Glomerular proteomic profiling of kidney biopsies with hypertensive nephropathy reveals a signature of disease progression

Mikkelsen, Havard

2023

Mikkelsen , H , Vikse , B E , Eikrem , O , Scherer , A , Finne , K , Osman , T & Marti , H-P 2023 , ' Glomerular proteomic profiling of kidney biopsies with hypertensive nephropathy reveals a signature of disease progression ' , Hypertension Research . https://doi.org/10.1038/s41440-022-01066-0

http://hdl.handle.net/10138/357004
https://doi.org/10.1038/s41440-022-01066-0

unspecified
acceptedVersion

Glomerular proteomic profiling of kidney biopsies with hypertensive nephropathy reveals a signature of disease progression

Authors: Håvard MIKKELSEN[a], Bjørn E. VIKSE[a,b], Oystein EIKREM[a,c], Andreas SCHERER[d,e], Kenneth FINNE[a], Tarig OSMAN[a], and Hans-Peter MARTI[a,c]

Institutions: [a]Department of Clinical Medicine, University of Bergen, Bergen, Norway, [b]Department of Medicine, Haugesund Hospital, Haugesund, Norway, [c]Department of Medicine, Haukeland University Hospital, Bergen, Norway, [d]Spheromics, Kontiolahti, Finland, [e]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland,

Running head: Glomerular proteomics of hypertensive nephropathy.

Different parts of the results have been presented in abstract form at following annual meetings: ESH in Milano 2019, ERA-EDTA in Berlin 2021, and ISH in Kyoto 2022.

Sources of support: Open project grants to Hans-Peter Marti (number F-12559) and to Bjørn Egil Vikse (number 911968) from the Western Norway Regional Health Authority.

Potential conflicts of interest: NONE

Correspondence: Hans-Peter Marti, phone: +47 940 30 929, e-Mail: hans-peter.marti@uib.no

Word count (including references, but not tables and legends): 5556

Number of Tables: n=2

Number of Figures: n= 4

Number of supplementary figues: 1

Link to the proteomics data: Data are available via ProteomeXchange with identifier PXD026709.

ABSTRACT

Hypertensive nephropathy (HN) requires a kidney biopsy as diagnostic gold-standard but histological findings are unspecific and specific prognostic markers are missing. We aimed at identifying candidate prognostic markers based on glomerular protein signatures.

We studied adult patients (n=17) with eGFR >30 ml/min/1.73m^2 and proteinuria <3g/d from the Norwegian Kidney Biopsy Registry, including subjects non progressing (NP, n=9), or progressing (P, n=8) to end-stage renal disease (ESRD) within an average follow-up of 22 years. Glomerular cross-sections from archival kidney biopsy sections were microdissected and processed for protein extraction. Proteomic analyses were performed using Q-exactive HF mass spectrometer and relative glomerular protein abundances were compared between P and NP patients. Immunohistochemistry (IHC) was used to validate selected data.

Amongst 1870 quality filtered proteins, 58 were differentially expressed in P and NP patients' glomeruli, with absolute fold changes (FC) ≥1.5, p≤0.05. Supervised classifier analysis (K nearest neighbour) identified a set of five proteins, including Gamma-butyrobetaine dioxygenase (BBOX1, O75936) and Cadherin 16 (CDH16, O75309), overexpressed in P, and Eosinophil peroxidase (EPX, P11678), DnaJ homolog subfamily B member 1 (DNAJB1, P25685) and Alpha-1-syntrophin (SNTA1, Q13424), overexpressed in NP glomeruli, correctly classifying 16/17 kidney biopsy samples. Geneset Enrichment Analysis (GSEA), showed that metabolic pathways were generally enriched in P, and structural cell pathways in NP. Pathway analysis identified Epithelial Adherens Junction Signaling as most affected canonical pathway. IHC analysis confirmed overexpression of BBOX1 and Cadherin 16 in glomeruli from P patients.

In conclusion, glomerular proteomic profiling can be used to discriminate P from NP HN patients.

Key words: Glomeruli, hypertension, hypertensive nephropathy, kidney biopsy, proteomics.

INTRODUCTION

The World Health Organization has identified hypertension as one of the most important health challenges to date. Despite treatment with currently available drugs, hypertension induces kidney tissue damage. Hypertension-associated chronic kidney disease (CKD), also known as hypertensive nephropathy (HN) or arterionephrosclerosis, is caused by long-standing periods of high blood pressure and is characterized by gradual loss of kidney function with various grades of proteinuria over a period of many years (1). HN represents one of the most common causes of end-stage renal disease (ESRD) worldwide and the most frequent, accounting for over 30% of all cases, in Norway (2). CKD, such as HN, is often diagnosed late because it typically remains asymptomatic with only biochemical alterations until severe loss of kidney function occurs. Patients may then progress to ESRD, requiring expensive dialysis and transplantation.

Despite its critical clinical significance, there has been very little advance and innovation in the field of hypertension-related kidney disease. Indeed, HN often represents an unspecific clinical diagnosis applied to non-diabetic CKD patients, presenting with low-level proteinuria and elevated blood pressure.

For over 50 years, clinicians have relied upon serum creatinine and, more lately, albuminuria as the sole biomarkers to monitor kidney health in CKD, including HN. These biomarkers are unspecific, general, CKD markers which do not allow for a specific prognostic evaluation. Kidney biopsy represents the current gold standard for HN diagnosis, but its findings are also unspecific and overlap with other types of CKD, e.g. with diabetic nephropathy. Current therapy guidelines advocate reduction of albuminuria and blood pressure control. However, this strategy, unmodified for decades, has frequently proven unsuccessful (3, 4).

Innovative advances in HN research are urgently needed. In particular, novel more reliable diagnostic and prognostic markers, as well as drugs targeting them to make current medical therapy more effective need to be identified.

Omics-related technologies are increasingly being used to help clarifying defined pathobiological aspects of human diseases. However, a large majority of these studies focus on cancers of diverse histological origin, and CKD are more rarely investigated (5).

In this pilot study we used proteomic technology to explore glomerular alterations detectable in early stage HN, aiming at the identification of protein signatures predicting progress to higher severity stages.. Therefore, we investigated glomerular proteomes of a Norwegian cohort of adult patients with HN, including patients without further severity progression ("Non-progressors", "NP"), and with progression towards ESRD ("Progressors", "P"). Importantly, carefully microdissected glomeruli were used as test samples, and, in particular, glomeruli with moderate sclerosis were excluded from our analysis, to avoid skewing results towards composition of glomerulosclerosis tissues. To the best of our knowledge, this is the first study on glomerular protein biomarkers of HN progression.

MATERIALS AND METHODS

Patients

Adult Caucasian patients (n=17) with biopsy-proven diagnosis of benign HN (angio-nephrosclerosis) from the Norwegian Kidney Biopsy Registry (NKBR) were included in the study based on the following criteria at the time of kidney biopsy: i) age > 40 years, ii) no other primary or systemic kidney disease, such as diabetes mellitus, glomerulonephritis or vasculitis, iii) eGFR > 30 ml/min/1.73 m^2 (CKD-EPI formula), iv) absence of a nephrotic or nephritic syndrome as clinical indication for the biopsy, **and v) diagnosis of treated or untreated arterial hypertension**. Due to easier access to stored biopsy material, we only included patients from Haukeland University Hospital. Importantly however, these patients are fully representative of the total adult Norwegian population receiving this diagnosis at kidney biopsy.

The patient cohort (n=17) was divided into two groups. The main clinical and laboratory characteristics of our patients are reported in Table 1. The first group included patients having progressed to end-stage renal disease (ESRD) within the follow-up period indicated below, which are referred to as progressors (P, n=8). Thereby, subjects with progression to ESRD within 3 years were excluded. We thought that these individuals most likely had an additional different disease entity that caused or at least contributed to a such rapid decrease of kidney function having displayed CKD stage 3 at time of diagnosis. The second group included patients with stable disease as defined by absence of progression to ESRD or death (non-progressors; NP, n=9). The follow-up time for both groups averaged 22 years (range 15-27). Patients were matched by GFR +/- 10ml/ min/1.73 m^2 and age +/- 10 years.

The NKBR receives reports from the treating physicians throughout the entire country when a patient reaches CKD 5 and/or starts renal replacement therapy. This way we could identify the progressors (P) through the Norwegian unique personal identification number, that were already included in the NKBR at the time of the kidney biopsy, and thereafter find matched non progressors. We were able to collect

information from both the patients electronic health records and paper charts from 14 out of 17 subjects throughout the study follow-up time. As far as we have been able to collect the respective data, there was an even distribution between the progressors (P) and non progressors (NP) also regarding smoking and alcohol intake, as well as the expected difference in serum creatine levels over time between the two study groups (data not shown). Furthermore, all patients were uniformly treated at the Division of Nephrology at Haukeland University Hospital according to our national guidelines on hypertension treatment published by the Helsedirektoratet, the Norwegian Ministry of Health (6). All subjects were on 1-3 antihypertensive medications with predominant use of angiotensin-converting enzyme inhibitors (ACEI) and angiotensin receptor blockers (ARB).

Laser capture microdissection and sample preparation

Kidney biopsy cores obtained between 1990 and 1998, that had been stored as formalin-fixed paraffin-embedded (FFPE) tissue were cut by microtome into ten micrometer thick sections. These sections were mounted on polyethylene naphthalate slides (MembraneSlide 1.0 PEN, Carl Zeiss MicroImaging GmbH, Göttingen, Germany), deparaffinized and stained with hematoxylin eosin (HE). Since PEN slides suffer from poor morphology due to lack of cover glass, parallel sections, stained with periodic acid Schiff (PAS) placed on standard glass slides with cover glass, were used to select tissues to be included in the study. Between 31 and 102 (mean, n=79) glomerular cross-sections without severe morphological damage (defined as not visible or only mild glomerulosclerosis on low resolution microscopy) were laser microdissected (PALM MicroBeam, Zeiss) and collected in the cap of specialized tubes (AdhesiveCap 500 clear, Zeiss), as described in a previous study (5).

Proteins were extracted and trypsinized using the filter-aided sample preparation (FASP) protocol, as described in detail elsewhere (7). Briefly, microdissected samples were suspended in 10 µl lysis buffer (0.1M Tris pH 8, 0.1M dithiothreitol [DTT], 4% sodium dodecyl sulfate [SDS]) and heated at 99°C for 60 min. Heating the samples with a strong detergent (SDS) is important as it increases protein yield from formalin-fixed paraffin-embedded samples (8). Next, we performed consecutive washing steps using a 30

kDa filter (Microcon YM-30, Millipore) to remove SDS, before digesting proteins with Trypsin overnight. Eluted digested peptides were de-salted using Oasis HLB µElution plates (Waters, Milford, MA, USA), freeze-dried and stored at -20°C until analysis.

Liquid chromatography and tandem mass spectrometry

Lyophilized peptides were dissolved in 2% ACN/0.1% FA and analyzed on a Q-Exactive HF (Thermo Scientific) mass spectrometer connected to a Dionex Ultimate NCR-3500RS LC system. Samples were trapped on the pre-column (Dionex, Acclaim PepMap 100, 2 cm x 75 µm i.d, 3 µm C18 beads) with loading buffer (0.1% TFA) at a flowrate of 5 µl/min for 5 minutes before separation by reverse phase chromatography (PepMap RSLC, 25cm x 75 µm i.d. EASY-spray column, packed with 2µm C18 beads) at a flow of 200 nL/min. Solvent A and B were 0.1% FA (vol/vol) in water and 100% ACN, respectively. Gradient composition was 5% B during trapping (5min) followed by 5-8 % B during 30 s, before 8–35 % B for the next 134.5min and 35-90 % B over 15 min. Washing and conditioning of the column were performed by 15 minutes treatment with 90 % B and then 20 min conditioning with 5% B. The MS instrument was equipped with an EASY-spray ion source (Thermo Scientific) and was operated in DDA-mode (data-dependent-acquisition). Instrument control was performed through Q-Excative HF Tune 2.4 and Xcalibur 3.0. MS spectra were acquired in the scan range 375-1500 m/z with resolution R = 120,000 at m/z 200, automatic gain control (AGC) target of 3e6 and a maximum injection time (IT) of 100ms.

The 12 most intense eluting peptides above intensity threshold 5e4, and charge states 2 or higher, were sequentially isolated to a target AGC value of 1e5, resolution R = 30,000, IT of 110 ms, and normalized collision energy of 28 %. The isolation window was set to 1.6 m/z with an isolation offset of 0.3 and a dynamic exclusion of 25 seconds. Lock-mass internal calibration was also used.

Label-free quantification

Raw proteomics data from the LC-MS analysis was processed using the freely available MaxQuant software (v1.6.8.0)(9), using recommended settings for label-free quantification (10). Identified features were searched against a human "reference proteome" database (downloaded 2019-06-25) from UniProt.org.

Data sharing

Mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE (11-13) partner repository with the dataset identifier PXD026709.

Statistical analysis

Raw data consisted of peptide information for 2934 proteins. Proteins with at least 2 unique peptides (n=2419) were considered for further analysis. We then excluded those proteins with more than 40% missing values in either P or NP or in both groups, leaving 1870 proteins with 4 or less missing values in the NP group (n=9), and 3 or less missing values in the P group (n=8) in the analysis.

The following steps were performed with JMP Genomics software (www.SAS.com). Raw data were log2 transformed, and missing values imputed per group using multivariate normal imputation with shrinkage estimation of covariance. Resulting data were normalized using quantile normalization. Univariate analysis was calculated by a One-way Analysis of Variance (ANOVA) with shrinking variances using Empirical Bayes, and without further standardization. The intial threshold for significance was a combined criterion of an absolute fold change (FC) ≥ 1.5, and a p-value≤ 0.05.

Pathway analysis was performed by Qiagen Ingenuity Pathway Analysis (IPA), content version 62089861). The IPA regulation z-score algorithm was used to predict the direction of change for a given function (increase or decrease). A positive z-score indicates a activated state, whereas a negative z-score is indicative of inhibition. IPA suggests a z-scores of larger 2 or smaller than -2 as significant.

Gene Set Enrichment Analysis (GSEA), using protein symbols, was run at https://www.gsea-msigdb.org/gsea/index.jsp. The permutation type was "Gene Set", with 1000 permutations, using Gene Ontology Database as knowledge base. Briefly, the gene set enrichment analysis is the enrichment score

(ES), which reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes, while the normalized enrichment score (NES) which is the primary statistic for examining gene set enrichment results, is calculated as the ES of a geneset compared to the mean of the ES of all genesets in the analysis.

KNN classification was performed using the module KNNXValidation within Genepattern (www.genepattern.org). Models with 1 to 10 features in steps of 1 were calculated. Signal to noise ratio (SNR) was used as feature selection method, with 3 neighbors and Euclidean Distance as distance measure. Graphs were generated using JMP Genomics 9.0 and Graphpad Prism 9.12 (www.graphpad.com).

Immunohistochemistry for BBOX1 and Cadherin 16

Tissue sections (3-5 μm thickness) were prepared from FFPE kidney biopsies from patients with stable (n=3) and progressive (n=3) HN from the Norwegian Kidney Biopsy Registry (NKBR). These samples were obtained from the six best-matched patients not included in the proteomics study. These patients were studied due to lack of tissue from the detection cohort. However, importantly, they also provided additional validation in another, unrelated, group of patients.

Sections were deparaffinized in xylene and rehydrated by submerging in diminishing alcohol concentrations. Endogenous peroxidase activity was quenched by incubation with 0.03% hydrogen peroxide for 10 min. Normal goat serum was used at 10% concentration to block potential unspecific binding upon 30 min incubation. Heat induced epitope retrieval was performed in a microwave oven (25 mins, pH 9). Sections were then incubated with either mouse monoclonal anti-BBOX1 (Abcam, ab23441, 1:100) or rabbit polyclonal anti-Cadherin16 (Sigma-Aldrich, HPA007600, 0.5 Ug/Ul) overnight at 4 C.

Positive reactions were visualized using Envision system (Dako) with DAB (3,3'-Diaminobenzidine). Hematoxyline was used for counterstaining before dehydrating the sections in ascending alcohol concentrations and xylene. Non-aqueous mounting medium was used for coverslipping. All washing steps were performed in TBS-T (Tris buffered saline with tween).

Ethics

All patients provided written informed consent. The study was performed in accordance with principles of the Declaration of Helsinki and was approved by the Ethics Review Boards of the Universities of Bergen (REK vest 2013/553).

RESULTS

Differential protein abundance

To identify proteins associated with HN progression, processed abundance data of 1870 quality filtered proteins (see Material and Methods) were subjected to statistical analysis (ANOVA). By applying the above detailed significance criteria, i.e. absolute fold change (aFC) ≥1.5 and p-value ≤0.05, we could select 58 proteins.

Among them, 25 were more abundant in P, and 33 in NP, with fold changes ranging between 5.58 and -3.00. The smallest p-value was 1.57E-4. The 20 proteins with the highest fold changes are listed in **Table 2A.** In particular, Cadherin 16 (CDH16, Uniprot ID O75309) and UDP-glucuronosyltransferase 2B7 (UGT2B7, Uniprot ID P16662) were the top two fold changing proteins (CDH16: FC P/NP 5.58, p-value 6.38E-04; UGT2B7: FC P/NP 4.66, p-value 6.00E-03), overexpressed in P glomeruli, whereas Eosinophil peroxidase (EPX: FC P/NP -3, p-value 1.57E-04) was the top overexpressed protein in NP glomeruli, as shown in **Table 2A.**

Importantly, visualization of a hierarchical cluster analysis of these proteins clearly identified two main clusters, albeit without a perfect separation between progressors (P) and non-progressors (NP) **(Figure 1A).** Indeed, although principal component analysis (PCA), aimed at visualizing sources of variation in a defined dataset, revealed a separation of samples along principal component 1, consistent with prognosis (i.e. P and NP), although a minor overlap was still observed, as shown in **Figure 1B**.

Notably, however, by increasing the FC threshold to ≥2 and p-value ≤0.05, yielding a total of 17 proteins, PCA unraveled a full sample separation (**Figures 1C-D**).

Pathway analysis

The set of 58 proteins with significantly different abundance was further analyzed for pathway enrichment. Pathways associated with cell-cell contact and cell-to-cell signaling were prominently involved

(Table 2B). In particular, three of them, Epithelial Adherens Junction Signaling, Sertoli Cell-Sertoli Cell Junction Signaling, and Tight Junction Signaling showed Benjamini-Hochberg adjusted p-values <0.05. Notably, Epithelial Adherens Junction Signaling was the most affected canonical pathway with five of six member proteins showing lower abundance in P.

Biomarkers for disease progression

To identify proteins of potential use as biomarkers of disease progression, possibly assisting in the early identification of high-risk patients, we used K nearest neighbors algorithm coupled with leave-one-out internal cross-validation.

Initially, the entire dataset of 1870 quality filtered proteins was used to identify markers with the best classifier performance by testing classifiers including 1 to 10 features. Within this context, the best performing model included 8 proteins but wrongly classified 2 samples.

In another approach, the previously identified 58 differentially abundant proteins were used as initial dataset. In this analysis, several classifier models performed equally well. Therefore, we decided to move forward with a model including the smallest possible number of proteins (n=5) as it might be technically more feasible to measure a smaller set of features in future prognostic studies. Notably, the five proteins included in this model were also part of the 8-protein classifier emerging from the entire dataset (see above).

This set of 5 markers correctly classified 16 of 17 samples (sensitivity 88.9%, specificity 100%) and included the following proteins: Gamma-butyrobetaine dioxygenase (BBOX1, O75936) and Cadherin 16 (CDH16, O75309), which were upregulated in P, and Eosinophil peroxidase (EPX, P11678), DnaJ homolog subfamily B member 1 (DNAJB1, P25685) and Alpha-1-syntrophin (SNTA1, Q13424), which were overexpressed in NP glomeruli (**Figure 2).** The only sample which was incorrectly allocated by using this signature was P02, a P sample. Despite all efforts, we were unable to identify clinical or histological background possibly underlying the wrong classification of this sample.

Immunohistochemical validation of key proteins

To validate our findings, we analyzed by immunohistochemistry (IHC) the expression of two classifier proteins in sections from a closely related local HN patient cohort from the NKBR. Interestingly, both Gamma-butyrobetaine dioxygenase (BBOX1, O75936), and Cadherin 16 (CDH16, O75309) proteins displayed similarly high tubular expression in sections from both NP and P patients **(Figure 3).** However, both markers were expressed to clearly higher levels in glomeruli from P, as compared to NP patients.

Geneset Enrichment Analysis (GSEA)

Differential expression of genes encoding the proteins identified by proteomic analysis was investigated by Geneset Enrichment Analysis (GSEA). We observed that the expression of genes encoding protein sets involved in cell metabolism, especially mitochondrial metabolism, was significantly enriched in HN P patients, whereas that of genes encoding proteins involved in cell-cell communication and structure was highly enriched in NP (**Table 2C),**

Epithelial-to-mesenchymal transition (EMT)

Previous studies have suggested a possible relationship between epithelial to mesenchymal transition (EMT) and HN (14). However, the PCA based on a previously reported well-defined EMT proteome (15) only showed a minor ability to separate P from NP samples, as depicted in **Figure S1.** Indeed, expression of FGF1, VIM, FN1, CDH1, CTNNB1, DSP, TJP1, and KRT18 EMT-related genes was also detected in our dataset, but only KRT18 gene expression was significantly upregulated in HN P samples (data not shown). Therefore EMT does not appear to play a major role in early HN stages in our group of patients.

Comparison of protein signatures of progressive HN and IgA nephropathy

IgA nephropathy (IgAN) is frequently associated with hypertension. Therefore we grew curious to test marker expression in glomeruli from benign HN and IgAN. To address this issue, we comparatively

evaluated protein signatures associated with HN progression with those associated with IgA nephropathy (IgAN) from a previously published dataset from our group (5, 16). Thereby, we had investigated adult IgAN patients (mean age 31.5 years) with an eGFR >45 ml/min/1.73m^2 and proteinurtia >1g/d (n=25) from the NKBR divided into a group having progressed to ESRD (n=9) during 10 years after diagnosis and into a stable group without such a disease progression (n=16); (5, 16).

In that study, a group of 100 proteins was found to be differentially abundant in glomeruli from patients with progressing or non progressing IgAN (data not shown). Interestingly, the expression of three proteins from the HN (n=58) and IgAN (n=100) datasets, namely Actin-related protein 2/3 complex subunit 1B (ARPC1B, O15143), Complement receptor type 1 (CR1, P17927), and Cathepsin G (CTSG, P08311), was altered in similar directions with similar fold changes in both databases (**Figure 4A-B).** However, three of the five HN progression classifier proteins, namely CDH16, BBOX1, and DNAJB1) were not differentially expressed in the IgAN dataset, and thereby qualified as HN specific **(Figure 4C).**

DISCUSSION

The aim of this study was the identification of glomerular protein signatures potentially predicting HN progression in the clinical setting. An early identification of a risk for later disease progression could motivate clinicians to adopt a series of preventive measures, including optimal patient care and follow-up with adequate blood pressure treatment. Indeed, 58 proteins met the criteria of differential abundance in glomeruli from patients with stable or progressive HN, as defined by deterioration of kidney function leading to ESRD within 20 years of follow-up. Moreover and most importantly, we were able to define a composite candidate classifier of potential clinical use, consisting of five proteins, including Gamma-butyrobetaine dioxygenase (BBOX1, O75936) and Cadherin 16 (CDH16, O75309), overexpressed in P, and Eosinophil peroxidase (EPX, P11678), DnaJ homolog subfamily B member 1 (DNAJB1, P25685) and Alpha-1-syntrophin (SNTA1, Q13424), overexpressed in NP glomeruli.

BBOX1, up-regulated in glomeruli from HN P, catalyzes defined steps in carnitine metabolism, playing an essential role in the transport of activated fatty acids across the mitochondrial membrane during mitochondrial beta-oxidation (www.genecards.org). Notably, this protein has recently been suggested to be overrepresented in urines from patients with diabetic kidney disease, potentially serving as a marker of diabetic kidney disease (17).

Cadherin 16 is a member of the cadherin family of cell adhesion proteins, produced in tubular epithelial cells in the adult kidney and in the developing genitourinary tract (18). It is dispensable for normal kidney development and its absence only causes a delay in maximal urinary concentrating ability (19). While the pathogenetic relevance of the glomerular upregulation of Cadherin 16 in HN P remains to be determined, its increased glomerular expression, also detectable by IHC, is clearly associated with HN progression in our series.

In contrast, three additional markers appeared to be down-regulated in progressing HN. EPX is a member of the peroxidase family expressed in eosinophils. A previous study showed that eosinophil

aggregates can frequently be observed in diabetic nephropathy, as compared to other glomerulopathies. Unfortunately however, HN was not investigated (20). Interestingly, peroxidase enzymes, such as EPX, have a pro-fibrotic capacity and are associated with fibrosis in several organs (21). Accordingly, EPX was recently shown to contribute to renal fibrosis development in the murine unilateral ureteral obstruction model (22). Instead, in our study, EPX is underrepresented in the patient group showing progressive HN. There is no obvious explanation for this finding which suggests that different mechanisms of fibrosis development may be at work in glomeruli, compared to tubulointerstitial tissue.

SNTA1 is the most common syntrophin isoform, involved in calcium signaling and actin cytoskeleton, and can be detected in many tissues (23). It has been shown that dexamethasone-mediated restoration of podocyte function in minimal change nephropathy, is associated with an increased SNTA1 expression (24). In our study , SNTA1 is expressed at a lower level in the P group, thus possibly suggesting a higher podocyte damage.

DNAJB1 is a member of the hsp40 (heat shock protein 40 kD) family of proteins and is also underrepresented in HN P glomeruli. DNAJB1 is a negative regulator of EGFR, a growth factor which has been found to attenuate diabetic nephropathy (24, 25). Accordingly, heat shock proteins are known to be renoprotective during injury and traumata (26).

Proteinuria, tubular hypertrophy, oxidative stress, collagen turnover, chronic inflammation, and vasoactive substances can contribute to the pathogenesis of hypertensive renal fibrosis (27). In particular, hypertension leads to oxidative stress, sustained by mitochondria, and resulting in tubulo-interstitial fibrosis (28). In our geneset enrichment analysis, focusing on genes encoding proteins identified by proteomic analysis, we observed that mitochondrial genesets were overrepresented in HN P glomeruli, possibly, at least in part, as a result of oxidative stress. In humans mitochondrial dysfunction induces aberrant tubular inflammation, representing a major profibrotic process. Moreover, alterations of endoplasmic reticulum–mitochondria crosstalk might contribute to the progression of acute kidney injury to CKD (29).

Accordingly, it has been suggested that inhibition of mitochondrial complex-1 activity by rotenone could slow the progression of acute kidney injury (AKI) to CKD (30).

A comparison of glomerular protein signature associated with HN progression with that observed in IgAN nephropathy progressors vs non-progressors (5, 16) shows that the number of differentially abundant proteins common to both datasets is minute (n=3), consistent with very different underlying biological and clinical processes. Whereas alterations in metabolism, cell-cell communication and mitochondrial function appear to be enriched in HN progressors glomeruli, immune- and complement-related alterations predominate in progressing IgA nephropathy. Therefore, importantly, the reported HN protein signature does not appear to be unspecifically associated with glomerular damage.

There is a lack of HN glomerular proteomics datasets in literature, usable for external validation of the classifier presented here. Most available datasets compare a "healthy" to a diseased state, whereas prognostic impact of defined glomerular protein signatures is not explored. Our study contributes to fill this knowledge gap and paves the way towards a wider use of glomerular proteomic profiling to predict HN progression, thereby favoring the development of earlier and more effective treatment options.

Limitations of our study even beyond the relatively low sample number should be acknowledged. Our study is based on a registry (NKBR) and contains therefore fewer clinical and laboratory data as compared to a prospective trial. Patients with HN who undergo a kidney biopsy probably differ from patients who do not, and the latter group remains much larger in current clinical practice. The most important difference is probably represented by a higher level of proteinuria and a more frequent presence of hematuria in patients undergoing kidney biopsy. However, to the best of our knowledge, in no study kidney biopsies have routinely been performed in all patients with suspected HN due to ethical and practical concerns. Therefore, kidney biopsies from patients who had indication for kidney biopsy provide a unique opportunity to study morphological changes and their consequences at the proteome or transcriptome level.

In addition, the investigation of another CKD beyond IgAN would help to further explore the potential specificity of our signature of HN progression.

Most obviously, further research is warranted to validate this data in larger cohorts and to clarify whether overexpression of the five classifier proteins identified here could also be investigated in liquid biopsies, such as serum and urine samples, thus potentially envisaging non-invasive, advanced disease monitoring, also targeting patients solely diagnosed on clinical criteria, without requiring a renal biopsy.

In conclusion, glomerular proteomic profiling at the time of the diagnostic kidney biopsy can be used to differentiate subsequent stable patients from patients developing ESRD as a result of hypertensive nephropathy.

Acknowledgments

We are thankful to Giulio Spagnoli, PhD, University of Basel, Switzerland, for correction of the manuscript and to Dagny Ann Sandnes for help with the immunohistological analyses as well as to Mohammad Ibrahim from our institution for his initial work of microdissection and sample preparation.

Source of funding

This work was supported by an open project grant (number F-12559) to Hans-Peter Marti and an open project grant (number 911968) to Bjørn Egil Vikse both from the Western Norway Regional Health Authority.

Conflict of Interests

All authors, HM, BEV, OE, AS, KF, TO, and HPM, report no conflict of interests in any respects, such as financial relationships with any type of institutions.

REFERENCES

1. Lusco MA, Najafian B, Alpers CE, Fogo AB**.** AJKD Atlas of Renal Pathology: Arterionephrosclerosis. Am J Kidney Dis. 2016;67(4):e21-2.
2. Reisaeter A, Aasberg A 2010;Pageshttp://www.nephro.no/nnr/AARSRAPPORT_NNR_2017.pdf.
3. Hart PD, Bakris GL**.** Hypertensive nephropathy: prevention and treatment recommendations. Expert Opin Pharmacother. 2010;11(16):2675-86.
4. Kidney Disease: Improving Global Outcomes CKDMBDUWG**.** KDIGO 2017 Clinical Practice Guideline Update for the Diagnosis, Evaluation, Prevention, and Treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD). Kidney Int Suppl (2011). 2017;7(1):1-59.
5. Paunas TIF, Finne K, Leh S, Marti HP, Mollnes TE, Berven F, et al. Glomerular abundance of complement proteins characterized by proteomic analysis of laser-captured microdissected glomeruli associates with progressive disease in IgA nephropathy. Clin Proteomics. 2017;14:30.
6. Heslse direktoratet Norway, Guideline of the treatment of arterial hypertension 2018; Pageshttps://www.helsedirektoratet.no/retningslinjer/forebygging-av-hjerte-og-karsykdom/legemidler-ved-primaerforebygging-av-hjerte-og-karsykdom/legemiddelbehandling-av-hoyt-blodtrykk#752b5a70-98c1-4210-9ca0-4dda34751375-praktisk
7. Wisniewski JR**.** Proteomic sample preparation from formalin fixed and paraffin embedded tissue. J Vis Exp. 2013(79).
8. Ikeda K, Monden T, Kanoh T, Tsujie M, Izawa H, Haba A, et al. Extraction and Analysis of Diagnostically Useful Proteins from Formalin-fixed, Paraffin-embedded Tissue Sections. Journal of Histochemistry & Cytochemistry. 1998;46(3):397-403.
9. Cox J, Mann M**.** MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26(12):1367-72.

10. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Mol Cell Proteomics. 2014;13(9):2513-26.
11. Jarnuczak AF, Vizcaino JA. Using the PRIDE Database and ProteomeXchange for Submitting and Accessing Public Proteomics Datasets. Curr Protoc Bioinformatics. 2017;59:13 31 1-13 31 12.
12. Deutsch EW, Bandeira N, Sharma V, Perez-Riverol Y, Carver JJ, Kundu DJ, et al. The ProteomeXchange consortium in 2020: enabling 'big data' approaches in proteomics. Nucleic Acids Res. 2020;48(D1):D1145-D52.
13. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019;47(D1):D442-D50.
14. Seccia TM, Caroccia B, Piazza M, Rossi GP. The Key Role of Epithelial to Mesenchymal Transition (EMT) in Hypertensive Kidney Disease. Int J Mol Sci. 2019;20(14).
15. Kalluri R, Neilson EG. Epithelial-mesenchymal transition and its implications for fibrosis. J Clin Invest. 2003;112(12):1776-84.
16. Paunas FTI, Finne K, Leh S, Osman TA, Marti HP, Berven F, et al. Characterization of glomerular extracellular matrix in IgA nephropathy by proteomic analysis of laser-captured microdissected glomeruli. BMC Nephrol. 2019;20(1):410.
17. Zhou LT, Lv LL, Qiu S, Yin Q, Li ZL, Tang TT, et al. Bioinformatics-based discovery of the urinary BBOX1 mRNA as a potential biomarker of diabetic kidney disease. J Transl Med. 2019;17(1):59.
18. Igarashi P. Following the expression of a kidney-specific gene from early development to adulthood. Nephron Exp Nephrol. 2003;94(1):e1-6.

19. Thomson RB, Dynia DW, Burlein S, Thomson B, Booth CJ, Knauf F, et al. Deletion of Cdh-16 Ksp-cadherin leads to a developmental delay in the ability to maximally concentrate urine in mouse. Am J Physiol Renal Physiol. 2021.

20. Dai DF, Sasaki K, Lin MY, Smith KD, Nicosia RF, Alpers CE, et al. Interstitial eosinophilic aggregates in diabetic nephropathy: allergy or not? Nephrol Dial Transplant. 2015;30(8):1370-6.

21. DeNichilo MO, Panagopoulos V, Rayner TE, Borowicz RA, Greenwood JE, Evdokiou A**.** Peroxidase enzymes regulate collagen extracellular matrix biosynthesis. Am J Pathol. 2015;185(5):1372-84.

22. Colon S, Luan H, Liu Y, Meyer C, Gewin L, Bhave G**.** Peroxidasin and eosinophil peroxidase, but not myeloperoxidase, contribute to renal fibrosis in the murine unilateral ureteral obstruction model. Am J Physiol Renal Physiol. 2019;316(2):F360-F71.

23. Bhat HF, Adams ME, Khanday FA**.** Syntrophin proteins as Santa Claus: role(s) in cell signal transduction. Cell Mol Life Sci. 2013;70(14):2533-54.

24. Jiang L, Hindmarch CC, Rogers M, Campbell C, Waterfall C, Coghill J, et al. RNA sequencing analysis of human podocytes reveals glucocorticoid regulated gene networks targeting non-immune pathways. Sci Rep. 2016;6:35671.

25. Park SY, Choi HK, Seo JS, Yoo JY, Jeong JW, Choi Y, et al. DNAJB1 negatively regulates MIG6 to promote epidermal growth factor receptor signaling. Biochim Biophys Acta. 2015;1853(10 Pt A):2722-30.

26. Chebotareva N, Bobkova I, Shilov E**.** Heat shock proteins and kidney disease: perspectives of HSP therapy. Cell Stress Chaperones. 2017;22(3):319-43.

27. Sun HJ**.** Current Opinion for Hypertension in Renal Fibrosis. Adv Exp Med Biol. 2019;1165:37-47.

28. Mennuni S, Rubattu S, Pierelli G, Tocci G, Fofi C, Volpe M**.** Hypertension and kidneys: unraveling complex molecular mechanisms underlying hypertensive renal damage. J Hum Hypertens. 2014;28(2):74-9.

29. Maekawa H, Inagi R**.** Pathophysiological Role of Organelle Stress/Crosstalk in AKI-to-CKD Transition. Semin Nephrol. 2019;39(6):581-8.

30. Zhang W, Sha Y, Wei K, Wu C, Ding D, Yang Y, et al. Rotenone ameliorates chronic renal injury caused by acute ischemia/reperfusion. Oncotarget. 2018;9(36):24199-208.

FIGURE LEGENDS

Figure 1. Hierarchical cluster and principal component analysis (PCA).

Hierarchical cluster analysis (Ward method) performed using 58 differentially abundant proteins with abs FC ≥1.5, p-value ≤0.05) yields incomplete separation of the two patient groups, with 2 progressor (P) samples (P-01 and P-02) clustering with non-progressor (NP) samples (A). In the PCA, the overlap of the 90 percentile contour ellipsoids also indicates that abundance values are not sufficient to fully explain the differences between the two sample groups (B). C and D: same analysis as in A and B, but with 17 proteins with absFC ≥2, p-value ≤0.05. This filter yields a complete separation of P and NP patient groups in hierarchical clustering (C) and PCA (D).

Figure 2. Classifier analysis. K nearest neighbor (KNN) analysis with leave-one-out internal cross validation was employed to develop a prognostic classifier signature of HN progression.

Normalized abundance values of 58 differentially abundant proteins were used to test the classification performance of sets of 1 to 10 proteins. A set of five proteins performed best, classifying 16 of 17 samples correctly. (A) Normalized abundance values of the five proteins included in the classifier. (B) Performance metrics and ROC of our classifier showing a high AUC of 0.99 to discriminate between NP and P. Abundance values of classifier proteins separate the two patient groups in hierarchical cluster analysis (C) and PCA (D). *NP = non-progressors, P = progressors.*

Figure 3. Immunohistochemistry (IHC).

IHC confirms the directionality of the abundance changes of the two differentially regulated proteins. Representative images show IHC staining in brown color for BBOX1 in the top row **(A,B)** and for cadherin 16 (CDH16) in the lower row **(C,D)**. BBOX1 signal is not-detected in stable HN glomeruli **(A)**, but is upregulated in glomeruli of HN progressors **(B)**, whereas it remains unmodified in tubuli. CDH16 staining is strong in tubuli but almost absent in stable, non progressing HN patients' glomeruli **(C)**, whereas it is increased in glomeruli during disease progression **(D)**. *Scale bar 200 μm.*

Figure 4. Comparison of signatures of HN progressors (P) vs nonprogressors (NP) vs IgAN (IGA) progressors (P) vs nonprogressors (NP).

(A) Protein signatures of HN and IgAN progression are different. (B) Three proteins are common to both lists of differentially abundant proteins: *O15143* (Actin-related protein 2/3 complex subunit 1B, ARPC1B), *P08311* (Cathepsin G, CTSG), and *P17927* (Complement receptor type 1, CR1). The original IgAN data (5) were reanalyzed using the same statistical process as for the HN data. (C) KNN classifier proteins of progression of HN are either not detected in early IgAN or are not significantly altered during progression of IgAN.

Supplementaery Figure 1. Principal component analysis (PCA) of epithelial-mesenchymal transition (EMT), as applied to HN data.

The PCA based on a previously described EMT proteome (15) fails to separate the progressor vs. non-progressor HN samples.

Table legends

Table 1. Clinical characteristics of patients with non-progressing and progressing HN at the time of biopsy. Key features are presented of all included subjects (n=17).

Table 2A. Differentially abundant proteins.

The twenty proteins with highest differential abundance, sorted by absolute fold change (FC) values. Positive FC mirror overabundance in P compared to NP, whereas negative FC indicate decreased abundance in P as compared to NP. *P = progressors, NP = non-progressors.*

Table 2B. Pathway enrichment analysis of the 58 differentially abundant proteins with expected activation in P and NP.

Canonical pathways with implications in cell-cell-contact and metabolism are prominently represented; z-scores cannot be determined in cases with insufficient number of analysis-ready molecules ("NA"), and z-score of 0 indicates no activation when compared to Ingenuity Knowledgebase.

Table 2C. Geneset Enrichment Analysis (GSEA).

Gene ontologies (GO), addressing genes encoding proteins identified by proteomic analysis, indicate that the expression of genes encoding proteins involved in metabolism, particularly mitochondrial metabolism, is enriched in progressors (top panel) while that of genes encoding proteins involved in cell-cell communication and structure is enriched in non-progressors (bottom panel). Significant gene ontologies genesets are sorted by decreasing normalized enrichment score.

Table 2D. Proteins associated with IgA nephropathy (IgAN) progressors (P) vs non-progressors (NP).

The twenty proteins showing the highest absolute fold change (FC) and p-value≤0.05 are reported. Positive FC indicate overabundance in P compared to NP, negative FC reflect decreased abundance in P as compared to NP.

Table 2E. Canonical pathways associated with IgA nephropathy (IgAN) progressors vs non-progressors.

Enriched canonical pathways in the set of 100 proteins with differential abundance (abs FC≥1, p-value ≤0.05) in IgAN progressors vs. non-progressors. A z-score >0 indicates activation in in progressors. NA: z-scores cannot be calculated due to low number of analysis-ready molecules in a pathway or the z-scores are very to close to 0. Thus, for these pathways no conclusion on the direction of change can be made.

Figures

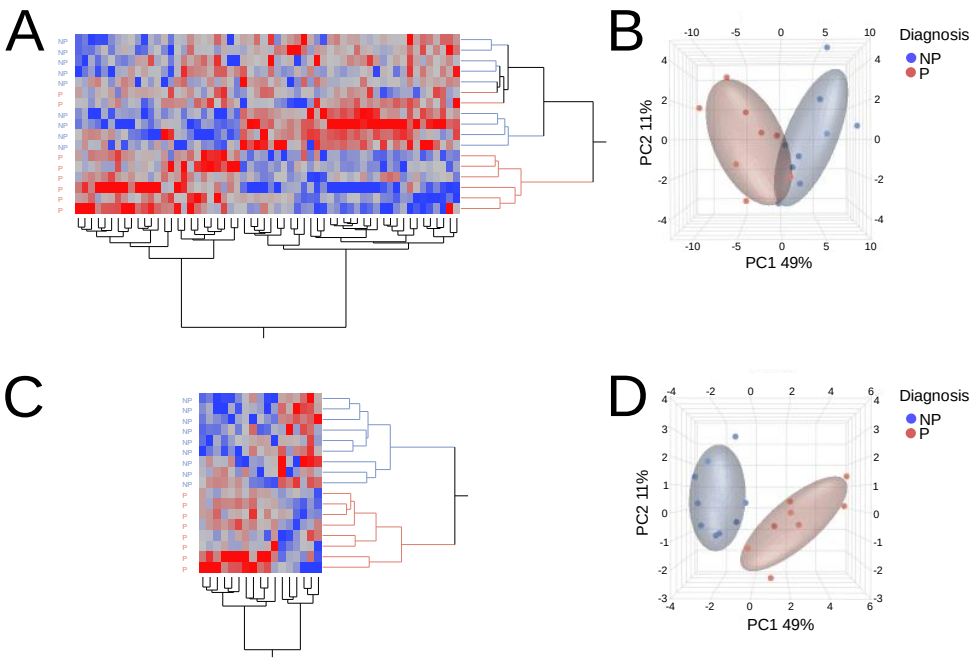


Figure 1

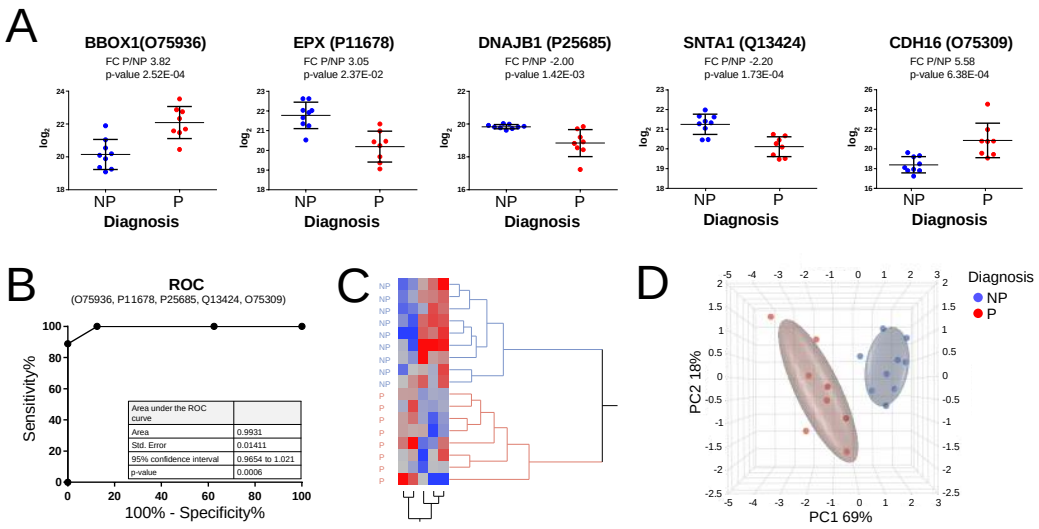


Figure 2

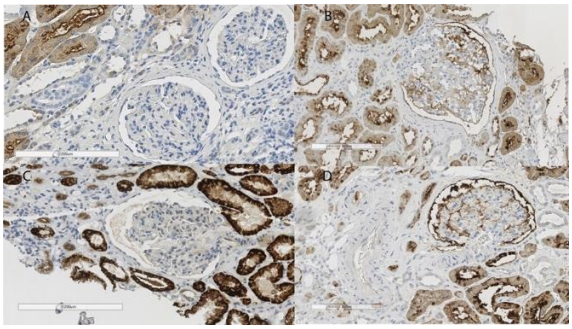
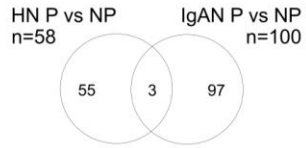


Figure 3

A



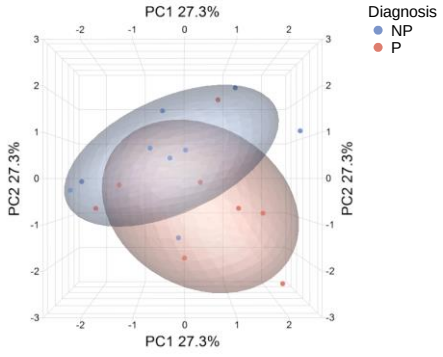
B

			HN			IGA	
Uniprot	Protein names	Symbol	fold change P/NP	p-value	adj. p-value	fold change P/NP	p-value	adj. p-value
O15143	Actin-related protein 2/3 complex subunit 1B	ARPC1B	1.58	3.57E-02	5.48E-01	1.64	1.71E-03	1.35E-01
P17927	Complement receptor type 1	CR1	-1.64	2.93E-02	5.07E-01	-1.65	5.04E-03	1.63E-01
P08311	Cathepsin G	CTSG	2.27	4.71E-02	5.77E-01	2.05	8.76E-03	1.81E-01

C

			HN			IGA	
Uniprot	Protein names	Symbol	fold change P/NP	p-value	adj. p-value	fold change P/NP	p-value	adj. p-value
O75309	Cadherin-16	CDH16	5.58	6.38E-04	2.98E-01			
O75936	Gamma-butyrobetaine dioxygenase	BBOX1	3.83	2.52E-04	1.57E-01			
P11678	Eosinophil peroxidase	EPX	-3.00	1.57E-04	1.57E-01			
Q13424	Alpha-1-syntrophin	SNTA1	-2.20	1.73E-04	1.57E-01	-1.02	9.17E-01	9.51E-01
P25685	DnaJ homolog subfamily B member 1	DNAJB1	-2.00	1.42E-03	4.39E-01	-1.08	2.74E-01	6.39E-01

Figure 4



Supplementaery Figure 1

Tables

	Non-progressing	Progressing	p-value
Patient number (N)	9	8	
Gender M/F	5/4	5/3	0.8
Age	57.3±10.8	56.9±7.5	0.9
Body height (cm.)	169±6	170±9	0.8
Body weight (kg.)	72.8±5.0	76.6±5.5	0.2
BMI	25.5±2.5	26.6±3.7	0.5
Serum creatinine (mcmol/l)	130±33	152±27	0.2
eGFR (ml/min/1.73m2)	51±9.8	44±13	0.2
Urinary protein (g/24h)	0.40±0.47	0.57±0.44	0.5
Systolic BP (mmHg)	157±36	146±21	0.5

Table 1. Clinical characteristics of patients with non-progressing and progressing HN at the time of biopsy. Key features are presented of all included subjects (n=17).

UniProt	Protein names	Symbol	FC P/NP	p-value	adj. p-value
O75309	Cadherin-16	CDH16	5,58	6,38E-04	2,98E-01
P16662	UDP-glucuronosyltransferase 2B7	UGT2B7	4,66	6,00E-03	4,55E-01
O75936	Gamma-butyrobetaine dioxygenase	BBOX1	3,83	2,52E-04	1,57E-01
P05783	Keratin, type I cytoskeletal 18	KRT18	3,67	1,62E-02	4,55E-01
P09110	3-ketoacyl-CoA thiolase, peroxisomal	ACAA1	3,05	2,37E-02	5,07E-01
P11678	Eosinophil peroxidase	EPX	-3,00	1,57E-04	1,57E-01
P49821	NADH dehydrogenase flavoprotein 1, mitochondrial	NDUFV1	2,77	1,54E-02	4,55E-01
Q96P63	Serpin B12	SERPINB12	-2,49	3,02E-02	5,14E-01
Q9Y617	Phosphoserine aminotransferase	PSAT1	2,48	1,30E-02	4,55E-01
O60547	GDP-mannose 4,6 dehydratase	GMDS	-2,31	1,43E-02	5,77E-01
P08311	Cathepsin G	CTSG	2,27	4,71E-02	4,55E-01
Q14894	Ketimine reductase mu-crystallin	CRYM	2,24	1,07E-02	4,55E-01
Q13424	Alpha-1-syntrophin	SNTA1	-2,20	1,73E-04	1,57E-01
Q92522	Histone H1x	H1FX	-2,04	1,45E-02	4,55E-01
P12429	Annexin A3	ANXA3	2,03	5,53E-03	4,55E-01
P25685	DnaJ homolog subfamily B, 1	DNAJB1	-2,00	1,42E-03	4,39E-01
Q86XE5	4-hydroxy-2-oxoglutarate aldolase, mitochondrial	HOGA1	2,00	2,54E-02	5,07E-01
P09936	Ubiquitin carboxyl-terminal hydrolase isozyme L1	UCHL1	1,97	1,30E-02	4,55E-01
P46939	Utrophin	UTRN	-1,92	2,53E-02	5,07E-01
Q9UPZ6	Thrombospondin type-1 domain-containing protein 7A	THSD7A	-1,90	9,00E-04	3,37E-01

Table 2A. Differentially abundant proteins.

The twenty proteins with highest differential abundance, sorted by absolute fold change (FC) values. Positive FC mirror overabundance in P compared to NP, whereas negative FC indicate decreased abundance in P as compared to NP. *P = progressors, NP = non-progressors.*

Canonical Pathways	z-score	p-value	adj. p-value
Epithelial Adherens Junction Signaling	NA	2,75E-06	4,90E-04
Sertoli Cell-Sertoli Cell Junction Signaling	NA	1,26E-05	1,12E-03
Tight Junction Signaling	NA	1,02E-03	6,03E-02
RhoGDI Signaling	NA	2,09E-03	9,12E-02
Protein Kinase A Signaling	0	3,31E-03	1,16E-01
4-hydroxyproline Degradation I	NA	5,01E-03	1,21E-01
GDP-L-fucose Biosynthesis I (from GDP-D-mannose)	NA	5,01E-03	1,21E-01
Sirtuin Signaling Pathway	NA	6,17E-03	1,21E-01
Cellular Effects of Sildenafil (Viagra)	NA	6,17E-03	1,21E-01
L-carnitine Biosynthesis	NA	7,41E-03	1,31E-01
Germ Cell-Sertoli Cell Junction Signaling	NA	9,12E-03	1,31E-01
Retinoate Biosynthesis II	NA	1,00E-02	1,31E-01
Serine Biosynthesis	NA	1,26E-02	1,31E-01
Eumelanin Biosynthesis	NA	1,26E-02	1,31E-01
Citrulline-Nitric Oxide Cycle	NA	1,26E-02	1,31E-01
Remodeling of Epithelial Adherens Junctions	NA	1,26E-02	1,31E-01
Agrin Interactions at Neuromuscular Junction	NA	1,32E-02	1,31E-01
ILK Signaling	NA	1,35E-02	1,31E-01
Agranulocyte Adhesion and Diapedesis	NA	1,66E-02	1,53E-01
Superpathway of Serine and Glycine Biosynthesis I	NA	1,74E-02	1,53E-01

Table 2B. Pathway enrichment analysis of the 58 differentially abundant proteins with expected activation in P and NP.

Canonical pathways with implications in cell-cell-contact and metabolism are prominently represented; z-scores cannot be determined in cases with insufficient number of analysis-ready molecules ("NA"), and z-score of 0 indicates no activation when compared to Ingenuity Knowledgebase.

Enriched in Progressors

Geneset	Size	ES	NES	p-val	adj. q-val
GO_Mitochondrial matrix	119	0,63	2,53	0	0
GO_Monocarboxylic acid catabolic process	43	0,72	2,48	0	0
GO_Oxidation reduction process	259	0,54	2,44	0	0
GO_Oxidoreductase activity	197	0,55	2,39	0	0
GO_Cellular modified amino acid metabolic process	52	0,67	2,36	0	0
GO_Monocarboxylic acid metabolic process	151	0,57	2,36	0	0
GO_Organic acid metabolic process	294	0,51	2,33	0	0
GO_Cellular respiration	75	0,62	2,29	0	0
GO_Small molecule catabolic process	126	0,57	2,29	0	0
GO_Organic acid catabolic process	86	0,59	2,28	0	0

Enriched in Nonprogressors

Geneset	Size	ES	NES	p-val	adj. q-val
GO_Cell junction organization	140	0,52	2,31	0	0
GO_Chromatin DNA binding	20	0,75	2,27	0	0,001
GO_Cell junction assembly	89	0,54	2,26	0	0,001
GO_Ficolin 1 rich granule membrane	23	0,72	2,24	0	0,001
GO_Nucleosome binding	14	0,81	2,22	0	0,001
GO_Action potential	18	0,74	2,18	0	0,001
GO_Chromattin assembly or disassembly	20	0,73	2,18	0	0,001
GO_Calmodulin binding	36	0,62	2,18	0	0,001
GO_Negative regulation of cellular protein localization	19	0,73	2,14	0	0,003
GO_Cell cell junction organization	49	0,57	2,12	0,002	0,004

Table 2C. Geneset Enrichment Analysis (GSEA).

Gene ontologies (GO), addressing genes encoding proteins identified by proteomic analysis, indicate that the expression of genes encoding proteins involved in metabolism, particularly mitochondrial metabolism, is enriched in progressors (top panel) while that of genes encoding proteins involved in cell-cell communication and structure is enriched in non-progressors (bottom panel). Significant gene ontologies genesets are sorted by decreasing normalized enrichment score.

UniProt	Protein names	Symbol	FC P/NP	p-value	adj. p-value
P08123	Collagen alpha-2(I) chain	COL1A2	-4,83	2,18E-02	2,51E-01
P98088	Mucin-5AC	MUC5AC	4,09	1,60E-03	1,35E-01
Q96CM8	Acyl-CoA synthetase family member 2, mitochondrial	ACSF2	-3,38	4,57E-02	3,27E-01
P10643	Complement component C7	C7	3,28	4,28E-04	9,21E-02
P08519	Apolipoprotein(a)	LPA	3,26	5,77E-03	1,69E-01
O95954	Formimidoyltransferase-cyclodeaminase	FTCD	-3,16	4,91E-03	1,63E-01
Q02413	Desmoglein-1	DSG1	-3,02	1,35E-02	2,08E-01
Q02985	Complement factor H-related protein 3	CFHR3	2,94	6,82E-03	1,81E-01
P47929	Galectin-7	LGALS7	2,83	4,27E-02	3,24E-01
P20711	Aromatic-L-amino-acid decarboxylase	DDC	-2,78	7,98E-03	1,81E-01
Q9BXR6	Complement factor H-related protein 5	CFHR5	2,72	3,90E-03	1,59E-01
P36980	Complement factor H-related protein 2	CFHR2	2,68	2,85E-04	9,21E-02
O75891	Cytosolic 10-formyltetrahydrofolate dehydrogenase	ALDH1L1	-2,66	8,21E-03	1,81E-01
Q6NZI2	Polymerase I and transcript release factor	PTRF	2,55	4,55E-04	9,21E-02
P07357	Complement component C8 alpha chain	C8A	2,54	7,28E-04	1,05E-01
P13671	Complement component C6	C6	2,53	6,88E-04	1,05E-01
P14780	Matrix metalloproteinase-9	MMP9	2,51	1,07E-02	1,91E-01
Q13011	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	ECH1	-2,50	9,34E-03	1,85E-01
P05976	Myosin light chain 1/3, skeletal muscle isoform	MYL1	2,48	1,01E-02	1,89E-01
P24158	Myeloblastin	PRTN3	2,45	4,35E-02	3,25E-01

Table 2D. Proteins associated with IgA nephropathy (IgAN) progressors (P) vs non-progressors (NP).

The twenty proteins showing the highest absolute fold change (FC) and p-value≤0.05 are reported. Positive FC indicate overabundance in P compared to NP, negative FC reflect decreased abundance in P as compared to NP.

Canonical Pathways	z-score	p-value	adj. p-value
Complement System	1,342	3,16E-20	5,01E-18
LXR/RXR Activation	2,111	5,01E-12	4,27E-10
FXR/RXR Activation	NA	1,62E-10	9,12E-09
Atherosclerosis Signaling	NA	8,91E-08	3,39E-06
Acute Phase Response Signaling	NA	1,00E-07	3,39E-06
Airway Pathology in Chronic Obstructive Pulmonary Disease	NA	1,17E-05	1,17E-05
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	2,449	1,82E-05	3,80E-04
Clathrin-mediated Endocytosis Signaling	NA	1,95E-05	3,80E-04
Hepatic Fibrosis / Hepatic Stellate Cell Activation	NA	2,00E-05	3,80E-04
IL-12 Signaling and Production in Macrophages	NA	2,63E-05	4,37E-04
Phenylalanine Degradation I (Aerobic)	NA	1,07E-04	1,66E-03
Xenobiotic Metabolism AHR Signaling Pathway	0	5,37E-04	7,41E-03
Serotonin Receptor Signaling	NA	8,32E-04	1,07E-02
Neuroprotective Role of THOP1 in Alzheimer's Disease	NA	1,66E-03	2,00E-02
Granzyme A Signaling	NA	2,95E-03	3,31E-02
Melatonin Degradation III	NA	4,27E-03	4,07E-02
Sorbitol Degradation I	NA	4,27E-03	4,07E-02
Dopamine Receptor Signaling	NA	4,37E-03	4,07E-02
Maturity Onset Diabetes of Young (MODY) Signaling	NA	4,57E-03	4,07E-02
Apelin Liver Signaling Pathway	NA	5,50E-03	4,57E-02

Table 2E. Canonical pathways associated with IgA nephropathy (IgAN) progressors vs non-progressors.

Enriched canonical pathways in the set of 100 proteins with differential abundance (abs FC≥1, p-value ≤0.05) in IgAN progressors vs. non-progressors. A z-score >0 indicates activation in in progressors. NA: z-scores cannot be calculated due to low number of analysis-ready molecules in a pathway or the z-scores are very to close to 0. Thus, for these pathways no conclusion on the direction of change can be made.