Diabetologia (2023) 66:174–189
https://doi.org/10.1007/s00125-022-05788-1

ARTICLE

Targeted proteomics identifies potential biomarkers of dysglycaemia, beta cell function and insulin sensitivity in Black African men and women

Amy E. Mendham[1,2] · Lisa K. Micklesfield[1] · Fredrik Karpe[3,4] · Andre Pascal Kengne[5] · Tinashe Chikowore[1] · Clement N. Kufe[1,6] · Maphoko Masemola[1] · Nigel J. Crowther[7] · Shane A. Norris[1,8] · Tommy Olsson[9] · Sölve Elmståhl[10,11] · Tove Fall[12] · Lars Lind[13] · Julia H. Goedecke[1,2,5]

Received: 4 January 2022 / Accepted: 9 June 2022 / Published online: 17 September 2022

Abstract

Aims/hypothesis Using a targeted proteomics approach, we aimed to identify and validate circulating proteins associated with impaired glucose metabolism (IGM) and type 2 diabetes in a Black South African cohort. In addition, we assessed sex-specific associations between the validated proteins and pathophysiological pathways of type 2 diabetes.

Methods This cross-sectional study included Black South African men (n=380) and women (n=375) who were part of the Middle-Aged Soweto Cohort (MASC). Dual-energy x-ray absorptiometry was used to determine fat mass and visceral adipose tissue, and fasting venous blood samples were collected for analysis of glucose, insulin and C-peptide and for targeted proteomics, measuring a total of 184 pre-selected protein biomarkers. An OGTT was performed on participants without diabetes, and peripheral insulin sensitivity (Matsuda index), HOMA-IR, basal insulin clearance, insulin secretion (C-peptide index) and beta cell function (disposition index) were estimated. Participants were classified as having normal glucose tolerance (NGT; n=546), IGM (n=116) or type 2 diabetes (n=93). Proteins associated with dysglycaemia (IGM or type 2 diabetes) in the MASC were validated in the Swedish EpiHealth cohort (NGT, n=1706; impaired fasting glucose, n=550; type 2 diabetes, n=210).

Results We identified 73 proteins associated with dysglycaemia in the MASC, of which 34 were validated in the EpiHealth cohort. Among these validated proteins, 11 were associated with various measures of insulin dynamics, with the largest number of proteins being associated with HOMA-IR. In sex-specific analyses, IGF-binding protein 2 (IGFBP2) was associated with lower HOMA-IR in women (coefficient –0.35; 95% CI –0.44, –0.25) and men (coefficient –0.09; 95% CI –0.15, –0.03). Metalloproteinase inhibitor 4 (TIMP4) was associated with higher insulin secretion (coefficient 0.05; 95% CI 0.001, 0.11; p for interaction=0.025) and beta cell function (coefficient 0.06; 95% CI 0.02, 0.09; p for interaction=0.013) in women only. In contrast, a stronger positive association between IGFBP2 and insulin sensitivity determined using an OGTT (coefficient 0.38; 95% CI 0.27, 0.49) was observed in men (p for interaction=0.004). A posteriori analysis showed that the associations between TIMP4 and insulin dynamics were not mediated by adiposity. In contrast, most of the associations between IGFBP2 and insulin dynamics, except for insulin secretion, were mediated by either fat mass index or visceral adipose tissue in men and women. Fat mass index was the strongest mediator between IGFBP2 and insulin sensitivity (total effect mediated 40.7%; 95% CI 37.0, 43.6) and IGFBP2 and HOMA-IR (total effect mediated 39.1%; 95% CI 31.1, 43.5) in men.

✉ Amy E. Mendham
amy.mendham@uct.ac.za

Extended author information available on the last page of the article

Research in context

What is already known about this subject?

- Studies have used targeted proteomics to better understand the pathophysiology of type 2 diabetes, and to identify convenient diagnostic and risk prediction models that may be used in a clinical setting
- Although the pathogenesis of type 2 diabetes appears to differ between Black Africans and Europeans, the diagnostic criteria and treatment strategies are the same

What is the key question?

- Is it possible to identify sex-specific circulating proteins associated with impaired glucose metabolism and type 2 diabetes in the Middle-Aged Soweto Cohort, and validate the results in a European population?

What are the new findings?

- We validated 34 proteins that were associated with type 2 diabetes, of which 11 were associated with measures of type 2 diabetes pathophysiology such as peripheral insulin sensitivity and beta cell function

How might this impact on clinical practice in the foreseeable future?

- Once validated in longitudinal studies, it is anticipated that these biomarkers, alongside the traditional glycaemic markers, may be used to develop more accurate sex-specific models for type 2 diabetes prediction

Conclusions/interpretation We validated 34 proteins that were associated with type 2 diabetes, of which 11 were associated with measures of type 2 diabetes pathophysiology such as peripheral insulin sensitivity and beta cell function. This study highlights biomarkers that are similar between cohorts of different ancestry, with different lifestyles and sociodemographic profiles. The African-specific biomarkers identified require validation in African cohorts to identify risk markers and increase our understanding of the pathophysiology of type 2 diabetes in African populations.

Keywords Adiposity · Beta cell function · Ethnicity · IGFBP2 · Impaired glucose metabolism · Insulin clearance · Insulin secretion · Insulin sensitivity · Obesity · TIMP4 · Type 2 diabetes

Abbreviations

DI	Disposition index
DXA	Dual-energy x-ray absorptiometry
FDR	False discovery rate
FMI	Fat mass index
IFG	Impaired fasting glucose
IGFBP2	IGF-binding protein 2
IGM	Impaired glucose metabolism
IGT	Impaired glucose tolerance
KIM1	Kidney injury molecule
MASC	Middle-Aged Soweto Cohort
NGT	Normal glucose tolerance
SAT	Subcutaneous adipose tissue
SCF	Stem cell factor
TIMP4	Metalloproteinase inhibitor 4
TNFRSF11A	TNF receptor superfamily member 11A
tPA	Tissue-type plasminogen activator
VAT	Visceral adipose tissue
VSIG2	V-set and immunoglobulin domain-containing protein 2

Introduction

Type 2 diabetes is a global health challenge, with Africa showing the highest projected relative increase over the next 25 years [1]. Africa is estimated to have the highest prevalence of undiagnosed diabetes (59.7%) and the highest proportion of deaths due to diabetes (73.1%) [1]. In 2016, type 2 diabetes was the second leading cause of death in South Africa overall, and the leading cause of death in women [2]. Although the exact mechanisms leading to the development of type 2 diabetes are unknown, studies have used targeted proteomics to better understand the pathophysiology of type 2 diabetes and to identify

convenient diagnostic and risk prediction models that may be used in a clinical setting [3–7]. However, these studies have been conducted in high-income populations; to our knowledge, no studies have been undertaken in low resourced African populations where sociocultural and lifestyle factors differ greatly.

Recent data suggest that there are differences in the pathogenesis of type 2 diabetes in populations of African ancestry, compared with their European counterparts [8–13]. However, it is unclear whether a similar profile of protein biomarkers is associated with the pathophysiology of type 2 diabetes in these populations as determined by insulin sensitivity, insulin secretion and insulin clearance. This is important to understand because currently the diagnostic criteria and treatment strategies used are the same for people of either African or European ancestry, due to limited data in African populations.

While many South African studies have included women, few have included data from men, preventing the exploration of potential sex differences in the factors associated with type 2 diabetes. The limited data available suggests that there are sex differences in the pathophysiology of type 2 diabetes, with preliminary evidence indicating that Black African women have higher insulin secretion and a higher disposition index (DI; estimated beta cell function) compared with men [14–16]. Furthermore, the inverse associations between body fatness and insulin sensitivity/beta cell function (DI) is stronger in Black African men compared with women [14]. but the insulin response decreases at a similar rate in both sexes as the population ages [16]. The reasons for these sex differences within African cohorts remain unclear, but they may reflect differences in total adiposity and body fat distribution [14, 16, 17]. Despite these differences, current guidelines for type 2 diabetes management do not differentiate between sexes or ancestry.

Therefore, the first aim of the current study was to use a targeted proteomics approach to identify circulating proteins associated with impaired glucose metabolism (IGM) and type 2 diabetes in a South African cohort, and to validate these against the Swedish EpiHealth cohort. The second aim was to use the South African cohort to assess sex-specific associations between the validated proteins and components of type 2 diabetes pathophysiology, namely insulin sensitivity, insulin secretion, insulin clearance and beta cell function.

Methods

Discovery cohort

Design, study population and setting This cross-sectional study includes the Middle-Aged Soweto Cohort (MASC) of Black South African men (n=502) and women (n=519) [18]. Participant recruitment and data collection took place between January 2017 and August 2018. Participants with known HIV or screen-detected using an HIV antibody test (One Step HIV-1/2, Guangzhou Wondfo Biotech, China) were excluded from this data analysis (n=185). In addition, participants who reported any form of ill health within the 7 days before testing (n=81) were also excluded. The final cohort included 755 participants (380 men and 375 women) (Fig. 1).

The study was performed in accordance with the Declaration of Helsinki and was approved by the Human Research Ethics Committee (Medical) of the University of the Witwatersrand (clearance certificate numbers M160604 and M160975). Written consent was obtained from all participants prior to their inclusion in the study. The testing procedures were performed at the South African Medical Research Council/WITS Developmental Pathways for Health Research Unit at the Chris Hani Baragwanath Hospital in Soweto, Johannesburg, South Africa.

Sociodemographic and medical questionnaire Interviewer-administered questionnaires were completed and the data were captured using REDCap [19]. Data collected included age, education (completed secondary education [>12 years], 10–12 years or ≤9 years of education), smoking (current smoker/non-smoker), alcohol (consumes/does not consume) and medication usage. Participants brought in all medications currently prescribed for type 2 diabetes, hypertension and hypercholesterolaemia, and details of these were recorded.

Anthropometry and body composition Weight was measured to the nearest 0.1 kg using a calibrated scale (TBF-410, Tanita, USA), and height was measured to the nearest 0.1 cm using a stadiometer (Holtain, UK). BMI was calculated and categorised according to the WHO criteria [20]. Waist circumference was measured at the midpoint between the last rib and the top of the iliac crest [21]. Dual-energy x-ray absorptiometry (DXA) (QDR 4500A, Hologic, USA) was used to measure subtotal fat mass (total body fat mass minus head) reported as kg. DXA data were analysed using APEX version 13.4.2.3 (Hologic). Fat mass index (FMI) was calculated as the subtotal fat mass (kg) divided by height squared (m^2). Abdominal visceral adipose tissue (VAT) and subcutaneous adipose tissue (SAT) were estimated using DXA [22].

Blood sampling, OGTT and biochemistry Fasting blood samples (10 ml) were collected after an overnight fast (10–12 h) for determination of glucose, insulin and C-peptide levels. Participants then completed a standard 2 h OGTT. After ingestion of glucose (75 g), venous blood

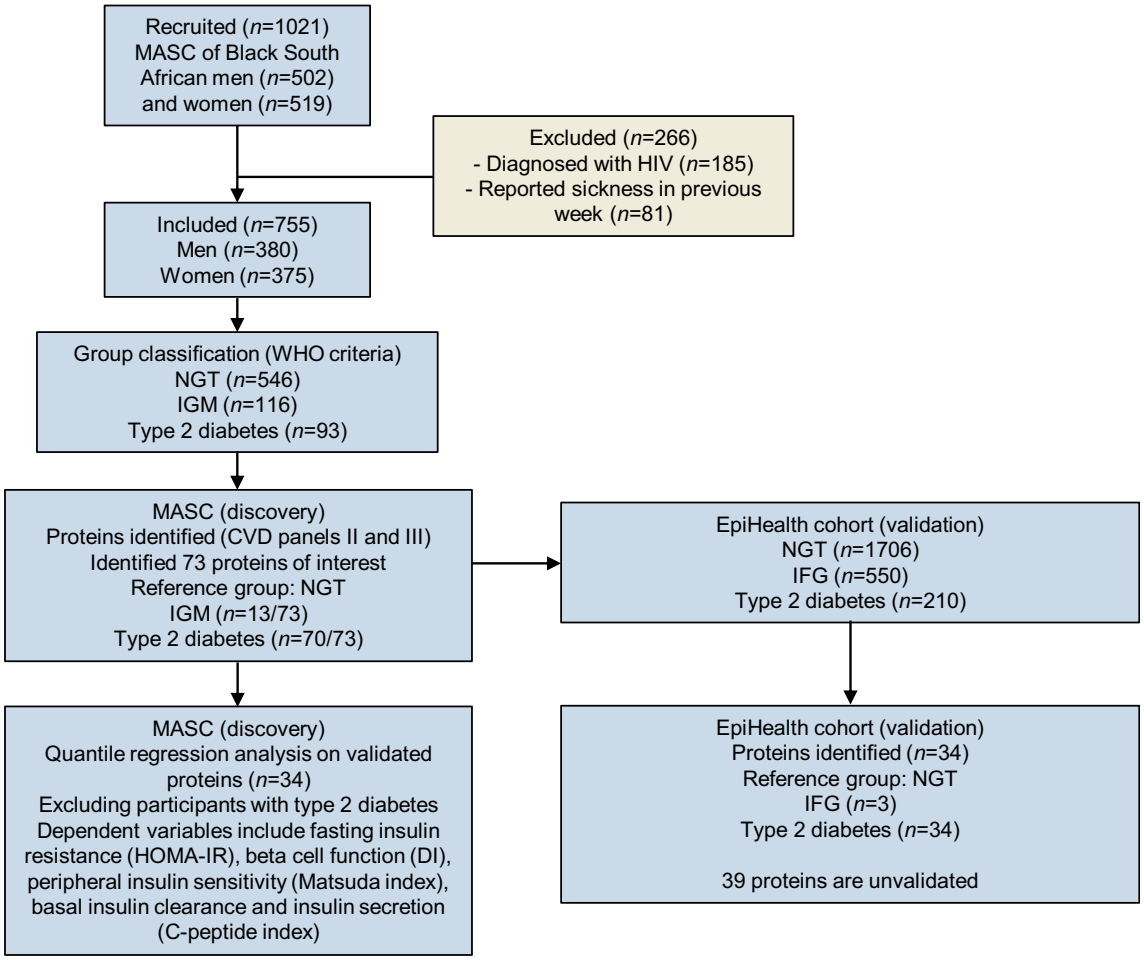


Fig. 1 Flow chart of study recruitment and analysis procedures

samples (approximately 5 ml) were drawn at 30, 60, 90 and 120 min to determine glucose, insulin and C-peptide levels. Participants with known diabetes did not complete the OGTT.

Plasma glucose concentrations were measured using an automated biochemistry analyser (RX Daytona, Randox, UK), as were serum insulin and C-peptide concentrations (IMMULITE 1000, Siemens, USA). The inter- and intra-assay CVs were 1.3% for glucose, 4.7% for insulin and 2.0% for C-peptide.

Proteomic analysis on plasma samples was performed using multiplex immunoassays (Olink Proseek cardiovascular disease panels CVD II and CVD III; Olink, Sweden), measuring a total of 184 pre-selected protein biomarkers that relate to the pathophysiology of type 2 diabetes (i.e. inflammation, immune response, chemotaxis, angiogenesis and the mitogen-activated protein kinase cascade; see https://www.olink.com/products-services/target/cardiometabolic-panel/). Two proteins with 36% (natriuretic peptides B, BNP) and 62% (spondin-1, SPON1) of data below the limit of detection were removed from the analysis, allowing the analysis of 182 proteins (from 0.9–9.3% below the limit of detection). The kits are based on proximity extension assay technology, and 92 oligonucleotide-labelled antibody probe pairs in each kit bind to their respective targets in the sample. All samples were randomised across plates, and the inter- and intra-assay CVs are based on control samples (pooled plasma samples) included on each plate. The intra-assay CV was 5% for both panels, and the inter-assay CVs were 12% and 16% for panels CVD II and CVD III, respectively. Data are presented as normalised protein expression (NPX) values, an arbitrary unit on a $\log_2$ scale (see www.olink.com/downloads).

Classification of glucose tolerance status Based on the fasting and 2 h OGTT plasma glucose results, participants were classified into groups according to WHO criteria: (1) normal glucose tolerance (NGT) if fasting plasma glucose was <6.1 mmol/l with a 2 h post-glucose load <7.8 mmol/l; (2) impaired fasting glucose (IFG) if fasting plasma glucose was between 6.1 and 6.9 mmol/l; (3) impaired glucose tolerance (IGT) if the 2 h post-glucose load was 7.8–11.0 mmol/l; (4) type 2 diabetes if fasting plasma glucose was ≥7.0 mmol/l and/or 2 h post-glucose load was ≥ 11.1 mmol/l and/or on diabetes medication

[23, 24]. The IFG and/or IGT participants were combined and described as having impaired glucose metabolism (IGM). Dysglycaemia is defined as the combination of IGM and type 2 diabetes.

OGTT calculations HOMA-IR was calculated from fasting glucose and insulin concentrations [25, 26], and the Matsuda index, calculated from the OGTT, was used to estimate peripheral insulin sensitivity [27]. Basal insulin clearance was calculated using the ratio of fasting C-peptide to insulin. Insulin secretion was estimated using the C-peptide index ($\Delta_{\text{C-peptide30min-0min}}/\Delta_{\text{Glucose30min-0min}}$) [28]. The DI, which is an estimate of beta cell function, was calculated using the ratio of the C-peptide index to insulin sensitivity [29, 30].

Validation cohort

Between 2011 and 2016, men and women aged 45–75 years were randomly selected from the population registry in the Swedish cities of Malmö and Uppsala, and invited to participate in the EpiHealth cohort study [31]. The Regional Ethical Review Board approved the study at Uppsala University (Dnr 2010/402), and all participants provided written informed consent [3]. Targeted proteomics analysis using the Olink CVD II and CVD III panels was completed on 2466 randomly selected fasting plasma samples from the Uppsala part of the EpiHealth cohort as previously described [3]. Data collection also included measures of fasting glucose, height, weight, waist circumference and fat mass. Fat mass was analysed using bioelectrical impedance (Tanita, Japan), and the FMI was calculated from these results. In addition, the participants completed a web-based questionnaire (see associated documentation; https://snd.gu.se/en/catalogue/study/ext0087#dataset; in Swedish), which included information on sex, age, alcohol intake (drinks per week), education (completed >12 years, 10–12 years or ≤9 years of education) and smoking (current smoker/non-smoker) [3]. Using the same glycaemic group definitions described above, participants were categorised into NGT (n=1706), IFG (n=550) or type 2 diabetes (n=210) groups (Fig. 1). GeneMANIA analytical software [32] was used to complete a separate network analysis for validated proteins. Each network analysis included an additional 20 dominant proteins that were predicted to be involved in the network.

Statistical analyses

Descriptive statistics A Shapiro–Wilk test was first used to assess the distribution of continuous data. Participant characteristics that are normally distributed are reported as means ± SD, data that are not normally distributed are presented as medians (IQR) and categorical variables are reported as n (%). In the discovery cohort, the Mann–Whitney U test was used to explore differences between sexes for all continuous variables, and the χ^2 test was used to explore sex differences for categorical variables. Cohort specific z scores were calculated for all proteins and measures of insulin dynamics (HOMA-IR, insulin sensitivity [Matsuda index], basal insulin clearance, insulin secretion [C-peptide index] and beta cell function [DI]) based on the raw score minus the population mean, divided by the population SD. This allows direct comparison of the magnitude of the effect sizes for each individual protein, as well as comparisons across different outcomes. Missing data were excluded pairwise for all analyses. Significance was set at p<0.05. Analyses were performed using IBM SPSS Statistics version 27 (USA). A flow chart of analytical procedures is shown in Fig. 1.

Logistic regression in the discovery cohort Logistic regression was used to identify the proteins associated with IGM and/or type 2 diabetes, using the NGT group as the reference, and adjusting for factors that are associated with the development of type 2 diabetes: age, sex, adiposity (waist circumference and FMI) and sociodemographic characteristics (alcohol, smoking and education). The level of significance was set at a false discovery rate (FDR) of 5%, and data are presented as OR and 95% CI.

Logistic regression in the validation cohort All proteins significantly associated with dysglycaemia in the discovery cohort (MASC) (n=73) were analysed in the validation cohort, and logistic regression was used to identify proteins associated with IFG and/or type 2 diabetes, using the NGT group as the reference and adjusting for the same covariates as in the MASC regressions. The level of significance was set at FDR of 5%, and data are presented as OR and 95% CI for each protein. Thirty-four of the proteins were validated, and the effects of chronic disease medication use (antihypertensive and lipid-lowering medication) were further explored. All logistic regression models were analysed using STATA SE version 15.1 (USA).

Quantile regressions on validated proteins Proteins validated in the EpiHealth cohort (n=34) were further explored in the MASC in association with measures of insulin dynamics using quantile regression analysis at the median (q=0.5). As the purpose of this regression analysis was to understand biomarkers related to the pathophysiology of type 2 diabetes, all participants with type 2 diabetes were excluded from the analysis. The z scores for all validated proteins were regressed against the z scores for all measures of insulin dynamics as the dependent variables. The model included all covariates

mentioned above, except for the inclusion of VAT rather than waist circumference, and addition of a protein × sex interaction term. Regressions were stratified by sex where a significant interaction was observed. The level of significance was set at FDR of 5% for each outcome, and data are reported as coefficients with 95% CI, with significance set at $p<0.05$. Quantile regression analyses were performed using IBM SPSS Statistics version 27.

A posteriori mediation analysis Increasing total and visceral adiposity are major determinants in the pathogenesis of type 2 diabetes. Our analysis found two proteins (metalloproteinase inhibitor 4 [TIMP4] and IGF-binding protein 2 [IGFBP2]) that were associated with multiple measures of insulin dynamics and have previously been shown to be partly expressed in adipose tissue [33, 34]. A posteriori simple mediation analysis was used to determine whether FMI or VAT mediated the relationship between these key proteins (independent variable) and measures of insulin dynamics (dependent variable). Models on the combined cohort of men and women were adjusted for age, sex, alcohol, smoking and education. Sex-stratified models were used where a significant sex interaction effect had been demonstrated in the quantile regressions. These included insulin secretion and DI for TIMP4, and HOMA-IR, insulin sensitivity and insulin secretion for IGFBP2. In each model, a dependent variable (measure of insulin dynamics), independent variable (protein) and mediator variable (FMI or VAT) was included to calculate the unstandardised regression coefficients and to generate total, direct and indirect effects. Full mediation was assigned when there were significant total and indirect effects and non-significant direct effects. When the total, indirect and direct effects were significant, partial mediation was considered to have occurred. Inconsistent mediation was present when only an indirect effect was significant. Bootstrapping of 5000 replicates was used to generate confidence intervals around the indirect effect. Analysis was completed using the PROCESS macro (version 4) for IBM SPSS Statistics version 27 [35].

Table 1 Descriptive characteristics of the discovery and validation cohorts

Variable	MASC (discovery cohort) (n=755)	EpiHealth cohort (validation cohort) (n=2466)
Age (years)	54 (50–60)	61 ± 8
Women	375 (49.7)	502 (20.3)
Current smoker[a]	204 (27.1)	1766 (71.6)
Consumes alcohol[a]	376 (49.9)	2440 (98.9)
Completed secondary education[b]	291 (39.2)	1308 (53.0)
Fasting glucose (mmol/l)	5.0 (4.5–5.6)	5.7 ± 1.0
2 h glucose (mmol/l)	5.9 (4.8–7.2)	–
IFG	49 (6.5)	550 (22.3)
IGT	94 (12.5)	–
IGM	116 (15.4)	–
Type 2 diabetes	93 (12.3)	210 (8.5)
Waist circumference (cm)	96.0 (87.0–105.0)	92.6 ± 11.8
FMI (kg/m^2)	10.2 (6.4–14.4)	8.2 ± 3.0
BMI (kg/m^2)	29.7 (24.6–34.8)	26.5 ± 3.8
BMI categories[c]		
Underweight	27 (3.6)	10 (0.4)
Normal weight	169 (22.4)	852 (34.5)
Overweight	195 (25.8)	1201 (48.7)
Obese	364 (48.2)	403 (16.3)

Continuous data are reported as means ± SD, skewed data reported as median (IQR) and categorical data are reported as n (%)

[a] Missing data (MASC): n=753

[b] Missing data (MASC): n=742

[c] BMI categories based on WHO criteria: underweight (<18.5 kg/m^2), normal weight (18.5–24.9 kg/m^2), overweight (25.0–29.9 kg/m^2) and obese (>30.0 kg/m^2)

Results

Descriptive characteristics

Table 1 presents the descriptive characteristics for the discovery cohort (MASC) and the validation cohort (EpiHealth). These results show that participants in the MASC were younger, with a higher percentage of women, and a lower proportion who consumed alcohol, were smokers and had completed secondary education. The MASC had a higher proportion of participants with obesity and type 2 diabetes and a lower proportion of IFG than in the EpiHealth cohort. Sex-specific characteristics in the MASC are shown in Table 2. A greater proportion of men smoked and consumed alcohol, while women showed a higher FMI, prevalence of obesity and type 2 diabetes. A higher proportion of women were also on antihypertensive medication. Fasting glucose was not different between sexes, however, women had higher fasting insulin and C-peptide concentrations and HOMA-IR, and lower basal insulin clearance than men. In addition, insulin sensitivity (Matsuda index) was lower and insulin secretion (C-peptide index) was higher in women compared with men, but estimates of beta cell function (DI) did not differ between sexes. Based on the glycaemic group definitions, participants were classified as having NGT (n=546), IGM (n=116) or type 2 diabetes (n=93) (Fig. 1).

Identified biomarkers of IGM and type 2 diabetes

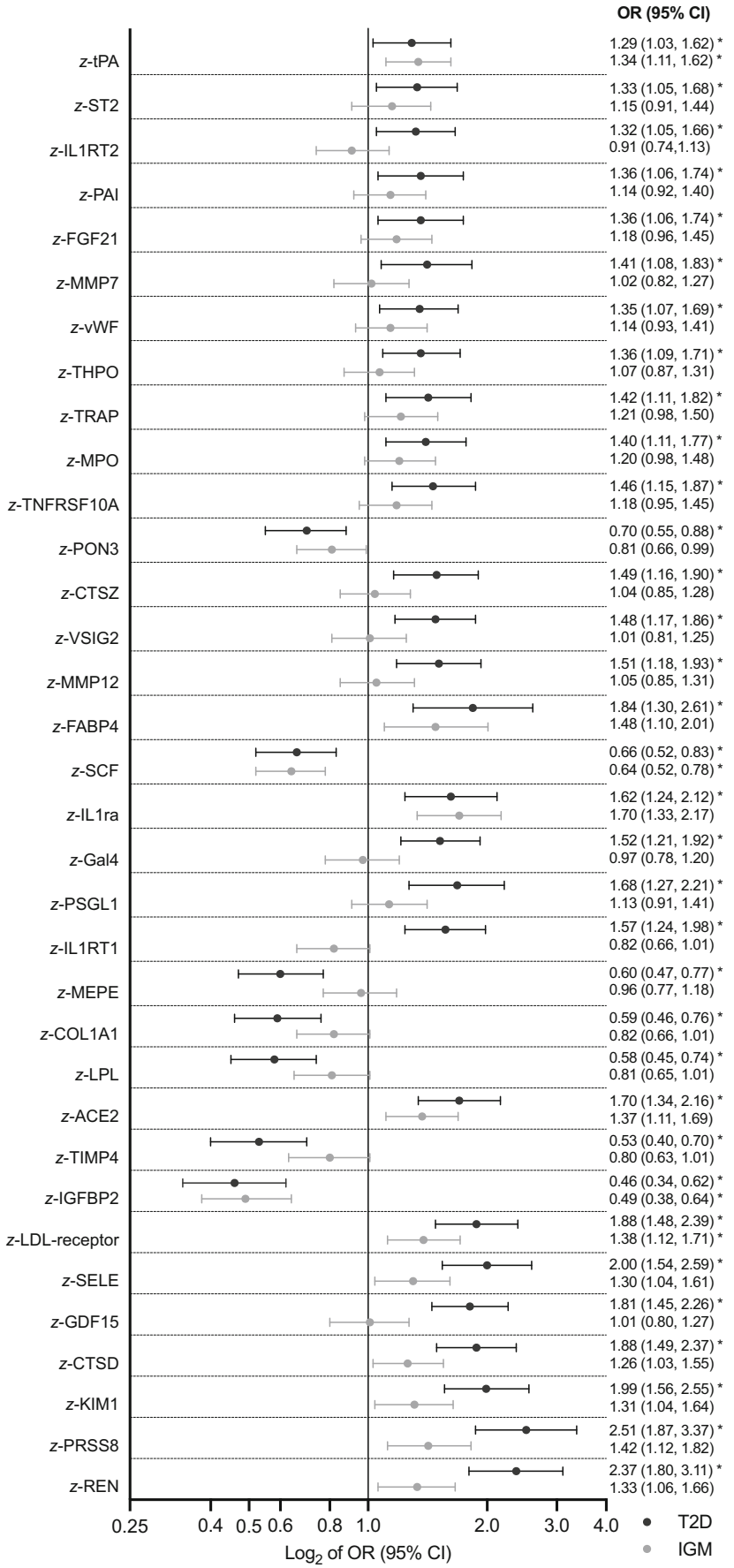
The MASC analysis identified 73 proteins of interest. Sixty proteins were significantly associated with only type 2 diabetes, ten were associated with both type 2 diabetes and

Table 2 Descriptive profile of the MASC (discovery cohort) and differences between men and women for behavioural, body composition and glycaemic characteristics

Variable	n	Men	n	Women
Age (years)	380	53 (48–59)	375	55 (51–59)
Current smoker	379	176 (46.4)	374	28 (7.5)
Consumes alcohol	379	272 (71.8)	374	104 (27.8)
Completed secondary education	376	150 (39.9)	366	141 (38.5)
BMI (kg/m^2)	380	26.0 (21.5–30.4)	375	33.3 (29.3–38.4)
BMI categories[a]				
Underweight		25 (6.6)		2 (0.5)
Normal weight		141 (37.1)		28 (7.5)
Overweight		113 (29.7)		82 (21.9)
Obese		101 (26.6)		263 (70.1)
FMI (kg/height^2)	368	6.5 (4.4–8.6)	359	14.3 (11.8–17.3)
VAT (cm^2)	363	84.2 (54.4–119.0)	358	108.4 (75.4–136.7)
SAT (cm^2)	363	204.6 (110.8–290.3)	358	472.5 (376.6–577.4)
VAT:SAT ratio	363	0.47 (0.35–0.63)	358	0.22 (0.18–0.28)
Fasting glucose (mmol/l)	380	5.0 (4.5–5.6)	375	5.0 (4.5–5.7)
2 h glucose (mmol/l)	355	5.8 (4.5–7.0)	338	6.0 (5.0–7.6)
Fasting insulin (pmol/l)	379	42.0 (15.8–83.4)	374	65.3 (36.4–104.6)
Fasting C-peptide (nmol/l)	379	0.53 (0.36–0.79)	374	0.60 (0.43–0.83)
HOMA-IR	378	1.4 (0.5–2.8)	374	2.2 (1.1–3.7)
Basal insulin clearance (ng/mlU)	378	0.28 (0.20–0.39)	374	0.20 (0.15–0.27)
Matsuda index	351	6.8 (3.5–12.4)	329	4.7 (2.8–8.0)
C-peptide index (ng/mmol)	314	2.3 (1.3–3.8)	291	2.7 (1.6–4.6)
DI (ng/mmol)	312	12.9 (6.7–25.0)	285	12.4 (5.4–25.9)
IGM	380	51 (13.4)	375	65 (17.3)
Type 2 diabetes	380	40 (10.5)	375	53 (14.1)
Glucose-lowering medication	380	21 (5.5)	375	33 (8.8)
Antihypertensive medication	380	102 (26.8)	375	136 (36.3)
Lipid-lowering medication	380	25 (6.6)	375	30 (8.0)

Continuous data are reported as medians (IQR) and categorical data are reported as n (%)

[a] BMI categories based on WHO criteria: underweight (<18.5 kg/m^2), normal weight (18.5–24.9 kg/m^2), overweight (25.0–29.9 kg/m^2) and obese (>30.0 kg/m^2)


OR (95% CI)
z-tPA
1.29 (1.03, 1.62) *
1.34 (1.11, 1.62) *
z-ST2
1.33 (1.05, 1.68) *
1.15 (0.91, 1.44)
z-IL1RT2
1.32 (1.05, 1.66) *
0.91 (0.74,1.13)
z-PAI
1.36 (1.06, 1.74) *
1.14 (0.92, 1.40)
z-FGF21
1.36 (1.06, 1.74) *
1.18 (0.96, 1.45)
z-MMP7
1.41 (1.08, 1.83) *
1.02 (0.82, 1.27)
z-vWF
1.35 (1.07, 1.69) *
1.14 (0.93, 1.41)
z-THPO
1.36 (1.09, 1.71) *
1.07 (0.87, 1.31)
z-TRAP
1.42 (1.11, 1.82) *
1.21 (0.98, 1.50)
z-MPO
1.40 (1.11, 1.77) *
1.20 (0.98, 1.48)
z-TNFRSF10A
1.46 (1.15, 1.87) *
1.18 (0.95, 1.45)
z-PON3
0.70 (0.55, 0.88) *
0.81 (0.66, 0.99)
z-CTSZ
1.49 (1.16, 1.90) *
1.04 (0.85, 1.28)
z-VSIG2
1.48 (1.17, 1.86) *
1.01 (0.81, 1.25)
z-MMP12
1.51 (1.18, 1.93) *
1.05 (0.85, 1.31)
z-FABP4
1.84 (1.30, 2.61) *
1.48 (1.10, 2.01)
z-SCF
0.66 (0.52, 0.83) *
0.64 (0.52, 0.78) *
z-IL1ra
1.62 (1.24, 2.12) *
1.70 (1.33, 2.17)
z-Gal4
1.52 (1.21, 1.92) *
0.97 (0.78, 1.20)
z-PSGL1
1.68 (1.27, 2.21) *
1.13 (0.91, 1.41)
z-IL1RT1
1.57 (1.24, 1.98) *
0.82 (0.66, 1.01)
z-MEPE
0.60 (0.47, 0.77) *
0.96 (0.77, 1.18)
z-COL1A1
0.59 (0.46, 0.76) *
0.82 (0.66, 1.01)
z-LPL
0.58 (0.45, 0.74) *
0.81 (0.65, 1.01)
z-ACE2
1.70 (1.34, 2.16) *
1.37 (1.11, 1.69)
z-TIMP4
0.53 (0.40, 0.70) *
0.80 (0.63, 1.01)
z-IGFBP2
0.46 (0.34, 0.62) *
0.49 (0.38, 0.64) *
z-LDL-receptor
1.88 (1.48, 2.39) *
1.38 (1.12, 1.71) *
z-SELE
2.00 (1.54, 2.59) *
1.30 (1.04, 1.61)
z-GDF15
1.81 (1.45, 2.26) *
1.01 (0.80, 1.27)
z-CTSD
1.88 (1.49, 2.37) *
1.26 (1.03, 1.55)
z-KIM1
1.99 (1.56, 2.55) *
1.31 (1.04, 1.64)
z-PRSS8
2.51 (1.87, 3.37) *
1.42 (1.12, 1.82)
z-REN
2.37 (1.80, 3.11) *
1.33 (1.06, 1.66)
0.25
0.4
0.5
0.6
0.8
1.0
2.0
3.0
4.0
Log2 of OR (95% CI)
T2D
IGM

◀ **Fig. 2** Forest plots showing the OR (95% CI) on a $\log_2$ scale of the proteins (*z* score) associated with type 2 diabetes (black) and IGM (grey) compared with NGT (reference). All 34 proteins were validated in the EpiHealth cohort and were associated with type 2 diabetes after adjustment for a 5% FDR. Only three proteins (tPA, SCF and IGFBP2) were also associated with IGM (indicated by an asterisk) after adjustment for a 5% FDR. Models were adjusted for age, waist circumference, FMI, sex, alcohol, smoking and education. COL1A1, collagen α-1 (I) chain; CTSD, cathepsin D; CTSZ, cathepsin Z; FABP4, fatty acid binding protein 4; FGF21, fibroblast growth factor 21; Gal4, galectin-4; GDF15, growth/differentiation factor 15; IGFBP2, IGF-binding protein 2; IL1ra, IL1 receptor antagonist protein; IL1RT1, IL1 receptor type 1; IL1RT2, IL1 receptor type 2; KIM1, kidney injury molecule; LPL, lipoprotein lipase; MEPE, matrix extracellular phosphoglycoprotein; MMP7, matrix metalloproteinase-7; MMP12, matrix metalloproteinase-12; MPO, myeloperoxidase; PAI, plasminogen activator inhibitor 1; PON3, paraoxonase; PRSS8, prostasin; PSGL1, P-selectin glycoprotein ligand 1; REN, renin; SCF, stem cell factor; SELE, E-selectin; ST2, interleukin 1 receptor-like 1; T2D, type 2 diabetes; THPO, thrombopoietin; TIMP4, metalloproteinase inhibitor 4; TNFRSF10A, TNF receptor superfamily member 10A; tPA, tissue-type plasminogen activator; TRAP, tartrate-resistant acid phosphatase type 5; VSIG2, V-set and immunoglobulin domain-containing protein 2; vWF, von Willebrand factor

IGM, and an additional three proteins were associated with IGM only. Of these 73 proteins of interest, 34 (47%) were validated in the EpiHealth cohort, of which three (IGFBP2, stem cell factor [SCF] and tissue-type plasminogen activator [tPA]) were associated with both IGM and type 2 diabetes, and 31 were associated with type 2 diabetes only (Fig. 2). When adjusting for antihypertensive and lipid-lowering medication, paraoxonase (PON3), matrix metalloproteinase-7 (MMP7), interleukin 1 receptor-like 1 (ST2) protein and V-set and immunoglobulin domain-containing protein 2 (VSIG2) were no longer associated with type 2 diabetes (see electronic supplementary material [ESM] Table 1). A network analysis using all validated proteins identified 526 total links between 54 proteins, including the 20 dominant proteins predicted to be involved in the network (ESM Results: GeneMANIA report). The network links show that co-expression (49.48%) and physical interactions (32.28%) are the dominant relationships between proteins within the network (ESM Fig. 1). The 39 proteins associated with IGM or type 2 diabetes in the MASC, but unvalidated in the EpiHealth cohort are shown in ESM Table 2. Of these 39 proteins, follistatin (OR 2.17; 95% CI 1.66, 2.84) and TNF receptor superfamily member 11A (TNFRSF11A) (OR 1.92; 95% CI 1.46, 2.52) were highly associated with type 2 diabetes ($p<0.0001$). Follistatin showed further associations with insulin secretion in the combined sample (coefficient –0.05; 95% CI –0.09, –0.01; p for interaction=0.561), and with HOMA-IR (coefficient –0.06; 95% CI –0.10, –0.01; p for interaction=0.018) and insulin sensitivity (coefficient 0.14; 95% CI 0.05, 0.23; p for interaction=0.004) in men only. TNFRSF11A was only associated with HOMA-IR in the combined sample (coefficient 0.09; 95% CI 0.01, 0.16; p for interaction=0.249) (ESM Table 3).

Associations between validated proteins and pathophysiological components of type 2 diabetes

When exploring the associations between validated proteins and the components of the pathophysiology of type 2 diabetes, participants with type 2 diabetes were excluded, and men and women were combined in the analysis (which included a protein × sex interaction term). In the combined sample of men and women from the MASC, HOMA-IR was positively associated with the concentration of seven proteins and inversely associated with the concentration of IGFBP2 (Fig. 3a). When exploring sex interactions, the positive association between HOMA-IR and cathepsin Z (CTSZ) was stronger in women (coefficient 0.12; 95% CI 0.03, 0.21) than men (coefficient 0.03, 95% CI –0.02, 0.08; p for interaction=0.010). The inverse association between HOMA-IR and IGFBP2 was greater in women (coefficient –0.35; 95% CI –0.44, –0.25) than men (coefficient –0.09; 95% CI –0.15, –0.03; p for interaction<0.001). When adjusting for antihypertensive and lipid-lowering medication, TIMP4, plasminogen activator inhibitor 1 (PAI) and kidney injury molecule (KIM1) showed a significant association with HOMA-IR (ESM Table 4). Insulin sensitivity was positively associated with IGFBP2 (coefficient 0.24; 95% CI 0.15, 0.32), with a stronger association in men (coefficient 0.38; 95% CI 0.27, 0.49) than in women (coefficient 0.22; 95% CI 0.14, 0.30; p for interaction=0.004), and remained significant when adjusting for chronic disease medication. Further, the inverse association between insulin sensitivity and matrix extracellular phosphoglycoprotein (MEPE) was stronger in men (coefficient –0.15; 95% CI –0.26, –0.04) than in women (coefficient –0.02; 95% CI –0.09, 0.06; p for interaction=0.008) (Fig. 3b). Basal insulin clearance was positively associated with IGFBP2 (coefficient 0.22; 95% CI 0.12, 0.32; $p<0.001$), and did not differ by sex (p for interaction=0.113) (Fig. 3c), and remained significant when adjusting for chronic disease medication. Insulin secretion showed a sex interaction with three proteins: IGFBP2 ($p=0.001$), TIMP4 ($p=0.025$) and VSIG2 ($p=0.045$), with a stronger association in women than in men (Fig. 3d). The DI was inversely associated with LDL-receptor (coefficient –0.04; 95% CI –0.07, –0.01; p for interaction=0.149) and positively associated with IGFBP2 (coefficient 0.07; 95% CI 0.04, 0.10; p for interaction=0.081). The relationship between DI and TIMP4 differed by sex (coefficient 0.06; 95% CI 0.03, 0.08; p for interaction=0.01), with a positive association in women (coefficient 0.06; 95% CI 0.02, 0.09) but not in men (coefficient –0.001; 95% CI –0.03, 0.02) (Fig. 3e). These results remained significant when adjusting for chronic disease medication (ESM Table 4).

A posteriori analysis

The results of mediation analyses are shown in Table 3. The associations between TIMP4 and measures of insulin dynamics were not mediated by adiposity, except for partial

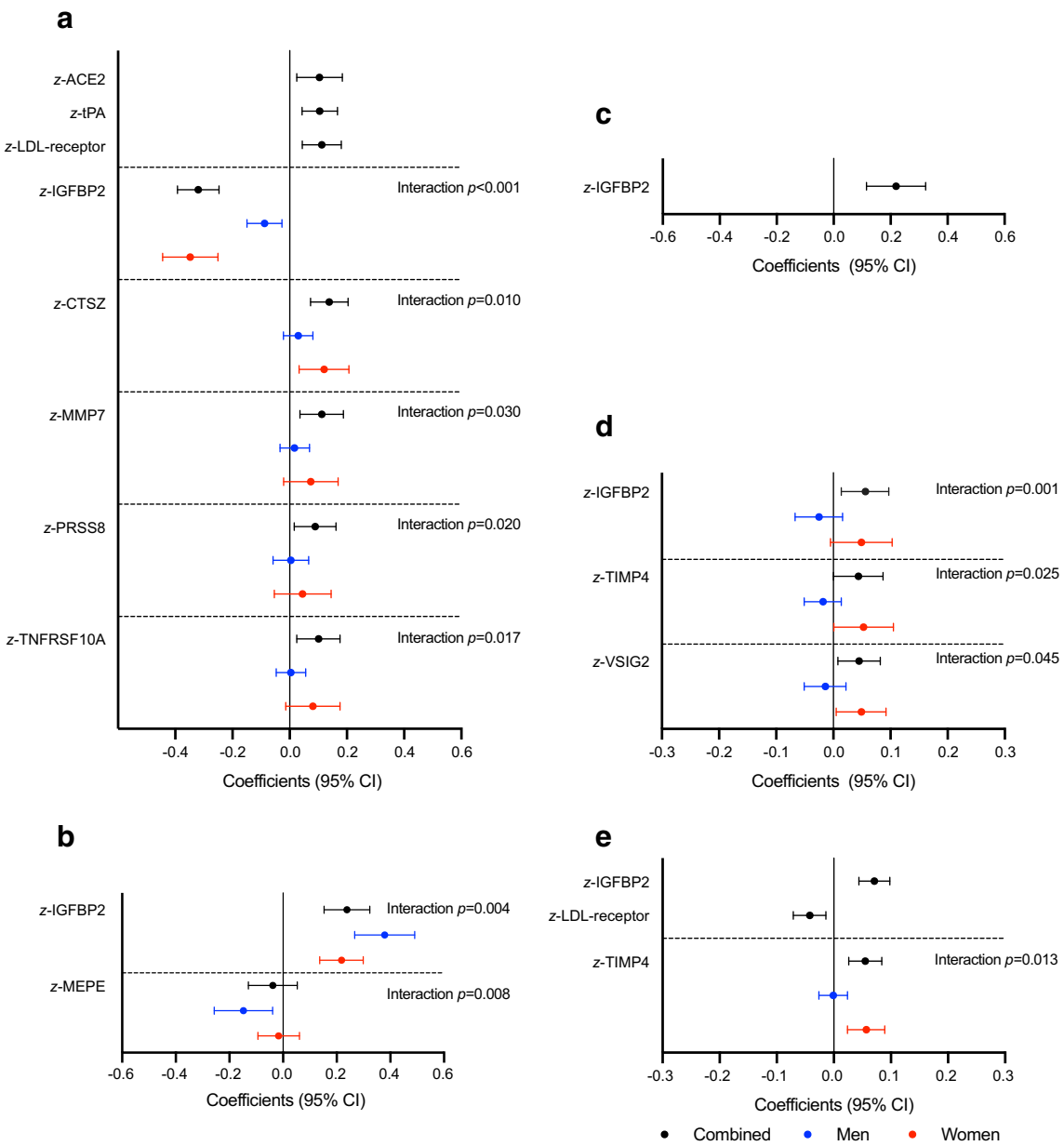


Fig. 3 Forest plots showing coefficients (95% CI) for the proteins (*z* score) significantly associated with (**a**) *z*-HOMA-IR, (**b**) *z*-insulin sensitivity, (**c**) *z*-basal insulin clearance, (**d**) *z*-insulin secretion (**e**) *z*-DI. Models were adjusted for age, VAT, FMI, alcohol, smoking, education, sex and a sex × protein interaction term. Blue, men; red, women; black, combined group. CTSZ, cathepsin Z; IGFBP2, IGF-binding protein 2; MEPE, matrix extracellular phosphoglycoprotein; MMP7, matrix metalloproteinase-7; PRSS8, prostasin; TIMP4, metalloproteinase inhibitor 4; TNFRSF10A, TNF receptor superfamily member 10A; tPA, tissue-type plasminogen activator; VSIG2, V-set and immunoglobulin domain-containing protein 2

mediation with DI in men only. In contrast, most of the associations between IGFBP2 and insulin dynamics, except for insulin secretion were mediated by either FMI or VAT in men and women.

Discussion

To our knowledge, this is the first study to use targeted proteomics to identify proteins associated with dysglycaemia in an African cohort. Specifically, we identified 73 proteins associated with IGM and/or type 2 diabetes in a cohort of middle-aged Black South African men and women, of which 34 proteins were validated in the EpiHealth cohort. Among the validated proteins, 11 were associated with components of type 2 diabetes pathophysiology such as fasting insulin resistance, peripheral insulin sensitivity and beta cell function. In particular, IGFBP2 and TIMP4 were associated with type 2 diabetes and multiple measures of insulin dynamics, which differed by sex and were only partially mediated by FMI and/or VAT.

Table 3 Mediation analysis to determine the contribution of FMI or VAT to the association between TIMP4 or IGFBP2 and measures of insulin dynamics in men and women

Dependent variable	Independent variable	Mediator	Total effect (95% CI)	Direct effect (95% CI)	Indirect effect (95% CI)	Proportion of total effect mediated (95% CI)
z-Insulin secretion (men)	*z*-TIMP4	FMI	–0.06 (–0.14, 0.02)	–0.06 (–0.14, 0.02)	–0.002 (–0.01, 0.006)	–
		VAT	–0.06 (–0.14, 0.02)	–0.06 (–0.14, 0.02)	0.002 (–0.003, 0.01)	–
z-Insulin secretion (women)	*z*-TIMP4	FMI	–0.02 (–0.22, 0.18)	–0.01 (–0.21, 0.19)	–0.009 (–0.05, 0.02)	–
		VAT	–0.02 (–0.22, 0.18)	–0.03 (–0.22, 0.17)	0.007 (–0.01, 0.03)	–
z-DI (men)	*z*-TIMP4	FMI	–0.05 (–0.15, 0.05)	–0.06 (–0.16, 0.04)	0.01 (0.0001, 0.03)*	24.3% (0.1, 58.0)‡
		VAT	–0.05 (–0.15, 0.05)	–0.06 (–0.16, 0.04)	0.01 (–0.003, 0.03)	–
z-DI (women)	*z*-TIMP4	FMI	–0.01 (–0.20, 0.17)	0.02 (–0.17, 0.20)	–0.03 (–0.09, 0.002)	–
		VAT	–0.01 (–0.20, 0.17)	–0.03 (–0.22, 0.15)	0.02 (–0.009, 0.06)	–
z-HOMA-IR (men)	*z*-IGFBP2	FMI	–0.35 (–0.47, -0.24)*	–0.22 (–0.35, –0.08)*	–0.14 (–0.21, –0.08)*	39.1% (31.1, 43.5)†
		VAT	–0.35 (–0.47, –0.24)*	–0.23 (–0.35, –0.10)*	–0.13 (–0.19, –0.08)*	36.5% (31.8, 40.1)†
z-HOMA-IR (women)	*z*-IGFBP2	FMI	–0.42 (–0.52, –0.31)*	–0.36 (–0.46, –0.25)*	–0.06 (–0.11, –0.03)*	14.4% (8.5, 20.3)†
		VAT	–0.42 (–0.0.52, –0.31)*	–0.33 (-0.44, –0.21)*	–0.09 (–0.16, –0.05)*	22.1% (14.7, 29.9)†
z-Basal insulin clearance	*z*-IGFBP2	FMI	0.33 (0.25, 0.41)*	0.28 (0.20, 0.37)*	0.05 (0.02, 0.08)*	14.8% (7.3,19.8)†
		VAT	0.33 (0.25, 0.41)*	0.31 (0.22, 0.40)*	0.02 (–0.02, 0.06)	–
z-Insulin secretion (men)	*z*-IGFBP2	FMI	–0.02 (–0.11, 0.06)	–0.02 (–0.11, 0.08)	–0.008 (-0.05, 0.03)	–
		VAT	–0.02 (–0.11, 0.06)	–0.04 (–0.13, 0.05)	0.02 (–0.01, 0.04)	–
z-Insulin secretion (women)	*z*-IGFBP2	FMI	0.08 (–0.11, 0.26)	0.06 (–0.13, 0.26)	0.01 (–0.04, 0.08)	–
		VAT	0.08 (–0.11, 0.26)	0.06 (–0.14, 0.26)	0.02 (–0.06, 0.09)	–
z-Insulin sensitivity (men)	*z*-IGFBP2	FMI	0.49 (0.37, 0.62)*	0.29 (0.15, 0.44)*	0.20 (0.14, 0.27)*	40.7% (37.0, 43.6)†
		VAT	0.49 (0.37, 0.62)*	0.29 (0.16, 0.43)*	0.20 (0.14, 0.28)*	40.4% (37.2, 44.7)†
z-Insulin sensitivity (women)	*z*-IGFBP2	FMI	0.37 (0.27, 0.47)*	0.31 (0.21, 0.41)*	0.06 (0.03, 0.10)*	16.1% (9.2, 21.8)†
		VAT	0.37 (0.27, 0.47)*	0.29 (0.19, 0.40)*	0.07 (0.03, 0.12)*	20.2% (11.7, 26.7)†
z-DI	*z*-IGFBP2	FMI	0.16 (0.06, 0.26)*	0.12 (0.02, 0.22)	0.04 (–0.003, 0.09)	–
		VAT	0.16 (0.06, 0.26)*	0.11 (0.004, 0.22)*	0.05 (0.007, 0.09)*	31.2% (11.5, 36.8)†

*Significant effect, $p<0.05$; † partial mediation, $p<0.05$; ‡ inconsistent mediation, $p<0.05$

'*z*-' indicates the *z* score

This study validated 34 proteins positively associated with prevalent type 2 diabetes, including three proteins associated with both IGM and type 2 diabetes in the African cohort. Of these, 23 proteins were associated with prevalent disease but not with any of the insulin traits commonly used to determine type 2 diabetes risk. Accordingly, the concentration of these proteins may simply reflect cardiovascular comorbidities and use of medication for chronic diseases associated with type 2 diabetes rather than a direct association with the disease. Insulin resistance and beta cell dysfunction are the two major processes in the pathophysiology of type 2 diabetes, with a decline in beta cell function evident prior to onset of IGT [36, 37, 38]. The influence of adiposity on proteins from Olink panel CVD II, as used in the current study, has been explored by Mendelian randomisation analysis in a European cohort, which showed that BMI causally influenced 32 of these proteins [39]. In our study, the association between nine of these 32 proteins and type 2 diabetes remained significant after adjusting for total and central adiposity (tPA, myeloperoxidase [MPO], growth/differentiation factor 15 [GDF15], fatty acid binding protein 4 [FABP4], SCF, cathepsin D [CTSD], IL1 receptor antagonist protein [IL1ra], KIM1 and E-selectin [SELE]), and only tPA was associated with measures of insulin dynamics. Further exploration is required to understand the insulin mechanisms that are independent of adiposity in the current cohort.

The results of the network analysis for all validated proteins identified functions related to endocrine processes, cytokine activity and response to IL1. Low-grade systemic inflammation is characterised by elevated concentrations of acute-phase proteins, cytokines and chemokines, which have been shown to be associated with altered beta cell function, insulin secretion and the overall occurrence of type 2 diabetes [40–42]. In particular, IL1β inhibits beta cell function and promotes beta cell apoptosis, while administration of a recombinant IL1

receptor antagonist protein, which inhibits IL1 receptor type 1 and causes interference with the IL1 pathway, has been shown to reduce expression of C-reactive protein and IL6, improve blood glucose levels and restore beta cell function in humans [43–48]. Our results show that IL1 receptor proteins are associated with prevalent type 2 diabetes, but not with insulin traits. This suggests that systemic concentrations of these proteins may reflect the disease endpoint rather than the pathophysiological processes involved in the development of type 2 diabetes.

We show the strongest relationships and largest number of proteins were associated with HOMA-IR, which predominantly reflects hepatic insulin resistance [49]. In addition to their metabolic roles, hepatocytes perform important immunological roles, which include the production of plasma proteins such as clotting factors, acute-phase proteins and inflammatory proteins [50], all pathways that are reflected in the targeted proteomics profile of the current study and other studies of this nature [3, 6]. Most notably, we showed that IGFBP2 was associated with all components of glycaemic control and insulin dynamics measured in this study, with clear sex differences in the associations with HOMA-IR and insulin sensitivity. These observations are supported by data from a population-based human cohort study that demonstrated a strong negative association between circulating IGFBP2 and type 2 diabetes risk, with epigenetic silencing of the *IGFBP2* gene predisposing to the development of type 2 diabetes [51]. In addition, circulating levels of IGFBP2 have been shown to be negatively associated with components of the metabolic syndrome and incident type 2 diabetes; however, no causal relationship was evident [52, 53]. The *IGFBP2* gene is predominantly expressed in the liver and white adipose tissue [33, 54], with a reduction in systemic IGFBP2 and *IGFBP2* mRNA expression in VAT, occurring with obesity [33]. Proteome-wide assessment of patients with type 2 diabetes has identified IGFBP2 as a novel prognostic marker of beta cell dysfunction [55], and a risk biomarker for early glycaemic disturbances [56], early dyslipidaemia via its independent association with VLDL-triglyceride [57], and hepatic energy metabolism and de novo lipogenesis [58]. When measured in an American cohort of European ancestry, the concentration of IGFBP2 was higher in men than in women [59]. Interestingly, we showed that the inverse association between IGFBP2 and fasting insulin resistance was stronger in women, whereas the positive association with peripheral insulin sensitivity was stronger in men. We further showed that visceral and total adiposity partially mediated the associations between IGFBP2 and all measures of insulin dynamics (except insulin secretion) in both men and women, and the sex differences may reflect sex-specific insulin traits that are independent of adiposity. Accordingly, further research is required to understand factors underlying sex-based differences in our cohort in order to provide mechanistic insight into the pathophysiology of type 2 diabetes.

Among the three proteins related to beta cell function in the current study, only the association with TIMP4 differed by sex, with a significant positive association in women only. This coincided with a significant positive association with insulin secretion in women only but no relationship with insulin sensitivity. The TIMP4 peptide is predominantly expressed in adipose tissue, promotes angiogenesis and is downregulated in obese mice [34, 60, 61]. In particular, *TIMP4* gene deletion in mice on a high-fat diet protected against hyperinsulinaemia and the accumulation of intramuscular triacylglycerol, but the mice showed defective lipid absorption that led to reduced adipocyte hypertrophy, fibrosis and diminished hepatic steatosis and dyslipidaemia [60]. Data on TIMP4 in humans are scarce, and our study shows that VAT and FMI are not mediators of the sex-specific associations with insulin secretion and beta cell function. Thus, the mechanisms by which TIMP4 affects insulin dynamics are uncertain and require further investigation.

Strengths and limitations

To our knowledge, the current study is the first in a sub-Saharan African cohort to report dysglycaemia and insulin traits in relation to a targeted proteomic profile, and provides an opportunity for follow-up research in this cohort. The cross-sectional nature of this study is a limitation when performing mediation analysis and when determining whether the identified proteins are the cause or consequence of the derangements in insulin dynamics that occur with dysglycaemia. Therefore, the present results must be confirmed in future mechanistic and prospective studies with incident cases of IGM and type 2 diabetes. Notably, follistatin and TNFRSF11A were strongly associated with type 2 diabetes only in the MASC, and may represent population-specific biomarkers that require further investigation, with a need for validation studies in different ethnic groups. However, with limited cohorts of this nature in Africa, and the cost–benefit limitations of proteomic studies in low-income communities, this approach is restricted. The proportions of men (26.6%) and women (70.1%) with obesity in the current study differ greatly, but reflect the current prevalence of obesity in South Africa [62]. Although the present study adjusts for FMI and VAT in the statistical modelling, the sex differences shown are not validated, and may be driven by the physiological consequences of the differences in adiposity, thus representing a limitation of the current study. A further limitation is that, although we considered the use of medication for chronic diseases, we cannot take into account prevalent cardiovascular disease as we did not have access to patient records to verify self-reported data. VAT was not measured in the EpiHealth cohort, and therefore waist circumference was used as a proxy in the validation phase. Regardless, adjusting for VAT in the MASC when exploring the associations between the validated proteins and components involved in the pathophysiology of type 2 diabetes remains a significant strength of the current study. The prevalence of IFG

and IGT vary between ethnic groups, with a low prevalence of IFG alongside a high rate of IGT and type 2 diabetes in Black South African cohorts studied previously [63] and in the current cohort. This emphasises the need to perform an OGTT to identify those with dysglycaemia when estimating diabetes risk in Black African cohorts. In the current study, participants were classified as having IGM based on fasting and OGTT measures in the discovery phase, whereas only fasting measures were available in the validation phase. Fasting glucose has been shown to be associated with incident type 2 diabetes in European cohorts [4]. Nonetheless, the classification of participants as having IGM based on the use of OGTT measures in a large African cohort remains a significant strength of this study in terms of identifying proteins associated with components relating to the pathophysiology of type 2 diabetes.

Conclusion

This study highlights biomarkers that are similar between cohorts of different ancestry, with different lifestyles and sociodemographic profiles. We show associations between 11 validated proteins and components of type 2 diabetes pathophysiology such as fasting insulin resistance, peripheral insulin sensitivity and beta cell function. Specifically, IGFBP2 and TIMP4 were associated with type 2 diabetes and multiple measures of insulin dynamics, which were only partially mediated by FMI and/or VAT. The African population-specific biomarkers identified require validation in African cohorts to identify risk markers and increase our understanding of the pathophysiology of type 2 diabetes in African populations.

Supplementary Information The online version contains peer-reviewed but unedited supplementary material available at https://doi.org/10.1007/s00125-022-05788-1.

Acknowledgements We are grateful to the participants in these cohorts, as well as the field staff of the Developmental Pathways for Health Research Unit. The graphical abstract was created with BioRender.com.

Data availability The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Funding The study was jointly funded by the South African Medical Research Council via the South African National Department of Health, the UK Medical Research Council (via the Newton Fund) and the GSK Africa Non-Communicable Disease Open Lab (via supporting grant project number ES/N013891/1) and the South African National Research Foundation (grant number UID:99108) and Västerbotten County Council, Sweden. TC is an International Training Fellow supported by Wellcome Trust grant number 214205/Z/18/Z. The study sponsors/funders were not involved in the design of the study, the collection, analysis and interpretation of data or writing the report, and did not impose any restrictions regarding publication of the report.

Authors' relationships and activities TO and TF are members of the Editorial Board of *Diabetologia*. The authors declare that there are no other relationships that might bias, or be perceived to bias, their work.

Contribution statement All authors have participated sufficiently in the work represented by the article and take public responsibility for the content. Participation included (i) conception or design of the work presented by the article, (ii) drafting the article or revising it for critically important content, and (iii) final approval of the version to be published. AEM is responsible for the integrity of the work as a whole.

References

1. International Diabetes Federation (2021) IDF Atlas, 10th edition. Available from https://diabetesatlas.org/. Accessed on 10 November 2021, DOI: https://doi.org/10.1111/jcpp.13557.
2. Statistics South Africa (2021) Mortality and causes of death in South Africa: Findings from death notification 2018. Available from https://www.statssa.gov.za/publications/P03093/P030932018.pdf. Accessed 10 November 2021, DOI: https://doi.org/10.1111/jcpp.13557.
3. Beijer K, Nowak C, Sundström J, Ärnlöv J, Fall T, Lind L (2019) In search of causal pathways in diabetes: a study using proteomics and genotyping data from a cross-sectional study. Diabetologia 62(11): 1998–2006. https://doi.org/10.1007/s00125-019-4960-8
4. Molvin J, Pareek M, Jujic A et al (2019) Using a targeted proteomics chip to explore pathophysiological pathways for incident diabetes – The Malmö Preventive Project. Sci Rep 9(1):1–7. https://doi.org/10.1038/s41598-018-36512-y
5. Abbasi A, Sahlqvist AS, Lotta L et al (2016) A systematic review of biomarkers and risk of incident type 2 diabetes: an overview of epidemiological, prediction and aetiological research literature. PLoS One 11(10):e0163721. https://doi.org/10.1371/journal.pone.0163721
6. Nowak C, Sundström J, Gustafsson S et al (2016) Protein biomarkers for insulin resistance and type 2 diabetes risk in two large community cohorts. Diabetes 65(1):276–284. https://doi.org/10.2337/db15-0881
7. Thorand B, Zierer A, Büyüközkan M et al (2021) A panel of 6 biomarkers significantly improves the prediction of type 2 diabetes in the MONICA/KORA study population. J Clin Endocrinol Metab 106(4):e1647–e1659. https://doi.org/10.1210/clinem/dgaa953
8. Goedecke JH, Olsson T (2020) Pathogenesis of type 2 diabetes risk in black Africans: a South African perspective. J Intern Med 288: 284–294. https://doi.org/10.1111/joim.13083
9. Hakim O, Bello O, Ladwa M et al (2019) Ethnic differences in hepatic, pancreatic, muscular and visceral fat deposition in healthy men of white European and black west African ethnicity. Diabetes Res Clin Pract 156:107866. https://doi.org/10.1016/j.diabres.2019.107866
10. Osei K, Schuster DP, Owusu SK, Amoah AGB (1997) Race and ethnicity determine serum insulin and C-peptide concentrations and hepatic insulin extraction and insulin clearance: comparative studies of three populations of West African ancestry and white Americans. Metabolism 46(1):53–58. https://doi.org/10.1016/S0026-0495(97)90167-0
11. Mohandas C, Bonadonna R, Shojee-Moradie F et al (2018) Ethnic differences in insulin secretory function between black African and white European men with early type 2 diabetes. Diabetes Obes Metab 20(7):1678–1687. https://doi.org/10.1111/dom.13283
12. Goedecke JH, Keswell D, Weinreich C et al (2015) Ethnic differences in hepatic and systemic insulin sensitivity and their associated determinants in obese black and white South African women. Diabetologia 58(11):2647–2652. https://doi.org/10.1007/s00125-015-3720-7
13. Osei K, Gaillard T (2017) Ethnic differences in glucose effectiveness and disposition index in overweight/obese African American and white women with prediabetes: a study of compensatory

mechanisms. Diabetes Res Clin Pract 130:278–285. https://doi.org/10.1016/j.diabres.2017.02.020
14. Kufe CN, Micklesfield LK, Masemola M et al (2022) Increased risk for type 2 diabetes in relation to adiposity in middle-aged black South African men compared to women. Eur J Endocrinol 186(5): 523–533. https://doi.org/10.1530/EJE-21-0527
15. Geer EB, Shen W (2009) Gender differences in insulin resistance, body composition, and energy balance. Gend Med 6:60–75. https://doi.org/10.1016/j.genm.2009.02.002
16. Goedecke JH, George C, Veras K et al (2016) Sex differences in insulin sensitivity and insulin response with increasing age in black South African men and women. Diabetes Res Clin Pract 122:207–214. https://doi.org/10.1016/j.diabres.2016.11.005
17. Nordström A, Hadrévi J, Olsson T, Franks PW, Nordström P (2016) Higher prevalence of type 2 diabetes in men than in women is associated with differences in visceral fat mass. J Clin Endocrinol Metab 101(10):3740–3746. https://doi.org/10.1210/jc.2016-1915
18. Goedecke JH, Nguyen KA, Kufe C et al (2022) Waist circumference thresholds predicting incident dysglycemia and type 2 diabetes in Black African men and women. Diabetes Obes Metab 24(5): 918–927. https://doi.org/10.1111/dom.14655
19. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG (2009) Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 42(2):377–381. https://doi.org/10.1016/j.jbi.2008.08.010
20. World Health Organization (2000) Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser 894:i-xii:1–253
21. World Health Organization (2011) Waist circumference and waist-hip ratio: report of a WHO expert consultation, Geneva, 8-11 December 2008. World Health Organization, Geneva, Switzerland
22. Micklesfield LK, Goedecke JH, Punyanitya M, Wilson KE, Kelly TL (2012) Dual-energy X-ray performs as well as clinical computed tomography for the measurement of visceral fat. Obesity 20(5): 1109–1114. https://doi.org/10.1038/oby.2011.367
23. American Diabetes Association (2019) Standards of medical care in diabetes 2019. Diabetes Care 42(Suppl 1):S124–S138. https://doi.org/10.2337/dc19-S011
24. World Health Organization (2006) Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. World Health Organization, Geneva, Switzerland.
25. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC (1985) Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28(7):412–419. https://doi.org/10.1007/BF00280883
26. Wallace TM, Levy JC, Matthews DR (2004) Use and abuse of HOMA modeling. Diabetes Care 27(6):1487–1495. https://doi.org/10.2337/diacare.27.6.1487
27. Matsuda M, DeFronzo RA (1999) Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 22(9):1462–1470. https://doi.org/10.2337/diacare.22.9.1462
28. Tura A, Kautzky-Willer A, Pacini G (2006) Insulinogenic indices from insulin and C-peptide: comparison of beta-cell function from OGTT and IVGTT. Diabetes Res Clin Pract 72(3):298–301. https://doi.org/10.1016/j.diabres.2005.10.005
29. Kahn SE, Prigeon RL, McCulloch DK et al (1993) Quantification of the relationship between insulin sensitivity and β-cell function in human subjects: evidence for a hyperbolic function. Diabetes 42(11):1663–1672. https://doi.org/10.2337/diab.42.11.1663
30. Ahrén B, Pacini G (2004) Importance of quantifying insulin secretion in relation to insulin sensitivity to accurately assess beta cell function in clinical studies. Eur J Endocrinol 150(2):97–104. https://doi.org/10.1530/eje.0.1500097
31. Lind L, Elmståhl S, Bergman E et al (2013) EpiHealth: a large population-based cohort study for investigation of gene–lifestyle interactions in the pathogenesis of common diseases. Eur J Epidemiol 28(2):189–197. https://doi.org/10.1007/s10654-013-9787-x
32. Warde-Farley D, Donaldson SL, Comes O et al (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 38(Suppl 2):W214–W220. https://doi.org/10.1093/nar/gkq537
33. Zhang X, Gu HF, Frystyk J, Efendic S, Brismar K, Thorell A (2019) Analyses of IGFBP2 DNA methylation and mRNA expression in visceral and subcutaneous adipose tissues of obese subjects. Growth Hormon IGF Res 45:31–36. https://doi.org/10.1016/j.ghir.2019.03.002
34. Mejia-Cristobal LM, Reus E, Lizarraga F et al (2015) Tissue inhibitor of metalloproteases-4 (TIMP-4) modulates adipocyte differentiation *in vitro*. Exp Cell Res 335(2):207–215. https://doi.org/10.1016/j.yexcr.2015.05.006
35. Hayes AF (2017) Introduction to mediation, moderation, and conditional process analysis. Guilford Press, New York, USA
36. Hudish LI, Reusch JEB, Sussel L (2019) β cell dysfunction during progression of metabolic syndrome to type 2 diabetes. J Clin Invest 129(10):4001–4008. https://doi.org/10.1172/JCI129188
37. Ferrannini E, Gastaldelli A, Miyazaki Y, Matsuda M, Mari A, DeFronzo RA (2005) β-cell function in subjects spanning the range from normal glucose tolerance to overt diabetes: a new analysis. J Clin Endocrinol Metab 90(1):493–500. https://doi.org/10.1210/jc.2004-1133
38. Mari A, Tura A, Natali A et al (2010) Impaired beta cell glucose sensitivity rather than inadequate compensation for insulin resistance is the dominant defect in glucose intolerance. Diabetologia 53(4):749–756. https://doi.org/10.1007/s00125-009-1647-6
39. Folkersen L, Gustafsson S, Wang Q et al (2020) Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. Nat Metab 2(10):1135–1148. https://doi.org/10.1038/s42255-020-00287-2
40. Kolb H, Mandrup-Poulsen T (2005) An immune origin of type 2 diabetes? Diabetologia 48(6):1038–1050. https://doi.org/10.1007/s00125-005-1764-9
41. Herder C, Baumert J, Thorand B et al (2006) Chemokines as risk factors for type 2 diabetes: results from the MONICA/KORA Augsburg study, 1984–2002. Diabetologia 49(5):921–929. https://doi.org/10.1007/s00125-006-0190-y
42. Herder C, Brunner EJ, Rathmann W et al (2009) Elevated levels of the anti-inflammatory interleukin-1 receptor antagonist precede the onset of type 2 diabetes: The Whitehall II Study. Diabetes Care 32(3):421–423. https://doi.org/10.2337/dc08-1161
43. Larsen CM, Faulenbach M, Vaag A et al (2007) Interleukin-1-receptor antagonist in type 2 diabetes mellitus. N Engl J Med 356(15):1517–1526. https://doi.org/10.1056/NEJMoa065213
44. Cavelti-Weder C, Babians-Brunner A, Keller C et al (2012) Effects of gevokizumab on glycemia and inflammatory markers in type 2 diabetes. Diabetes Care 35(8):1654–1662. https://doi.org/10.2337/dc11-2219
45. Rissanen A, Howard CP, Botha J, Thuren T (2012) Effect of anti-IL-1β antibody (canakinumab) on insulin secretion rates in impaired glucose tolerance or type 2 diabetes: results of a randomized, placebo-controlled trial. Diabetes Obes Metab 14(12):1088–1096. https://doi.org/10.1111/j.1463-1326.2012.01637.x
46. Eguchi K, Nagai R (2017) Islet inflammation in type 2 diabetes and physiology. J Clin Invest 127(1):14–23. https://doi.org/10.1172/JCI88877
47. Donath MY, Schumann DM, Faulenbach M, Ellingsgaard H, Perren A, Ehses JA (2008) Islet inflammation in type 2 diabetes: from metabolic stress to therapy. Diabetes Care 31(Suppl 2):S161–S164. https://doi.org/10.2337/dc08-s243

48. Dinarello CA (2000) The role of the interleukin-1-receptor antagonist in blocking inflammation mediated by interleukin-1. N Engl J Med 343(10):732–734. https://doi.org/10.1056/NEJM200009073431011
49. Abdul-Ghani MA, Matsuda M, Balas B, DeFronzo RA (2007) Muscle and liver insulin resistance indexes derived from the oral glucose tolerance test. Diabetes Care 30(1):89–94. https://doi.org/10.2337/dc06-1519
50. Robinson MW, Harmon C, O'Farrelly C (2016) Liver immunology and its role in inflammation and homeostasis. Cell Mol Immunol 13(3):267–276. https://doi.org/10.1038/cmi.2016.3
51. Wittenbecher C, Ouni M, Kuxhaus O et al (2019) Insulin-like growth factor binding protein 2 (IGFBP-2) and the risk of developing type 2 diabetes. Diabetes 68(1):188–197. https://doi.org/10.2337/db18-0620
52. Gudmundsdottir V, Zaghlool SB, Emilsson V et al (2020) Circulating protein signatures and causal candidates for type 2 diabetes. Diabetes 69(8):1843–1853. https://doi.org/10.2337/db19-1070
53. Elhadad MA, Wilson R, Zaghlool SB et al (2021) Metabolic syndrome and the plasma proteome: from association to causation. Cardiovasc Diabetol 20(1):1–13. https://doi.org/10.1186/s12933-021-01299-2
54. Hedbacker K, Birsoy K, Wysocki RW et al (2010) Antidiabetic effects of IGFBP2, a leptin-regulated gene. Cell Metab 11(1):11–22. https://doi.org/10.1016/j.cmet.2009.11.007
55. Belongie KJ, Ferrannini E, Johnson K, Andrade-Gordon P, Hansen MK, Petrie JR (2017) Identification of novel biomarkers to monitor β-cell function and enable early detection of type 2 diabetes risk. PLoS One 12(8):e0182932. https://doi.org/10.1371/journal.pone.0182932
56. Noordam R, van Heemst D, Suhre K, Krumsiek J, Mook-Kanamori DO (2020) Proteome-wide assessment of diabetes mellitus in Qatari identifies IGFBP-2 as a risk factor already with early glycaemic disturbances. Arch Biochem Biophys 689:108476. https://doi.org/10.1016/j.abb.2020.108476
57. Carter S, Li Z, Lemieux I et al (2014) Circulating IGFBP-2 levels are incrementally linked to correlates of the metabolic syndrome and independently associated with VLDL triglycerides. Atherosclerosis 237(2):645–651. https://doi.org/10.1016/j.atherosclerosis.2014.09.022
58. Fahlbusch P, Knebel B, Hörbelt T et al (2020) Physiological disturbance in fatty liver energy metabolism converges on IGFBP2 abundance and regulation in mice and men. Int J Mol Sci 21(11):4144. https://doi.org/10.3390/ijms21114144
59. Lau ES, Paniagua SM, Guseh JS et al (2019) Sex differences in circulating biomarkers of cardiovascular disease. J Am Coll Cardiol 74(12):1543–1553. https://doi.org/10.1016/j.jacc.2019.06.077
60. Sakamuri SSVP, Watts R, Takawale A et al (2017) Absence of tissue inhibitor of metalloproteinase-4 (TIMP4) ameliorates high fat diet-induced obesity in mice due to defective lipid absorption. Sci Rep 7(1):1–13. https://doi.org/10.1038/s41598-017-05951-4
61. Maquoi E, Munaut C, Colige A, Collen D, Lijnen HR (2002) Modulation of adipose tissue expression of murine matrix metalloproteinases and their tissue inhibitors with obesity. Diabetes 51(4):1093–1101. https://doi.org/10.2337/diabetes.51.4.1093
62. National Department of Health, Statistics South Africa, South African Medical Research Council, ICF (2017) South Africa – Demographic and Health Survey 2016: key findings. National Department of Health, Statistics South Africa, South African Medical Research Council and ICF, Pretoria, South Africa
63. Peer N, Steyn K, Lombard C, Lambert EV, Vythilingum B, Levitt NS (2012) Rising diabetes prevalence among urban-dwelling black South Africans. PLoS One 7(9):e43336. https://doi.org/10.1371/journal.pone.0043336

Affiliations

Amy E. Mendham[1,2] · Lisa K. Micklesfield[1] · Fredrik Karpe[3,4] · Andre Pascal Kengne[5] · Tinashe Chikowore[1] · Clement N. Kufe[1,6] · Maphoko Masemola[1] · Nigel J. Crowther[7] · Shane A. Norris[1,8] · Tommy Olsson[9] · Sölve Elmståhl[10,11] · Tove Fall[12] · Lars Lind[13] · Julia H. Goedecke[1,2,5]

1 South African Medical Research Council/WITS Developmental Pathways for Health Research Unit, Department of Paediatrics, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

2 Health through Physical Activity, Lifestyle and Sport Research Centre, International Federation of Sports Medicine (FIMS), International Collaborating Centre of Sports Medicine, Division of Physiological Sciences, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

3 Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

4 National Institute for Health and Care Research, Oxford Biomedical Research Centre, Oxford University Hospitals Foundation Trust, Oxford, UK

5 Biomedical Research and Innovation Platform and Non-Communicable Diseases Research Unit, South African Medical Research Council, Cape Town, South Africa

6 Epidemiology and Surveillance Section, National Institute for Occupational Health, National Health Laboratory Service, Johannesburg, South Africa

7 Department of Chemical Pathology, National Health Laboratory Service and University of the Witwatersrand Faculty of Health Sciences, Johannesburg, South Africa

8 School of Human Development and Health, University of Southampton, Southampton, UK

9 Department of Public Health and Clinical Medicine, Medicine, Umeå University, Umeå, Sweden

10 Department of Clinical Sciences in Malmö, Division of Geriatric Medicine, Lund University, Lund, Sweden

11 Clinical Research Centre, Skåne University Hospital, Malmö, Sweden

12 Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

13 Department of Medical Sciences, Uppsala University Hospital, Uppsala University, Uppsala, Sweden