medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Title: Plasma proteomic signatures of adiposity are associated with cardiovascular risk factors and type 2 diabetes risk in a multi-ethnic Asian population.

Authors: Charlie G.Y. Lim[1], Bige Ozkan[2,3], Yujian Liang[1], Jiali Yao[1], Nang Ei Ei Khaing[1], Mary R. Rooney[2], Chiadi E. Ndumele[2,3], E Shyong Tai[1,4], Josef Coresh[2], Xueling Sim[1,*], and Rob M. van Dam[1,5,*]

Authors affiliations:

[1] Saw Swee Hock School of Public Health, National University of Singapore, 12 Science Drive 2, Singapore 117549 (C.G.Y.L., Y.L., J.Y., N.E.E.K., E.S.T., X.S., R.M.vD.)

[2] Department of Epidemiology and the Welch Center for Prevention, Epidemiology, and Clinical Research, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA (B.O., M.R.R., C.E.N., J.C.)

[3] Ciccarone Center for the Prevention of Cardiovascular Disease, Division of Cardiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA (B.O., C.E.N.)

[4] Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, 10 Medical Dr, Singapore 117597 (E.S.T.)

[5] Departments of Exercise and Nutrition Sciences and Epidemiology, Milken Institute School of Public Health, The George Washington University, 950 New Hampshire Ave, NW, Washington, DC 20052 (R.M.vD.)

[*]These authors jointly supervised this work

Author contributions

Concept and design: C.G.Y.L., X.S., and R.M.vD. Acquisition, analysis, or interpretation of data: C.G.Y.L., X.S., R.M.vD, B.O., M.R.R., C.E.N., J.C., J.Y., Y.L., E.S.T., N.E.E.K. Drafting of manuscript: C.G.Y.L., X.S., R.M.vD. Critical revision of manuscript: E.S.T., B.O., M.R.R., C.E.N., J.C.

Disclaimers: The National University of Singapore has signed a collaboration agreement with SomaLogic to conduct SomaScan of MEC stored samples at no charge in exchange for the rights to analyze linked MEC phenotype data.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Conflict of interest: All authors report no conflict of interest.

Corresponding Authors:

Xueling Sim
Saw Swee Hock School of Public Health, National University of Singapore
12 Science Drive 2, Singapore 117549
Tel: +65 6601 5096
E-mail: ephsx@nus.edu.sg

Rob M. van Dam
Department of Exercise and Nutrition Sciences, Milken Institute School of Public Health, The George Washington University
950 New Hampshire Ave NW #2, Washington DC 20052
Tel: +1-202-994-9412
E-mail: rvandam@gwu.edu

Data availability: Summary statistics for all measured proteins are provided in the Supplementary Tables. Data from the Multi-Ethnic Cohort study can be requested by researchers for scientific purposes through an application process at the listed website (https://blog.nus.edu.sg/sphs/data-and-samples-request/). Data will be shared through an institutional data sharing agreement.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Abstract

The molecular mechanisms connecting obesity and cardiometabolic diseases are not clearly understood. We evaluated the associations between body mass index (BMI), waist circumference (WC), and ~5,000 plasma proteins in the Singapore Multi-Ethnic Cohort (MEC1). Among 410 BMI-associated and 385 WC-associated proteins, we identified protein signatures of BMI and WC and validated them in an independent dataset across two time points and externally in the Atherosclerosis Risk in Communities (ARIC) study. Among participants with more precise adiposity measurements, the BMI- and WC-protein signatures were highly correlated with total and visceral body fat, respectively. Both protein signatures were significantly associated with cardiometabolic risk factors. In prospective analyses, both protein signatures were strongly associated with type 2 diabetes (T2D) risk in both MEC1 and ARIC and explained the associations between anthropometric measurements and T2D risk. Our protein signatures have potential uses for the monitoring of metabolically unhealthy obesity and the changes therein during interventions.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Introduction

Obesity is a metabolic disorder resulting from a complex interplay between genetic, psychosocial, and environmental factors (1). Although an overwhelming body of evidence point to obesity as a major risk factor for metabolic diseases such as type 2 diabetes (T2D) and cardiovascular diseases (CVDs) (2), the molecular markers and mechanisms connecting obesity and these metabolic diseases are not fully understood.

Proteins are the interface between genetics, environment, and phenotype (3) and can therefore serve as important biomarkers for the identification of biological pathways linking genetic and environmental exposures to disease risk (4). Recent advances in high-throughput proteomic technologies have enabled the measurement of a vast number of proteins across a wide range of concentrations with good reproducibility (5), unlocking the potential for large-scale interrogation of the human proteome to uncover novel interactions between the human proteome and health. Recent studies in European populations have identified proteins associated with body mass index (BMI) (6,7), highlighting pathways involved in lipid metabolism and inflammations that may contribute to obesity-related diseases (7). However, these studies were based on platforms that covered a limited number of proteins (~1100 proteins) and used BMI as the only measure of adiposity in participants of predominantly European ancestry. The latter may be particularly important given the well-recognized differences in the association of BMI in populations of Asian Ancestry compared to those of European ancestry and the potential role of visceral adiposity in the pathophysiology of type 2 diabetes in Asia (8). Moreover, these studies did not examine whether the obesity-related proteins may mediate the cardiometabolic effects of obesity.

We evaluated the associations between ~5000 plasma proteins, and BMI and waist circumference (WC) in a multi-ethnic Asian population comprising of Chinese, Malay, and Indian adults. Using machine learning techniques, we developed protein signatures of BMI and WC and evaluated their associations with total and visceral body fat, CVD risk factors, and the incidence of T2D.

Results

The study design and baseline characteristics of study participants are shown in **Figure 1** and **Supplementary Table 1**. Participants were male (44.0%) and female (56.0%) adults with a mean (± SD) age of 47.6 (± 12.0) years. The ethnic distribution was 34.8% Chinese, 32.6% Malay, and 32.7% Indian.

Association with demographic characteristics and stability over time

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We evaluated the associations between plasma protein levels and demographic characteristics in non-overlapping participants from the Population Set and controls in the type 2 diabetes case control set (T2DCC Set). The Population Set was a random sample of 631 participants from the Singapore Multi-Ethnic Cohort Phase 1 (MEC1) (9), and the T2DCC set was a nested case-control study (616 cases, 1,200 controls) in the MEC1. At Bonferroni threshold for ~5,000 proteins tested (alpha = 1.00×10^{-05}), 25.8% (n = 1,286) of the proteins were significantly associated with sex, 17.7% (n = 879) were associated with age, and 27.3% (n = 1,358) were associated with ethnicity (**Supplementary Table 2**). Of the significant associations in our study, 76.6% of the associations with age and 84.2% of the associations with sex were also significant and directionally consistent with a recent proteome-wide association study among 35,559 Icelanders (10) (**Supplementary Figure 1**).

We also evaluated the reproducibility of protein levels among this group of individuals without T2D and measured at both baseline and follow-up (n = 1,085) over a mean duration of 6.4 (± 1.6) years. About 59% of the proteins had median fold change values ranging from 0.95 to 1.05 (median relative change ≤ 5% from baseline to follow-up). A fifth of the proteins (20.9%) had good reproducibility over time [intraclass correlation coefficient (ICC) 0.75–1.00], 41.7% of the proteins had moderate reproducibility (ICC 0.50–0.75), and 18.8% of the proteins had poor reproducibility (ICC 0–0.25) (**Supplementary Table 3**). Proteins with a low ICC tended to have high median fold changes values (**Supplementary Figure 2**).

Discovery of the general and abdominal adiposity signatures

Using the Population Set as discovery, we evaluated the ethnic-specific associations between plasma protein levels with (i) BMI, and (ii) WC (**Supplementary Table 4**). In the meta-analysis adjusted for age and sex, 410 proteins were associated with BMI, and 385 proteins were associated with WC, with 359 proteins in common between BMI and WC (**Supplementary Table 4 and Figure 2**). Across the three ethnic groups, there were no evidence of statistical heterogeneity ($P_{het} > 1.00\times10^{-5}$) and all associations were in the same direction except one protein (complement factor D) for WC which was excluded from further analyses. All 359 overlapping protein-BMI and protein-WC associations were directionally consistent. Proteins that were associated with only one of the adiposity measures at Bonferroni threshold were directionally consistent and nominally associated for the other adiposity measure ($P < 0.05$; n = 51 for BMI and n = 25 for WC). The top four most strongly associated proteins were the same for BMI and WC, namely leptin (LEP), heart-type fatty acid binding protein (FABP3), insulin-like growth factor-binding protein 1 (IGFBP1), and growth hormone receptor (GHR). Proteins known to be associated with adiposity such as C-reactive protein (CRP), adiponectin (ADIPOQ), and insulin (INS) were also significantly associated with BMI and WC in our study.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We performed a look-up of the 410 BMI-associated proteins in three recent proteomics association studies in European adults, namely the DIOGenes study (n = 494) adults, the INTERVAL study (n = 3,301), and the KORA study (n = 4,600) (3,6,7). We considered our observed protein-BMI association to be replicated if it was directionally consistent and significant at the Bonferroni threshold ($P < 0.05/410$) in the corresponding study. Out of the 410 BMI-associated proteins observed in our study, 101 proteins were not measured in any of the other three studies, and 263 (85%) out of the remaining 309 BMI-associated proteins were replicated in at least one study (**Supplementary Table 4**).

Using the proteins associated with BMI and WC in our study, we performed feature selection using elastic net regression. We identified 124 proteins with non-zero coefficients for BMI (here after referred to as BMI-protein signature) and 125 proteins with non-zero coefficients for WC (WC-protein signature) (**Figure 3**), with 60 overlapping proteins in both signatures. The coefficients for the proteins included in the BMI and WC signatures are reported in **Supplementary Table 5**. We compared our protein signatures with the published results for 11,471 participants of the Fenland Study (11). Of the 124 proteins in our BMI-protein signature, 28 (22.3%) were included in the Fenland Study signature for percentage of body fat. Similarly, of the 125 proteins in our WC-protein signature, 34 (27.2%) were included in the Fenland Study signature for visceral fat mass (**Supplementary Table 5)**.

Internal validation of the general and abdominal adiposity signatures

Validation of the protein signatures identified in the Population Set was done using the T2DCC Set. As expected, the BMI-protein signature was strongly correlated with BMI at baseline and follow-up (r = 0.842 and 0.823) (**Table 1 and Supplementary Figure 3**). Similarly, the WC-protein signature was strongly correlated with participants' WC at baseline and follow-up (r = 0.778 and 0.783). Among 814 participants with both baseline and follow-up data, the changes in BMI (r = 0.632) and changes in WC (r = 0.438) over a ~6-year period were directly correlated with changes in the respective protein signatures over the same period (**Supplementary Figure 4**). We also examined the correlations between changes in anthropometric measurements and changes in individual proteins in the respective signatures (**Supplementary Table 6**). The Pearson correlation coefficients for the correlations between changes in BMI and proteins in the BMI-protein signature ranged from 0.438 (leptin) to -0.469 (adiponectin), and the correlation coefficients for changes in WC and proteins in the WC-protein signature ranged from 0.343 (growth hormone receptor) and -0.326 for (sex hormone-binding globulin). All correlation coefficients reported here were significant ($P < 0.001$).

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We further evaluated the relationship between the anthropometric measurements, protein signatures, and direct measures of general and visceral adiposity among a subset of ethnic Chinese participants [n = 207 for Dual-Energy X-ray Absorptiometry (DXA) and n = 151 for Computed Tomography (CT) measurements] (**Table 1 and Supplementary Table 7**). As expected, BMI and WC were directly correlated with total body fat mass (r = 0.872) and visceral fat area (r = 0.799), respectively. The BMI-protein signature was directly and strongly correlated with total body fat mass (r = 0.782), android fat mass (r = 0.787), and trunk fat mass (r = 0.801), and moderately correlated with subcutaneous adipose tissue (SAT) (r = 0.596) and visceral adipose tissue (VAT) (r = 0.676). In comparison with the BMI-protein signature, the WC-protein signature was less strongly correlated with total body fat mass (r = 0.619) and SAT (r = 0.433) and more strongly correlated with VAT (r = 0.735). Both the BMI-protein and WC-protein signatures were inversely correlated with lean mass percentage (r = -0.488 and r = -0.232 respectively). The scatterplots suggested a direct linear relationship between total body fat mass and the BMI-protein signature (beta = 4.98 kg per SD; $P < 0.001$) and VAT area with the WC-protein signature (beta = 48.1 cm^2 per SD; $P < 0.001$) (**Figure 4**).

Association between protein signatures of adiposity and cardiometabolic risk factors

Using the T2DCC Set, we examined the associations between CVD risk factors and the protein signatures. All adiposity-related predictors (BMI, BMI-protein signature, WC, and WC-protein signature) were significantly associated with higher systolic blood pressure, low-density lipoprotein (LDL) cholesterol, triglycerides, fasting glucose, hemoglobin A1c and lower high-density (HDL) lipoprotein cholesterol, after adjustment for age, sex, and ethnicity. (**Supplementary Table 8**). A comparison of effect sizes revealed significantly stronger associations of the protein signatures with HDL-cholesterol, triglycerides, hemoglobin A1C, compared to the anthropometric measurements.

We also evaluated our protein signatures' ability to differentiate between metabolically healthy and metabolically healthy obesity in the T2DCC set. We defined metabolically unhealthy obesity as having 2 out of 4 metabolic abnormalities: (i) hypertension; (ii) elevated fasting plasma glucose; (iii) elevated triglycerides; (iv) reduced HDL cholesterol. Among 1,262 overweight (BMI ≥ 23 kg/m^2) participants in the T2DCC set, 717 (56.8%) were metabolically unhealthy. After adjusting for age, sex, ethnicity, and BMI, having a higher BMI-protein signature (per SD increment, OR = 2.36, 95% CI 1.93 to 2.88) or WC-protein signature (OR = 2.75, 95% CI 2.28 to 3.32) was associated with a higher odds of being metabolically unhealthy. Results were similar when we used the international BMI cut-off of 25.0 kg/m^2 for overweight definition.

Association between protein signatures of adiposity and T2D incidence

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We evaluated the associations between the incidence of T2D and the protein signatures within the T2DCC Set. All predictor variables were standardized to facilitate the comparison of effect sizes (**Figure 5 and Supplementary Table 9**). Both BMI (OR 1.91, 95% CI 1.70 to 2.14) and the BMI-protein signature (OR 2.44, 95% CI 2.15 to 2.77) were significantly associated with T2D incidence, adjusted for age, sex, and ethnicity. Mediation analysis was used to evaluate if the effects of BMI on T2D incidence were mediated through the proteins in the BMI-protein signature. There was no evidence of a direct effect of BMI ($OR_{BMI\text{-}direct}$ 0.90, 95% CI 0.74 to 1.09), suggesting that the effect of BMI on T2D incidence was mediated through the proteins in the general adiposity signature ($OR_{BMI\text{-}indirect}$ 2.13, 95% CI 1.81 to 2.50). Similarly, both WC (OR 2.09, 95% CI 1.86 to 2.36) and the WC-protein signature (OR 2.84, 95% CI 2.47 to 3.25) were significantly associated with T2D incidence, adjusted for age, sex, and ethnicity. The direct effect of WC on T2D incidence was non-significant ($OR_{WC\text{-}direct}$ 1.07, 95% CI 0.90 to 1.27), whereas the indirect effect of WC on T2D incidence was significant ($OR_{WC\text{-}direct}$ 1.96, 95% CI 1.72 to 2.23), suggesting that the effect of WC on T2D incidence was mediated through proteins in the WC-protein signature.

As the effects of adiposity on insulin resistance may be modified by ethnicity (12,13), we conducted further mediation analyses stratified by ethnicity (**Supplementary Table 9**). We found a stronger association between the BMI-protein signature and T2D among Malays (OR 2.96, 95% CI 2.33 to 3.77) and Chinese (OR 2.37, 95% CI 1.94 to 2.88) compared with Indians (OR 1.91, 95% CI 1.55 to 2.35) (*P*-interaction = 0.019). Similarly, the WC-protein signature was more strongly associated with T2D incidence among Malays (OR 3.59, 95% CI 2.77 to 4.67) and Chinese (OR 2.72, 95% CI 2.20 to 3.37) compared with Indians (OR 2.21, 95% CI 1.76 to 2.77) (*P*-interaction = 0.028).

Finally, we evaluated the discrimination of our protein signatures in predicting T2D incidence using the area under the receiver operating characteristic curve (AUC), which is an indicator of the model's ability to differentiate between individuals that develop T2D versus those who do not (**Supplementary Table 10**). After accounting for age, sex, and ethnicity, the AUC for using the BMI-protein signature to predict T2D (AUC = 0.709, 95% CI 0.685 to 0.734) was significantly higher than using BMI to predict T2D (AUC = 0.665, 95% CI 0.639, 0.691, *P*-value for difference in AUC <0.001). Similarly, the AUC for using the WC-protein signature to predict T2D (AUC = 0.736, 95% CI 0.712 to 0.759) was significantly higher than WC (AUC = 0.682, 95% CI 0.657, 0.708, *P*-value for difference in AUC <0.001).

External validation in the ARIC study

We performed external validation of our results in the ARIC study (n = 8,428), a population-based U.S. cohort study. Participants in the ARIC study had a mean age of 56.7 (± 5.7) years, more likely to

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

be female (57%), and were white (81%) and African Africans (19%) (**Supplementary Table 11**). After Bonferroni adjustment ($P < 0.05/124$ for BMI and 0.05/125 for WC), 118 (95.2%) out of 124 proteins in the BMI-protein signature and 124 (99.2%) out of 125 of the proteins in the WC-protein signature were directionally consistent and significantly associated with BMI and WC, respectively (**Supplementary Table 5 and Supplementary Figure 5**). Concordant with our findings, the BMI-protein signature was highly correlated with BMI (r = 0.822) and the WC-protein signature was highly correlated with WC (r = 0.773). A higher BMI-protein signature (HR per SD increment = 1.80, 95% CI 1.72 to 1.89) and WC-protein signature (HR = 1.97, 95% CI 1.87 to 2.07) at baseline was strongly associated with the incidence of T2D over a median follow-up of 19.4 years.

Pathway enrichment analysis

In our pathway enrichment analysis, we identified 36 significant annotations from the BMI-protein signature, and 28 significant annotations from the WC-protein signature (**Supplementary Table 12**). Here, we highlight enriched pathways associated with the BMI- or WC-protein signature. Pathways related to post-translational protein phosphorylation (REAC:R-HSA-8957275), the regulation of insulin-like growth factor (IGF) transport and uptake by insulin-like growth factor binding proteins (IGFBPs) (REAC:R-HSA-381426), and complement and coagulation cascades (KEGG:04610) were significantly enriched ($P_{FDR} < 0.05$) in both BMI- and WC-protein signatures. In additional, pathways involved in adenosine monophosphate-activated protein kinase (AMPK) signalling (KEGG:04152), extracellular matrix (ECM) receptor interaction (KEGG:04512), and cell adhesion molecules (CAMs) (KEGG:04514) were significantly enriched ($P_{FDR} < 0.05$) in the WC-protein signature.

Discussion

We conducted the largest study in populations of Asian ancestry involving ~5000 plasma proteins. We identified 410 proteins that were associated with BMI and 385 proteins associated with waist circumference of which 359 were shared. These associations were consistent across all three ethnic groups in our study population and many replicated findings in white, African American, and European populations highlighting the robustness of these findings to ethnic and geographical variation. In addition, we identified associations with age and sex with effect sizes highly correlated with findings from populations of European ancestry. Almost 2/3 of the proteins examined showed <5% change within individuals over a period 6 years while about 20% showed poor reproducibility over this time frame which may be helpful to evaluate associations with disease or to monitor changes in physiology over time.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Of the 359 proteins associated with BMI and WC, we replicated several candidate proteins that were recently identified to be obesity-associated proteins in European populations. These proteins include kallistatin (SERPINA4) (3,6,7,14), E-selectin (SELE) (3,6,7,11,15,16), seizure-6-like protein (SEZ6L) (17), reticulon-4 receptor (RTN4R) (3,7,18), and neuronal growth regulator 1 (NEGR1) (3,19,20). Previous studies suggested that SERPINA4, SELE, and SEZ6L may be involved in the regulation of inflammation and the immune system (16,21–23), and RTN4R and NEGR1 may be involved in the regulation of energy metabolism through AMPK activation. We also identified novel proteins that were not previously identified to be related to adiposity, most likely by virtue of the number of proteins represented on this version of the assay. Some of these include serine protease high temperature requirement A1 (HTRA1), syndecan-3 (SDC3), and complexin-2 (CPLX2). HTRA1 was previously shown to regulate the availability of IGFs by cleaving IGFBPs (24) and may modulate IGF-1 signalling (25). In addition, HTRA1 was recently suggested to inhibit adipogenesis by regulating the formation of adipocytes and may potentially be an indicator of adipose tissue dysfunction (26). Polymorphisms in the SDC3 gene have been associated with obesity in Korean (27) and Taiwanese adults (28). SDC3 is involved in the regulation of appetite through the melanocortin system (29) and may therefore be a potential therapeutic target to treating obesity. CPLX2 was recently suggested to play a role in the translocation of glucose transporter 4 (GLUT4) which plays critical roles in glucose uptake regulation (30), although it should be noted that GLUT4 translocation is an intracellular process and it is unclear whether plasma levels of CPLX2 are related to this process. In addition, CPLX2 may also have a role in the regulation of the immune system (31). We believe that at least a proportion of these findings are replicable and biologically relevant and thus form a novel resource for the study of obesity and its effects on human physiology.

Using supervised machine learning techniques, we identified protein signatures of BMI and WC which were validated externally in the ARIC study. These signatures showed consistent associations with BMI and WC at different time points, and the changes in the signatures were correlated with changes in BMI and WC over time. The ability of these protein signatures to capture dynamic changes in adiposity may provide the ability to better evaluate the response to treatment in overweight/obese individuals. In contrast to the associations with BMI and WC, where >85% of associations overlapped between these two anthropometric measures, slightly less than half of the proteins included in these machine learning derived signatures overlapped. We further showed that the BMI-protein signature was highly correlated with total body fat mass whereas the WC-protein signature was more strongly correlated with VAT area and less strongly correlated with total body fat mass, suggesting that the two signatures were able to capture the differences in plasma proteins

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

associated with adiposity in different anatomical compartments. To our knowledge, this has not been demonstrated previously.

Compared to the anthropometric measures, both the BMI- and WC-protein signatures were more strongly associated with HDL-cholesterol, triglycerides, and hemogloblin A1c levels. In prospective analyses, both protein signatures were significantly associated with the incidence of T2D and fully explained the associations between anthropometric measurements and diabetes risk. In addition, our protein signatures were able to better differentiate those at a higher risk of T2D compared to the anthropometric measurements. These findings suggest that the signatures are likely to capture physiologically relevant changes that are not captured by traditional measures of adiposity.

Our observed differences in plasma protein signatures between metabolically healthy and metabolically unhealthy obesity may inform potential therapeutic targets and biological pathways behind the development of metabolically unhealthy obesity (32). Pathways related to complement and coagulation cascades, IGF and IGFBPs regulation, and post-translational protein phosphorylation were significantly enriched in both the BMI-and WC-protein signatures. The complement system is part of the human innate immune system, and dysregulation of the complement system from an overexpression of cytokines and adipose tissue-derived factors may lead to chronic inflammation and the development of metabolic disorders such as insulin resistance and T2D (33). Similarly, imbalances in levels of IGFs and IGFBPs has been linked to the dysregulation of fat metabolism and insulin resistance (34–36), and the regulation of the bioactivity of IGFBP-1 through post-translational phosphorylation may play an important role in mediating the effects of IGFBP-1 on glucose metabolism (37). In addition, pathways involving AMPK signalling, ECM-receptor interaction, and CAMs were enriched in the WC-protein signature. AMPK is the key enzyme involved in the regulation of nutrient metabolism (38), and the dysregulation of AMPK has been linked to insulin resistance and T2D (39). The over-activation of ECM receptor signalling pathways may lead to an increased deposition of ECM that may further trigger inflammatory pathways downstream (40). Elevation of CAMs have been previously associated with obesity and T2D (41,42), and is suggested to play an important role in inflammation and atherogenesis (43). Taken together, the pathways enriched in the adiposity-related protein signatures depict a state of dysregulated cell signalling, systemic inflammation, and impaired glucose and fat metabolism. In addition, several of the pathways (e.g., complement pathways, ECM-receptor interaction, and IGFBP regulation) enriched in our adiposity-related protein signatures were also enriched in a set of T2D-associated proteins identified by a recent study on diabetes and plasma proteins among European adults (44), implying that pathways associated with obesity may have a central role in the pathogenesis of diabetes.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Previous studies have reported ethnic differences in the association between obesity and insulin resistance (12,45). Khoo et al. (12) demonstrated that the association of body fat percentage with insulin resistance was weaker among ethnic Indians than Chinese and Malays residing in Singapore. Furthermore, Retnakaran et al. (45) found that pre-pregnancy BMI was more strongly associated with insulin resistance in East Asian women than South Asian women in Canada. In our study, while the associations between plasma proteins and anthropometric measurements were consistent across different populations, we observed stronger associations between the BMI- and WC-protein signatures and T2D risk among Chinese and Malays compared with Indians. This suggested that the observed ethnic differences in the associations between obesity and T2D may be explained by potential differential effects of obesity-related proteins on T2D incidence.

The strengths of our study include the use of BMI and WC to capture both general and abdominal adiposity signatures, the rigorous validation of our protein signatures which included (i) replication of associations between proteins and anthropometric measures in an independent sample cross-sectionally and longitudinally, (ii) validation against more precise measures of adiposity with DXA and CT scans, and (iii) external replication of associations of proteins with anthropometric measures and T2D incidence in a non-Asian population. Potential limitations of our study include the use of BMI and WC as the proxy for general adiposity and abdominal adiposity in our training dataset. Among participants with DXA and CT scans, BMI was highly correlated with total body fat mass ($r = 0.87$) and WC was highly correlated ($r = 0.80$) with VAT area, suggesting that BMI and WC were indicative of general and abdominal adiposity in our sample. Second, the plasma proteome includes both secreted proteins and leakage proteins, and the concentration of leakage proteins in the plasma may not be a direct reflection of its biological activity in the cells. Nonetheless, the plasma proteome including both secreted and leakage proteins have been shown to be informative about the current and future state of health (11,46). Third, we were not able to infer causality as the associations between plasma proteins and anthropometric measurements were cross-sectional. Even though our participants did not have T2D at baseline, it is possible that early manifestations of diabetes may have influenced plasma protein levels. Last, while studies have demonstrated the specificity of the SomaScan assay (3,46), it remains possible that some proteins may not be correctly identified by the SomaScan platform used in our study.

In conclusion, our findings on adiposity-related proteins are robust across ethnic and geographically diverse groups. Our protein signatures have potential uses for the assessment of metabolically adverse adiposity and to monitor changes therein during interventions. While further studies are needed to evaluate the causal relationship between the identified proteins and adiposity, the large extent to which the proteins appear to mediate associations between body fatness and the risk of

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

type 2 diabetes can lead to a better understanding of the mechanisms relating obesity and diabetes and why sub-groups of the population appear to be more affected by the effects of obesity on metabolic health than others.

Online Methods

Study design

An overview of this study is shown in **Figure 1**. Participants in this study were sampled from the Singapore Multi-Ethnic Cohort Phase 1 (MEC1), a prospective cohort study initiated to investigate the impact of lifestyle factors and molecular biomarkers on health and chronic diseases. Details about the MEC1 have been previously documented (9). Briefly, the MEC1 is a population-based cohort comprising of 14,465 male and female adults recruited between 2004 to 2010 when they were 21 years and older. Oversampling of ethnic minorities (Malay and Indians) was done to achieve a good representation of three major Asian ethnic groups in Singapore—Chinese, Malay, and Indian. Between 2011 to 2016 (mean follow-up of 6.4 years), 6,112 participants agreed to a follow-up visit. For both the baseline and follow-up study, participants completed a standardized interviewer-administered questionnaire on socio-demographic characteristics, lifestyle, and medical history. Participants were also invited to a health screening facility for a physical examination conducted by trained research staff. Height was measured without shoes on a portable stadiometer while weight was measured on a digital scale. WC was measured by trained research staff using a stretch-resistant tape at the mid-point between the last rib and iliac crest. Blood was drawn for biomarker measurements and biobanking. For the latter, the blood samples were transported at 4–8°C to biobank laboratories where they were processed and stored frozen at -80°C on the same day.

At follow-up, a subset of Chinese participants also underwent Dual-Energy X-ray Absorptiometry (DXA) and Computed Tomography (CT) scan for a more comprehensive assessment of body fat composition and intra-abdominal fat distribution, as previously described (47). For the DXA measurements, body composition measurements including total body fat mass, trunk fat mass, android fat mass, and lean body mass were assessed using a medium speed total body acquisition mode (Discovery Wi, Hologic, Bedford, MA, and software Hologic Apex 3.01). For the CT scan, images at the inter-vertebral space L2/L3 and L4/L5 levels were identified by a single observer (sliceOmatic version 5.0, Tomovision, Magog, Canada). The images were subsequently analyzed by two independent readers using a workstation (eFilm Workstation version 4.0, Hartland, USA) and the readings for visceral adipose tissue (VAT) area and subcutaneous adipose tissue (SAT) area were averaged. The average inter-observer coefficient of variation was 2.34% for SAT and 3.22% for VAT at L2/L3 level. For the present analysis, we used SAT and VAT area at the L2/L3 inter-vertebral space

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

as VAT area at L2/L3 was found to be most strongly correlated with VAT volume and cardiovascular risk factors among Chinese adults (48).

We selected two groups of participants to be profiled on the SomaScan proteomic assay. The first group was a random sample of 720 participants (hereafter referred as the Population Set) with plasma samples available at both baseline and follow-up and equal numbers in each sex and ethnic group (n = 120 each). For the second group, we conducted a nested case-control study including 759 incident type 2 diabetes (T2D) cases and 1,484 matched controls, hereafter referred as the T2D case-control (T2DCC) set. Incident diabetes cases were ascertained through record linkage with a national health care database, self-reports of a physician-diagnosis of diabetes at follow-up, or having a fasting blood glucose ≥ 7mmol/L or random blood glucose ≥ 11mmol/L or glycated hemoglobin A1c ≥ 6.5% at follow-up (49). Controls were selected using risk-set sampling and were matched 1:2 according to age (± 5 years), sex, ethnicity, and date of blood collection (± 2 years). Of 4,897 samples from 2,981 participants sent to SomaLogic, 609 samples were not measured due to shipment issues. We excluded samples that did not pass SomaLogic quality control (n = 82), duplicate samples at baseline (n = 4), participants from ethnic groups other than Chinese, Malay, or Indian (n = 2), participants with physician diagnosed heart diseases, stroke, or T2D at baseline (n = 26), and those with missing anthropometric data (n = 10). The final number of participants for analyses was 631 from the Population Set, and 616 cases and 1,200 controls from the T2DCC Set. (**Supplementary Table 13**). Among these participants, 598 of the Population set, 327 cases, and 487 controls had both baseline and follow-up data. Written consent was obtained from all participants and this study was approved by the National University of Singapore Institutional Review Board (reference code: N-18-059).

We externally replicated our findings in the Atherosclerosis Risk in Communities (ARIC) Study, a population-based cohort study in the USA. Participants were male and female adults and white and African American. Details on the ARIC study have been reported elsewhere (50). After exclusion of participants with prevalent diabetes (defined based on self-report diagnosis, medication use, fasting glucose ≥ 7mmol/L, or HbA1c ≥ 6.5%) or heart diseases at baseline, we included a total of 8,428 ARIC participants with proteomic data measured on the SomaScan V4 platform during Visit 2 (1990-1992) in our analyses. Over a median follow-up duration of 19.4 years, 1,800 diabetes incident cases were ascertained based on self-reported diagnosis by a healthcare provider or glucose-lowering medication use reported by participants during annual follow-up telephone calls or clinic visits.

Proteomic assay

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Relative protein abundances were measured in plasma samples using an aptamer-based technology (SomaScan V4 assay) by SomaLogic, Inc. (Boulder, Colorado, US). Details of the SomaScan assay (51,52), and the reproducibility and specificity of the SomaScan assay have been reported previously (3,11,46,53,54).

For this study, relative fluorescence units (RFU) for 5,284 SOMAmers were obtained (**Supplementary Table 2)**. After the exclusion of 298 SOMAmers that targeted non-human proteins, 7 deprecated SOMAmers and 1 SOMAmer that targeted a protein whose annotation was withdrawn by the National Center for Biotechnology Information (55), 4,978 SOMAmers targeting 4,775 human proteins remained. A total of 82 samples (2%) were excluded from the subsequent analyses as the scale factors exceeded the SomaLogic's recommended range of 0.4–2.5.

As an additional quality check, we compared SomaLogic's measurement of four proteins (C-reactive protein, insulin, adiponectin, and interleukin-1 receptor antagonist) against the values measured by our laboratories using traditional assays. The Spearman rank correlation coefficients were: 0.97–0.98 for C-reactive protein, 0.47–0.58 for insulin, 0.88–0.90 for adiponectin and 0.60 for interleukin-1 receptor antagonist protein. More details of the assays are reported in **Supplementary Table 14**.

Statistical analysis

All data analyses were performed using R version 4.2.0 and were based on $\log_2$ transformed RFU. For all analyses except for the intraclass correlation (ICC) and median fold change over time, protein levels in the log-scale were winsorized at ±5 standard deviations to reduce the impact of extreme outliers. All *P*-values reported are from two-sided tests.

Protein stability over time and association with demographic characteristics

To evaluate the reliability of protein levels over a ~6-year period among participants who did not develop T2D during the period, we computed the median fold change over time (ratio between protein levels at baseline and follow-up) and ICC among participants from the Population Set and controls from the T2DCC Set with both baseline and follow-up data (n = 1,085). We used a two-way mixed-effects model for absolute agreement ICC (56). ICC values above 0.75 suggest that the intra-individual variability in protein level is low, while values between 0.50–0.75 suggest moderate variability over time, and values below 0.50 suggest high variability over time (56).

We also evaluated the associations between protein levels and age, sex, and ethnicity using multivariable linear regression models that mutually adjusted for the mentioned demographic predictors. The overall significance of ethnicity was determined using a likelihood ratio rest. A

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Bonferroni corrected threshold (0.05/4978 = 1.00×10^{-5}) was used to determine statistical significance.

Discovery and internal validation of protein signatures of general (BMI) and abdominal (WC) adiposity

We used BMI as a proxy of general adiposity and WC as a proxy of abdominal adiposity. Using data from the Population Set at baseline as the discovery dataset, we evaluated ethnic-specific associations between protein levels and adiposity (BMI or WC) adjusted for age and sex. The ethnic-specific results were pooled using fixed-effects inverse-variance meta-analysis.

To identify a set of proteins for the general and abdominal adiposity signatures, we first selected proteins that were consistently associated across the ethnic groups using the following criteria: (i) significantly associated with the trait of interest (BMI or WC) at the Bonferroni corrected threshold ($P < 1.00\times10^{-05}$), (ii) had the same direction of association across three ethnic groups, and (iii) had no evidence of statistical heterogeneity using the Cochran Q test ($P_{\text{het}} > 1.00\times10^{-05}$). For proteins targeted by multiple SOMAmers, the SOMAmer with the smallest *P*-value was selected. Using the associated proteins, we performed 100 replicates of 10-fold cross validation elastic net regression analysis (57). The elastic net regression is a regularized regression method that enables automatic variable selection and variance reduction (57). The output from the elastic net regression analysis was a selected list of proteins with non-zero beta coefficients. The BMI-protein and WC-protein signatures were calculated as the weighted sum of the concentration of the proteins, with the weights being the beta coefficients from the elastic net regression model.

Using the T2DCC Set as the validation dataset, we first examined the cross-sectional correlations between the protein signatures and BMI and WC at baseline and follow-up, and the correlations between changes in the protein signatures and changes in BMI and WC over a ~6-year period. On a subset of Chinese participants with body composition measurements (total body fat mass, trunk fat mass, android fat mass, and percentage of lean mass) (n = 207) and VAT and SAT measurements (n = 151), we also computed the Pearson correlation coefficients.

Association between adiposity signatures and CVD risk factors

We evaluated the associations between the adiposity signatures and CVD risk factors (systolic blood pressure, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, fasting glucose, and glycated hemoglobin A1c) among participants in the T2DCC Set. We used linear regression models adjusted for age, sex, and ethnicity and standardized all predictor variables to facilitate the comparison of effect sizes with BMI and WC. The difference in standardized beta

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

coefficients was evaluated using the Z-test. Participants on anti-hypertensive medication were excluded from analyses involving systolic blood pressure, and participants on lipid-lowering medications were excluded from analyses involving blood lipids.

We also evaluated our protein signatures' ability to differentiate between metabolically healthy and metabolically healthy obesity (58). For this analysis, we focused on overweight or obese individuals (BMI ≥ 23.0 kg/m^2) in the T2DCC set. We defined metabolically unhealthy as having 2 or more of the 4 metabolic abnormalities: (i) hypertension defined as having a systolic blood pressure ≥ 130 mmHg or diastolic blood pressure ≥ 85 mmHg; (ii) elevated fasting plasma glucose defined as ≥ 5.6 mmol/L, (iii) elevated triglycerides defined as ≥ 1.7 mmol/L, and (iii) reduced high density lipoprotein cholesterol defined as < 1.0 mmol/L for males and < 1.3 mmol/L for females. We used logistic regression models with metabolic health status as outcome and standardized protein signatures as predictors, adjusted for age, sex, ethnicity, and BMI.

Association between adiposity signatures and T2D incidence

We applied the BMI and WC protein signatures on the T2DCC Set to examine the association with T2D incidence using logistic regression models adjusted for age, sex, and ethnicity. Variables were expressed as per 1 standard deviation (SD) to facilitate comparison with BMI and WC. Mediation analysis using natural effects models (Medflex R package version 0.6-7) (59) was used to evaluate if the association between the BMI and WC and T2D incidence was mediated by the respective protein signatures. We assessed the discrimination of the models using the AUC calculated using the pROC R package version 1.18.2.

External validation of protein signatures in the ARIC study

Data from the ARIC study was used for the external validation of our findings. First, we evaluated the associations between anthropometric measurements and each protein in the BMI- and WC-protein signatures, adjusted for age, sex, race, and study center. We consider our adiposity-protein associations replicated if it was statistically significant at the Bonferroni threshold ($P < 0.05/124$ for BMI and $P < 0.05/125$ for WC) and directionally consistent with the ARIC study. Next, we computed the BMI- and WC-protein signature for each ARIC participant using the weights identified in the Singapore cohort and assessed their correlations with BMI and WC. Finally, we evaluated the associations between T2D incidence and both protein signatures using Cox proportional hazards model, adjusted for age, sex, race, and study center.

Functional annotation and pathway enrichment analysis

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We performed functional annotation of the proteins using the UniProt Knowledgebase (60). We also queried the Gene Ontology (GO) (61), Kyoto Encyclopedia of Genes and Genomes (KEGG) (62), and Reactome (63) databases using the gprofiler2 R package version 0.2.1 (64) to identify functional annotations and pathways that are significantly overexpressed in the sets of proteins in the BMI- and WC-protein signatures using the hypergeometric test. We used all proteins measured by the SomaScan V4 as the custom background to avoid false positives due to the functional properties of the assay (65). Annotations were considered to be significant at a false discovery rate (FDR) corrected *P*-value of 0.05 (66).

Acknowledgements

We thank all participants, study team and investigators for their contributions to research. The MEC study is supported by individual research and clinical scientist award schemes from the Singapore National Medical Research Council (NMRC, including MOH-000271-00) and the Singapore Biomedical Research Council (BMRC), the Singapore Ministry of Health (MOH), the National University of Singapore (NUS) and the Singapore National University Health System (NUHS). The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract nos. (75N92022D00001, 75N92022D00002, 75N92022D00003, 75N92022D00004, 75N92022D00005). The funders had no role in the design, implementation, analysis, and interpretation of the data.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

References

1. Hruby A, Hu FB. The Epidemiology of Obesity: A Big Picture. Pharmacoeconomics. 2015 Dec 4;33(7):673–89.

2. Riaz H, Khan MS, Siddiqi TJ, Usman MS, Shah N, Goyal A, et al. Association Between Obesity and Cardiovascular Outcomes: A Systematic Review and Meta-analysis of Mendelian Randomization Studies. JAMA Netw Open. 2018 Nov 2;1(7):e183788–e183788.

3. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018 Jun 6;558(7708):73–9.

4. Hanash S. Disease proteomics. Nature. 2003 Mar 13;422(6928):226–32.

5. Haslam DE, Li J, Dillon ST, Gu X, Cao Y, Zeleznik OA, et al. Stability and reproducibility of proteomic profiles in epidemiological studies: comparing the Olink and SOMAscan platforms. Proteomics. 2022 Jul 1;22(13–14):2100170.

6. Carayol J, Chabert C, Di Cara A, Armenise C, Lefebvre G, Langin D, et al. Protein quantitative trait locus study in obesity during weight-loss identifies a leptin regulator. Nat Commun. 2017 Dec 1;8(1).

7. Zaghlool SB, Sharma S, Molnar M, Matías-García PR, Elhadad MA, Waldenberger M, et al. Revealing the role of the human blood plasma proteome in obesity using genetic drivers. Nat Commun. 2021 Feb 24;12(1):1–13.

8. Ma RCW, Chan JCN. Type 2 diabetes in East Asians: similarities and differences with populations in Europe and the United States. Ann N Y Acad Sci. 2013;1281(1):64–91.

9. Tan KHX, Tan LWL, Sim X, Tai ES, Lee JJM, Chia KS, et al. Cohort Profile: The Singapore Multi-Ethnic Cohort (MEC) study. Int J Epidemiol. 2018;47(3):699-699j.

10. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and disease. Nat Genet. 2021 Dec 2;53(12):1712–21.

11. Williams SA, Kivimaki M, Langenberg C, Hingorani AD, Casas JP, Bouchard C, et al. Plasma protein patterns as comprehensive indicators of health. Nat Med. 2019 Dec 2;25(12):1851–7.

12. Khoo CM, Khee-Shing Leow M, Sadananthan SA, Lim R, Venkataraman K, Khoo EYH, et al. Body Fat Partitioning Does Not Explain the Interethnic Variation in Insulin Sensitivity Among Asian Ethnicity: The Singapore Adults Metabolism Study. Diabetes. 2014 Mar 1;63(3):1093–

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

102.

13. Khoo CM, Sairazi S, Taslim S, Gardner D, Wu Y, Lee J, et al. Ethnicity Modifies the Relationships of Insulin Resistance, Inflammation, and Adiponectin With Obesity in a Multiethnic Asian Population. Diabetes Care. 2011 May 1;34(5):1120–6.

14. Zhu H, Chao J, Kotak I, Guo D, Parikh SJ, Bhagatwala J, et al. Plasma kallistatin is associated with adiposity and cardiometabolic risk in apparently healthy African American adolescents. Metabolism. 2013 May;62(5):642.

15. Lee CH, Kuo FC, Tang WH, Lu CH, Su SC, Liu JS, et al. Serum E-selectin concentration is associated with risk of metabolic syndrome in females. PLoS One. 2019 Sep 1;14(9).

16. Zanni M V., Stanley TL, Makimura H, Chen CY, Grinspoon SK. Effects of TNF-alpha antagonism on E-selectin in obese subjects with metabolic dysregulation. Clin Endocrinol (Oxf). 2010;73(1):48.

17. Harlid S, Myte R, Van Guelpen B. The Metabolic Syndrome, Inflammation, and Colorectal Cancer Risk: An Evaluation of Large Panels of Plasma Protein Markers Using Repeated, Prediagnostic Samples. Mediators Inflamm. 2017;2017.

18. Wang X, Yang Y, Zhao D, Zhang S, Chen Y, Chen Y, et al. Inhibition of high-fat diet-induced obesity via reduction of ER-resident protein Nogo occurs through multiple mechanisms. J Biol Chem. 2022 Feb 1;298(2):101561.

19. Willer CJ, Speliotes EK, Loos RJF, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009 Jan;41(1):25–34.

20. Bauer F, Elbers CC, Adan RAH, Loos RJF, Onland-Moret NC, Grobbee DE, et al. Obesity genes identified in genome-wide association studies are associated with adiposity measures and potentially with nutrient-specific food preference. Am J Clin Nutr. 2009 Oct 1;90(4):951–9.

21. Frühbeck G, Gómez-Ambrosi J, Rodríguez A, Ramírez B, Valentí V, Moncada R, et al. Novel protective role of kallistatin in obesity by limiting adipose tissue low grade inflammation and oxidative stress. Metabolism. 2018 Oct 1;87:123–35.

22. Qiu WQ, Luo S, Ma SA, Saminathan P, Li H, Gunnersen JM, et al. The Sez6 Family Inhibits Complement by Facilitating Factor I Cleavage of C3b and Accelerating the Decay of C3 Convertases. Front Immunol. 2021 Apr 15;12:1171.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

23. Shah T, Zabaneh D, Gaunt T, Swerdlow DI, Shah S, Talmud PJ, et al. Gene-centric analysis identifies variants associated with interleukin-6 levels and shared pathways with other inflammation markers. Circ Cardiovasc Genet. 2013 Apr;6(2):163–70.

24. Hou J, Clemmons DR, Smeekens S. Expression and characterization of a serine protease that preferentially cleaves insulin-like growth factor binding protein-5. J Cell Biochem. 2005 Feb 15;94(3):470–84.

25. Jacobo SMP, DeAngelis MM, Kim IK, Kazlauskas A. Age-Related Macular Degeneration-Associated Silent Polymorphisms in HtrA1 Impair Its Ability To Antagonize Insulin-Like Growth Factor 1. Mol Cell Biol. 2013 May 1;33(10):1976.

26. Tiaden AN, Bahrenberg G, Mirsaidi A, Glanz S, Blüher M, Richards PJ. Novel Function of Serine Protease HTRA1 in Inhibiting Adipogenic Differentiation of Human Mesenchymal Stem Cells via MAP Kinase-Mediated MMP Upregulation. Stem Cells. 2016 Jun 1;34(6):1601–14.

27. Ha E, Kim MJ, Choi BK, Rho JJ, Oh DJ, Rho TH, et al. Positive Association of Obesity with Single Nucleotide Polymorphisms of Syndecan 3 in the Korean Population. J Clin Endocrinol Metab. 2006 Dec 1;91(12):5095–9.

28. Chang BCC, Hwang LC, Huang WH. Positive Association of Metabolic Syndrome with a Single Nucleotide Polymorphism of Syndecan-3 (rs2282440) in the Taiwanese Population. Int J Endocrinol. 2018;2018.

29. Reizes O, Benoit SC, Strader AD, Clegg DJ, Akunuru S, Seeley RJ. Syndecan-3 Modulates Food Intake by Interacting with the Melanocortin/AgRP Pathway. Ann N Y Acad Sci. 2003 Jun 1;994(1):66–73.

30. Pavarotti MA, Tokarz V, Frendo-Cumbo S, Bilan PJ, Liu Z, Zanni-Ruiz E, et al. Complexin-2 redistributes to the membrane of muscle cells in response to insulin and contributes to GLUT4 translocation. Biochem J. 2021 Jan 29;478(2):407–22.

31. Tsuru E, Oryu K, Sawada K, Nishihara M, Tsuda M. Complexin 2 regulates secretion of immunoglobulin in antibody-secreting cells. Immunity, Inflamm Dis. 2019 Dec 1;7(4):318–25.

32. Iacobini C, Pugliese G, Blasetti Fantauzzi C, Federici M, Menini S. Metabolically healthy versus metabolically unhealthy obesity. Metabolism. 2019 Mar 1;92:51–60.

33. Shim K, Begum R, Yang C, Wang H. Complement activation in obesity, insulin resistance, and type 2 diabetes mellitus. World J Diabetes. 2020 Jan 1;11(1):1.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

34. Friedrich N, Thuesen B, Jrøgensen T, Juul A, Spielhagen C, Wallaschofksi H, et al. The Association Between IGF-I and Insulin Resistance: A general population study in Danish adults. Diabetes Care. 2012 Apr;35(4):768.

35. Haywood NJ, Slater TA, Matthews CJ, Wheatcroft SB. The insulin like growth factor and binding protein family: Novel therapeutic targets in obesity & diabetes. Mol Metab. 2019 Jan 1;19:86.

36. Clemmons DR. Role of IGF-binding proteins in regulating IGF responses to changes in metabolism. J Mol Endocrinol. 2018 Jul 1;61(1):T139–69.

37. Heald AH, Siddals KW, Fraser W, Taylor W, Kaushal K, Morris J, et al. Low Circulating Levels of Insulin-Like Growth Factor Binding Protein-1 (IGFBP-1) Are Closely Associated With the Presence of Macrovascular Disease and Hypertension in Type 2 Diabetes. Diabetes. 2002 Aug 1;51(8):2629–36.

38. Herzig S, Shaw RJ. AMPK: guardian of metabolism and mitochondrial homeostasis. Nat Rev Mol Cell Biol 2017 192. 2017 Oct 4;19(2):121–35.

39. Coughlan KA, Valentine RJ, Ruderman NB, Saha AK. AMPK activation: a therapeutic target for type 2 diabetes? Diabetes, Metab Syndr Obes Targets Ther. 2014 Jun 24;7:241.

40. Lin D, Chun TH, Kang L. Adipose extracellular matrix remodelling in obesity and insulin resistance. Biochem Pharmacol. 2016 Nov 1;119:8–16.

41. Hegazy GA, Awan Z, Hashem E, Al-Ama N, Abunaji AB. Levels of soluble cell adhesion molecules in type 2 diabetes mellitus patients with macrovascular complications. J Int Med Res. 2020;48(4):1–11.

42. Miller MA, Cappuccio FP. Cellular adhesion molecules and their relationship with measures of obesity and metabolic syndrome in a multiethnic population. Int J Obes 2006 308. 2006 Mar 7;30(8):1176–82.

43. Hope SA, Meredith IT. Cellular adhesion molecules and cardiovascular disease. Part I. Their expression and role in atherogenesis. Intern Med J. 2003 Aug 1;33(8):380–6.

44. Gudmundsdottir V, Zaghlool SB, Emilsson V, Aspelund T, Ilkov M, Gudmundsson EF, et al. Circulating Protein Signatures and Causal Candidates for Type 2 Diabetes. Diabetes. 2020 Aug 1;69(8):1843.

45. Retnakaran R, Hanley AJG, Connelly PW, Sermer M, Zinman B. Ethnicity Modifies the Effect of

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Obesity on Insulin Resistance in Pregnancy: A Comparison of Asian, South Asian, and Caucasian Women. J Clin Endocrinol Metab. 2006 Jan 1;91(1):93–7.

46. Emilsson V, Ilkov M, Lamb JR, Finkel N, Gudmundsson EF, Pitts R, et al. Co-regulatory networks of human serum proteins link genetics to disease. Science. 2018 Aug 8;361(6404):769.

47. Khaing NEE, Shyong TE, Lee J, Soekojo CY, Ng A, Van Dam RM. Epicardial and visceral adipose tissue in relation to subclinical atherosclerosis in a Chinese population. PLoS One. 2018 Apr 1;13(4):e0196328.

48. Ng AC, Wai DC, Tai ES, Ng KM, Chan LL. Visceral adipose tissue, but not waist circumference is a better measure of metabolic risk in Singaporean Chinese and Indian men. Nutr Diabetes. 2012 Aug 6;2(8):e38–e38.

49. American Diabetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2021. Diabetes Care. 2021 Jan 1;44(Supplement_1):S15–33.

50. Steffen BT, Tang W, Lutsey PL, Demmer RT, Selvin E, Matsushita K, et al. Proteomic analysis of diabetes genetic risk scores identifies complement C2 and neuropilin-2 as predictors of type 2 diabetes: the Atherosclerosis Risk in Communities (ARIC) Study. Diabetologia. 2022 Jan 1;66(1):105–15.

51. Rohloff JC, Gelinas AD, Jarvis TC, Ochsner UA, Schneider DJ, Gold L, et al. Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. Mol Ther Nucleic Acids. 2014;3(10):e201.

52. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, et al. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. PLoS One. 2010;5(12):e15004.

53. Kim CH, Tworoger SS, Stampfer MJ, Dillon ST, Gu X, Sawyer SJ, et al. Stability and reproducibility of proteomic profiles measured with an aptamer-based platform. Sci Reports 2018 81. 2018 May 30;8(1):1–10.

54. Tin A, Yu B, Ma J, Masushita K, Daya N, Hoogeveen RC, et al. Reproducibility and Variability of Protein Analytes Measured Using a Multiplexed Modified Aptamer Assay. J Appl Lab Med. 2019 Jul 1;4(1):30.

55. NCBI. LOC652493 ig kappa chain V-I region HK102-like [Homo sapiens (human)] - Gene - NCBI [Internet]. 2011.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

56. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016 Jun 1;15(2):155.

57. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Ser B (Statistical Methodol. 2005 Apr 1;67(2):301–20.

58. Blüher M. Metabolically Healthy Obesity. Endocr Rev. 2020 Jun 1;41(3):405–20.

59. Steen J, Loeys T, Moerkerke B, Vansteelandt S. medflex: An R Package for Flexible Mediation Analysis using Natural Effect Models. J Stat Softw. 2017 Feb 27;76(1):1–46.

60. Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023 Jan 6;51(D1):D523–31.

61. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000 May;25(1):25.

62. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000 Jan 1;28(1):27–30.

63. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2018 Jan 1;46(Database issue):D649.

64. Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H. gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. F1000Research. 2020 Jul 15;9:709.

65. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc. 2019 Feb 1;14(2):482.

66. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 1995 Jan 1;57(1):289–300.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Tables and Figures

Table 1. Pearson correlation coefficients between body mass index (BMI) and waist circumference (WC) protein signatures and anthropometric, DXA, and CT adiposity measurements in the T2DCC Set.

	BMI signature			WC signature		
	Cases	Controls	Combined	Cases	Controls	Combined
Anthropometric measurements at baseline[1]						
BMI, kg/m^2	0.815	0.847	0.842	0.675	0.731	0.726
WC, cm	0.639	0.717	0.716	0.675	0.798	0.778
Anthropometric measurements at follow-up[2]						
BMI, kg/m^2	0.809	0.829	0.823	0.696	0.712	0.713
WC, cm	0.712	0.742	0.752	0.726	0.791	0.783
DXA measurements[3]						
Total body fat mass, kg	0.756	0.789	0.782	0.628	0.583	0.619
Lean mass, %	-0.482	-0.484	-0.488	-0.270	-0.171	-0.232
Android fat mass, kg	0.755	0.787	0.787	0.706	0.689	0.717
Trunk fat mass, kg	0.764	0.812	0.801	0.685	0.665	0.691
CT measurements[4]						
SAT volume, cm^2	0.499	0.641	0.596	0.418	0.378	0.433
VAT volume, cm^2	0.542	0.719	0.676	0.603	0.776	0.735

All correlation coefficients were statistically significant at P <0.001. Abbreviations: BMI, body mass index; Computed Tomography, CT; DXA, Dual-Energy X-ray Absorptiometry; SAT, subcutaneous adipose tissue; T2DCC, type 2 diabetes case control; VAT, visceral adipose tissue; WC, waist circumference.

[1] Measured among all participants in the T2DCC Set at baseline; n = 616 for cases, 1,200 for controls, and 1,816 when combined.

[2] Measured among all participants in the T2DCC Set at follow-up; n = 327 for cases, 487 for controls, and 814 when combined.

[3] Measured in a subset of Chinese participants at follow-up; n = 69 for cases, 138 for controls, and 207 when combined.

[4] Measured in a subset of Chinese participants at follow-up; n = 51 for cases, 100 for controls, and 151 when combined.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

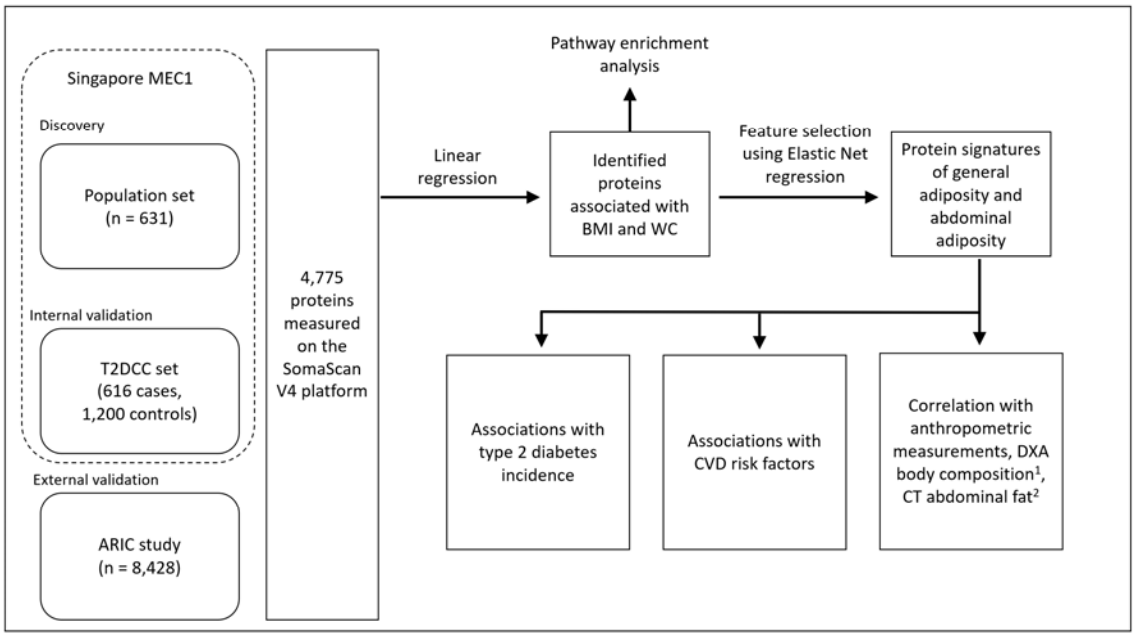

Figure 1. Overview of the study design. [1] Measured in a subset of Chinese participants in the MEC1 (n = 207) at follow-up.[2] Measured in a subset of Chinese participants in the MEC1 (n = 151) at follow-up. Abbreviations: ARIC, Atherosclerosis Risk in Communities; BMI, body mass index; CT, computed tomography; CVD, cardiovascular disease; DXA, dual-energy X-ray absorptiometry; MEC1, Multi-Ethnic Cohort Phase 1; T2D, type 2 diabetes; T2DCC, T2D case-control; WC, waist circumference.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

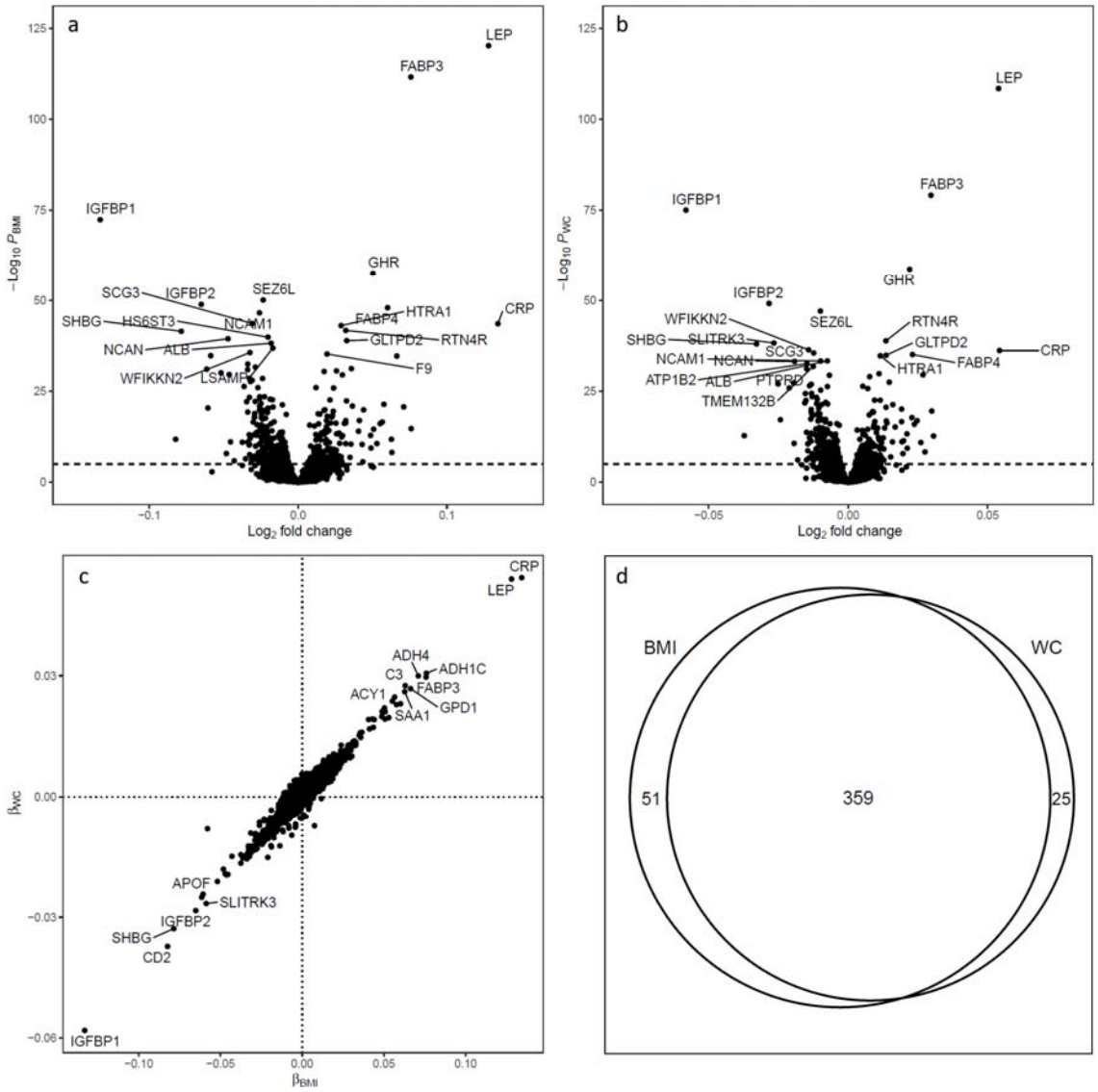

Figure 2. (a) Volcano-plot of -$\log_{10}(P_{BMI})$ against $\log_2$(fold change) per kg/m^2 increment of BMI. (b) Volcano-plot of -$\log_{10}(P_{WC})$ against $\log_2$(fold change) per cm increment of WC. Dashed lines represent the threshold for statistical significance after Bonferroni correction. (c) Scatterplot of the beta coefficients for WC against the beta coefficients for BMI. Text annotations are Entrez Gene symbols and the full protein names are provided in Supplementary Table 4. (d) Venn diagram of the number of proteins that were significantly associated with BMI and WC at the Bonferroni threshold. Abbreviations: BMI, body mass index; WC, waist circumference.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

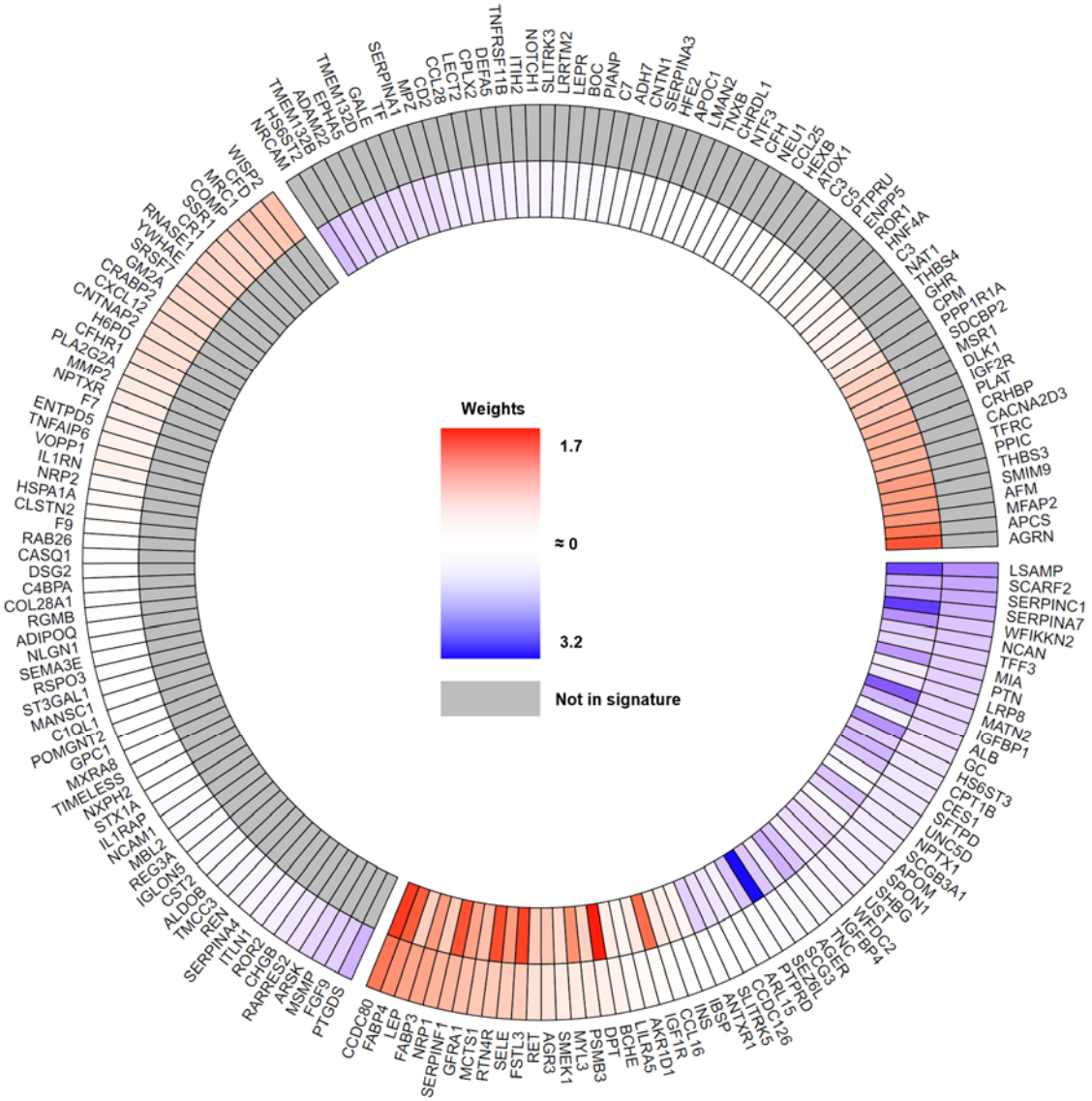

Figure 3. Proteins in the BMI-protein signature (outer circle) and WC-protein signature (inner circle). The colours represent the direction (red = increase, blue = decrease, grey = not in signature) while the intensity of the colours represents the magnitudes of the beta coefficients from elastic net regression. The bottom right segment represents the shared proteins between the BMI-protein and WC-protein signature. Moving anticlockwise, the top right segment represents proteins that are only in the WC-protein signature. Finally, the last segment on the top left represents proteins that are only in the BMI-protein signature. Text annotations are Entrez Gene symbols and the full protein names are provided in Supplementary Table 5. Abbreviations: BMI, body mass index; WC, waist circumference.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

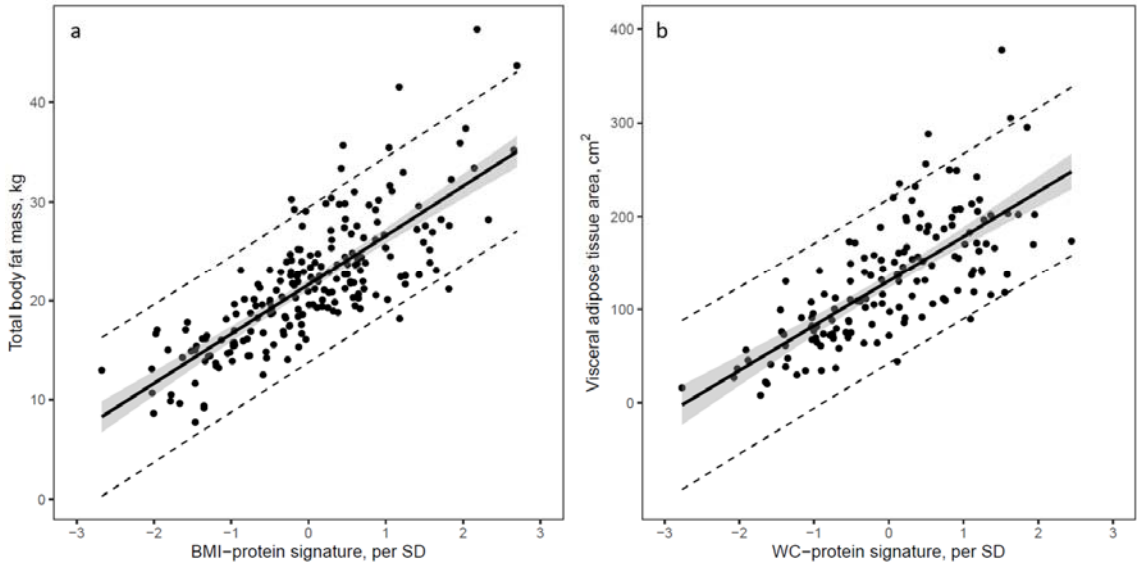

Figure 4. Scatterplots of (a) total body fat mass against BMI-protein signature in 207 Chinese participants with DXA and (b) visceral adipose tissue volume against the WC-protein signature in 151 Chinese participants with CT. Linear regression was fitted with the shaded region representing the 95% confidence intervals and the dotted lines representing the 95% prediction intervals. Abbreviations: CT, computed tomography; DXA, dual-energy X-ray absorptiometry; SD, standard deviation.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.25.23293136; this version posted July 27, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

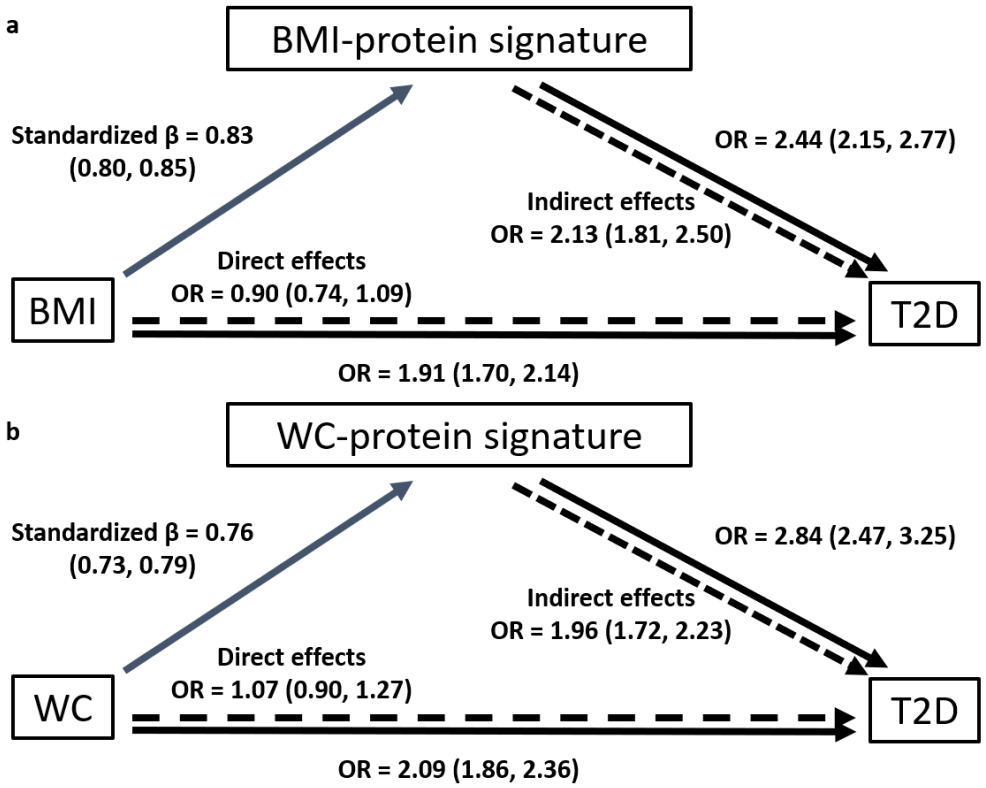

Figure 5. Directed acyclic graphs of the associations between T2D incidence and (a) BMI and the BMI-protein signature, and (b) WC and the WC-protein signature. All models were adjusted for age, sex, and ethnicity, and all predictor variables were expressed as per 1 standard deviation increment to facilitate the comparison of effect sizes. The direct and indirect effects were estimated using mediation analysis with natural effect models. Direct effects are the independent effects of BMI/WC on T2D incidence after controlling for the protein signatures. Indirect effects are the estimated effects of BMI or WC on T2D incidence that is mediated through the respective protein signatures. Abbreviations: BMI, body mass index; WC, waist circumference.