Identification of Comprehensive Metabotypes Associated with Cardiometabolic Diseases in the Population-Based KORA Study

Anna Riedl, Nina Wawro, Christian Gieger, Christa Meisinger, Annette Peters, Michael Roden, Florian Kronenberg, Christian Herder, Wolfgang Rathmann, Henry Völzke, Martin Reincke, Wolfgang Koenig, Henri Wallaschofski, Hans Hauner, Hannelore Daniel, and Jakob Linseisen*

Scope: "Metabotyping" describes the grouping of metabolically similar individuals. We aimed to identify valid metabotypes in a large cohort for targeted dietary intervention, for example, for disease prevention.

Methods and results: We grouped 1729 adults aged 32–77 years of the German population-based KORA F4 study (2006–2008) using k-means cluster analysis based on 34 biochemical and anthropometric parameters. We identified three metabolically distinct clusters showing significantly different biochemical parameter concentrations. Cardiometabolic disease status was determined at baseline in the F4 study and at the 7 year follow-up termed FF4 (2013/2014) to compare disease prevalence and incidence between clusters. Cluster 3 showed the most unfavorable marker profile with the highest prevalence of cardiometabolic diseases. Also, disease incidence was higher in cluster 3 compared to clusters 2 and 1, respectively, for hypertension (41.2%/25.3%/18.2%), type 2 diabetes (28.3%/5.1%/2.0%), hyperuricemia/gout (10.8%/2.3%/0.7%), dyslipidemia (19.2%/18.3%/5.6%), all metabolic (54.5%/36.8%/19.7%), and all cardiovascular (6.3%/5.5%/2.3%) diseases together.

Conclusion: Cluster analysis based on an extensive set of biochemical and anthropometric parameters allows the identification of comprehensive metabotypes that were distinctly different in cardiometabolic disease occurrence. As a next step, targeted dietary strategies should be developed with the goal of preventing diseases, especially in cluster 3.

A. Riedl, Dr. N. Wawro, Prof. C. Meisinger, Prof. J. Linseisen
Independent Research Group Clinical Epidemiology
Helmholtz Zentrum München
German Research Center for Environmental Health
Ingolstädter Landstr. 1, 85764 Neuherberg, Germany
E-mail: anna.riedl@helmholtz-muenchen.de

A. Riedl, Dr. N. Wawro, Dr. C. Gieger, Prof. C. Meisinger, Prof. A. Peters, Prof. J. Linseisen
Institute of Epidemiology II
Helmholtz Zentrum München
German Research Center for Environmental Health
Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

A. Riedl, Dr. N. Wawro, Dr. C. Gieger, Prof. A. Peters, Prof. M. Roden, Prof. C. Herder, Prof. W. Rathmann, Prof. H. Völzke
German Center for Diabetes Research
Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

A. Riedl, Dr. N. Wawro, Prof. C. Meisinger, Prof. J. Linseisen
Chair of Epidemiology
Ludwig-Maximilians-Universität München
at UNIKA-T (Universitäres Zentrum für Gesundheitswissenschaften am Klinikum Augsburg)
Neusässer Str. 47, 86156 Augsburg, Germany

Dr. C. Gieger
Research Unit of Molecular Epidemiology
Helmholtz Zentrum München
German Research Center for Environmental Health (GmbH)
Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

Prof. M. Roden
Division of Endocrinology and Diabetology
Medical Faculty
Heinrich Heine University Düsseldorf
Auf'm Hennekamp 65, 40225 Düsseldorf, Germany

Prof. M. Roden, Prof. C. Herder
Institute for Clinical Diabetology
German Diabetes Center
Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf
Auf'm Hennekamp 65, 40225 Düsseldorf, Germany

Prof. F. Kronenberg
Division of Genetic Epidemiology
Department of Medical Genetics
Molecular and Clinical Pharmacology
Medical University of Innsbruck
Schöpfstr. 41, 6020 Innsbruck, Austria

1. Introduction

Intrinsic and environmental influences like genetics, epigenetics, body constitution, or the gut microbiome in interplay with lifestyle lead to differences in metabolism, dietary requirements, and the health–disease trajectory.[1–5] Metabotyping or metabolic phenotyping describes the process of defining homogeneous subgroups of individuals by means of their metabolic or phenotypic characteristics. These subgroups are termed metabotypes or metabolic phenotypes. Individuals within a subgroup show a higher metabolic similarity compared to individuals from different subgroups.[6–11] Targeted dietary recommendations at the metabolic subgroup level are assumed to be more effective with regard to individual health benefits and health care costs than the currently used general dietary advice.[1,4,6,12,13]

To date, several attempts have been made to identify metabotypes in humans.[6,14] However, there exists no uniform definition of the term "metabotype." Studies available reveal a large heterogeneity in the samples/populations studied, the statistical methods applied, the number of metabotypes identified, and especially in the number and type of clustering variables used. Most studies grouped individuals based on a few metabolites or single specific metabolite subclasses, and thereby considered only small parts of human metabolism, resulting in a restricted definition.[7,15–30] The metabotype definition could become more refined and robust by the inclusion of a broad range of parameters covering different metabolic pathways and processes or by the inclusion of further phenotypic data.[14]

Therefore, the present study examined whether cluster analysis based on an extensive set of biochemical and anthropometric parameters can define comprehensive and valid metabotypes in a population-based cohort study. Currently, little is known about disease risks associated with metabotypes defined in previous studies. Only a few selected diseases were considered there,[15–20,28,29,31,32] whereas in the present study, we analyzed the occurrence of diet-related diseases, namely hypertension, type 2 diabetes, hyperuricemia/gout, dyslipidemia, myocardial infarction, stroke and cancer, in clusters defining metabotypes in the well-phenotyped population-based Cooperative Health Research in the Region of Augsburg (KORA) cohort. These metabotypes may be used later to develop targeted and precise strategies for cardiometabolic disease prevention and treatment on this metabotype group level.

Prof. W. Rathmann
Institute for Biometrics and Epidemiology
German Diabetes Center
Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf
Auf'm Hennekamp 65, 40225 Düsseldorf, Germany
Prof. H. Völzke, Prof. W. Koenig
DZHK – German Centre for Cardiovascular Research
Partner Site Munich Heart Alliance
Pettenkoferstr. 8a & 9, 80336 Munich, Germany
Prof. H. Völzke
Institute for Community Medicine
University Medicine Greifswald
Walther-Rathenau-Str. 48, 17475 Greifswald, Germany
Prof. M. Reincke
Medizinische Klinik und Poliklinik IV
Klinikum der Universität München
Ludwig-Maximilians-Universität München
Ziemssenstr. 1, 81377 Munich, Germany
Prof. W. Koenig
Deutsches Herzzentrum München
Technische Universität München
Lazarettstr. 36, 80636 Munich, Germany
Prof. W. Koenig
Department of Internal Medicine II-Cardiology
University of Ulm Medical Center
Albert-Einstein-Allee 23, 89081 Ulm, Germany
Prof. H. Wallaschofski
Institute of Clinical Chemistry and Laboratory Medicine
University Medicine Greifswald
Ferdinand-Sauerbruch-Str., 17489 Greifswald, Germany
Prof. H. Hauner
Else Kröner-Fresenius Centre for Nutritional Medicine
Technical University of Munich
Gregor-Mendel-Str. 2, 85354 Freising-Weihenstephan, Germany
Prof. H. Hauner, Prof. J. Linseisen
ZIEL – Institute for Food and Health
Technical University of Munich
Weihenstephaner Berg 1, 85354 Freising, Germany
Prof. H. Hauner
Institute of Nutritional Medicine
Klinikum rechts der Isar
Technical University of Munich
Uptown München Campus D, Georg-Brauchle-Ring 60/62, 80992 Munich, Germany
Prof. H. Daniel
Technical University of Munich
Gregor-Mendel-Str. 2, 85354 Freising-Weihenstephan, Germany

2. Experimental Section

2.1. Study Population

Analyses are based on data from the population-based KORA F4 (2006–2008) and FF4 (2013/2014) studies, both follow-ups of the KORA S4 study conducted between 1999 and 2001 in the region of Augsburg in Southern Germany.[33] Briefly, S4 included 4261 participants aged 25–74 years, whereof 3080 individuals and 2279 individuals, respectively, took part in the subsequent first follow-up (F4) after 7 years and in the second follow-up examination (FF4) after another 7 years. In total, 2161 individuals participated in both follow-ups. Reduction of sample size from F4 to FF4 occurred because participants had died ($n = 168$), moved away ($n = 97$), refused to be contacted ($n = 67$) or could not be contacted ($n = 48$), were unable to come ($n = 332$), or refused to participate ($n = 207$) in the second follow-up. All participants answered self-administered questionnaires, and underwent a standardized physical examination as well as a computer-assisted personal interview by trained staff at the study center. Details have been described elsewhere.[34] Written informed consent was obtained from all participants and the studies were approved by the Ethics Committee of the Bavarian Medical Association and conducted in accordance with the Declaration of Helsinki.

2.2. Assessment of Biochemical and Anthropometric Parameters for Metabotyping

In addition to BMI, all 33 fasting biochemical parameters available from KORA F4 were used to identify comprehensive

metabotypes. These were assessed by standard methods in a subsample of 1768 participants and are described in the following.

In serum, HDL, LDL, and total cholesterol (TC) quantification was performed by enzymatic, colorimetric methods (CHOD-PAP method; AHDL/ALDL/CHOL Flex, Dade Behring, Marburg, Germany). TC and HDL were additionally combined by dividing TC by HDL denoted as TC/HDL. Triglycerides (TGs) (GPO-PAP; TGL Flex, Dade Behring), glucose (hexokinase method; GLU Flex, Dade Behring), uric acid (uricase method; URCA Flex, Dade Behring) as well as the liver enzymes gamma-glutamyltransferase (GGT), glutamate-pyruvate transaminase (GPT), glutamate-oxaloacetate transaminase (GOT), and alkaline phosphatase (AP) (IFCC method; Roche Diagnostics, Mannheim, Germany) were quantified by enzymatic color tests. ELISA kits were used to determine insulin (Invitrogen; Camarillo, CA), IL-18 (MBL; Nagoya, Japan) and thyroid peroxidase (TPO) antibodies (VARELISA; Elias Medizintechnik GmbH, Freiburg, Germany). High-sensitive C-reactive protein (hs-CRP) and cystatin C were measured by immunonephelometry on a BN II analyzer using reagents from Siemens (Erlangen, Germany). Levels of thyroid-stimulating hormone (TSH) and free thyroxine (FT4) were analyzed by immunochemiluminescence procedures (Dimension Vista System, Siemens, Germany).

In plasma, concentrations of lipoprotein(a) (Lp(a)), apolipoprotein A-IV (ApoA-IV), afamin (Division of Genetic Epidemiology, Innsbruck, Austria), leptin (Mercodia, Uppsala, Sweden), and sex-hormone-binding globulin (SHBG; IBL International GmbH, Hamburg, Germany) were assessed using ELISA kits. Renin was quantified by a chemiluminescence assay (LIASON Direct Renin; DiaSorin, Dietzenbach, Germany) and aldosterone by an in-house immunofluorescence assay. Nonesterified fatty acids (NEFAs) were measured by an enzymatic colorimetric assay (NEFA-C Kit; Wako, Richmond, VA, USA). Insulin-like growth factor-I (IGF-I) and insulin-like-growth-factor-binding-protein-3 (IGFBP3) were measured using an automated chemiluminescence immunoassay (IDS-iSYS; Immunodiagnostic Systems, Munich, Germany).

Fresh venous whole EDTA blood was used to determine the leukocyte count by impedance measurements (Coulter LH 750; Beckman Coulter Inc., USA). The average telomere length in leukocytes was measured using a quantitative PCR-based technique.[35] Glycated hemoglobin (HbA1c) was determined with a reverse-phase, cation-exchange, HPLC method (HA 8160; Menarini, Florence, Italy).

Creatinine and albumin (CREATININ-JK; Greiner, Bahlingen, Germany) were determined from frozen urine by nephelometry on a BN II analyzer (Siemens, Erlangen, Germany) and a Cobas Mira analyzer (Roche Diagnostics, Mannheim, Germany), respectively.

The BMI was calculated as the ratio of weight in kilogram to height in square meter, both measured in light clothing without shoes at a physical examination by trained staff. The BMI was used continuously in the identification of metabotypes and additionally categorized into underweight (BMI <18.5 kg m^{-2}), normal (BMI 18.5 to <25 kg m^{-2}), overweight (BMI 25 to <30 kg m^{-2}), and obese (BMI $\geq$ 30 kg m^{-2}) to describe the metabotypes.

2.3. Assessment of Variables for the Description of Metabotypes

Information on socio-demographics and lifestyle was collected in a standardized personalized computer-assisted interview and a self-administered questionnaire in F4. Socio-demographic data included sex, age (in years), and education (<10 years, 10 to <12 years, $\geq$ 12 years). Additionally, data on physical activity (active, inactive), and smoking (nonsmoker, ex-smoker, smoker) were used. Alcohol consumption was categorized by gender into no (0 g d^{-1}), moderate (women: >0 to <20 g d^{-1}; men: >0 to <40 g d^{-1}), and high (women: ≥ 20 g d^{-1}; men: ≥ 40 g d^{-1}) alcohol consumption.

The study participants were also interviewed about the presence (yes/no) of the following cardiometabolic diseases in F4 as well as in FF4: hypertension was defined by a blood pressure of $\geq$140/90 mmHg, or as the existence of a known, drug-controlled hypertension. Type 2 diabetes was defined by either current intake of an antidiabetic medication or a self-reported diagnosis validated by the respective treating physician. Cancer history and inpatient treatment of myocardial infarction or stroke were based on self-report. Dyslipidemia and hyperuricemia/gout were assessed by the current intake of lipid-lowering drugs and hyperuricemia/gout medication, respectively.

2.4. Statistical Analysis

We performed all statistical analyses using the statistical software package RStudio version 1.0.136 that uses R version 3.2.2 (R Development Core Team, 2010, http://www.r-project.org). p-values < 0.05 were considered statistically significant.

2.4.1. Preprocessing

Of the 1768 participants with information on biochemical and anthropometric parameters, 23 participants were excluded from analysis as they were not fasting at least 8 h before blood collection and 16 participants were excluded due to more than 10% missing values of all clustering variables. The remaining missing values (0.0–8.8% regarding subjects and 0.0–3.1% regarding biochemical and anthropometric parameters) were imputed in the resulting dataset ($n = 1729$, $v = 34$) using the R package "mice" (multivariate imputation by chained equations) version 2.25,[36] since complete data are a prerequisite for clustering. In total, five imputed datasets were generated with ten iterations each. Before clustering, all biochemical and anthropometric parameters were z-standardized to avoid bias by different scales and units in the clustering process.

2.4.2. Metabotyping

Subsequently, metabotyping of 1729 F4 participants was performed simultaneously for all five imputed datasets using the k-means cluster algorithm (employing Euclidean distance as distance measure) of the R package "miclust" (multiple imputation in cluster analysis) version 1.2.5.[37] BMI and all 33 biochemical parameters available at F4, but no further variables were

included. A series of cluster analyses were conducted with different cluster numbers ($k = 2$–8). We applied 26 cluster validity criteria of the R package "NbClust" version 3.0[38] to determine the most appropriate cluster number and to validate the cluster solutions. These revealed the best classification into two or three clusters (Figure S1, Supporting Information). We present results based on the 3-cluster solution. Three clusters allow for a more precise metabolic characterization of individuals than the 2-cluster solution. Results of the 2-cluster solution and a flow chart showing the overall analysis strategy are provided in Figure S2 and Tables S1–S4, Supporting Information. The metabotypes were further validated by the analysis of differences in the biochemical and anthropometric parameter values between clusters to identify metabolically homogeneous subgroups.

2.4.3. Descriptive Statistics of Metabotypes

Descriptive data are shown as median, 25th and 75th percentiles for continuous variables and as total and relative frequencies for categorical variables. Differences in characteristics across clusters were detected using Kruskal–Wallis test (with subsequent Kruskal–Wallis post hoc test with Bonferroni correction) for continuous variables and Pearson's chi-squared test (with subsequent chi-squared post hoc test with Bonferroni correction) for categorical variables. The prevalence of cardiometabolic diseases was assessed in total and by cluster only in participants with information on the respective disease status at F4 available. Data on the incidence of cardiometabolic diseases during the 7 year follow-up from F4 to FF4 were given for the whole study population and for each cluster separately; only participants with a disease diagnosis in FF4 and absence of the respective disease at baseline in F4 were considered as incident cases. The cardiometabolic diseases were analyzed individually and summarized into metabolic (hypertension, type 2 diabetes, hyperuricemia/gout, and dyslipidemia) and cardiovascular diseases (myocardial infarction and stroke). The presence of the summarized diseases was defined as suffering from at least one of the respective individual diseases.

3. Results

Table 1 summarizes the demographic characteristics of the study population at F4 in total and for each of the three metabotype clusters. In the total study population, participants were 32–77 years old and the proportion of men and women was approximately equal. About 44% of participants were assigned to the largest cluster 1 ($n = 760$), about 36% to cluster 2 ($n = 616$), and about 20% to the smallest cluster 3 ($n = 353$). In cluster 1 about one-third of the participants were male, whereas in clusters 2 and 3 about two-thirds were men. Persons in cluster 3 were characterized by the highest median age of 66.0 years (range = 41–77 years) and BMI of 31.8 kg m^{-2} (range = 21.8–56.0 kg m^{-2}). Participants in cluster 3 were also those who were most inactive (54.4%). Compared to clusters 1 and 2 they had the highest percentage of alcohol abstainers, as well as the lowest proportions of current smokers and highly educated individuals.

The median values of the 34 biochemical and anthropometric parameters used as clustering variables are presented in total and stratified by cluster in **Table 2**. Except for TPO antibodies ($p = 0.08$) and aldosterone ($p = 0.21$), all other parameters were found to be significantly different across clusters. Thus, all three clusters were well separated and showed a distinct marker profile. The median values were highest in 23 parameters in cluster 3 and lowest in 24 parameters in cluster 1, mostly in the same parameters. Subjects in cluster 3 showed the highest median concentrations of glucose, HbA1c, uric acid, leukocytes, GPT, GOT, GGT, NEFAs, afamin, leptin, cystatin C, hs-CRP, albumin, IL-18, renin, FT4, and insulin (all $p < 0.05$). Cluster 2 was characterized by subjects with higher median values of TC/HDL, TC, LDL, Lp(a), IGF-I, and IGFBP3, but lower median concentrations of ApoA-IV and FT4 compared to the two other clusters (all $p < 0.05$).

Table 3 displays the prevalence of cardiometabolic diseases at F4 and the incidence at the 7 year follow-up FF4, in total and for each cluster. Except for cancer ($p = 0.13$), cluster 3 showed the highest proportions of prevalent cases compared to cluster 2 and cluster 1, respectively. This holds for metabolic diseases (type 2 diabetes [31.7% vs. 3.2% vs. 2.9%], hyperuricemia/gout [13.1% vs. 3.7% vs. 1.6%], dyslipidemia [33.0% vs. 11.7% vs. 13.2%] and hypertension [75.6% vs. 48.0% vs. 29.9%], and all these four diseases together [87.8% vs. 52.5% vs. 36.5%]) as well as for cardiovascular diseases (myocardial infarction [9.3% vs. 2.3% vs. 2.0%], stroke [5.7% vs. 2.3% vs. 1.7%], and these two cardiovascular diseases together [13.9% vs. 4.4% vs. 3.6%]) (all $p < 0.05$). Clusters 1 and 2 showed differences in the prevalence of hypertension, hyperuricemia/gout, and all metabolic diseases together with higher proportions of affected individuals in cluster 2.

Regarding the disease incidence, no differences were observed for stroke ($p = 0.06$) and cancer ($p = 0.24$) between clusters. Again, cluster 3 showed the higher incidence compared to cluster 2 and cluster 1, respectively, regarding hypertension (41.2% vs. 25.3% vs. 18.2%), type 2 diabetes (28.3% vs. 5.1% vs. 2.0%), hyperuricemia/gout (10.8% vs. 2.3% vs. 0.7%), dyslipidemia (19.2% vs. 18.3% vs. 5.6%), all metabolic (54.5% vs. 36.8% vs. 19.7%), and all cardiovascular (6.3% vs. 5.5% vs. 2.3%) diseases together.

4. Discussion

Three metabolically distinct clusters ("metabotypes") were identified in the German population-based KORA F4 study using k-means cluster analysis based on 34 biochemical and anthropometric parameters. One of these clusters (cluster 3) was identified as a "high-risk" cluster with the highest prevalence and incidence of cardiometabolic diseases. Specifically, persons in cluster 3 showed the highest median concentrations of glucose and HbA1c, which are diagnostic criteria for insulin resistance/prediabetes and type 2 diabetes according to the American Diabetes Association.[39] The combination of established risk factors for type 2 diabetes including age, obesity, physical inactivity, hypertension ,and history of cardiovascular diseases[39,40] as well as emerging risk factors for type 2 diabetes such as afamin[41] may explain the finding of the highest type 2 diabetes prevalence and incidence in this cluster. Diabetes, age, obesity, and physical inactivity could be as well associated with the highest hypertension prevalence and incidence in cluster 3. All of

Table 1. Demographic baseline characteristics of the total study population and across the three clusters, KORA F4 study.

	Total	Metabotypes			*p*-value
		Cluster 1	Cluster 2	Cluster 3	
	$N = 1729$	$N = 760$	$N = 616$	$N = 353$	
Socio-demography					
Sex					
Men	839 (48.5)	253 (33.3)[a]	376 (61.0)[b]	210 (59.5)[b]	**<0.0001**
Women	890 (51.5)	507 (66.7)[a]	240 (39.0)[b]	143 (40.5)[b]	
Age in years					
Median (25th,75th)	61.0 (54.0, 68.0)	58.0 (52.0, 66.0)[a]	61.0 (54.0, 67.0)[b]	66.0 (58.0, 72.0)[c]	**<0.0001**
Education					
< 10 years	173 (10.0)	70 (9.2)[a]	54 (8.8)[a]	49 (13.9)[b]	**0.02**
10 to < 12 years	901 (52.2)	385 (50.8)[a]	323 (52.5)[a]	193 (54.7)[b]	
≥ 12 years	652 (37.8)	303 (40.0)[a]	238 (38.7)[a]	111 (31.4)[b]	
Anthropometry					
BMI in kg m^{-2}					
Median (25th, 75th)	27.5 (24.8, 30.7)	25.2 (23.1, 27.7)[a]	28.6 (26.1, 30.8)[b]	31.8 (28.8, 35.6)[c]	**<0.0001**
Underweight	1 (0.1)	1 (0.1)[a]	0 (0.0)[b]	0 (0.0)[c]	**<0.0001**
Normal	456 (26.5)	356 (46.9)[a]	82 (13.3)[b]	18 (5.2)[c]	
Overweight	740 (42.9)	321 (42.3)[a]	324 (52.6)[b]	95 (27.3)[c]	
Obese	526 (30.5)	81 (10.7)[a]	210 (34.1)[b]	235 (67.5)[c]	
Lifestyle					
Physical activity					
Inactive	731 (42.3)	266 (35.0)[a]	273 (44.3)[b]	192 (54.4)[c]	**<0.0001**
Active	997 (57.7)	493 (65.0)[a]	343 (55.7)[b]	161 (45.6)[c]	
Smoking status					
Nonsmoker	765 (44.3)	371 (48.9)[a]	254 (41.2)[b]	140 (39.7)[c]	**0.0002**
Ex-smoker	711 (41.1)	282 (37.2)[a]	255 (41.4)[b]	174 (49.3)[c]	
Smoker	252 (14.6)	106 (14.0)[a]	107 (17.4)[b]	39 (11.0)[c]	
Alcohol consumption					
0 g d^{-1}	507 (29.3)	222 (29.2)[a]	157 (25.5)[b]	128 (36.3)[a]	**0.001**
> 0–< 40/20 g d^{-1} (M/F)	872 (50.5)	369 (48.6)[a]	346 (56.2)[b]	157 (44.5)[a]	
≥ 40/20 g d^{-1} (M/F)	349 (20.2)	168 (22.1)[a]	113 (18.3)[b]	68 (19.3)[a]	

Median (25th, 75th percentile) for continuous variables and *n* (column %) for categorical variables; Different superscript letters between clusters indicate a significant difference between clusters based on Kruskal–Wallis test (with subsequent Kruskal–Wallis post hoc test with Bonferroni correction) for continuous variables and Pearson's chi-squared test (with subsequent chi-squared post hoc test with Bonferroni correction) for categorical variables; Significant results ($p < 0.05$) are highlighted in bold; Due to missing information, reduced datasets for education $n = 1726$, BMI $n = 1723$, physical activity $n = 1728$, smoking status $n = 1728$, and alcohol consumption $n = 1728$; KORA, Cooperative Health Research in the Region of Augsburg.

the above mentioned variables are risk factors for cardiovascular diseases,[42,43] which were also more frequent in this cluster. Although dyslipidemia, which is another risk factor for cardiovascular diseases,[42,43] showed the highest prevalence and incidence in cluster 3, rather low levels of TC/HDL, TC, and LDL were observed in this cluster. This may be explained by the fact that dyslipidemia was assessed by the intake of lipid-lowering medication. This was not observed for hyperuricemia/gout, which was also assessed by the intake of specific medication. As low education is shown to be associated with poor health,[44] the highest proportion of participants with an education period of <10 years in cluster 3 may also have an influence on the highest disease burden in this cluster. The highest percentage of alcohol abstainers and the lowest proportion of current smokers in cluster 3 may be due to the highest disease prevalence in this cluster. Suffering from those diseases may induce lifestyle changes like reducing alcohol consumption and smoking. For smoking, this assumption could be confirmed by the highest number of ex-smokers in this cluster.

Cluster 2 represented an "intermediate-risk" cluster. In particular, the highest values of TC/HDL, TC, and LDL were related to the high dyslipidemia incidence in this cluster, which was almost as high as in cluster 3. Cluster 1 can be termed as a "low-risk" cluster with subjects showing the most favorable study characteristics. Overall, the metabolic differences in subjects between clusters show that we identified metabolically homogeneous subgroups and, thus, valid metabotypes. These metabolic differences can also explain the distinct increase in prevalence and incidence

Table 2. Median baseline values of biochemical and anthropometric parameters of the total study population and across the three clusters, KORA F4 study.

	Total	Metabotypes			*p*-value
		Cluster 1	Cluster 2	Cluster 3	
	$N = 1729$	$N = 760$	$N = 616$	$N = 353$	
Clustering variables					
BMI [kg m^{-2}]	27.6 (24.8, 30.7)	25.2 (23.1, 27.7)[a]	28.6 (26.1, 30.8)[b]	**31.8** (28.8, 35.5)[c]	**<0.0001**
Glucose [mg dL^{-1}]	96 (90, 106)	92 (86, 98)[a]	98 (92, 104)[b]	**114** (101, 133)[c]	**<0.0001**
TC/HDL	3.98 (3.30, 4.78)	3.32 (2.89, 3.73)[a]	**4.86** (4.35, 5.50)[b]	4.20 (3.59, 4.82)[c]	**<0.0001**
TC [mmol L^{-1}]	5.68 (5.01, 6.36)	5.45 (4.83, 6.02)[a]	**6.25** (5.68, 6.93)[b]	5.19 (4.63, 5.79)[c]	**<0.0001**
HbA1c [%]	5.5 (5.3, 5.8)	5.4 (5.2, 5.6)[a]	5.5 (5.3, 5.8)[b]	**5.9** (5.6, 6.6)[c]	**<0.0001**
HDL [mmol L^{-1}]	1.42 (1.19, 1.68)	**1.65** (1.42, 1.91)[a]	1.27 (1.11, 1.45)[b]	1.21 (1.06, 1.42)[b]	**<0.0001**
Uric acid [μmol L^{-1}]	313 (255, 375)	263 (226, 313)[a]	340 (293, 394)[b]	**370** (316, 428)[c]	**<0.0001**
LDL [mmol L^{-1}]	3.57 (3.00, 4.19)	3.28 (2.79, 3.75)[a]	**4.21** (3.70, 4.73)[b]	3.18 (2.71, 3.70)[a]	**<0.0001**
TGs [mmol L^{-1}]	1.26 (0.88, 1.78)	0.88 (0.70, 1.17)[a]	1.65 (1.27, 2.30)[b]	**1.73** (1.20, 2.38)[b]	**<0.0001**
Leukocytes [nL^{-1}]	5.7 (4.8, 6.7)	5.3 (4.6, 6.2)[a]	5.8 (4.9, 6.7)[b]	**6.4** (5.4, 7.6)[c]	**<0.0001**
GPT [μkat L^{-1}]	0.37 (0.27, 0.51)	0.31 (0.24, 0.39)[a]	0.40 (0.31, 0.54)[b]	**0.53** (0.38, 0.78)[c]	**<0.0001**
GOT [μkat L^{-1}]	0.42 (0.36, 0.50)	0.39 (0.34, 0.46)[a]	0.43 (0.37, 0.50)[b]	**0.49** (0.41, 0.64)[c]	**<0.0001**
GGT [μkat L^{-1}]	0.47 (0.31, 0.76)	0.35 (0.25, 0.52)[a]	0.54 (0.38, 0.84)[b]	**0.70** (0.46, 1.19)[c]	**<0.0001**
AP [μkat L^{-1}]	1.13 (0.94, 1.35)	1.04 (0.88, 1.24)[a]	1.20 (1.00, 1.40)[b]	**1.24** (1.06, 1.47)[b]	**<0.0001**
Average telomere length in leukocytes (ratio of the telomere repeat copy number to a single copy gene)	1.77 (1.59, 1.98)	**1.82** (1.63, 2.03)[a]	1.75 (1.57, 1.95)[b]	1.70 (1.52, 1.93)[b]	**<0.0001**
NEFAs [mg dL^{-1}]	6.92 (5.34, 8.74)	6.69 (5.19, 8.58)[a]	6.71 (5.16, 8.37)[a]	**7.92** (6.17, 9.92)[b]	**<0.0001**
Lp(a) [mg dL^{-1}]	12.2 (5.5, 31.7)	11.9 (5.5, 28.3)[a]	**14.6** (6.8, 38.2)[b]	9.1 (3.5, 25.4)[c]	**<0.0001**
ApoA-IV [mg dL^{-1}]	15.1 (12.6, 17.9)	15.5 (12.7, 18.0)[a]	14.7 (12.4, 17.0)[b]	**15.7** (12.9, 19.5)[a]	**<0.0001**
Afamin [mg L^{-1}]	71.4 (61.3, 82.8)	63.5 (55.9, 72.7)[a]	75.0 (65.3, 85.5)[b]	**83.5** (72.4, 97.1)[c]	**<0.0001**
Leptin [ng mL^{-1}]	13.4 (6.1, 26.5)	10.6 (4.8, 19.8)[a]	12.6 (6.3, 25.7)[b]	**24.1** (12.5, 46.9)[c]	**<0.0001**
TPO antibodies [IU mL^{-1}]	12 (11, 15)	12 (11, 15)[a]	**13** (11, 15)[a]	12 (11, 14)[a]	0.08
Cystatin C [mg L^{-1}]	0.74 (0.68, 0.83)	0.71 (0.65, 0.78)[a]	0.75 (0.69, 0.83)[b]	**0.84** (0.74, 0.95)[c]	**<0.0001**
hs-CRP [mg L^{-1}]	1.28 (0.63, 2.65)	0.85 (0.44, 1.56)[a]	1.51 (0.77, 2.84)[b]	**2.47** (1.17, 4.75)[c]	**<0.0001**
Albumin [mg L^{-1}]	7.5 (3.9, 16.8)	6.4 (3.4, 13.3)[a]	7.5 (4.0, 16.4)[b]	**13.0** (5.6, 29.6)[c]	**<0.0001**
Creatinine [g L^{-1}]	1.31 (0.84, 1.92)	1.21 (0.77, 1.80)[a]	**1.49** (0.89, 2.02)[b]	1.44 (0.96, 2.00)[b]	**<0.0001**
IL-18 [pg mL^{-1}]	299 (231, 389)	262 (206, 334)[a]	305 (248, 389)[b]	**386** (302, 483)[c]	**<0.0001**
IGF-1 [ng mL^{-1}]	126 (100, 156)	126 (102, 156)[a]	**133** (107, 167)[b]	109 (80, 137)[c]	**<0.0001**
Renin [μIU mL^{-1}]	10.8 (5.7, 19.1)	9.2 (5.0, 15.7)[a]	10.6 (6.3, 18.3)[b]	**16.3** (7.9, 39.8)[c]	**<0.0001**
Aldosterone [pg mL^{-1}]	38 (26, 56)	38 (26, 56)[a]	38 (26, 56)[a]	**42** (26, 62)[a]	**0.21**
IGFBP3 [ng mL^{-1}]	3390 (2840, 3970)	3260 (2768, 3770)[a]	**3700** (3170, 4240)[b]	3092 (2420, 3880)[a]	**<0.0001**
SHBG [nmol L^{-1}]	30.4 (23.1, 42.1)	**37.6** (28.1, 50.0)[a]	26.7 (19.8, 34.7)[b]	26.8 (20.5, 35.2)[b]	**<0.0001**
TSH [mIU L^{-1}]	1.23 (0.81, 1.80)	1.19 (0.76, 1.73)[a]	1.24 (0.83, 1.80)[a,b]	**1.32** (0.83, 1.98)[b]	**0.01**
FT4 [pmol L^{-1}]	14.0 (12.7, 15.4)	14.0 (12.8, 15.4)[a]	13.7 (12.5, 15.1)[b]	**14.4** (12.9, 16.1)[c]	**<0.0001**
Insulin [μIU mL^{-1}]	4.4 (2.9, 7.9)	3.3 (2.2, 4.7)[a]	4.8 (3.5, 7.9)[b]	**10.7** (5.3, 21.0)[c]	**<0.0001**

Median (25th, 75th percentile) calculated by the means of medians, 25th and 75th percentiles over all five imputed datasets; Kruskal–Wallis test (and Kruskal–Wallis post hoc test with Bonferroni correction) was calculated for each of the five imputed datasets and the *p*-value was determined from the mean of the Kruskal–Wallis test statistics over the five imputed datasets; Significant results ($p < 0.05$) are highlighted. Different superscript letters between clusters indicate a significant difference between clusters, whereas the same superscript letters between clusters indicate no significant difference between clusters; Underlined values represent the lowest value across the clusters; the **bold** values represent the highest; AP, alkaline phosphatase; ApoA-IV, apolipoprotein A-IV; FT4, free thyroxine; GGT, gamma-glutamyltransferase; GOT, glutamate-oxaloacetate transaminase; GPT, glutamate-pyruvate transaminase; HbA1c, glycated hemoglobin; hs-CRP, high-sensitive C-reactive protein; IGF-I, insulin-like growth factor-I; IGFBP3, insulin-like-growth-factor-binding-protein-3; KORA, Cooperative Health Research in the Region of Augsburg; Lp(a), lipoprotein(a); NEFAs, nonesterified fatty acids; SHBG, sex-hormone-binding globulin; TC, total cholesterol; TGs, triglycerides; TPO, thyroperoxidase; TSH, thyroid-stimulating hormone.

Table 3. Disease prevalence and incidence in the total study population and across the three clusters, KORA F4/FF4 study.

	Total	Metabotypes			*p*-value
		Cluster 1	Cluster 2	Cluster 3	
	N = 1729	*N* = 760	*N* = 616	*N* = 353	
Prevalence of diseases in KORA F4			% (*n*)		
Hypertension	45.7 (789)	29.9 (227)[a]	48.0 (295)[b]	75.6 (267)[c]	**<0.0001**
Type 2 diabetes	8.9 (154)	2.9 (22)[a]	3.2 (20)[a]	31.7 (112)[b]	**<0.0001**
Hyperuricemia/gout	4.7 (81)	1.6 (12)[a]	3.7 (23)[b]	13.1 (46)[c]	**<0.0001**
Dyslipidemia	16.7 (288)	13.2 (100)[a]	11.7 (72)[a]	33.0 (116)[b]	**<0.0001**
Any of the above metabolic diseases	52.7 (909)	36.5 (277)[a]	52.5 (323)[b]	87.8 (309)[c]	**<0.0001**
Myocardial infarction	3.6 (62)	2.0 (15)[a]	2.3 (14)[a]	9.3 (33)[b]	**<0.0001**
Stroke	2.7 (47)	1.7 (13)[a]	2.3 (14)[a]	5.7 (20)[b]	**0.001**
Any of the above cardiovascular diseases	6.0 (103)	3.6 (27)[a]	4.4 (27)[a]	13.9 (49)[b]	**<0.0001**
Cancer	9.2 (159)	8.8 (67)[a]	8.1 (50)[a]	11.9 (42)[a]	0.13
Incidence of diseases in KORA FF4			% (*n*)		
Hypertension	22.7 (166)	18.2 (76)[a]	25.3 (62)[b]	41.2 (28)[c]	**<0.0001**
Type 2 diabetes	6.8 (78)	2.0 (11)[a]	5.1 (22)[b]	28.3 (45)[c]	**<0.0001**
Hyperuricemia/gout	2.9 (35)	0.7 (4)[a]	2.3 (10)[a]	10.8 (21)[b]	**<0.0001**
Dyslipidemia	12.4 (131)	5.6 (28)[a]	18.3 (74)[b]	19.2 (29)[b]	**<0.0001**
Any of the above metabolic diseases	27.6 (177)	19.7 (75)[a]	36.8 (84)[b]	54.5 (18)[b]	**<0.0001**
Myocardial infarction	2.0 (24)	0.9 (5)[a]	3.2 (14)[b]	2.5 (5)[a,b]	**0.02**
Stroke	2.4 (29)	1.4 (8)[a]	2.7 (12)[a]	4.2 (9)[a]	0.06
Any of the above cardiovascular diseases	4.1 (49)	2.3 (13)[a]	5.5 (24)[b]	6.3 (12)[b]	**0.01**
Cancer	6.8 (77)	7.6 (40)[a]	5.1 (21)[a]	8.0 (16)[a]	0.24

Column % (*n*); Pearson's chi-squared test (Fisher's exact test if expected frequencies were too low) and chi-squared post hoc test; Significant results ($p < 0.05$) are highlighted. Different superscript letters between clusters indicate a significant difference between clusters, whereas the same superscript letters between clusters indicate no significant difference between clusters; Prevalence: Due to missing information, reduced datasets for all metabolic diseases $n = 1725$, hypertension $n = 1726$, and all other diseases $n = 1728$; Incidence: Due to missing information, reduced datasets for hypertension $n = 730$, type 2 diabetes $n = 1153$, hyperuricemia/gout $n = 1190$, dyslipidemia $n = 1059$, all metabolic diseases $n = 641$, myocardial infarction $n = 1201$, stroke $n = 1222$, all cardiovascular diseases $n = 1183$, and cancer $n = 1136$; KORA, Cooperative Health Research in the Region of Augsburg.

of metabolic and cardiovascular diseases from cluster 1 to cluster 3. It is interesting that we found such a clear distinction in disease prevalence and incidence between clusters using an unsupervised clustering method and that we identified metabotypes that differentiate subjects with regard to their cardiometabolic disease risk. The greatest differences in the incidence between clusters were found for type 2 diabetes and hyperuricemia/gout. Compared to clusters 1 and 2, respectively, participants in cluster 3 showed a 14.2-fold and a 5.5-fold higher incidence of type 2 diabetes as well as a 15.4-fold and a 4.7-fold higher incidence of hyperuricemia/gout.

There have been numerous studies establishing metabotypes by statistical methods,[6,14] proposing that cluster analysis seems to be a suitable method. As we aimed to identify general metabotypes initially independent of diseases, we preferred such an unsupervised learning algorithm to supervised methods like regression analysis, which are optimal for disease prediction. Most previous studies grouped individuals based on a few metabolites, mainly parameters of lipid and carbohydrate metabolism,[7,15–27] or single specific metabolite classes like plasma lipoproteins[28,29] or fatty acids.[30] Few studies have included several parameters originating from different metabolic pathways,[31,45] extensive metabolic data,[32,46,47] or "-omics" data,[48] whereas anthropometric data were included in many studies in the clustering process.[18–20,24,25,27,31,47] We expected that such extensive data may provide a more comprehensive characterization of individuals in the metabotyping process and, thus, a more refined and generalized metabotype definition. Our study is among the first to use metabotyping on such an extensive range of phenotypic data.

We described metabolic differences between the metabotypes defined in detail. Our study is the first to investigate prevalence and, especially, incidence of various cardiometabolic diseases across clusters. This allowed us to identify disease-related metabotypes. Studies examining the prevalence were limited to at most three diseases, mainly diabetes, hypertension, or the metabolic syndrome,[15–20,28,29,31] and only one previous study examined the risk of incident hypertension.[32] Van Bochove et al.,[28] for example, identified three plasma lipoprotein clusters with varying lipid response to fenofibrate therapy. They presented, inter alia, prevalence of diabetes and hypertension across clusters, but these showed smaller differences between the clusters than in our study. This is possibly due to using lipoproteins in the clustering process only.

The identification of comprehensive metabotypes associated with diet-related diseases offers the opportunity to develop and establish targeted dietary recommendations based on this metabotype group level. This may be especially relevant for cluster 3. Furthermore, as not only physically inactive participants or participants in higher age groups or with a high BMI are assigned to this cluster, we could also identify additional individuals at high cardiometabolic risk through metabotyping. These individuals may not have been identified as "at risk" otherwise, although they might benefit from targeted prevention strategies. As shown in previous studies,[45–47] selected diet–disease associations stratified by metabotype should be investigated to identify differing responsiveness to diet across metabotypes. This may lead to targeted recommendations to responsive subgroups that mostly benefit from this type of intervention, see, for example, refs. 45 and 47. Further, these metabotypes may also be relevant in the development of other preventive lifestyle approaches like targeted physical activity interventions.

We defined metabotypes in a large population-based sample. Thus, our results may be generalizable to the adult population in Southern Germany. However, these findings should be replicated in other cohorts. An extensive set of biochemical and anthropometric parameters was used to get a precise characterization of individuals and, thus, comprehensive and valid metabotypes for the development and establishment of targeted and precise disease prevention strategies.[14] These parameters were only measured once, not considering intraindividual variation over time due to, for example, stress or subclinical diseases.[1,49,50] As these differences were shown to be smaller than interindividual differences,[51] we do not expect our results to change when including repeated measurements. Further, it was shown that some metabolic differences between individuals are visible only through metabolic challenge tests like an oral glucose tolerance test, and, therefore, would remain undetected using fasting blood values as we did in our study.[52] To date, only a few studies defined metabotypes based on data of metabolic response to metabolic challenge tests and interventions.[7,26,48,53,54] Usually, not all fasting parameters used for metabotyping in the present study and even less data of metabolic response to metabolic challenge tests are routinely measured in the clinical setting. The future assignment of whole populations to the metabotypes defined is therefore difficult. For this, we propose to establish a risk score with a reduced set of the biochemical and anthropometric parameters used here, achieving a similar assignment of individuals to metabotypes to enable the final implementation of these results in whole populations. To describe the clustering result by disease incidence between metabotypes, a relatively long follow-up duration of 7 years was available. A limitation in the assessment of the disease status at baseline and follow-up is, however, that it was mostly based on self-report, thus we cannot rule out misreporting. Information on dyslipidemia and hyperuricemia/gout was only based on medication intake, therefore, possibly leading to underestimating the true disease prevalence and incidence.

In conclusion, cluster analysis based on an extensive set of biochemical and anthropometric parameters allowed the identification of three valid and comprehensive metabotypes showing a different disease incidence in a 7 year follow-up. Further studies are needed to examine the reproducibility of our clustering result in other cohorts and to test these newly identified metabotypes for differing dietary responses with respect to the cardiometabolic risk variables. If successful, targeted and precise dietary recommendations could be developed on the metabotype level for cardiometabolic disease prevention.

Acknowledgements

A.R. conceived and conducted the data analyses, interpreted the data, and wrote the manuscript; N.W. and C.G. contributed to data analyses and interpretation, and revised the manuscript; C.M., A.P., M.Ro., F.K., C.H., W.R., H.V., M.Re., W.K., and H.W. were involved in the study organization, provided data, and reviewed the manuscript; H.H. and H.D. contributed to data interpretation and revised the manuscript; J.L. conceived and designed the data analyses, contributed to data analyses and interpretation, and revised the manuscript. All authors have read and approved the final manuscript.

The Cooperative Health Research in the Region of Augsburg (KORA) studies were initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. The German Diabetes Center was supported by the Ministry of Culture and Science of the State of North Rhine-Westphalia and the German Federal Ministry of Health. This study was supported in part by a grant from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD). The preparation of this paper was supported by the *enable* Cluster and is catalogued by the *enable* Steering Committee as *enable* 011 (http://enable-cluster.de). This work was funded by a grant of the German Ministry for Education and Research (BMBF) FK 01EA1409E.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

cardiometabolic disease, cluster analysis, *enable*-Cluster, metabolic phenotype, metabotype

[1] E. Holmes, I. D. Wilson, J. K. Nicholson, *Cell* **2008**, *134*, 714.
[2] J. K. Nicholson, *Mol. Syst. Biol.* **2006**, *2*, 52.
[3] R. D. Beger, W. Dunn, M. A. Schmidt, S. S. Gross, J. A. Kirwan, M. Cascante, L. Brennan, D. S. Wishart, M. Oresic, T. Hankemeier, D. I. Broadhurst, A. N. Lane, K. Suhre, G. Kastenmüller, S. J. Sumner, I. Thiele, O. Fiehn, R. Kaddurah-Daouk, *Metabolomics* **2016**, *12*, 149.
[4] B. de Roos, *Proc. Nutr. Soc.* **2013**, *72*, 48.
[5] J. K. Nicholson, E. Holmes, *Discov. Med.* **2006**, *6*, 63.

[6] C. B. O'Donovan, M. C. Walsh, M. J. Gibney, E. R. Gibney, L. Brennan, *Proc. Nutr. Soc.* **2016**, *75*, 106.
[7] C. Morris, C. O'Grada, M. Ryan, H. M. Roche, M. J. Gibney, E. R. Gibney, L. Brennan, *PLoS One* **2013**, *8*, e72890.
[8] J. K. Nicholson, E. Holmes, J. M. Kinross, A. W. Darzi, Z. Takats, J. C. Lindon, *Nature* **2012**, *491*, 384.
[9] J. Kaput, *Curr. Opin. Biotechnol.* **2008**, *19*, 110.
[10] L. Brennan, *Proc. Nutr. Soc.* **2008**, *67*, 404.
[11] L. Brennan, *Curr. Opin. Biotechnol.* **2016**, *44*, 35.
[12] C. Celis-Morales, K. M. Livingstone, C. F. Marsaux, A. L. Macready, R. Fallaize, C. B. O'Donovan, C. Woolhead, H. Forster, M. C. Walsh, S. Navas-Carretero, R. San-Cristobal, L. Tsirigoti, C. P. Lambrinou, C. Mavrogianni, G. Moschonis, S. Kolossa, J. Hallmann, M. Godlewska, A. Surwillo, I. Traczyk, C. A. Drevon, J. Bouwman, B. van Ommen, K. Grimaldi, L. D. Parnell, J. N. Matthews, Y. Manios, H. Daniel, J. A. Martinez, J. A. Lovegrove, E. R. Gibney, L. Brennan, W. H. Saris, M. Gibney, J. C. Mathers, Food4Me Study, *Int. J. Epidemiol.* **2017**, *46*, 578.
[13] K. M. Livingstone, C. Celis-Morales, S. Navas-Carretero, R. San-Cristobal, A. L. Macready, R. Fallaize, H. Forster, C. Woolhead, C. B. O'Donovan, C. F. Marsaux, S. Kolossa, L. Tsirigoti, C. P. Lambrinou, G. Moschonis, M. Godlewska, A. Surwillo, C. A. Drevon, Y. Manios, I. Traczyk, E. R. Gibney, L. Brennan, M. C. Walsh, J. A. Lovegrove, W. H. Saris, H. Daniel, M. Gibney, J. A. Martinez, J. C. Mathers, Food4Me Study, *Am. J. Clin. Nutr.* **2016**, *104*, 288.
[14] A. Riedl, C. Gieger, H. Hauner, H. Daniel, J. Linseisen, *Br. J. Nutr.* **2017**, *117*, 1631.
[15] C. B. O'Donovan, M. C. Walsh, A. P. Nugent, B. McNulty, J. Walton, A. Flynn, M. J. Gibney, E. R. Gibney, L. Brennan, *Mol. Nutr. Food Res.* **2015**, *59*, 377.
[16] N. Zubair, C. W. Kuzawa, T. W. McDade, L. S. Adair, *Asia Pac. J. Clin. Nutr.* **2012**, *21*, 271.
[17] N. Zubair, C. W. Kuzawa, N. R. Lee, T. W. McDade, L. S. Adair, *Asia Pac. J. Clin. Nutr.* **2014**, *23*, 148.
[18] M. A. Wilcox, D. F. Wyszynski, C. I. Panhuysen, Q. Ma, A. Yip, J. Farrell, L. A. Farrer, *BMC Genet.* **2003**, *4*, S15.
[19] M. Wilcox, Q. Li, Y. Sun, P. Stang, J. Berlin, D. Wang, *BMC Proc.* **2009**, *3*, S53.
[20] C. R. Tzeng, Y. C. Chang, Y. C. Chang, C. W. Wang, C. H. Chen, M. I. Hsu, *Fertil. Steril.* **2014**, *101*, 1404.
[21] V. Bermúdez, J. Rojas, J. Salazar, R. Anez, A. Toledo, L. Bello, V. Apruzzese, R. González, M. Chacín, M. Cabrera, C. Cano, M. Velasco, J. López-Miranda, *J. Diabetes Res.* **2015**, *2015*, 750265.
[22] R. N. Baumgartner, R. M. Siervogel, A. F. Roche, *Am. J. Hum. Biol.* **1989**, *1*, 43.
[23] C. M. Bucci, W. E. Legnani, R. L. Armentano, *J. Phys. Conf. Ser.* **2016**, *705*, 012033.
[24] R. C. Huang, T. A. Mori, V. Burke, J. Newnham, F. J. Stanley, L. I. Landau, G. E. Kendall, W. H. Oddy, L. J. Beilin, *Diabetes Care* **2009**, *32*, 695.
[25] A. K. Ventura, E. Loken, L. L. Birch, *Pediatrics* **2006**, *118*, 2434.
[26] S. Krishnan, J. W. Newman, T. A. Hembrooke, N. L. Keim, *Nutr. Metab.* **2012**, *9*, 26.
[27] R. Micciolo, *Hum. Biol.* **1992**, *64*, 539.
[28] K. van Bochove, D. B. van Schalkwijk, L. D. Parnell, C. Q. Lai, J. M. Ordovás, A. A. de Graaf, B. van Ommen, D. K. Arnett, *PLoS One* **2012**, *7*, e38072.
[29] A. C. Frazier-Wood, S. Glasser, W. T. Garvey, E. K. Kabagambe, I. B. Borecki, H. K. Tiwari, M. Y. Tsai, P. N. Hopkins, J. M. Ordovas, D. K. Arnett, *Lipids Health Dis.* **2011**, *10*, 237.
[30] K. Li, L. Brennan, B. A. McNulty, J. F. Bloomfield, D. J. Duff, N. F. Devlin, M. J. Gibney, A. Flynn, J. Walton, A. P. Nugent, *Mol. Nutr. Food Res.* **2016**, *60*, 2043.
[31] P. A. Andreeva-Gateva, V. D. Simeonov, R. T. Georrgieva-Nikolova, R. K. Tafradjiiska-Hadjiolova, *Endocrinologia* **2014**, *19*, 168.
[32] W. Qureshi, L. Wagenknecht, S. Watkins, F. Chilton, J. Rotter, L. Carlos, D. Herrington, *Circulation* **2014**, *129*, A23.
[33] R. Holle, M. Happich, H. Löwel, H. E. Wichmann, MONICA/KORA Study Group, *Gesundheitswesen* **2005**, *67*, S19.
[34] W. Rathmann, B. Haastert, A. Icks, H. Löwel, C. Meisinger, R. Holle, G. Giani, *Diabetologia* **2003**, *46*, 182.
[35] R. M. Cawthon, *Nucleic Acids Res.* **2002**, *30*, e47.
[36] S. van Buuren, K. Groothuis-Oudshoorn, *J. Stat. Softw.* **2011**, *45*, 1.
[37] X. Basagana, J. Barrera-Gómez, M. Benet, J. M. Antó, J. Garcia-Aymerich, *Am. J. Epidemiol.* **2013**, *177*, 718.
[38] M. Charrad, N. Ghazzali, V. Boiteau, A. Niknafs, *J. Stat. Softw.* **2014**, *61*, 1.
[39] S. Genuth, K. G. Alberti, P. Bennett, J. Buse, R. Defronzo, R. Kahn, J. Kitzmiller, W. C. Knowler, H. Lebovitz, A. Lernmark, D. Nathan, J. Palmer, R. Rizza, C. Saudek, J. Shaw, M. Steffes, M. Stern, J. Tuomilehto, P. Zimmet, *Diabetes Care* **2003**, *26*, 3160.
[40] World Health Organization, *Global Report on Diabetes*, World Health Organization, Geneva **2016**.
[41] B. Kollerits, C. Lamina, C. Huth, P. Marques-Vidal, S. Kiechl, I. Seppälä, J. Cooper, S. C. Hunt, C. Meisinger, C. Herder, L. Kedenko, J. Willeit, B. Thorand, D. Dähnhardt, D. Stöckl, K. Willeit, M. Roden, W. Rathmann, B. Paulweber, A. Peters, M. Kähönen, T. Lehtimäki, O. T. Raitakari, S. E. Humphries, P. Vollenweider, H. Dieplinger, F. Kronenberg, *Diabetes Care* **2017**, *40*, 1386.
[42] World Health Organization, *Cardiovascular Diseases (CVDs)*, World Health Organization, Geneva **2017**. http://www.who.int/mediacentre/factsheets/fs317/en/. Accessed December 1, 2017.
[43] World Health Organization, *A Global Brief on Hypertension*, World Health Organization, Geneva **2013**.
[44] World Health Organization, *The Determinants of Health*, World Health Organization, Geneva **2018**. http://www.who.int/hia/evidence/doh/en/. Accessed May 7, 2018.
[45] A. O'Sullivan, M. J. Gibney, A. O. Connor, B. Mion, S. Kaluskar, K. D. Cashman, A. Flynn, F. Shanahan, L. Brennan, *Mol. Nutr. Food Res.* **2011**, *55*, 679.
[46] A. A. Moazzami, A. Shrestha, D. A. Morrison, K. Poutanen, H. Mykkänen, *J. Nutr.* **2014**, *144*, 807.
[47] R. Vázquez-Fresno, R. Llorach, A. Perera, R. Mandal, M. Feliz, F. J. Tinahones, D. S. Wishart, C. Andres-Lacueva, *J. Nutr. Biochem.* **2016**, *28*, 114.
[48] J. Bouwman, J. T. Vogels, S. Wopereis, C. M. Rubingh, S. Bijlsma, B. Ommen, *BMC Med. Genomics* **2012**, *5*, 1.
[49] S. Rezzi, F. P. Martin, C. Alonso, M. Guilarte, M. Vicario, L. Ramos, C. Martínez, B. Lobo, E. Saperas, J. R. Malagelada, J. Santos, S. Kochhar, *J. Proteome Res.* **2009**, *8*, 4799.
[50] V. Ghini, E. Saccenti, L. Tenori, M. Assfalg, C. Luchinat, *J. Proteome Res.* **2015**, *14*, 2951.
[51] M. Assfalg, I. Bertini, D. Colangiuli, C. Luchinat, H. Schäfer, B. Schütz, M. Spraul, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 1420.
[52] S. Krug, G. Kastenmüller, F. Stückler, M. J. Rist, T. Skurk, M. Sailer, J. Raffler, W. Römisch-Margl, J. Adamski, C. Prehn, T. Frank, K. H. Engel, T. Hofmann, B. Luy, R. Zimmermann, F. Moritz, P. Schmitt-Kopplin, J. Krumsiek, W. Kremer, F. Huber, U. Oeh, F. J. Theis, W. Szymczak, H. Hauner, K. Suhre, H. Daniel, *FASEB J.* **2012**, *26*, 2607.
[53] T. T. Wang, A. J. Edwards, B. A. Clevidence, *J. Nutr. Biochem.* **2013**, *24*, 1538.
[54] E. C. Chua, G. Shui, I. T. Lee, P. Lau, L. C. Tan, S. C. Yeo, B. D. Lam, S. Bulchand, S. A. Summers, K. Puvanendran, S. G. Rozen, M. R. Wenk, J. J. Gooley, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14468.