

Published in final edited form as:

Eur Heart J. 2012 September ; 33(18): 2342–2350. doi:10.1093/eurheartj/ehs185.

Urinary proteome analysis in hypertensive patients with left ventricular diastolic dysfunction

Tatiana Kuznetsova¹, Harald Mischak^{2,3}, William Mullen³, and Jan A. Staessen^{1,*}

¹Studies Coordinating Centre, Division of Hypertension and Cardiovascular Rehabilitation, Department of Cardiovascular Sciences, University of Leuven, Campus Sint Rafaël, Kapucijnenvoer 35, block d, level 00, Box 7001, B 3000 Leuven, Belgium

²Mosaiques Diagnostic and Therapeutics AG, Hannover, Germany

³BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, UK

Abstract

Aims—Despite the significant heart failure (HF) burden on society, easily applicable screening techniques, particularly for the early detection of asymptomatic left ventricular (LV) dysfunction, are lacking. The present study aimed to identify and test a set of urinary polypeptides that may indicate early LV diastolic dysfunction as defined by echocardiography in hypertensive patients in a cross-sectional case–control study nested within the FLEMish study on ENvironment, Genes and Health Outcome (FLEMENGHO).

Methods and results—To identify potentially discriminating urinary biomarkers for LV diastolic dysfunction, we applied capillary electrophoresis coupled to mass spectrometry. In the discovery set, we compared 19 hypertensive patients with asymptomatic LV diastolic dysfunction with 19 healthy controls. In the absence of adjustment for multiple testing, 85 urinary peptides were different between cases and controls at a *P*-value of <0.033. With adjustment for multiple testing, three potential biomarkers remained significantly different between cases and controls (*P* = 0.02). We combined the 85 potential biomarkers in a high-dimensional model (classifier), which we applied in a blinded manner to an independent test set of 16 hypertensive patients with symptomatic HF and 16 healthy controls. Upon unblinding, the area under the receiver operating characteristic curve (AUC) of the HF classification was 0.84 (95% CI: 0.70–0.98; *P* = 0.001).

Conclusion—In asymptomatic hypertensive patients with LV diastolic dysfunction, we identified a set of urinary polypeptides specific for essential hypertension with LV diastolic dysfunction that subsequently distinguished hypertensive patients with overt HF from healthy controls.

Keywords

Diastolic dysfunction; Urinary proteomics

© The Author 2012.

*Corresponding author. Tel: +32 16 34 7104, Fax: +32 16 34 7106, jan.staessen@med.kuleuven.be.

Supplementary material Supplementary material is available at *European Heart Journal* online.

Conflict of interest: H.M. is the founder and co-owner of Mosaiques Diagnostics, which developed the CE-MS technology for clinical application.

Introduction

Heart failure (HF) is a progressive condition that begins with risk factors for left ventricular (LV) dysfunction (e.g. hypertension), proceeds to asymptomatic changes in cardiac structure (e.g. LV hypertrophy) and function (e.g. impaired relaxation), and then evolves into clinically overt HF, disability and death.¹ The 5-year mortality rate for symptomatic HF is ~60%. Heart failure can clinically manifest with predominantly systolic or diastolic LV dysfunction, or both. In a randomly selected population sample, the frequency of asymptomatic echocardiographically diagnosed diastolic LV dysfunction (early stage) is as high as 27%.² This constitutes a large pool of subjects at high risk for diastolic HF. Despite its significant burden on society, the diagnosis of LV diastolic dysfunction and diastolic HF remains difficult.

Easily applicable screening techniques for the early detection of asymptomatic stages of LV dysfunction are lacking. Clinical proteomics aims at defining novel biomarkers for the early diagnosis of the disease, their prognostic assessment, and the selection and monitoring of proper treatment modalities. In this study, we focused on the urinary proteome, because urine can be easily sampled and contains an array of low-molecular-weight proteins and peptides that can be analysed without additional manipulation, such as proteolysis.³ Capillary electrophoresis coupled to mass spectrometry (CE-MS)^{4,5} is now an established method, which allowed identifying peptide markers specific for a variety of diseases, such as diabetes and diabetic nephropathy,⁶ graft rejection,⁷ and coronary artery disease.⁸

One of the core features of cardiac remodelling in hypertensive patients at risk of developing HF and patients with overt HF is related to increased interstitial deposition and cross-linking of type-I collagen. Thus, an impairment of the normal extracellular matrix (ECM) turnover might result in fibrosis of the myocardium, and therefore lead to increased LV stiffness. Capillary electrophoresis coupled to mass spectrometry might be used for the identification and validation of effective screening biomarkers including collagen polypeptides for the risk stratification of patients with an early stage of LV dysfunction. Achieving this goal would meet a pressing clinical need. Thus, the aim of the present study was to identify and validate a set of urinary polypeptides in hypertensive patients that may indicate LV diastolic dysfunction as defined by echocardiography in a cross-sectional case-control study nested within the FLEMish study on ENvironment, Genes and Health Outcome (FLEMENGHO).

Methods

Study population

This study is a sub-analysis of a large-scale family-based study on the genetic epidemiology of cardiovascular phenotypes (FLEMENGHO), for which recruitment started in 1985 and continued through 2008.⁹ At each contact, standardized and validated questionnaires were completed to collect detailed information about each participant's personal and familial medical history, use of medication, smoking habits, intake of alcohol, physical activity, and lifestyle. High-fidelity phenotyping at the examination centre included clinical and anthropometric measurements, echocardiography, blood sampling, and 24-h urine collection. Biological samples were stored in a biobank for later analysis. The Ethics Committee of the University of Leuven approved the FLEMENGHO study. All participants gave their written informed consent.

For urinary biomarker discovery, we selected 19 hypertensive patients with echocardiographically diagnosed *asymptomatic* LV diastolic dysfunction and 19 age-matched healthy controls. The detailed echocardiographic protocol is available in the Supplementary material online, *Methods*. We defined cases as having hypertension (systolic

blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg or the use of antihypertensive drugs) and moderate diastolic dysfunction based on age-specific echocardiographic criteria (Supplementary material online, *Methods*).² Subjects with a previous history of coronary heart disease, valvular heart diseases, renal failure, or diabetes were excluded from the discovery phase.

For biomarker validation, an additional test set of 16 hypertensive patients with mild-to-moderate symptomatic HF (NYHA class II-III) and 16 healthy controls were selected from the FLEMENGHO cohort. We defined cases as having hypertension (systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg or the use of antihypertensive drugs) and moderate/severe LV diastolic dysfunction based on age-specific echocardiographic criteria (Supplementary material online, *Methods*) and symptoms that might be attributable to HF, such as dyspnoea, peripheral oedema, pulmonary congestion, and chest pain. To collect the information on the HF symptoms during the follow-up visit, we administered the standardized London School of Hygiene cardiovascular, dyspnoea, and respiratory questionnaires.¹⁰

Sample preparation and capillary electrophoresis coupled to mass spectrometry analysis

All participants collected 24-urine samples within 1 week of the echocardiographic examinations. Aliquots were stored at -80°C . Urine (0.7 mL) was thawed immediately before analysis and diluted with 0.7 mL of 2 M urea, 10 mM NH_4OH containing 0.02% SDS.¹¹ To remove higher molecular mass proteins, such as albumin and immunoglobulin G, the sample was ultrafiltered using Centriscart ultracentrifugation devices (20 kDa MWCO; Sartorius, Goettingen, Germany) at 3000 g relative centrifugal force until 1.1 mL of filtrate was obtained. This filtrate was then applied onto a PD-10 desalting column (GE Healthcare, Uppsala, Sweden) equilibrated in 0.01% NH_4OH in HPLC-grade H_2O (Roth, Germany) to decrease matrix effects by removing urea, electrolytes, and salts, and to enrich polypeptides. Finally, all samples were lyophilized, stored at 4°C , and suspended in HPLC-grade H_2O shortly before CE-MS analyses.¹²

Capillary electrophoresis coupled to mass spectrometry analyses were performed using a P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton, USA) online coupled to a micrO-TOF MS (Bruker Daltonic, Bremen, Germany).¹² The ESI sprayer (Agilent Technologies, Palo Alto, CA, USA) was grounded, and the ion spray interface potential was set between -4 and -4.5 kV. Data acquisition and mass spectrometry acquisition methods were automatically controlled by the capillary electrophoresis via contact-close relays. Spectra were accumulated every 3 s, over a range of charge states (m/z) 350–3000. Accuracy, precision, selectivity, sensitivity, reproducibility, and stability of the CE-MS measurements have been published elsewhere.¹³ The quality control criteria were as follows:¹⁴ The sample must contain a minimum of 950 features, a minimum mass resolution of 8000, and a minimum migration time interval of 10 min. After calibration, migration time deviation must be <0.35 min.

Data processing

Mass spectra were processed using MosaiquesVisu software, including peak picking, deconvolution, and de-isotoping.¹⁵ Migration time and peak intensity were normalized using internal polypeptide standards.¹⁶ These fragments are believed to be the result of normal biological processes and appear to be unaffected by any disease state studied to date on the basis of 13 000 samples in our database.^{17,18} The resulting peak list characterizes each polypeptide by its molecular mass, normalized capillary electrophoresis migration time, and normalized signal intensity. All detected polypeptides were deposited, matched, and

annotated in a Microsoft SQL database, allowing further analysis and comparison of multiple patient groups.

Statistical methods and identification of biomarkers

We compared means and proportions of clinical and echocardiographic characteristics of the discovery and test samples by means of a t -test and the χ^2 statistics, respectively, using SAS software, version 9.1.3 (SAS Institute, Cary, NC, USA).

In the discovery phase, we compared the natural logarithm-transformed signal amplitude of the CE-MS urinary polypeptide profile between patients and controls using the Wilcoxon rank sum test.¹⁹ This non-parametric test is suitable for skewed proteomic data. We tested the null hypothesis that patients and controls have the same continuous distribution of signal amplitude of the CE-MS urinary polypeptide profile. The signal amplitude represents the calibrated counts (intensity) recorded by the mass spectrometry device. Statistical adjustment for multiple testing was performed by applying Benjamini-Hochberg.²⁰ This correction allows controlling the false discovery rate at a level of 0.05.²⁰ We searched for a cluster of urinary polypeptides discriminating between cases and controls based on the distribution of biomarkers in individual subjects. For each case and each control, the selected polypeptides were combined into a single summary variable, using the support-vector machine-based MosaCluster software, version 1.6.5 (Supplementary material online, *Methods*). In the test set, researchers blinded to the clinical condition of the study participants measured the cluster polypeptides. After breaking the code, we calculated the sensitivity and the specificity based on tabulating the number of *correctly classified* samples in the test set, using receiver operating characteristic (ROC) plots. The area under the ROC curve (AUC) provides a single measure of overall accuracy that is independent of any particular threshold.²¹

Results

Tables 1 and 2 list the clinical and echocardiographic characteristics of cases and controls in the discovery and test sets. Among cases in the discovery set, 8 (42.1%), 10 (52.6%), and 2 (10.5%) were on treatment with diuretics, beta-blockers, or renin system inhibitors (angiotensin-converting enzyme inhibitors or angiotensin receptor blockers). Among the 16 patients in the test set, these numbers were 9 (56.2%), 8 (50.0%), and 7 (43.7%), respectively. All urine samples were successfully analysed with CE-MS and passed all quality control criteria.

Discovery set

The compiled urinary proteomics data from cases and controls in the discovery set are shown in Figure 1A and B. In the absence of adjustment for multiple testing, 85 peptides were found different between cases and controls at a nominal P -value of <0.033 . The distribution of these 85 biomarkers in the 19 cases and 19 controls is shown in Figure 1C and D. To classify urinary polypeptides according to whether they were down- or up-regulated in cases vs. controls, we calculated a ratio (R) of $(\ln \text{signal amplitude} \times \text{frequency/number of participants})$ in controls to $(\ln \text{signal amplitude} \times \text{frequency/number of participants})$ in cases. Supplementary material online, *Table S1* lists further details on these 85 polypeptides ordered by ascending R. The information provided includes polypeptide identifier (SQL number), mass, CE-migration time, percentage of samples, in which the polypeptide could be detected, signal amplitude of the polypeptides, R and P -values for comparisons of the polypeptide signal amplitude distribution between cases and controls. With adjustment for multiple testing applied, three potential biomarkers (IDs 8725, 40737, and 61984) remained significantly different between cases and controls ($P = 0.02$).

Test set

We performed additional CE-MS analysis using an independent test set of 16 hypertensive patients with symptomatic HF and 16 healthy controls in a blinded manner and the data were subjected to the classifier. Upon unblinding, 11 out of 16 HF samples and 15 out of 16 healthy control samples (Figure 2A) were predicted correctly using a classification threshold of the d-score = -0.38 with a sensitivity of 69% (95% CI: 41–89) and a specificity of 94% (95% CI: 70–100). The AUC of the HF classification was 0.84 (95% CI: 0.70–0.98; $P=0.001$; Figure 2B).

Sequencing

Table 3 lists information on 70 of 85 polypeptides from the classifier, in which the sequence data (Supplementary material online, *Methods*) were determined. First, the three polypeptides, which remained statistically significant after adjustment for multiple testing, were collagen alpha-1(V) (ID 8725), WW domain-binding protein 11 (WBP11) (ID 61984), and isoform 1 of collagen alpha-1(XXVI) (ID 40737). All three polypeptides were significantly down-regulated in patients with LV dysfunction compared with the controls (Table 3). Secondly, we noted in patients with LV dysfunction a down-regulation of five other collagen alpha-1(V) fragments and the majority of collagen alpha-1(I) fragments (9 of 14; Table 3). In contrast, 11 of 13 fragments of collagen alpha-1(III) were found to be up-regulated in hypertensive patients with LV diastolic dysfunction (Table 3).

Discussion

The key finding of our study was that by applying the CE-MS methodology to easily obtainable urine samples, it was possible to derive sets of urinary polypeptides specific for LV diastolic dysfunction in hypertensive patients. These pilot findings suggest that the urinary proteome might provide a clinically useful tool for screening and subsequent monitoring of subclinical LV dysfunction in hypertensive patients.

Under physiological conditions, the urinary protein content originates ~70% from the kidney and the urinary tract, while 30% is derived from plasma.²² ~60% of the total mass of urinary peptides and proteins represent fragments of collagen.¹⁷ Indeed, in our current study, most of the marker polypeptides originated from the metabolism of collagen (Table 3).

The collagen in the ECM is maintained by an equilibrium of synthesis and degradation, which is disturbed in chronic HF.²³ The cardiac ECM is predominantly composed of fibrillar collagen type I (85%) and type III (11%). Furthermore, small amounts of collagen types IV and V co-distribute with collagen I. The prime location of collagen V is at the fibril core and is important in nucleating collagen I-containing fibrils.²⁴ Left ventricular diastolic dysfunction and HF associated with hypertension are characterized by increased interstitial deposition and cross-linking of type I collagen. The modified ECM proteins lead to LV stiffness.²³ Animal models of the ageing heart revealed reduced proteolytic activity and degradation of collagen I.²⁵ Moreover, elevated levels of serum markers of collagen synthesis (amino-terminal peptide of procollagen type III, PIIINP) and degradation (carboxyl-terminal telopeptide of collagen type I, C1P) independently predicted mortality and cardiovascular disease in population studies.^{26,27} Thus, an impairment of the normal ECM turnover might result in fibrosis of the myocardium, which is a core feature of cardiac remodelling in patients at risk of developing HF.

Our study showed in hypertensive patients with asymptomatic LV diastolic dysfunction a down-regulation of several fragments of collagens type I and type V as well as an up-regulation of fragments of collagen type III. Along similar lines, Zürgbig *et al.*²⁸ reported a decreased urinary excretion of collagen alpha-1(I) fragments with ageing in humans. Down-

regulation of collagen- α 1(I) and fibrinogen fragments also occurs in chronic kidney disease.^{6,29} However, the later fragments are partially different from those in the current study, indicating differential pathways in collagen breakdown in kidney and heart disease. In addition, the down-regulation of collagen V and the up-regulation of collagen III fragments are not a prominent feature in chronic kidney disease.^{6,29}

We also noted a down-regulation of polypeptide (ID 61984) associated with WBP11 in LV dysfunction hypertensive patients, which remained statistically significant after adjustment for multiple testing. There are two major functions of WBP11: a splicing factor and a modulator (inhibitor) of protein phosphatase-1 (PP-1).³⁰ In cardiomyocytes, PP-1 plays a crucial role for calcium handling and relaxation via dephosphorylation of phospholamban.³¹ The activity of PP-1 is enhanced in HF. Therefore, reduced levels of WBP11, which we observed in patients with LV dysfunction in our study, might result in increased activity of PP-1. In turn, this would delay LV relaxation and results in the progression of LV dysfunction towards symptomatic HF in hypertensive patients.

The present study must be interpreted within the context of its limitations. We combined 85 potential biomarkers in a high-dimensional classifier, since the accuracy and robustness of a high-dimensional model improves with the number of biomarkers.¹⁹ In addition, it is by now accepted that a single biomarker may generally not be able to display a complex disease such as HF.⁴ However, the current study had a small sample size. Only three biomarkers remained significant upon the implementation of rigorous statistical testing. Nevertheless, conforming to the recommendations for clinical proteomics,³² we have tested the classification of independent samples from cases and controls, resulting in an ROC ~0.84. These observations suggest that urinary biomarkers might accurately identify LV dysfunction in hypertensive patients, once replicated in a substantially large population. Another limitation of the study is that the identified set of polypeptides might be specific only for essential hypertension with LV diastolic dysfunction. Further investigation is required to validate this set of polypeptides in patients with HF caused by other pathologies, for example in patients with aortic valve stenosis. On the other hand, hypertension is the major risk factor for the development of HF and contributes a large proportion of HF cases in the general population. In the Framingham Heart Study,³³ the population-attributable risk for congestive HF associated with hypertension was 39% in men and 59% in women.

Future research should expand those preliminary findings by applying urinary proteome analysis to test the prognostic value of the derived set of urinary polypeptides against well-established clinical and laboratory prognostic indices in a large cohort. The long-term goal should be to develop the urinary proteomics approach into a clinically useful tool for the early detection of subclinical LV dysfunction and to monitor treatment. Insight in the urinary proteome will also provide clues to the pathogenetic mechanisms driving LV dysfunction and possibly identify targets for treatment.

In conclusion, from the discovery set in asymptomatic hypertensive patients, we derived a panel of urinary polypeptides that may be specific for essential hypertension with LV diastolic dysfunction. Furthermore, this set distinguished hypertensive patients with clinically overt HF from healthy controls in the test (validation) set. Thus, the urinary proteome might be useful as a tool for screening and subsequent monitoring of subclinical LV dysfunction in hypertensive patients.

Acknowledgments

Funding The European Union (grants IC15-CT98-0329-EPOGH, LSHM-CT-2006-037093, HEALTH-F4-2007-201550, HEALTH-2011-278249-EU-MASCARA, and ERC Advanced Grant-2011-294713-EPLORE)

and the Fonds voor Wetenschappelijk Onderzoek Vlaanderen, Ministry of the Flemish Community, Brussels, Belgium (grants G.0575.06 and G.0734.09) supported the Studies Coordinating Centre (Leuven, Belgium).

References

1. Hunt SA, Abraham WT, Chin MH, et al. ACC/AHA 2005 guideline update for the diagnosis and management of chronic heart failure in the adult. A report of the American College of Cardiology/American Heart Association Task Force on practical guidelines. *Circulation*. 2005; 112:e154–e235. [PubMed: 16160202]
2. Kuznetsova T, Herbots L, López B, Jin Y, Richart T, Thijs L, González A, Herregods MC, Fagard RH, Díez J, Staessen JA. Prevalence of left ventricular diastolic dysfunction in a general population. *Circ Heart Fail*. 2009; 2:105–112. [PubMed: 19808325]
3. Mischak H, Julian BA, Novak J. High-resolution proteome/peptidome analysis of peptides and low-molecular-weight proteins in urine. *Proteomics Clin Appl*. 2007; 1:792–804. [PubMed: 20107618]
4. Kaiser T, Wittke S, Just I, Krebs R, Bartel S, Fliser D, Mischak H, Weissinger EM. Capillary electrophoresis coupled to mass spectrometer for automated and robust polypeptide determination in body fluids for clinical use. *Electrophoresis*. 2004; 25:2044–2055. [PubMed: 15237405]
5. Schiffer E, Mischak H, Novak J. High resolution proteome/peptidome analysis of body fluids by capillary electrophoresis coupled with MS. *Proteomics*. 2006; 6:5615–5627. [PubMed: 16991199]
6. Rossing K, Mischak H, Dakna M, Zürbig P, Novak J, Julian BA, Good DM, Coon JJ, Tarnow L, Rossing P, PREDICTIONS Network. Urinary proteomics in diabetes and CKD. *J Am Soc Nephrol*. 2008; 19:1283–1290. [PubMed: 18448586]
7. Weissinger EM, Schiffer E, Hertenstein B, Ferrara JL, Holler E, Stadler M, Kolb HJ, Zander A, Zürbig P, Kellmann M, Ganser A. Proteomic patterns predict acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. *Blood*. 2007; 109:5511–5519. [PubMed: 17339419]
8. Zimmerli LU, Schiffer E, Zürbig P, Good DM, Kellmann M, Mouis L, Pitt AR, Coon JJ, Schmieder RE, Peter KH, Mischak H, Kolch W, Delles C, Dominiczak AF. Urinary proteomic biomarkers in coronary artery disease. *Mol Cell Proteomics*. 2008; 7:290–298. [PubMed: 17951555]
9. Li Y, Zagato L, Kuznetsova T, Tripodi G, Zerbini G, Richart T, Thijs L, Manunta P, Wang JG, Bianchi G, Staessen JA. Angiotensin-converting enzyme I/D and alpha-adducin Gly460Trp polymorphisms: from angiotensin-converting enzyme activity to cardiovascular outcome. *Hypertension*. 2007; 49:1291–1298. [PubMed: 17452507]
10. Rose, GA.; Blackburn, H. *Cardiovascular Survey Methods*. WHO; Geneva: 1968. p. 162-168.
11. Theodorescu D, Fliser D, Wittke S, Mischak H, Krebs R, Walden M, Ross M, Eltze E, Bettendorf O, Wulfing C, Semjonow A. Pilot study of capillary electrophoresis coupled to mass spectrometry as a tool to define potential prostate cancer biomarkers in urine. *Electrophoresis*. 2005; 26:2797–2808. [PubMed: 15981297]
12. Theodorescu D, Wittke S, Ross MM, Walden M, Conaway M, Just I, Mischak H, Frierson HF. Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. *Lancet Oncol*. 2006; 7:230–240. [PubMed: 16510332]
13. Haubitz M, Good DM, Woywodt A, Haller H, Rupprecht H, Theodorescu D, Dakna M, Coon JJ, Mischak H. Identification and validation of urinary biomarkers for differential diagnosis and evaluation of therapeutic intervention in antineutrophil cytoplasmic antibody-associated vasculitis. *Mol Cell Proteomics*. 2009; 8:2296–2307. [PubMed: 19564150]
14. Good DM, Zürbig P, Argilés A, Bauer HW, Behrens G, Coon JJ, Dakna M, Decramer S, Delles C, Dominiczak AF, Ehrlich JH, Eitner F, Fliser D, Frommberger M, Ganser A, Girolami MA, Golovko I, Gwinner W, Haubitz M, Herget-Rosenthal S, Jankowski J, Jahn H, Jerums G, Julian BA, Kellmann M, Kliem V, Kolch W, Krolewski AS, Luppi M, Massy Z, Melter M, Neusüss C, Novak J, Peter K, Rossing K, Rupprecht H, Schanstra JP, Schiffer E, Stolzenburg JU, Tarnow L, Theodorescu D, Thongboonkerd V, Vanholder R, Weissinger EM, Mischak H, Schmitt-Kopplin P. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Mol Cell Proteomics*. 2010; 9:2424–2437. [PubMed: 20616184]
15. Neuhoff N, Kaiser T, Wittke S, Krebs R, Pitt A, Burchard A, Sundmacher A, Schlegelberger B, Kolch W, Mischak H. Mass spectrometry for the detection of differentially expressed proteins: a

- comparison of surface-enhanced laserdesorption/ionization and capillary electrophoresis/mass spectrometry. *Rapid Commun Mass Spectrom*. 2004; 18:149–156. [PubMed: 14745763]
16. Jantos-Siwy J, Schiffer E, Brand K, Schumann G, Rossing K, Delles C, Mischak H, Metzger J. Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. *J Proteome Res*. 2009; 8:268–281. [PubMed: 19012428]
 17. Coon JJ, Zürbig P, Dakna M, Dominiczak AF, Decramer S, Fliser D, Frommberger M, Golovko I, Good DM, Herget-Rosenthal S, Jankowski J, Julian BA, Kellmann M, Kolch W, Massy Z, Novak J, Rossing K, Schanstra JP, Schiffer E, Theodorescu D, Vanholder R, Weissinger EM, Mischak H, Schmitt-Kopplin P. CE-MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics. *Proteomics Clin Appl*. 2008; 2:964. [PubMed: 20130789]
 18. Siwy J, Mullen W, Golovko I, Franke J, Zürbig P. Human urinary peptide database for multiple disease biomarker discovery. *Proteomics Clin Appl*. 2011; 5:367–374. [PubMed: 21591268]
 19. Dakna M, Harris K, Kalousis A, Carpentier S, Kolch W, Schanstra JP, Haubitz M, Vlahou A, Mischak H, Girolami M. Addressing the challenge of defining valid proteomic biomarkers and classifiers. *BMC Bioinformatics*. 2010; 11:594. [PubMed: 21208396]
 20. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc B (Methodological)*. 1995; 57:125–133.
 21. DeLeo, JM. Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainty. *Proceedings of the Second International Symposium on Uncertainty Modeling and Analysis*; IEEE Computer Society Press; 1993. p. 318-325.
 22. Pieper R, Gatlin CL, McGrath AM, Makusky AJ, Mondal M, Seonarain M, Field E, Schatz CR, Estock MA, Ahmed N, Anderson NG, Steiner S. Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. *Proteomics*. 2004; 4:1159–1174. [PubMed: 15048996]
 23. Burlew BS, Weber KT. Cardiac fibrosis as a cause of diastolic dysfunction. *Herz*. 2002; 27:92–98. [PubMed: 12025467]
 24. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol*. 2008; 20:495–501. [PubMed: 18640274]
 25. Annoni G, Luvarà G, Arosio B, Gagliano N, Fiordaliso F, Santambrogio D, Jeremic G, Mircoli L, Latini R, Vergani C, Masson S. Age-dependent expression of fibrosis-related genes and collagen deposition in the rat myocardium. *Mech Ageing Dev*. 1998; 101:57–72. [PubMed: 9593313]
 26. Velagaleti RS, Gona P, Sundström J, Larson MG, Siwik D, Colucci WS, Benjamin EJ, Vasan RS. Relations of biomarkers of extracellular matrix remodeling to incident cardiovascular events and mortality. *Arterioscler Thromb Vasc Biol*. 2010; 30:2283–2288. [PubMed: 20798380]
 27. Barasch E, Gottdiener JS, Aurigemma G, Kitzman DW, Han J, Kop WJ, Tracy RP. The relationship between serum markers of collagen turnover and cardiovascular outcome in the elderly: the cardiovascular health study. *Circ Heart Fail*. 2011; 4:733–739. [PubMed: 21900186]
 28. Zürbig P, Decramer S, Dakna M, Jantos J, Good DM, Coon JJ, Bandin F, Mischak H, Bascands JL, Schanstra JP. The human urinary proteome reveals high similarity between kidney aging and chronic kidney disease. *Proteomics*. 2009; 9:2108–2117. [PubMed: 19296547]
 29. Alkhalaf A, Zürbig P, Bakker SJ, Bilo HJ, Cerna M, Fischer C, Fuchs S, Janssen B, Medek K, Mischak H, Roob JM, Rossing K, Rossing P, Rychlík I, Sourij H, Tiran B, Winklhofer-Roob BM, Navis GJ, PREDICTIONS Group. Multicentric validation of proteomic biomarkers in urine specific for diabetic nephropathy. *PLoS One*. 2010; 5:e13421. [PubMed: 20975990]
 30. Llorian M, Beullens M, Andrés I, Ortiz JM, Bollen M. SIPP1, a novel pre-mRNA splicing factor and interactor of protein phosphatase-1. *Biochem J*. 2004; 378:229–238. [PubMed: 14640981]
 31. Neumann J. Altered phosphatase activity in heart failure, influence on Ca²⁺ movement. *Basic Res Cardiol*. 2002; 97(Suppl 1):I91–I95. [PubMed: 12479241]
 32. Mischak H, Allmaier G, Apweiler R, Attwood T, Baumann M, Benigni A, Bennett SE, Bischoff R, Bongcam-Rudloff E, Capasso G, Coon JJ, D’Haese P, Dominiczak AF, Dakna M, Dihazi H, Ehrlich JH, Fernandez-Llama P, Fliser D, Frokiaer J, Garin J, Girolami M, Hancock WS, Haubitz

- M, Hochstrasser D, Holman RR, Ioannidis JP, Jankowski J, Julian BA, Klein JB, Kolch W, Luider T, Massy Z, Mattes WB, Molina F, Monsarrat B, Novak J, Peter K, Rossing P, Sánchez-Carbayo M, Schanstra JP, Semmes OJ, Spasovski G, Theodorescu D, Thongboonkerd V, Vanholder R, Veenstra TD, Weissinger E, Yamamoto T, Vlahou A. Recommendations for biomarker identification and qualification in clinical proteomics. *Sci Transl Med.* 2010; 2:46.
33. Levy D, Larson MG, Vasan RS, Kannel WB, Ho KK. The progression from hypertension to congestive heart failure. *JAMA.* 1996; 275:1557–1562. [PubMed: 8622246]

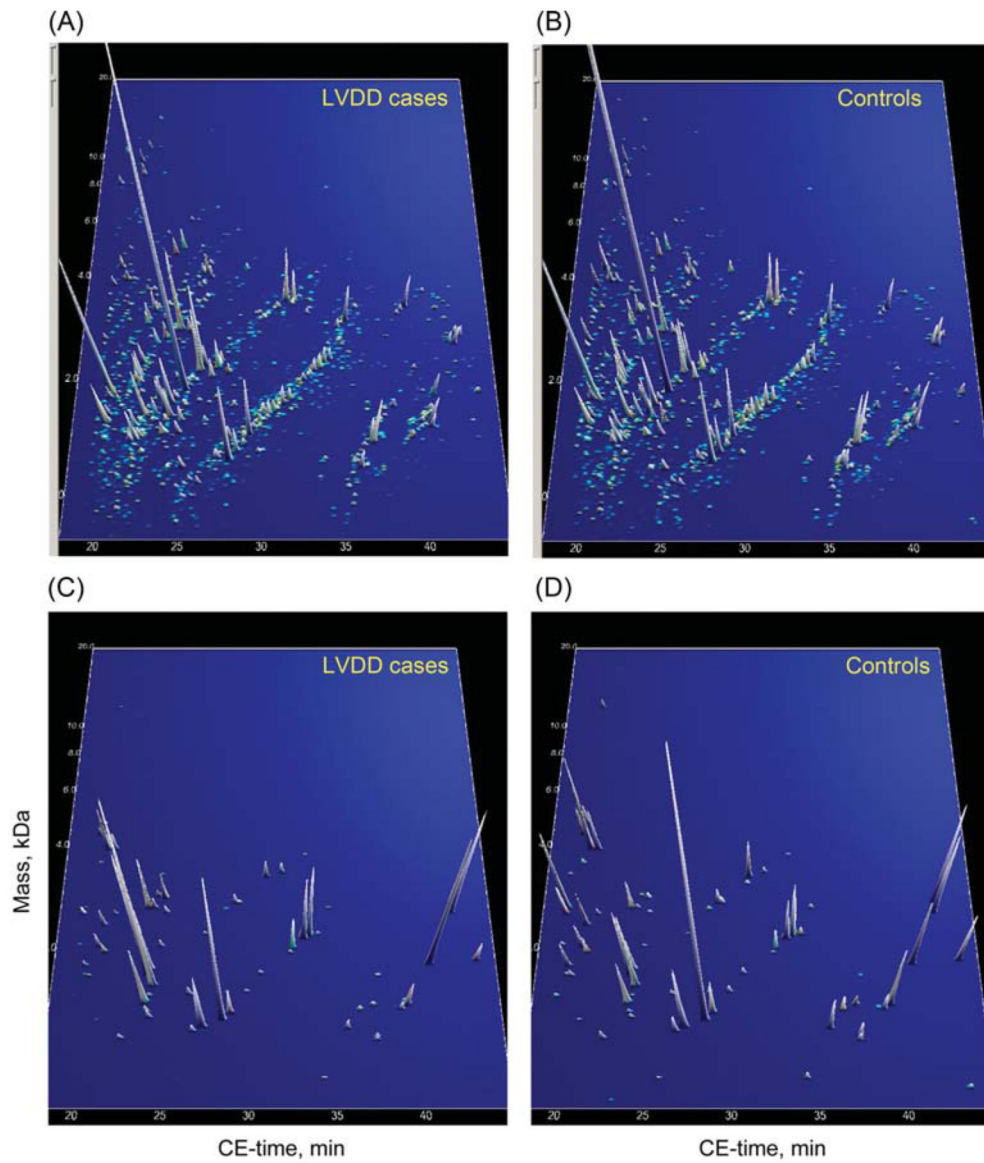


Figure 1. (A and B) Compiled polypeptide patterns of the 19 hypertensive patients with LV diastolic dysfunction and 19 controls examined in the validation set. The molecular mass (0.8–25 kDa, on a logarithmic scale) is plotted against normalized CE migration time (18–45 min). The signal intensity is encoded by the peak height and colour. (C and D) Distribution of the 85 selected polypeptides for LV diastolic dysfunction in the same patients and controls. The signal intensity is magnified five-fold in comparison with (A) and (B).

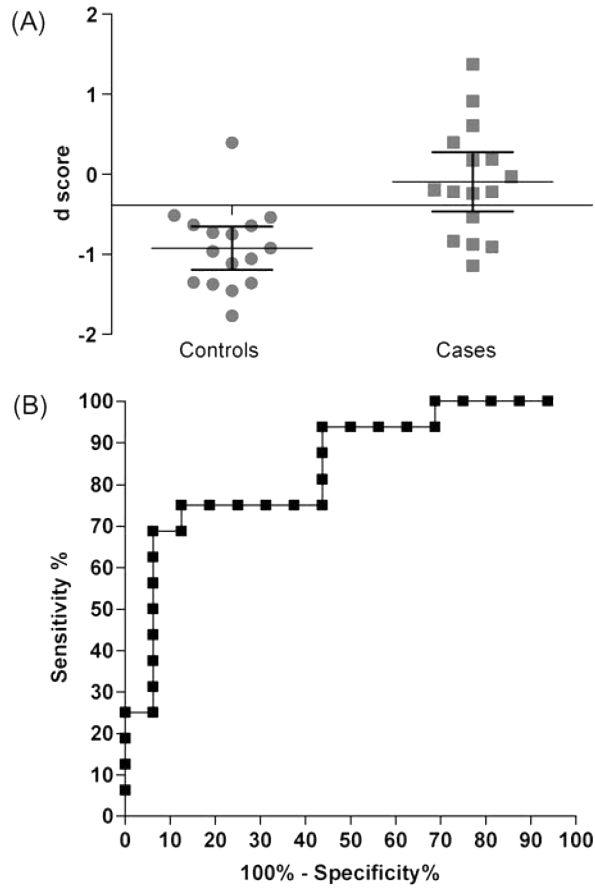


Figure 2.

(A) Mean \pm 95% CI of the d-score used for the classification of the test set. (B) Receiver operator curve (ROC) analysis for urinary biomarker detection of LV dysfunction. The observed area under the curve for left ventricular dysfunction using a panel of 85 biomarkers = 0.84 (95% CI: 0.70–0.98; $P=0.0010$).

Table 1
Clinical characteristics of participants

Characteristic	Discovery set			Test set		
	Control (<i>n</i> = 19)	Case (<i>n</i> = 19)	<i>P</i> -value	Control (<i>n</i> = 16)	Case (<i>n</i> = 16)	<i>P</i> -value
Age, years	63.2 ± 7.5	67.0 ± 7.9	0.13	62.9 ± 4.2	74.5 ± 8.0	< 0.0001
Women, <i>n</i> (%)	9 (47.4)	11 (57.9)	0.52	7 (43.7)	9 (56.3)	0.48
Body mass index, kg/m ²	24.9 ± 2.4	29.5 ± 4.1	0.002	24.7 ± 2.8	31.8 ± 7.1	0.001
Systolic pressure, mmHg	120.5 ± 8.4	151.6 ± 12.0	<0.0001	123.5 ± 9.7	146.7 ± 21.6	0.001
Diastolic pressure, mmHg	74.9 ± 5.3	82.0 ± 11.5	0.02	74.3 ± 4.7	79.7 ± 9.9	0.06
Heart rate, b.p.m.	62.1 ± 5.9	60.7 ± 9.0	0.84	62.9 ± 10.2	69.1 ± 14.1	0.17
Questionnaire data						
Current smoking, <i>n</i> (%)	6 (31.6)	4 (21.0)	0.47	5 (31.2)	1 (6.2)	0.07
Drinking alcohol, <i>n</i> (%)	9 (47.4)	5 (26.3)	0.18	8 (50.0)	6 (37.5)	0.49
Hypertensive, <i>n</i> (%)	0	19 (100)	—	0	16 (100)	—
Treated for hypertension, <i>n</i> (%)	0	13 (68.4)	—	0	14 (87.5)	—
NYHA class II-III, <i>n</i> (%)	0	0	—	0	16 (100)	—
Coronary heart disease, <i>n</i> (%)	0	0	—	0	7 (43.7)	—
Diabetes, <i>n</i> (%)	0	0	—	0	3 (18.7)	—
Biochemical data						
Serum creatinine, μmol/L	84.5 ± 10.4	82.7 ± 10.2	0.63	87.6 ± 10.9	104.1 ± 32.0	0.07
Glomerular filtration rate, mL/min	74.6 ± 9.7	73.4 ± 9.1	0.70	72.6 ± 11.6	60.9 ± 20.7	0.07
Blood glucose, mmol/L	4.86 ± 0.33	4.81 ± 0.55	0.72	4.94 ± 0.47	5.85 ± 1.17	0.06

Values are mean (± SD), or number of subjects (%). NYHA, The New York Heart Association classification.

Table 2
Echocardiographic characteristics of participants

Characteristic	Discovery set			Test set		
	Control (<i>n</i> = 19)	Case (<i>n</i> = 19)	<i>P</i> -value	Control (<i>n</i> = 16)	Case (<i>n</i> = 16)	<i>P</i> -value
Conventional echocardiography						
Left atrium diameter, cm	3.84 ± 0.44	4.34 ± 0.48	0.002	3.96 ± 0.51	4.70 ± 0.73	0.003
LV internal diameter, cm	4.83 ± 0.41	4.96 ± 0.50	0.43	5.06 ± 0.57	5.38 ± 0.49	0.13
Interventricular septum, cm	0.96 ± 0.12	1.17 ± 0.20	0.001	0.99 ± 0.16	1.12 ± 0.15	0.03
Posterior wall, cm	0.89 ± 0.10	1.05 ± 0.16	0.001	0.91 ± 0.15	1.03 ± 0.09	0.01
LV mass index, g/m ²	87.7 ± 12.7	114.6 ± 26.4	0.0009	95.8 ± 21.9	120.6 ± 22.8	0.008
LV hypertrophy, <i>n</i> (%)	0	7 (41.2)	—	0	8 (50.0)	—
Ejection fraction, %	71.8 ± 5.7	71.7 ± 9.5	0.96	68.3 ± 6.1	55.7 ± 11.5	0.003
Ejection fraction <50%, <i>n</i> (%)	0	0	—	0	7 (43.8)	—
Transmitral Doppler data						
E peak, cm/s	71.6 ± 12.6	73.3 ± 15.6	0.73	69.2 ± 8.4	65.9 ± 24.9	0.62
A peak, cm/s	67.2 ± 11.3	87.3 ± 17.7	0.0003	69.0 ± 10.0	79.9 ± 23.8	0.10
<i>E/A</i> ratio	1.09 ± 0.25	0.88 ± 0.27	0.02	1.02 ± 0.19	1.02 ± 0.87	0.99
Tissue Doppler velocities						
<i>E'</i> peak, cm/s	10.5 ± 2.11	6.77 ± 1.72	< 0.0001	10.5 ± 1.79	5.35 ± 1.60	< 0.0001
<i>A'</i> peak, cm/s	11.3 ± 2.03	10.8 ± 2.01	0.46	11.3 ± 1.71	8.87 ± 2.79	0.01
<i>E'/A'</i> ratio	0.98 ± 0.33	0.66 ± 0.25	0.002	0.96 ± 0.25	0.72 ± 0.51	0.14
<i>EE'</i> ratio	6.97 ± 1.27	11.1 ± 1.73	< 0.0001	6.66 ± 0.72	13.2 ± 4.63	< 0.0001
Diastolic dysfunction, <i>n</i> (%)	0	19 (100)	—	0	16 (100)	—

Values are mean (± SD), or number of subjects (%).

Table 3
List of polypeptides with available information on sequencing

ID	Sequence of polypeptide	Protein name	Cases		Controls		R
			%	MA	%	MA	
8725	GDAGSKGpmV	Collagen alpha-1(V)	0.05	1.94	0.63	2.28	0.067
123106	RDVEEEEEEEGLEEDAELLTELQEVLG	Coiled-coil and C2 domain-containing protein 1B	0.05	1.98	0.47	2.63	0.080
1577	KGDTGPPpGP	Collagen alpha-1(III)	0.05	1.65	0.47	1.85	0.095
103493	DEAGSEADHEGTHSTKRGHAKSRPV	Fibrinogen alpha	0.05	3.36	0.47	3.29	0.109
44146	DDFDAHKALEDDE	Isoform 1 of Histone-lysine N-methyltransferase MLL2	0.11	1.91	0.58	2.49	0.146
4845	GGSGAmGSmD	Immunoglobulin-like and fibronectin type III domain-containing protein 1	0.16	1.55	0.63	2.44	0.161
37610	GPpGpPGSpGEQGPSG	Collagen alpha-1(I)	0.11	1.73	0.53	1.87	0.192
83441	GAVGEKGEPGEAGEpGLpGEGGPpG	Collagen alpha-1(V)	0.11	3.45	0.53	3.56	0.201
74703	KSSSHQDSSRmSSVGDYNT	Bone morphogenetic protein 5	0.11	2.64	0.53	2.7	0.203
101157	GPpGADGQpGAKGEpGDAGAKGDAGpPGPA	Collagen alpha-1(I)	0.11	1.97	0.53	1.98	0.207
103022	FNINNLDNNWLKMHFWFYFA	Dermatan-sulfate epimerase-like protein	0.16	2.52	0.68	2.56	0.232
46091	KGETGDVGMGppGPP	Collagen alpha-1(V)	0.16	2.08	0.53	2.24	0.280
32022	TYFPFDLSHG	Haemoglobin subunit alpha	0.21	1.99	0.58	2.21	0.326
82708	GRTGDAGPVPpGPPpGppGpPGPPS	Collagen alpha-1(I)	0.32	2.57	0.84	2.68	0.365
98089	DEAGSEADHEGTHSTKRGHAKSRP	Fibrinogen alpha	0.32	2.97	0.84	3	0.377
61984	DQDKHDDSTDDSDTDK	WW domain-binding protein 11	0.53	2.64	1	3.12	0.448
46369	GPpGEKGGQPPGpQQp	Collagen alpha-1(V)	0.32	2.78	0.68	2.84	0.461
143947	DQGPVGRTEVGA VGPpGFAGEKGPSGEAGTAGPpGTpGPQG	Collagen alpha-2(I)	0.37	2.26	0.79	2.24	0.472
39275	DGVGQpGLPpPGPpG	Collagen alpha-1(XVIII)	0.47	2.59	0.79	2.96	0.520
56493	KGDEGEAGDPGDDNNDI	Collagen alpha-1(VI)	0.47	2.56	0.79	2.74	0.556
41972	EQGLpGAAGQDGPpGP	Isoform D pre-pro-protein of collagen alpha-1 (XI)	0.53	2.75	0.84	2.95	0.588
24168	GPpGPPGSSNQG	Collagen alpha-6(IV)	0.58	2.8	0.84	3.26	0.593
107858	VSESSIHIIGVSLGAHVGGmVGLFQQ	Isoform 2 of phospholipase A1 member A	0.63	2.36	0.89	2.69	0.621
23356	GPpGPPGSSNQG	Collagen alpha-6(IV)	0.58	2.63	0.84	2.9	0.626
97599	LGSHSQDEEDEDTEYFDAMEDS	101 kDa protein	0.58	2.59	0.89	2.66	0.634
23697	DDGEAGKpGRpG	Collagen alpha-1(I)	0.68	2.8	1	2.88	0.661
36566	EEEDSSDSSDSE	Isoform 1 of AP-3 complex subunit beta-1	0.58	2.77	0.74	3.27	0.664
26670	GQDGRpGPpGPpG	Collagen alpha-1(I)	0.63	3.02	0.84	3.3	0.686
58050	GPpGEAGKpGEQGVPGDLG	Collagen alpha-1(I)	0.63	2.57	0.84	2.79	0.691
28005	TYFPFDLSHG	Haemoglobin subunit alpha	0.68	3.08	0.84	3.4	0.733
69979	KGSpGSDGpKGEKGDGPpEGP	Isoform 2C2A' of collagen alpha-2(VI) chain	0.79	2.86	0.95	3.17	0.750

ID	Sequence of polypeptide	Protein name	Cases		Controls		R
			%	MA	%	MA	
40737	GPpGPAGNpGpSpNSP	Isoform 1 of collagen alpha-1(XXVI)	0.84	3.33	1	3.68	0.760
65368	WIDAPDDVIFYMATEET	Metastasis-associated protein MTA1 79 kDa protein	0.79	3.17	0.89	3.61	0.779
73434	ADGSDDLA VSHGSmDSGHGTH	C-myc promoter-binding protein isoform 1	0.84	3.1	1	3.28	0.794
108574	DmGPpGPQGPpGKDGPPGVKGENGHPGSP	Isoform 2 of collagen alpha-1 (XIII)	0.79	3.56	0.89	3.85	0.821
90344	GKNGDDGEAGKpGRpGERGPpGPQ	Collagen alpha-1(I)	0.89	3.12	0.95	3.46	0.845
36759	PpGPpGFPGDpGPpG	Collagen alpha-3(V)	0.89	2.94	0.95	3.18	0.866
28561	SpGPDGKTGPpGPA	Collagen alpha-1(I)	0.89	3.36	0.89	3.79	0.886
107460	KNGETGPQPPGPTGPGGDKGDTGPpGpQG	Collagen alpha-1(III)	0.95	2.91	1	3.11	0.889
32171	ApGDRGEpGPpGPA	Collagen alpha-1(I)	0.95	4.07	1	4.27	0.905
39322	GPpGPpGFPGDpGPpG	Collagen alpha-3(V)	1	3.2	1	3.49	0.917
35339	ApGEDGRpGPpGPQ	Collagen alpha-1(II)	1	3.36	1	3.53	0.952
81196	NGApGNDGAKGDAGApGApGSQGApG	Collagen alpha-1(I)	1	3.72	1	3.59	1.036
41601	DGQPGAKGEpGDAGAK	Collagen alpha-1(I)	1	3.72	1	3.56	1.045
62866	SGpQGppGSEGFTGPPGPQG	Collagen alpha-2(IV)	1	3.89	1	3.71	1.048
99021	QQEQLQQQFQQQEQLQQQ	Zinc finger protein 853	1	3.88	1	3.7	1.049
79136	AGPpGEAGKpGEQGVpGDLGApGP	Collagen alpha-1(I)	1	3.74	1	3.49	1.072
50840	DGApGKNGERGGpGGpGP	Collagen alpha-1(III)	0.95	4.17	0.95	3.86	1.080
72533	PpGEAGKpGEQGVpGDLGApGP	Collagen alpha-1(I)	0.95	3.49	0.95	3.21	1.087
57537	NDGApGKNGERGGpGGpGP	Collagen alpha-1(III)	1	4.02	0.95	3.82	1.108
50212	VGGGEQPPPAPAPRRE	Xylosyltransferase 1	0.89	2.7	0.89	2.43	1.111
60149	GNDGApGKNGERGGpGGpGP	Collagen alpha-1(III)	1	3.72	0.95	3.47	1.128
103198	ERGEAGIpGVpGAKGEDGKDGSpGEpGA	Collagen alpha-1(III)	0.89	2.94	0.89	2.47	1.190
104786	NRGERGSESGHPGQpGppGpPGAPGP	Collagen alpha-1(III)	1	3.58	0.89	3.34	1.204
33135	GAPGPRGRDGEpGT	Isoform 1 of collagen alpha-1(II)	1	2.86	0.89	2.65	1.213
73291	DDKDEEDSPRPRSPGGPD	Zinc finger and BTB domain-containing protein 46	0.84	2.75	0.79	2.37	1.234
45021	RDGEPGTPGNpGPpGP	Isoform 1 of collagen alpha-1(II)	1	2.82	0.89	2.55	1.243
99475	DDILASPPRLPEPQYPGAPHSS	Collagen alpha-1(XVIII)	0.95	2.78	0.89	2.38	1.246
40294	DEPPQSPWDRVK	Apolipoprotein A-I	1	2.85	0.84	2.62	1.295
35424	AMFGPKGFGRGGAE	Cysteine-rich protein 1	0.95	2.79	0.79	2.56	1.311
131294	PGEDGEpGRNGNPGEVGFAGSpGARGFPGAPGLPGL	Collagen alpha-2(V)	1	2.87	0.79	2.71	1.341
111564	ERGEAGIpGVpGAKGEDGKDGSpGEpGANG	Collagen alpha-1(III)	0.89	3.21	0.79	2.67	1.354
104195	NRGERGSESGHPGQPGpGppGApGP	Collagen alpha-1(III)	0.89	2.61	0.74	2.29	1.371
28747	SpGERGETGPpGP	Collagen alpha-1(III)	1	3.44	0.74	3.32	1.400
44802	GGAGEpGKNGAKGEpGp	Isoform 1 of collagen alpha-1(III)	0.79	2.51	0.63	2.1	1.499
113452	NGEAGSAGPpGppGLRGSpGSRGLPGADGRAG	Collagen alpha-2(I)	0.89	2.47	0.58	2.29	1.655

ID	Sequence of polypeptide	Protein name	Cases		Controls		R
			%	MA	%	MA	
69681	SNGNpGPpGPSGSpGKDGpGP	Collagen alpha-1(III)	0.84	2.43	0.42	2.51	1.936
55516	RSGGGGGGGGGQGSTNYGKS	Isoform 3 of heterogeneous nuclear ribonucleoprotein A/B	0.79	2.54	0.42	2.39	1.999
80360	ISVPGPMGPSGPRGLpGPpGApGP	Collagen alpha-1(I)	0.68	2.74	0.26	2.73	2.625
82784	ADGQpGakGEpGDAGAKGDAGPpGP	Collagen alpha-1(III)	0.63	2.28	0.21	2.31	2.961

ID, polypeptide identifier (SQL number); %, percentage of samples, in which the polypeptide could be detected; MA, mean signal amplitude of the polypeptides. R was calculated as $\frac{(\ln \text{ signal amplitude} \times \text{ frequency/number of participants}) \text{ in controls}}{(\ln \text{ signal amplitude} \times \text{ frequency/number of participants}) \text{ in cases}}$. The polypeptides were ordered by ascending R.