ORIGINAL ARTICLE

Proteomic characterization of obesity-related nephropathy

Ralph Wendt[1,*], Tianlin He[2,*], Agnieszka Latosinska[2,*], Justyna Siwy[2,*], Harald Mischak[2,*] and Joachim Beige [1,3,*]

[1]Department of Nephrology and Kuratorium for Dialysis and Transplantation Renal Unit, Hospital St Georg, Leipzig, Germany, [2]Mosaiques Diagnostics , Hannover, Germany and [3]Department of Nephrology, Martin-Luther-University Halle/Wittenberg, Halle, Germany

Correspondence to: Joachim Beige; E-mail: Joachim.Beige@sanktgeorg.de
*All authors contributed equally to this work.

ABSTRACT

Background. Nephropathy related to obesity lacks a pathophysiological understanding and definite diagnostic pathways by biomarkers.

Methods.. In this study we investigated the association between urinary peptides and body mass index (BMI) and renal function in proteome data sets from 4015 individuals.

Results. A total of 365 urinary peptides were identified to be significantly associated with BMI. The majority of these peptides were collagen fragments. In addition, most of the peptides also demonstrated a significant concordant association with estimated glomerular filtration rate (eGFR) in the investigated cohort, with the presence of diabetes exhibiting no significant association. A new classifier was developed, based on 150 urinary peptides, that enabled the distinction of non-obese subjects with preserved kidney function from obese, non-diabetic subjects with eGFR $>45\,\text{mL/min/1.73}\,\text{m}^2$ in an independent cohort, with an area under the curve of 0.93.

Conclusions. On a molecular level, the data strongly suggest a link between obesity and fibrosis, which may be a major cause of obesity-related nephropathy.

Keywords: diabetes, glomerular filtration rate, obesity, peptide, renal risk, urinary

INTRODUCTION

In recent years there has been an increase in the prevalence of secondary diseases associated with obesity. This has led to a change in research focus on obesity in the context of abnormally high body weight and chronic kidney disease (CKD) [1, 2]. In clinical practice, CKD associated with obesity and improved renal function associated with weight loss are commonly seen, and trials in obesity treatment have yielded promising results in terms of improved renal function [3, 4], but a detailed understanding of the pathogenesis and epidemiology of nephropathy related to obesity, in the absence of diabetes, is rather limited [5]. Obesity causes various structural, haemodynamic and metabolic alterations in the kidney [6]. There is evidence of focal glomerulosclerosis in renal tissue samples obtained from patients with

obesity [7], which points to a role of hyperfiltration in nephropathy related to obesity. However, these histological features are different from the typical histological picture seen in diabetic nephropathy, even if similar to features seen in diabetic nephropathy in the presence of co-occurring obesity. 'Diabetoid' changes (focal mesangial sclerosis and focal thickening of glomerular and tubular basement membranes) can also be seen in obese patients without diabetes [7]. Thus the concurrent clinical manifestation of obesity and nephropathy in patients without diabetes leads to the hypothetical consideration of 'obesity-related nephropathy' (ORN), a condition that is characterized by renal injury in the presence of obesity, but in the absence of diabetes [8]. Such a 'non-diabetic and haemodynamic' mechanism underlying ORN pathophysiology could be related to a disrupted balance between adipokines mediating renal damage and protective adipokines secreted by adipose tissue in the setting of obesity [9, 10]. Therefore the question arises of whether there are specific biomarkers (different to those of diabetic nephropathy with obesity) that characterize nephropathy associated with obesity.

The urinary proteome contains a wealth of information on diseases and their pathophysiology. Proteomic changes are a major cause of non-communicable diseases, can precede and trigger the development of clinically relevant diseases and are targets for drug interventions. On this basis, numerous urinary proteins and peptides are considered as valuable biomarkers in the management of various diseases, including CKD. Several studies have investigated the use of urinary proteomic changes in the early detection of CKD, i.e. before the onset of irreversible kidney damage. Of particular interest is the CKD273 classifier, which is a panel of 273 urinary peptides, mostly collagen fragments. The CKD273 classifier was initially developed in a cross-sectional cohort and has subsequently been tested in many studies, several of which involved >1000 subjects [11]. The compiled data indicate that a major component of pathophysiology that is effectively identified by the CKD273 classifier, via the analysis of collagen degradation products, is fibrosis. The technique used to generate the CKD273 score was capillary electrophoresis coupled to mass spectrometry (CE-MS), which allows the assessment of thousands of urinary peptides in a single analysis. Therefore this approach using CE-MS on a database containing >70 000 urine proteomic data sets was considered highly suitable for investigating molecular changes in the context of ORN.

In light of the above, this present work aimed to investigate the impact of body mass index (BMI) on urinary peptides and to characterize as yet unspecified urinary peptide patterns that may be indicative of ORN, with a view to elucidating possible molecular mechanisms underlying ORN pathophysiology. We used urine samples from previous research work to assess for the presence of diabetes and associations with BMI, estimated glomerular filtration rate (eGFR) and age. Defining urinary peptide patterns could help provide insights into the pathophysiology of ORN that would in turn help in further characterization of ORN.

MATERIALS AND METHODS

Proteomic investigation and data retrieval

For this study, urinary proteome data stored in the Human Urine Proteome Database obtained by CE-MS were assessed. The CE-MS technique applied has been described in detail [12], including reproducibility, repeatability, sample preparation, data evaluation and normalization. This database contains urinary proteome/peptidome information on >70 000 samples, assessing a data space comprising a total of >100 000 peptides

and proteins [13], as well as anonymized clinical information on participants enrolled in the studies.

All data sets ($n = 4015$) that contained data on proteome, age (year), BMI (kg/m^2) and eGFR (mL/min/1.73 m^2) were extracted and no additional selection criteria were included. The presence of underlying kidney disease was identified by the presence or absence of diabetes and nephropathy ($n = 1435$). The presence of other renal diseases ($n = 80$) was not fully assessed because of a missing definite diagnosis in 2500 cases. The eGFR was calculated from serum creatinine levels using the Chronic Kidney Disease Epidemiology Collaboration equation [14].

Patient population, data acquisition, statistics and generation of classifier

In silico results of CKD273 proteomic profiling and all peptides detectable in >30% of all samples ($n = 771$), along with clinical data, were exported into an SPSS 22.0 data set (IBM, Armonk, New York, USA) and analysed further for possible associations of peptides and biomarker patterns with BMI and eGFR. Due to the large sample size, a P-value <0.01 was considered significant.

To generate a classifier that would potentially detect ORN, multiple peptides significantly associated with obesity and reduced kidney function were combined using support vector machines (SVMs). This methodology is described in further details in Mischak *et al.* [12]. Briefly, analysis was limited to patients without diabetes, with cases defined as patients with BMI >30 kg/m^2 and eGFR <45 mL/min/1.73 m^2, and controls defined as patients with BMI <27 kg/m^2 and eGFR >60 mL/min/1.73 m^2. Using the Excel 'RAND' function (Microsoft, Redmond, WA, USA), cases and controls were randomly split into training and test sets, based on the 2/3 and 1/3 rule, respectively. There was no significant difference in age, sex, BMI, eGFR and CKD273 scoring between the training and test sets.

A univariable linear regression model was used to examine the association between peptide abundance and continuous outcomes, including eGFR and BMI. Based on the hypothesis that the outcome (eGFR or BMI) Y is linearly related to individual peptide abundance X, the parameters α and β in the linear equation $Y_i = \alpha X_i + \beta + \varepsilon_i$ were estimated using the ordinary least squares method, so the sum of squares of error ε can be minimized. Graphically, α corresponds to the slope of the trend line drawn from the peptide abundance X and its predicted outcome Y in the fitted linear model. The coefficient of determination (R^2) is the square of the correlation between the predicted outcome and the actual outcome, which ranges from 0 (no correlation) to 1 (full correlation). Peptide abundance was log-transformed and undetectable abundance was omitted.

Ethics

All included studies were conducted to conform to regulations on the protection of subjects participating in medical research and in accordance with the principles of the Declaration of Helsinki (2013) and received ethical approval from their respective institutional review boards. Written informed consent was obtained from all participants at the time of sampling. All data sets received were anonymized. This study was approved by the ethics committee of the Hannover Medical School, Germany (no. 3116-2016)

RESULTS

All data sets with information on age, sex, eGFR, BMI, CKD273 scoring and CE-MS were extracted from the database (Table 1).

Table 1. Clinical characteristics of subjects providing urine samples

Characteristics	BMI >30 kg/m^2 ($n = 1234$)	BMI <30 kg/m^2 ($n = 2781$)	P-value
CKD273 scoring	−0.186 ± 0.634	−0.386 ± 0.598	<0.0001
Age (years)	60 ± 11.2	58 ± 13.8	<0.0001
eGFR <45 mL/min/1.73 m^2, n (%)	112 (9.1)	206 (7.4)	0.07
Diabetes, n (%)	589 (47.7)	846 (30.4)	<0.0001
BMI (kg/m^2)	34 ± 4.0	25 ± 2.7	<0.0001
eGFR (mL/min/1.73 m^2)	80 ± 24.9	81 ± 23.0	<0.0001
Albuminuria (mg/L)	216 ± 744	220 ± 827	>0.05

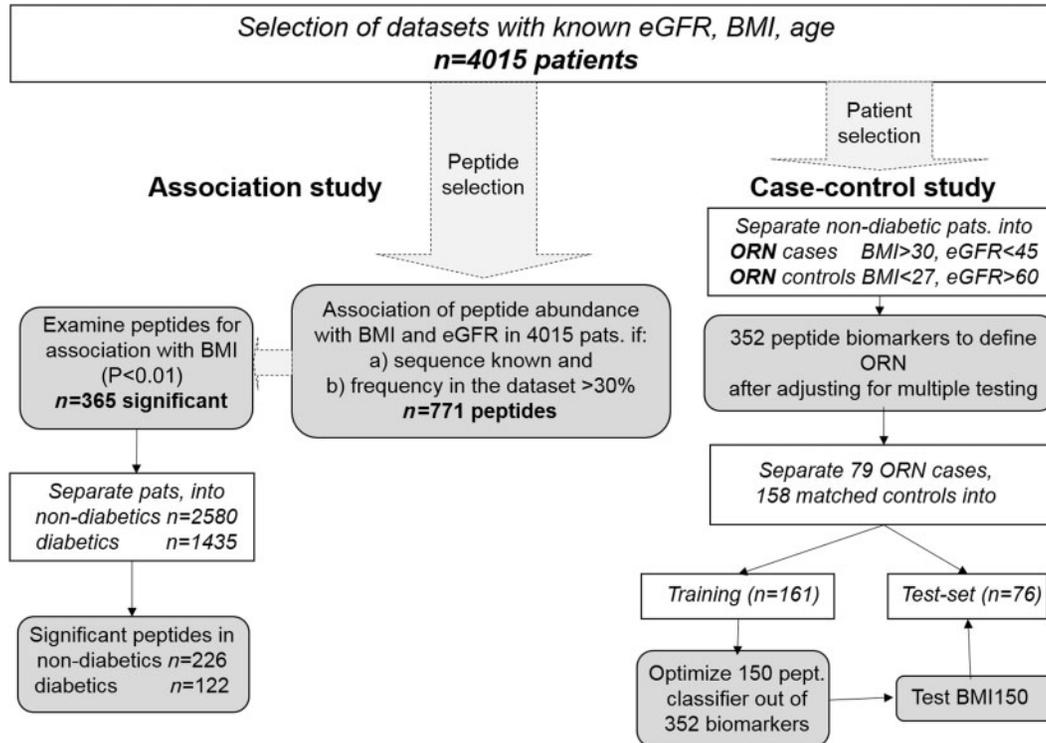Values presented as mean ± SD unless stated otherwise.

FIGURE 1: Flow chart showing the study design and use of samples for the different investigations. The full cohort of 4015 data sets was used to identify BMI-associated peptides. Separating the cohort into diabetic versus non-diabetic patients resulted in the identification of biomarkers for BMI relevant only to the specific subpopulation. To generate a classifier that can potentially detect ORN, biomarkers were defined in a subcohort of non-diabetic subjects and combined into a classifier, called BMI150.

Values presented as mean ± SD unless stated otherwise.

The study design and use of samples are shown in Figure 1. As expected, in the first step, a significantly strong association between eGFR and CKD273 score was observed ($\rho = 0.4748$, not shown). Although the association between eGFR and BMI in the overall cohort did not reach the pre-specified level of significance ($\rho = 0.064$; P = 0.013), a direct and significant association between BMI and CKD273 score was detectable (Figure 2). These data showed that urinary peptides reflect BMI at the molecular level.

To further examine any correlation between BMI and molecular features, we investigated the distribution of known (sequenced) urinary peptides associated with BMI. Only peptides that were observed in at least 30% of all data sets were included, which resulted in a total of 771 peptides that were assessed. Of these, 365 showed associations with BMI, reaching the target significance level of P < 0.01, with association strengths R ranging from 0.1 to 0.28. The 10 peptides showing the most significant association, based on R^2, are listed in Table 2.

The vast majority of significantly associated peptides are specific collagen fragments, mainly from collagen type I, and most of these peptides (in fact, all of the 10 most significant peptides) were found to be less abundant with increasing BMI. An example of the correlation between a single peptide and BMI is presented in Figure 3.

A large proportion ($n = 458$) of the peptides associated with BMI also showed an association with eGFR, but with an opposing trend, as expected, also reaching the same significance level of P < 0.01. This close concordant (concordance based on opposing association of urinary peptides with BMI and with eGFR, with low BMI and/or high eGFR being beneficial) correlation was clearly demonstrated when the slopes for the association of all investigated peptides with eGFR and BMI were evaluated. The highly significant inverse correlation between slopes of eGFR and BMI is shown in Figure 4.

In addition, we identified peptides that showed a discordant association with BMI and eGFR, depicted in Table 3.
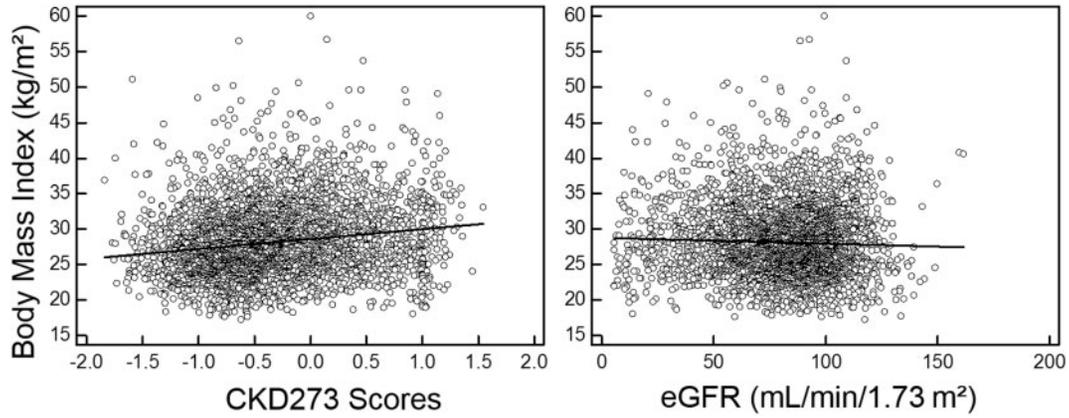
FIGURE 2: Association of the established multipeptide renal risk marker CKD273 (left panel) and eGFR (right panel) with BMI in 4015 subjects. Due to the large sample size, a P-value <0.01 was considered significant. While the association with CKD273 reached the predefined significance level [$\rho = 0.17$ (95% CI 0.14–0.20); P < 0.0001], a significant association between BMI and eGFR could not be established [$\rho = -0.039$ (95% CI –0.07 to –0.008); P = 0.013].

Table 2. Ten peptides showing the most significant association with BMI and eGFR

Sequence	Symbol	StartAA	StopAA	Slope (BMI)	P (BMI)	R^2 (BMI)	Slope (eGFR)	P (eGFR)	R^2 (eGFR)
ApGDRGEpGPpGPA	COL1A1	798	811	−3.4858	< 1E-20	0.0784	14.7468	< 1E-20	0.0641
GEAGHPGPPGPpGP	COL5A1	1555	1568	−3.0459	< 1E-20	0.0620	9.8429	< 1E-20	0.0348
SpGPDGKTGPpGPA	COL1A1	546	559	−2.9182	< 1E-20	0.0619	11.6615	< 1E-20	0.0464
DDGEAGKpGRpG	COL1A1	231	242	−3.0385	< 1E-20	0.0588	12.1982	< 1E-20	0.0487
GKNGDDGEAGKPGRpGERGPpGpQ	COL1A1	227	250	−2.8479	< 1E-20	0.0582	5.6369	2.39E-10	0.0107
ESGREGApGAEGSPGRDGSpGAKGDRGETGpA	COL1A1	1011	1042	−2.6923	< 1E-20	0.0576	5.9004	2.40E-12	0.0133
DGRpGPpGPpG	COL1A1	562	572	−2.5862	< 1E-20	0.0538	7.7913	6.41E-17	0.0266
DGEAGKpGRpG	COL1A1	232	242	−2.7042	< 1E-20	0.0520	6.9681	5.82E-09	0.0195
EAGRDGNpGNDGppGRDGQPGHKGERGYpG	COL1A2	923	952	−2.6878	< 1E-20	0.0519	2.6682	0.0065	0.0024
ADGQpGAKGEPGDAGAKGDAGppGPA	COL1A1	819	844	−2.6864	< 1E-20	0.0516	6.2465	1.38E-12	0.0128

AA, amino acid.

The slope was calculated using the linear regression model with the outcome (BMI or eGFR) as the dependent variable and individual peptide abundance as the independent variable. All peptides listed showed a negative association with BMI and lower abundance with higher BMI and a positive association with eGFR and higher abundance with reduced kidney function.
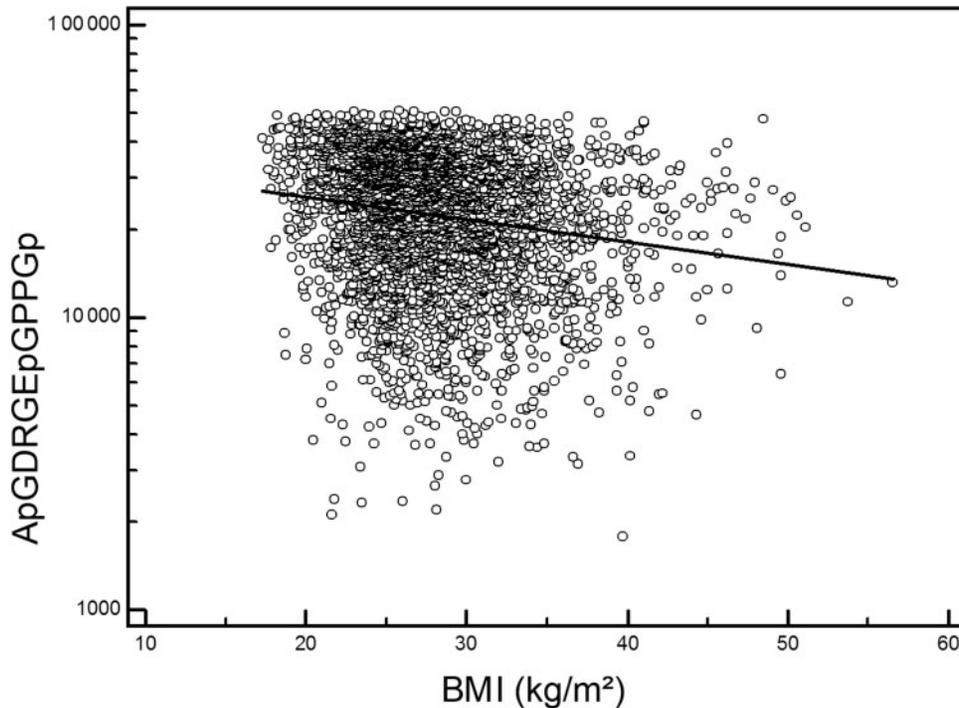

FIGURE 3: Correlation of a single peptide (ApGDRGEpGPPGp) with BMI. As most peptides were derived from collagen type I, this peptide showed a significant inverse correlation with BMI ($\rho = -0.17$; P < 0.0001).

Since we identified multiple peptides associated with BMI, we investigated if a classifier could be developed that detects ORN. Such a classifier would be evaluated in a subsequent study for its value in predicting disease progression and response to therapy, similar to the CKD273 classifi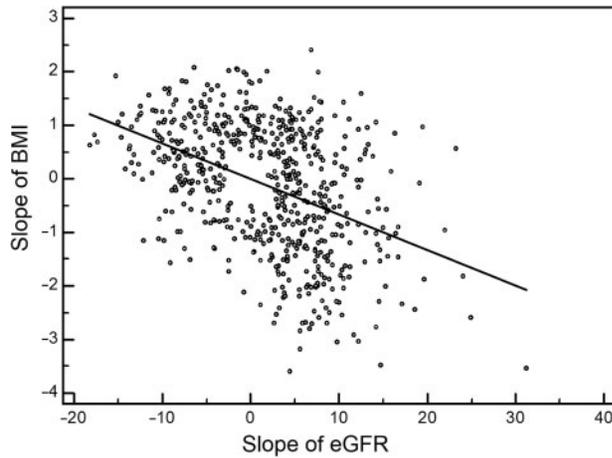er. Cases were limited to patients without diabetes and defined as patients with BMI $>30\,\text{kg/m}^2$ and eGFR $<45\,\text{mL/min/1.73 m}^2$ and controls as subjects with BMI $<27\,\text{kg/m}^2$ and eGFR $>60\,\text{mL/min/1.73 m}^2$. In the data set, 79 cases were identified, with 158 age-matched controls. The data set was split into a training set consisting of 105 controls and 56 cases, and an independent test set consisting of 53 controls and 23 cases. Patient characteristics are presented in Table 4. In the training set, a total of 352 biomarkers (minimum frequency 30%; $P < 0.01$ after adjusting for multiple testing) with known sequence were identified. These biomarkers were combined into an SVM-based classifier that was further optimized using a take-one-out procedure to ultimately generate a classifier based on 150 peptides, called BMI150. Applying this classifier to the independent test set of 76 samples resulted in very high accuracy to identify the ORN phenotype, with an area under the receiver operating characteristics curve (Figure 5A) of 0.929 (95% CI 0.846–0.975). On analysing all data sets available (excluding those used as part of the training set and patients with a BMI between 27 and 30 kg/m^2 or eGFR between 45 and 60 mL/min/1.73 m^2), the BMI150 classifier significantly distinguished between obese and non-obese patients with preserved kidney function, although this did not reach significance in CKD patients, likely because of the small number of subjects in that group (Figure 5B).

We investigated whether diabetes may have a significant impact on the association of urinary peptides with BMI, and so

FIGURE 4: Peptide associations (ρ) with eGRF and BMI. The plots show the slopes representing the association of 604 peptides with eGFR and BMI.

Table 3. Peptides showing discordant correlation (negative correlations marked in grey)

Protein name	Slope (BMI)	P (BMI)	R^2 (BMI)	Slope (eGFR)	P (eGFR)	R^2 (eGFR)
Neurosecretory protein VGF	−15.115	2.00E-20	0.0297	−68.366	4.22E-14	0.0246
Ubiquitin-like protein ATG12	15.101	7.99E-04	0.0239	72.403	7.66E-07	0.0335
Fibrinogen α chain	24.040	3.00E-20	0.0219	68.850	1.08E-03	0.0085
POTE ankyrin domain family member F	−15.612	1.48E-10	0.0191	−91.036	<1E-20	0.0307
Collagen α-1(III) chain	19.845	4.17E-14	0.0184	76.884	9.66E-10	0.0133
Collagen α-3(IV) chain	13.571	4.27E-12	0.0173	28.485	0.000178052	0.0037
Collagen α-1(III) chain	13.209	3.54E-05	0.0127	45.624	3.09E-01	0.0070
Collagen α-1(I) chain	14.398	3.99E-02	0.0123	78.565	2.14E-04	0.0171
Collagen α-2(I) chain	−13.082	1.49E-05	0.0112	−40.607	7.02E+00	0.0049
Collagen α-3(IX) chain	−11.549	7.25E-01	0.0107	−121.361	<1E-20	0.0523

Table 4. Characteristics of patient cohort used to generate a new classifier for the detection of ORN

Variable	Training set (56 cases/105 controls)	Test set (23 cases/53 controls)	P-value, training set versus test set	Group 1: cases ($n = 79$)	Group 2: controls ($n = 158$)	P-value, group 1 versus group 2
Median age (IQR; years)	74.0 (11.3)	74.0 (9.0)	0.5436[a]	74 (10.8)	74 (11.0)	0.8202[a]
Median eGFR (IQR; mL/min/1.73 m^2)	70.1 (46.2)	66.4 (37.8)	0.8028[a]	31.1 (17.5)	79.5 (20.0)	**0.0001**[a]
Median BMI (IQR; kg/m^2)	25.6 (7.6)	25.7 (8.5)	0.6975[a]	33.1 (3.8)	24.4 (3.0)	**0.0001**[a]
Median CKD273 score (IQR)	−0.10 (1.05)	−0.18 (1.00)	0.9611[a]	0.57 (0.82)	−0.43 (0.74)	**0.0001**[a]
Median albumin (IQR; mg/L)	277.1 (1792.6); $n = 32$	1540.2 (3484.0); $n = 16$	0.8013[a]	414.9 (2231.0); $n = 30$	119.5 (3762.0); $n = 18$	0.8646[a]
Sex (male/female)	94/54; $n = 148$	43/31; $n = 74$	0.5257[b]	39/35; $n = 74$	98/50; $n = 148$	0.0709[b]

[a]Mann–Whitney test.
[b]Chi-squared test.
IQR, interquartile range. Bold values indicate statistical significance.
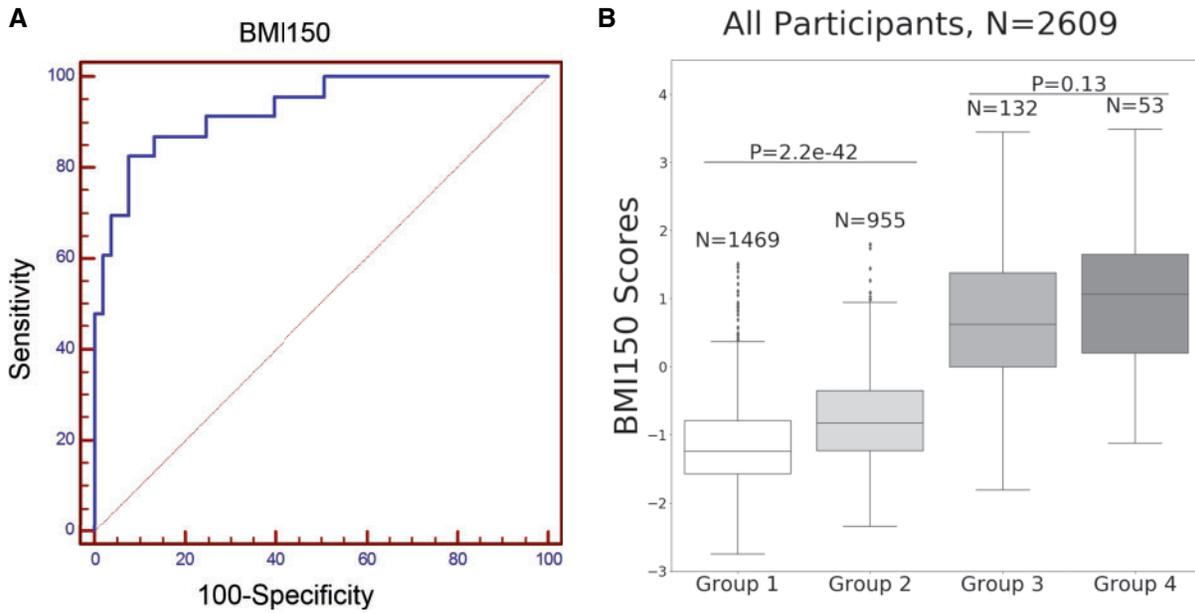
A

B

FIGURE 5: (**A**) Performance of a newly developed classifier based on 150 peptides, called BMI150, for the detection of ORN (arbitrarily defined as BMI $>$30 kg/m^2 and eGFR $<$45 mL/min/1.73 m^2, with the exclusion of diabetes) in an independent sample of 76 subjects. The figure shows highly significant results (area under the curve = 0.93) obtained from the receiver operating characteristics analysis. Only patients without diabetes were included in the analysis. Cases were defined as patients with BMI $>$30 kg/m^2 and eGFR $<$45 mL/min/1.73 m^2 and controls as BMI $<$27 kg/m^2 and eGFR $>$60 mL/min/1.73 m^2. (**B**) Comparison of classifiers in different patient groups in cases versus controls to detect the ORN phenotype. Group 1: eGFR $>$60 mL/min/1.73 m^2 and BMI $<$27 kg/m^2; group 2: eGFR $>$60 mL/min/1.73 m^2 and BMI $>$30 kg/m^2; group 3: eGFR $<$45 mL/min/1.73 m^2 and BMI $<$27 kg/m^2; and group 4: eGFR $<$45 mL/min/1.73 m^2 and BMI $>$30 kg/m^2.
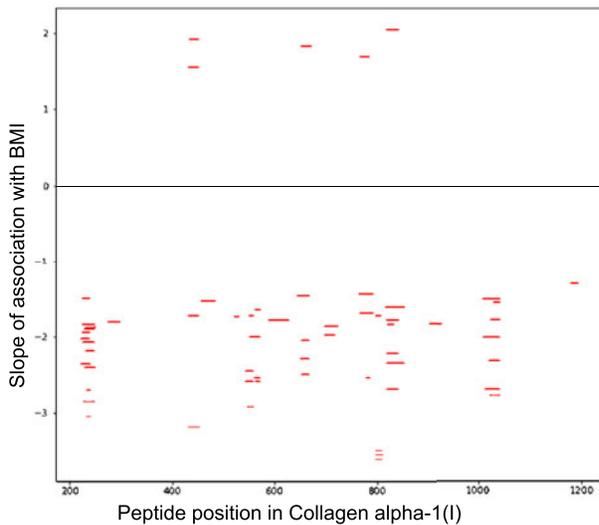

FIGURE 6: Apparent clustering of peptides associated with BMI. The figure shows the position (start and stop amino acids) in the collagen α-1(I) protein of the 60 peptides that showed the most significant association with BMI, as well as the slope of the association. The data argue against a random distribution and suggest that specific hot spots in the collagen α-1(I) protein, specifically at amino acid positions 220, 800 and 1020, are affected by BMI. Also evident is the negative association of most peptides with BMI.

we restricted the analysis to include only the 1435 diabetic subjects, irrespective of the type of diabetes. We detected a significant association with BMI for 122 of the 365 peptides, all concordant with respect to regulation (i.e. up- or downregulation) in the entire cohort of 4015 subjects. On investigating the 2580 non-diabetic subjects, 236 of the 365 biomarkers also showed a significant association with BMI, of which 235 were concordant and only one showed a significant inverse

correlation. When comparing the peptides significantly associated with BMI in the diabetic and non-diabetic populations, we found a very large overlap of 90 peptides (73.8% of all peptides were found significant in the smaller group comprising the diabetic patients). These data suggest that an impact of diabetes on the association of urinary peptides with BMI could be ruled out. The lower number of significantly associated peptides in each of the subcohorts was almost certainly a result of the reduced power due to the smaller sample.

While a large number of peptides, other than collagen-derived peptides, were found to be associated with eGFR, consistent with many previous studies, the fraction of non-collagen peptides associated with BMI was substantially lower. Of the 100 peptides most significantly associated with BMI, only 12 were not collagen-derived. Sixty of the collagen-derived fragments were from collagen α-1(I) protein. To obtain first insights into the association between collagen peptides and BMI, we analysed the distribution of the 60 collagen α-1(I)-derived peptides that were most significantly associated with BMI. As shown in Figure 6, there was a non-uniform peptide distribution throughout the collagen alpha-1(I) protein and these findings suggest that certain regions in the collagen α-1(I) protein are specifically affected/impacted by BMI.

DISCUSSION

Our analyses of urine samples from a large cohort of patients with known eGFR, BMI and diabetes status showed a complex multilevel association between peptide fragments and the risk of renal disease in obesity. Of particular note, however, the overall cohort demonstrated an almost negative eGFR–BMI gross correlation, presumably due to the origin of the data set derived from a non-obesity-directed approach. Most peptides associated with eGFR and BMI belonged to the collagen

superfamily and were inversely associated with BMI but directly with eGFR.

This concordant observation indicates a reduction of collagen degradation with increasing BMI and/or reduced kidney function. It is tempting to speculate that collagen homeostasis and BMI, on the one hand, and kidney function, on the other, are substantially interconnected.

The very strict association between BMI and collagen fragments was surprising, but it also indicates specificity. This led to the development of a well-performing new classifier based on 150 peptides (called BMI150) to identify patients with BMI >30 and eGFR <45 mL/min/1.73 m^2 without diabetes mellitus (Figure 4), thus confirming the entity of ORN on a urinary proteome level. In both groups of patients with either preserved or reduced kidney function, molecular information included in the BMI150 classifier was able to distinguish between obese and non-obese subjects with nephropathy. In advanced stages of CKD (eGFR <45 mL/min/1.73 m^2), this distinction failed to reach significance, likely due to the small number of patients included in this subgroup (based on the given results, a total of 238 patients, instead of the available 129, would have been needed to reach the significance level of 0.01 to distinguish ORN from non-obese nephropathy). These findings highlight the need to include study populations of >300 nephropathic patients with obesity in future research.

The renal interstitial matrix is normally composed of collagen types I, III, V, VI, VII and XV. On the other hand, mainly collagen types I and III are deposited in the early stages of renal fibrosis [15], as confirmed in animal models that also showed increased expression of α-1(I) mRNA in the interstitial space [16]. Collagen type I co-localizes with the extracellular matrix proteoglycans decorin and biglycan, which influence collagen type I fibrillogenesis. They accumulate in glomeruli, the tubulointerstitial space and arterial walls in human renal fibrosing disease [17]. Accumulation of collagen type I in fibrosis is due to both decreased collagen degradation and increased collagen synthesis [18, 19]. Among the proteins that were differentially expressed, there were significantly reduced quantities of fragments of collagen type I in the urine of diabetic patients [20], which suggests decreased collagen proteolysis, probably due to cross-linking rendering the collagens resistant to proteolytic cleavage or increased protease inhibitor expression [20].

The association of urinary peptides (as part of the CKD273 classifier) with the degree of renal fibrosis has been demonstrated in renal biopsies in CKD patients [21]. The CKD273 classifier consists mainly of collagen peptide fragments, with seven fibrosis-associated peptides displaying a negative association with the degree of fibrosis. These peptides are most probably an indicator of collagen degradation. The process of collagen degradation seems to be attenuated in CKD, leading to collagen accumulation in renal tissue, and consequently renal fibrosis. Markers of renal impairment (serum creatinine, blood urea nitrogen and serum cystatin C) are highly non-specific to the underlying disease process. In contrast, urinary peptides are specific indicators of fibrosis, which precedes progressive renal dysfunction, and thus would prove extremely valuable.

In obesity-related kidney disease, inhibition of the 5′-AMP-activated protein kinase (AMPK) pathway has been established as a critical pathway regulating inflammation and especially fibrosis [22]. Activation of AMPK was shown to reduce mesangial matrix expansion, as well as urinary levels of transforming growth factor β1 (TGF-β1) [23], and also to markedly reduce glomerular TGF-β1 and collagen accumulation in several mouse models of diabetic kidney disease [24]. In diet-induced obesity,

there is induction of TGF-β1 in the kidney, in association with an upregulation of extracellular matrix components, including collagen type I [25].

Another key player in the context of obesity is glycation, whereby excessive food intake results not only in increased BMI, but hypothetically also in an increase in glycation reactions from diabetes-dependent oxidation, as well as from external (non-diabetic) alimentary sources [26, 27]. Glycation, in turn, results in increased collagen cross-linking, which leads to resistance of collagen to physiological degradation. A correlation between the level of advanced glycation end products (AGEs) and collagen cross-linking was noted in a diabetic rat model, suggesting the involvement of advanced glycation in cross-linking [28]. Such processes with altered collagen fibrillar organization have also been shown under the influence of increased sugar exposure in non-diabetic models [29]. While there is convincing evidence that AGEs can directly induce collagen cross-linking, it is hypothesized that these AGEs, once formed, could induce collagen cross-linking even in the absence of glucose [30].

Increased cross-linking of collagen and decreased collagen degradation are most likely demonstrated by a reduction in the abundance of collagen fragments. Attenuation of physiological degradation will result in an increase in tissue collagen, and consequently fibrosis. Fibrosis has been increasingly recognized as a major player in adipose tissue dysfunction [31]. Following weight loss as a result of bariatric surgery, there is evidence of major collagen remodelling in subcutaneous adipose tissue, with increased collagen degradation and, importantly, decreased collagen cross-linking [32].

A highly relevant finding from this study is an overall very strong inverse correlation between peptides, mostly derived from collagen fragments, and BMI and eGFR. For most peptides, this highly similar association in our very large cohort indicates that there is a similar underlying molecular mechanism in both obesity and CKD. Based on our data and hypotheses on the collagen fragments presented above, we believe fibrosis may be a key overlapping element. However, it is important to highlight that some peptides were identified and showed a direct (discordant) association with BMI and eGFR. Of the 10 discordant peptides (Table 3), 6 were collagen derived, of which only one belonged to the collagen α-1(I) chain, with the other peptides belonging to other collagen subtypes. Different collagen subtypes have been shown to be differentially expressed and distributed in different structural parts of the kidney [33].

The peptides that showed a discordant association are the neurosecretory protein VGF, the ubiquitin-like protein ATG12, the fibrinogen α chain and POTE ankyrin domain family member F. VGF was found to be reduced in obesity, both in mouse models and in humans [34]. In a different study, a VGF peptide was reported to be negatively associated with eGFR [35]. As such, the discordant association of the VGF peptide with BMI and eGFR is in line with previous, albeit isolated, findings. Expression of the autophagy-related gene *ATG12* was reported to be attenuated in kidney disease [36], consistent with the positive correlation with eGFR found in our study. In an independent study, a significant positive correlation between *ATG12* expression and BMI was observed [37]. To the best of our knowledge, the role of POTE ankyrin domain family member F has not been investigated in the context of obesity. However, studies have shown that disruption of the *ankyrin repeat domain 26* gene (*ANKRD26*), the precursor gene of the *POTE* genes, results in obesity and gigantism [38], supporting a negative association with BMI, as observed in our study. We could not identify a study that investigated a potential association between POTE ankyrin

domain family member F and kidney function. The discordance observed with specific fibrinogen and collagen fragments may be a result of differential activation of proteases in obesity, in comparison with kidney disease, and not due to a general discordance in the regulation of these proteins.

STUDY LIMITATIONS

A major limitation of the study is that it is not based on a well-controlled prospective sampling of specific cases defining the condition under investigation but rather represents a meta-analysis of earlier cases. In particular, while diabetes was well represented in our cohort ($n = 1435$), other renal diseases ($n = 80$) were underrepresented due to a lack of histological or definitive serological diagnostic parameters. However, the very large number of data sets (>4000) from >20 different clinical centres should ensure the validity of our findings with regard to diabetes, obesity and nephropathy and reduce the risk of bias. Further studies should ensure data are available on renal diagnosis from histological or other investigations, although in patients with obesity, kidney biopsy would represent a considerable procedural challenge, as well as involve risks of complication. This reasoning underscores the need for non-invasive, risk-free means of obtaining a diagnosis, particularly in obese patients, with proteomic investigations being one of several (mainly serological) options compared with histology. Another limitation is the lack of follow-up data. As such, we cannot estimate the predictive value of the identified associations or the newly developed classifier. However, our data strongly suggest the potential for a significant predictive value. In analogy to the CKD273 classifier, we aimed to investigate the predictive value of the BMI150 classifier in patients with obesity and CKD, especially in those without diabetes, and in predicting response to intervention.

ACKNOWLEDGEMENTS

AUTHORS' CONTRIBUTIONS

J.B. and H.M. were responsible for the idea, study design, data analysis, manuscript writing and manuscript revision. H.M., T.H., A.L. and J.S. were responsible for proteomic analysis, manuscript revision and statistical analysis. R.W. was responsible for manuscript writing and manuscript revision.

CONFLICT OF INTEREST STATEMENT

H.M. is a co-founder and co-owner of Mosaiques Diagnostics, who developed the CE-MS technique. T.H., J.S. and A.L. are employed by Mosaiques Diagnostics. All other authors have no conflicts of interest to disclose.

REFERENCES

1. Kovesdy CP, Furth SL, Zoccali C, on behalf of the World Kidney Day Steering Committee. Obesity and kidney disease. *Can J Kidney Health Dis* 2017; 4: 205435811769866
2. Stenvinkel P, Zoccali C, Ikizler AT. Obesity in CKD—what should nephrologists know? *J Am Soc Nephrol* 2013; 24: 1727–1736
3. Bilha SC, Nistor I, Nedelcu A *et al*. The effects of bariatric surgery on renal outcomes: a systematic review and meta-analysis. *Obes Surg* 2018; 28: 3815–3833
4. Chang AR, Chen Y, Still C *et al*. Bariatric surgery is associated with improvement in kidney outcomes. *Kidney Int* 2016; 90: 164–171
5. Whaley-Connell A, Sowers JR. Obesity and kidney disease: from population to basic science and the search for new therapeutic targets. *Kidney Int* 2017; 92: 313–323
6. Tsuboi N, Okabayashi Y, Shimizu A *et al*. The renal pathology of obesity. *Kidney Int Rep* 2017; 2: 251–260
7. D'Agati VD, Kaskel FJ, Falk RJ. Focal segmental glomerulosclerosis. *N Engl J Med* 2011; 365: 2398–2411
8. Praga M, Morales E. The fatty kidney: obesity and renal disease. *Nephron* 2017; 136: 273–276
9. Briffa JF, McAinch AJ, Poronnik P *et al*. Adipokines as a link between obesity and chronic kidney disease. *Am J Physiol Renal* 2013; 305: F1629–F1636
10. Rüster C, Wolf G. Adipokines promote chronic kidney disease. *Nephrol Dial Transplant* 2013; 28: iv8–iv14
11. Schanstra JP, Zürbig P, Alkhalaf A *et al*. Diagnosis and prediction of CKD progression by assessment of urinary peptides. *J Am Soc Nephrol* 2015; 26: 1999–2010
12. Mischak H, Vlahou A, Ioannidis J. Technical aspects and inter-laboratory variability in native peptide profiling: the CE–MS experience. *Clin Biochem* 2013; 46: 432–443
13. Latosinska A, Siwy J, Mischak H *et al*. Peptidomics and proteomics based on CE-MS as a robust tool in clinical application: the past, the present, and the future. *Electrophoresis* 2019; 40: 2294–2308
14. Levey AS, Stevens LA, Schmid CH *et al*. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; 150: 604–612
15. Johnson TS, Haylor JL, Thomas GL *et al*. Matrix metalloproteinases and their inhibitions in experimental renal scarring. *Nephron Exp Nephrol* 2002; 10: 182–195
16. Sharma AK, Mauer MS, Kim Y *et al*. Interstitial fibrosis in obstructive nephropathy. *Kidney Int* 1993; 44: 774–788
17. Stokes MB, Holler S, Cui Y *et al*. Expression of decorin, biglycan, and collagen type I in human renal fibrosing disease. *Kidney Int* 2000; 57: 487–498
18. Eddy A. Molecular insights into renal interstitial fibrosis. *J Am Soc Nephrol* 1996; 7: 2495–2508
19. Mason RM, Wahab N. Extracellular matrix metabolism in diabetic nephropathy. *J Am Soc Nephrol* 2003; 14: 1358–1373
20. Rossing K, Mischak H, Dakna M *et al*. Urinary proteomics in diabetes and CKD. *J Am Soc Nephrol* 2008; 19: 1283–1290
21. Magalhães P, Pejchinovski M, Markoska K *et al*. Association of kidney fibrosis with urinary peptides: a path towards non-invasive liquid biopsies? *Sci Rep* 2017; 7: 16915
22. Sharma K. Obesity, oxidative stress, and fibrosis in chronic kidney disease. *Kidney Int Suppl* 2014; 4: 113–117
23. Declèves A-E, Zolkipli Z, Satriano J *et al*. Regulation of lipid accumulation by AMK-activated kinase in high fat diet–induced kidney injury. *Kidney Int* 2014; 85: 611–623

24. Dugan LL, You Y-H, Ali SS *et al*. AMPK dysregulation promotes diabetes-related reduction of superoxide and mitochondrial function. *J Clin Invest* 2013; 123: 4888–4899

25. Declèves A-E, Mathew AV, Cunard R *et al*. AMPK mediates the initiation of kidney disease induced by a high-fat diet. *J Am Soc Nephrol* 2011; 22: 1846–1855

26. Vlassara H, Uribarri J. Advanced glycation end products (AGE) and diabetes: cause, effect, or both? *Curr Diab Rep* 2014; 14: 453

27. Vlassara H, Striker GE. AGE restriction in diabetes mellitus: a paradigm shift. *Nat Rev Endocrinol* 2011; 7: 526–539

28. Sajithlal G, Chithra P, Chandrakasan G. Effect of curcumin on the advanced glycation and cross-linking of collagen in diabetic rats. *Biochem Pharmacol* 1998; 56: 1607–1614

29. Bai P, Phua K, Hardt T *et al*. Glycation alters collagen fibril organization. *Connect Tissue Res* 1992; 28: 1–12

30. Sajithlal GB, Chithra P, Chandrakasan G. Advanced glycation end products induce crosslinking of collagen in vitro. *Biochim Biophys Acta* 1998; 1407: 215–224

31. Sun K, Tordjman J, Clément K *et al*. Fibrosis and adipose tissue dysfunction. *Cell Metab* 2013; 18: 470–477

32. Liu Y, Aron-Wisnewsky J, Marcelin G *et al*. Accumulation and changes in composition of collagens in subcutaneous adipose tissue after bariatric surgery. *J Clin Endocrinol Metab* 2016; 101: 293–304

33. Genovese F, Manresa AA, Leeming D *et al*. The extracellular matrix in the kidney: a source of novel non-invasive biomarkers of kidney fibrosis? *Fibrogenesis Tissue Repair* 2014; 7: 4

34. D'Amato F, Noli B, Angioni L *et al*. VGF peptide profiles in type 2 diabetic patients' plasma and in obese mice. *PLoS One* 2015; 10: e0142333

35. Øvrehus MA, Zürbig P, Vikse BE *et al*. Urinary proteomics in chronic kidney disease: diagnosis and risk of progression beyond albuminuria. *Clin Proteom* 2015; 12: 21

36. Matboli M, Eissa S, Ibrahim D *et al*. Caffeic acid attenuates diabetic kidney disease via modulation of autophagy in a high-fat diet/streptozotocin- induced diabetic rat. *Sci Rep* 2017; 7: 2263

37. Xu Q, Mariman ECM, Roumans NJT *et al*. Adipose tissue autophagy related gene expression is associated with glucometabolic status in human obesity. *Adipocyte* 2018; 7: 12–19

38. Bera TK, Liu X-F, Yamada M *et al*. A model for obesity and gigantism due to disruption of the *Ankrd26* gene. *Proc Natl Acad Sci USA* 2008; 105: 270–275