) and the 'Commission Nationale de l'Informatique et des Libertés' (CNIL 334641/907094) and was conducted according to the Declaration of Helsinki guidelines.

### Baseline data collection

At inclusion, self-administered questionnaires about sociodemographic characteristics, smoking status, medication use, health status, and family history of cancer were filled-in by participants. During a baseline clinical examination, all participants underwent anthropometric measurements (height and weight) and a blood draw (occurring after a 12-hour fasting period) by study nurses and physicians. These 35 mL venous blood samples were collected in sodium heparin Vacutainer Tubes (Becton Dickinson,

son.). After centrifugation at +4°C, plasma aliquots were immediately prepared and stored frozen at –20°C during less than 2 days and then stored in liquid nitrogen.

### Case ascertainment

Health events that occurred during follow-up (from 1994 to 2007) were self-reported by participants. Medical data were gathered through participants, physicians, and hospitals and reviewed by an independent physician expert committee. Pathologic reports were used to validate cases and extract cancer characteristics. Cancers were classified by using the International Classification of Diseases, 10th Revision, Clinical Modification (16).

### Nested case-control study

This analysis initially included 215 participants with a first incident invasive breast cancer matched 1:1 with controls for baseline age (35–39 years/40–44 years/45–49 years/50–54 years/55–59 years/>60 years), menopausal status (pre-/postmenopause status at baseline and at diagnosis), BMI (<18.5 kg/m<sup>2</sup>, underweight; ≥18.5–<25 kg/m<sup>2</sup>, normal weight; and ≥25 kg/m<sup>2</sup>, overweight/obese), arm assignment to the initial SU.VI.MAX clinical trial (placebo/supplemented), smoking status (current smokers and non/former smokers at baseline), and season of blood draw (*a priori*-defined periods: October–November/December–January–February/March–April–May). Controls selection was based on the classical density sampling method (17), that is, every time a case was diagnosed, one control was selected from other participants from the cohort with no breast cancer at that time.

### Metabolomic analyses

Metabolomic profiling was conducted using ultrahigh performance liquid chromatography/mass spectroscopy with the Metabolic Profiler Platform (Bruker Daltonique). Details on chemicals and reagents, biological samples preparation, metabolic profiling, raw data extraction, and metabolite identification are available in Supplementary Materials and Methods and Supplementary Fig. S1. Samples were randomized within the analytic sequence based on a Williams Latin Square strategy defined according to the main factors of the study. The stability of the analytic system was monitored using pooled samples as quality control and analyzed in the same sequence than participant samples (injected one time at the beginning of each sequence and then after each set of 12 samples). This metabolomic discovery approach allows semiquantitative measurements, representing relative ion intensities in arbitrary unit. However, even if it is not absolute quantitation, there are linear relations between ion intensity for each metabolite and its concentration level in the samples.

### Statistical analysis

Participants' baseline characteristics were compared between cases and controls using  $\chi^2$  tests for categorical variables and Student *t* test for continuous variables. Less than 5% of values were missing for all covariates and missing values were replaced by the mode.

Prospective associations between each ion's signal intensity and breast cancer risk were characterized using multivariable conditional logistic regression models. OR for an increment of 1 SD for each metabolite and their 95% confidence intervals (95%

CI) were computed, as well as tests for linear trend. Multivariable models were adjusted for the following baseline characteristics: age (continuous), BMI (continuous), smoking status (current, former, and nonsmokers), height (continuous), alcohol intake (continuous), physical activity (low, moderate, and intense), education level (primary, secondary, and superior), family history of breast cancer, number of children, and use of hormone replacement therapy for menopause; all these covariates constituting the reference model. Models were adjusted for matching factors only when the variable was more precise than the one used for the matching, to avoid residual confounding. The models were not additionally adjusted for menopausal status (pre-/postmenopause status at baseline and at diagnosis), season of blood draw (October–November/December–January–February/March–April–May), and arm assignment to the initial SU.VI. MAX clinical trial (yes/no) because cases and controls were already matched for these variables with the same coding. To limit the false negative rate (type II error = risk of not detecting real associations) in this exploratory untargeted metabolomic study, a FDR level of 0.2 was retained from the Benjamini Hochberg procedure (ref. 18; 1,221 tests: 530 ions in Electrospray (ESI)-positive mode and 691 ions in ESI-negative mode). Linearity assumption was checked by plotting the ion's intensity as a function of logit(predicted values of breast cancer risk). ORs for categorized intensities in tertiles were also computed, as well as a Spearman correlation matrix of the identified associated metabolites.

Sensitivity analysis was performed on the same models by excluding cases diagnosed during the first year of follow-up, as well as the matched controls to prevent reverse causality bias. Exploratory stratified and heterogeneity (interaction) analyses were also performed according to menopausal status at diagnosis and according to the median follow-up duration: 5.8 years (to investigate differentially short- or long-term associations with breast cancer risk). Interaction was tested as the produce of each ion (continuous/SD) by these two factors (two categories each).

A multivariate analysis using a principal component analysis (PCA) with rotated factors by orthogonal transformation was performed on ions associated with breast cancer from logistic regressions. The four first factors were retained according to the plot of the total explained variance (Cattel scree plots; ref. 19). The multivariable conditional logistic regression models were run on these factors (with the same adjustments as described above) to investigate the associations between metabolomic patterns (combinations of metabolites variations) and breast cancer risk.

Analyses were performed using SAS software (v9.3) and plots were generated using R software (v3.5.2). All of the statistical tests were two-sided.

#### Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

#### Informed consent

Informed consent was obtained from all individual participants included in the study.

## Results

The baseline characteristics of breast cancer cases and controls in the study population are summarized in Table 1. The cases were more likely to be taller with more frequent family history of breast cancer than controls. A total of 25% of cases were diagnosed before 3.5 years, 50% before 5.8 years, and 75% before 9.4 years. The maximum length of follow-up was 12.8 years.

A total of 3,565 ions were detected and 1,221 were retained for statistical analysis. From logistic regression analyses, 53 ions were positively associated with breast cancer risk and 20 ions were inversely associated ( $P \leq 0.01$ ; FDR  $\leq 0.2$ ). Results of all associated ions are shown in Supplementary Table S1. Among these ions, some have been identified using available databases (in-house database, Metlin; <https://metlin.scripps.edu>), Human Metabolome Database ([www.hmdb.ca](http://www.hmdb.ca)), Kyoto Encyclopedia of Genes and Genomes database, and analytic data (as MS2); the associations between these ions ( $N = 9$ ) and breast cancer risk are displayed in Table 2. A lower plasma level of O-succinyl-homoserine (OR=0.70, 95%CI=[0.55-0.89]) and higher plasma levels of valine/norvaline [1.45 (1.15–1.83)], glutamine/isoglutamine [1.33 (1.07–1.66)], 5-aminovaleric acid [1.46 (1.14–1.87)], phenylalanine [1.43 (1.14–1.78)], tryptophan [1.40 (1.10–1.79)],  $\gamma$ -glutamyl-threonine [GGT; 1.39 (1.09–1.77)], acetyl tributyl citrate ATBC [1.41 (1.10–1.79)], and the putative pregnene-triol sulfate [1.38 (1.08–1.77)] were associated with an increased risk of developing breast cancer during follow-up. Linearity assessments, as well as analyses performed with ions categorized in tertiles were provided in Supplementary Fig. S2A and S2B and Supplementary Table S2. The Spearman correlation matrix of the identified associated metabolites is provided in Supplementary Table S3. Positive correlations ( $P < 0.05$ ) were found for pregnene-triol sulfate with 5-aminovaleric acid, tryptophan, and GGT; for phenylalanine with valine, tryptophan, 5-aminovaleric acid, and glutamine; for valine with 5-aminovaleric acid, tryptophan, and glutamine; for 5-aminovaleric acid with GGT; for tryptophan with glutamine; and for GGT with glutamine. Inverse correlations were found for O-succinyl-homoserine with 5-aminovaleric acid and ATBC; and for 5-aminovaleric acid with ATBC. Results of sensitivity, stratification, and heterogeneity analyses are shown in Supplementary Tables S4, S5, and S6. All identified ions remained associated with breast cancer when excluding cases diagnosed before 1 year of follow-up. No interaction was found between breast cancer risk and duration of follow-up. Higher plasma level of glutamine was associated with an increased risk of breast cancer only in premenopausal women ( $P_{\text{interaction}} = 0.003$ ; FDR = 0.2).

All ions associated with breast cancer risk in logistic regressions were then analyzed by PCA. The four first PCA factors (metabolomic patterns) explained more than 36% of total variability. Loading values are displayed in Supplementary Table S7. The first factor was notably characterized by higher plasma levels of valine and 5-aminovaleric acid, the second factor was in particular characterized by higher plasma levels of phenylalanine, valine, 5-aminovaleric acid, and tryptophan, the third by lower plasma level of pregnene-triol sulfate, and the fourth by higher plasma levels of phenylalanine, valine 5-aminovaleric acid, GGT, and glutamine (loading values  $\geq |0.20|$ ). Moreover, several unidentified ions also characterized these four factors. These data allowed distinguishing three metabolomic patterns (factors 1, 2, and 4)

**Table 1.** Baseline characteristics of breast cancer cases and controls<sup>a</sup>, SU.VI. MAX cohort, France (1994–2007)

	Breast cancer cases (N = 211)	Controls (N = 211)	P <sup>b</sup>
Age at baseline (y)	49.0 ± 6.0	48.9 ± 6.0	0.9
BMI (kg/m <sup>2</sup> )	23.2 ± 4.0	23.6 ± 4.2	0.3
<18.5 kg/m <sup>2</sup> (underweight)	9 (4.3)	6 (2.8)	0.7
≥18.5–<25 kg/m <sup>2</sup> (normal weight)	148 (70.1)	151 (71.6)	
≥25 kg/m <sup>2</sup> (overweight)	54 (25.6)	54 (25.6)	
Height (cm)	163.0 ± 6.2	160.9 ± 5.9	0.0004
Intervention group			1
Placebo	105 (49.8)	105 (49.8)	
Antioxidants	106 (50.2)	106 (50.2)	
Smoking status			1
Never and former	164 (77.7)	164 (77.7)	
Current smoker	47 (22.3)	47 (22.3)	
Physical activity			0.8
Irregular	67 (31.8)	60 (28.4)	
<1 h/d walking equivalent	69 (32.7)	72 (34.1)	
≥1 h/d walking equivalent	75 (35.5)	79 (37.4)	
Educational level			0.3
Primary	37 (17.5)	44 (20.9)	
Secondary	82 (38.9)	91 (43.1)	
Superior	92 (43.6)	76 (36.0)	
Number of biological children	1.8 ± 1.2	2 ± 1.2	0.1
Hormonal treatment for menopause (yes)	76 (36.0)	75 (35.5)	0.9
Menopausal status at baseline			1
Premenopausal	129 (61.1)	129 (61.1)	
Postmenopausal	82 (38.9)	82 (38.9)	
Menopausal status at diagnosis			1
Premenopausal	78 (37.0)	78 (37.0)	
Postmenopausal	133 (63.0)	133 (63.0)	
Family history of breast cancer <sup>c</sup> (yes)	35 (16.6)	21 (10.0)	0.04
Alcohol intake (g/d)	11.2 ± 12.0	11.5 ± 13.0	0.8
Month of blood draw			1
March–April–May	77 (36.5)	77 (36.5)	
October–November	31 (14.7)	32 (15.2)	
December–January–February	103 (48.8)	102 (48.3)	

<sup>a</sup>Values are means ± SD or n (%).

<sup>b</sup>P for the comparison between breast cancer cases and controls using  $\chi^2$  tests for categorical variables or Student *t* test (from analysis of variance models) for continuous variables.

<sup>c</sup>Among first-degree female relatives.

associated with an increased risk of breast cancer in multivariable conditional logistic regression models and one (factor 3) with a decreased risk (FDR < 0.05; Table 3).

## Discussion

To our knowledge, this prospective study was the first fully untargeted and nonoriented LC/MS-based analysis aiming to investigate the associations between prediagnosis metabolic profiles and long-term breast cancer risk. Untargeted metabolomics, by the detection of a large number of metabolites, allowed discriminating early phenotypes and revealing a plasma metabolic signature announcing the evolution toward breast cancer in the subsequent 13 years. This could be relevant for detecting specific plasmatic profiles of women at higher risk for this disease and improve personalized prevention. A profile (signature) of circulating levels of multiple metabolites provides insights to draw an overall picture of the different pathways involved in the

etiology of this cancer (20). Even if the origin of metabolites variations (exogenous or endogenous) cannot be determined in this study, our results also illustrate that metabolomics integrates information regarding the influence of both extrinsic and intrinsic factors, such as lifestyle exposures and basal metabolism, and consequently address the complex issue of environment and health interaction.

In this study, we found that a lower plasma level of O-succinyl-homoserine and higher plasma levels of valine/norvaline, glutamine/isoglutamine, 5-aminovaleic acid, phenylalanine, tryptophan,  $\gamma$ -glutamyl-threonine, ATBC, and pregnene-triol sulfate were associated with an increased risk of developing breast cancer during the following decade. Three metabolomic patterns characterized by several concurrently metabolites variations were associated with an increased risk of breast cancer, as well as another metabolomic pattern associated with a decreased risk. These univariate analysis (analysis of one by one metabolite variation) and multivariate analysis (analysis of combinations of simultaneous metabolites variations) allowed getting complementary overviews of metabolic signatures discriminating phenotypes at higher risk for this disease well before breast cancer clinical onset.

We previously highlighted a higher plasma level of glutamine and valine in the future breast cancer cases with NMR analyses on the same population study (13), which was consistent with our current results. An increase in plasmatic levels of valine and three  $\gamma$ -glutamyl dipeptides was also recently found to be positively associated with BMI and estrogen receptor-positive breast cancer risk in postmenopausal women through a semiuntargeted MS study (12). The pregnene-triol sulfate is a steroid hormone and belongs to progestative family compounds. In line with our results, two semiuntargeted MS studies found a positive association between steroid hormones and breast cancer risk in postmenopausal women (11, 12). Four metabolomic studies reported that an increased level of phenylalanine was positively associated with hepatocellular carcinoma (HCC) risk (21–24) and three of these studies found that a higher level of glutamine was associated with a decreased risk of HCC (21–23). Concerning valine, findings were more contrasted; one of these studies showed a positive association with HCC risk (21), whereas another observed an inverse association (22). An increase in plasmatic valine level was also reported in two metabolomic studies to be positively associated with pancreatic adenoma risk (25) and inversely associated with prostate cancer risk (26). In three metabolomic studies an inverse association was found between phenylalanine (26), tryptophan (27),  $\gamma$ -glutamyl-histidine (28), and  $\gamma$ -glutamyl-phenylalanine (26, 27) and prostate cancer risk, whereas a positive association was reported between  $\gamma$ -glutamyl-glutamine and nonaggressive prostate cancer risk (27). The comparisons of our results with other prospective metabolomic studies should be interpreted with caution, taking into account methodologic and technological variations (as the wide range of metabolomics platforms and types of analyses), because the etiology may vary according to cancer locations.

The findings highlighted in this study provide evidence for circulating metabolites as putative risk biomarkers for breast cancer. Given that metabolites concerned by the observed variations are not specific of one metabolic pathway and as their variations could arise from several types of exposure and endogenous reactions, precaution is needed in the mechanistic interpretation of our epidemiologic findings. Even so, we propose

**Table 2.** Identified ions associated with breast cancer risk from multivariable conditional logistic regression<sup>a</sup>, SU.VI.MAX cohort, France (1994–2007)

Name (level of confidence for identification <sup>b</sup> )	Annotation	Measured mass m/z	Retention time (minute)	Cases/controls		Mean ± SD	OR (95% CI)	P <sub>trend</sub>
				Mode	controls			
L-phenylalanine (1)	[M+H] <sup>+</sup>	166.08618	10.11	ESI <sup>+</sup>	211/211	227,718 ± 42,535	1.43 (1.14–1.78)	0.002 <sup>c</sup>
L-valine/norvaline (1)	[M+H-CH <sub>2</sub> O <sub>2</sub> ] <sup>+</sup>	72.08089	3.08	ESI <sup>+</sup>	211/211	115,033 ± 24,435	1.45 (1.15–1.83)	0.002 <sup>c</sup>
5-aminovaleric acid (1)	[M+Na] <sup>+</sup>	140.06816	2.29	ESI <sup>+</sup>	211/211	10,219 ± 5,139	1.46 (1.14–1.87)	0.003 <sup>d</sup>
O-succinyl-L-homoserine (1)	[M+Na-2H] <sup>-</sup>	240.04755	2.86	ESI <sup>-</sup>	203/203	919 ± 400	0.70 (0.55–0.89)	0.004 <sup>d</sup>
ATBC (acetyl tributyl citrate) (1)	[M+Na] <sup>+</sup>	425.21694	19.31	ESI <sup>+</sup>	211/211	2,198 ± 964	1.41 (1.10–1.79)	0.006 <sup>d</sup>
L-tryptophan (1)	[M+Na] <sup>+</sup>	227.07887	11.09	ESI <sup>+</sup>	211/211	799 ± 180	1.40 (1.10–1.79)	0.007 <sup>d</sup>
L-γ-glutamyl-L-threonine (1)	[M+H-H <sub>2</sub> O] <sup>+</sup>	231.1451	2.52	ESI <sup>+</sup>	211/211	1,054 ± 601	1.39 (1.09–1.77)	0.008 <sup>d</sup>
L-glutamine/L-isoglutamine (1)	[M+H-C <sub>4</sub> H <sub>8</sub> O] <sup>+</sup>	90.0551	2.32	ESI <sup>+</sup>	211/211	13,187 ± 2,982	1.33 (1.07–1.66)	0.009 <sup>d</sup>
Pregnene-triol sulfate (2)	C <sub>21</sub> H <sub>33</sub> O <sub>6</sub> S	413.19771	16.11	ESI <sup>-</sup>	203/203	2,441 ± 1,254	1.38 (1.08–1.77)	0.01 <sup>d</sup>

<sup>a</sup>Multivariable models were adjusted for age (continuous), BMI (continuous), smoking status (current smokers, former smokers, and nonsmokers), season of blood draw (*a priori*-defined periods: October–November/December–January–February/March–April–May), height (continuous), alcohol intake (continuous), physical activity (irregular/<1 h/d walking equivalent/≥1 h/d walking equivalent), education level (primary/secondary/superior), family history of breast cancer (yes/no), number of children (continuous), and use of hormone replacement therapy for menopause (yes/no; reference model). Tests for linear trend were performed using the continuous variables. ORs were presented for an increment of 1 SD for each metabolite.

<sup>b</sup>Metabolites were classified according to Sumner and colleagues (46) concerning the levels of confidence in the identification process: identified (level 1: confirmed by standard), putatively annotated (level 2: based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries).

<sup>c</sup>FDR after Benjamini Hochberg correction for multiple testing = 0.1.

<sup>d</sup>FDR after Benjamini Hochberg correction for multiple testing = 0.2.

interpretation hypotheses for some presumed pathways, summarized in Fig. 1. Among the possible interpretations, our results could suggest a perturbation of energy and amino acids metabolism, a set-up of conditions favoring cell proliferation, inflammation, and oxidative stress in women who will develop breast cancer involving environmental exposure and microbiota metabolism. Glutamine and valine (one of the branched chain amino acids; BCAA) positively correlated in our study, are two glycogenic amino acids; their plasma increase could suggest an amplified phenomenon of gluconeogenesis. BCAAs were previously suggested as markers of cardiovascular disease risk (29), insulin resistance (30), and diabetes (31). Moreover there is convincing evidence on the effect of obesity on an increased risk of post-

menopausal breast cancer, and pathways related to insulin resistance is one of the number of plausible mechanisms that may explain this relationship (32). Moreover, valine and glutamine are regulators of many cell signaling pathways including activators of the mTOR pathway; one of the main cell proliferation routes (31). ATBC is an alternative plasticizer to phthalates (33) widely used in products such as food wrap, vinyl toys, and pharmaceutical excipients. Although this contaminant does not seem to accumulate in tissues (34), recent data revealed potential biological activity on tissue growth (35). The combination of plasmatic increase in both ATBC and in a progestative sulfate steroid may interact either with PPAR signaling involved in cell proliferation and inflammation, and/or with steroid signaling (36). Indeed, ATBC has been shown to influence steroid levels (35) and is known to interact with several signaling pathways as a ligand for steroid xenobiotic receptor or sexual hormone-binding globulin involved in steroid signaling (33). Phenylalanine and tryptophan are aromatic amino acids (AAA, correlated in our study;  $R = 0.27$ ) and high level of AAAs in blood have been considered as an inflammatory marker in patients with liver cirrhosis (37). Disruptions in AAAs and BCAAs levels are found in several pathologies with chronic low-grade inflammation such as obesity and diabetes (38, 39); in our study valine (BCAA) is highly positively correlated with AAAs (phenylalanine;  $R = 0.47$  and tryptophan;  $R = 0.24$ ). Data on the kynurenine/tryptophan ratio, unfortunately unavailable in this study, would be useful to support this hypothesis, as it reflects IDO1 gene expression, which increases in situations of chronic low-grade inflammation. 5-aminovaleric acid is a catabolism product from lysine; its increase could witness either the presence of cellular necrosis or microbiota metabolism. In mice, higher urinary levels of 5-aminovaleric acid were associated with intestinal inflammation (40). Concerning identified metabolites, the second PCA-derived factor highlighted in our study could represent an inflammatory pattern, as it was notably characterized by higher plasma levels of phenylalanine, valine, 5-aminovaleric acid, and tryptophan. Our findings could also reflect an influence of gut microbiota, because O-succinyl-homoserine is a metabolite exclusively synthesized by bacteria. Another metabolite synthesized by gut microbiota, 2-amino-4-

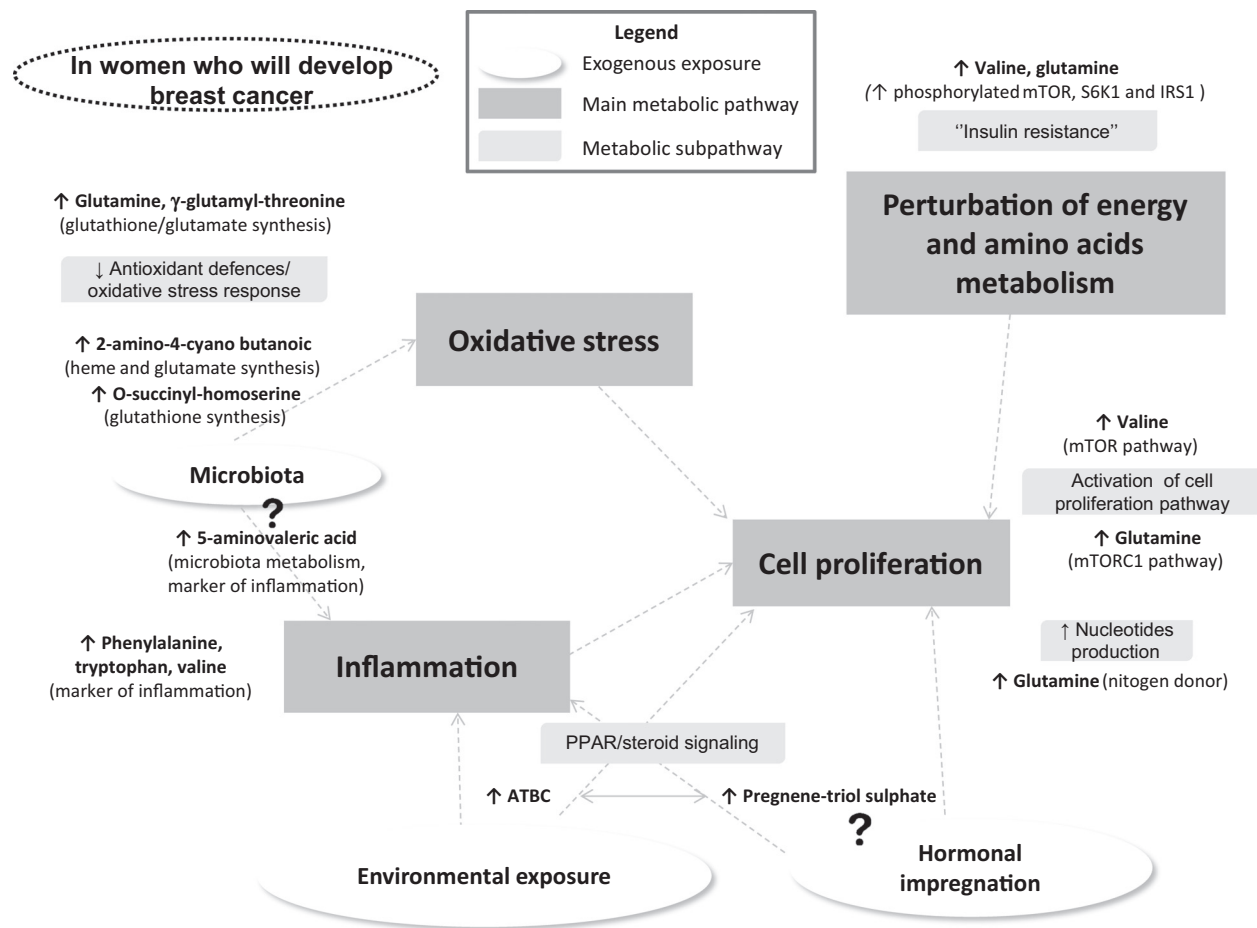
**Table 3.** Associations between the metabolomic patterns derived by PCA and breast cancer risk from multivariable conditional logistic regression<sup>a</sup>, SU.VI.MAX cohort, France (1994–2007)

PCA Factors <sup>b</sup>	Cases/controls	OR (95% CI)	P <sub>trend</sub>	Corrected
				P <sub>trend</sub> <sup>c</sup>
Factor 1	200/200	1.38 (1.11–1.73)	0.004	0.004
Factor 2	200/200	1.47 (1.16–1.85)	0.001	0.002
Factor 3	200/200	0.49 (0.37–0.65)	<0.0001	<0.0001
Factor 4	200/200	1.49 (1.17–1.89)	0.001	0.002

<sup>a</sup>Multivariable models were adjusted for age (continuous), BMI (continuous), smoking status (current smokers, former smokers, and nonsmokers), season of blood draw (*a priori*-defined periods: October–November/December–January–February/March–April–May), height (continuous), alcohol intake (continuous), physical activity (irregular/<1 h/d walking equivalent/≥1 h/d walking equivalent), education level (primary/secondary/superior), family history of breast cancer (yes/no), number of children (continuous), and use of hormone replacement therapy for menopause (yes/no; reference model). Tests for linear trend were performed using the continuous variables. ORs are computed for a 1 point increment of the continuous variable.

<sup>b</sup>The four first PCA factors (metabolomic patterns) explained more than 36% of total variability. The first factor was notably characterized by higher plasma levels of valine and 5-aminovaleric acid, the second factor was in particular characterized by higher plasma levels of phenylalanine, valine, 5-aminovaleric acid, and tryptophan, the third by lower plasma level of pregnene-triol sulfate, and the fourth by higher plasma levels of phenylalanine, valine 5-aminovaleric acid, GGT, and glutamine (loading values ≥|0.20|). Moreover, several unidentified ions also characterized these four factors.

<sup>c</sup>P<sub>trend</sub> after Benjamini Hochberg correction for multiple testing.



**Figure 1.**

Summary of proposed mechanistic interpretations for some putative pathways regarding observed metabolites variations in plasma of women who will develop breast cancer in subsequent years. Among the possible interpretations, our results could suggest a perturbation of energy and amino acids metabolism, a set-up of conditions favoring cell proliferation, inflammation, and oxidative stress in women who will develop breast cancer involving environmental exposure and microbiota metabolism.

cyano butanoic acid, was associated with increased breast cancer risk in this study before FDR correction only [OR = 1.28 (1.03–1.58);  $P = 0.02$ ; FDR = 0.3; data not tabulated). The variation of these two microbiota-related metabolites could be interpreted as an oxidative stress response given their potential role in glutamate and glutathione synthesis. Indeed, humans' metabolism can produce cysteine (one component of glutathione) using O-succinyl-homoserine instead of homocysteine (41). 2-amino-4-cyano butanoic acid is the substrate of nitrilase responsible for the *de novo* synthesis of glutamate in bacteria and fungus (42). Moreover it is also implied in heme synthesis, recognized as an oxidative biocompound (43). Likewise, the increase in plasmatic glutamine and in  $\gamma$ -glutamyl-threonine (positively correlated in our study) can suggest both a decrease in antioxidant defenses and an oxidative stress response ( $\gamma$ -glutamyl-dipeptides may indicate increased glutathione synthesis). The data available in this study do not allow to draw firm conclusions about these mechanistic interpretations, which need further investigations for their validation, such as complementary analyses of other metabolites variation involved in the same pathway (but not detected with the current analytic technic), fluxomic analyses to determine the

rates of metabolic reactions, or lipidomic and proteomic analyses to cover a wide range of biggest molecules.

The major strengths of this study pertained to the fully untargeted MS metabolomics approach and the prospective design. However, several limitations should be acknowledged. First, the relatively small sample size was probably sufficient for main analyses, but it was certainly limited for secondary stratified analyses. For instance, a larger sample size could have help to investigate more precisely the relation between the metabolites and the duration before cancer development (e.g., short, medium, and long term). Second, we cannot rule out the possibility of residual or unmeasured confounding despite the large spectrum of covariates taken into account, and we cannot exclude the possibility of overfitting of the models. Third, because the size of the sample did not allow keeping apart a validation set, we did not compute predictive performance such as AUC. However the main interest of our study lied in a better understanding of breast cancer etiology rather than in an improvement in discrimination models. Then, only LC/MS and not gas chromatography–mass spectroscopy was applied. This may have limited our ability to detect some categories of metabolites which could have

supported certain mechanistic interpretations. Moreover, as the metabolic profiles were established a decade before the diagnosis, some small variations may not have been detected. Some of the metabolites highlighted in our study are at the cross-roads of several metabolic pathways, thus these mechanistic interpretations resulting from literature should be considered with caution and need further investigations to be validated. Other interpretations could also be considered. Furthermore, this metabolomic analysis was based on a single blood draw so we could not investigate the stability of the profiles. Nevertheless, several studies showed a good stability and reproducibility of metabolomic measurements for most of metabolites (44, 45). Finally, as it is usually the case in a fully untargeted metabolomic approach, several associated ions could not be identified to date. However, these findings will be useful in the future in a context of strong development of metabolites databases and to compare our results with other future LC/MS studies.

In conclusion, this prospective study highlighted associations between baseline untargeted metabolomic profiles and subsequent long-term breast cancer risk. These findings generate mechanistic hypotheses regarding the etiology of this complex disease, which were consistent with those reported in our NMR analysis, and include a possible implication of gut microbiota activity and environmental contamination/exposure. If replicated in other independent prospective studies and after quantification of key metabolites to set discrimination thresholds, these results might contribute to develop screening strategies for the identification, before tumor set up, of women at higher risk of developing a breast cancer in the decade following the blood draw.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Disclaimer

The funders had no role in the design, analysis, or writing of this article.

## References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- Lindon JC, Nicholson JK. The emergent role of metabolic phenotyping in dynamic patient stratification. *Expert Opin Drug Metab Toxicol* 2014;10:915–9.
- Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wagele B, et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011;477:54–60.
- Norvig P, Relman DA, Goldstein DB, Kammen DM, Weinberger DR, Aiello LC, et al. 2020 visions. *Nature* 2010;463:26–32.
- Holmes E, Loo RL, Stampler J, Bictash M, Yap IK, Chan Q, et al. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 2008;453:396–400.
- Claudio WM, Quattrone A, Biganzoli L, Pestrin M, Bertini I, Di LA. Metabolomics: available results, current research projects in breast cancer, and future applications. *J Clin Oncol* 2007;25:2840–6.
- Budczies J, Denkert C, Muller BM, Brockmoller SF, Klauschen F, Györfy B, et al. Remodeling of central metabolism in invasive breast cancer compared to normal breast tissue - a GC-TOFMS based metabolomics study. *BMC Genomics* 2012;13:334.

## Authors' Contributions

**Conception and design:** L. Lécuyer, M. Touvier

**Development of methodology:** L. Lécuyer, E. Pujos-Guillot, M. Touvier

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** B. Lyan, M. Petera, D. Centeno, P. Galan, S. Hercberg, E. Kesse-Guyot, M. Touvier, M. Lagree

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** L. Lécuyer, C. Dalle, B. Lyan, A. Demidem, A. Rossary, M.-P. Vasson, M. Petera, D. Centeno, M. Deschasaux, V. Partula, B. Srour, P. Latino-Martel, N. Druésne-Pecollo, S. Durand, E. Pujos-Guillot, M. Touvier

**Writing, review, and/or revision of the manuscript:** L. Lécuyer, C. Dalle, B. Lyan, A. Demidem, A. Rossary, M.-P. Vasson, M. Petera, T. Ferreira, D. Centeno, P. Galan, S. Hercberg, M. Deschasaux, V. Partula, P. Latino-Martel, E. Kesse-Guyot, N. Druésne-Pecollo, S. Durand, E. Pujos-Guillot, M. Touvier

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** L. Lécuyer, E. Pujos-Guillot, M. Touvier

**Study supervision:** L. Lécuyer, A. Demidem, P. Galan, S. Hercberg, M. Touvier

## Acknowledgments

The authors thank Younes Esseddik, Frédéric Coffinieres, Thi Hong Van Duong, Paul Flanzky, Régis Gatibelza, Jagatjit Mohinder, and Maithyly Sivapalan (computer scientists), Rachida Mehroug and Frédérique Ferrat (logistic assistants), Nathalie Arnault, Véronique Gourlet, PhD, Fabien Szabo, PhD, Julien Allegre, and Laurent Bourhis (data-manager/statisticians), and Cédric Agaesse (dietitian) for their technical contribution to the SU.VI.MAX study. We also thank Nathalie Druésne-Pecollo, PhD (operational coordination), as well as all participants of the SU.VI.MAX study. The authors thank Dan Chaltiel for his help in R programming. This work was conducted in the framework of the French network for Nutrition And Cancer Research (NACRe network), [www.inra.fr/nacre](http://www.inra.fr/nacre) and received the NACRe Partnership Label. Metabolomic analysis was performed within the metaboHUB French infrastructure (ANR-INBS-0010). This work was supported by the French National Cancer Institute (grant no. INCa\_8085 for the project, and PhD grant no. INCa\_11323 for L. Lécuyer), the Federative Institute for Biomedical Research IFRB Paris 13, and the Cancéropole Ile-de-France/Région Ile de France.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 5, 2019; revised April 12, 2019; accepted May 28, 2019; published first June 4, 2019.

- Hennekes C, Bullinger D, Fux R, Friese N, Seeger H, Neubauer H, et al. Prediction of breast cancer by profiling of urinary RNA metabolites using Support Vector Machine-based feature selection. *BMC Cancer* 2009;9:104.
- Bro R, Kamstrup-Nielsen MH, Engelsen SB, Savorani F, Rasmussen MA, Hansen L, et al. Forecasting individual breast cancer risk using plasma metabolomics and biocontours. *Metabolomics* 2015;11:1376–80.
- Brauer HA, Makowski L, Hoadley KA, Casbas-Hernandez P, Lang LJ, Roman-Perez E, et al. Impact of tumor microenvironment and epithelial phenotypes on metabolism in breast cancer. *Clin Cancer Res* 2013;19:571–85.
- Playdon MC, Ziegler RG, Sampson JN, Stolzenberg-Solomon R, Thompson HJ, Irwin ML, et al. Nutritional metabolomics and breast cancer risk in a prospective study. *Am J Clin Nutr* 2017;106:637–49.
- Moore SC, Playdon MC, Sampson JN, Hoover RN, Trabert B, Matthews CE, et al. A metabolomics analysis of body mass index and postmenopausal breast cancer risk. *J Natl Cancer Inst* 2018;110:djx244.
- Lécuyer L, Victor Bala A, Deschasaux M, Bouchemal N, Nawfal Triba M, Vasson M-P, et al. NMR metabolomic signatures reveal predictive plasma metabolites associated with long-term risk of developing breast cancer. *Int J Epidemiol* 2018;47:484–94.
- Hercberg S, Galan P, Preziosi P, Bertrais S, Mennen L, Malvy D, et al. The SU.VI.MAX study: a randomized, placebo-controlled trial of the health effects

- of antioxidant vitamins and minerals. *Arch Intern Med* 2004;164:2335–42.
15. Hercberg S, Preziosi P, Briancon S, Galan P, Triol I, Malvy D, et al. A primary prevention trial using nutritional doses of antioxidant vitamins and minerals in cardiovascular diseases and cancers in a general population: the SU.VI.MAX study—design, methods, and participant characteristics. *Supplement en Vitamines et Mineraux Antioxydants. Control Clin Trials* 1998;19:336–51.
  16. World Health Organisation. ICD-10, International classification of diseases and related health problems. 10th Revision. Geneva, Switzerland: World Health Organization; 1993.
  17. Vandenbroucke JP, Pearce N. Case-control studies: basic concepts. *Int J Epidemiol* 2012;41:1480–9.
  18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995;57:289–300.
  19. Cattell RB. The scree test for the number of factors. *Multivar Behav Res* 1966;1:245–76.
  20. Ramautar R, Berger R, van der Greef J, Hankemeier T. Human metabolomics: strategies to understand biology. *Curr Opin Chem Biol* 2013;17:841–6.
  21. Assi N, Thomas DC, Leitzman M, Stepien M, Chajes V, Philip T, et al. Are metabolic signatures mediating the relationship between lifestyle factors and hepatocellular carcinoma risk? Results from a nested case-control study in EPIC. *Cancer Epidemiol Biomarkers Prev* 2018;27:531–40.
  22. Fages A, Duarte-Salles T, Stepien M, Ferrari P, Fedirko V, Pontoizeau C, et al. Metabolomic profiles of hepatocellular carcinoma in a European prospective cohort. *BMC Med* 2015;13:242.
  23. Stepien M, Duarte-Salles T, Fedirko V, Floegel A, Barupal DK, Rinaldi S, et al. Alteration of amino acid and biogenic amine metabolism in hepatobiliary cancers: findings from a prospective cohort study. *Int J Cancer* 2016;138:348–60.
  24. Assi N. A statistical framework to model the meeting-in-the-middle principle using metabolomic data: application to hepatocellular carcinoma in the EPIC study. *Mutagenesis* 2015;30:743–53.
  25. 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;20:1193–8.
  26. 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;10:1036–41.
  27. 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;115:1087–95.
  28. 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;137:2124–32.
  29. Magnusson M, Lewis GD, Ericson U, Orho-Melander M, Hedblad B, Engstrom G, et al. A diabetes-predictive amino acid score and future cardiovascular disease. *Eur Heart J* 2013;34:1982–9.
  30. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* 2009;9:311–26.
  31. O'Connell TM. The complex role of branched chain amino acids in diabetes and cancer. *Metabolites* 2013;3:931–45.
  32. Pischon T, Nimptsch K. Obesity and risk of cancer: an introductory overview. *Recent Results Cancer Res* 2016;208:1–15.
  33. Takeshita A, Igarashi-Migitaka J, Nishiyama K, Takahashi H, Takeuchi Y, Koibuchi N. Acetyl tributyl citrate, the most widely used phthalate substitute plasticizer, induces cytochrome p450 3a through steroid and xenobiotic receptor. *Toxicol Sci* 2011;123:460–70.
  34. Bui TT, Giovanoulis G, Cousins AP, Magnér J, Cousins IT, de Wit CA. Human exposure, hazard and risk of alternative plasticizers to phthalate esters. *Sci Total Environ* 2016;541:451–67.
  35. Rasmussen LM, Sen N, Vera JC, Liu X, Craig ZR. Effects of *in vitro* exposure to dibutyl phthalate, mono-butyl phthalate, and acetyl tributyl citrate on ovarian antral follicle growth and viability. *Biol Reprod* 2017;96:1105–17.
  36. Kambia N, Farce A, Belarbi K, Gressier B, Luyckx M, Chavatte P, et al. Docking study: PPARs interaction with the selected alternative plasticizers to di(2-ethylhexyl) phthalate. *J Enzyme Inhib Med Chem* 2016;31:448–55.
  37. Amathieu R, Nahon P, Triba M, Bouchemal N, Trinchet JC, Beaugrand M, et al. Metabolomic approach by 1H NMR spectroscopy of serum for the assessment of chronic liver failure in patients with cirrhosis. *J Proteome Res* 2011;10:3239–45.
  38. Zhenyukh O, Civantos E, Ruiz-Ortega M, Sánchez MS, Vázquez C, Peiró C, et al. High concentration of branched-chain amino acids promotes oxidative stress, inflammation and migration of human peripheral blood mononuclear cells via mTORC1 activation. *Free Radic Biol Med* 2017;104:165–77.
  39. Wang Q, Holmes MV, Davey Smith G, Ala-Korpela M. Genetic support for a causal role of insulin resistance on circulating branched-chain amino acids and inflammation. *Diabetes Care* 2017;40:1779–86.
  40. Lin H-M, Barnett MPG, Roy NC, Joyce NI, Zhu S, Armstrong K, et al. Metabolomic analysis identifies inflammatory and noninflammatory metabolic effects of genetic modification in a mouse model of Crohn's disease. *J Proteome Res* 2010;9:1965–75.
  41. Messerschmidt A, Worbs M, Steegborn C, Wahl MC, Huber R, Laber B, et al. Determinants of enzymatic specificity in the Cys-Met-metabolism PLP-dependent enzymes family: crystal structure of cystathionine gamma-lyase from yeast and intrafamilial structure comparison. *Biol Chem* 2003;384:373–86.
  42. Strobel GA. 4-Amino-4-cyanobutyric acid as an intermediate in glutamate biosynthesis. *J Biol Chem* 1967;244:3265–9.
  43. Jeney V, Balla J, Yachie A, Varga Z, Vercellotti GM, Eaton JW, et al. Pro-oxidant and cytotoxic effects of circulating heme. *Blood* 2002;100:879–87.
  44. Carayol M, Licaj I, Achaintre D, Sacerdote C, Vineis P, Key TJ, et al. Reliability of serum metabolites over a two-year period: a targeted metabolomic approach in fasting and non-fasting samples from EPIC. *PLoS One* 2015;10:e0135437.
  45. Floegel A, Drogan D, Wang-Sattler R, Prehn C, Illig T, Adamski J, et al. Reliability of serum metabolite concentrations over a 4-month period using a targeted metabolomic approach. *PLoS One* 2011;6:e21103.
  46. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 2007;3:211–21.