

Research Article

Open Access

Sex Dependency of Human Metabolic Profiles Revisited

Sabine Siebert^{1*}, Zhonghao Yu², Rui Wang-Sattler², Thomas Illig^{2,3}, Jerzy Adamski⁴, Stefan Schreiber^{5,6,7}, Michael Krawczak^{7,8}, Michael Nothnagel⁹ and Ute Nöthlings^{1,7,9}

¹Institute of Experimental Medicine, Section of Epidemiology, Christian-Albrechts University Kiel, Kiel, 24105, Germany

²Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, 85764, Germany

³Hannover Unified Biobank, Hannover Medical School, Hannover, 30625, Germany

⁴Genome Analysis Center, Institute of Experimental Genetics, Helmholtz Zentrum München, Neuherberg, 84764, Germany

⁵Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Christian-Albrechts-University Kiel, Kiel, 24105, Germany

⁶Department of General Internal Medicine, University Hospital Schleswig-Holstein, Christian-Albrechts-University Kiel, Kiel, 24105, Germany

⁷PopGen Biobank, University Hospital Schleswig-Holstein, Kiel, 24105, Germany

⁸Institute of Medical Informatics and Statistics, Christian-Albrechts-University Kiel, Kiel, 24105, Germany

⁹Department of Nutrition and Food Sciences, Nutritional Epidemiology, University of Bonn, Bonn, 53115, Germany

Abstract

Background: Human metabolic profiles based on the four compound classes acylcarnitines, amino acids, hexoses and phospho and sphingolipids were found to exhibit a significant sex difference in a previous study. We set out to verify this result in a geographically distinct cohort with an adequately sized sample by analyzing the same set of biomarkers and various additional biogenic amines not hitherto considered in such studies.

Methods: The study population comprised 165 individuals (101 males, 64 females) for whom the serum concentrations of 138 different metabolites were measured. Sex differences were analyzed both at the level of individual metabolites by linear regression analysis and at the level of joint consideration of metabolites by partial least-squares discriminant analysis (PLS-DA).

Results: The concentration of 60 metabolites (43.5%) showed a nominally significant sex difference in a linear regression analysis, 11 of which (8.0%) remained significant after correction for multiple testing. Among the previously studied markers, the most significant sex dependency was observed for lyso-phosphatidylcholine acyl C18:2 (adjusted $p=0.001$) and octadecadienyl-L-carnitine ($p=0.004$). Among the newly analyzed biogenic amines, only creatinine (adjusted $p<10^{-4}$) and total dimethylarginine ($p=0.017$) showed a significant sex difference. PLS-DA confirmed the sex dependency of metabolic profiles.

Conclusion: Various previously reported sex differences in human serum metabolite concentration were confirmed in an independent and slightly different cohort. In addition, the concentrations of at least two biogenic amines were found to be sex-dependent as well. In the light of an increased interest in and an increased availability of large-scale metabolic data, our study strongly emphasizes the need for sex stratification or sex adjustment in epidemiological and molecular studies based upon such profiles.

Keywords: Case-control study; Metabolomics; Partial least-square discriminant analysis; Sex adjustment

Introduction

Sex is known to influence the manifestation of many phenotypes in humans, including diseases [1-3]. Correspondingly, either sex stratification or sex adjustment is commonly advised in the epidemiological and molecular analysis of such traits [4-6]. For example, sex differences have been reported recently to characterize even well-established disease associations between single nucleotide polymorphisms and both Coronary Artery Disease and Crohn Disease [7]. Many metabolites are known to play an intermediate, gene- and environment-dependent role in the etiology of complex traits. Thus, metabolic profiles can also be expected to be sex-dependent. Indeed, a recent two-tiered study of the KORA F3 and F4 cohorts by Mittelstrass et al. [8] revealed significantly different serum concentrations in men and women for 63 of 131 metabolic markers tested. However, various environmental factors, such as nutritional status, physical activity level, medication and stress, are well known to also strongly cause variation in single metabolite concentrations [9]. Moreover, since genetic variation influences metabolic pathways, it is conceivable that there is some metabolic differentiation within Germany, given the observed minor genetic differentiation in Germany [10]. We therefore set out to investigate the reported sex dependence of metabolic profiles in an independent population-based study population from Northern Germany with a slightly different genetic composition

and environmental exposition than the previously analyzed South German samples. Moreover, in order to gain further insight into the sex specificity of metabolic activity, we expanded the set of investigated metabolites by the inclusion of various biogenic amines.

Methods

Study design and population

We investigated the metabolic profiles of selected members of the PopGen population-based control cohort, which was established in the city of Kiel, Northern Germany, between June 2005 and February 2006 [11]. Briefly, PopGen participants were ascertained through the local

***Corresponding author:** Sabine Siebert, Institute of Experimental Medicine, Section of Epidemiology, Christian-Albrechts University Kiel, Niemannsweg 11, 24105 Kiel, Germany, Tel: +49-431-597-8685; Fax: +49-431-597-1887; E-mail: s.siebert@iem.uni-kiel.de

Received June 29, 2012; Accepted August 18, 2012; Published August 20, 2012

Citation: Siebert S, Yu Z, Wang-Sattler R, Illig T, Adamski J, et al. (2012) Sex Dependency of Human Metabolic Profiles Revisited. Metabolomics 2:115. doi:10.4172/2153-0769.1000115

Copyright: © 2012 Siebert S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

population registry and invited to visit the study center at the local university hospital. EDTA blood samples were collected at baseline and fasting serum samples stored. All cohort members completed a questionnaire on lifestyle factors and underwent physical examinations, including measurement of waist and hip circumference. Body mass index (BMI) was calculated from self-reported weight and height data. For 230 of the 747 control individuals in PopGen, information on the fasting serum concentration of 186 metabolites was available from a previous study on fatty liver disease (FLD). For the current study, all 115 participants without FLD and 50 FLD patients were randomly selected from the cohort population, thereby mimicking the 30% prevalence of FLD in Germany [12]. All PopGen cohort members were of German descent.

Ethics statement

All PopGen cohort members had given written informed consent prior to the study. Use of PopGen data for the present purpose was also approved by the ethics committee of the Medical Faculty of the Christian-Albrechts University, Kiel, Germany.

Measurement of metabolites and quality control

The serum concentration of 186 metabolites was measured with the AbsoluteIDQ™ p180 Kit (BIOCRATES Life Sciences AG, Innsbruck, Austria) as described [13]. Further details on the assays and reagents used can be found in the AbsoluteIDQ™ p180 Kit manual (www.biocrates.com). Measurements of metabolite concentrations were performed on three different plates, each time with the same kit and the same set of three negative controls. Five additional positive controls ('QC samples') were included on each plate for further quality control. Only metabolites with an average coefficient of variation <25% over the 15 QC samples, and with a detection rate >90% in all 165 serum samples combined, were analyzed further. Detection thresholds ('Limits of Detection', LODs) for single metabolites were taken from the Analytical Specifications AbsoluteIDQ™ p180 Kit manual (AS-p180, available upon request via www.biocrates.com).

Statistical analysis

We tested for sex differences of non-metabolic characteristics by using a χ^2 test for categorical variables and a Student *t*-test for continuous variables.

We analyzed log-transformed single metabolite concentrations using linear regression models with sex as the predictor while adjusting for age, waist-to-hip ratio (WHR), HDL cholesterol level and smoking status. Since most metabolites followed a right-skewed distribution, log-transformation of the metabolites was applied to ensure applicability of the linear regression framework by achieving a better fit to a Gaussian distribution. Given the strong partial correlations between metabolites, we applied the Westfall and Young Step-Down MaxT procedure [14] to correct *p* values for multiple testing. *P* values below 0.05 were considered statistically significant. Prior to regression modeling, we performed a formal power analysis based upon Cohen's effect size measure γ [15].

In order to compare the strengths of single metabolite sex differences between our study and that of Mittelstrass et al. [8], we also calculated the relative sex difference (Δ) for each metabolite, being defined as the difference between the sex-specific mean concentrations divided by mean concentration in men.

Given the strong partial correlations between metabolites, we also performed a partial least-square discriminant analysis (PLS-DA) [16].

PLS-DA focuses upon maximizing the covariance between predictors and response when estimating the parameters of a linear regression model, rather than trying to maximize the variance of the predictors alone. It thus represents a regression extension of unsupervised principal component analysis. Here, we modeled metabolites as predictors and sex as the sole response variable. We quantified the contribution of individual metabolites to the PLS-DA model by the 'Variable Importance in the Projection (VIP)' score [17]. With this score, the average of squared VIP scores over all variables equals unity and metabolites with a VIP score exceeding unity are therefore considered more important for discriminating between sexes than the other metabolites. The results of the PLS-DA were visualized by a scatter plot of the first two PLS components of each metabolic profile. Receiver operator characteristic (ROC) analyses were carried out to assess the predictive potential of the PLS-DA for a given number of PLS components considered. We imputed missing values for leucine (*n*=1), histamine (*n*=9), SDMA (*n*=8) and taurine (*n*=12) by the corresponding sex-specific sample means. For PLS-DA, concentrations of individual metabolites were standardized to unit variance and zero mean.

All analyses were performed using the R statistical software (v. 2.14.1) [18]. For PLS-DA, we used R package *pls* (v. 2.3.0) [19], whereas the ROC analysis was performed using R package *pROC* [20].

Results

The present study included 165 individuals (101 men, 64 women). Male and female participants were found to differ significantly in terms of their waist-to-hip ratio (WHR), HDL cholesterol level and proportion of ever smokers (Table 1). However, while we were analyzing a non-healthy but representative study population no significant differences between men and women could be observed for medical conditions, including FLD, hypertension, cancer, chronic disease, any form of diabetes, gallstones, heart attack, inflammatory bowel disease and neuropathy (Table 1).

Of total 186 measured metabolites 138 (74.19%) passed quality control and entered our analysis, including 68 phosphatidylcholines (89.47%), 21 amino acids (100%), 14 acylcarnitines (35.00%), 14 sphingomyelins (93.33), 11 biogenic amines (57.89%), 9 lyso-phosphatidylcholines (64.29%) and 1 hexose (100%). This marker set comprised 116 of the 131 metabolites originally studied by Mittelstrass et al. [8].

Single-metabolite analysis

Linear regression analysis of the log-transformed metabolite concentrations on sex without further adjustment for any covariates revealed a nominally significant association with sex for 42 of the 138 metabolites (30.4%; Supplementary Table S1). However, given the many phenotypic differences noted between male and female study subjects (Table 1), all models were subsequently adjusted for age, WHR, HDL cholesterol level and smoking status. In the adjusted analysis, 60 of the 138 metabolites (43.5%) showed a nominally significant association with sex, and 11 of these associations (8.0% of the total) remained significant after correction for multiple testing (Table 2). Twenty-two of the 60 nominally significant associations in our study were previously found in the KORA study [8]. Creatinine ($p < 10^{-4}$), lyso-phosphatidylcholine acyl C18:2 ($p = 0.001$), octadecadienyl-L-carnitine (C18:2, $p = 0.004$) and valerylcarnitine (C5, $p = 0.007$) showed the smallest *P* values in our study with lyso-phosphatidylcholine acyl C18:2 also

	Men	Women	p
N	101	64	
Age, years	60 (6)	62 (7)	0.084
Range	50 - 76	50 - 77	
Body mass index, kg/m²	26.19 (3.80)	25.85 (4.69)	0.612
Waist-hip ratio	0.93 (0.07)	0.84 (0.09)	1.04×10 ⁻¹²
HDL cholesterol, mg/dl	59.36 (14.43)	76.78 (21.07)	2.61×10 ⁻⁹
LDL cholesterol, mg/dl	143.38 (35.12)	147.20 (37.64)	0.508
Triglyceride, mg/dl	147.59 (94.08)	128.22 (72.66)	0.163
Ever smoker, n (%)	64 (63.37%)	27 (42.19%)	0.012
Alcohol intake¹, n (%)	40 (54.79)	17 (38.64)	0.133
Alcohol intake, C2-units	0.64 (0.88)	0.38 (0.60)	0.085
High alcohol intake ² , n (%)	8 (10.96)	3 (6.82)	0.677
Fatty liver disease, n (%)	30 (29.70)	20 (31.25)	0.971
Hypertension, n (%)	66 (65.36)	34 (53.13)	0.161
Prevalent diseases³, n (%)			
0	42 (41.58)	18 (28.13)	0.092
1	38 (37.62)	31 (48.44)	
2	17 (16.83)	8 (12.5)	
3	2 (1.98)	6 (9.38)	
4	2 (1.98)	1 (1.56)	

Data are given as mean (standard deviation) unless indicated otherwise. P values refer to a χ^2 test or a Student *t* test, as appropriate.

¹Data are based upon 117 individuals because of missing values (28 men, 20 female).

²High alcohol intake was defined as consuming more than 4 C2-units for men and 2 C2-units for women

³Number of prevalent diseases including cancer, chronic disease, any form of diabetes, gallstones, heart attack, inflammatory bowel disease and neuropathy

Table 1: Participant Characteristics.

Metabolite	Unadjusted model		Adjusted model		P _{adj} ¹	Δ (%) ²
	β	p	β	p		
Acylcarnitines						
C12	0.17	0.025	0.29	0.002	0.033	16.23
C14:2	0.19	0.019	0.32	0.002	0.029	13.49
C16	0.19	0.004	0.26	0.002	0.025	16.57
C18	0.21	0.001	0.28	4.5×10 ⁻⁴	0.010	18.36
C18:2	0.25	4.1×10 ⁻⁴	0.33	1.7×10 ⁻⁴	0.004	18.62
C5	0.38	2.8×10 ⁻⁹	0.28	2.7×10 ⁻⁴	0.007	31.62
Amino Acids						
Glutamine	0.17	0.001	0.22	0.001	0.011	12.70
Biogenic Amines						
Creatinine	0.28	1.1×10 ⁻⁸	0.30	10 ⁻⁸	<10 ⁻⁴	21.42
total DMA	0.13	0.004	0.19	0.001	0.017	9.47
Lyso-Phosphatidylcholines						
lysoPC a C18:1	0.18	0.015	0.28	0.002	0.029	12.49
lysoPC a C18:2	0.28	0.001	0.41	3.9×10 ⁻⁵	0.001	19.39

P values and regression coefficients (β) derived from a linear regression analysis of log-transformed metabolite concentrations measured in 165 study participants (101 men, 64 women). All models were also additionally adjusted for age, waist-to-hip ratio, HDL cholesterol and smoking status

¹ P_{adj} denotes the p value adjusted for multiple testing by the Westfall and Young Step-Down MaxT procedure with 10,000 permutations

² Δ denotes relative sex difference, defined as the difference in mean concentration between men and women, divided by the mean concentration in men

Table 2: Significant Sex Differences in Serum Metabolite Concentration (after correction for multiple testing).

found to be significantly associated with sex in the study by Mittelstrass et al. [8].

While most of the phosphatidylcholines and sphingomyelins tended to be higher in women than in men in the unadjusted analysis, this relation was reversed more often than not when HDL cholesterol was adjusted for (Supplementary Table S1). Moreover association with sex of some phosphatidylcholines and sphingomyelins became nominally significant in both the univariate and the multivariate regression analysis only after including HDL cholesterol (Supplementary Table

S1 and Figure 1). Calculation of Pearson's correlation coefficient confirmed weak positive linear relationship between the lipoprotein and almost half of the phospholipids (28 phosphatidylcholines and 8 sphingomyelins with 0.25 ≤ *r* ≤ 0.5).

For the vast majority of metabolites, regression results did not substantially change after adjustment for BMI instead of WHR or vice versa (Supplementary Table S2). However, lyso-phosphatidylcholines acyl C17:0 showed a significant sex difference after adjustment for WHR, but not after adjustment for BMI,

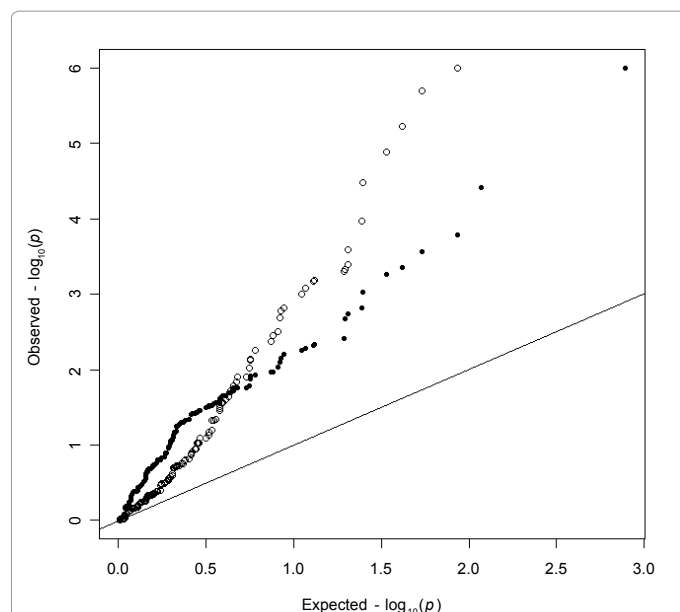


Figure 1: Q-Q Plot of p values from single-metabolite linear regression analyses. P values for the association between sex and single metabolite concentrations were obtained from linear regression models of log-transformed metabolites regressed on sex, while, at a later stage, additionally adjusting for age, waist-to-hip ratio, HDL cholesterol level and smoking status. Open circles: unadjusted model; filled circles: adjusted model.

thereby corroborating previous findings by Mittelstrass et al. [8]. Moreover, for the most part results of our linear regression analyses were independent of further adjustment for lipid parameters (LDL cholesterol, triglycerides; Supplementary Table S2).

Because information on alcohol consumption was available for only 117 study participants (73 men, 44 women), a separate analysis was conducted on this subset. Although not significant likely due to the limited sample size, the absolute number of alcohol consumers and the average intake (in C2-units) differed among male and female study participants (Table1, $p=0.133/p=0.085$). Single regression analysis restricted to this sample subset yielded very similar results for the metabolites, although associations between sex and single metabolites became less significant due to the smaller sample size. However, when both, the former unadjusted and adjusted model, were additionally adjusted for alcohol intake, results were not influenced and significant sex-effects remained significant (data not shown).

Next we ruled out biases possibly arising from analyzing a non-healthy but representative study population. In any case, the results of our linear regression analyses changed only marginally when adjusted for FLD, hypertension and other prevalent diseases. Moreover, the major conclusion that many metabolite concentrations show a pronounced sex difference also did not change when only the 115 healthy controls were analyzed (data not shown).

Almost all acylcarnitines, amino acids, biogenic amines, hexose and lyso-phosphatidylcholines showed higher concentrations in men than in woman, while phosphatidylcholines and sphingomyelins tended to be higher in women than in men (Supplementary Table S1). The largest relative sex difference was observed for valerylcarnitine (C5, $\Delta=31.6\%$), alpha-Aminoadipic acid (alpha-AAA, $\Delta=27.0\%$), phosphatidylcholines diacyl C 32:2 ($\Delta=-25.7\%$), C 34:3 ($\Delta=-23.7\%$) and C 36:6 ($\Delta=-21.9\%$).

Joint-metabolites analysis

Joint discriminatory analysis (PLS-DA) of all 138 metabolites discriminated well, although not perfectly, between sexes (Figure 2). Leave-one-out cross-validation identified the first five PLS components as providing optimal discriminatory power (minimal root mean squared error of prediction: 38.7%). Taken together, these five components explained 65% of the variation in the predictors (metabolites) and 68% of the variation in the response (sex). VIP scores based upon the first five PLS components indicated acylcarnitines (particularly C5, C3, C0 and C18) and amino acids (particularly leucine, glycine, isoleucine, proline, valine and ornithine) as more important for discriminating between sexes in the PLS-DA model than the other metabolite classes (Supplementary Table S1). Subsequent ROC analysis yielded an area under curve (AUC) of 99.0% (95% CI: [98.0; 100]) for the first five PLS components, with a specificity of 95.3% and a sensitivity of 96.0% obtained at the optimal threshold for the dependent variable of the regression model (Supplementary Figure S1). Ten-fold cross-validation with an equal number of men and women in each partition yielded a misclassification rate estimate of 19%.

Discussion

In this study, we could confirm a previously reported sex dependency of human metabolic profiles, comprising acylcarnitines, amino acids, biogenic amines, hexose, phosphatidylcholines, lyso-phosphatidylcholines and sphingomyelins [8], in an independent German cohort of slightly different genetic compositions [10] and a slightly different set of environmental factor values. Whilst joint consideration of metabolites in a PLS-DA revealed an overall sex difference in terms of the measured serum concentrations, linear regression analysis of individual metabolites revealed various statistically significant sex dependencies. Acylcarnitines, with 6 out

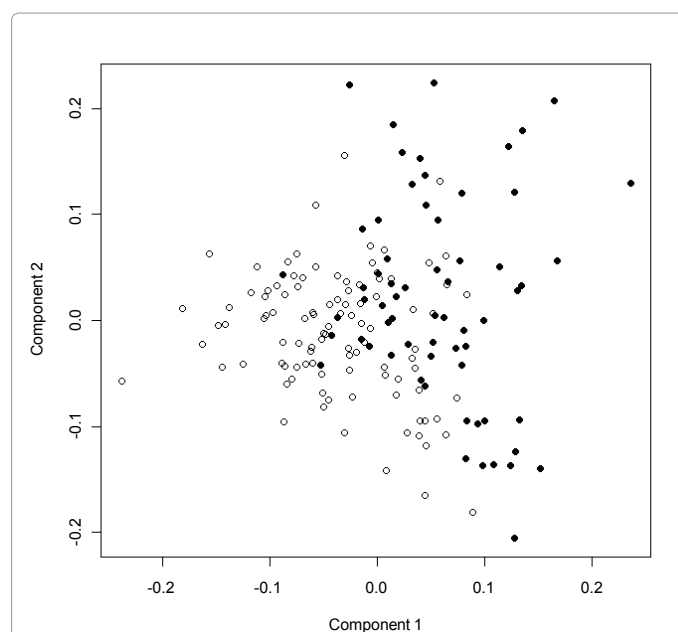


Figure 2: Scatter plot of the two PLS components from a partial least-square discriminant analysis (PLS-DA). Components were obtained from a joint analysis of 138 serum metabolite concentrations measured in 165 individuals (101 male, 64 female) using PLS-DA. Open circles: males; filled circles: females.

of 14 metabolites yielding a multiplicity-corrected significant p value, represented the most discriminatory group.

More specifically, serum concentration of individual metabolites showed a similar trend of sex dependency in the PopGen and the KORA cohorts [8] and, in most cases, outcome difference of the same magnitude. Mean serum concentrations were generally higher in men than in women in both studies, with the exception of the phosphatidylcholines and sphingomyelins. Glycine was the only amino acid with a marked, albeit not significantly, lower serum level in men than women, and this was observed in both PopGen ($\Delta=-8.2\%$) and KORA F4 ($\Delta=-14\%$). The serum concentration of hexose, the sum of C6 sugars, was consistently higher in men than women (PopGen $\Delta=14.1\%$) even though the difference failed to attain statistical significance in our study after correction for multiple testing.

Mittelstrass et al. [8] found 102 of the 131 metabolite concentrations (77.9%) analyzed in KORA F4 to differ significantly between men and women. Sixty-three of them (61.8%, or 48.1% of the total) were also found to be significantly associated with sex in another (verification) KORA cohort, KORA F3 [8]. We were able to confirm this result for 22 of the 63 metabolites (34.9%) through a nominally significant association with sex in PopGen. Interestingly, C18 was the only one of 12 acylcarnitines for which the sex difference discovered in KORA F4 could be verified in KORA F3. In contrast, our study corroborated a significant sex difference for six of the acylcarnitines (C12, C14:2, C16, C18, C18:2 and C5) after correction for multiple testing.

Sex differences for some phosphatidylcholines and sphingomyelins became nominally significant in our study only after adjustment for HDL cholesterol. HDL is known to be higher on average in women than in men [21], which was also the case here (Table 1). Since phospholipids are a predominant component (42-51%) of the lipid moiety of HDL [22], it was not surprising that the observed association between phospholipids and sex was modified by the inclusion of HDL cholesterol in the regression model. Lifestyle factors, like smoking behavior or alcohol intake, might also be confounders of sex-specific metabolic profiles. However, we found sex differences to be largely invariant to adjustments for smoking status and alcohol consumption in our study. Adjustment for lipid parameters (LDL cholesterol, triglycerides) as well as medical conditions pointed towards independence, and the strength of association was only slightly affected by adjustment for WHR or BMI.

A recent study emphasized properly taking into account the high variability of serum metabolite concentrations apparent in living adults even without any obvious metabolic disorder [23]. The authors referred to parameters like age, sex, genetic background, ethnicity, diurnal variation, diet, and health status or activity level as causing factors for these relatively large ranges of metabolite concentrations. Thus, results may have been affected by differences among the PopGen and KORA cohort, especially by means of differing regional habits (Northern vs. Southern Germany) such as diet or other lifestyle factors. In particular, the cohorts differed in mean age, BMI, WHR; values of blood lipids as well as proportion of smokers and high alcoholic intake (Table 1 and [8]). In addition, a recently SNP-based analysis of genetic substructure in the German population revealed genetic differentiation between the samples along a North-South gradient within Germany, namely between PopGen and KORA subjects [10]. Nevertheless, even if the number of significant findings differed between our study and that of Mittelstrass et al. [8], possibly due to the lower sample size of our study, the overall finding (sex difference in human metabolic profiles) was the same and, more specifically, the predominant proportion of

comparisons showed a similar trend and outcome differences of the same magnitude.

Joint PLS-DA of all 138 metabolites discriminated well and subsequent ROC analysis yielded an AUC of 99.0% for the first five PLS components. However, since this analysis used the same data as training and as test set, this impressive value likely overestimates the predictive power provided by the metabolites. We therefore performed ten-fold cross-validation with an equal number of men and women in each partition to obtain a less biased estimate of the classification error. Based on the first five PLS components of individual metabolic profiles, 22 men were classified as women, 10 women as men and 133 individuals (79 men, 54 women) were classified correctly, corresponding to a notably low misclassification rate of 19%.

A potential limitation of our study may have been that the study samples were taken from a sex- and age-matched case-control study of fatty liver disease (FLD). Since the liver is a metabolically active organ and FLD-related changes in metabolite concentrations are likely, the sex differences detected in our study could have been due to the presence of one or more FLD-associated confounders. To reduce the risk of confounding, however, 50 patients with sonographically diagnosed FLD were chosen at random to complement the 115 non-FLD samples available, thereby mimicking the 30% prevalence of FLD reported for the German population [12]. Moreover, the major conclusion that many metabolite concentrations show a pronounced sex difference also did not change when adjusting for medical conditions or when analyzing only the 115 healthy controls. While physical activity and nutritional status have previously been reported to likely influence the serum metabolic profile [9,23], these variables could not be included in our analyses due to the degree of missing information. Another potential limitation of our study was the small sample size. However, formal power analysis revealed that our sample had sufficient power (80%) to detect effect sizes γ of 0.45 or higher, justifying replication of at least modest sex effects.

In addition to confirming previous results in an independent sample of slightly different genetic composition and environmental exposition, our study revealed a sex difference in the serum concentration of metabolites not hitherto analyzed. In particular, the AbsoluteIDQ™ p180 kit used in our study comprises an additional set of 19 biogenic amines that were not included in the AbsoluteIDQ™ p150 kit used by Mittelstrass et al. [8] and of which 11 entered our analysis. We found that the serum concentrations of these metabolites are also likely to be sex-dependent, especially in the case of creatinine ($p<0.001$) and total dimethylarginine (total DMA, $p=0.017$). Creatinine, whose serum concentration is known to increase with muscle mass, has been shown to be sex-dependent before [24]. This notwithstanding, whilst the possibility of a sex difference in metabolic profiles comprising acylcarnitines, amino acids, glycerophospholipids and sphingomyelins has been reviewed systematically already [8], further studies are required to better clarify the sex-specific role of biogenic amines.

In conclusion, our study has confirmed a previously discovered sex difference in human metabolic profiles. Given the expanding scale of both interest and availability of metabolic profiles in biomedical research, our results emphasize the importance of taking sex into account when analyzing such profiles to avoid spurious results in the identification of metabolic risk factors in epidemiological and molecular studies. On the other hand, even though our results are also strongly indicative of sex differences for various metabolite concentrations, a considerable overlap between male and female profiles was still immanent, both at the level of individual markers and

in the multivariate PLS-DA. The use of the term “sexual dimorphism” by Mittelstrass et al. [8] to characterize an otherwise incontrovertible sex difference appears therefore somewhat inappropriate to us and should be avoided in the future.

Acknowledgement

We thank all PopGen participants and the PopGen team for their support. The authors are also most grateful to staff at the Genome Analysis Center and Metabolomics Core Facility, Helmholtz Centre Munich, particularly Werner Römisch-Margl, Julia Scarpa, Katharina Sckell and Arsin Sabuchi, for their help with the sample logistics, data collection and metabolite quantification.

References

- Barnes LL, Wilson RS, Bienias JL, Schneider JA, Evans DA, et al. (2005) Sex differences in the clinical manifestations of Alzheimer disease pathology. *Arch Gen Psychiatry* 62: 685-691.
- Reeves MJ, Bushnell CD, Howard G, Gargano JW, Duncan PW, et al. (2008) Sex differences in stroke: epidemiology, clinical presentation, medical care, and outcomes. *Lancet Neurol* 7: 915-926.
- Mostertz W, Stevenson M, Acharya C, Chan I, Walters K, et al. (2010) Age- and sex-specific genomic profiles in non-small cell lung cancer. *JAMA* 303: 535-543.
- Geller SE, Adams MG, Carnes M (2006) Adherence to federal guidelines for reporting of sex and race/ethnicity in clinical trials. *J Womens Health (Larchmt)* 15: 1123-1131.
- Kim AM, Tinggen CM, Woodruff TK (2010) Sex bias in trials and treatment must end. *Nature* 465: 688-689.
- Tingen CM, Kim AM, Wu PH, Woodruff TK (2010) Sex and sensitivity: the continued need for sex-based biomedical research and implementation. *Womens Health (Lond Engl)* 6: 511-516.
- Liu LY, Schaub MA, Sirota M, Butte AJ (2012) Sex differences in disease risk from reported genome-wide association study findings. *Hum Genet* 131: 353-364.
- Mittelstrass K, Ried JS, Yu Z, Krumsiek J, Gieger C, et al. (2011) Discovery of sexual dimorphisms in metabolic and genetic biomarkers. *PLoS Genet* 7: e1002215.
- Holmes E, Nicholson JK (2007) Human metabolic phenotyping and metabolome wide association studies. *Ernst Schering Found Symp Proc* 227-249.
- Steffens M, Lamina C, Illig T, Bettecken T, Vogler R, et al. (2006) SNP-based analysis of genetic substructure in the German population. *Hum Hered* 62: 20-29.
- Krawczak M, Nikolaus S, von Eberstein H, Croucher PJ, El Mokhtari NE, et al. (2006) PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet* 9: 55-61.
- Baumeister SE, Volzke H, Marschall P, John U, Schmidt CO, et al. (2008) Impact of fatty liver disease on health care utilization and costs in a general population: a 5-year observation. *Gastroenterology* 134: 85-94.
- Illig T, Gieger C, Zhai G, Römisch-Margl W, Wang-Sattler R, et al. (2010) A genome-wide perspective of genetic variation in human metabolism. *Nat Genet* 42: 137-141.
- Westfall PH, Young SS (1993) Resampling-based multiple testing: Examples and methods for p-value adjustment. John Wiley & Sons.
- Cohen J (1969) Statistical Power Analysis for the Behavioral Sciences. (2nd edn), Academic Press, New York, London.
- Jong SD (1993) SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* 18: 251-263.
- Wold S, Johansson E, Cocchi M (1993) 3D QSAR in Drug Design; Theory, Methods and Applications. Leiden: Escom Science Publisher.
- Team RDC (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Mevik B-H, Wehrens R, Hovde Liland K (2011) pls: Partial Least Squares and Principal Component regression. R package version 2.3-0.
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, et al. (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12: 77.
- Davis CE, Williams DH, Oganov RG, Tao SC, Rywik SL, et al. (1996) Sex difference in high density lipoprotein cholesterol in six countries. *Am J Epidemiol* 143: 1100-1106.
- Robichaud JC, van der Veen JN, Yao Z, Trigatti B, Vance DE (2009) Hepatic uptake and metabolism of phosphatidylcholine associated with high density lipoproteins. *Biochim Biophys Acta* 1790: 538-551.
- Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, et al. (2011) The human serum metabolome. *PLoS One* 6: e16957.
- Jones CA, McQuillan GM, Kusek JW, Eberhardt MS, Herman WH, et al. (1998) Serum creatinine levels in the US population: third National Health and Nutrition Examination Survey. *Am J Kidney Dis* 32: 992-999.