GENETICS/GENOMES/PROTEOMICS/METABOLOMICS

Deciphering the Plasma Proteome of Type 2 Diabetes

Mohamed A. Elhadad,[1,2,3] Christian Jonasson,[4,5] Cornelia Huth,[2,6] Rory Wilson,[1,2] Christian Gieger,[1,2,6] Pamela Matias,[1,2,3] Harald Grallert,[1,2,6] Johannes Graumann,[7,8] Valerie Gailus-Durner,[9] Wolfgang Rathmann,[6,10] Christine von Toerne,[11] Stefanie M. Hauck,[11] Wolfgang Koenig,[3,12,13] Moritz F. Sinner,[3,14] Tudor I. Oprea,[15,16,17] Karsten Suhre,[18] Barbara Thorand,[2,6] Kristian Hveem,[4,5] Annette Peters,[2,3,6,19] and Melanie Waldenberger[1,2,3]

Diabetes 2020;69:2766–2778 | https://doi.org/10.2337/db20-0296

With an estimated prevalence of 463 million affected, type 2 diabetes represents a major challenge to health care systems worldwide. Analyzing the plasma proteomes of individuals with type 2 diabetes may illuminate hitherto unknown functional mechanisms underlying disease pathology. We assessed the associations between type 2 diabetes and >1,000 plasma proteins in the Cooperative Health Research in the Region of Augsburg (KORA) F4 cohort (*n* = 993, 110 cases), with subsequent replication in the third wave of the Nord-Trøndelag Health Study (HUNT3) cohort (*n* = 940, 149 cases). We computed logistic regression models adjusted for age, sex, BMI, smoking status, and hypertension. Additionally, we investigated associations with incident type 2 diabetes and performed two-sample bidirectional Mendelian randomization (MR) analysis to prioritize our results. Association analysis of prevalent type 2 diabetes revealed 24 replicated proteins, of which 8 are novel. Proteins showing association with incident type 2 diabetes were aminoacylase-1, growth hormone receptor, and insulin-like growth factor–binding protein 2. Aminoacylase-1 was associated with both prevalent and incident type 2 diabetes. MR analysis yielded nominally significant causal effects of type 2 diabetes on cathepsin Z and rennin, both known to have roles in the pathophysiological pathways of cardiovascular disease, and of sex hormone–binding globulin on type 2 diabetes. In conclusion, our high-throughput proteomics study replicated previously reported type 2 diabetes–protein associations and identified new candidate proteins possibly involved in the pathogenesis of type 2 diabetes.

[1]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[2]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[3]German Research Center for Cardiovascular Disease (DZHK), Partner Site Munich Heart Alliance, Munich, Germany
[4]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health, Norwegian University of Science and Technology, Trondheim, Norway
[5]HUNT Research Center, Department of Public Health, Norwegian University of Science and Technology, Levanger, Norway
[6]German Center for Diabetes Research (DZD), München-Neuherberg, Germany
[7]Biomolecular Mass Spectrometry, Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany
[8]The German Centre for Cardiovascular Research (DZHK), Partner Site Rhine-Main, Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany
[9]German Mouse Clinic, Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[10]Institute of Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[11]Research Unit Protein Science, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[12]Deutsches Herzzentrum München, Technische Universitat München, Munich, Germany
[13]Institute of Epidemiology and Medical Biometry, University of Ulm, Ulm, Germany
[14]Department of Medicine I, University Hospital Munich, Ludwig Maximilian University, Munich, Germany
[15]Department of Internal Medicine and UNM Comprehensive Cancer Center, University of New Mexico School of Medicine, Albuquerque, NM
[16]Department of Rheumatology and Inflammation Research, Institute of Medicine, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
[17]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[18]Department of Biophysics and Physiology, Weill Cornell Medicine - Qatar, Education City, Doha, Qatar
[19]Institute of Medical Information Sciences, Biometry and Epidemiology, Ludwig Maximilian University, Munich, Germany

Corresponding authors: Melanie Waldenberger, waldenberger@helmholtz-muenchen.de, and Mohamed A. Elhadad, mohamed.elhadad@helmholtz-muenchen.de

Received 2 April 2020 and accepted 28 August 2020

This article contains supplementary material online at https://doi.org/10.2337/figshare.12896426.

© 2020 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at https://www.diabetesjournals.org/content/license.

Type 2 diabetes is a significant cause of morbidity and mortality, with an estimated worldwide prevalence of 463 million patients, one-half of whom are undiagnosed (1). It is a complex, multifactorial disease characterized by an interplay of both genetic and nongenetic factors that lead to insulin resistance and hyperinsulinemia (1,2). Moreover, type 2 diabetes causes widespread microvascular and macrovascular complications, resulting in significant health care expenditure (1).

The proteomics of type 2 diabetes, the investigation of a set of proteins within different tissues of diabetic animal models, and the comparison of patients with diabetes with healthy control subjects have enabled the discovery of new protein–type 2 diabetes associations (3–5). Examples of associations include adiponectin (3), leptin (5), and insulin-like growth factor–binding protein 2 (IGFBP-2) (4). Of particular clinical interest is the study of type 2 diabetes associations with plasma proteins, which reflect systemic effects and may serve as predictive biomarkers (3,5–7).

The integration of genetic and proteomic knowledge has provided new insight into the pathophysiology of type 2 diabetes. The best example is Mendelian randomization (MR), a method used to infer causality in observational study settings (4,8). Previous MR studies of biomarkers and type 2 diabetes have suggested causal protective roles for proteins like adiponectin, β-carotene, N-terminal proB-type natriuretic peptide, and sex hormone–binding globulin (SHBG) as well as causal harmful roles of delta-6 desaturase and ferritin (7).

Here, we use a highly multiplexed aptamer-based proteomics platform to analyze the associations between prevalent type 2 diabetes and 1,095 plasma proteins in the Cooperative Health Research in the Region of Augsburg (KORA) study. We replicate our results in the independent Nord-Trøndelag Health Study (HUNT) study and investigate associations with incident type 2 diabetes using follow-up data from KORA and HUNT. Moreover, we test the performance of our newly discovered biomarkers to predict incident type 2 diabetes when added to an adapted version of the updated German Diabetes Risk Score (GDRS) (9). We then evaluate these newly identified proteins using the protein-protein interaction resource STRING (10). Finally, we applied two-sample bidirectional MR analysis (11) to assess causality and prioritize the newly discovered relationships.

RESEARCH DESIGN AND METHODS

Study Populations

KORA Cohort

The KORA study comprises independent samples from Augsburg in southern Germany (12). In the current study, we used a subsample of 1,000 individuals randomly selected from the participants of the KORA F4 survey (N = 3,080, performed 2006–2008) with deep phenotyping data (n = 1,800) (13). Detailed clinical and sociodemographic information was collected. Data from the KORA FF4 survey (performed 2013–2014) represents the 7-year follow-up of KORA F4. The ethics committee of the Bavarian Medical Association (Berlin, Germany) reviewed and approved the study, and all participants gave written informed consent.

HUNT Cohort

HUNT is a prospective population-based cohort from Nord-Trøndelag County in Norway (14). We used the HUNT3 survey (n = 1,117 with proteomics measurements, performed 2006–2008) for the validation of the KORA study results. The HUNT study collected detailed sociodemographic and clinical information. We used linked primary care and hospital registries for information on diabetes status at 9 years follow-up. All study participants provided written informed consent.

Proteomics Measurement

Proteins were measured in fasting and nonfasting plasma samples in KORA and HUNT, respectively, using the SOMAscan platform as described previously (13,15). In summary, plasma and bead-coupled aptamers, each of which has a high affinity toward a specific protein, were incubated. After washing steps, bead-bound proteins were biotinylated, and complexes comprising biotinylated target proteins and fluorescence-labeled aptamers were photocleaved off the bead support and pooled. Following recapture on streptavidin beads and further washing steps, aptamers were eluted and quantified as a proxy to protein concentration by hybridization to custom arrays of aptamer-complementary oligonucleotides. On the basis of standard samples included on each plate, the resulting raw intensities were processed using a data analysis workflow that included hybridization normalization, median signal normalization, and signal calibration to control for interplate differences (16). Raw intensities are reported in relative florescence units.

In KORA, one sample failed SOMAscan quality control, leaving 999 samples for analysis. Of the 1,129 SOMAmer probes (SOMAscan assay version 3.2), 29 failed SOMAscan quality control. We also removed the five probes recommended by the SOMAscan assay change log issued on 22 December 2016, leaving 1,095 probes for analysis. For replication, we used the HUNT probes that passed quality control.

Definition of Outcome and Model Covariates

In KORA, type 2 diabetes was defined as self-reported disease validated by the responsible physician or medical chart review or as current use of glucose-lowering medication. All participants without known diabetes were assigned to receive a standard 75-g oral glucose tolerance test (3). Prevalent type 2 diabetes refers to participants with the disease at the time of blood sample collection, and incident refers to those developing type 2 diabetes after that time point within a 7- and 9-year follow-up period in KORA and HUNT, respectively.

In HUNT, prevalent type 2 diabetes was self-reported, which we validated using clinical data from hospitals and primary care registries using the ICD-10 code E11 and the International Classification of Primary Care, Second Edition, code T90. We identified incident cases of type 2 diabetes from the same registries using identical codes.

We classified participants of both cohorts who participated in leisure time physical activity for at least 1 h/week as physically active (more details are available in the Supplementary Material). Current hypertension was defined in KORA as having a systolic blood pressure ≥140 mmHg, diastolic blood pressure ≥90 mmHg, and/or use of antihypertensive medication. In HUNT, we used hospital and primary care data and ICD-10 codes I10–I15 and International Classification of Primary Care, Second Edition, codes K86 or K87 to identify participants with hypertension.

Drugs were assessed in KORA by asking the participants to bring the packages of their medication and supplements with them to their study center visit. Using database software (17), medications were identified using Anatomical Therapeutic Chemical codes, medication identifier bar code, or product name.

Statistical Analysis

Preprocessing of the quality controlled SOMAscan data was the same in both cohorts and involved $\log_2$ transformation and (0, 1) standardization by subtracting the per-cohort mean and dividing by the per-cohort SD to allow easier interpretation of the odds ratios (ORs) per SD of the protein.

Proteome-Wide Analysis

Using logistic regression, we ran two proteome-wide analyses in KORA: associations between proteins with prevalent type 2 diabetes and with incident type 2 diabetes. For each of the two outcomes, we ran one model per protein (i.e., 1,095 models per outcome) and adjusted for the potential confounders age, sex, BMI, smoking status, and current hypertension at baseline. We then replicated the results in HUNT using the same model. We excluded participants from both cohorts with missing values for the confounders, which led to the sample sizes of 993 and 940 for KORA and HUNT, respectively. For the analysis with incident type 2 diabetes, we further excluded all participants with prevalent type 2 diabetes, resulting in sample sizes of 881 and 794 for KORA and HUNT, respectively. We used the false discovery rate (FDR) Benjamini-Hochberg method separately for the outcomes to account for multiple testing. An association was considered statistically significant at FDR <0.05.

We replicated significant results in HUNT using the same model. We considered proteins replicated at FDR <0.05, with FDR calculated on the basis of the number of significant proteins in KORA. To examine whether antidiabetic drug intake influenced the replicated associations, we ran sensitivity analyses by including the drugs of interest as confounders one at a time.

Data Analytics of Replicated Proteins

The candidate proteins were processed through the Pharos (18) platform, experimental Gene Ontology (19) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) (20) pathways, human disease association data from the GWAS Catalog (21), Online Mendelian Inheritance in Man (22), and the text-mining DISEASES platform (23) as well as through phenotype data from the corresponding mouse ortholog knockouts (24). We mined these resources for data on the potential associations between the candidate proteins and type 2 diabetes.

Prediction of Incident Diabetes

We applied a biomarker discovery strategy to investigate whether proteins significantly associated with incident type 2 diabetes in KORA could be used for prediction of incident type 2 diabetes. These were 10 proteins, of which 1 failed quality control in HUNT. We used KORA as a training data set and HUNT as a test data set and used an adapted version of GDRS, a diabetes risk score that was trained in 21,845 participants of the European Prospective Investigation into Cancer and Nutrition Potsdam (EPIC-Potsdam) study with a mean follow-up time of 7 years, as a benchmark (9). More details on the GDRS are available in the Supplementary Material.

Adaptation of the GDRS was necessary because some of the variables were missing from one or both cohorts. We defined smoking status using only information on current and former smoking per se without regard to the number of cigarettes smoked. We used the average of the original GDRS score weights for each smoking category to represent our combined categories (former: [15 + 45] / 2 = 30; current: [23 + 77] / 2 = 50). Family history of diabetes was defined in KORA as having at least one parent or sibling with diabetes and in HUNT, as having at least one parent, sibling, or child with diabetes. We calculated the risk of a positive family history by averaging the original GDRS of having one parent, both parents, or at least one sibling with type 2 diabetes ([56 + 106 + 48] / 3 = 70). Our final adapted GDRS was calculated as follows: 5.1 × age in years + 7.6 × waist circumference in cm – 2.7 × height in cm + 47 × hypertension status − 2 × physical activity (at least 1 h/week) + 30 × former smoking + 50 × current smoking + 70 × family history of type 2 diabetes.

We tested prediction performance of all models using the receiver operating characteristic area under the curve (AUC) and applied the DeLong test to compare AUCs of nested models. First, we added the proteins to the adapted GDRS model and applied the least absolute shrinkage and selection operator (LASSO) (25) for model selection in KORA. LASSO shrinks the sum of the absolute values of the regression coefficients, forcing some to be set to 0, thus performing a form of model selection. The LASSO λ was chosen by cross-validation using the squared error for Gaussian models. The GDRS was calculated in each cohort and used as a score that was fixed by setting its penalty factor

to 0 to prevent any shrinkage by LASSO. We then compared the performance of the LASSO protein-extended GDRS model to the adapted GDRS model. We assessed the calibration of the LASSO-selected model using calibration plots (26). Moreover, we tested the performance of the proteins as single predictors on top of the adapted GDRS model.

MR

We attempted to infer causality of the replicated proteins associated with type 2 diabetes by applying two-sample bidirectional MR. Figure 1 shows the summary of the pipeline for the causal inference analysis. In summary, we extracted single nucleotide polymorphisms (SNPs) as instrumental variables (IVs) from published genome-wide association study (GWAS) summary statistics of European ancestry if they passed the Bonferroni threshold of $P <$ 5e-8. We extracted the IVs from the meta-analysis of type 2 diabetes GWAS studies by Xue et al. (27) (N = 455,607) and the GWAS studies of SOMAscan-measured proteins by Sun et al. (28) (N = 3,301), Suhre et al. (13) (N = 1,000), and Emilsson et al. (29) (N = 5,457) for proteins. We identified ambiguous palindromic SNPs, defined as SNPs with A/T or G/C alleles and an effect allele frequency of ~0.5 using the cutoff points defined by the TwoSampleMR package in R (30). We replaced these with a proxy SNP, defined as a SNP with r^2 >0.85 with the SNP in question, when available, or excluded them from further analyses

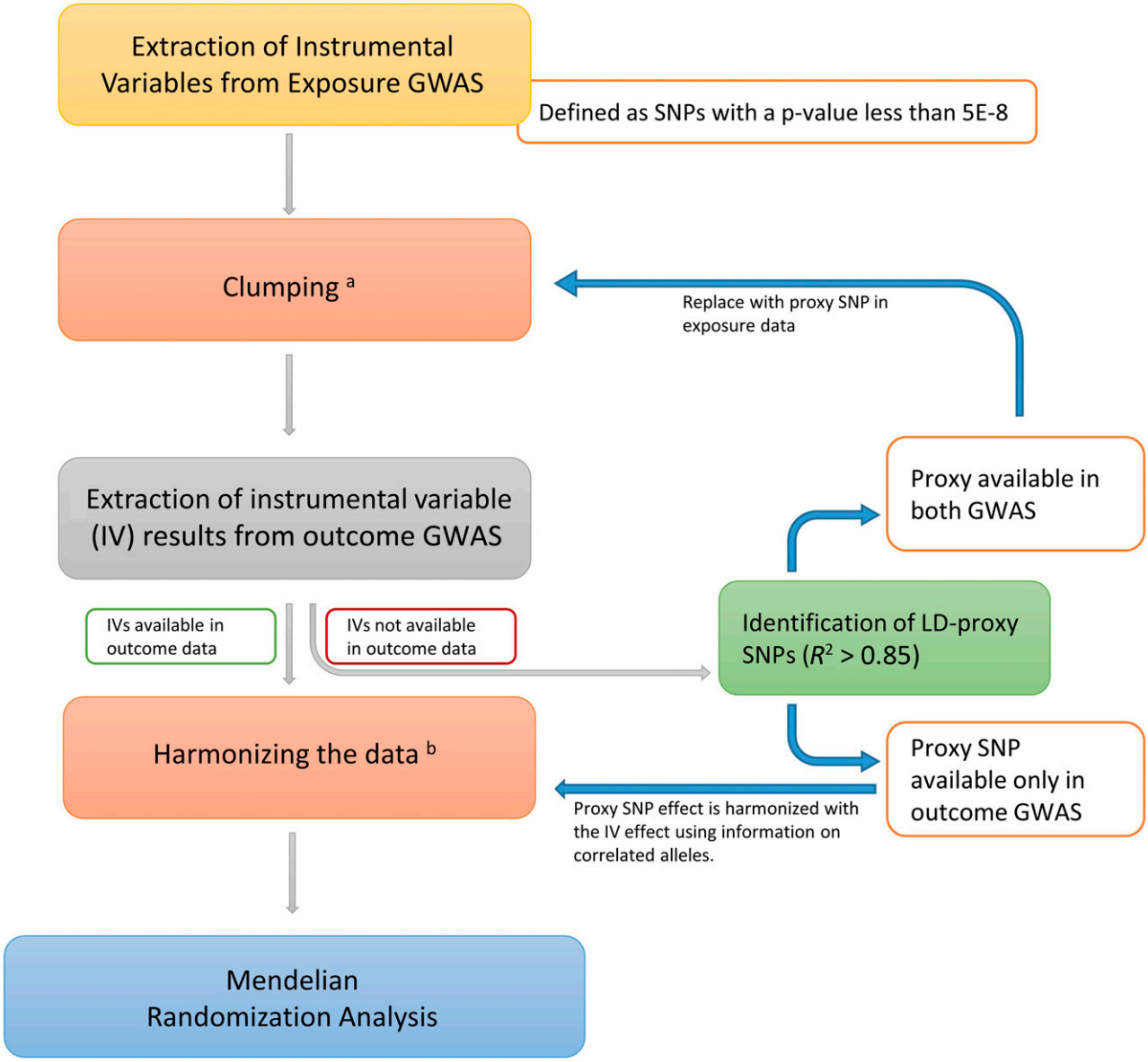

Figure 1—MR analysis flowchart. [a]Clumping refers to the process of selecting only the independent IVs (i.e., those that are not in linkage disequilibrium [LD] with one another) using the cutoff LD r^2 >0.001. [b]Harmonizing the data refers to ensuring that the effects of the IV on the exposure and the outcome reflect the same strand effect.

Table 1—Baseline characteristics of the prevalent study populations

Variable	KORA (n = 993)	HUNT (n = 940)	P value*
Age (years)	59.31 (43–79)	69.03 (31.6–99.4)	<0.001
Sex female	514 (51.8)	245 (26.1)	<0.001
BMI (kg/m^2), mean (SD)	27.79 (4.58)	28.36 (3.96)	0.003
Waist circumference (cm), mean (SD)	94.51 (13.81)	100.01 (11.01)	<0.001
Physical inactivity	376 (37.9)	472 (49.2)	<0.001
Smoking status			<0.001
Never	423 (42.6)	234 (24.9)	
Former	422 (42.5)	504 (53.6)	
Current	148 (14.9)	202 (21.5)	
Family history of diabetes	312 (38.1)	280 (31.6)	0.005
Hypertension	396 (39.9)	389 (41.4)	0.531

Data are mean (range) or n (%) unless otherwise indicated. *Continuous variables were tested for a difference between the two populations using t tests and categorical variables with χ^2 tests with continuity correction.

(31). We then clumped the SNPs, which implies removing SNPs in linkage disequilibrium with the lead SNP using the r^2 cutoff 0.001. We did not manually prune the final list of IVs. Furthermore, IVs selected for proteins needed to be in *cis* (i.e., within 1 Mb of the protein-coding gene as per Human Genome Assembly GRCh37.p13).

We proceeded to extract the results of these IVs or of one of their proxies from the outcome's GWAS. For proteins, priority was given to results from Sun et al. (28) because of the larger sample size, followed by Suhre et al. (13) dependent on availability.

We used the Wald ratio to check for causality (32). In cases of more than one IV, we used the random effects model of the inverse variance–weighted meta-analysis to combine the Wald ratio estimates of all IVs (8,32). For sensitivity analyses, whenever there was more than one IV, we ran the MR-Egger regression model to look for horizontal pleiotropy in our causal models (33) and leave-one-out analysis and forest plots to identify outliers among these IVs that would be driving the results in a certain direction and examined scatter plots to check for outliers.

Analytical steps are summarized in Supplementary Fig. 1. All analyses were done using R version 3.5.1 software (The R Foundation for Statistical Computing). For MR analysis, the TwoSampleMR package of R version 0.4.22 was used (30).

Data and Resource Availability

Informed consents given by KORA study participants do not cover data posting in public databases. However, the KORA data are available given approval of online requests at the KORA Project Application Self-Service Tool (https://epi.helmholtz-muenchen.de). The HUNT data can be accessed given approval of applications to the HUNT Research Centre (https://www.ntnu.edu/hunt/data). The data used in the MR analysis are publicly available and can be accessed through https://cnsgenomics.com/content/data (Xue et al. [27]), https://www.phpc.cam.ac.uk/ceu/proteins (Sun et al. [28]), https://metabolomics.helmholtz-muenchen.de/pgwas/ (Suhre et al. [13]), and www.sciencemag.org/content/361/6404/769/suppl/DC1 (Emilsson et al. [29]). Example code for the analytic steps of the article can be accessed at https://github.com/maelhadad/T2D_SOMAscan_Proteomics.

RESULTS

Descriptive Statistics of the Study Populations

Table 1 and Supplementary Table 1 show the baseline characteristics of both cohorts and their follow-up subsets, respectively. HUNT participants were on average older and comprised more men.

Association Results of Plasma Proteins With Type 2 Diabetes

The proteome-wide analysis with prevalent type 2 diabetes yielded 85 FDR-significant proteins (Supplementary Table 2), of which 24 successfully replicated in HUNT (Table 2 and Fig. 2*A*). Of these, osteomodulin was most strongly associated (on the basis of KORA P value) with an OR-per-SD increase in protein level of 0.61 (95% CI 0.47–0.77) in KORA and of 0.65 (0.53–0.79) in HUNT. Among the positively associated proteins, peptide YY (PYY) had the strongest association (1.34 [1.1–1.62] in KORA and 1.58 [1.32–1.92] in HUNT).

To assess whether the proteome panel was associated with future type 2 diabetes, we performed a proteome-wide analysis with incident type 2 diabetes using the same model, which yielded 10 FDR-significant protein associations (Supplementary Table 3). Of these, aminoacylase-1, growth hormone receptor, and IGFBP-2 replicated in HUNT (Table 3 and Fig. 2*B*). Adiponectin failed quality control in HUNT, and thus, replication was not possible. Among the replicated proteins, aminoacylase-1 showed the strongest association (OR 1.78 [95% CI 1.34–2.37] in KORA and 1.6 [1.26–2.05] in HUNT). Interestingly, aminoacylase-1 overlapped between the replicated results of both prevalent and incident type 2 diabetes.

Table 2—Results of the proteome-wide analysis with prevalent type 2 diabetes in KORA, for those proteins statistically significant in the discovery (FDR <0.05), and their replicated results (FDR <0.05, same direction of effect) in HUNT

			KORA (n = 993)			HUNT (n = 940)		
Protein full name	Protein short name	UniProt identifier	OR (95% CI)	*P* value	FDR *P* value	OR (95% CI)	*P* value	FDR *P* value
α-L-Iduronidase	IDUA	P35475	1.48 (1.2–1.84)	3.04e-04	1.07e-02	1.44 (1.19–1.74)	1.59e-04	1.42e-03
Aminoacylase-1	Aminoacylase-1	Q03154	2.1 (1.64–2.71)	5.62e-09	2.05e-06	1.49 (1.22–1.84)	1.26e-04	1.26e-03
Apolipoprotein B	Apo B	P04114	0.48 (0.37–0.61)	4.19e-09	2.05e-06	0.7 (0.57–0.84)	2.87e-04	1.86e-03
Cathepsin Z	CATZ	Q9UBR2	1.41 (1.13–1.77)	2.27e-03	3.35e-02	1.33 (1.1–1.62)	3.20e-03	1.37e-02
Cerebral dopamine neurotrophic factor	ARMEL	Q49AH0	0.64 (0.48–0.82)	7.25e-04	1.85e-02	0.7 (0.55–0.87)	1.83e-03	8.62e-03
Complement C2	C2	P06681	2.01 (1.37–3.04)	6.63e-04	1.81e-02	1.47 (1.2–1.82)	3.03e-04	1.86e-03
Galectin-3–binding protein	LG3BP	Q08380	1.6 (1.27–2.01)	5.04e-05	2.51e-03	1.43 (1.2–1.72)	9.47e-05	1.08e-03
Gelsolin	Gelsolin	P06396	0.55 (0.43–0.69)	4.31e-07	9.44e-05	0.66 (0.54–0.81)	4.31e-05	6.88e-04
Hepatocyte growth factor receptor	Met	P08581	0.62 (0.49–0.78)	4.89e-05	2.51e-03	0.78 (0.65–0.92)	4.93e-03	1.88e-02
Kallikrein-7	Kallikrein 7	P49862	0.59 (0.46–0.75)	1.47e-05	1.46e-03	0.67 (0.54–0.82)	1.95e-04	1.56e-03
Lysosomal protective protein	Cathepsin A	P10619	1.54 (1.24–1.92)	8.51e-05	3.88e-03	1.32 (1.09–1.6)	5.48e-03	1.91e-02
Matrilin-2	MATN2	O00339	0.62 (0.49–0.77)	2.77e-05	2.17e-03	0.7 (0.57–0.86)	7.17e-04	3.82e-03
Osteomodulin	OMD	Q99983	0.61 (0.47–0.77)	3.89e-05	2.34e-03	0.64 (0.52–0.78)	1.22e-05	3.26e-04
Peptide YY	PYY	P10082	1.34 (1.1–1.62)	3.36e-03	4.59e-02	1.53 (1.27–1.86)	9.26e-06	3.26e-04
Periostin	Periostin	Q15063	0.54 (0.43–0.68)	1.52e-07	4.16e-05	0.75 (0.62–0.92)	5.50e-03	1.91e-02
Plasma protease C1 inhibitor	C1-esterase inhibitor	P05155	0.67 (0.53–0.84)	5.39e-04	1.59e-02	0.76 (0.61–0.93)	9.66e-03	3.22e-02
Renin	Renin	P00797	1.61 (1.32–1.99)	5.48e-06	6.67e-04	1.45 (1.21–1.74)	5.16e-05	6.88e-04
RGM domain family member B	RGMB	Q6NW40	0.64 (0.49–0.81)	3.52e-04	1.20e-02	0.73 (0.59–0.9)	3.25e-03	1.37e-02
Sex hormone-binding globulin	SHBG	P04278	0.62 (0.47–0.8)	2.41e-04	9.11e-03	0.63 (0.51–0.77)	1.12e-05	3.26e-04
SLIT and NTRK-like protein 5	SLIK5	O94991	0.6 (0.47–0.76)	3.81e-05	2.34e-03	0.78 (0.66–0.93)	4.38e-03	1.75e-02
Transforming growth factor-β receptor type 3	TGFbR3	Q03167	0.58 (0.46–0.73)	4.45e-06	6.09e-04	0.74 (0.61–0.89)	1.40e-03	6.99e-03
Trypsin-1	Trypsin	P07477	0.63 (0.5–0.78)	4.06e-05	2.34e-03	0.7 (0.58–0.84)	2.19e-04	1.59e-03
Tumor necrosis factor–inducible gene 6 protein	TSG-6	P98066	0.58 (0.45–0.74)	2.27e-05	2.07e-03	0.7 (0.57–0.85)	3.96e-04	2.26e-03
Wnt inhibitory factor 1	WIF1	Q9Y5W5	0.5 (0.37–0.66)	2.97e-06	4.75e-04	0.65 (0.54–0.79)	2.21e-05	4.43e-04

Sorted alphabetically. All analyses were adjusted for age, sex, BMI, smoking status, and hypertension.

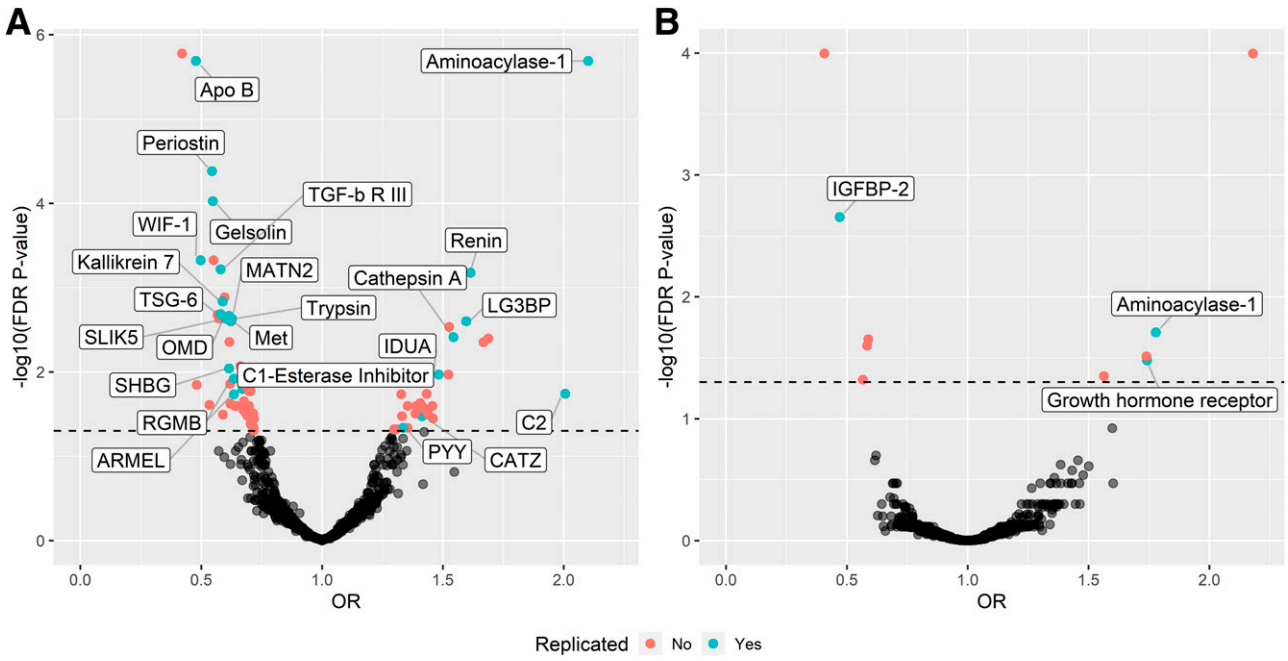

Figure 2—Volcano plot of type 2 diabetes results in KORA, where proteins that replicated in HUNT are labeled. *A*: Results of the proteome-wide analysis with prevalent type 2 diabetes in KORA. *B*: Results of the proteome-wide analysis with incident type 2 diabetes in KORA.

Additionally, we assessed the concordance of the effect estimates across the cohorts. Of 85 KORA FDR-significant proteins associated with prevalent type 2 diabetes, only 7 had different effect directions, but none of these was nominally significant in HUNT (Fig. 3*A*). For incident type 2 diabetes, two proteins showed opposite effect directions, with neither of these reaching nominal statistical significance (Fig. 3*B*).

Overlap With Known Type 2 Diabetes Genetic and Protein Associations

To assess the overlap between our results and known type 2 diabetes associations, we compared our results to gene-based results described by Xue et al. (27). Furthermore, we compared our replicated proteins with protein lists of interest published by the Human Diabetes Proteome Project, namely the 1,000 diabetes-related proteins, the human islet of Langerhans proteome, the rodent β-cell proteome, and the human blood glycated proteome (34). Of the 26 unique replicated proteins, 18 overlapped with at least one list. Eight proteins have not been previously found to be related to type 2 diabetes (Supplementary Table 4).

Data Analytics of Replicated Proteins

Supplementary Table 5 shows information extracted from Pharos for our replicated proteins. α-L-Iduronidase, cathepsin A, and cathepsin Z shared the same lysosomal pathway association according to KEGG (20).

Investigating Potential Effects of Drugs on Type 2 Diabetes–Protein Associations in KORA

None of the replicated protein-incident type 2 diabetes associations showed loss of significance when adjusting for any of the investigated drugs. On the other hand, three of the replicated associations with prevalent type 2 diabetes lost statistical significance when adjusting for antidiabetic medication intake (Supplementary Table 6 and Supplementary Fig. 2). All the associations retained the same direction of effect apart from PYY, which showed an opposite effect after adjusting for antidiabetic medication and, more specifically, metformin.

Prediction of Incident Type 2 Diabetes

Starting with the nine proteins associated with incident type 2 diabetes in KORA available in HUNT, we evaluated whether a subset of them selected using LASSO would improve the predictive performance of the adapted GDRS benchmark model (9). LASSO selected five proteins, namely transforming growth factor-β receptor type 3 (TGFbR3), tartrate-resistant acid phosphatase type 5, pappalysin-1, afamin, and scavenger receptor cysteine-rich type 1 protein M130 (sCD163). The LASSO-selected protein-enhanced model showed improvement in both KORA and HUNT (GDRS protein-extended AUC 0.84 [95% CI 0.79–0.89] and 0.67 [0.61–0.72]; GDRS-only AUC 0.77 [0.71–0.83] and 0.66 [0.60–0.72], respectively); however, according to the Delong test, the AUC improvement in HUNT was not statistically significant ($P = 0.72$) (Supplementary Fig. 3). The calibration plot of the LASSO-selected model in HUNT yielded an intercept of 0.23 and a slope of 0.53 (Supplementary Fig. 3). The intercept of the calibration plot examines the difference of means of predicted and observed risk. In HUNT, it is >0, thus showing higher observed type 2 diabetes cases in HUNT as those predicted. This could be attributed to longer follow-up in HUNT (9 years vs. 7 years in KORA)

Table 3—Results of the proteome-wide analysis with incident type 2 diabetes in KORA, for those proteins statistically significant in the discovery (FDR <0.05), and their replicated results (FDR <0.05, same direction of effect) in HUNT

			KORA (n = 881)			HUNT (n = 794)		
Protein full name	Protein short name	UniProt identifier	OR (95% CI)	*P* value	FDR *P* value	OR (95% CI)	*P* value	FDR *P* value
Aminoacylase-1	Aminoacylase-1	Q03154	1.78 (1.34–2.37)	7.15e-05	1.96e-02	1.6 (1.26–2.04)	1.27e-04	1.14e-03
Growth hormone receptor	Growth hormone receptor	P10912	1.74 (1.31–2.38)	2.43e-04	3.32e-02	1.42 (1.07–1.88)	1.37e-02	4.11e-02
Insulin-like growth factor–binding protein 2	IGFBP-2	P18065	0.47 (0.34–0.65)	6.07e-06	2.22e-03	0.57 (0.42–0.77)	2.91e-04	1.31e-03

Sorted alphabetically. All analyses were adjusted for age, sex, BMI, smoking status, and hypertension.

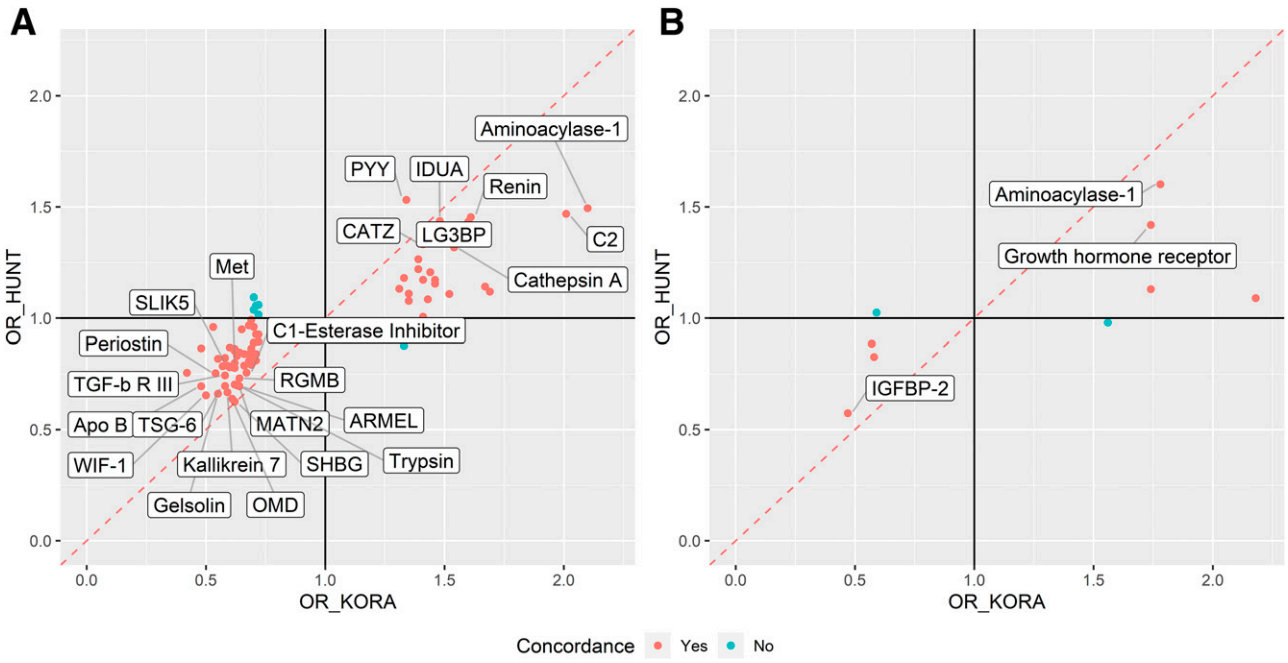

Figure 3—Coefficient concordance between KORA and HUNT for prevalent type 2 diabetes (*A*) and incident type 2 diabetes (*B*). Proteins that replicated in HUNT are labeled.

and to the fact that HUNT is older than KORA and therefore has more cases of type 2 diabetes. The slope of calibration is 0.53 in HUNT, which indicates a possible overfitting of the model or the need for coefficient shrinkage in HUNT that could also be attributed to the heterogeneity between the study populations in terms of patient characteristics and outcome definition. The training data set used gold standard screening to define type 2 diabetes, where HUNT did not apply a similar definition and would therefore have hidden cases and measurement error. Therefore, the outcome being predicted for HUNT (and defined by KORA) is slightly different from the outcome observed.

We further tested the performance of individual proteins as predictors of incident type 2 diabetes in KORA and validated our models in HUNT (Supplementary Fig. 4). The following proteins showed relatively similar performance in both cohorts: aminoacylase-1 (KORA AUC 0.78 [95% CI 0.73–0.84]; HUNT AUC 0.71 [0.65–0.77]), growth hormone receptor (KORA AUC 0.77 [0.71–0.83]; HUNT AUC 0.70 [0.64–0.76]), and IGFBP-2 (KORA AUC 0.78 [0.72–0.84]; HUNT AUC 0.73 [0.68–0.79]).

MR Analysis of Replicated Plasma Proteins and Type 2 Diabetes in KORA

Using up to 120 SNPs as genetic instruments, we investigated whether type 2 diabetes had a causal effect on the 26 replicated proteins from both the prevalence and the incidence analyses (Supplementary Table 7 and Supplementary Fig. 5). For cathepsin Z (MR inverse variance–weighted $\beta = 0.13$; $P = 2.00\text{e-}03$) and renin (0.08; $P = 3.15\text{e-}02$), a nominally significant causal effect of prevalent type 2 diabetes was observed, each with the same direction of effect as its observational results. MR-Egger analyses to test for the presence of horizontal pleiotropy showed no significant results for either protein (intercept $P = 0.17$ and 0.1 for cathespin Z and renin, respectively). Tests and plots to check for outliers in the IVs showed no significant aberrations (Supplementary Figs. 6 and 7).

We also ran MR to investigate whether any of the proteins had a causal effect on type 2 diabetes. We analyzed 13 proteins for which we found independent *cis*-acting IVs (Supplementary Table 8 and Supplementary Fig. 5). We observed a nominally significant causal effect of SHBG on type 2 diabetes, with the same direction of effect as its observed association (MR Wald $\beta = -0.09$; $P = 2.95\text{e-}02$). None of the associations for either direction survived Bonferroni multiple testing correction.

DISCUSSION

We report a proteome-wide analysis of type 2 diabetes in KORA and replication in HUNT using aptamer-based affinity proteomics. Our analysis yielded 26 unique replicated significant protein associations. Of these, 24 replicated exclusively with prevalent type 2 diabetes, 2 replicated exclusively with incident type 2 diabetes, and aminoacylase-1 replicated with both.

Aminoacylase-1 is a zinc-dependent peptidase involved in amino acid metabolism (35). The protein has not been described in the context of type 2 diabetes before but has been reported to be overexpressed in obese liver tissue, thus linking it to obesity and inflammation (36). A further study found aminoacylase-1 to be downregulated in obese omental fat, which the authors hypothesized to be due to adipocyte dysfunction caused by obesity (35). Moreover,

aminoacylase-1 is associated with arginine production according to KEGG (20). Plasma levels of arginine were found to be higher in patients with type 2 diabetes (37).

In addition to aminoacylase-1, incident type 2 diabetes results included an inverse association with IGFBP-2 and a positive association with growth hormone receptor. IGFBP-2 was reported to have type 2 diabetes protective effects and has been shown to reverse hyperglycemia in insulin and leptin deficiency (38). These associations highlight the role of the growth hormone axis in the early pathophysiology of type 2 diabetes. Both growth hormone and IGF-I are known to play roles in the insulin receptor cascade, leading to insulin resistance (39).

The analysis of prevalent type 2 diabetes confirmed previously known proteomic associations like gelsolin (40), renin (41), SHBG (42), and hepatocyte growth factor receptor and revealed promising new candidate proteins, including osteomodulin, matrilin-2, Wnt inhibitory factor-1 (WIF1), tumor necrosis factor–inducible gene 6 protein (TNFAIP6), cerebral dopamine neurotrophic factor (CDNF), RGM domain family member B, TGFbR3, and SLIT and NTRK-like protein 5, which were downregulated in type 2 diabetes cases, and lysosomal protective protein, galectin-3 binding protein (LGALS3BP), and PYY, which were upregulated.

Our results overlap and complement results of mass spectrometry studies on obesity. Plasma levels of apolipoprotein B, LGALS3BP, and SHBG were found to be altered by sustained weight loss (43) and gastric bypass surgery–induced weight loss (44), with the latter affecting also plasma protease C1 inhibitor, complement C2, and gelsolin.

New protein associations with prevalent type 2 diabetes included proteins previously reported in association with complications of type 2 diabetes. Increased circulating levels of LGALS3BP were linked to nonalcoholic fatty liver disease (45) and acute venous thrombosis (46), and TGFbR3 was reported to be associated with diabetic nephropathy (47). TNFAIP6 and CDNF were shown to have protective effects, while WIF1, TGFbR3, and PYY were reported to have harmful effects, in the development and progress of cardiovascular atherosclerotic diseases (48–52). Along this line, members of the complement family like plasma protease C1 inhibitor and complement C2 were downregulated and upregulated, respectively, in our results, and proteins from the renin-angiotensin and kallikrein-kinin systems included the upregulated renin and downregulated kallikrein-7.

Although our study cohorts were different regarding the fasting status of their samples, most proteins (78 of 85 for prevalent and 8 of 10 for incident type 2 diabetes) showed concordant effects between cohorts, while none of the nonconcordant proteins were statistically significant in the replication (Fig. 3). Nonetheless, fasting has significant metabolic consequences that are expected to be reflected in the plasma proteome and could have contributed to nonreplication in HUNT. However, there are multiple other potential explanations for the nonreplication, perhaps differing from one protein to another. Importantly, while plasma protein levels differ between fasting and nonfasting samples, this does not necessarily match the variance in the protein levels caused by the disease status. As such, disease-related variance would still be apparent despite differences in fasting status. For example, some of our examined proteins were reported to show differences in their levels according to fasting status, like SHBG (53), PYY (53), and soluble CD163 (54), yet their associations with disease status were replicated in our study. Of the proteins that failed replication in HUNT, MMP2 (53) and pappalysin-1 (55) have been found to be affected by food intake. However, their effect sizes were similar in both cohorts, suggesting that fasting status may not be the primary reason for nonreplication for most proteins.

Sensitivity analyses into the potential effect of drugs on the type 2 diabetes–plasma protein associations showed loss of significance of some associations after adjusting for antidiabetic medication. The effect direction of all the resultant associations remained the same except for PYY, which showed a change of direction after adjusting for metformin intake; however, because its effect estimate was not significant after adjustment, it is difficult to draw any conclusions from this.

Additionally, we evaluated the significant proteins' ability to predict incident type 2 diabetes. The protein-extended models showed improved performance over the adapted GDRS benchmark model (9) in both the KORA discovery and the HUNT replication, although the improvement was very small and not statistically significant ($P = 0.72$) for the latter. Moreover, we tested the performances of individual proteins on top of the adapted benchmark model. The best performances in the replication cohort came from aminoacylase-1, growth hormone receptor, and IGFBP-2, each of which achieved approximately equal performance in HUNT compared with KORA, results that warrant validation in clinical trials using commercially available ELISA kits. Because the KORA samples were taken from individuals in a fasting state (≥8 h) and HUNT samples were taken nonfasting, these results seem to indicate that fasting status is largely irrelevant with regard to type 2 diabetes prediction for these candidate biomarkers. However, fasting may potentially be relevant for other markers, since the AUC was much smaller in HUNT compared with KORA for some of the other measured biomarkers in combination with the GDRS.

Our investigations into the causal framework governing the relation between plasma proteins and type 2 diabetes showed suggestive harmful causal effects of SHBG on type 2 diabetes. SHBG has been previously reported to be associated with type 2 diabetes (42) and may be implicated in the development of insulin resistance (42). We demonstrated it to be negatively associated with type 2 diabetes, a causal direction suggested by the MR analysis as well.

Causal inference analysis showed suggestive causal effects of type 2 diabetes on both cathepsin Z and renin. In line with previous observations, we demonstrated renin to be

positively associated with type 2 diabetes in both observational and MR analysis results (41). The association is an indicator of the upregulated renin-angiotensin-aldosterone system, which is activated in obesity and type 2 diabetes, thus contributing to cardiovascular disease complications (41,56). Cathepsin Z is a member of the peptidase C1 family that plays a role in lysosomal function, which might explain its connection to diabetes through β-cell failure driven by lysosomal degradation (57).

Study Strengths

We applied a high-throughput proteomics platform on samples from population-based cohorts for our analyses, which enabled us to test a large number of proteins with a wide concentration range and to generalize our results to our samples' respective populations. We used samples from plasma, which is easily accessible and is the usual medium of biomarkers. Additionally, the plasma proteome reflects on the levels of proteins originating from a broad range of tissues, thus giving us insight into systemic pathways. Finally, we were able to test for the causal relationship in both directions using publicly available data on genetic associations with both type 2 diabetes and proteins.

Study Limitations

We are aware of several limitations to our study. First, aptamer-based proteomics is susceptible to potential probe cross-reactivity and nonspecific binding (28,29). However, we verified that none of the proteins identified have been flagged for such issues (validation data presented in Supplementary Material, Supplementary Table 10, and Supplementary Figs. 10 and 11). Because of the lack of oral glucose tolerance test data in HUNT, the rigorous definition of type 2 diabetes used in KORA could not be extended, and the discrepancy in fasting status between the cohorts may have contributed to the limited replication of our results. Our prediction models do not reflect the dynamic changes in the proteome, which would require a more detailed investigation. This is also true for the MR results, which reflect the lifelong genetic risk rather than point change in single protein levels in relation to disease status. Although, there is an overlap between the participants of the genetic data sets used for type 2 diabetes and proteins through KORA, none of the associations tested using such data were significant.

Conclusion

Our proteome-wide analysis of type 2 diabetes replicated known associations and revealed novel candidate proteins. Associations with incident type 2 diabetes included aminoacylase-1, which overlapped with prevalent type 2 diabetes associations. New associations with prevalent type 2 diabetes included TNFAIP6, CDNF, WIF1, TGFbR3, and PYY, all of which are believed to play a role in the development of cardiovascular complications, like atherosclerosis. MR suggested a causal role of SHBG on type 2 diabetes, which is in line with previous observational and MR analysis results. It also suggested a causal effect of type 2 diabetes on cathepsin Z and renin, both of which are known to play a role in type 2 diabetes complications. Our results offer insight into proteins involved in the pathogenesis of type 2 diabetes and its complications, proteins that could be valuable drug targets for all levels of prevention.

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 and by the State of Bavaria. This work was supported by Deutsche Forschungsgemeinschaft grant WA 4081/1-1 and by the Bundesministerium für Bildung und Forschung within the framework of the European Union Joint Programming Initiative "A Healthy Diet for a Healthy Life" (DIMENSION grant number 01EA1902A). This work was further supported by the Biomedical Research Program at Weill Cornell Medicine in Qatar, a program funded by the Qatar Foundation. The HUNT3 study was funded by the Norwegian Ministry of Health, Norges Teknisk-Naturvitenskapelige Universitet, Norges Forskningsråd, Helse Midt-Norge, the Nord-Trøndelag County Council, and the Norwegian Institute of Public Health. M.A.E. is supported by the Deutsches Zentrum für Herz-Kreislaufforschung. T.I.O. is supported by National Institutes of Health grants U24-CA-224370, U24-TR-002278, and U01-CA-239108. K.S. is supported by Qatar National Research Fund grant NPRPC11-0115-180010.

Duality of Interest. C.J. has received personal fees for research consultancy work from Pfizer and Bayer outside the submitted work. T.I.O. has received honoraria or consulted for Abbott, AstraZeneca, Chiron, Genentech, Infinity Pharmaceuticals, Merz Pharmaceuticals, Merck Darmstadt, Mitsubishi Tanabe, Novartis, Ono Pharmaceuticals, Pfizer, Roche, Sanofi, and Wyeth outside the submitted work. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. M.A.E. analyzed the data, interpreted the results, and wrote and revised the manuscript. M.A.E. and M.W. designed the study. R.W., V.G.-D., and T.I.O. helped with the analyses. C.J., C.H., R.W., C.G., P.M., H.G., J.G., W.R., C.v.T., S.M.H., W.K., M.F.S., K.S., B.T., K.H., A.P., and M.W. were involved in the data collection, data management, and preparation of their respective cohorts. All authors contributed to the writing of the article, critically reviewed it, and approved the final version for submission. M.A.E. and M.W. are the guarantors of this work and, as such, had full access to all the data in the study and take full responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the AHA Epidemiology and Prevention–Lifestyle and Cardiometabolic Health 2020 Scientific Sessions, Phoenix, AZ, 3–6 March 2020. This study has been published previously in abstract form in the following publication: Elhadad MA, Jonasson C, Huth C, et al. Deciphering the plasma proteome of type 2 diabetes. Circulation 2020;141(Suppl. 1):A21.

References

1. International Diabetes Federation. IDF Diabetes Atlas, 9th edition, 2019. Brussels, Belgium, International Diabetes Federation. Accessed 23 March 2020. Available from https://www.diabetesatlas.org
2. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol 2018;14:88–98
3. Huth C, von Toerne C, Schederecker F, et al. Protein markers and risk of type 2 diabetes and prediabetes: a targeted proteomics approach in the KORA F4/FF4 study. Eur J Epidemiol 2019;34:409–422
4. Beijer K, Nowak C, Sundström J, Ärnlöv J, Fall T, Lind L. In search of causal pathways in diabetes: a study using proteomics and genotyping data from a cross-sectional study. Diabetologia 2019;62:1998–2006
5. Nowak C, Sundström J, Gustafsson S, et al. Protein biomarkers for insulin resistance and type 2 diabetes risk in two large community cohorts. Diabetes 2016;65:276–284

6. von Toerne C, Huth C, de Las Heras Gala T, et al. MASP1, THBS1, GPLD1 and ApoA-IV are novel biomarkers associated with prediabetes: the KORA F4 study. Diabetologia 2016;59:1882–1892
7. Abbasi A. Mendelian randomization studies of biomarkers and type 2 diabetes. Endocr Connect 2015;4:249–260
8. Zheng J, Baird D, Borges MC, et al. Recent developments in Mendelian randomization studies. Curr Epidemiol Rep 2017;4:330–345
9. Paprott R, Mühlenbruch K, Mensink GBM, et al. Validation of the German Diabetes Risk Score among the general adult population: findings from the German Health Interview and Examination Surveys. BMJ Open Diabetes Res Care 2016;4:e000280
10. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019;47:D607–D613
11. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658–665
12. Holle R, Happich M, Löwel H, Wichmann HE; MONICA/KORA Study Group. KORA–a research platform for population based health research. Gesundheitswesen 2005;67(Suppl. 1):S19–S25
13. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome [published correction appears in Nat Commun 2017;8:15345]. Nat Commun 2017;8:14357
14. Krokstad S, Langhammer A, Hveem K, et al. Cohort profile: the HUNT study, Norway. Int J Epidemiol 2013;42:968–977
15. Ganz P, Heidecker B, Hveem K, et al. Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. JAMA 2016;315:2532–2541
16. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One 2010;5:e15004
17. Mühlberger N, Behrend C, Stark RG, Holle R. Datenbankgestutzte Online-Erfassung von Arzneimitteln im Rahmen gesundheitswissenschaftlicher Studien - Erfahrungen mit der IDOM-Software. München, Germany, Urban & Fischer, 2003
18. Nguyen DT, Mathias S, Bologa C, et al. Pharos: collating protein information to shed light on the druggable genome. Nucleic Acids Res 2017;45:D995–D1002
19. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res 2019;47:D330–D338
20. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. Nucleic Acids Res 2019;47:D590–D595
21. Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 2019;47:D1005–D1012
22. Amberger JS, Bocchini CA, Scott AF, Hamosh A. OMIM.org: leveraging knowledge across phenotype-gene relationships. Nucleic Acids Res 2019;47:D1038–D1043
23. Pletscher-Frankild S, Pallejà A, Tsafou K, Binder JX, Jensen LJ. DISEASES: text mining and data integration of disease-gene associations. Methods 2015;74:83–89
24. Koscielny G, Yaikhom G, Iyer V, et al. The International Mouse Phenotyping Consortium Web Portal, a unified point of access for knockout mice and related phenotyping data. Nucleic Acids Res 2014;42:D802–D809
25. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc B 1996;58:267–288
26. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–138
27. Xue A, Wu Y, Zhu Z, et al.; eQTLGen Consortium. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat Commun 2018;9:2941
28. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. Nature 2018;558:73–79
29. Emilsson V, Ilkov M, Lamb JR, et al. Co-regulatory networks of human serum proteins link genetics to disease. Science 2018;361:769–773
30. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. eLife 2018;7:e34408
31. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol 2016;45:1717–1726
32. Teumer A. Common methods for performing Mendelian randomization. Front Cardiovasc Med 2018;5:51
33. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512–525
34. Topf F, Schvartz D, Gaudet P, et al. The Human Diabetes Proteome Project (HDPP): from network biology to targets for therapies and prevention. Transl Proteom 2013;1:3–11
35. Pérez-Pérez R, García-Santos E, Ortega-Delgado FJ, et al. Attenuated metabolism is a hallmark of obesity as revealed by comparative proteomic analysis of human omental adipose tissue. J Proteomics 2012;75:783–795
36. Caira S, Iannelli A, Sciarrillo R, et al. Differential representation of liver proteins in obese human subjects suggests novel biomarkers and promising targets for drug development in obesity. J Enzyme Inhib Med Chem 2017;32:672–682
37. Cao Y-F, Li J, Zhang Z, et al. Plasma levels of amino acids related to urea cycle and risk of type 2 diabetes mellitus in Chinese adults. Front Endocrinol (Lausanne) 2019;10:50
38. Hedbacker K, Birsoy K, Wysocki RW, et al. Antidiabetic effects of IGFBP2, a leptin-regulated gene. Cell Metab 2010;11:11–22
39. Kim S-H, Park M-J. Effects of growth hormone on glucose metabolism and insulin resistance in human. Ann Pediatr Endocrinol Metab 2017;22:145–152
40. Khatri N, Sagar A, Peddada N, et al. Plasma gelsolin levels decrease in diabetic state and increase upon treatment with F-actin depolymerizing versions of gelsolin. J Diabetes Res 2014;2014:152075
41. Joseph JJ, Echouffo Tcheugui JB, Effoe VS, Hsueh WA, Allison MA, Golden SH. Renin-angiotensin-aldosterone system, glucose metabolism and incident type 2 diabetes mellitus: MESA. J Am Heart Assoc 2018;7:e009890
42. Le TN, Nestler JE, Strauss JF III, Wickham EP III. Sex hormone-binding globulin and type 2 diabetes mellitus. Trends Endocrinol Metab 2012;23:32–40
43. Geyer PE, Wewer Albrechtsen NJ, Tyanova S, et al. Proteomics reveals the effects of sustained weight loss on the human plasma proteome. Mol Syst Biol 2016;12:901
44. Wewer Albrechtsen NJ, Geyer PE, Doll S, et al. Plasma proteome profiling reveals dynamics of inflammatory and lipid homeostasis markers after Roux-en-Y gastric bypass surgery. Cell Syst 2018;7:601–612.e3
45. Niu L, Geyer PE, Wewer Albrechtsen NJ, et al. Plasma proteome profiling discovers novel proteins associated with non-alcoholic fatty liver disease. Mol Syst Biol 2019;15:e8793
46. DeRoo EP, Wrobleski SK, Shea EM, et al. The role of galectin-3 and galectin-3-binding protein in venous thrombosis. Blood 2015;125:1813–1821
47. Reeves WB, Andreoli TE. Transforming growth factor β contributes to progressive diabetic nephropathy. Proc Natl Acad Sci U S A 2000;97:7667–7669
48. Watanabe R, Sato Y, Ozawa N, Takahashi Y, Koba S, Watanabe T. Emerging roles of tumor necrosis factor-stimulated gene-6 in the pathophysiology and treatment of atherosclerosis. Int J Mol Sci 2018;19:465
49. Zhang G-L, Wang L-H, Liu X-Y, et al. Cerebral dopamine neurotrophic factor (CDNF) has neuroprotective effects against cerebral ischemia that may occur through the endoplasmic reticulum stress pathway. Int J Mol Sci 2018;19:1905
50. Ress C, Paulweber M, Goebel G, et al. Circulating Wnt inhibitory factor 1 levels are associated with development of cardiovascular disease. Atherosclerosis 2018;273:1–7
51. Sun F, Li X, Duan W-Q, et al. Transforming growth factor-β receptor III is a potential regulator of ischemia-induced cardiomyocyte apoptosis. J Am Heart Assoc 2017;6:e005357

52. Smith RM, Klein R, Kruzliak P, Zulli A. Role of peptide YY in blood vessel function and atherosclerosis in a rabbit model. Clin Exp Pharmacol Physiol 2015; 42:648–652
53. Pellis L, van Erk MJ, van Ommen B, et al. Plasma metabolomics and proteomics profiling after a postprandial challenge reveal subtle diet effects on human metabolic status. Metabolomics 2012;8:347–359
54. Dencker M, Gårdinger Y, Björgell O, Hlebowicz J. Effect of food intake on 92 biomarkers for cardiovascular disease. PLoS One 2017;12:e0178656
55. Dencker M, Gårdinger Y, Björgell O, Hlebowicz J. Effect of food intake on biomarkers for cardiovascular disease and inflammation analyzed with the Proseek Multiplex CVD II kit. Genet Mol Res 2018;17:gmr16039884
56. Bahtiyar G, Gutterman D, Lebovitz H. Heart failure: a major cardiovascular complication of diabetes mellitus. Curr Diab Rep 2016;16:116
57. Pasquier A, Vivot K, Erbs E, et al. Lysosomal degradation of newly formed insulin granules contributes to β cell failure in diabetes. Nat Commun 2019;10: 3312