

ORIGINAL ARTICLE

Linking diet, physical activity, cardiorespiratory fitness and obesity to serum metabolite networks: findings from a population-based study

A Floegel¹, A Wientzek¹, U Bachlechner¹, S Jacobs², D Drogan¹, C Prehn³, J Adamski³, J Krumsiek⁴, MB Schulze², T Pischon⁵ and H Boeing¹

OBJECTIVE: It is not yet resolved how lifestyle factors and intermediate phenotypes interrelate with metabolic pathways. We aimed to investigate the associations between diet, physical activity, cardiorespiratory fitness and obesity with serum metabolite networks in a population-based study.

METHODS: The present study included 2380 participants of a randomly drawn subcohort of the European Prospective Investigation into Cancer and Nutrition-Potsdam. Targeted metabolomics was used to measure 127 serum metabolites. Additional data were available including anthropometric measurements, dietary assessment including intake of whole-grain bread, coffee and cake and cookies by food frequency questionnaire, and objectively measured physical activity energy expenditure and cardiorespiratory fitness in a subsample of 100 participants. In a data-driven approach, Gaussian graphical modeling was used to draw metabolite networks and depict relevant associations between exposures and serum metabolites. In addition, the relationship of different exposure metabolite networks was estimated.

RESULTS: In the serum metabolite network, the different metabolite classes could be separated. There was a big group of phospholipids and acylcarnitines, a group of amino acids and C6-sugar. Amino acids were particularly positively associated with cardiorespiratory fitness and physical activity. C6-sugar and acylcarnitines were positively associated with obesity and inversely with intake of whole-grain bread. Phospholipids showed opposite associations with obesity and coffee intake. Metabolite networks of coffee intake and obesity were strongly inversely correlated (body mass index (BMI): $r = -0.57$ and waist circumference: $r = -0.59$). A strong positive correlation was observed between metabolite networks of BMI and waist circumference ($r = 0.99$), as well as the metabolite networks of cake and cookie intake with cardiorespiratory fitness and intake of whole-grain bread ($r = 0.52$ and $r = 0.50$; respectively).

CONCLUSIONS: Lifestyle factors and phenotypes seem to interrelate in various metabolic pathways. A possible protective effect of coffee could be mediated via counterbalance of pathways of obesity involving hepatic phospholipids. Experimental studies should validate the biological mechanisms.

International Journal of Obesity (2014) 38, 1388–1396; doi:10.1038/ijo.2014.39

Keywords: metabolomics; physical activity; diet; network analysis; systems epidemiology

INTRODUCTION

Targeted metabolomics, which simultaneously studies many well-defined substrates and products of metabolism, is a powerful approach to detect metabolic alterations that are linked to human behavior such as diet and physical activity.^{1–3} Furthermore, metabolite levels in body fluids may also reflect consequent phenotypes including cardiorespiratory fitness and obesity.^{4–7} Therefore, targeted metabolomics may particularly help unravel biological mechanisms and identify pathways and networks through which behavioral factors and phenotypes are linked to metabolism. Knowledge about these mechanisms may offer great potential in terms of prevention of chronic diseases, which are strongly linked to human behavior and phenotypes.^{8,9} Consequently, adequate dietary and lifestyle recommendations could be generated.

Within the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study, a well-phenotyped cohort with comprehensive behavior assessment, we applied a targeted metabolomics platform that covers 127 serum metabolites, including amino acids, acylcarnitines, choline-containing phospholipids and hexose. In preceding analyses in the EPIC-Potsdam study, we have studied cross-sectional associations of diet, physical activity, cardiorespiratory fitness and obesity measures with these serum metabolites.^{10–13} In the present study, we attempted to provide a comprehensive picture of the associations of these human behaviors and consequent phenotypes with serum metabolite networks. To this end, we visualized these associations in our population-based cohort by applying an innovative network approach, and evaluated

¹Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; ²Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; ³Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; ⁴Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany and ⁵Molecular Epidemiology Group, Max Delbrück Center for Molecular Medicine (MDC) Berlin-Buch, Berlin-Buch, Germany. Correspondence: Dr A Floegel, Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, Nuthetal 14558, Germany.
E-mail: anna.floegel@dife.de

Received 6 January 2014; revised 17 February 2014; accepted 20 February 2014; accepted article preview online 10 March 2014; advance online publication, 1 April 2014

relationships between different lifestyle and phenotype networks.

SUBJECTS AND METHODS

Study population

EPIC-Potsdam is part of the multicenter EPIC-study, a prospective cohort study conducted in 23 European centers, with the scope of investigating diet and cancer risk that also includes other risk factors and chronic diseases. EPIC-Potsdam is the largest German study center and recruited 27 548 adults of the Potsdam area aged mainly 35–65 years between 1994 and 1998.¹⁴ The EPIC-Potsdam study was approved by the ethics committee of the Medical Society of the State of Brandenburg. Participants provided written informed consent before they participated in the baseline examination, which included anthropometric and blood pressure measurements, a blood sample collection as well as a personal interview on medical history, a sociodemographic and lifestyle questionnaire and a self-administered food frequency questionnaire; further details are given below.¹⁵ All measurements were conducted by qualified staff following standardized procedures.¹⁵ During the blood sample collection, 30 ml of venous blood were drawn, rapidly processed and fractionated into serum, plasma, 'buffy coat' (leukocytes) and erythrocytes, which were stored in straws at -196°C until metabolomic analysis.¹⁵

Of all the EPIC-Potsdam participants who had provided blood samples at baseline ($n=26\,444$), a subcohort of 2500 participants was randomly drawn for biomarker measurements.¹⁶ By randomly selecting the subcohort, it was assured that the results are representative of the full cohort. For the present analysis, the EPIC-Potsdam subcohort was used, and participants with implausible energy intakes (< 800 and > 6000 kcal per day) or missing information on serum metabolite concentrations and covariates were excluded ($n=120$). Baseline characteristics of the included participants have been published previously.¹⁰ In brief, participants were 61% women with mean age of 49.8 years and mean body mass index (BMI) of 26.1 kg m^{-2} . Of the participants, 4.5% had prevalent type 2 diabetes, and those were more likely to have increased hexose levels.

In 2007, a subgroup of 208 EPIC-Potsdam participants took part in a validation study for physical activity assessment, which also included repeated objective measurement of physical activity and cardiorespiratory fitness and a blood sample collection at two time points 4 months apart. Note that this subgroup was independent of the previously described subcohort. Blood was drawn from participants who had fasted overnight by qualified staff using monovette tubes with coagulation inhibitor. Serum was fractionated by centrifugation and stored in a freezer at -80°C until metabolomic analysis. Of the 208 participants, 50 men and 50 women with blood samples available were randomly selected for metabolomic measurements.

Dietary assessment

At baseline, diet was assessed by a self-administered semiquantitative food frequency questionnaire containing 148 items. It captured average food and beverage intake during the preceding 12 months with details on frequencies and portion sizes. The estimation of the portion sizes was facilitated by providing photos of standard portion sizes and household measures of different foods. In addition, fat content of dairy products, fat quality of fat spreads as well as preparation of certain foods were assessed. The average intake of each food item was calculated as the product of frequency and portion size and reported in g per day. Previously, the food frequency questionnaire was validated and reproducibility of the results was demonstrated.^{17–20}

Anthropometric measurements

During the baseline examination, anthropometric measurements were conducted in the study center by trained staff following standardized protocols.¹⁵ These included measurement of weight and height in light underwear and without shoes. BMI was calculated as the ratio of weight (kg) to squared height (m^2). Waist circumference in cm was measured at the midpoint between the lower ribs and the iliac crest.

Objective measurement of physical activity in a substudy

In the EPIC-Potsdam validation study for physical activity assessment instruments, objectively measured cardiorespiratory fitness and physical activity were measured. This study has been described in detail previously.²¹

In brief, cardiorespiratory fitness was assessed by an 8-min step test (200-mm step; Reebok, Lancaster, UK). It was obtained by extrapolating a regression line between an age-related maximum heart rate and workload and was reported as $\text{VO}_{2\text{max}}$ in $\text{ml kg}^{-1}\text{ min}^{-1}$.²² Physical activity was measured by a combined heart rate and movement sensor (Actiheart, CamNtech, Cambridge, UK) with a sampling frequency of 32 Hz.²³ The Actiheart was worn by the participants with two electrocardiography electrodes attached to the chest at two time points continuously for 4 days. Therefore, activity intensity ($\text{J min}^{-1}\text{ kg}^{-1}$) was estimated from heart rate and acceleration. Physical activity energy expenditure was calculated by summing the activity intensity time series and reported in kJ kg^{-1} per day. Finally, the mean value of cardiorespiratory fitness and physical activity energy expenditure from both time points was calculated, which was weighted by the test duration and probability of wear, respectively.

Measurement of serum metabolites

Metabolite concentrations were measured in 10 μl baseline serum samples of the EPIC-Potsdam subcohort and in serum samples of 100 participants of the validation study for physical activity assessment with the AbsoluteIDQ p150 kit (BIOCRATES, Innsbruck, Austria).^{24,25} The targeted metabolomics method simultaneously determined concentrations of 163 predefined metabolites including acylcarnitines (Cx:y), amino acids, hexose (sum of six-carbon monosaccharides, including glucose, but without distinction of isomers) and choline-containing phospholipids (lyso-, diacyl- and acyl-alkyl-phosphatidylcholines and sphingomyelins). For the lipid derivatives, fatty-acid side chains were abbreviated Cx:y, where x represented the number of carbon atoms and y the number of double bonds. For the phosphatidylcholines, only the total number of carbon atoms and double bonds across two fatty-acid side chains could be determined. Lyso-phosphatidylcholines, sphingomyelins and acylcarnitines contained a single fatty-acid side chain that was detected. All samples were analyzed at the Genome Analysis Center (Helmholtz Zentrum München) between 2009 and 2010. Sample preparation and metabolite quantification of these cohort samples has been described previously in full detail.^{26,27} In brief, an automated robotic system (Hamilton ML Star, Bonaduz, Switzerland) conducted the following procedure: 10 μl of serum was pipetted onto filters with stable isotope-labeled internal standards in 96-well plates and dried in nitrogen stream. Amino acids were derivatized with 5% phenylisothiocyanat reagent and the plates were dried again. The other metabolites and internal standards were extracted using 5 mm ammonium acetate in methanol, and then centrifuged and filtrated. Final extracts were diluted with mass spectrometry running solvent and analyzed using an API 4000 triple quadrupole mass spectrometer (AB Sciex, Darmstadt, Germany). Metabolites were quantified by multiple reaction monitoring in combination with internal standards, and metabolite concentrations were calculated with the Met/Q software package (BIOCRATES). The metabolomics method has been validated by the manufacturer according to the Food and Drug Administration guideline 'Guidance for industry—Bioanalytical Method Validation, May 2011', which implies proof of reproducibility in a certain error range. The limit of detection for the individual metabolites was set to three times the values of the buffer-only-containing samples.

The median coefficients of analytical variation were 7.3% within plate and 11.3% between plates for the EPIC-Potsdam samples.²⁶ The coefficients of variation, limit of detection and lower and upper quantification limits for the individual metabolites in this data set have been published previously²⁶ and are available online (<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0021103#s5>).

Those metabolites below the limit of detection and with high analytical variation (mainly hydroxyacylcarnitines) were excluded and, consequently, the final metabolite set comprised 127 metabolites (17 acylcarnitines, 14 amino acids, 1 hexose, 34 diacyl-phosphatidylcholines, 37 acyl-alkyl-phosphatidylcholines, 10 lyso-phosphatidylcholines and 14 sphingomyelins).

Statistical analysis

Within the subcohort, the metabolite network was calculated using Gaussian graphical modeling.²⁸ In brief, Gaussian graphical modeling is a data-driven method that uses the high degree of correlation between metabolites to construct metabolite networks. Each node in the network represents one metabolite and each edge between two nodes represents the dependency of two metabolites reflected by their partial correlation. Pearson's partial correlation coefficients between each possible pair of metabolites were calculated with adjustment for all the other metabolites,

to eliminate the indirect effects, and the covariates age and sex, in accordance with previous studies.^{28,29} As we used 127 metabolites, the metabolite network contained 127 nodes and could possibly contain $\binom{127}{2} = 8001$ edges. We chose a cutoff of >0.20 for the absolute value of the partial correlation coefficient between metabolites for the edges to be depicted. Therefore, the number of edges was reduced to 206. All calculations were done with SAS version 9.2 (SAS Institute Inc, Cary, NC, USA). Metabolite networks were then visualized with the yEd graph editor (yWorks GmbH, Tuebingen; <http://www.yworks.com>).

In the next step, we considered different exposures. To reflect dietary behavior, we selected three foods that were previously found to be linked to serum metabolites and chronic disease risk in this population.^{10,30} This included low intake of whole-grain bread, as an established risk factor, as well as low intakes of coffee and cake and cookies as less-established risk factors.³⁰ Physical activity was reflected by objectively measured physical activity energy expenditure. To mirror consequent phenotypes, we selected cardiorespiratory fitness and two obesity measures, BMI to reflect total body fat and waist circumference to reflect abdominal fat. The exposure information—namely, the respective exposure-metabolite association measures—was included as colors in the metabolite network. Therefore, we used the association measures of diet, physical activity, cardiorespiratory fitness and obesity with the single metabolites that we previously obtained.^{10–12} These were β -coefficient from linear mixed models regression analysis for physical activity and cardiorespiratory fitness, and Spearman's partial correlation coefficients for diet and obesity measures. Physical activity and cardiorespiratory fitness data were additionally z-transformed to make the β -coefficients directly comparable. All of the association measures were multivariable adjusted for age, sex, alcohol intake, smoking, education and, if applicable, prevalent diseases and the other exposures. Note that the adjustment models differed slightly for different exposures as they were based on our previous studies.^{10–12} We colored the network based on the direction and strength of association between each exposure and the metabolites, and thus ended up with one metabolite network for each exposure.

To compare the metabolite networks of different exposures, we additionally calculated the correlation of the association measures between the different exposures and 127 serum metabolites, so essentially 'the correlation of the correlation coefficients'. As an example to obtain the correlation of the metabolite networks of BMI and coffee, we calculated the correlation of the correlation coefficients between BMI and 127 serum metabolites with the correlation coefficients between coffee intake and 127 serum metabolites. As these were normally distributed for all exposures, we calculated Pearson's correlation coefficients. Therefore, we obtained the correlation of metabolite networks of different exposures.

RESULTS

On the basis of their partial correlations, the 127 serum metabolites measured in 2380 EPIC-Potsdam participants were depicted in a metabolite network (Figure 1). In the metabolite network, there was a big cluster of phospholipids and acylcarnitines, one cluster of amino acids and C6-sugar. Within the network of the phospholipids, subgroups of sphingomyelins, lyso-phosphatidylcholines, diacyl-phosphatidylcholines and acyl-alkyl-phosphatidylcholines with different chain length and fatty-acid desaturation emerged. The center of the network was dominated by diacyl-phosphatidylcholines and their hydrolysis products (lyso-phosphatidylcholines). These were also partly connected to acyl-alkyl-phosphatidylcholines, sphingomyelins and acylcarnitines. Amino acids and C6-sugar were not connected to the phospholipids. Seven individual metabolites were not connected to any other metabolite, namely, hexose, ornithine, proline and C9-acyl-carnitine, as well as diacyl-phosphatidylcholine C42:4, and acyl-alkyl-phosphatidylcholines C42:1 and C44:3. Free carnitine and propionyl-carnitine were linked as a pair. Although this network approach was purely data-driven, those phospholipids that are known to be only one reaction step apart were often strongly positively correlated, and thus, neighboring in the network, for example, sphingomyelins C18:0 and C18:1 (desaturase) and C16:1 and C18:1 (elongase) or amino acids glycine and serine (hydroxyl-methyl-transferase).

When looking at the association between diet, physical activity, cardiorespiratory fitness and obesity with the serum metabolite network, it was observed that the phenotypes (obesity and cardiorespiratory fitness) were more strongly linked to the metabolite network compared with the behavioral characteristics (diet and physical activity; Figures 2–4). The amino acids were particularly positively linked to cardiorespiratory fitness and physical activity, and differentially with the other exposures. C6-sugar was positively associated with BMI and waist circumference. It was inversely associated with cardiorespiratory fitness and physical activity, as well as intake of whole-grain bread and cake and cookies. Acylcarnitines were positively associated with obesity measures, and inversely associated with intake of whole-grain bread, coffee and cake and cookies. Sphingomyelins were particularly positively related to cardiorespiratory fitness and coffee intake, and inversely with the intake of whole-grain bread. Acyl-alkyl-phosphatidylcholines were positively linked to cardiorespiratory fitness and coffee intake, and inversely with obesity measures. Most lyso-phosphatidylcholines were positively associated with cardiorespiratory fitness, physical activity and coffee intake, and inversely with the intake of whole-grain bread and cake and cookies. Diacyl-phosphatidylcholines were positively associated with obesity measures and cardiorespiratory fitness, and inversely with intake of whole-grain bread, coffee and cake and cookies.

We additionally calculated the correlation of metabolite networks of different exposures (Table 1). The strongest positive correlation was observed between the metabolite networks of BMI and waist circumference ($r=0.99$). The metabolite networks of these two obesity measures were strongly inversely associated with the metabolite network of coffee intake (BMI: $r=-0.57$ and waist circumference: $r=-0.59$). The metabolite network of cardiorespiratory fitness was positively correlated with the metabolite networks of intake of cake and cookies ($r=0.52$) and physical activity ($r=0.46$). The metabolite network of intake of cake and cookies was positively correlated with the metabolite network of intake of whole-grain bread ($r=0.50$).

DISCUSSION

The present study used an innovative approach to generate multiple metabolite networks for different human behaviors and phenotypes, and compared these networks in the setting of a population study. Although the generation of the metabolite network was purely data driven, the metabolites were separated by their different biochemical classes. This was particularly achieved for amino acids, C6-sugar and acylcarnitines. The cluster of different phospholipids could not be completely differentiated, for example, there were still some direct connections between sphingomyelins and phosphatidylcholines. Biosynthesis of sphingomyelins includes the transfer of phosphocholine from phosphatidylcholines to ceramide,³¹ so there is a biological plausible link of these two groups. However, the direct link of these two different classes could also be a result of the limitations of the metabolomics measurements. The manufacturer of the metabolomics kit reported about possible interferences between individual metabolites.³² This could, for example, have affected hydroxy-sphingomyelin C14:1 and acyl-alkyl-phosphatidylcholine C30:1, two metabolites that were inversely correlated and neighboring in our metabolite network. Nevertheless, it was often the case that metabolites that are one reaction step apart in biological pathways were direct neighbors in our metabolite network. Krumsiek *et al.*²⁸ previously showed in the Cooperative Health Research in the Region of Augsburg (KORA) cohort that the statistical method Gaussian graphical modeling could discriminate a pathway distance of one for most of these phospholipids with sensitivity and specificity ranging between 0.75 and 1.0. This previous study also showed that a pathway distance of two is indicated by strong negative correlations. This was also observed

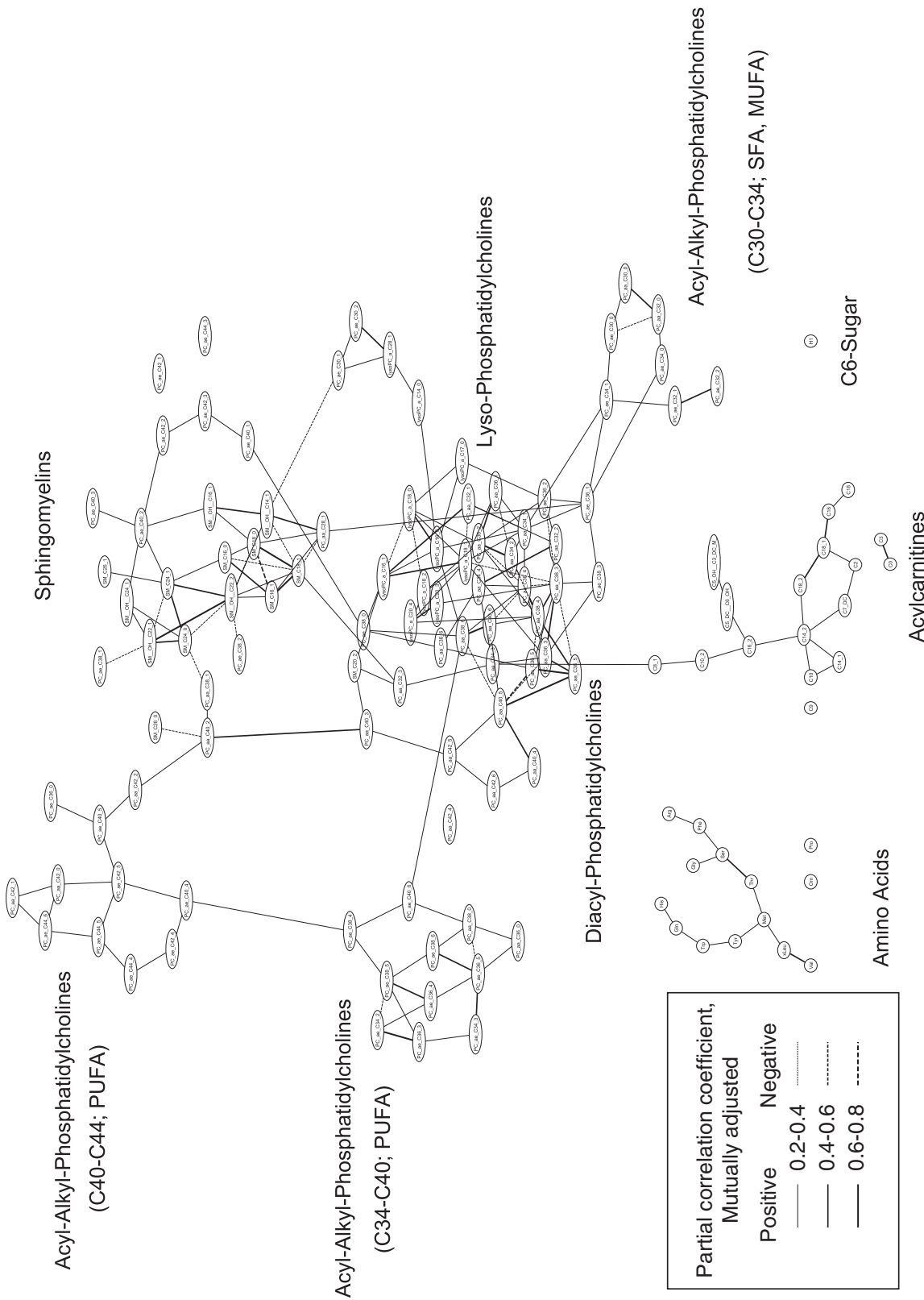


Figure 1. Serum metabolite network of the EPIC-Potsdam subcohort ($n = 2380$). Each node represents one metabolite and each edge between two nodes represents the partial correlation between two metabolites mutually adjusted for the other metabolites. Solid line represents positive correlation and dotted line negative correlation. The thickness of the line indicates the strength of the correlation.

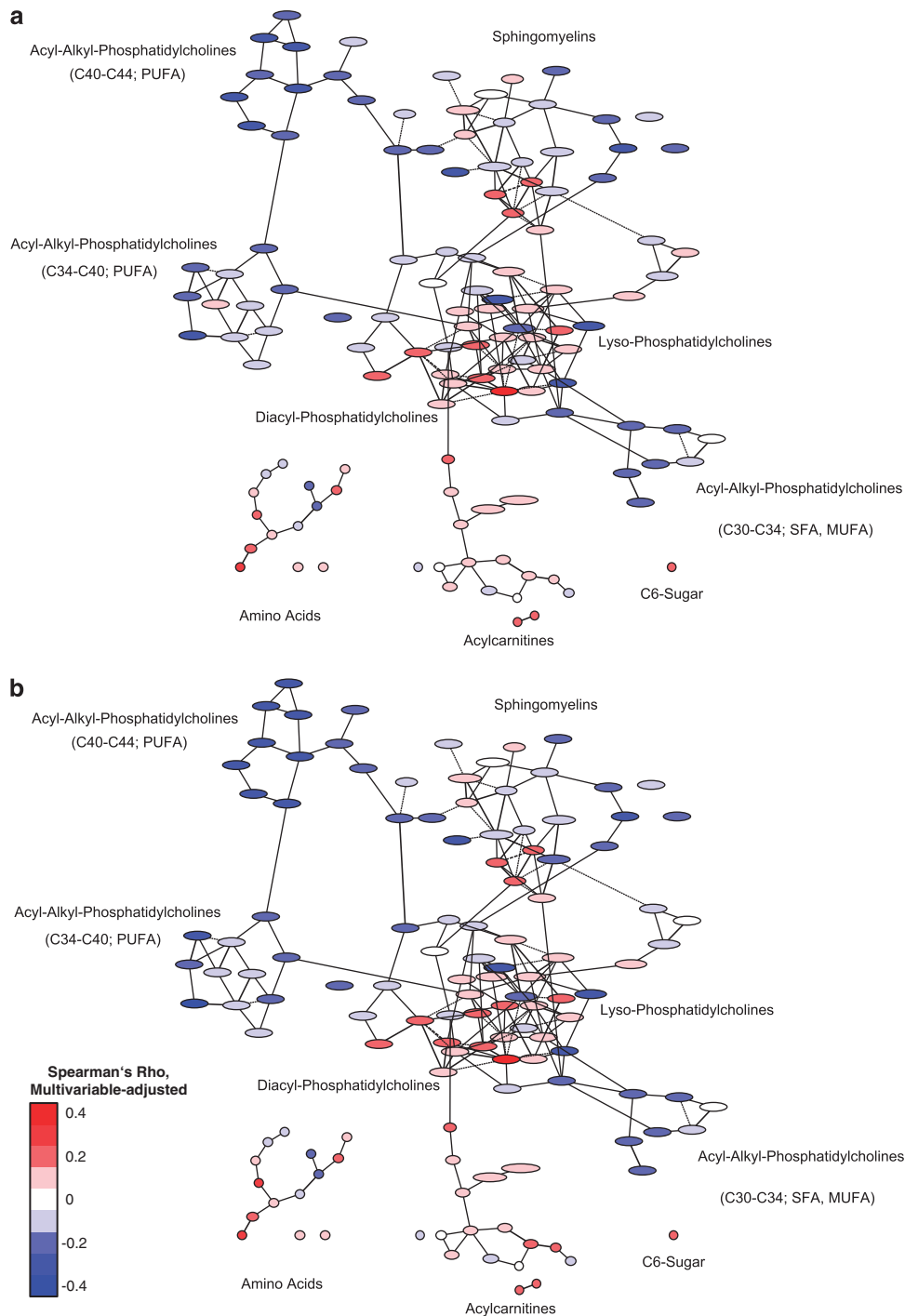


Figure 2. Association between BMI (**a**) and waist circumference (**b**) and the serum metabolite network of the EPIC-Potsdam subcohort. Presented are partial correlation coefficients adjusted for age, sex, education, alcohol consumption, smoking and physical activity, adapted from Bachlechner *et al.*¹² Red color implies positive association and blue color inverse association between exposure and metabolite. Intensity of the color reflects the strength of association.

in our study, for example, sphingomyelins C16:1 and C18:0 (desaturation and elongation), and diacyl-phosphatidylcholines C36:5 and C40:5 (two elongations). In summary, the metabolite network of our study population that was drawn with the data-driven method Gaussian graphical modeling could differentiate the metabolite classes and seemed to generally reflect biological plausible pathways.

In our study, we observed several associations between lifestyle factors and phenotypes with serum metabolite networks. As our

study is of an observational design, we cannot prove biological mechanisms or demonstrate tissue origins of systemic metabolite changes. Nevertheless, our study may give some first ideas about possible mechanisms and pathways, which are discussed in the following paragraphs, but should be cautiously interpreted and confirmed in experimental studies.

We observed that obesity and cardiorespiratory fitness were more strongly linked to serum metabolite networks compared with diet and physical activity. Our results are in agreement with a

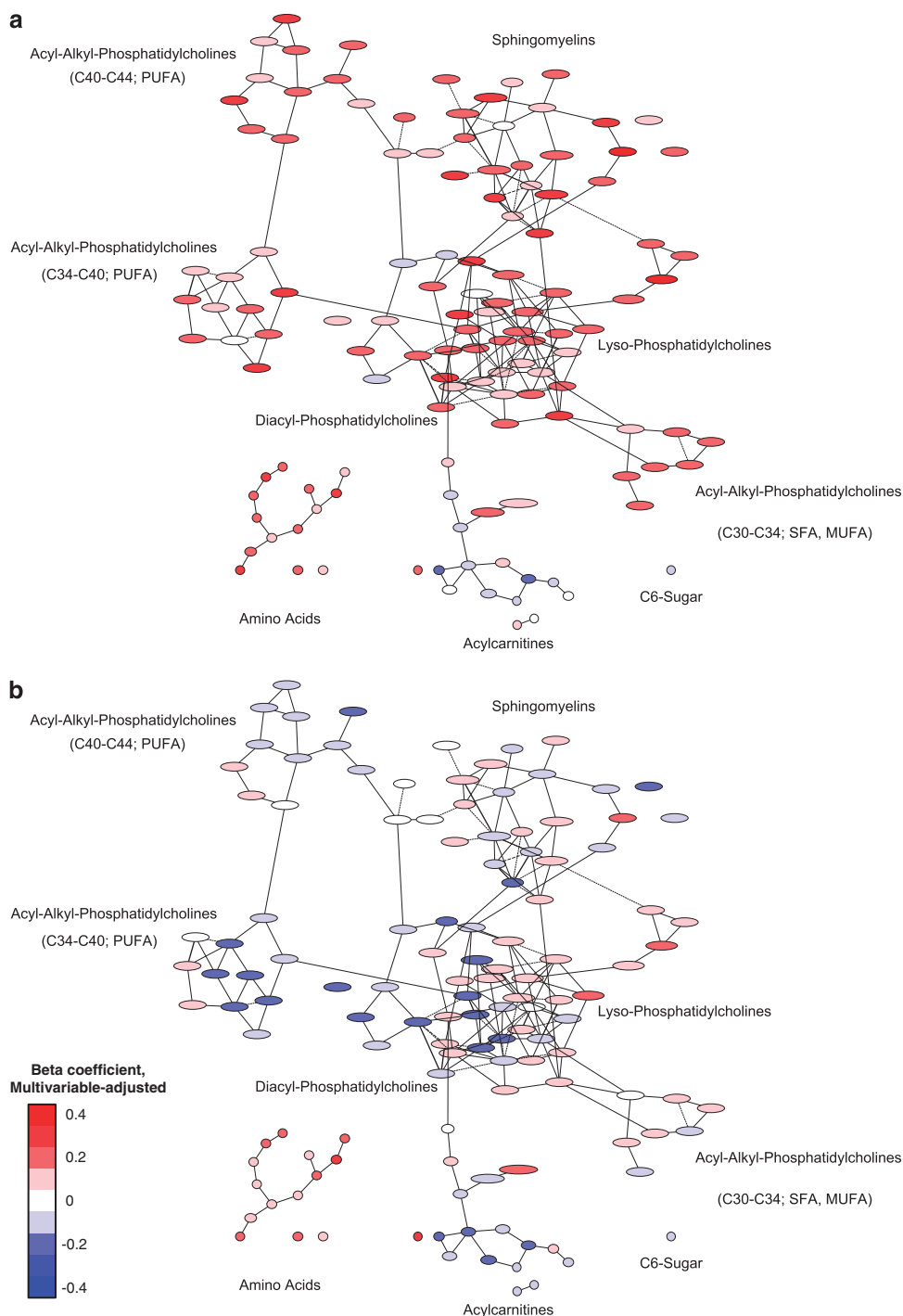


Figure 3. Association between cardiorespiratory fitness (**a**) and physical activity energy expenditure (**b**), and the serum metabolite network of the EPIC-Potsdam subcohort. Presented are β -coefficients adjusted for age, sex, education, alcohol consumption, smoking, BMI, waist circumference and measurement occasion, adopted from Wientzek *et al.*¹¹ Red color implies positive association and blue color inverse association between exposure and metabolite. Intensity of the color reflects the strength of association.

previous study,⁴ which reported that low cardiorespiratory fitness and high levels of visceral adipose tissue were strong predictors for impaired cardiometabolic health. This included altered lipoprotein-lipid profiles, deranged glucose-insulin homeostasis and increased inflammatory markers. However, our results have to be interpreted with caution as the different exposures were assessed with different methods. Particularly, diet was not measured objectively but self-reported in a food frequency questionnaire. Thus, it cannot be ruled out that imprecision of

the dietary assessment may lead to underestimation of the true effect of diet. In contrast, both cardiorespiratory fitness and physical activity were measured objectively with heart rate and movement sensors and may be better comparable. Therefore, our results indicate that cardiorespiratory fitness may be stronger linked to metabolite networks than physical activity. Physical activity represents a lifestyle, whereas cardiorespiratory fitness can be considered a phenotype, which may not only be influenced by physical activity or inactivity but also by past physical activity,

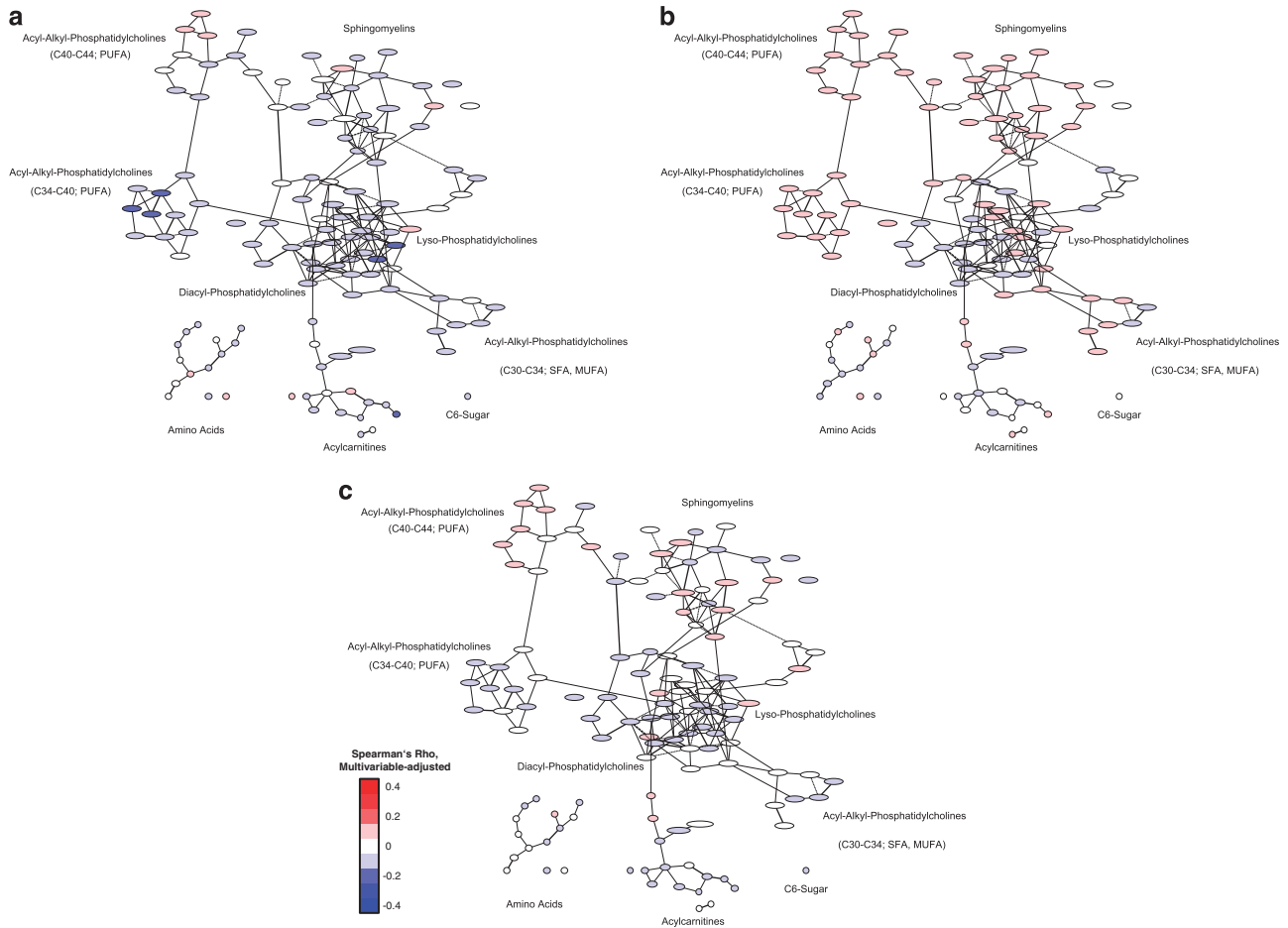


Figure 4. Association between intake of whole-grain bread (**a**), coffee (**b**), cake and cookies (**c**), and the serum metabolite network of the EPIC-Potsdam subcohort. Presented are partial correlation coefficients adjusted for age, sex, education, alcohol consumption, smoking, physical activity, BMI, waist circumference, prevalent hypertension and prevalent diabetes. Red color implies positive association and blue color inverse association between exposure and metabolite. Intensity of the color reflects the strength of association.

Table 1. Correlation of metabolite networks of different exposures

| | Fitness network | Physical activity network | Whole-grain bread network | Coffee network | Cake & cookies network | BMI network | Waist circumference network |
|-----------------------------|-----------------|---------------------------|---------------------------|---------------------|------------------------|----------------------|-----------------------------|
| Fitness network | 1.00 | 0.46 (0.31; 0.59) | 0.28 (0.11; 0.43) | 0.02 (-0.15; 0.20) | 0.52 (0.38; 0.63) | -0.29 (-0.43; -0.12) | -0.28 (-0.43; -0.11) |
| Physical activity network | | 1.00 | 0.36 (0.20; 0.50) | -0.02 (-0.20; 0.15) | 0.31 (0.14; 0.46) | -0.14 (-0.31; 0.03) | -0.11 (-0.28; 0.06) |
| Whole-grain bread network | | | 1.00 | -0.09 (-0.26; 0.08) | 0.50 (0.35; 0.62) | -0.23 (-0.39; -0.06) | -0.21 (-0.37; -0.04) |
| Coffee network | | | | 1.00 | 0.15 (-0.03; 0.32) | -0.57 (-0.67; -0.44) | -0.59 (-0.69; -0.46) |
| Cake & cookies network | | | | | 1.00 | -0.24 (-0.39; -0.06) | -0.24 (-0.39; -0.07) |
| BMI network | | | | | | 1.00 | 0.99 (0.99; 1.00) |
| Waist circumference network | | | | | | | 1.00 |

Abbreviation: BMI, body mass index. Presented are Pearson's correlation coefficients (95% confidence interval).

other environmental and lifestyle factors and genetics.³³ Therefore, cardiorespiratory fitness may reflect long-term exposure and be stronger linked to serum metabolites than physical activity.

On the level of metabolite classes, we observed that amino acids were particularly positively linked to cardiorespiratory fitness and physical activity. Morris *et al.*⁵ found that amino-acid excretion was reduced in participants with high fitness as indicated by lower urinary levels of amino acids. In contrast, in this previous study also plasma levels of many amino acids were lower in the high fitness group. With high energy expenditure, oxidation of branched chain amino acids is generally increased.³⁴ However, as response to exercise, amino-acid biosynthesis as well as protein

breakdown in skeletal muscle may increase.^{35,36} This could possibly increase the systemic pool of amino acids.

In the present study, C6-sugar was particularly positively linked to obesity measures. This may be a result of impaired glucose metabolism and insulin resistance in obese individuals. It has previously been reported that obesity correlated positively with levels of fasting plasma glucose.³⁷ A recent intervention study could show that people with abdominal obesity were able to improve their levels of fasting plasma glucose with a combined physical activity and dietary intervention.³⁸ This is in line with our findings that C6-sugar was further inversely correlated with cardiorespiratory fitness, physical activity and whole-grain bread

intake. In our study population, we previously found that increased levels of C6-sugars were linked to a higher risk of type 2 diabetes.³⁹ Thus, via this pathway, lifestyle modifications may offer potential for prevention of chronic diseases.

Acylcarnitines showed similar correlations with lifestyle factors and phenotypes as C6-sugar. Adams *et al.*⁴⁰ reported increased levels of acylcarnitines in African Americans with type 2 diabetes, which may be a result of incomplete fatty-acid oxidation and via pro-inflammatory pathways trigger development of insulin resistance. In rats, fed with a 'cafeteria diet' with high-fat and low-fiber content, a serum accumulation of acylcarnitines was observed reflective of systemic mitochondrial dysfunction.⁴¹ This eventually led to tissue inflammation and development of obesity and metabolic syndrome. Thus, increased serum acylcarnitines may be reflective of inflammatory processes and oxidative stress involved in obesity.

Among the phospholipids, sphingomyelins, lyso- and acyl-alkyl-phosphatidylcholines were positively associated with coffee intake, and the latter class was inversely associated with obesity. It was previously observed that levels of acyl-alkyl-phosphatidylcholines were reduced in obese individuals.⁴² In addition, these phospholipid classes have been linked to lower risk of type 2 diabetes in this study population, whereas diacyl-phosphatidylcholines were linked to higher risk.³⁹ In the present study, diacyl-phosphatidylcholines were positively associated with obesity and inversely with coffee intake. All of these phospholipids are primarily synthesized in the liver and secreted as part of blood lipoproteins.⁴³ In a previous study, we could show that acyl-alkyl-phosphatidylcholines were positively correlated with high-density lipoprotein cholesterol, whereas diacyl-phosphatidylcholines were positively correlated with plasma triglycerides.³⁹ Coffee intake has been associated with decreased risk of chronic diseases in prospective cohort studies, and particularly with lower risk of type 2 diabetes.^{30,44} However, the biological mechanisms are not yet understood. In the present study, it was additionally observed that the metabolite network of coffee was strongly inversely correlated with the metabolite network of obesity. Thus, our results indicate that the suggested protective effect of coffee could be mediated via counterbalance of pathways of obesity, in particular, on the level of phospholipids that are supplied from hepatic synthesis. Thus, our findings may point toward a crucial role of the liver. These assumptions may be supported by the fact that higher coffee intake has also been associated with lower risk of fibrosis in patients with non-alcoholic fatty liver disease.⁴⁵ In addition, in an experimental study, coffee prevented induced liver cirrhosis in rats.⁴⁶ Finally, coffee consumption has also been associated with lower risk of liver cancer in prospective cohort studies.⁴⁷

The metabolite network of cake and cookies was positively associated with the metabolite networks of cardiorespiratory fitness and intake of whole-grain bread. This finding was rather unexpected. Nevertheless, it is in line with the observation that cake and cookie intake was inversely associated with risk of chronic diseases in this study population.³⁰ It is still to debate and see in other studies whether this observation has a biological background or is the result of some type of reporting bias, that is, that people with lower levels of cardiorespiratory fitness and low intake of whole-grain bread tend to under-report their cake and cookie intake compared with their counterparts. At this point, we can only conclude that the observation is interesting and that this issue should be investigated further.

The strengths of the present study include that we provided a comprehensive overview of the associations between phenotypes and lifestyle characteristics with serum metabolite networks. We conducted metabolomics analyses in a large population-based cohort measuring 127 metabolites in more than 2300 participants. In addition, we integrated data on objectively measured physical activity and cardiorespiratory fitness, obesity measures and questionnaire-based dietary assessment with metabolomics data using an innovative network approach. Study limitations include

that we used a commercial metabolomics kit and could only investigate those metabolites that were included by the manufacturer. Therefore, we might have missed associations of other important metabolites, for example, other amino acids or lipid derivatives. It would be very interesting to study a broader range of metabolites with different targeted or untargeted metabolomics platforms in future studies. Furthermore, our main focus in the discussion was on the level of metabolite classes. Nevertheless, in some conditions, metabolite changes were different for individual metabolites within the same class, which is to be expected, as individual metabolites, for example, different amino acids, may carry very distinct metabolic functions. Another limitation of the study is that the assessment instruments of exposures were different. Therefore, they may be differentially prone to measurement error and, consequently, association measures between different exposures and serum metabolites may not be directly comparable. In addition, we used a cross-sectional design, which illustrates associations but not necessarily temporality and causality. However, knowledge is scarce about the association of these lifestyle factors and phenotypes with metabolite networks in a population setting.

In summary, we could visualize a metabolite network that reflected biological pathways with a data-driven approach based on 2380 EPIC-Potsdam participants. Our results support the concept that different lifestyle factors and phenotypes are reflected in metabolic pathways. Amino-acid metabolism seems to be particularly linked to cardiorespiratory fitness and physical activity, whereas levels of acylcarnitines and C6-sugar were linked to obesity and low intake of whole-grain bread. In addition, the metabolite networks of coffee and obesity were strongly inversely correlated. Therefore, a possible protective effect of coffee on chronic disease risk may be particularly mediated by hepatic phospholipids. Experimental studies need to further validate these findings.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We specially thank Ellen Kohlsdorf and Wolfgang Bernigau for data management and Martin Floegel for his support with figure formatting. We also thank Julia Scarpa, Arsin Sabunchi and Werner Römisch-Margl for metabolomics measurements performed at the Helmholtz Centrum München, Genome Analysis Center, Metabolomics Core Facility. We thank all the EPIC-Potsdam study participants for their devoted participation in the study. This study was supported by the Federal Ministry of Science, Germany (Grant number 01 EA 9401), the European Union (Grant number SOC 95 201408 05F02) and the German Cancer Aid (Grant number 70-2201-Bo2), and by a grant from the Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD eV).

REFERENCES

- 1 Gibney MJ, Walsh M, Brennan L, Roche HM, German B, van Ommen B. Metabolomics in human nutrition: opportunities and challenges. *Am J Clin Nutr* 2005; **82**: 497–503.
- 2 Primrose S, Draper J, Elsom R, Kirkpatrick V, Mathers JC, Seal C *et al*. Metabolomics and human nutrition. *Br J Nutr* 2011; **105**: 1277–1283.
- 3 Krug S, Kastenmuller G, Stuckler F, Rist MJ, Skurk T, Sailer M *et al*. The dynamic range of the human metabolome revealed by challenges. *FASEB J* 2012; **26**: 2607–2619.
- 4 Rheaume C, Arsenault BJ, Dumas MP, Perusse L, Tremblay A, Bouchard C *et al*. Contributions of cardiorespiratory fitness and visceral adiposity to six-year changes in cardiometabolic risk markers in apparently healthy men and women. *J Clin Endocrinol Metab* 2011; **96**: 1462–1468.
- 5 Morris C, Grada CO, Ryan M, Roche HM, De Vito G, Gibney MJ *et al*. The relationship between aerobic fitness level and metabolic profiles in healthy adults. *Mol Nutr Food Res* 2013; **57**: 1246–1254.

- 6 Bye A, Vettukattil R, Aspenes ST, Giskeodegard GF, Gribbestad IS, Wisloff U et al. Serum levels of choline-containing compounds are associated with aerobic fitness level: the HUNT-study. *PLoS One* 2012; **7**: e42330.
- 7 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**: 311–326.
- 8 Schulze MB, Hoffmann K, Boeing H, Linseisen J, Rohrmann S, Mohlig M et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. *Diabetes Care* 2007; **30**: 510–515.
- 9 Ford ES, Bergmann MM, Kroger J, Schienkiewitz A, Weikert C, Boeing H. Healthy living is the best revenge: findings from the European Prospective Investigation Into Cancer and Nutrition-Potsdam study. *Arch Intern Med* 2009; **169**: 1355–1362.
- 10 Floegel A, von Ruesten A, Drogan D, Schulze MB, Prehn C, Adamski J et al. Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. *Eur J Clin Nutr* 2013; **67**: 1100–1108.
- 11 Wientzek A, Floegel A, Knüppel S, Vigil M, Drogan D, Adamski J et al. Serum metabolites related to cardiorespiratory fitness, physical activity energy expenditure, sedentary time and vigorous activity. *Int J Sport Nutr Exerc Metab* 2013; e-pub ahead of print 13 November 2013.
- 12 Bachlechner U, Floegel A, Prehn C, Adamski J, Pischon T, Boeing H. Associations between anthropometric parameters and serum metabolites using a targeted metabolomics approach. *Obesity* 2014 (submitted).
- 13 Jacobs S, Kröger J, Floegel A, Boeing H, Drogan D, Pischon T et al. Evaluation of various biomarkers as potential mediators of the association between coffee consumption and incident type 2 diabetes in the EPIC-Potsdam Study. *Am J Clin Nutr* 2014 (submitted).
- 14 Boeing H, Korfmann A, Bergmann MM. Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. *Ann Nutr Metab* 1999; **43**: 205–215.
- 15 Boeing H, Wahrendorf J, Becker N. EPIC-Germany: a source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. *Ann Nutr Metab* 1999; **43**: 195–204.
- 16 Ford ES, Schulze MB, Bergmann MM, Thamer C, Joost HG, Boeing H. Liver enzymes and incident diabetes: findings from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. *Diabetes Care* 2008; **31**: 1138–1143.
- 17 Kroke A, Klipstein-Grobusch K, Voss S, Moseneder J, Thielecke F, Noack R et al. Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* 1999; **70**: 439–447.
- 18 Boeing H, Bohlscheid-Thomas S, Voss S, Schneeweiss S, Wahrendorf J. The relative validity of vitamin intakes derived from a food frequency questionnaire compared to 24-hour recalls and biological measurements: results from the EPIC pilot study in Germany. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* 1997; **26**: S82–S90.
- 19 Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of energy and macronutrient intake of a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* 1997; **26**: S71–S81.
- 20 Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* 1997; **26**: S59–S70.
- 21 Wientzek A, Tormo Diaz MJ, Castaño JM, Amiano P, Arriola L, Overvad K et al. Cross-sectional associations of objectively measured physical activity, cardiorespiratory fitness and anthropometry in European adults. *Obesity* (Silver Spring) 2013; e-pub ahead of print 26 June 2013; doi:10.1002/oby.20530.
- 22 Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol* 2001; **37**: 153–156.
- 23 Brage S, Brage N, Franks PW, Ekelund U, Wareham NJ. Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur J Clin Nutr* 2005; **59**: 561–570.
- 24 Romisch-Margl W, Prehn C, Bogumil R, Rohring C, Suhre K, Adamski J. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. *Metabolomics* 2012; **8**: 133–142.
- 25 Illig T, Gieger C, Zhai G, Romisch-Margl W, Wang-Sattler R, Prehn C et al. A genome-wide perspective of genetic variation in human metabolism. *Nat Genet* 2010; **42**: 137–141.
- 26 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.
- 27 Floegel A, Stefan N, Yu Z, Mühlenbruch K, Drogan D, Joost HG et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 2012; **62**: 639–648.
- 28 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.
- 29 Krumsiek J, Suhre K, Evans AM, Mitchell MW, Mohney RP, Millburn MV et al. Mining the unknown: a systems approach to metabolite identification combining genetic and metabolic information. *PLoS Genet* 2012; **8**: e1003005.
- 30 von Ruesten A, Feller S, Bergmann MM, Boeing H. Diet and risk of chronic diseases: results from the first 8 years of follow-up in the EPIC-Potsdam study. *Eur J Clin Nutr* 2013; **67**: 412–419.
- 31 Gault CR, Obeid LM, Hannun YA. An overview of sphingolipid metabolism: from synthesis to breakdown. *Adv Exp Med Biol* 2010; **688**: 1–23.
- 32 AbsoluteIDQTM Kit. *Analytical Specifications p150*. BIOCRATES Life Sciences: Innsbruck, Austria, 2006.
- 33 Bouchard C, Rankinen T. Individual differences in response to regular physical activity. *Med Sci Sports Exerc* 2001; **33**: S446–S451.
- 34 Shimomura Y, Murakami T, Nakai N, Nagasaki M, Harris RA. Exercise promotes BCAA catabolism: effects of BCAA supplementation on skeletal muscle during exercise. *J Nutr* 2004; **134**: 1583S–1587SS.
- 35 Rennie MJ, Tipton KD. Protein and amino acid metabolism during and after exercise and the effects of nutrition. *Annu Rev Nutr* 2000; **20**: 457–483.
- 36 Van Hall G, Saltin B, Wagenmakers AJ. Muscle protein degradation and amino acid metabolism during prolonged knee-extensor exercise in humans. *Clin Sci* 1999; **97**: 557–567.
- 37 McAdams MA, Van Dam RM, Hu FB. Comparison of self-reported and measured BMI as correlates of disease markers in US adults. *Obesity* 2007; **15**: 188–196.
- 38 Soon HK, Saad HA, Taib MN, Rahman HA, Mun CY. Effects of combined physical activity and dietary intervention on obesity and metabolic parameters in adults with abdominal obesity. *Southeast Asian J Trop Med Public Health* 2013; **44**: 295–308.
- 39 Floegel A, Stefan N, Yu Z, Mühlenbruch K, Drogan D, Joost HG et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 2013; **62**: 639–648.
- 40 Adams SH, Hoppel CL, Lok KH, Zhao L, Wong SW, Minkler PE et al. Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. *J Nutr* 2009; **139**: 1073–1081.
- 41 Sampey BP, Freemerman AJ, Zhang J, Kuan PF, Galanko JA, O'Connell TM et al. Metabolomic profiling reveals mitochondrial-derived lipid biomarkers that drive obesity-associated inflammation. *PLoS One* 2012; **7**: e38812.
- 42 Wallner S, Schmitz G. Plasmalogens the neglected regulatory and scavenging lipid species. *Chem Phys Lipids* 2011; **164**: 573–589.
- 43 Cole LK, Vance JE, Vance DE. Phosphatidylcholine biosynthesis and lipoprotein metabolism. *Biochim Biophys Acta* 2011; **1821**: 754–761.
- 44 Huxley R, Lee CM, Barzi F, Timmermeister L, Czernichow S, Perkovic V et al. Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: a systematic review with meta-analysis. *Arch Intern Med* 2009; **169**: 2053–2063.
- 45 Molloy JW, Calcagno CJ, Williams CD, Jones FJ, Torres DM, Harrison SA. Association of coffee and caffeine consumption with fatty liver disease, nonalcoholic steatohepatitis, and degree of hepatic fibrosis. *Hepatology* 2012; **55**: 429–436.
- 46 Moreno MG, Chavez E, Aldaba-Muruato LR, Segovia J, Vergara P, Tsutsumi V et al. Coffee prevents CCl(4)-induced liver cirrhosis in the rat. *Hepatology* 2011; **5**: 857–863.
- 47 Bravi F, Bosetti C, Tavani A, Gallus S, La Vecchia C. Coffee reduces risk for hepatocellular carcinoma: an updated meta-analysis. *Clin Gastroenterol Hepatol* 2013; **11**: 1413–1421.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>