Supplementary Tables S4, S5 and S6 – **Plasma metabolomic signatures associated with long-term breast cancer risk in the SU.VI.MAX prospective cohort.**

Supplementary results of sensitivity, stratification and heterogeneity analyses

Results of sensitivity analysis after exclusion of cases diagnosed during the first year of follow-up are shown in **Supplementary Table S4**. All identified ions remained associated with breast cancer (FDR <0.05). No interaction was found between breast cancer risk and duration of follow-up **(Supplementary Table S5)**. **Supplementary Table S6** showed that higher plasma level of glutamine was associated with an increased risk of breast cancer risk only in premenopausal women (P-interaction=0.003, FDR=0.2).

The directions of all these associations were the same than the ones for the main analysis.

Supplementary Table S4 Association between identified ions and breast cancer risk from multivariable conditional logistic regression[a] after the exclusion of cases diagnosed during the first year of follow-up - SU.VI.MAX cohort, France (1994-2007)

Name (Level of confidence for identification[b])	Cases/ controls	OR [95% CI]	P-trend	Corrected P-trend[c]
L-valine/norvaline (1)	197/197	1.45 [1.13-1.85]	0.003	0.009
L-glutamine / L-isoglutamine (1)	197/197	1.29 [1.04-1.60]	0.02	0.02
5-aminovaleric acid (1)	197/197	1.40 [1.09-1.80]	0.009	0.01
L-phenylalanine (1)	197/197	1.43 [1.15-1.79]	0.002	0.007
L-tryptophan (1)	197/197	1.36 [1.06-1.74]	0.02	0.02
L-γ-glutamyl-L-threonine (1)	197/197	1.41 [1.09-1.83]	0.009	0.01
O-succinyl-L-homoserine (1)	189/189	0.72 [0.56-0.91]	0.007	0.01
ATBC (acetyl tributyl citrate) (1)	197/197	1.41 [1.10-1.80]	0.007	0.01
pregnene-triol sulfate (2)	189/189	1.38 [1.07-1.77]	0.01	0.02

[a] Multivariable models were adjusted for age (continuous), BMI (continuous), smoking status (current smokers, former smokers and non-smokers), season of blood draw (a priori-defined periods: October-November/December-January-February/March-April-May), height (continuous), alcohol intake (continuous), physical activity (irregular/< 1h/d walking equivalent/≥ 1h/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.

[b] Metabolites were classified accordingly to Sumner et al (1) 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).

[c] P-trend after Benjamini Hochberg correction for multiple testing.

Supplementary Table S5 Association between identified ions and breast cancer risk in the women with a follow-up duration < or ≥ median (5.8 years) from multivariable conditional logistic regression[a] - SU.VI.MAX Cohort, France (1994-2007)

Name (Level of confidence for identification[b])	Follow-up duration	Cases/ controls	OR [95% CI]	P-trend	Corrected P-trend[c]	P-interaction	Corrected P-interaction[c]
L-valine/norvaline (1)	< median	105/105	1.36 [0.97-1.91]	0.08	0.1	0.5	0.9
	≥ median	106/106	1.68 [1.15-2.45]	0.008	0.05	0.5	0.9
L-glutamine / L-isoglutamine (1)	< median	105/105	1.60 [1.15-2.24]	0.006	0.04	0.2	0.8
	≥ median	106/106	1.16 [0.83-1.63]	0.4	0.4	0.2	0.8
5-aminovaleric acid (1)	< median	105/105	1.89 [1.24-2.89]	0.003	0.04	0.2	0.8
	≥ median	106/106	1.30 [0.92-1.83]	0.1	0.2	0.2	0.8
L-phenylalanine (1)	< median	105/105	1.40 [1.00-1.96]	0.05	0.09	0.6	0.9
	≥ median	106/106	1.54 [1.10-2.16]	0.01	0.06	0.6	0.9
L-tryptophan (1)	< median	105/105	1.53 [1.06-2.22]	0.02	0.06	0.6	0.9
	≥ median	106/106	1.35 [0.96-1.90]	0.09	0.2	0.6	0.9
L-γ-glutamyl-L-threonine (1)	< median	105/105	1.43 [1.03-2.00]	0.03	0.07	0.5	0.9
	≥ median	106/106	1.54 [1.00-2.38]	0.05	0.1	0.5	0.9
O-succinyl-L-homoserine (1)	< median	102/102	0.61 [0.41-0.91]	0.02	0.06	0.7	0.9
	≥ median	101/101	0.66 [0.48-0.93]	0.02	0.06	0.7	0.9
ATBC (acetyl tributyl citrate) (1)	< median	105/105	1.36 [0.97-1.92]	0.07	0.1	1.0	1.0
	≥ median	106/106	1.46 [1.00-2.13]	0.05	0.1	1.0	1.0
pregnene-triol sulfate (2)	< median	102/102	1.62 [1.05-2.51]	0.03	0.07	0.5	0.9
	≥ median	101/101	1.24 [0.89-1.74]	0.2	0.3	0.5	0.9

[a] Multivariable models were adjusted for age (continuous), BMI (continuous), smoking status (current smokers, former smokers and non-smokers), season of blood draw (a priori-defined periods: October-November/December-January-February/March-April-May), height (continuous), alcohol intake (continuous), physical activity (irregular/< 1h/d walking equivalent/≥ 1h/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.

[b] Metabolites were classified accordingly to Sumner et al (1) 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).

[c] P after Benjamini Hochberg correction for multiple testing.

Supplementary Table S6 Association between identified ions and breast cancer risk in premenopausal or postmenopausal women (at the time of cancer diagnosis) from multivariable conditional logistic regression[a] - SU.VI.MAX cohort, France (1994-2007)

Name (Level of confidence for identification[b])	Menopausal status at cancer diagnosis	Cases/ controls	OR [95% CI]	P-trend	Corrected P-trend[c]	P-interaction	Corrected P-interaction[c]
L-valine/norvaline (1)	premenopausal	78/78	1.54 [1.00-2.38]	0.05	0.1	0.9	1.0
	postmenopausal	133/133	1.50 [1.10-2.05]	0.01	0.03	0.9	1.0
L-glutamine / L-isoglutamine (1)	premenopausal	78/78	2.49 [1.49-4.17]	0.0005	0.04	0.003	0.2
	postmenopausal	133/133	1.03 [0.78-1.35]	0.9	0.9	0.003	0.2
5-aminovaleric acid (1)	premenopausal	78/78	1.57 [1.01-2.45]	0.05	0.1	0.8	1.0
	postmenopausal	133/133	1.39 [1.03-1.87]	0.03	0.06	0.8	1.0
L-phenylalanine (1)	premenopausal	78/78	1.66 [1.10-2.49]	0.01	0.1	0.4	0.7
	postmenopausal	133/133	1.30 [0.98-1.71]	0.07	0.1	0.4	0.7
L-tryptophan (1)	premenopausal	78/78	1.67 [1.01-2.76]	0.04	0.1	0.3	0.7
	postmenopausal	133/133	1.32 [0.99-1.77]	0.06	0.1	0.3	0.7
L-γ-glutamyl-L-threonine (1)	premenopausal	78/78	1.45 [0.96-2.20]	0.08	0.2	0.9	1.0
	postmenopausal	133/133	1.25 [0.90-1.74]	0.2	0.2	0.9	1.0
O-succinyl-L-homoserine (1)	premenopausal	78/78	0.81 [0.52-1.27]	0.4	0.4	0.1	0.7
	postmenopausal	125/125	0.62 [0.44-0.86]	0.004	0.02	0.1	0.7
ATBC (acetyl tributyl citrate) (1)	premenopausal	78/78	1.28 [0.83-1.98]	0.3	0.3	0.5	0.9
	postmenopausal	133/133	1.48 [1.09-2.02]	0.01	0.03	0.5	0.9
pregnene-triol sulfate (2)	postmenopausal	133/133	1.27 [0.96-1.67]	0.03	0.1	0.3	0.7
	premenopausal	78/78	1.67 [1.06-2.60]	0.2	0.2	0.3	0.7
	postmenopausal	125/125	1.25 [0.92-1.69]	0.05	0.1	0.9	1.0

[a] Multivariable models were adjusted for age (continuous), BMI (continuous), smoking status (current smokers, former smokers and non-smokers), season of blood draw (a priori-defined periods: October-November/December-January-February/March-April-May), height (continuous), alcohol intake (continuous), physical activity (irregular/< 1h/d walking equivalent/≥ 1h/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.

[b] Metabolites were classified accordingly to Sumner et al (1) 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).

[c] P after Benjamini Hochberg correction for multiple testing.

REFERENCES

1. 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.