Effects of Metformin on Metabolite Profiles and LDL Cholesterol in Patients With Type 2 Diabetes

DOI: 10.2337/dc15-0658

Tao Xu,[1,2] Stefan Brandmaier,[1,2] Ana C. Messias,[3] Christian Herder,[4,5] Harmen H.M. Draisma,[6,7,8] Ayse Demirkan,[9,10] Zhonghao Yu,[1,2] Janina S. Ried,[11] Toomas Haller,[12] Margit Heier,[2] Monica Campillos,[13] Gisela Fobo,[13] Renee Stark,[14] Christina Holzapfel,[15] Jonathan Adam,[1,2] Shen Chi,[1,2] Markus Rotter,[1,2] Tommaso Panni,[2] Anne S. Quante,[11,16] Ying He,[17,18] Cornelia Prehn,[19] Werner Roemisch-Margl,[13] Gabi Kastenmüller,[13] Gonneke Willemsen,[6,7] René Pool,[6,7] Katarina Kasa,[9] Ko Willems van Dijk,[10] Thomas Hankemeier,[20] Christa Meisinger,[2] Barbara Thorand,[2] Andreas Ruepp,[13] Martin Hrabé de Angelis,[21,22,23] Yixue Li,[17,18] H.-Erich Wichmann,[2,16,24] Bernd Stratmann,[25] Konstantin Strauch,[11,16] Andres Metspalu,[12] Christian Gieger,[1,2] Karsten Suhre,[13,26,27] Jerzy Adamski,[19,22,23] Thomas Illig,[1,28,29] Wolfgang Rathmann,[30] Michael Roden,[4,5,31] Annette Peters,[1,2,23,32] Cornelia M. van Duijn,[8,33] Dorret I. Boomsma,[6,7] Thomas Meitinger,[34,35] and Rui Wang-Sattler[1,2,23]

[1]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany
[2]Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany
[3]Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany
[4]Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University, Düsseldorf, Germany

Corresponding author: Rui Wang-Sattler, rui.wang-sattler@helmholtz-muenchen.de.

Received 30 March 2015 and accepted 24 June 2015.

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc15-0658/-/DC1.

T.X., S.B., A.C.M., C.He., H.H.M.D., and A.D. contributed equally. C.M.v.D. and D.I.B. contributed equally.

OBJECTIVE

Metformin is used as a first-line oral treatment for type 2 diabetes (T2D). However, the underlying mechanism is not fully understood. Here, we aimed to comprehensively investigate the pleiotropic effects of metformin.

RESEARCH DESIGN AND METHODS

We analyzed both metabolomic and genomic data of the population-based KORA cohort. To evaluate the effect of metformin treatment on metabolite concentrations, we quantified 131 metabolites in fasting serum samples and used multivariable linear regression models in three independent cross-sectional studies (n = 151 patients with T2D treated with metformin [mt-T2D]). Additionally, we used linear mixed-effect models to study the longitudinal KORA samples (n = 912) and performed mediation analyses to investigate the effects of metformin intake on blood lipid profiles. We combined genotyping data with the identified metformin-associated metabolites in KORA individuals (n = 1,809) and explored the underlying pathways.

RESULTS

We found significantly lower (P < 5.0E-06) concentrations of three metabolites (acyl-alkyl phosphatidylcholines [PCs]) when comparing mt-T2D with four control groups who were not using glucose-lowering oral medication. These findings were controlled for conventional risk factors of T2D and replicated in two independent studies. Furthermore, we observed that the levels of these metabolites decreased significantly in patients after they started metformin treatment during 7 years' follow-up. The reduction of these metabolites was also associated with a lowered blood level of LDL cholesterol (LDL-C). Variations of these three metabolites were significantly associated with 17 genes (including *FADS1* and *FADS2*) and controlled by AMPK, a metformin target.

CONCLUSIONS

Our results indicate that metformin intake activates AMPK and consequently suppresses FADS, which leads to reduced levels of the three acyl-alkyl PCs and LDL-C. Our findings suggest potential beneficial effects of metformin in the prevention of cardiovascular disease.

Type 2 diabetes (T2D) is a chronic disease with diminished response to insulin and relative insulin deficiency (1). Patients with T2D mostly take metformin as first-line oral treatment to lower their glucose levels and to improve insulin sensitivity (2). Despite metformin's use as an antihyperglycemic agent for more than 50 years, its

primary mode of action is not yet completely understood (3). Inside a cell, metformin apparently inhibits complex I of the mitochondrial electron transport chain and thereby reduces the cellular energy status and upregulates the cytoplasmic 5′-AMPK pathway (3). Activated AMPK stimulates catabolic processes (glycolysis and fatty acid oxidation) and inhibits anabolic pathways (gluconeogenesis and fatty acid synthesis). So far, six metformin targets are documented in the DrugBank (4) database, including the AMPK complex and five metformin transporters. Furthermore, metformin was reported to have several possible pleiotropic effects, resulting in reduced risks for both cancer (5) and cardiovascular disease (CVD) (6), as well as reduced levels of LDL cholesterol (LDL-C) (7,8).

Metabolomic studies have detected metabolite profile changes during the development of T2D (9–12) and identified concentration differences caused by various physiological and environmental factors such as age (13), sex (14), smoking status (15), and alcohol consumption (16). Several metabolomic studies attempted to unravel the physiological effects of metformin (17–21). However, they either used technologies covering only small sets of metabolites or examined relatively few participants (e.g., 20 healthy volunteers [18], 15 patients [17,19], 31 patients [20], and 24 patients treated with glipizide and 23 patients with metformin [21]). As interindividual genetic variations contribute to diverse metabolite profiles and different drug responses, combining metabolomics and genomics may help to understand the mechanisms underlying the action of medications (22–25).

In this study, we discovered metformin treatment–associated metabolites in the Cooperative Health Research in the Region of Augsburg (KORA) cohort (26,27). We confirmed our finding in longitudinal KORA data and replicated them in two independent studies: the Erasmus Rucphen Family study (ERF) (28) and the Netherlands Twin Register (NTR) (29). The biologically relevant pathways for the identified metabolites and their associated genes were further analyzed in organ-specific protein-metabolite interaction networks (30,31). Additionally, we assessed the effects of metformin treatment on LDL-C levels.

RESEARCH DESIGN AND METHODS

An overview of the analysis work flow is shown in Fig. 1.

Ethics Statement

All participants gave written informed consent. The KORA study was approved by the ethics committee of the Bavarian Medical Association, Germany; the ERF study by the medical ethics board of the Erasmus MC Rotterdam, the Netherlands; and the NTR study by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center, Amsterdam, the Netherlands.

KORA Cohort

KORA is a population-based cohort study conducted in Southern Germany (26). The baseline survey 4 (KORA S4) consists of 4,261 individuals (aged 25–74 years) examined between 1999 and 2001. During the years 2006–2008, 3,080 participants took part in the follow-up survey 4 (KORA F4). Clinical data for each participant were retrieved from medical records. Based on physician-validated and self-reported diagnosis (9,26), fasting glucose and 2-h postglucose load, and information on medications (Table 1), we excluded *1*) patients suffering from type 1 and steroid-induced diabetes (n = 9), *2*) patients with T2D treated with both metformin and insulin (n = 15), *3*) patients taking glucose-lowering oral medication other than metformin (n = 25), and *4*) patients lacking clear information on treatment (n = 1). Furthermore, participants with overnight nonfasting blood samples (n = 16) or isolated impaired fasting glucose (n = 112) were excluded. We previously showed that impaired fasting glucose and impaired glucose tolerance (IGT) should be considered two different phenotypes (9). In KORA F4, we focused on five groups: *1*) patients

[5]German Center for Diabetes Research, Düsseldorf, Germany
[6]Department of Biological Psychology, Faculty of Psychology and Education, VU University Amsterdam, Amsterdam, the Netherlands
[7]EMGO Institute for Health and Care Research, Amsterdam, the Netherlands
[8]Neuroscience Campus Amsterdam, Amsterdam, the Netherlands
[9]Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands
[10]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands
[11]Institute of Genetic Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany
[12]Estonian Genome Center, University of Tartu, Tartu, Estonia
[13]Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, Neuherberg, Germany
[14]Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, Neuherberg, Germany
[15]Else Kroener-Fresenius-Center for Nutritional Medicine, Faculty of Medicine, Technische Universität München, Munich, Germany
[16]Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany
[17]Shanghai Center for Bioinformation Technology, Shanghai, China
[18]Bioinformatics Center, Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China
[19]Institute of Experimental Genetics, Genome Analysis Center, Helmholtz Zentrum München, Neuherberg, Germany
[20]Faculty of Science, Leiden Academic Centre for Drug Research, Analytical BioSciences, the Netherlands
[21]Institute of Experimental Genetics, Helmholtz Zentrum München, Neuherberg, Germany
[22]Chair of Experimental Genetics, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Freising, Germany
[23]German Center for Diabetes Research, Neuherberg, Germany
[24]Institute of Medical Statistics and Epidemiology, Technische Universität München, Munich, Germany
[25]Heart and Diabetes Center NRW, Diabetes Center, Ruhr-University Bochum, Bad Oeynhausen, Germany
[26]Faculty of Biology, Ludwig-Maximilians-Universität, Planegg-Martinsried, Germany
[27]Department of Physiology and Biophysics, Weill Cornell Medical College in Qatar, Education City, Qatar Foundation, Doha, Qatar
[28]Hannover Unified Biobank, Hannover Medical School, Hannover, Germany
[29]Institute for Human Genetics, Hannover Medical School, Hanover, Germany
[30]Institute for Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University, Düsseldorf, Germany
[31]Department of Endocrinology and Diabetology, Medical Faculty, Düsseldorf, Germany
[32]Department of Environmental Health, Harvard School of Public Health, Boston, MA
[33]Center for Medical Systems Biology, Leiden, the Netherlands
[34]Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany
[35]Institute of Human Genetics, Technische Universität München, Munich, Germany

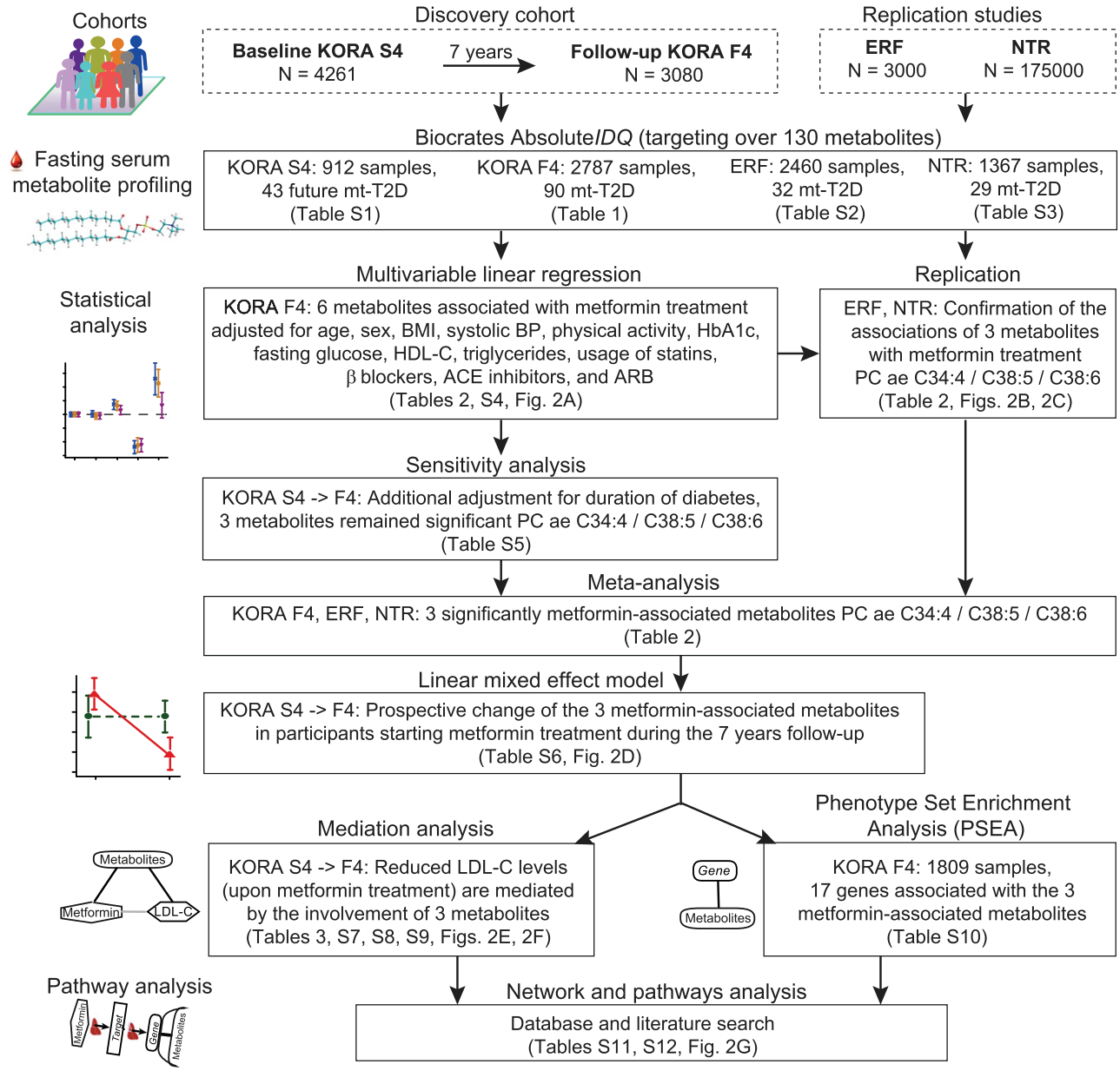


Figure 1—Flowchart of the study design.

with metformin-treated T2D (mt-T2D), *2*) patients with T2D with insulin treatment (it-T2D), *3*) patients with T2D without glucose-lowering treatment (non–antidiabetes drug treated [ndt-T2D]), *4*) participants with prediabetes with IGT, and *5*) healthy individuals with normal glucose tolerance (NGT) (Table 1).

Replication Studies

The ERF includes 3,000 living descendants of 22 couples who had at least six children baptized in the community church around 1850–1900. The participants are not selected based on any disease or other outcome. Details about the genealogy of the population have previously been provided (28).

The NTR recruits twins and their family members to study the causes of individual differences in health, behavior, and lifestyle. Participants are followed longitudinally; details about the cohort have previously been published (29). A subsample of unselected twins and their family members has taken part in the NTR-Biobank (32) in which biological samples, including DNA and RNA, were collected in a standardized manner after overnight fasting.

Duration of diabetes and 2-h postglucose levels were not available in either the ERF or NTR study. The diagnosis of patients with diabetes in both ERF and NTR studies was based on self-report. Owing to the limited number of it-T2D patients in these two replication studies ($n = 3$ and $n = 9$, respectively), this group is not included in the statistical analyses in these two replication studies.

Initially, we had contacted a third potential replication study, the Estonian Genome Center of the University of Tartu (EGCUT). However, only two mt-T2D participants with available metabolomics data were available in this cohort; results from the EGCUT study are therefore not shown.

Table 1—Characteristics of the KORA F4 cross-sectional study population

Clinical parameters	NGT	IGT	ndt-T2D	mt-T2D	it-T2D
n	2,129	375	169	90	24
Age, years	52.8 (12.6)	63.9 (11.0)	66.3 (9.7)	66.8 (8.7)	69.2 (9.8)
Male	46	49	62	59	54
BMI, kg/m^2	26.6 (4.3)	29.7 (4.9)	30.8 (4.4)	31.7 (5.4)	32.2 (5.9)
Waist, cm	90.5 (12.9)	99.7 (14.3)	104.6 (11.4)	106.3 (1.27)	107.2 (12.4)
Physical activity, >1 h per week	58	50	47	33	17
High alcohol intake†	17	17	18	20	8
Smoker	21	8	12	13	8
Systolic BP, mmHg	119.1 (17.4)	127.5 (18.5)	133.7 (18.6)	131.3 (18.9)	135.6 (22.7)
HDL-C, mg/dL	57.6 (14.4)	54.1 (14.0)	47.8 (12.1)	50.6 (10.5)	48.0 (9.6)
LDL-C, mg/dL	134.9 (34.3)	143.7 (35.4)	138.5 (36.5)	122.9 (29.0)	120.0 (31.6)
Triglycerides, mg/dL	110.6 (73.0)	146.0 (86.2)	175.1 (127.0)	174.4 (132.2)	142.1 (73.2)
HbA_{1c}, %	5.4 (0.3)	5.6 (0.3)	6.3 (0.9)	6.9 (1.1)	7.3 (1.1)
HbA_{1c}, mmol/mol	36 (3.3)	38 (3.3)	45 (9.8)	52 (12.0)	56 (12.0)
Fasting glucose, mg/dL	91.7 (7.6)	100.1 (10.6)	125.7 (29.1)	144.1 (37.1)	141.9 (39.0)
2-h postglucose load, mg/dL	97.7 (20.8)	161.7 (17.1)	214.5 (50.7)¥	—	—
Time since diagnosis, years	—	—	1.0 (3.1)#	7.7 (7.1)	16.7 (7.4)
Insulin, μIU/mL	6.9 (25.9)	13.1 (64.0)	16.6 (30.1)	10.4 (10.4)	32.2 (77.8)
Statin usage	8	16	24	38	33
β-Blocker usage	12	31	43	41	63
ACE inhibitor usage	8	21	31	43	58
ARB usage	6	9	15	13	8
Metformin usage	0	0	0	100	0
Insulin therapy	0	0	0	0	100

Percentages of individuals or means (SD) are shown for each variable and each group (NGT, IGT, ndt-T2D, mt-T2D, and it-T2D). †≥20 g/day for women; ≥40 g/day for men. ¥*n* = 121. #For newly diagnosed T2D patients (*n* = 112), years since T2D diagnosis was defined as 0.

Blood Sampling

In the KORA cohort study, blood was drawn into *S*-Monovette serum tubes (Sarstedt AG & Co., Nümbrecht, Germany) in the morning between 8:00 A.M. and 10:30 A.M. after at least 8 h of fasting. Tubes were gently inverted twice, followed by 30 min resting at room temperature to obtain complete coagulation. For serum collection, blood was centrifuged at 2,750 g at 15°C for 10 min. Serum was filled into synthetic straws, which were stored in liquid nitrogen (−196°C) until the metabolomics analyses (9,23).

In the ERF and NTR, the overnight fasting serum samples were drawn for metabolite profiling. Details about the sampling in these two cohorts were described in previous publications (28,32).

Metabolomics Measurement

The serum samples from participants in the baseline KORA S4 and follow-up KORA F4 study were measured with the Absolute*IDQ*p180 and Absolute*IDQ*p150 kits (Biocrates Life Sciences AG, Innsbruck, Austria), respectively. The assay procedures were previously described in detail (27). For KORA S4 and F4, identical quality-control procedures (9,13), which are explained in details in our previous publications, were used. In KORA F4, 131 metabolites of the initially targeted 163 metabolites passed all quality-control criteria: hexose (H1), 24 acylcarnitines, 14 amino acids, 13 sphingomyelines, 34 phosphatidylcholines (PCs), diacyl (aa), 37 PCs acyl-alkyl (ae), and 8 lysoPCs. In total, 124 metabolites overlapped between KORA S4 and F4, including H1, 21 acylcarnitines, 14 amino acids, 13 sphingomyelines, 33 PC aas, and 34 PC aes, as well as 8 lysoPCs.

The metabolite measurements for both replication studies (ERF and NTR) were performed using the same platform (Absolute*IDQ*p150 kit) as in the KORA F4 study. Additionally, in ERF, PC ae C36:4, PC ae C38:5, and PC ae C38:6 were measured in the full set of serum samples by a targeted liquid chromatography–mass spectrometry method. The measurement is performed on a UPLC-ESI-Q-TOF (Agilent 6530; Agilent Technologies, San Jose, CA) mass spectrometer using reference mass correction. Chromatographic separation was achieved on an ACQUITY UPLC HSS T3 column (1.8 μm, 2.1 * 100 mm) with a flow of 0.4 mL/min over a 16-min gradient. The metabolites were detected in full scan in the positive-ion mode. The raw data were processed using Agilent MassHunter Quantitative Analysis software (version B.04.00; Agilent Technologies).

Measured concentration values of all analyzed metabolites are reported in micromolar (μM) and were natural-log transformed, and the distributions were subsequently standardized with mean of zero and an SD of 1 for all analyses unless otherwise indicated.

Single Nucleotide Polymorphism Genotyping, Imputation, and Genes

In KORA F4, we carried out genotyping using the Affymetrix 6.0 GeneChip array (Affymetrix, Santa Clara, CA). Imputation was performed with Impute (http://mathgen.stats.ox.ac.uk/impute/), version 0.4.2 (reference HapMap phase 2, release 22). We only used autosomal single nucleotide polymorphisms (SNPs) with a

minor allele frequency >5%, call rate >95%, and imputation quality >0.4. For the phenotype set enrichment analysis (PSEA), we only mapped those SNPs to a gene that were either in its transcribed region or in its flanking region (110 kb upstream, 40 kb downstream). Gene information was downloaded from the UCSC (University of California, Santa Cruz) genome browser (http://genome.ucsc.edu). The SNP gene mapping was described in detail previously (25). In total, 20,801 genes were analyzed.

Statistical Analysis

To evaluate the effect of metformin treatment on metabolites, we used multivariable linear regression models with the metabolite concentration values as outcome and the grouping variable as predictor. Each metabolite was assessed individually. To include potential confounders, we adjusted for two sets of covariates: *1*) age and sex as the crude model and *2*) age, sex, BMI, physical activity, alcohol intake, smoking, systolic blood pressure (BP), levels of HDL cholesterol (HDL-C), triglycerides, HbA_{1c}, and fasting glucose, as well as the use of statins, β-blockers, ACE inhibitors, and angiotensin receptor blockers (ARBs) as the full model (Table 1). To account for multiple testing, we used Bonferroni correction and considered only those metabolites with a $P < 0.05/131 = 3.8E-04$ to be statistically significantly different in KORA F4. Meta-analysis of the three studies was performed using random effect models, using a restricted maximum-likelihood estimator.

In the KORA S4 to F4 longitudinal study, we used linear mixed-effect models. We adjusted for the two sets of covariates as described above while assigning a random offset to each of the individual participant in the longitudinal study. Additionally, using linear regression models on the KORA data set, including two time points (S4 n = 1,335 and F4 n = 2,763) (9), we calculated the residues of the metabolite concentrations adjusted for age, sex, BMI, physical activity, alcohol intake, smoking, systolic BP, HDL-C, triglyceride, fasting glucose, and HbA_{1c}. The significance of the changes in the metabolite concentrations between the two time points (S4 and F4) was tested using a linear mixed-effect model with the covariates at two time points.

PSEA is a gene-based approach to analyze the associations of genome-wide SNP data with multiple phenotypes in a combined way (25). The significance of enrichment was calculated based on 10,000 permutations (limited by computational restrictions), while setting the significance level at $P < 1.0E-04$ (lowest possible P value owing to the permutation number).

Mediation analysis (33) was conducted to model the identified metabolites as mediators for the association between metformin treatment and LDL-C and total cholesterol in the longitudinal KORA data. The mediation effects of each single metabolite and their summed concentration were tested with crude and fully adjusted multivariable linear regression models.

All statistical analyses were performed in R (version 3.0.1 [http://cran.r-project.org/]).

Pathway Analysis

With use of a bioinformatical approach, a network was constructed by retrieving pairwise connections between candidate metabolites, PSEA-identified genes, intermediate proteins, and known metformin target genes (9,31). Information on protein-protein interactions was extracted from STITCH (30). Known metformin target genes were retrieved from the DrugBank (4). In our network, we only considered the shortest paths (allowing one intermediate protein, confidence score >0.7) connecting the protein encoded by the genes identified in PSEA with the metformin target genes.

RESULTS

Metabolite Profiles in Three Cohorts

We quantified >130 metabolites in fasting serum samples from the KORA S4 and F4, ERF, and NTR studies (Fig. 1). The discovery study, KORA F4, includes 2,129 NGT, 375 IGT, 169 ndt-T2D, 90 mt-T2D, and 24 it-T2D subjects (characteristics shown in Table 1). In the longitudinal study, we used samples from 912 participants without metformin treatment at baseline (KORA S4); 43 of them were treated with metformin at follow-up (KORA F4 [Supplementary Table 1]). In reference to the two replication cohorts, ERF contained 29 ndt-T2D and 32 mt-T2D patients (characteristics shown in Supplementary Table 2), while NTR included 73 ndt-T2D and 29 mt-T2D patients (characteristics shown in Supplementary Table 3).

In general, patients with T2D in the three studies were older and more frequently men, with higher BMI, and took more nonantihyperglycemic medications than the participants without diabetes. Among the five groups in KORA F4, people on statin treatment had significantly lower LDL-C levels than those who were not taking statins (Supplementary Fig. 1). When comparing mt-T2D with ndt-T2D, lower levels of LDL-C were observed both in the cross-sectional (KORA F4, ERF, and NTR) and in the longitudinal KORA studies. Following the 43 patients, who started metformin treatment after the baseline, we did not observe significant changes in the levels of HbA_{1c} and fasting glucose but, however, observed significant changes for LDL-C and total cholesterol (Supplementary Table 1).

Metabolites Associated With Metformin Treatment

We found six metabolites including three acyl-alkyl PCs, two diacyl (aa) PCs, and one amino acid to have significantly lower concentrations in the 90 mt-T2D patients compared with the 169 ndt-T2D individuals in KORA F4 (Table 2). For example, for the metabolite PC ae C36:4, we observed that the fully adjusted effect estimate was −0.66 with P = 4.92E-07; i.e., the PC ae C36:4 level in the mt-T2D group was 0.66 SD lower than the ndt-T2D group.

We further investigated whether the observed differences are specifically for metformin treatment or just reflect the progress of T2D in general. The concentrations of the six metabolites are significantly lower in mt-T2D than in the NGT and IGT groups (Supplementary Table 4). In contrast, none of the six metabolites showed a significantly different concentration in the pairwise comparisons among the four groups without metformin treatment, i.e., NGT, IGT, ndt-T2D, and it-T2D (Supplementary Table 4).

For sensitivity analysis, we tested the associations of the six metabolites after adding the duration of T2D to the fully adjusted model. The three acyl-alkyl PCs (PC ae C36:4, PC ae C38:5, and PC ae C38:6), which are composed of at least one polyunsaturated fatty acid (PUFA), remained significantly different in the

Table 2—Metabolites associated with metformin treatment in patients with T2D

	Discovery KORA		Replication ERF		Replication NTR		Meta-analysis	
Metabolite	Effect estimate (95% CI)	*P*	Effect estimate (95% CI)	*P*	Effect estimate (95% CI)	*P*	Effect estimate (95% CI)	*P*
Crude model								
PC ae C36:4	−0.66 (−0.90, −0.43)	**1.07E-07**	−1.03 (−1.36, −0.69)	**2.65E-08**	−0.97 (−1.38, −0.56)	**9.21E-06**	−0.86 (−1.11, −0.60)	**5.24E-11**
PC ae C38:5	−0.65 (−0.90, −0.41)	**2.92E-07**	−1.04 (−1.37, −0.71)	**1.94E-08**	−0.72 (−1.12, −0.32)	**5.75E-04**	−0.79 (−1.04, −0.55)	**2.86E-10**
PC ae C38:6	−0.58 (−0.82, −0.35)	**2.39E-06**	−0.76 (−1.11, −0.42)	**3.53E-05**	−0.61 (−1.02, −0.19)	**4.44E-03**	−0.63 (−0.81, −0.45)	**4.11E-12**
PC aa C36:0	−0.50 (−0.75, −0.25)	**9.53E-05**	−0.30 (−0.78, 0.18)	2.31E-01	−0.54 (−1.08, 0.01)	5.43E-02	−0.47 (−0.68, −0.26)	**7.48E-06**
PC aa C38:0	−0.73 (−0.97, −0.49)	**1.48E-08**	−0.31 (−0.88, 0.27)	2.98E-01	−0.56 (−1.12, −0.01)	**4.49E-02**	−0.64 (−0.87, −0.40)	**8.22E-08**
Ornithine	−0.57 (−0.80, −0.33)	**4.42E-06**	−0.03 (−0.18, 0.11)	6.49E-01	0.01 (−0.45, 0.46)	9.78E-01	−0.21 (−0.59, 0.17)	2.72E-01
Full model								
PC ae C36:4	−0.66 (−0.91, −0.41)	**4.92E-07**	−1.08 (−1.33, −0.83)	**8.37E-05**	−0.84 (−1.27, −0.42)	**1.67E-04**	−0.79 (−1.02, −0.55)	**5.27E-11**
PC ae C38:5	−0.62 (−0.88, −0.36)	**3.81E-06**	−1.11 (−1.36, −0.85)	**6.18E-05**	−0.62 (−1.04, −0.20)	**4.65E-03**	−0.73 (−0.99, −0.47)	**3.73E-08**
PC ae C38:6	−0.58 (−0.82, −0.34)	**4.94E-06**	−0.74 (−1.00, −0.48)	**6.15E-03**	−0.49 (−0.91, −0.07)	**2.15E-02**	−0.58 (−0.78, −0.39)	**2.96E-09**
PC aa C36:0	−0.57 (−0.84, −0.30)	**4.25E-05**	−0.37 (−0.62, −0.12)	1.41E-01	−0.31 (−0.86, 0.24)	2.68E-01	−0.49 (−0.71, −0.27)	**9.79E-06**
PC aa C38:0	−0.68 (−0.94, −0.42)	**5.25E-07**	−0.35 (−0.64, −0.05)	2.47E-01	−0.34 (−0.90, 0.22)	2.32E-01	−0.56 (−0.82, −0.30)	**1.85E-05**
Ornithine	−0.58 (−0.80, −0.19)	**3.82E-05**	−0.01 (−0.09, 0.07)	8.80E-01	−0.13 (−0.62, 0.36)	6.02E-01	−0.25 (−0.62, 0.13)	2.00E-01

Effect estimates and *P* values for the comparison between mt-T2D and ndt-T2D were calculated using multivariable linear regression analysis adjusted for crude (age and sex) and full model, which includes age, sex, BMI, physical activity, alcohol intake, smoking, systolic BP, HDL-C, triglycerides, HbA_{1c}, fasting glucose, and use of statins, β-blockers, ACE inhibitors, and ARBs. Lipid side chain composition is abbreviated as Cx:y, where x denotes the number of carbons in the side chain and y the number of double bonds. Statistically significant *P* values appear in boldface.

comparison between mt-T2D and ndt-T2D ($P < 3.8E-04$) (Fig. 2*A*), whereas the other three metabolites were not significantly different anymore (Supplementary Table 5). After adjustment of the full model for *1*) waist, *2*) LDL-C, and *3*) the combination of LDL-C and insulin, the effect estimates of the six metabolites were almost unchanged (Supplementary Table 5).

Replication and Meta-analysis

For the three acyl-alkyl PCs, we observed consistent results in both replication studies (ERF and NTR); i.e., significantly lower levels were observed in mt-T2D patients compared with ndt-T2D individuals ($P < 0.05$) (Table 2 and Fig. 2*B* and *C*). Additionally, a meta-analysis of the three studies (KORA F4, ERF, and NTR) yielded significant results for the three replicated metabolites ($P < 3.8E-04$) (Table 2). We refer to these three highly intercorrelated metabolites, which are not associated with fasting glucose or HbA_{1c}, as metformin associated in the following paragraphs (Supplementary Fig. 2).

In the longitudinal examination, we found significantly decreased levels of the three metformin-associated metabolites in patients who underwent metformin treatment during the follow-up ($P < 3.8E-04$ using the fully adjusted model) (Supplementary Table 6). Consistent results for the three acyl-alkyl PCs were observed in a sensitivity analysis with a subgroup of 55 ndt-T2D patients at the KORA S4, of whom 19 were ndt-T2D patients and 36 were mt-T2D patients in KORA F4 in the fully adjusted model ($P < 0.05$) (Fig. 2*D* and Supplementary Table 6). These prospective findings confirmed our observations in the cross-sectional study.

Relationship Between Metformin Treatment, the Three Metabolites, and LDL-C Levels

To investigate a potentially mediating effect of the three acyl-alkyl PCs on the associations between metformin treatment and lipid profiles, we explored the prospective data of 912 KORA participants (Supplementary Table 1). We found that metformin treatment accounts for a significant decrease of LDL-C and total cholesterol levels, while its influence on HDL-C and triglycerides was not significant in both crude and

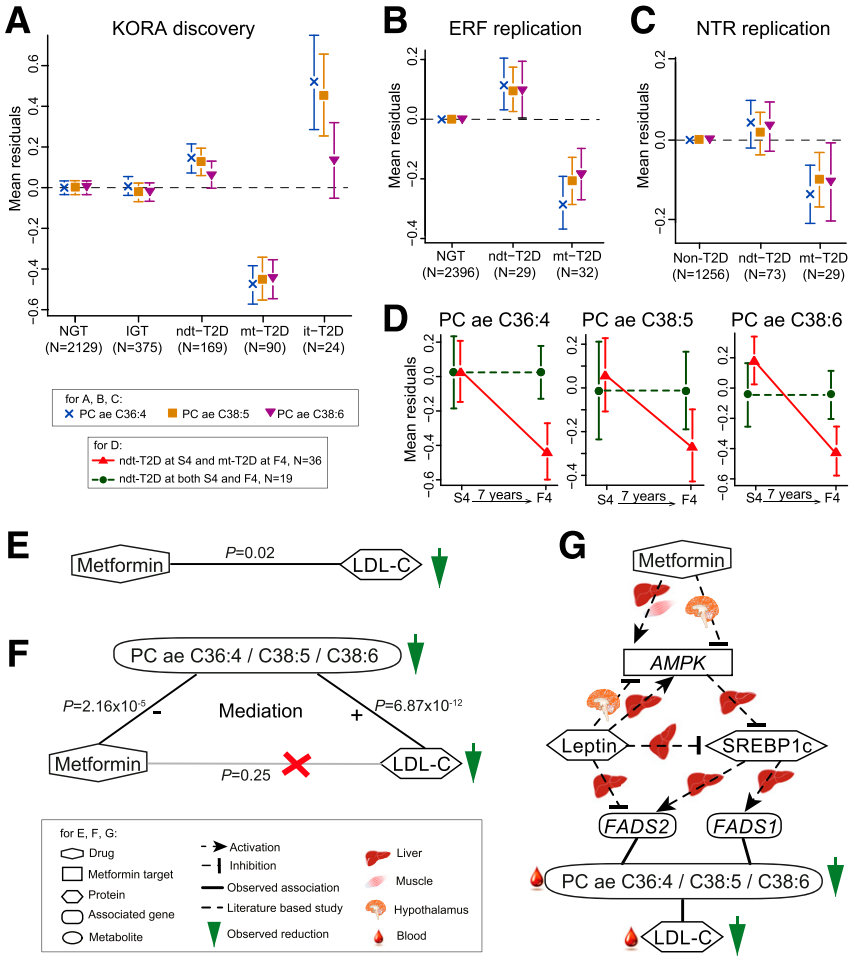


Figure 2—Differences in metabolite concentrations, mediation effect, and organ-specific pathways. Mean residuals of the concentrations (with SEs) of three identified acyl-alkyl PC metabolites for the NGT, IGT, ndt-T2D, mt-T2D, and it-T2D groups derived in cross-sectional analysis of the KORA F4 are shown in *A*. The mean residuals of the same metabolites in ERF are illustrated for the NGT, ndt-T2D, and mt-T2D groups in *B* and in NTR for the non-T2D, ndt-T2D, and mt-T2D groups in *C*, respectively. *D* refers to the longitudinal setting of the KORA study and shows the mean residuals of the concentrations (with SEs) of the three metabolites with respect to changes within the 7 years between baseline and follow-up study when people were treated with metformin. Residuals were calculated from linear regression model with the full adjustment. *E*: The association between metformin and LDL-C without consideration of the three metformin-associated metabolites. *F*: The results of the mediation analysis; the red cross indicates that the direct association between metformin and LDL-C is not significant anymore. *G*: An overview of the involved pathways. The connections indicated by liver, hypothalamus, muscle, and blood show organ specificity between genes, pathway-related proteins, and metformin drug targets as well as metformin. The metabolites (ellipses) were connected to metformin treatment (straight side hexagons) through genes (rounded rectangles), proteins (hexagons), and metformin targets (rectangles). The activation or inhibition is indicated. Plus or minus symbol next to the line indicate positive or negative association. For further information, see Table 3 and Supplementary Tables 8 and 12.

fully adjusted models ($P < 0.05$) (Fig. 2*E* and Supplementary Table 7). In particular, metformin was associated with a decrease in LDL-C levels of 11.83 mg/dL. We therefore focused on the analysis of LDL-C and total cholesterol.

After adding the three metabolites to the full model, the direct association between metformin treatment and LDL-C levels was not significant anymore ($P = 0.25$) (Fig. 2*F* and Supplementary Table 8*A*). Based on longitudinal analysis, we found consistent results as reported above (Table 2); i.e., significantly reduced levels of the three metabolites in the metformin-treated patients were observed (e.g., for the summed metabolite concentration $P = 2.16E-05$) (Fig. 2*F* and Supplementary Table 8*B*). Furthermore, we found significant positive associations between LDL-C and each of the three metabolites, as well as their summed concentration after adjusting for metformin treatment (e.g., for the summed metabolite concentration $P = 6.87E-12$) (Fig. 2*F* and Supplementary Table 8*C*). This means that these associations of the metformin-associated metabolites with LDL-C are independent of metformin treatment. Finally, for each of the three metformin-associated metabolites (and their summed concentration), the mediation effects on the association between metformin treatment and the LDL-C levels were significant in both models (Table 3). For instance, the summed concentration of the metabolites mediates 3.43 mg/dL reduction in LDL-C level, which accounts for 29% of the total effect of metformin on LDL-C (Table 3).

To rule out the potential effect of statin intake, we performed a sensitivity analysis by excluding individuals taking statin at baseline KORA S4 and/or follow-up F4. The mediation effects of the summed concentration were also significant for the associations between metformin and LDL-C level (Supplementary Table 9*A*). However, although the crude and full model showed similarly significant mediation effects for total cholesterol (Supplementary Table 8*D* and *E* and Table 3), after excluding statin users from the analysis, the effects on total cholesterol were not significant anymore with respect to the fully adjusted model ($P < 0.05$) (Supplementary Table 9*B*).

Seventeen Genes Are Linked to Metformin-Associated Metabolites and Pathway Analysis

To identify genes associated with the three metabolites, we applied PSEA on these metabolites in a subset of KORA F4 individuals (n = 1,809) with available genotyping data and metabolite profiles. We found 17 genes with an enrichment of SNPs in their transcribed or flanking region ($P < 1.0E-4$) (Supplementary Table 10). These genes belong to five clusters, one of them containing 12 genes located on chromosome 11. A literature search revealed disease phenotypes associated with these 17 genes. Six genes, namely, *FADS1*, *FADS2*, *FADS3*, *MYRF*, *BEST1* and *RAB3IL1*, are associated with T2D or its comorbidities, including retinopathy and

Table 3—Mediation effects of the three metabolites for the association between metformin treatment and reduction of LDL-C and total cholesterol

	Crude model			Full model		
	Effect estimate (95% CI)	*P*	Explained effect (%)	Effect estimate (95% CI)	*P*	Explained effect (%)
LDL-C						
PC ae C36:4	−3.05 (−4.38, −1.71)	2.21E-04	25.74	−2.51 (−3.71, −1.31)	1.33E-03	21.22
PC ae C38:5	−2.94 (−4.21, −1.67)	2.65E-04	24.82	−2.36 (−3.50, −1.22)	1.94E-03	19.97
PC ae C38:6	−5.25 (−8.11, −2.40)	1.34E-05	44.40	−4.02 (−6.59, −1.44)	4.55E-04	33.95
Summed concentration†	−4.37 (−6.37, −2.37)	1.52E-05	36.92	−3.43 (−5.19, −1.67)	2.95E-04	28.99
Total cholesterol						
PC ae C36:4	−5.00 (−7.77, −2.23)	2.63E-05	26.1	−3.08 (−4.72, −1.45)	7.42E-04	26.5
PC ae C38:5	−4.99 (−7.71, −2.26)	2.33E-05	26.0	−2.77 (−4.25, −1.30)	1.38E-03	23.8
PC ae C38:6	−6.99 (−11.86, −2.13)	8.96E-06	36.4	−4.16 (−6.98, −1.35)	5.11E-04	35.8
Summed concentration†	−6.63 (−10.55, −2.71)	2.76E-06	34.6	−3.87 (−6.05, −1.69)	2.41E-04	33.3

The estimates of the mediation effects and *P* values were calculated using the longitudinal (KORA S4→F4) mediation analysis adjusted for the crude and full model. The mediation effects for the three metformin-associated metabolites and the summed concentration are shown. †The summed concentration refers to the overall concentration of the three metabolites (PC ae C36:4, PC ae C38:5, and PC ae C38:6).

coronary artery diseases (for references, see Supplementary Table 10).

To explore potentially related pathways, we used a bioinformatics approach, integrating the 17 identified genes with 6 known metformin target genes (4) into a protein-protein interaction network (9,30,31). For 3 of the 17 genes, there was no record for *Homo sapiens* in the STITCH (30); therefore, we investigated the interaction of the remaining 14 genes with the 6 metformin targets (Supplementary Table 11). *AMPK* was found to be linked to *FADS1* and *FADS2* through interacting proteins (leptin and sterol regulatory element–binding protein 1c [SREBP1c]). A manual evaluation of these interactions in a literature research showed organ specificity, mainly referring to liver and hypothalamus (Fig. 2*G*). The AMPK complex is inhibited by leptin and metformin in the hypothalamus, whereas it is activated by metformin and leptin in the liver. (References for each interaction are provided in the Supplementary Table 12).

CONCLUSIONS

We found significant concentration differences for three metabolites (PC ae C36:4, PC ae C38:5, and PC ae C38:6) in the blood of patients with T2D under metformin treatment and replicated them in two independent studies. We identified SNP variations in 17 genes (including *FADS1* and *FADS2*) that were associated with the three metabolites. Based on these genes, we built an interaction network to investigate the underlying mechanisms of metformin treatment and identified the organ-specific AMPK pathway. We further found that the reduced LDL-C levels in metformin-treated patients with T2D were mediated partially by the three acyl-alkyl PCs. Sensitivity analyses were performed to consider the duration of diabetes and statin use.

The levels of metabolites depend on multiple modifiable factors, such as lifestyle and environment (9–11,13–16). We therefore considered a number of confounding effects, e.g., physiological parameters (age, sex, BMI, and systolic BP), lifestyle (physical activity, alcohol intake, and smoking), glucose levels (HbA_{1c} and fasting glucose), lipid levels (HDL-C and triglycerides), and medication usage (statins, β-blockers, ACE inhibitors, and ARBs). Additionally, intermediates or end products of metabolism are influenced by underlying genetic factors (23,24). In our study, phenotypes and genotypes are available for each person (n = 1,809); we thus used phenotype set enrichment analysis (25). Our combined analysis of genetic and metabolomic data enabled us to identify genes associated with the three metabolites and supported the identification of an organ-specific pathway. The observation of significantly lower levels of the three metformin-associated metabolites (polyunsaturated acyl-alkyl PCs) in the mt-T2D patients can be explained by metformin's effects on AMPK in the liver (Fig. 2*G* and Supplementary Table 12). In the hepatocyte, metformin increases the AMP-to-ATP ratio and thus leads to the activation of AMPK. Activated AMPK blocks SREBP1c, a transcription factor controlling enzymes involved in the fatty acid synthesis and inhibiting the synthesis of FADS1 and FADS2 (22). This results in a reduced synthesis of unsaturated fatty acids and consequently lower acyl-alkyl PCs concentrations. Leptin occupies a central position in the network (Fig. 2*G*) and affects the FADS complex via three different interactions. In the liver, leptin not only activates AMPK, thereby suppressing SREBP1c and downregulating FADS1 and FADS2, but can also directly inhibit both SREBP1c and FADS2 (34). Metformin and leptin exert opposite effects in the hypothalamus and in the liver (for references, see Supplementary Table 12), but further studies are required to better understand the organ-specific metformin effects in humans.

Recently, clinical practice guidelines have recommended the usage of metformin as first-line therapy in patients with T2D with heart failure (1,2). Our observation of lower blood levels of LDL-C in metformin-treated patients points toward a beneficial effect of metformin for the prevention of CVD. A meta-analysis of randomized clinical trials shows that metformin treatment results in lowered LDL-C levels in newly diagnosed T2D patients (8). Similar results were also reported in patients without T2D in an epidemiological study (7). Here, we observed that metformin treatment leads to lowered LDL-C levels, an effect mediated most likely through metformin-mediated reduction of FADS

activity and consequently reduction of the levels of PUFA, namely arachidonic acid (35). It has been suggested that lower levels of arachidonic acid leads to an increased membrane fluidity, thus increasing LDL-C receptor recycling (35). This hypothesis is especially strong, given that genetic variants assigned to lower activity of FADS1 and -2 were significantly associated with lower LDL-C levels (36). While certain PCs can indeed exert antidiabetic effects (37), further mechanistic studies are required to test whether lowering of these circulating lipids contributes directly to the prevention of CVDs or merely by its antidiabetic effect (1).

Beyond its common antihyperglycemic action and its effect in lowering LDL-C, metformin can potentially reduce the risk of cancer mortality and diminish the progression of cancer (38). In the current study, we have found the three metformin-associated metabolites significantly associated with two genes, *FEN1* and *C20orf94*, which are involved in DNA repair (39,40). This may partly explain that metformin has been shown to influence the prevalence of different types of carcinoma, such as gastrointestinal cancers (39) and leukemia (40).

The strength of our study is that we used three independent cohort studies to discover and replicate our observations. Importantly, all results presented in this study were independent of physiological parameters, lifestyle, glucose levels, lipid levels, and medication. We combined metabolomics and genomics data, broad literature research, and organ-specific information from animal studies to deepen the insight into the underlying mechanisms.

Our findings are limited by the observational nature of cohort studies, and the applied methods, such as the mediation analysis, are of purely statistical character, but they offer the opportunity to raise new questions for experimental confirmation studies, such as randomized controlled clinical trials to investigate, for instance, the effect of metformin on blood lipid levels of patients without diabetes.

In the present studies (KORA, ERF, NTR), the duration of T2D is based on self-reported information. Moreover, neither data on the dosage nor data on duration and compliance of the metformin treatment were available. Furthermore, it has to be mentioned that the degree of diabetes severity presumably discriminates the different groups of patients (ndt-T2D, mt-T2D, and it-T2D), which is reflected by their HbA_{1c} and fasting glucose values (Table 1). Although the investigated metabolite panel does not represent the whole human metabolome, the comprehensive analysis of >130 metabolites from different classes represents a considerable improvement compared with previous technologies.

We found three metformin-associated metabolites, which showed no overlap with the findings of previous studies (17–21). This is likely to result from the use of different sampling matrices (plasma vs. serum), unmeasured metabolites (asymmetric dimethylarginine), or study design (glipizide treatment). Additional, our study considered considerably more potential cofounding effects in a comparably larger number of individuals than previous studies (17–21).

In conclusion, we observed that metformin treatment reduced levels of the three acyl-alkyl PC metabolites in patients with T2D. This change in the metabolic profiles may mediate lowered blood levels of LDL-C. The underlying mechanism is most likely the metformin-induced activation of AMPK and the consequent suppression of SREBP1c and FADS, which leads to reduced levels of PUFA and LDL-C. Our findings suggest a pharmaco-epidemiologic mechanism by which metformin may exert beneficial effects to prevent CVD. More importantly, our study suggests a novel approach to identify pleiotropic effects of medication using multilevel omics data.

Acknowledgments. The authors express their appreciation to all KORA study participants for donating their blood and time. The authors thank the field staff in Augsburg conducting the KORA studies. The authors thank the staff from the Institute of Epidemiology at the Helmholtz Zentrum München and the Genome Analysis Center Metabolomics Platform, who helped in the sample logistics, data and straw collection, and metabolomic measurements, and especially J. Scarpa, K. Faschinger, F. Scharl, N. Lindemann, H. Chavez, A. Sabunchi, A. Schneider, A. Ludolph, S. Jelic, and B. Langer. The authors are grateful to all general practitioners involved in ERF for their contributions. The authors thank all participants in the NTR.

Funding. The KORA study was 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. Part of this project was supported by The European Community's Seventh Framework Programme grants HEALTH-2009-2.2.1-3/242114 (Project OPTiMiS) and HEALTH-2013-2.4.2-1/602936 (Project CarTarDis). The German Diabetes Center is funded by the German Federal Ministry of Health (Berlin, Germany) and The Ministry of Innovation, Science and Research of the State of North Rhine-Westphalia (Düsseldorf, Germany). The diabetes part of the KORA F4 study was funded by a grant from the German Research Foundation (DFG) RA 459/3-1. This study was supported in part by a grant from The BMBF to the German Center for Diabetes Research (DZD e. V.). W.R.-M. is funded by The German Federal Ministry of Education and Research grant 03IS2061B (project Gani_Med). K.Su. is supported by Biomedical Research Program funds at Weill Cornell Medical College in Qatar, a program funded by the Qatar Foundation. The EGCUT received support from The European Community's Seventh Framework Programme grant BBMRI-LPC 313010, targeted financing from Estonian Government IUT20-60 and IUT24-6, Estonian Research Roadmap through the Estonian Ministry of Education and Research (3.2.0304.11-0312), the Center of Excellence in Genomics (EXCEGEN), and Development Fund from the University of Tartu (SP1GVARENG), and from an EFSD New Horizons grant. This work was also supported by the U.S. National Institutes of Health (R01DK075787). The ERF was supported by grants from the Netherlands Organisation for Scientific Research (NWO), and Erasmus MC and the Centre for Medical Systems Biology (CMSB). Telomere length assessment was supported through funds from The European Community's Seventh Framework Programme (FP7/2007-2013), grant agreement HEALTH-F4-2007-201413 (ENGAGE). Research for the NTR was funded by The Netherlands Organisation for Scientific Research (NWO) (MagW/ZonMW grants 904-61-090, 985-10-002, 904-61-1, 480-04-004, 400-05-717; Addiction-31160008 Middelgroot-911-09-032; and Spinozapremie 56-464-14192), the Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL) (184·021·007), the VU University's Institute for Health and Care Research (EMGO+), and The European Community's Seventh Framework Program (FP7/2007-2013), grant HEALTH-F4-2007-201413 (ENGAGE).

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. T.X., S.B., A.C.M., C.He., H.H.M.D., A.D., Z.Y., J.S.R., and T.Hal. analyzed the data and interpreted the results. T.X., S.B., A.C.M., C.He., M.Rod., T.M., and R.W.-S. wrote the manuscript. M.H., R.S., C.Ho., J.Adam, S.C., M.Rot., T.P., Y.H., G.K., C.M., B.T., A.R., M.H.d.A., Y.L., H.-E.W., B.S., A.M., C.G., K.St., and W.R. assisted in manuscript generation and revision. M.C., G.F., A.S.Q., G.W., R.P., K.K., and K.W.v.D. interpreted the results. C.P., W.R.-M., K.Su., T.Han., J.Adams., and T.I. performed the metabolic profiling. T.I., A.P., C.M.v.D., D.I.B.,

T.M., and R.W.-S. conceived and designed the study. R.W.-S. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

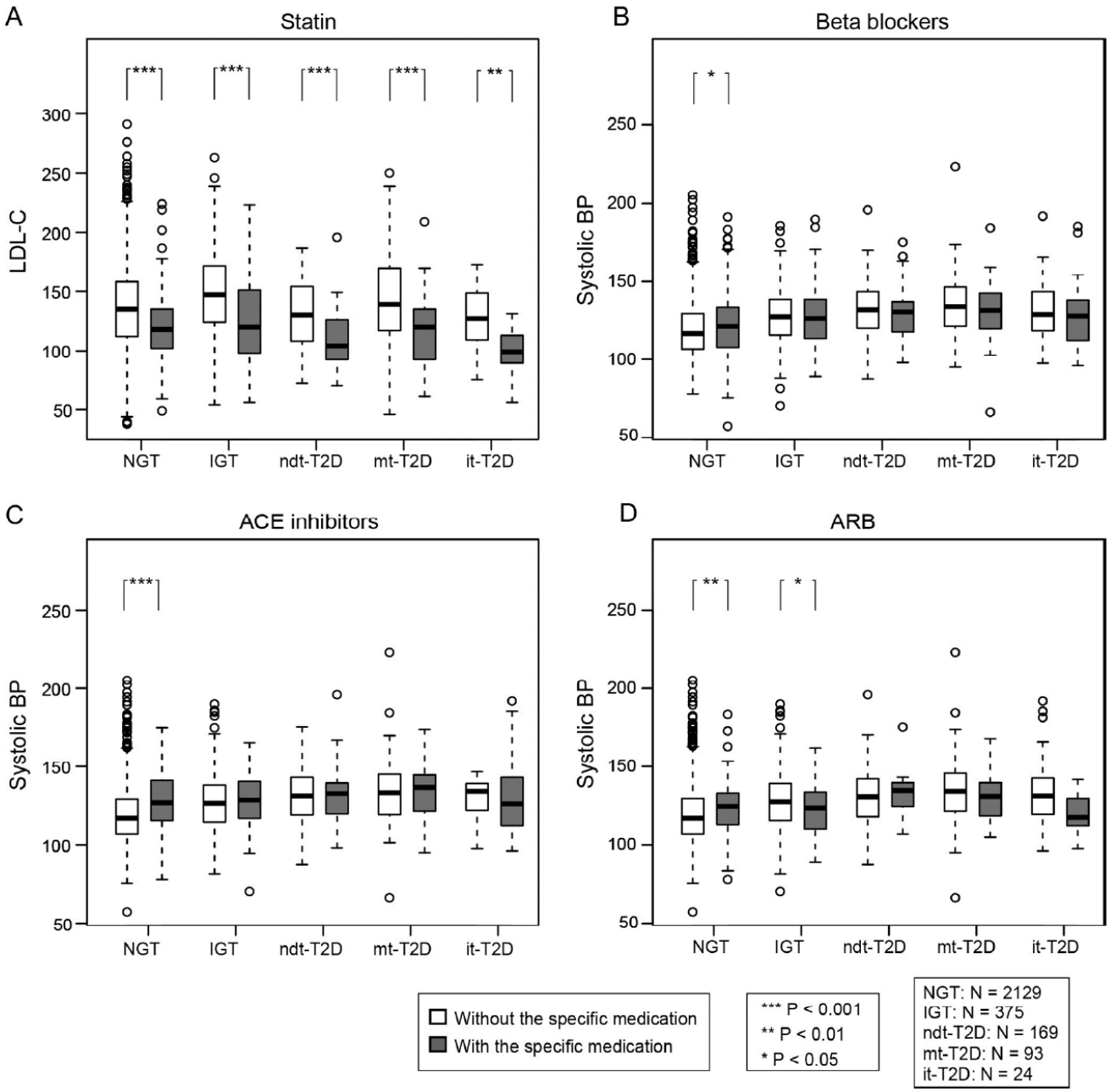
1. Foretz M, Guigas B, Bertrand L, Pollak M, Viollet B. Metformin: from mechanisms of action to therapies. Cell Metab 2014;20:953–966
2. ADA. Executive summary: Standards of medical care in diabetes–2013. Diabetes Care 2013; 36(Suppl. 1):S4–S10
3. Madiraju AK, Erion DM, Rahimi Y, et al. Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase. Nature 2014;510:542–546
4. Law V, Knox C, Djoumbou Y, et al. DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res 2014;42:D1091–D1097
5. Decensi A, Puntoni M, Goodwin P, et al. Metformin and cancer risk in diabetic patients: a systematic review and meta-analysis. Cancer Prev Res (Phila) 2010;3:1451–1461
6. Norwood DK, Chilipko AA, Amin SM, Macharia D, Still KL. Evaluating the potential benefits of metformin in patients with cardiovascular disease and heart failure. Consult Pharm 2013;28:579–583
7. Glueck CJ, Fontaine RN, Wang P, et al. Metformin reduces weight, centripetal obesity, insulin, leptin, and low-density lipoprotein cholesterol in nondiabetic, morbidly obese subjects with body mass index greater than 30. Metabolism 2001; 50:856–861
8. Wulffelé MG, Kooy A, de Zeeuw D, Stehouwer CD, Gansevoort RT. The effect of metformin on blood pressure, plasma cholesterol and triglycerides in type 2 diabetes mellitus: a systematic review. J Intern Med 2004;256:1–14
9. Wang-Sattler R, Yu Z, Herder C, et al. Novel biomarkers for pre-diabetes identified by metabolomics. Mol Syst Biol 2012;8:615
10. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. Nat Med 2011;17:448–453
11. Floegel A, Stefan N, Yu Z, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes 2013;62:639–648
12. Newgard CB, An J, Bain JR, 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
13. Yu Z, Zhai G, Singmann P, et al. Human serum metabolic profiles are age dependent. Aging Cell 2012;11:960–967
14. Mittelstrass K, Ried JS, Yu Z, et al. Discovery of sexual dimorphisms in metabolic and genetic biomarkers. PLoS Genet 2011;7:e1002215
15. Xu T, Holzapfel C, Dong X, et al. Effects of smoking and smoking cessation on human serum metabolite profile: results from the KORA cohort study. BMC Med 2013;11:60
16. Jaremek M, Yu Z, Mangino M, et al. Alcohol-induced metabolomic differences in humans. Transl Psychiatry 2013;3:e276
17. Marchetti P, Masiello P, Benzi L, et al. Effects of metformin therapy on plasma amino acid pattern in patients with maturity-onset diabetes. Drugs Exp Clin Res 1989;15:565–570
18. Cai S, Huo T, Li N, Xiong Z, Li F. Lysophosphatidylcholine–biomarker of Metformin action: studied using UPLC/MS/MS. Biomed Chromatogr 2009;23:782–786
19. Huo T, Cai S, Lu X, Sha Y, Yu M, Li F. Metabonomic study of biochemical changes in the serum of type 2 diabetes mellitus patients after the treatment of metformin hydrochloride. J Pharm Biomed Anal 2009;49:976–982
20. Asagami T, Abbasi F, Stuelinger M, et al. Metformin treatment lowers asymmetric dimethylarginine concentrations in patients with type 2 diabetes. Metabolism 2002;51:843–846
21. Zhang Y, Hu C, Hong J, et al. Lipid profiling reveals different therapeutic effects of metformin and glipizide in patients with type 2 diabetes and coronary artery disease. Diabetes Care 2014;37:2804–2812
22. Gieger C, Geistlinger L, Altmaier E, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet 2008;4:e1000282
23. Illig T, Gieger C, Zhai G, et al. A genome-wide perspective of genetic variation in human metabolism. Nat Genet 2010;42:137–141
24. Suhre K, Shin SY, Petersen AK, et al.; CARDIoGRAM. Human metabolic individuality in biomedical and pharmaceutical research. Nature 2011;477:54–60
25. Ried JS, Döring A, Oexle K, et al. PSEA: Phenotype Set Enrichment Analysis–a new method for analysis of multiple phenotypes. Genet Epidemiol 2012;36:244–252
26. Rathmann W, Kowall B, Heier M, et al. Prediction models for incident type 2 diabetes mellitusin the older population: KORA S4/F4 cohort study. Diabet Med 2010;27:1116–1123
27. Jourdan C, Petersen AK, Gieger C, et al. Body fat free mass is associated with the serum metabolite profile in a population-based study. PLoS One 2012;7:e40009
28. Aulchenko YS, Heutink P, Mackay I, et al. Linkage disequilibrium in young genetically isolated Dutch population. Eur J Hum Genet 2004; 12:527–534
29. Boomsma DI, de Geus EJ, Vink JM, et al. Netherlands Twin Register: from twins to twin families. Twin Res 2006;9:849–857
30. Kuhn M, Szklarczyk D, Pletscher-Frankild S, et al. STITCH 4: integration of protein-chemical interactions with user data. Nucleic Acids Res 2014;42:D401–D407
31. Lechner M, Höhn V, Brauner B, et al. CIDeR: multifactorial interaction networks in human diseases. Genome Biol 2012;13:R62
32. Willemsen G, de Geus EJ, Bartels M, et al. The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin Res 2010;13:231–245
33. Bauer DJ, Preacher KJ, Gil KM. Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: new procedures and recommendations. Psychol Methods 2006;11:142–163
34. Lim CT, Kola B, Korbonits M. AMPK as a mediator of hormonal signalling. J Mol Endocrinol 2010;44:87–97
35. Sone Y, Kido T, Ainuki T, et al. Genetic variants of the fatty acid desaturase gene cluster are associated with plasma LDL cholesterol levels in Japanese males. J Nutr Sci Vitaminol (Tokyo) 2013;59:325–335
36. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45:1274–1283
37. Lee JM, Lee YK, Mamrosh JL, et al. A nuclear-receptor-dependent phosphatidylcholine pathway with antidiabetic effects. Nature 2011;474:506–510
38. Pernicova I, Korbonits M. Metformin–mode of action and clinical implications for diabetes and cancer. Nat Rev Endocrinol 2014;10:143–156
39. Liu L, Zhou C, Zhou L, et al. Functional FEN1 genetic variants contribute to risk of hepatocellular carcinoma, esophageal cancer, gastric cancer and colorectal cancer. Carcinogenesis 2012;33:119–123
40. Meissner B, Bartram T, Eckert C, et al. Frequent and sex-biased deletion of SLX4IP by illegitimate V(D)J-mediated recombination in childhood acute lymphoblastic leukemia. Hum Mol Genet 2014;23:590–601

SUPPLEMENTARY DATA

Effects of metformin on metabolite profiles and LDL cholesterol in type 2 diabetes patients

Supplementary Figure 1. Influence of non-antihyperglycemic agents on type 2 diabetes risk factors

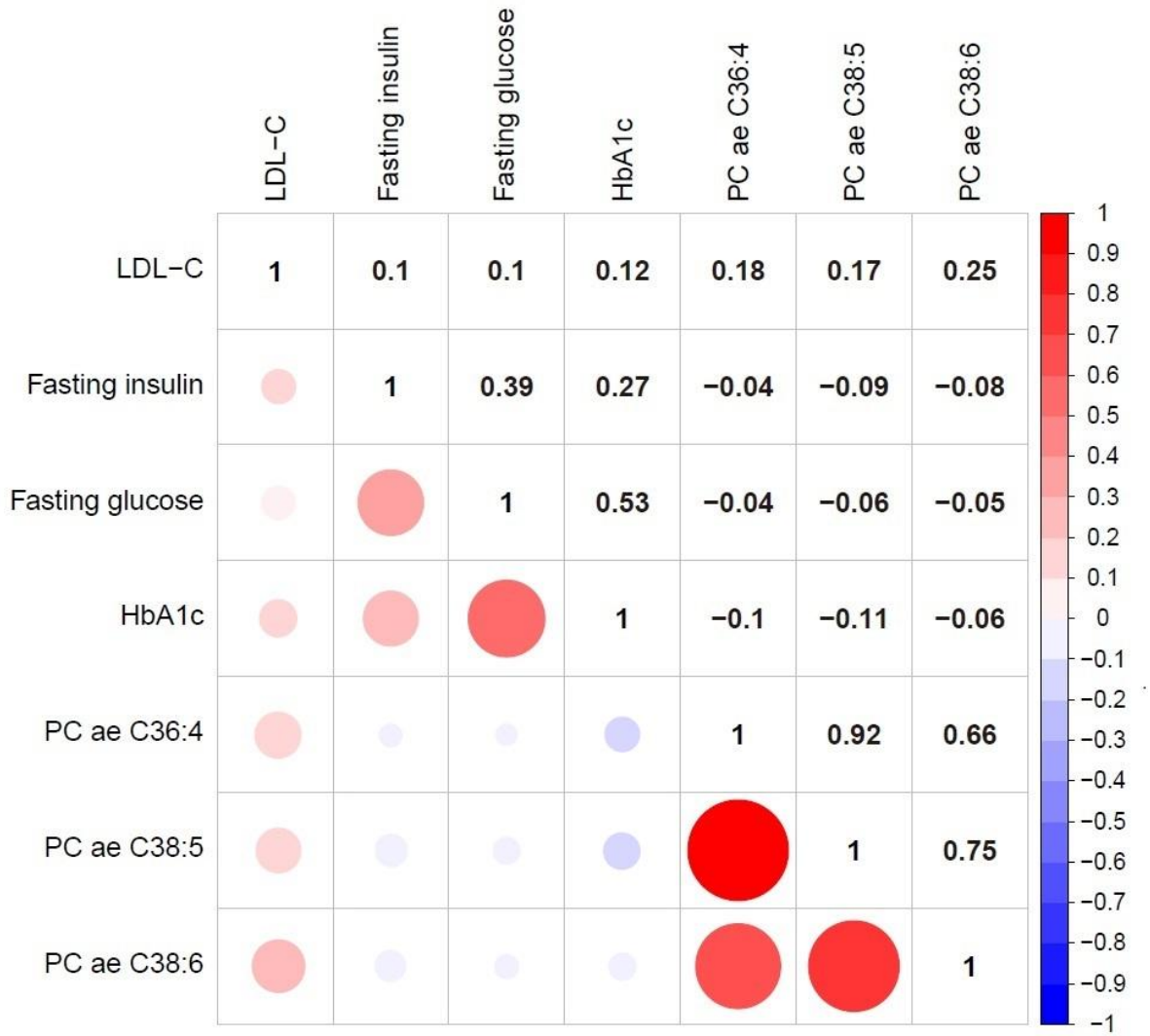
Statin usage was associated with lower levels of low density lipoprotein cholesterol (LDL-C) in all five groups: NGT, IGT, ndt-T2D, mt-T2D and it-T2D (plot A); treatment with beta blockers (plot B), angiotensin-converting-enzyme (ACE) inhibitors (plot C), or angiotensin receptor blockers (ARB) (plot D) are associated with different levels of systolic blood pressure (BP). NGT, normal glucose tolerance; IGT, impaired glucose tolerance; ndt-T2D, non-drug treated type 2 diabetes; mt-T2D, metformin-treated type 2 diabetes, it-T2D, insulin-treated type 2 diabetes.



SUPPLEMENTARY DATA

Supplementary Figure 2. Correlation of the three metformin-associated metabolites with risk factors of type 2 diabetes

The spearman correlation coefficients between the three metformin-associated metabolites and conventional risk factors of type 2 diabetes including LDL-C, fasting insulin, fasting glucose, and HbA1c, in the cross-sectional KORA F4 study are shown. Both the size of the cycle and intensity of color indicate the degree of correlation between the metabolites and the risk factors. The numeric values of spearman correlation coefficients are shown in the upper triangle.



SUPPLEMENTARY DATA

Supplementary Table 1. Characteristics of the KORA S4 → F4 prospective study samples

Percentages of individuals or means (SD) are shown for each variable and each group. Abbreviations: w/o, without; w/, with; ndt-T2D, non-anti-diabetic drug treated type 2 diabetes; mt-T2D, metformin-treated type 2 diabetes; BMI, body mass index; h, hour; BP, blood pressure; HDL-C, high-density lipoprotein cholesterol; LDL, low-density lipoprotein cholesterol; ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blockers. Significant *P* values are indicated bold (*P* value<0.05).

	Baseline S4: w/o metformin Follow-up F4: w/o metformin (n=869)			**Baseline S4: w/o metformin Follow-up F4: w/ metformin (n=43)**		
	S4	**F4**	***P*-value**	**S4**	**F4**	***P*-value**
Age	63.1 (5.4)	70.2 (5.4)	-	63.9 (4.4)	71.1 (4.6)	-
Male	51	51	-	55	55	-
Weight, kg	77.0 (12.5)	77.6 (13.3)	**8.9E-05**	89.6 (14.2)	87.1 (14.8)	**3.4E-04**
BMI, kg/m^2	28.0 (4.0)	28.3 (4.2)	**6.4E-09**	32.6 (4.2)	31.6 (4.4)	**2.8E-03**
Physical activity, > 1h per week	52.8	47.5	**2.2E-03**	67.4	58.2	0.29
High alcohol intake[†]	20.0	17.2	**0.05**	27.9	20.9	0.37
Smoker	11.9	7.4	**1.8E-08**	18.6	11.6	0.37
Systolic BP, mmHg	133.5 (18.9)	127.9 (19.4)	**1.5E-13**	144.0 (17.7)	130.8 (17.6)	**4.7E-05**
HDL-C, mg/dL	59.3 (16.5)	56.7 (14.1)	**2.6E-12**	53.2 (11.6)	52.7 (8.9)	0.66
LDL-C, mg/dL	155.5 (39.5)	141.8 (36.4)	**1.0E-21**	143.1 (37.2)	125.5 (24.4)	**5.6E-03**
Total cholesterol, mg/dL	245.4 (41.2)	223.7 (40.8)	**1.5E-26**	233.9 (41.7)	203.37(33.91)	**3.5E-04**
Triglycerides, mg/dL	127.4 (71.1)	127.1 (71.4)	0.76	171.8 (165.3)	151.2 (148.3)	0.08
HbA_{1C}, %	5.6 (0.4)	5.7 (0.5)	**0.01**	6.4 (0.9)	6.6 (0.7)	0.35
HbA_{1C}, mmol/mol	38 (4.4)	39 (5.5)	**0.01**	46 (9.8)	49 (7.7)	0.35
Fasting glucose, mg/dL	99.8 (11.1)	99.8 (17.4)	0.58	134.7 (33.1)	130.9 (27.3)	0.35
Statin usage	9.4	22.7	**1.5E-21**	9.3	30.2	**2.7E-02**
Beta blocker usage	18.1	31.5	**1.2E-18**	20.9	37.2	**4.6E-02**
ACE inhibitor usage	9.4	24.4	**9.3E-23**	20.9	51.2	**3.6E-03**
ARB usage	3.7	12.4	**1.6E-15**	2.3	13.9	0.07

[†] ≥40g/day in men; ≥20g/day in women

SUPPLEMENTARY DATA

Supplementary Table 2. Characteristics of the ERF study samples

Percentages of individuals or means (SD) are shown for each variable and each group. NGT, normal glucose tolerance; ndt-T2D, non-anti-diabetic drug treated type 2 diabetes; mt-T2D, metformin-treated type 2 diabetes; it-T2D, insulin-treated type 2 diabetes; BMI, body mass index; h, hour; BP, blood pressure; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; ACE, angiotensin-converting-enzyme; ARB, angiotensin receptor blockers.

Clinical parameters	**NGT**	**ndt-T2D**	**mt-T2D**	**it-T2D**
n	2396	29	32	3
Age, years	47.9 (14.2)	55.7 (11.4)	60.7 (11.8)	54.4 (15.5)
Males	43.8	51.0	58.8	20.0
BMI, kg/m^2	26.7 (4.5)	28.6 (4.5)	32.4 (6.1)	28.1 (5.8)
Physical activity, > 1h per week	43.3	77.0	70.6	66.6
High alcohol intake†	7.2	10.3	8.7	33.3
Active Smokers	39.3	40	42.9	100
Systolic BP, mmHg	138.9 (19.7)	154.4 (22.1)	151.5 (19.2)	152.50 (35.91)
HDL-C, mg/dL	49.4 (13.9)	47.1 (13.9)	39.4 (9.3)	40.15 (10.42)
LDL-C, mg/dL	144.8 (28.2)	145.2 (42.1)	108.5 (30.9)	122.39 (38.99)
Triglycerides, mg/dL	115.0 (64.6)	161.1 (107.1)	163.7 (92.0)	151.32 (77.87)
HbA_{1C}, %	NA	NA	NA	NA
HbA_{1C}, mmol/mol	NA	NA	NA	NA
Fasting glucose, mg/dL	79.8 (11.4)	145.0 (21.8)	139.3 (49.0)	170.09 (78.91)
2-h post-glucose load, mg/dL	NA	NA	NA	NA
Statin usage	17.5	34.4	58.8	100
Beta blocker usage	15.7	18.8	29.9	66.6
ACE inhibitor usage	5.6	9.4	19.6	66.6
ARB usage	7.8	5.9	17.6	33.3
Metformin usage	0	0	100	0
Insulin therapy	0	0	0	100

† ≥ 20g/day for women; ≥ 40g/day for men

SUPPLEMENTARY DATA

Supplementary Table 3. Characteristics of the NTR study samples

Percentages of individuals or means (SD) are shown for each variable and each group. Non-T2D, non-type 2 diabetes; ndt-T2D, non-anti-diabetic drug treated type 2 diabetes; mt-T2D, metformin-treated type 2 diabetes; it-T2D, insulin-treated type 2 diabetes; BMI, body mass index; h, hour; BP, blood pressure; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; ACE, angiotensin-converting-enzyme; ARB, angiotensin receptor blockers.

Clinical Parameters	**non-T2D**	**ndt-T2D**	**mt-T2D**	**it-T2D**
n	1256	73	29	9
Age, years	50.8 (14.1)	56.6 (11.0)	62.6 (6.5)	49.3 (18.9)
Male	66	75	59	67
BMI, kg/m^2	25.7 (3.6)	28.6 (4.2)	30.1 (5.4)	25.1 (3.8)
Physical activity, > 1h per week	NA	NA	NA	NA
High alcohol intake	NA	NA	NA	NA
Smoker	22	19	17	0
Systolic BP, mmHg	NA	NA	NA	NA
HDL-C, mg/dL	52.1 (14.5)	44.5 (14.0)	44.6 (9.7)	56.3 (16.5)
LDL-C, mg/dL	123.1 (35.7)	120.5 (39.3)	100.1 (32.9)	97.7 (22.8)
Triglycerides, mg/dL	131.9 (74.5)	190.1 (102.0)	156.1 (97.9)	77.7 (29.2)
HbA_{1C}, %	5.2 (1.1)	6.0 (1.3)	6.0 (0.8)	6.7 (1.04)
HbA_{1C}, mmol/mol	33 (12.0)	42 (14.2)	42 (8.7)	50 (11.4)
Fasting glucose, mg/dL	98.4 (10.0)	151.8 (35.6)	144.6 (34.1)	146.6 (41.6)
2-h post-glucose load, mg/dL	NA	NA	NA	NA
Statin usage	10	27	55	33
Beta blocker usage	9	18	28	22
ACE inhibitor usage	4	15	38	44
ARB usage	4	10	14	11
Metformin usage	0	0	100	0
Insulin therapy	0	0	0	100

SUPPLEMENTARY DATA

Supplementary Table 4. Variances of the six metabolites in other pair-wise comparisons in the discovery KORA F4 study

Effect estimates were calculated using multivariable linear regression analysis with the fully adjusted model. Significant *P* values are indicated bold (*P* value<3.8E-04).

Metabolite	mt-T2D vs. NGT		mt-T2D vs. IGT		mt-T2D vs. it-T2D	
	Effect estimate (95% CI)	***P*-value**	**Effect estimate (95% CI)**	***P*-value**	**Effect estimate (95% CI)**	**P-value**
PC ae C36:4	-0.64(-0.93,-0.35)	**1.65E-05**	-0.55(-0.85,-0.24)	4.99E-04	-0.73(-1.21,-0.25)	3.53E-03
PC ae C38:5	-0.59(-0.88,-0.30)	**7.69E-05**	-0.49(-0.80,-0.18)	2.08E-03	-0.65(-1.14,-0.16)	1.14E-02
PC ae C38:6	-0.77(-1.05,-0.49)	**1.09E-07**	-0.66(-0.95,-0.36)	**1.60E-05**	-0.32(-0.77,0.14)	1.76E-01
PC aa C36:0	-0.74(-1.03,-0.45)	**4.47E-07**	-0.69(-1.02,-0.37)	**3.84E-05**	0.25(-0.24,0.74)	3.29E-01
PC aa C38:0	-0.65(-0.94,-0.36)	**8.75E-06**	-0.64(-0.96,-0.32)	**1.07E-04**	-0.21(-0.70,0.28)	4.05E-01
Ornithine	-0.70(-0.98,-0.43)	**5.71E-07**	-0.41(-0.74,-0.09)	1.38E-02	-0.59(-1.12,-0.07)	2.85E-02

	it-T2D vs. NGT		it-T2D vs. IGT		it-T2D vs. ndt-T2D	
	Effect estimate (95% CI)	***P*-value**	**Effect estimate (95% CI)**	***P*-value**	**Effect estimate (95% CI)**	***P*-value**
PC ae C36:4	0.04(-0.47,0.55)	8.78E-01	0.15(-0.43,0.73)	6.11E-01	-0.08(-0.57,0.41)	7.57E-01
PC ae C38:5	0.01(-0.49,0.51)	9.76E-01	0.12(-0.46,0.70)	6.91E-01	-0.09(-0.57,0.38)	7.03E-01
PC ae C38:6	-0.44(-0.93,0.05)	8.12E-02	-0.32(-0.88,0.23)	2.57E-01	-0.24(-0.69,0.22)	3.11E-01
PC aa C36:0	-0.61(-1.10,-0.11)	1.60E-02	-0.44(-1.05,0.18)	1.63E-01	-0.54(-1.03,-0.05)	3.26E-02
PC aa C38:0	-0.19(-0.68,0.31)	4.55E-01	-0.17(-0.78,0.43)	5.77E-01	-0.34(-0.82,0.14)	1.69E-01
Ornithine	-0.29(-0.76,0.19)	2.34E-01	0.20(-0.42,0.82)	5.27E-01	0.05(-0.44,0.54)	8.39E-01

	ndt-T2D vs. NGT		ndt-T2D vs. IGT		NGT vs. IGT	
	Effect estimate (95% CI)	***P*-value**	**Effect estimate (95% CI)**	***P*-value**	**Effect estimate (95% CI)**	***P*-value**
PC ae C36:4	0.04(-0.16,0.24)	6.80E-01	-0.02(-0.23,0.18)	8.24E-01	-0.09(-0.21,0.02)	1.20E-01
PC ae C38:5	0.03(-0.17,0.22)	7.81E-01	-0.04(-0.25,0.16)	6.78E-01	-0.11(-0.22,0.00)	6.04E-02
PC ae C38:6	0.16(-0.03,0.35)	1.01E-01	0.05(-0.14,0.25)	6.04E-01	-0.14(-0.25,-0.03)	1.56E-02
PC aa C36:0	0.08(-0.11,0.28)	4.13E-01	-0.02(-0.23,0.20)	8.91E-01	-0.17(-0.28,-0.05)	3.99E-03
PC aa C38:0	-0.08(-0.27,0.12)	4.27E-01	-0.10(-0.31,0.11)	3.63E-01	-0.09(-0.20,0.03)	1.29E-01
Ornithine	0.24(0.05,0.43)	1.16E-02	-0.02(-0.24,0.19)	8.23E-01	**-0.21(-0.32,-0.10)**	**1.82E-04**

SUPPLEMENTARY DATA

Supplementary Table 5. Sensitivity analysis for the comparison between mt-T2D and ndt-T2D in the discovery KORA F4

Effect estimates and *P* values were calculated with multivariable linear regression analysis adjusted for the full set of covariates and additionally 1) the years since the T2D diagnosis; 2) waist; 3) LDL-C; and 4) LDL-C with insulin. All newly diagnosed T2D patients in the follow up were assigned with 0 years for the duration of diabetes. CI denotes confidence interval. Significant *P* values are indicated bold (*P* value<3.8E-04).

Metabolite	Full model + duration of diabetes		Full model + waist		Full model + LDL-C		Full model + LDL-C + insulin	
	Effect estimate (95% CI)	*P*-value	Effect estimate (95% CI)	*P*-value	Effect estimate (95% CI)	*P*-value	Effect estimate (95% CI)	*P*-value
PC ae C36:4	-0.69(-1.06,-0.31)	**6.08E-05**	-0.66(-0.91,-0.41)	**6.54E-07**	-0.60(-0.85,-0.35)	**6.09E-06**	-0.61(-0.87,-0.36)	**4.21E-06**
PC ae C38:5	-0.56(-0.97,-0.15)	**9.02E-05**	-0.62(-0.88,-0.36)	**4.79E-06**	-0.54(-0.80,-0.29)	**5.10E-05**	-0.55(-0.81,-0.29)	**5.23E-05**
PC ae C38:6	-0.64(-1.04,-0.25)	**1.82E-04**	-0.58(-0.82,-0.33)	**5.69E-06**	-0.47(-0.71,-0.23)	**1.42E-04**	-0.48(-0.72,-0.24)	**1.04E-04**
PC aa C36:0	-0.69(-1.12,-0.26)	2.33E-03	-0.56(-0.83,-0.29)	**5.57E-05**	-0.49(-0.76,-0.22)	**4.64E-04**	-0.50(-0.77,-0.23)	**3.51E-04**
PC aa C38:0	-0.74(-1.15,-0.32)	7.2E-04	-0.68(-0.94,-0.42)	**6.59E-07**	-0.57(-0.82,-0.32)	**1.72E-05**	-0.57(-0.83,-0.32)	**1.67E-05**
Ornithine	-0.26(-0.68,0.16)	0.23	-0.58(-0.85,-0.32)	**3.10E-05**	-0.57(-0.84,-0.29)	**6.81E-05**	-0.58(-0.86,-0.30)	**6.11E-05**

SUPPLEMENTARY DATA

Supplementary Table 6. Longitudinal analysis of the effect of metformin on metabolite profiles
Linear mixed effect model (with metabolite as dependent and the group as independent variable) adjusted for both the crude and the full set of covariates in the longitudinal study of 912 participants with no anti-diabetic medical treatment at baseline KORA S4. Of these participants, 43 started metformin treatment after the baseline KORA S4 study. A sensitivity analysis was conducted in a subset of 55 participants who were ndt-T2D patients at KORA S4, 36 of them took metformin during the follow-up. Significant *P* values are indicated bold (*P* value<0.05).

Metabolites	Crude model Effect estimate (95% CI)	*P*-value	Full model Effect estimate (95% CI)	*P*-value
With 912 participants				
PC ae C36:4	-0.65 (-0.92,-0.36)	**2.38E-06**	-0.67 (-0.94,-0.41)	**1.01E-06**
PC ae C38:5	-0.65 (-0.91,-0.37)	**1.53E-06**	-0.60 (-0.87,-0.35)	**5.18E-06**
PC ae C38:6	-0.60 (-0.85,-0.35)	**2.32E-06**	-0.54 (-0.78,-0.30)	**1.30E-05**
Sensitivity analysis with subset of 55 participants				
PC ae C36:4	-0.69 (-1.13, -0.25)	**3.7E-03**	-0.70 (-1.15, -0.24)	**4.35E-03**
PC ae C38:5	-0.43 (-0.86, 0.01)	0.06	-0.45 (-0.89, -0.01)	**0.05**
PC ae C38:6	-0.36 (-0.80, 0.06)	0.10	-0.45 (-0.86, -0.03)	**0.04**

SUPPLEMENTARY DATA

Supplementary Table 7. Associations between metformin treatment and change in lipid profile in the prospective KORA S4→F4 study

The associations were calculated using linear mixed effect models of 912 participants with no anti-diabetic medical treatment at baseline KORA S4. Of these participants, 43 started metformin treatment after the baseline KORA S4 study. Statin users (n = 247) in either S4 or F4 were excluded from the analysis to rule out any potential influence of statin use. In total, 757 participants with longitudinal KORA S4→F4 data were used for the sensitivity analysis, 28 of them took metformin only during the follow-up. The full models were modified for investigation of the HDL-C and triglycerides. The linear mixed effect models (with the lipid levels as outcome and the grouping variable as predictor) were adjusted for age and sex (crude model). The full model was additionally adjusted for BMI, physical activity, alcohol intake, smoking, systolic BP, HbA_{1c}, fasting glucose, and usage of statin, beta blockers, ACE inhibitors, and ARB. Associations for LDL-C and total cholesterol were additionally adjusted for HDL-C and triglycerides, whereas the associations for HDL-C and triglycerides were additionally adjusted for LDL-C and total cholesterol. Significant *P* values are indicated bold (*P* value<0.05).

	Crude model		**Full model**	
	Effect estimate (95% CI)	***P*-value**	**Effect estimate (95% CI)**	***P*-value**
With 912 participants				
LDL-C	-13.14 (-22.88, -3.40)	**8.34E-03**	-11.83 (-21.51, -2.15)	**0.02**
HDL-C	0.05 (-3.03, 3.12)	0.98	0.38 (-2.60, 3.36)	0.80
Total cholesterol	-19.16 (-29.77, -8.55)	**4.23E-04**	-16.18 (-26.03, -6.34)	**0.02**
Triglycerides	-7.44 (-23.45, 8.57)	0.36	-2.01 (-17.54, 13.52)	0.80
Sensitivity analysis with subset of 757 participants				
LDL-C	-9.77(-19.49, -0.05)	**0.04**	-8.68(-19.92, -0.85)	**0.04**
HDL-C	-0.78(-4.62, 3.05)	0.69	0.40(-3.35, 4.15)	0.83
Total cholesterol	-14.03(-24.47, -3.58)	**0.01**	-11.63(-21.23, -2.04)	**0.02**
Triglycerides	-1.59(-20.35, 17.17)	0.87	-5.18(-23.35, 12.99)	0.58

SUPPLEMENTARY DATA

Supplementary Table 8. Mediation analysis of the associations between metformin, LDL-C and the three metabolites

The associations were calculated using multivariable linear regression models with crude and full adjustments for 912 participants with no anti-diabetic medical treatment at baseline KORA S4. Of these participants, 43 started metformin treatment after the baseline KORA S4 study. Significant *P* values are indicated bold (*P* value<0.05).

	Crude model		**Full model**	
	Effect estimate (95% CI)	***P*-value**	**Effect estimate (95% CI)**	***P*-value**
A) Association between metformin treatment and LDL-C adjusted for different metabolites (Direct effect of metformin on LDL-C)				
Adjusted for PC ae C36:4	-10.1 (-19.74,-0.47)	**0.04**	-5.71 (-14.39,2.96)	0.2
Adjusted for PC ae C38:5	-9.98 (-19.62,-0.33)	**0.04**	-5.65 (-14.34,3.04)	0.2
Adjusted for PC ae C38:6	-7.74 (-17.21,1.73)	0.11	-3.41 (-11.95,5.12)	0.43
Adjusted for summed concentration*	-9.16 (-18.76,0.43)	0.06	-5.04 (-13.69,3.62)	0.25
B) Association between metformin treatment and metabolites				
PC ae C36:4	-0.62 (-0.88,-0.36)	**2.51E-06**	-0.53 (-0.78,-0.28)	**3.72E-05**
PC ae C38:5	-0.63 (-0.89,-0.37)	**1.97E-06**	-0.49 (-0.74,-0.25)	**9.80E-05**
PC ae C38:6	-0.60 (-0.85,-0.35)	**2.76E-06**	-0.45 (-0.69,-0.22)	**1.71E-04**
Summed concentration*	-1.85 (-2.56,-1.14)	**3.12E-07**	-1.47 (-2.15,-0.79)	**2.16E-05**
C) Association between metabolites and LDL-C adjusted for metformin treatment				
PC ae C36:4	4.81 (3.19,6.44)	**7.22E-09**	4.76 (2.94,6.57)	**2.95E-07**
PC ae C38:5	4.59 (2.97,6.21)	**3.22E-08**	4.77 (2.94,6.60)	**3.64E-07**
PC ae C38:6	8.59 (6.97,10.21)	**9.57E-25**	8.81 (6.95,10.66)	**2.69E-20**
Summed concentration*	2.31 (1.72,2.90)	**1.49E-14**	2.33 (1.67,3.00)	**6.87E-12**
D) Association between metformin treatment and total cholesterol adjusted for different metabolites (Direct effect of metformin on total cholesterol)				
Adjusted for PC ae C36:4	-16.45 (-26.93,-5.96)	**2.13E-03**	-8.54 (-17.67,0.59)	0.07
Adjusted for PC ae C38:5	-16.63 (-27.10,-6.16)	**1.87E-03**	-8.57 (-17.70,0.56)	0.07
Adjusted for PC ae C38:6	-15.45 (-25.76,-5.13)	**3.38E-03**	-6.66 (-15.55,2.23)	0.14
Adjusted for summed concentration*	-15.71 (-26.11,-5.31)	**3.09E-03**	-7.10 (-16.12,1.91)	0.12
E) Association between metabolites and total cholesterol adjusted for metformin treatment				
PC ae C36:4	8.02 (6.32,9.73)	**4.45E-20**	5.81 (3.86,7.76)	**5.87E-09**
PC ae C38:5	7.94 (6.24,9.63)	**7.76E-20**	5.60 (3.64,7.57)	**2.59E-08**
PC ae C38:6	11.69 (10.01,13.37)	**5.52E-41**	9.19 (7.19,11.1)	**3.24E-19**
Summed concentration*	3.58 (2.97,4.19)	**3.61E-30**	2.64 (1.92,3.35)	**5.25E-13**

*The summed concentration refers to the overall concentration of the three metabolites (PC ae C36:4, PC ae C38:5, PC ae C38:6)

SUPPLEMENTARY DATA

Supplementary Table 9. Mediation effects of the three metabolites for the association between metformin treatment and LDL-C reduction after removing statin users

Statin users (n = 247) in either S4 or F4 were excluded from the analysis to rule out any potential influence of statin use. In total, 757 participants with longitudinal KORA S4→F4 data were used for the sensitivity mediation analysis. 28 of these participants took metformin only during the follow-up. The mediation effects for three metformin-associated metabolites were assessed individually and as a combined total concentration using both the crude model and the full model. Significant *P* values are indicated bold (*P* value<0.05).

	Crude model		**Full model**	
	Effect estimate (95% CI)	***P*-value**	**Effect estimate (95% CI)**	***P*-value**
A) Mediation of the effect of metformin on LDL-C reduction				
PC ae C36:4	-3.34 (-4.97, -1.71)	**2.59E-04**	-1.11 (-1.74, -0.48)	**0.04**
PC ae C38:5	-3.28 (-4.85, -1.71)	**2.45E-04**	-0.84 (-1.29, -0.39)	0.07
PC ae C38:6	-5.42 (-8.81, -2.03)	**3.80E-05**	-2.77 (-4.73, -0.81)	**0.006**
Summed concentration	-4.18 (-6.28, -2.08)	**5.46E-05**	-1.73 (-2.75, -0.71)	**0.016**
B) Mediation of the effect of metformin on total cholesterol reduction				
PC ae C36:4	-2.85 (-5.48, -0.21)	**0.01**	-1.10 (-1.87, -0.33)	0.08
PC ae C38:5	-2.63 (-4.90, -0.36)	**0.01**	-0.71 (-1.16, -0.26)	0.14
PC ae C38:6	-4.15 (-8.51, -1.76)	**0.01**	-1.46 (-2.89, -0.03)	0.09
Summed concentration	-3.74 (-7.26, -0.22)	**0.01**	-1.26 (-2.20, -0.32)	0.07

SUPPLEMENTARY DATA

Supplementary Table 10. Genes with associated common variants linked to the three acyl-alkyl PCs identified by PSEA
Abbreviations, clusters, functions, associated diseases of 17 genes identified by PSEA are reported in the first four columns, respectively. The chromosome number and location (minimum and maximum) are listed in the following columns. The considered number of SNPs considered by using PSEA is shown in the last column. All listed genes showed a P<1.0E-04 in a permutation test in the PSEA.

Gene	Gene cluster	Functions	Associated diseases	Chr.	Position (min. – max.)	No. of SNPs
C7orf42	1	Unclear		7	65913652-66100806	83
SLC26A4	2	Transport (chloride, iodide, sulfate)	Deafness (1)	7	106978397-107183738	146
LOC286002[§]	2	Transcription		7	107045357-107198952	114
DAGLA	3	Insulin secretion, lipid metabolism	Spinocerebellar ataxia (2)	11	61094821-61309777	122
C11orf9	3	Transcription	Retinopathy (3)	11	61166755-61352140	104
FEN1	3	DNA repair, DNA transcription	Cancer (breast (4; 5), ovarian gastrointestinal (6), lung (7))	11	61210766-61360086	96
DKFZP434K028	3	Unclear		11	61238622-61391664	106
FADS2	3	Fatty acid metabolism	Retinopathy (8), Cornary artery disease (9)	11	61244266-61430694	131
C11orf10	3	Unclear		11	61273486-61426522	114
MIR611[§]	3	Transcription		11	61277244-61426522	112
FADS1	3	Fatty acid metabolism	Diabetes Retinopathy (10-13)	11	61287413-61448917	119
MIR1908[§]	3	Transcription		11	61300075-61448917	112
FADS3	3	Fatty acid metabolism	Diabetes (14), Coronary artery disease (15), Hyperlipidemia (16)	11	61358484-61525549	145
BEST1	3	Transport (chloride, calcium)	Retinopathy (17)	11	61365899-61527475	138
RAB3IL1	3	Transport (protein)	Retinopathy (18)	11	61381461-61551489	144
MANSC1	4	Unclear	Cancer (prostate, breast, lymphoma, leukemia) (19; 20)	12	12333540-12502043	159
C20orf94 (*SLX4IP*)	5	Unclear (possibly DNA repair)	Childhood acute lymphoblastic leukemia (21)	20	10254387-10591850	432

[§]Gene has no proteins in *Homo sapiens* in STITCH (22).

SUPPLEMENTARY DATA

Supplementary Table 11. Known metformin target genes

Known metformin target genes were retrieved from DrugBank (23).

Metformin target gene	Full biochemical name
AMPK	AMP-activated protein kinase
SLC22A1 (also known as *OCT1*)	solute carrier family 22, member 1
SLC22A2 (also known as *OCT2*)	solute carrier family 22, member 2
SLC22A3 (also known as *OCT3*)	solute carrier family 22, member 3
SLC47A1 (also known as *MATE1*)	solute carrier family 47, member 1
SLC29A4 (also known as *hENT4* and *PMAT*)	solute carrier family 29, member 4

Supplementary Table 12. Interactions between metformin target genes, pathway related proteins and PSEA identified genes

The first two columns show the names of the pairwise interaction partners in the investigated network. The following columns describe the observed interaction with literature and the underlying species.

Action 1	Action 2	Interaction	Species
Metformin	AMPK	In liver and muscle: Metformin →AMPK (24; 25); In hypothalamus: Metformin -| AMPK (26)	Rat Rat
AMPK	SREBP1c	In liver: AMPK -| SREBP1c (27)	Mouse
Leptin	AMPK	In liver: Leptin →AMPK (28); In hypothalamus: Leptin -| AMPK (29)	Mouse Mouse
SREBP1c	FADS2	In liver: SREBP1c → FADS2 (30)	Mouse
SREBP1c	FADS1	In liver: SREBP1c → FADS1 (30)	Mouse
Leptin	SREBP1c	In liver: Leptin -| SREBP1c (31)	Mouse
Leptin	FADS2	In liver: Leptin -| FADS2 (32)	Rat

SUPPLEMENTARY DATA

Reference

1. Reiisi S, Sanati MH, Tabatabaiefar MA, Ahmadian S, Reiisi S, Parchami S, Porjafari H, Shahi H, Shavarzi A, Hashemzade Chaleshtori M: The Study of SLC26A4 Gene Causing Autosomal Recessive Hearing Loss by Linkage Analysis in a Cohort of Iranian Populations. International journal of molecular and cellular medicine 2014;3:176-182

2. Knight MA, Hernandez D, Diede SJ, Dauwerse HG, Rafferty I, van de Leemput J, Forrest SM, Gardner RJ, Storey E, van Ommen GJ, Tapscott SJ, Fischbeck KH, Singleton AB: A duplication at chromosome 11q12.2-11q12.3 is associated with spinocerebellar ataxia type 20. Human molecular genetics 2008;17:3847-3853

3. Stohr H, Marquardt A, White K, Weber BH: cDNA cloning and genomic structure of a novel gene (C11orf9) localized to chromosome 11q12-->q13.1 which encodes a highly conserved, potential membrane-associated protein. Cytogenetics and cell genetics 2000;88:211-216

4. Abdel-Fatah TM, Russell R, Albarakati N, Maloney DJ, Dorjsuren D, Rueda OM, Moseley P, Mohan V, Sun H, Abbotts R, Mukherjee A, Agarwal D, Illuzzi JL, Jadhav A, Simeonov A, Ball G, Chan S, Caldas C, Ellis IO, Wilson DM, 3rd, Madhusudan S: Genomic and protein expression analysis reveals flap endonuclease 1 (FEN1) as a key biomarker in breast and ovarian cancer. Molecular oncology 2014;8:1326-1338

5. Lv Z, Liu W, Li D, Liu L, Wei J, Zhang J, Ge Y, Wang Z, Chen H, Zhou C, Yuan Q, Zhou L, Yang M: Association of functional FEN1 genetic variants and haplotypes and breast cancer risk. Gene 2014;538:42-45

6. Liu L, Zhou C, Zhou L, Peng L, Li D, Zhang X, Zhou M, Kuang P, Yuan Q, Song X, Yang M: Functional FEN1 genetic variants contribute to risk of hepatocellular carcinoma, esophageal cancer, gastric cancer and colorectal cancer. Carcinogenesis 2012;33:119-123

7. Yang M, Guo H, Wu C, He Y, Yu D, Zhou L, Wang F, Xu J, Tan W, Wang G, Shen B, Yuan J, Wu T, Lin D: Functional FEN1 polymorphisms are associated with DNA damage levels and lung cancer risk. Human mutation 2009;30:1320-1328

8. Meyers KJ, Mares JA, Igo RP, Jr., Truitt B, Liu Z, Millen AE, Klein M, Johnson EJ, Engelman CD, Karki CK, Blodi B, Gehrs K, Tinker L, Wallace R, Robinson J, LeBlanc ES, Sarto G, Bernstein PS, SanGiovanni JP, Iyengar SK: Genetic evidence for role of carotenoids in age-related macular degeneration in the Carotenoids in Age-Related Eye Disease Study (CAREDS). Investigative ophthalmology & visual science 2014;55:587-599

9. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, Penninx BW, Janssens AC, Wilson JF, Spector T, Martin NG, Pedersen NL, Kyvik KO, Kaprio J, Hofman A, Freimer NB, Jarvelin MR, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks AA, Falchi M, Willemsen G, Hottenga JJ, de Geus EJ, Montgomery GW, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands EJ, Uitterlinden AG, Witteman JC, Oostra BA, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Doring A, Wichmann HE, Smit JH, McCarthy MI, van Duijn CM, Peltonen L, Consortium E: Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nat Genet 2009;41:47-55

10. Goodarzi MO, Guo X, Cui J, Jones MR, Haritunians T, Xiang AH, Chen YD, Taylor KD, Buchanan TA, Hsueh WA, Raffel LJ, Rotter JI: Systematic evaluation of validated type 2 diabetes and glycaemic trait loci for association with insulin clearance. Diabetologia 2013;56:1282-1290

11. Neale BM, Fagerness J, Reynolds R, Sobrin L, Parker M, Raychaudhuri S, Tan PL, Oh EC, Merriam JE, Souied E, Bernstein PS, Li B, Frederick JM, Zhang K, Brantley MA, Jr., Lee AY, Zack DJ, Campochiaro B, Campochiaro P, Ripke S, Smith RT, Barile GR, Katsanis N, Allikmets R, Daly MJ, Seddon JM: Genome-wide association study of advanced age-related macular degeneration identifies a

role of the hepatic lipase gene (LIPC). Proceedings of the National Academy of Sciences of the United States of America 2010;107:7395-7400
12. Fauser S, Smailhodzic D, Caramoy A, van de Ven JP, Kirchhof B, Hoyng CB, Klevering BJ, Liakopoulos S, den Hollander AI: Evaluation of serum lipid concentrations and genetic variants at high-density lipoprotein metabolism loci and TIMP3 in age-related macular degeneration. Investigative ophthalmology & visual science 2011;52:5525-5528
13. Meyers KJ, Johnson EJ, Bernstein PS, Iyengar SK, Engelman CD, Karki CK, Liu Z, Igo RP, Jr., Truitt B, Klein ML, Snodderly DM, Blodi BA, Gehrs KM, Sarto GE, Wallace RB, Robinson J, LeBlanc ES, Hageman G, Tinker L, Mares JA: Genetic determinants of macular pigments in women of the Carotenoids in Age-Related Eye Disease Study. Investigative ophthalmology & visual science 2013;54:2333-2345
14. Huang T, Sun J, Chen Y, Xie H, Xu D, Huang J, Li D: Genetic variants in desaturase gene, erythrocyte fatty acids, and risk for type 2 diabetes in Chinese Hans. Nutrition 2014;30:897-902
15. Kwak JH, Paik JK, Kim OY, Jang Y, Lee SH, Ordovas JM, Lee JH: FADS gene polymorphisms in Koreans: association with omega6 polyunsaturated fatty acids in serum phospholipids, lipid peroxides, and coronary artery disease. Atherosclerosis 2011;214:94-100
16. Plaisier CL, Horvath S, Huertas-Vazquez A, Cruz-Bautista I, Herrera MF, Tusie-Luna T, Aguilar-Salinas C, Pajukanta P: A systems genetics approach implicates USF1, FADS3, and other causal candidate genes for familial combined hyperlipidemia. PLoS genetics 2009;5:e1000642
17. Alapati A, Goetz K, Suk J, Navani M, Al-Tarouti A, Jayasundera T, Tumminia SJ, Lee P, Ayyagari R: Molecular diagnostic testing by eyeGENE: analysis of patients with hereditary retinal dystrophy phenotypes involving central vision loss. Investigative ophthalmology & visual science 2014;55:5510-5521
18. Bogdanov P, Corraliza L, J AV, Carvalho AR, Garcia-Arumi J, Ramos D, Ruberte J, Simo R, Hernandez C: The db/db mouse: a useful model for the study of diabetic retinal neurodegeneration. PLoS One 2014;9:e97302
19. Haiman CA, Han Y, Feng Y, Xia L, Hsu C, Sheng X, Pooler LC, Patel Y, Kolonel LN, Carter E, Park K, Le Marchand L, Van Den Berg D, Henderson BE, Stram DO: Genome-wide testing of putative functional exonic variants in relationship with breast and prostate cancer risk in a multiethnic population. PLoS genetics 2013;9:e1003419
20. Haferlach C, Bacher U, Kohlmann A, Schindela S, Alpermann T, Kern W, Schnittger S, Haferlach T: CDKN1B, encoding the cyclin-dependent kinase inhibitor 1B (p27), is located in the minimally deleted region of 12p abnormalities in myeloid malignancies and its low expression is a favorable prognostic marker in acute myeloid leukemia. Haematologica 2011;96:829-836
21. Meissner B, Bartram T, Eckert C, Trka J, Panzer-Grumayer R, Hermanova I, Ellinghaus E, Franke A, Moricke A, Schrauder A, Teigler-Schlegel A, Dorge P, von Stackelberg A, Basso G, Bartram CR, Kirschner-Schwabe R, Bornhauser B, Bourquin JP, Cazzaniga G, Hauer J, Attarbaschi A, Izraeli S, Zaliova M, Cario G, Zimmermann M, Avigad S, Sokalska-Duhme M, Metzler M, Schrappe M, Koehler R, Te Kronnie G, Stanulla M: Frequent and sex-biased deletion of SLX4IP by illegitimate V(D)J-mediated recombination in childhood acute lymphoblastic leukemia. Human molecular genetics 2014;23:590-601
22. Kuhn M, Szklarczyk D, Pletscher-Frankild S, Blicher TH, von Mering C, Jensen LJ, Bork P: STITCH 4: integration of protein-chemical interactions with user data. Nucleic acids research 2014;42:D401-407
23. Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, Pon A, Banco K, Mak C, Neveu V, Djoumbou Y, Eisner R, Guo AC, Wishart DS: DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic acids research 2011;39:D1035-1041

24. Lim CT, Kola B, Korbonits M: AMPK as a mediator of hormonal signalling. Journal of molecular endocrinology 2010;44:87-97
25. Zhou G, Myers R, Li Y, Chen Y, Shen X, Fenyk-Melody J, Wu M, Ventre J, Doebber T, Fujii N, Musi N, Hirshman MF, Goodyear LJ, Moller DE: Role of AMP-activated protein kinase in mechanism of metformin action. The Journal of clinical investigation 2001;108:1167-1174
26. Chau-Van C, Gamba M, Salvi R, Gaillard RC, Pralong FP: Metformin inhibits adenosine 5'-monophosphate-activated kinase activation and prevents increases in neuropeptide Y expression in cultured hypothalamic neurons. Endocrinology 2007;148:507-511
27. Awazawa M, Ueki K, Inabe K, Yamauchi T, Kaneko K, Okazaki Y, Bardeesy N, Ohnishi S, Nagai R, Kadowaki T: Adiponectin suppresses hepatic SREBP1c expression in an AdipoR1/LKB1/AMPK dependent pathway. Biochemical and biophysical research communications 2009;382:51-56
28. Miyamoto L, Ebihara K, Kusakabe T, Aotani D, Yamamoto-Kataoka S, Sakai T, Aizawa-Abe M, Yamamoto Y, Fujikura J, Hayashi T, Hosoda K, Nakao K: Leptin activates hepatic 5'-AMP-activated protein kinase through sympathetic nervous system and alpha1-adrenergic receptor: a potential mechanism for improvement of fatty liver in lipodystrophy by leptin. The Journal of biological chemistry 2012;287:40441-40447
29. Minokoshi Y, Alquier T, Furukawa N, Kim YB, Lee A, Xue B, Mu J, Foufelle F, Ferre P, Birnbaum MJ, Stuck BJ, Kahn BB: AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus. Nature 2004;428:569-574
30. Matsuzaka T, Shimano H, Yahagi N, Amemiya-Kudo M, Yoshikawa T, Hasty AH, Tamura Y, Osuga J, Okazaki H, Iizuka Y, Takahashi A, Sone H, Gotoda T, Ishibashi S, Yamada N: Dual regulation of mouse Delta(5)- and Delta(6)-desaturase gene expression by SREBP-1 and PPARalpha. Journal of lipid research 2002;43:107-114
31. Tobe K, Suzuki R, Aoyama M, Yamauchi T, Kamon J, Kubota N, Terauchi Y, Matsui J, Akanuma Y, Kimura S, Tanaka J, Abe M, Ohsumi J, Nagai R, Kadowaki T: Increased expression of the sterol regulatory element-binding protein-1 gene in insulin receptor substrate-2(-/-) mouse liver. The Journal of biological chemistry 2001;276:38337-38340
32. Gallardo N, Bonzon-Kulichenko E, Fernandez-Agullo T, Molto E, Gomez-Alonso S, Blanco P, Carrascosa JM, Ros M, Andres A: Tissue-specific effects of central leptin on the expression of genes involved in lipid metabolism in liver and white adipose tissue. Endocrinology 2007;148:5604-5610