



OPEN Identification of mechanistic CKD biomarkers in a rat SNx kidney fibrosis model by transcriptomics and proteomics detectable in biofluids

Karin Barnouin^{1,6}✉, Elisa Tonoli², Clare Coveney³, John Atkinson^{1,7}, Margarida Sancho^{1,8}, Andrew Skelton¹, David J. Boocock³, Linghong Huang^{1,9}, Joseph Shephard¹, Timothy S. Johnson^{1,5,9}, Elisabetta A. M. Verderio^{2,4}✉ & Breda Twomey¹✉

The rat sub-total nephrectomy (SNx) is a functional model of general chronic kidney disease (CKD) where the main pathological driver is glomerular hypertension representative of several subtypes of CKD. Comprehensive transcriptomics and proteomics analyses on the SNx rats were performed to identify biomarkers in plasma or urine that correlate with kidney disease and functional kidney loss. Kidneys were subjected to collagen I and III staining for fibrosis scoring, SWATH-MS proteomics and bulk RNA-sequencing transcriptomics, with SWATH-MS also performed on plasma and urine. Differential expression analysis demonstrated significant dysregulation of genes and proteins involved in fibrosis, metabolism, and immune response in the SNx rats compared to controls. Gene ontology analysis of the intersecting genes and proteins from both studies demonstrated common biology between animal cohorts that reached the predefined kidney disease thresholds (serum creatinine > two-fold or proteinuria > three-fold increase over sham-operated). Thirteen significantly differential molecules were detected with consistent directional changes in both omics datasets. These molecules were detected independently in kidney (both RNA and protein) and urine (protein only), but not in plasma. Bioinformatics analysis enabled the identification of mechanistic CKD biomarkers including lumican and collagen alpha-1(III) chain, whose co-expression has previously been both implicated in fibrosis and detected in urine in CKD patients.

Chronic kidney disease (CKD) affects up to 16% of the population across the world and it is associated with significant mortality and morbidity¹. CKD is strongly linked to diabetes and hypertension. However, environmental factors and genetic background also concur with the aetiology of this complex condition. A common feature of failing kidneys is glomerulosclerosis and tubulointerstitial fibrosis that ultimately causes progressive kidney failure. Fibrotic remodelling is characterised by excess fibrous connective tissue and proliferation of fibroblasts, which is believed to be aberrant wound healing in response to chronic insult from the primary disease. During fibrotic remodelling, fibroblasts not only proliferate but assume an α -SMA positive myofibroblastic phenotype, epithelial cells can either die or undergo epithelial mesenchymal transformation leading to a more stromal extracellular matrix (ECM) producing phenotype while pericytes proliferate and take on myofibroblast characteristics². Overexpression of ECM components, accompanied by metabolic reprogramming is a histopathological hallmark of tissue fibrosis. The deposition, accumulation, increased crosslinking and stabilisation of collagens and fibronectin result in the hardening and scarring of tissues.

¹UCB Pharma, Slough SL1 3WE, UK. ²School of Science and Technology, Centre for Systems Health and Integrated Metabolic Research (SHiMR), Nottingham Trent University, Nottingham NG11 8NS, UK. ³John Van Geest Cancer Research Centre, Nottingham Trent University, Nottingham NG11 8NS, UK. ⁴Department of Biological, Geological, and Environmental Sciences, University of Bologna, BIGEA, 40126 Bologna, Italy. ⁵Experimental Renal Medicine, Oncology & Metabolism, University of Sheffield, Sheffield S10 2RZ, UK. ⁶Present address: MSD, London EC2M 6UR, UK. ⁷Present address: Gilead Sciences, Oxford OX4 4GE, UK. ⁸Present address: GSK, London TW8 9GS, UK. ⁹Present address: Mestag Therapeutics, Cambridge CB10 1XL, UK. ✉email: kbarnouin@gmail.com; elisabetta.verderio-edwards@ntu.ac.uk; Breda.Twomey@ucb.com

These changes cause pathological remodelling of the organ which lead to its failure^{3,4}. The development of tubulointerstitial fibrosis is the best predictor of progression⁵.

Initial detection of CKD is typically by use of protein dipstick in the urine, with subsequent characterisation and tracking of the stage of disease (CKD stage 1–5) determined by assessment of kidney function. This normally uses serum creatinine levels to estimate the glomerular filtration rate (eGFR) using the Modification of Diet in Renal Disease (MDRD) or CKD Epidemiology Collaboration (EPI) equations. More detailed assessment of stage also incorporates albuminuria measurements with high albuminuria being indicative of more rapidly progressing disease^{6,7}.

Kidney biopsy remains the gold standard to diagnose fibrosis within the kidney; however, it not only causes risk and discomfort to the patient but is an imperfect test, susceptible of sample bias and only useful when fibrosis is fully established. Moreover, the biopsy approach neither provides an early diagnosis nor a prognosis response. Importantly, it also limits the development of new anti-fibrotics due to the long time required to determine the effectiveness of a therapeutic fibrosis development.

Serum creatinine and proteinuria levels constitute the main biofluid biomarkers for detection and monitoring of CKD. However, these do not reflect all aspects of kidney disease pathophysiology (for example, kidney tubular health), do not always predict GFR decline, and are late disease markers requiring significant damage to occur. Several recent studies have tried to identify earlier, better, more predictive, and reliable markers of kidney disease in serum and urine to enable both clinical management and clinical trials^{8–12}. The identification of progression or mechanistic biomarkers and the underlying pathobiological processes in experimental animal models could enable their use in drug discovery and the translation of pre-clinical to clinical research.

The application of experimental models of CKD are useful for pre-clinical evaluation. However, because CKD progresses in humans over many years with a complex mix of genetics, environmental, age, sex, and lifestyle elements as underlying risk factors, it has been extremely challenging to mimic disease progression with animal models, that at the same time allow a fast and reliable assessment of disease progression or stability. Many experimental models have been developed that can be performed within acceptable time frames, have functional measurements relatable to human kidney disease, and that induce morphological and biochemical changes at the organ level translatable to the clinic, but all have some limitations. The models that enable readout of functional parameters while also displaying the collagen ECM accumulation in kidney tissue are most relevant for clinical translation.

Several kidney damage models such as the mouse unilateral ureteral obstruction (UUO), mouse and rat Adriamycin, mouse aristolochic acid nephropathy (AAN) and rat subtotal nephrectomy (SNx) models have been employed by us and others to simulate CKD^{13–20}. The SNx model is generally accepted as a general model that mimics many aspects of human CKD such as focal segmental glomerular sclerosis (FSGS), diabetic nephropathy (DN), membranous nephropathy (MN), IgA nephropathy (IgAN), membranoproliferative glomerulonephritis (MPGN), polycystic kidney disease (PKD), membranous glomerulonephritis (MGN) and rapidly progressive glomerular nephritis (RPGN). It displays some of the histological and pathobiological features typical of changes of focal segmental glomerular sclerosis²¹ with an increase in proteinuria and loss of kidney function associated with tubulointerstitial involvement. CKD in this model is driven by a combination of glomerular hypertension and inflammation, that are key factors in human disease progression. SNx rats also have cardiac defects, one of the co-morbidities for human CKD^{20,22}. The model therefore provides a platform with ready access to biofluids and tissues enabling the identification of disease relevant biomarkers.

In this investigation we have therefore aimed to use the SNx model to identify mechanistic markers of CKD development (fibrosis) that are detectable in biofluids and associated with changes in tissue biology, rather than those of plasma that appear in biofluids resulting from fibrotic renal structure deterioration such as currently used proteinuria and serum creatinine. It is anticipated that these could reflect the active nature of disease in humans and be employed as early diagnostic or progression biomarkers.

We have performed two independent rat SNx experiments and tracked disease progression by proteinuria and serum creatinine. Plasma, urine, and kidney tissue were analysed at predefined disease criteria, by label-free quantitative proteomics using the data-independent (DIA) analysis technique called sequential window acquisition of all theoretical fragmentation spectra (SWATH) mass spectrometry (MS) (SWATH-MS). This type of MS acquisition method provides greater depth in protein identification and better quantification than data-dependent analysis (DDA) analysis, especially when the data is interrogated by an in silico generated library^{23,24}. Kidney tissues were concurrently analysed by bulk RNA-seq transcriptomics. The readouts of both omics' experiments were compared and integrated, to identify strong reproducible disease biomarkers symptomatic of disease development in the rat SNx model of CKD.

Results

Preparation of diseased rat kidneys for omics analyses

As the objective of this work was identification of kidney disease biomarkers representative of progressive phenotypes that could correlate with biomarkers in the more accessible serum and urinary compartments, we thought it imperative to synchronise the animals for disease development, with tissue recovery occurring when disease thresholds had been reached. Two separate SNx studies (named study 1 and study 2) were performed at two independent sites with similar pre-defined disease monitoring criteria as summarised in Fig. 1a. Animals were only terminated when they reached the disease threshold functional levels of twofold increase in serum creatinine or threefold increases in proteinuria over sham-operated controls. However, the time to reach these disease thresholds varied between the studies where study 1 animals were terminated between days 81–116 and study 2 animals between days 52–98, reflecting variability not only between groups but also between individual animals within the same study. All the animals in study 1 were included in the analysis. However, only six out of eighteen rats reached the predetermined threshold in study 2. Representative kidney sections, stained

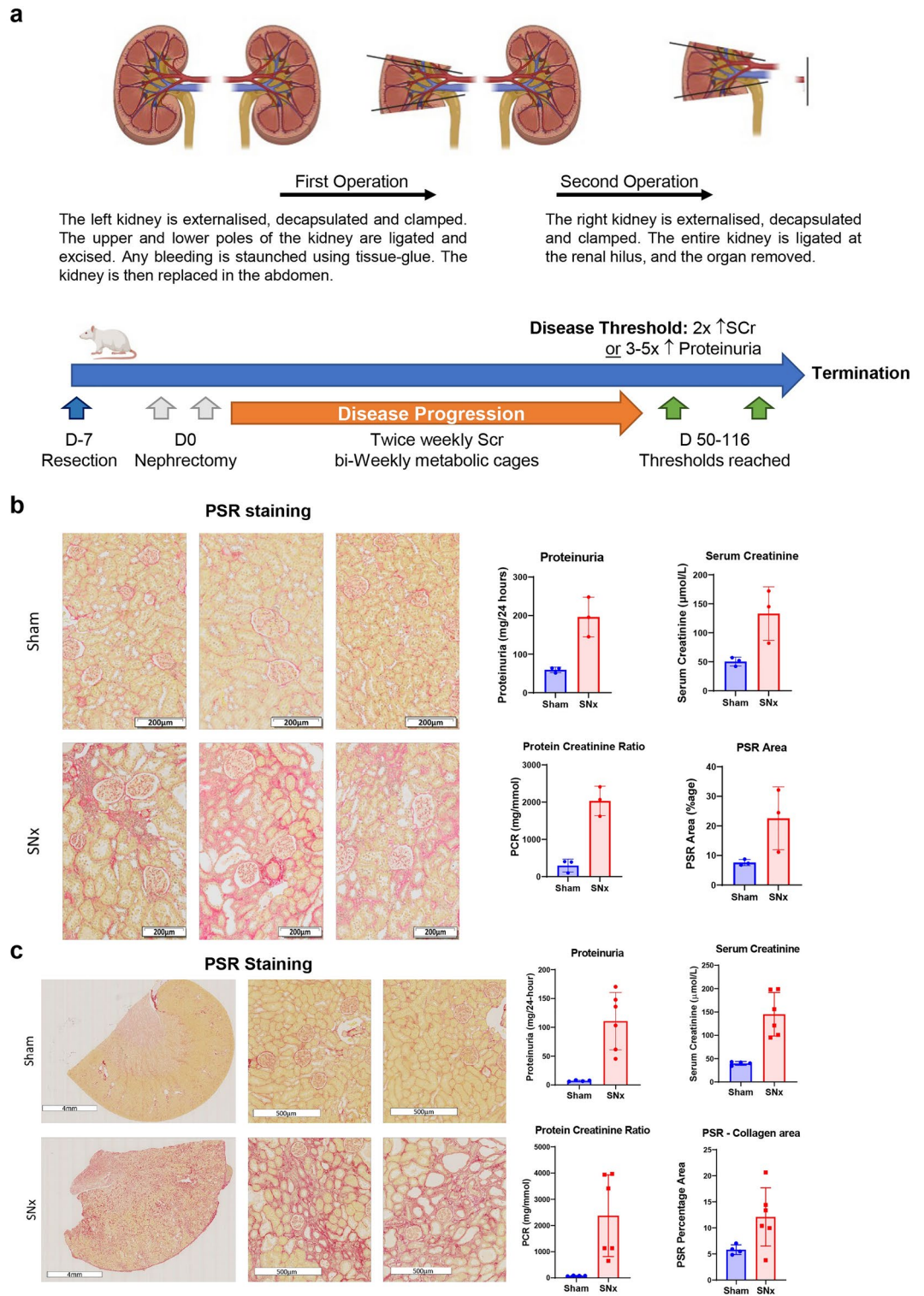


Fig. 1. Characterisation of rat SNx study samples. **(a)** The surgical procedure in rats termed sub-total nephrectomy (SNx) where 5/6 of the functioning kidneys are removed and post-surgery the animals develop failing kidney function from day 50 onwards, depicted on a timeline with monitoring of the failing kidneys using a combination of serum creatinine and proteinuria measurements. **(b)** Study 1 kidney sections from either sham ($n = 3$) or SNx rats ($n = 3$), stained with Picosirius Red (PSR) and the PSR-stained kidneys were scanned, the percentage of PSR area plotted alongside the other functional kidney readouts, Proteinuria, Serum creatinine and Protein creatinine ratio. Scale bar: 200 μm . **(c)** Study 2 kidney sections from 1/4 sham rats or 2/6 SNx rats, stained with PSR and PSR-stained kidneys from all rats were scanned, percentage of PSR area plotted alongside the other functional kidney readouts, Proteinuria, Serum creatinine and Protein creatinine ratio. Scale bar: 4 mm or 500 μm , as indicated.

for collagens I and III content using pico-sirius red (PSR) from both study 1 and 2, are shown in Figs. 1b, c, respectively. Urine and serum samples from both studies collected before termination were analysed for serum creatinine, proteinuria, and urinary proteinuria creatinine ratio as measurements of kidney function and glomerular damage (Figs. 1b, c), as well as processed for proteomics. Kidney sample preparation for transcriptomic and proteomic analyses were similar between the two studies.

RNA-Seq transcriptomics–global observations

Bulk-RNA seq were performed on rat SNx, Adriamycin, and UO kidney tissues. About 12,000 genes were profiled in each RNA-sequencing experiment (Table 1, Supplementary Fig. 1, and Supplementary data 1). Unsupervised PCA analysis of the two SNx studies transcriptomics data showed that, despite intra- and inter-study variability, the diseased SNx animals could be separated from sham (healthy) controls (Fig. 2a-I, b-I). In study 1, over 4000 proteins changed significantly (FDR < 0.05) after SNx treatment with over 1700 genes down- and up-regulated (Fig. 2a-II, Table 1). In study 2, over 4700 genes were significantly modulated (FDR < 0.05) with over 2800 genes that were down- and up-regulated (Fig. 2b-II, Table 1). Gene Pathway Ontology analysis showed that the most abundantly upregulated genes are represented by gene products involved in complement, immunological, and STAT signalling, associated with response to inflammation and ECM organisation pathways. Most of the significantly down-regulated genes are involved in the metabolism of amino acids, fatty acids or in mitochondrial processes (Fig. 2a-III, b-III).

Heatmap comparison of transcriptomics profiles of rat SNx with Adriamycin and UO kidney disease rat models revealed that although most of the genes that are up- or down-regulated in each experiment are similar, the magnitude of change differs between them. Overall, genes involved in T-cell inflammatory processes are upregulated in both studies, whereas down-regulated genes are involved in metabolism. Expectedly, ECM modelling pathways implicated in kidney fibrosis regulation are more upregulated in the aggressive UO and Adriamycin models than in both SNx studies (Supplementary Fig. 1).

DIA-SWATH proteomics – global observations

Over 6000 proteins were identified in the rat kidneys, and more than 450 proteins in study 2 plasma and urine were identified (Table 1, Fig. 3, Supplementary Fig. 2, and Supplementary data 1). As with the transcriptomics data, PCA analysis of the protein expression could clearly separate Sham (Healthy) and SNx rats for both kidney and urine datasets. In contrast, the separation was less distinct for plasma (Supplementary Fig. 3).

Gene ontology analysis of proteins

To gain a deeper understanding of the changes in the biological processes and molecular functions as well as in which cellular compartment these occur in CKD, Gene Ontology (GO) analysis was performed on the up- and down-regulated genes or proteins that were identified in both the proteomic and transcriptomic data. The web-based application Rivigo, that employs an algorithm that clusters GO terms according to semantic similarity, was used to summarise the data. The data was visualised as TreeMaps where the larger the square, the larger the log₁₀ p-value, the more significant the enrichment. For the up-regulated kidney proteins, post-SNx processes involved in stress and structural pathways were the most significantly enriched, for example, regulation of collagen fibril organization and cellular response to reactive nitrogen species (Supplementary Fig. 6a and data 2). The most enriched GO biological processes for down-regulated kidney proteins post SNx included responses to translation, oxidative stress, detoxification, cell matrix adhesion and metabolic pathways (e.g., glucose 6-phosphate metabolic process) (Supplementary Fig. 6b and data 2). In urine the identified down-regulated proteins have been shown to play roles in immune response, proteolysis, transport, or cell adhesion. Up-regulated proteins in urine are involved in endopeptidase activity, small molecule transport, plasminogen activation or cell structure while in plasma up-regulated proteins are implicated in complement activation, lipid metabolism, transport, and cytolysis (Supplementary Fig. 4 and data 2).

We then performed GO analysis of the intersecting proteins. The most significantly enriched biological processes in the up-regulated proteins in proteomics kidney studies were those involved in catabolic, metabolic, cell polarity and immune responses. Notably, the most CKD relevant significantly enriched GO biological processes were the proteins and genes that were in common between study 1 and 2 in the kidney proteomics and transcriptomics. Those proteins are involved in cell adhesion, cell differentiation, matrix organisation, cell proliferation, phagocytosis, and immune system processes. Furthermore, the cellular compartments were similar in study 1 and 2 kidneys in the proteomics and transcriptomics data and mainly corresponded to the localisation of proteins that play a role in cell shape and structure. Please refer to TreeMaps and tables for more details (Fig. 4, Supplementary Fig. 5 and 6, and data 2).

Differential expression	Proteins				Transcripts	
	Kidney study 1	Kidney study 2	Plasma study 2	Urine study 2	Kidney study 1	Kidney study 2
Total	6329	6871	488	632	12,714	11,533
Unchanged	5151	2593	444	184	8618	5667
Downregulated	220	659	2	207	1797	2812
Upregulated	958	3619	42	241	2299	3054

Table 1. Number of total, unchanged and significantly changing FDR < 0.05 proteins in SWATH-DIA and genes in RNA-seq experiments.

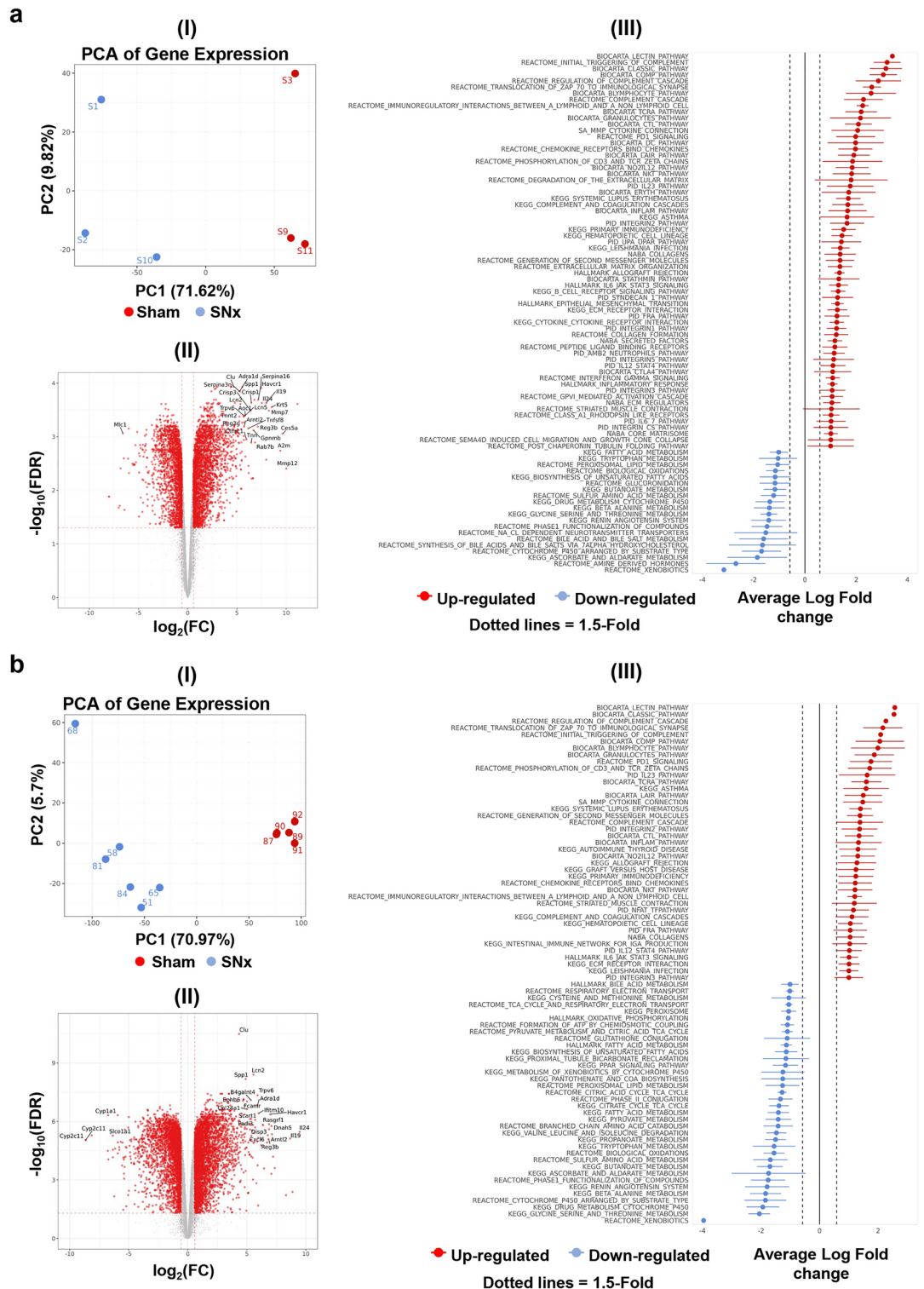


Fig. 2. RNA seq analysis of kidneys from SNx compared to sham rats. PCA (I), differential gene analysis Volcano plot (II) and Gene Pathway Ontology CAMERA Functional enrichment analysis (III) were performed for both study 1 (a) and study 2 (b).

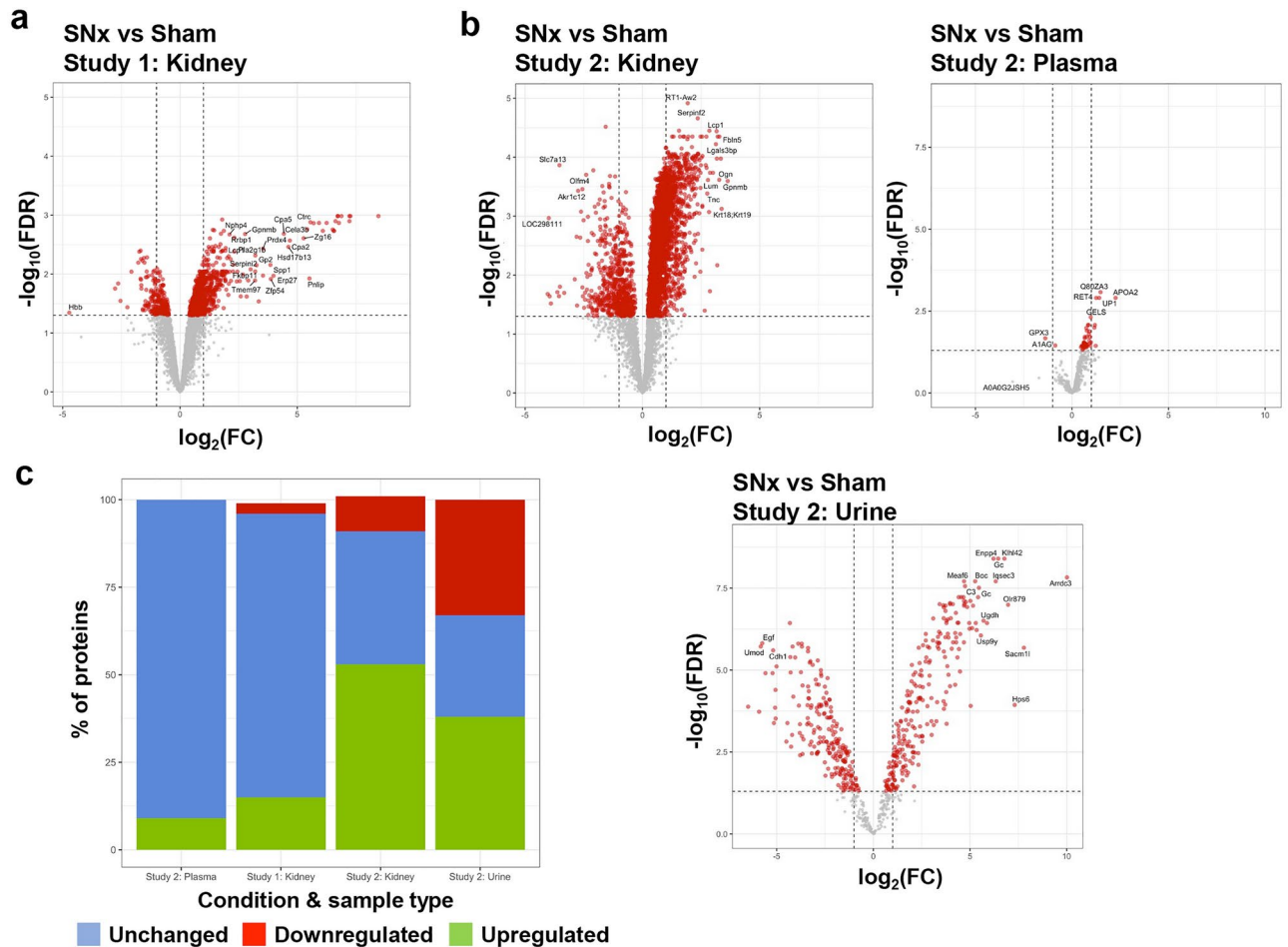


Fig. 3. Significantly changed proteins in SNx compared to sham rat kidneys. Volcano plots showing up- and down-regulated proteins in study 1 kidney (a) and study 2 kidney, plasma and urine (b). (c) Summary of protein counts expressed as percentages of downregulated (red), upregulated (green) and unchanged (blue) proteins compared to total number of proteins quantified in each sample.

Translatable CKD biomarker identification by data integration

Upset plots were used to identify the intersecting proteins and genes between the up- or down-regulated genes or proteins in both proteomics and transcriptomics methods (Fig. 5, Supplementary Fig. 5). Thirteen proteins were significantly differentially regulated (six upregulated, seven downregulated) at both protein and transcript levels in kidney and urine in all the analyses. The upregulated proteins were lumican (Lum), profilin 1 (Pfn1), collagen alpha-1(III) chain (Col3a1), complement component C6 (C6), collagen type VIII alpha1 chain (Col8a1) and Immunoglobulin Heavy Constant Mu (Ighm) (Fig. 5a, b, Supplementary Fig. 7a). The downregulated proteins were solute carriers Slc7a13 and Slc3a1, gamma-glutamyl transferase (Ggt1), the hydrolases alpha/beta hydrolase domain containing 14b (Abhd14b), Erythrocyte Membrane Protein Band 4.1 Like 3 (Epb41l3), trimethyllysine hydroxylase (Tmlhe) and the dipeptidyl peptidase 4 (Dpp4) (Fig. 5c, Supplementary Fig. 7b). Although these proteins are commonly detected in plasma, none of them passed the FDR 0.05 threshold for significant fold change in this biofluid. All these proteins except for DPP4 were also found to be similarly differentially regulated on the transcript level in the kidneys of Adriamycin and UOO rat models (Supplementary transcriptomic and proteomics Limma analysis data).

Discussion

CKD animal models are of value in the preclinical setting for identifying biomarkers and surrogate target engagement biomarkers²⁵. Several models have been developed but inherently display differences in their capacity to monitor functional kidney data alongside quantitative gene and protein changes. Selecting the right model of CKD is crucial in pre-clinical research to test interventions aimed at controlling CKD. Transcriptomic comparison of different models performed here has shown variation in the quantitative and qualitative gene expression profiles of TGF beta and SMAD pathways known to be modulated in fibrosis and extracellular matrix remodelling (Supplementary Fig. 1). Although none of the models are perfect predictors of human disease, the choice of model and rodent depends on the specific project and therapeutic target requirements, such as the option for functional correlation, representation of distinct pathway changes, and model duration, which can



Fig. 4. GO of up-regulated intersecting proteins and genes in kidney. TreeMaps showing enriched Molecular Functions (a), Biological processes (b) and Cellular compartments (c) of the significantly up-regulated proteins/genes in the kidneys of both studies.

distinguish acute versus chronic phases of disease progression. The ideal model for biomarker discovery would enable monitoring of the molecular changes non-invasively as the disease evolves.

The SNx model was selected because it is the best characterised and has the capacity for real-time functional monitoring of disease progression using serum creatinine and proteinuria measurements before termination. Typically, widespread glomerulosclerosis and tubulointerstitial fibrosis gradually develop, mimicking the

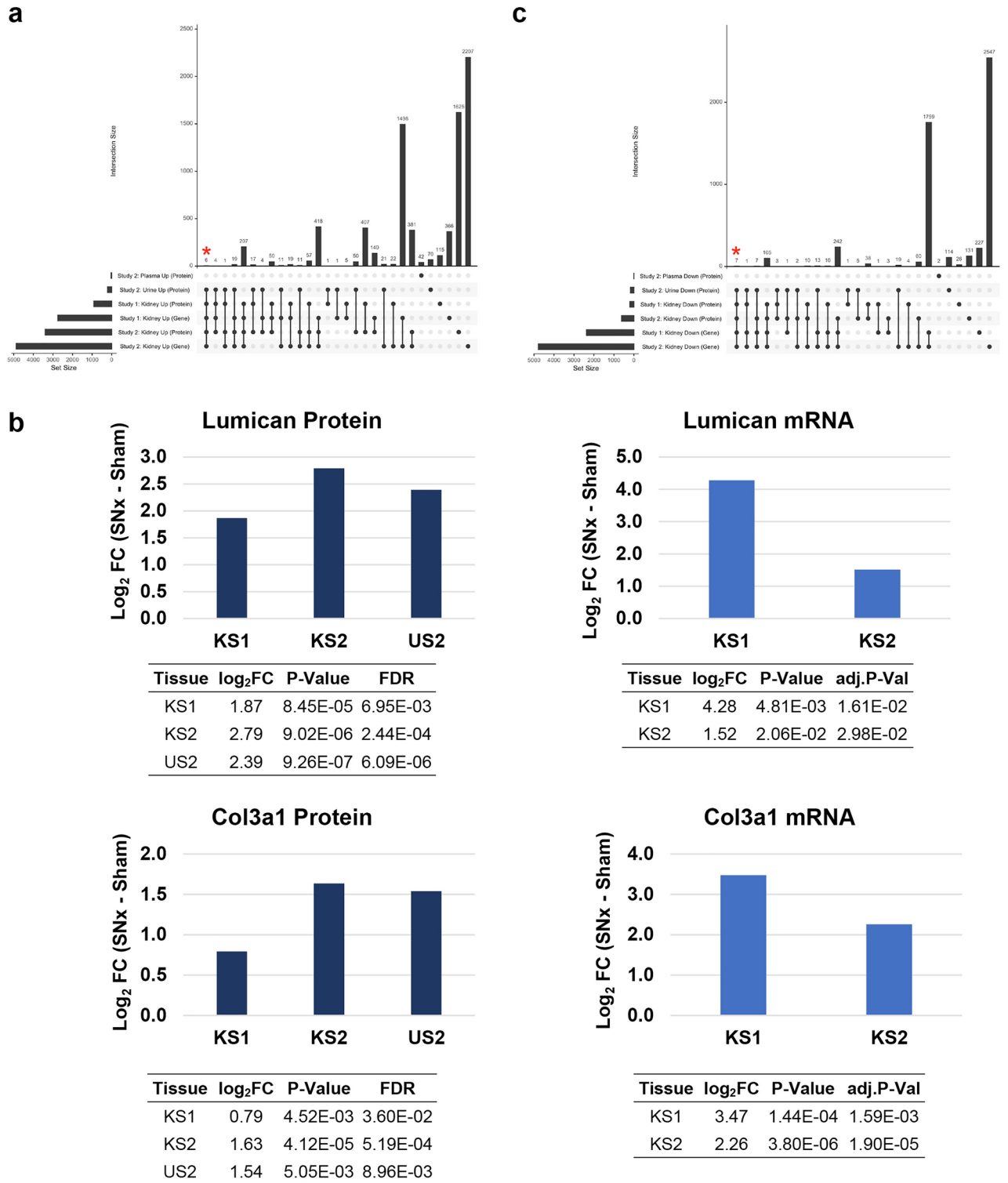


Fig. 5. Proteomics and gene expression intersections. (a) Upset plots showing up-regulated proteins/genes in common between study 1 and study 2 proteomics-transcriptomics data (kidneys, plasma, urine; FDR < 0.05). Red asterisk indicates proteins/genes up-regulated in all analyses apart from plasma. (b) Protein and gene expression of lumican and collagen type III alpha1 chain (Col3a1). (c) Upset plots showing down-regulated proteins/genes in common between study 1 and study 2 proteomics-transcriptomics data (kidneys, plasma, urine; FDR < 0.05). Red asterisk indicates proteins/genes down-regulated in all analyses apart from plasma.

progressive renal failure in humans consequent to loss of functional renal mass. However, the accurate interpretation of the data generated from the models is not straightforward because of the high variability of the response of the rats to the subtotal nephrectomy surgery. The analysis of the two SNx models independently generated in this study demonstrated both inter- and intra-model differences in the time it took to reach a specific disease threshold determined by functional analyses. The variability between animals was evident in the Principal Component Analysis (PCA) plots of both the RNA-seq and proteomics, where the healthy rats cluster close together, whereas the SNx rats are in comparison further apart (Fig. 2a, b and Supplementary Fig. 3). Integrating the omics data acquired from the two independent SNx studies has enabled the identification of CKD and fibrosis-specific biomarkers. Many genes were differentially expressed at both protein and transcript levels in kidney tissue in the two SNx studies. Many proteins also changed significantly in the urine, but not in plasma. Among those changing, six proteins were upregulated at both the protein and transcript level in the kidney and urine in all the analyses, including in the RNAseq studies performed in UUO and Adriamycin kidney tissues: Lum, Col3a1, Col8a1, Pfn1, C6, and Ighm. Interestingly, Lum, a proteoglycan, and Col3a1, co-express and regulate keratan, TGFbeta2, and FasL that in turn, modulate MDM2-p53, apoptosis, as well as TGF-beta signalling pathway²⁶. These pathways have previously been implicated in growth suppression and cell adhesion (KEGG pathway proteoglycans in cancer), lung disease²⁷, kidney (Supplementary Fig. 8), and other fibrotic diseases²⁸.

Notably, Lum mRNA levels has been found to be present in significantly higher levels in RPGN, DN, IgAN, FSGS, MGN, lupus nephritis (LN) and hypertensive nephropathy (HT). In the same study, lumican protein stained positive in DN patients²⁹. Interestingly, lumican was only detected in the urine of stage 3 CKD patients and not in healthy controls in a glycoproteomics study indicating that it could be used as an early biomarker³⁰. Furthermore, Lum was recently identified as an early diagnostic biomarker for children with Alport Syndrome³¹. Col3A1 kidney mRNA and protein in urine was also found to be elevated in a puromycin aminonucleoside (PAN)- induced nephropathy rat model³². It is thought that during ECM remodeling fragments of collagen degradation products are secreted in urine preceding proteinuria, and therefore could be used as an early marker of progressive renal fibrosis³³. Indeed, 55 fragments of Col3A1 were detected in a urine proteomic study of CKD patients³⁴. Downregulation of Pfn1 has been shown to be regulated by CK2 and has a protective effect against kidney fibrosis³⁵. In contrast, C6, Col8a1, and Ighm are all involved in the complement pathway, and their upregulation is related to a general immune response to injury and disease (Figs. 4b and 7a) and have also previously been shown to contribute to CKD progression³⁶.

Proteins and genes downregulated in urine as well as kidneys (but not plasma) at both the protein and transcript levels were Slc7a13, Slc3a1, Ggt1, Abhd14b, Epb41l3, Tmlhe and Dpp4. Slc7a13 and Slc3a1 are both involved in amino acid transport and are highly expressed in kidneys. Downregulation or dysfunction of solute carriers has previously been observed in kidney disease³⁷. Ggt1 catalyses the transfer of glutathione to small molecules and generally plays a protective role against toxins and oxidative stress, but also specifically in kidneys^{38,39}. Tmlhe is involved in amine, polyamine, and carnitine biosynthesis. Increased levels of L-carnitine have been shown to attenuate kidney fibrosis development⁴⁰. Interestingly, Dpp4 acts as a serine endopeptidase and inhibitors against its activity are used to treat diabetes and has also been shown to metabolise collagen in kidney⁴¹. Epb41l3 is an adaptor protein that interacts with cytoskeletal proteins and is believed to modulate the activity of methyl transferases such as PRMT3 and PRMT5^{42,43}. DNA-methylation is thought to contribute to epigenetic causes of diabetic kidney disease⁴⁴. Taken together, all these CKD biomarkers identified herein have been shown to be modulated during the development of fibrosis.

The aim of our analysis was to identify proteins that were detectable in all the transcriptomics and proteomics datasets that could be used as candidate mechanistic CKD markers. From the CKD biomarker candidates detectable in urine in the proteomics study conducted by Kim et al.⁴⁵ we also found significantly increased levels of Pros1 and Tf, but only in study 2 kidney and urine proteomics. Interestingly, established CKD marker NGAL1 (LCN2) transcript was significantly increased (\log_2 5-fold change (FC) (or 32 FC) in both studies 1 and 2). Protein levels were up by \log_2 1.3 FC (or 2.5 FC) and \log_2 2.3 FC (or 5 FC) in studies 1 and 2, respectively, but not in urine. Whereas KIM1 (Havcr1) was only significantly up in study 2 (\log_2 2.7 FC or 6.5 FC) kidney proteomics but not in urine proteomics or kidney transcriptomic. These observations indicate that KIM1 and NGAL1 may not be suitable as urine prognostic markers.

In summary, the data presented here is, to our knowledge, the first time that SWATH-MS data has been collected on SNx rat kidney tissue, plasma, and urine and systematically compared. In addition, there are no previous reports that compare the proteome profile with mRNA expression levels in SNx models. Whilst mRNA and protein expression levels do not always correlate, analysis of the differential expression of the proteins and genes in all the datasets and enriched gene ontologies have enabled the identification of Lum and Col3a1 as relevant translatable mechanistic CKD biomarkers that can be detected in urine as a single assay. The possibility of developing a multiplexed assay that tracks changes in several of the identified proteins could offer a platform for detecting early signs of progressive human renal disease before glomerular lesions and tubulointerstitial lesions become measurable. Furthermore, it could provide a signature to rapidly monitor response to anti-fibrotic therapies in clinical studies or the effectiveness of existing therapies earlier in clinical care. Finally, this unique multi-omics profiling data of renal tissue and biofluids post-SNx also serves as a repository together with others⁴⁶ for scientists in the field to obtain information that can be reanalysed and validated using orthogonal or targeted methods.

Methods

Subtotal nephrectomy rat, UUO mouse, and adriamycin mouse model generation

Subtotal Nephrectomy (SNx) was performed on male Wistar Han rats (200–400 g Charles River, UK) essentially as described^{22,47} using 2/3 resection of the right kidney followed two weeks later by left kidney nephrectomy.

Study 1 (performed at UCB Pharma) contained 3 non-op and 3 SNx rats. Study 2 (performed at the University of Sheffield) had 5 sham and 18 SNx operated rats. Serum creatinine was measured weekly and proteinuria biweekly with animals culled using a Schedule 1 method when proteinuria reached >3–fivefold increase or serum creatinine reached >twofold increase over sham-operated. At termination, 24-h urine and plasma samples were collected and snap frozen at -80°C . Following Schedule 1 killing, kidney tissue was recovered, with 1/4 fixed in NBF for histology and the three other 1/4 snap frozen in liquid nitrogen for SWATH or Transcriptomics.

For comparison to other models, UO, and Adriamycin mouse models of CKD were generated alongside the SNx at UCB Pharma under project licence (30/3051). Procedures were performed as described in^{48,49}, respectively. Detailed methods for the generation of the models and histology analyses can be found in the Supplementary Information 1 methods section.

The studies were carried out under license according to the regulations of housing conditions, husbandry, and procedures laid down by Her Majesty's Government, UK (Animals Scientific Procedures Act 1986). Study 1 was performed at UCB Biopharma under project license 30/3051 and study 2 at the University of Sheffield under project license (PBE09C70E). This study was reported according to ARRIVA guidelines.

Kidney clinical measurements

SNx rats were placed in metabolic cages to collect 24-h urine samples. Urine was centrifuged to remove any sediment from food or faeces prior to analysis. Protein content in urine was determined using a modified Lowry method⁵⁰. Urine creatinine was measured using an enzymatic creatinine assay (Crystal Chem, USA, Cat: 80,340) and serum creatinine a StatSensor Xpress (Nova Biomedical, USA) with creatinine Stat Strips. The protein:creatinine ratio was calculated as mg protein per mmol of creatinine.

Bulk RNA sequencing

Material was quantified using RiboGreen (Invitrogen) on the FLUOstar OPTIMA plate reader (BMG Labtech) and the size profile and integrity analysed on the 2200 or 4200 TapeStation (Agilent, RNA ScreenTape). RIN estimates for all samples were between 1.3 and 7.8. Input material was normalised to 30ng prior to library preparation. The RNA was depleted of ribosomal RNA using NEBNext rRNA Depletion Kit (NEB, Human/Mouse/Rat) following manufacturer's instructions. Strand specific library preparation was completed using NEBNext Ultra II mRNA kit (NEB) following manufacturer's instructions. Libraries were amplified (14 cycles) on a Tetrad (Bio-Rad) using in-house unique dual indexing primers (based on⁵¹). Individual libraries were normalised using Qubit, and the size profile was analysed on the 2200 or 4200 TapeStation. Individual libraries were normalised and pooled together accordingly. The pooled library was diluted to ~ 10 nM for storage. The 10 nM library was denatured and further diluted prior to loading on the sequencer. Paired end sequencing was performed using a NovaSeq6000 platform (Illumina, NovaSeq 6000 S2/S4 reagent kit v1.5300 cycles).

RNA-seq data analysis

Raw sequencing data was initially assessed with the FastQC and MultiQC utilities. Transcripts were quantified using the Rn6 reference transcriptome, and pre-built selective alignment index from RefGenie. Quantifications were summarised to the length scaled TPM unit via tximport and voom transformed with the Limma framework. Differential expression analysis was performed with Limma, and functional analyses with the CAMERA function and Mitch framework, using MSigDB reference gene sets⁵².

Proteomics sample preparation and LCMS SWATH analysis

SNx and sham kidney sections were processed as previously described⁵³ with minor modifications. Specifically, after kidneys were lysed in lysis buffer [9.5 M urea, 2% (w/v) DTT, 1% (w/v) *N*-octyl- β -glucopyranoside, protease inhibitors; 1 mL of buffer every 100 mg of tissue], protein extracts (75 μg) were vacuum concentrated to dryness and stored at -20°C for 16 h. Samples were then resuspended in solubilisation buffer [5% SDS (w/v), 100 mM triethylammonium bicarbonate (TEAB), pH 7.55] and proteins digested by S-Trap (PROTIFI, US) according to manufacturer's instructions. Urine samples were thawed on ice and supplemented with protease inhibitors (Roche), centrifuged at $5000\times g$, passed through a 0.45 μm syringe filter, concentrated and desalted using centrifugal concentrator tubes (Vivaspin 500, 5 kDa MWCO, Sigma-Aldrich). Urinary proteins underwent reduction, alkylation, and trypsin digestion for 16 h as previously described⁵³. Plasma samples were thawed on ice and supplemented with protease inhibitors (Roche), then cleared using Spin-X[®] Centrifuge Tube Filters (Corning) and 50 μg processed as described above. The resulting peptides were vacuum dried and stored at -20°C until reconstitution for MS analysis.

6.6 μg of sample was injected and trapped onto a YMC Triart-C18 pre-column (5×0.3 mm, 3 μm) at a flow rate of 10 $\mu\text{L}/\text{min}$ for 2 min prior to switching the valve to run the gradient elution through a YMC Triart-C18 analytical column (150×0.3 mm, 3 μm , 5 $\mu\text{L}/\text{min}$) in line with the Sciex TripleTOF 6600 via a DuoSpray Source using a 50 μm electrode in positive mode, +5500 V. The following linear gradients were used for SWATH-MS acquisition, mobile phase B increasing from 3 to 30% over 38 min, 30% to 40% over 5 min, 40% to 80% over 2 min for wash and re-equilibration (total run time 57 min). Data independent (SWATH) acquisition was performed using 100 variable SWATH windows (optimised on sample type) with the following parameters: TOFMS m/z 400–1250, MSMS m/z 100–1500, 25 ms accumulation time; 2.6 s total cycle time.

Label-free SWATH mass spectrometry data analysis

SWATH mass spectrometry data was converted to MZML using MSConvert (Proteowizard MS Covert Version: 3.0.21075-41cb77451) and processed using DIA-NN (ver1.8) (Demichev, 2020). Default settings in DIA-NN were used except that Match Between Runs (MBR) was also selected. In addition, to increase the depth of our proteome analysis, an in silico generated PROSIT library²⁴ of the annotated human proteome was generated within the

software and used to directly interrogate the DIA data. DIA-NN aggregates peptide level quantifications to the protein level. The resulting CSV was pivoted within R and further analysed in Perseus⁵⁴. Protein abundances were Log_2 transformed, and proteins filtered for a minimum of three valid values per sample. Missing values were imputed following a normal distribution using the default values within Perseus. The dataset was then normalised using median abundances. The resulting spreadsheet was exported for further analysis and data visualisation in the R environment.

Differential protein and transcript expression, upset plots, and gene ontology analysis

Limma Bioconductor R package was used to determine fold changes between groups. Volcano plots were generated by R package ggplot2⁵⁵. UpsetR package⁵⁶ was used to visualise and determine the numbers of unique genes in each sample set as well as intersecting proteins and their identities between datasets. Gene ontology analysis was then performed on the intersection proteins to determine the biological processes, molecular function, and cellular components and pathways involved in CKD using the Gene Ontology Database for Annotation, Visualization, and Integrated Discovery (DAVID) Bioinformatics Resource (v2022q3). Rat protein accession numbers were converted to gene names and imported into the database. Rivigo web server (<http://revigo.irb.hr/>) was subsequently employed to reduce the complexity and redundancy of the GO annotations. TreeMap R scripts generated in Rivigo were exported, and figures generated in R Studio after modification of the scripts to include $-\log_{10} p$ - values on the maps.

Data availability

The mass spectrometry proteomics data generated during the current study is available in ProteomeXchange Consortium via the PRIDE⁵⁷ partner repository with the dataset identifier PXD048564. The transcriptomics data discussed in this publication have been deposited in NCBI Gene Expression Omnibus^{58,59} and are accessible through GEO Series accession numbers (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSExxx>). The following accession numbers have been assigned to the data: GSE253009 (encompasses both study 1 and 2 kidney datasets), GSE253008 (subseries kidney study 1 dataset) and GSE253007 (subseries kidney study 2 dataset). All data analysed during this study are included in this published article (and its Supplementary Information files).

Received: 20 January 2024; Accepted: 10 March 2025

Published online: 02 April 2025

References

1. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **395**, 709–733. [https://doi.org/10.1016/S0140-6736\(20\)3045-3](https://doi.org/10.1016/S0140-6736(20)3045-3) (2020).
2. Sun, K.-H., Chang, Y., Reed, N. I. & Sheppard, D. α -Smooth muscle actin is an inconsistent marker of fibroblasts responsible for force-dependent TGF β activation or collagen production across multiple models of organ fibrosis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **310**, L824–L836. <https://doi.org/10.1152/ajplung.00350.2015> (2016).
3. Wynn, T. A. & Ramalingam, T. R. Mechanisms of fibrosis: Therapeutic translation for fibrotic disease. *Nat. Med.* **18**, 1028–1040. <https://doi.org/10.1038/nm.2807> (2012).
4. Burhan, I. et al. Interplay between transglutaminases and heparan sulphate in progressive renal scarring. *Sci. Rep.* **6**, 31343. <https://doi.org/10.1038/srep31343> (2016).
5. Genovese, F., Manresa, A. A., Leeming, D. J., Karsdal, M. A. & Boor, P. The extracellular matrix in the kidney: A source of novel non-invasive biomarkers of kidney fibrosis?. *Fibrogenesis Tissue Repair* **7**, 4. <https://doi.org/10.1186/1755-1536-7-4> (2014).
6. Ammirati, A. L. Chronic Kidney Disease. *Rev. Assoc. Med. Bras.* **66**(Suppl1), 3–9. <https://doi.org/10.1590/1806-9282.66.S1.3> (2020).
7. Charles, C. & Ferris, A. H. Chronic kidney disease. *Prim. Care* **47**, 585–595. <https://doi.org/10.1016/j.pop.2020.08.001> (2020).
8. Niewczas, M. A. et al. A signature of circulating inflammatory proteins and development of end-stage renal disease in diabetes. *Nat. Med.* **25**, 805–813. <https://doi.org/10.1038/s41591-019-0415-5> (2019).
9. Good, D. M. et al. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Mol. Cell. Proteomics* **9**, 2424–2437. <https://doi.org/10.1074/mcp.M110.001917> (2010).
10. Yu, Z. et al. Polygenic risk scores for kidney function and their associations with circulating proteome, and incident kidney diseases. *J. Am. Soc. Nephrol.* **32**, 3161–3173. <https://doi.org/10.1681/ASN.2020111599> (2021).
11. Tepus, M., Tonoli, E. & Verderio, E. A. M. Molecular profiling of urinary extracellular vesicles in chronic kidney disease and renal fibrosis. *Front. Pharmacol.* <https://doi.org/10.3389/fphar.2022.1041327> (2023).
12. Da Silva, L. M., Pullen, N., Pereira, M. & Johnson, T. S. Urinary levels of pro-fibrotic transglutaminase 2 (TG2) may help predict progression of chronic kidney disease. *PLoS One* **17**, e0262104. <https://doi.org/10.1371/journal.pone.0262104> (2022).
13. Vielhauer, V. et al. Obstructive nephropathy in the mouse: progressive fibrosis correlates with tubulointerstitial chemokine expression and accumulation of CC chemokine receptor 2- and 5-positive leukocytes. *J. Am. Soc. Nephrol.* **12**, 1173–1187. <https://doi.org/10.1681/ASN.V1261173> (2001).
14. Bryant, C., Cianciolo, R., Govindarajan, R. & Agrawal, S. Adriamycin-induced nephropathy is robust in N and modest in J substrain of C57BL/6. *Front. Cell. Dev. Biol.* **10**, 924751. <https://doi.org/10.3389/fcell.2022.924751> (2022).
15. Fragiadaki, M. et al. Interstitial fibrosis is associated with increased COL1A2 transcription in AA-injured renal tubular epithelial cells in vivo. *Matrix Biol.* **30**, 396–403. <https://doi.org/10.1016/j.matbio.2011.07.004> (2011).
16. Debelle, F. D. et al. Aristolochic acids induce chronic renal failure with interstitial fibrosis in salt-depleted rats. *J. Am. Soc. Nephrol.* **13**, 431–436. <https://doi.org/10.1681/ASN.V132431> (2002).
17. Huang, L., Scarpellini, A., Funck, M., Verderio, E. A. M. & Johnson, T. S. Development of a chronic kidney disease model in C57BL/6 mice with relevance to human pathology. *Nephron Extra* **3**, 12–29. <https://doi.org/10.1159/000346180> (2013).
18. Morrison, A. B. Experimentally induced chronic renal insufficiency in the rat. *Lab. Invest.* **11**, 321–32 (1962).
19. Simic, I., Tabatabaeifar, M. & Schaefer, F. Animal models of nephrotic syndrome. *Pediatr. Nephrol.* **28**, 2079–2088. <https://doi.org/10.1007/s00467-012-2376-5> (2013).
20. Yang, H.-C., Zuo, Y. & Fogo, A. B. Models of chronic kidney disease. *Drug Discov. Today Dis. Models* **7**, 13–19. <https://doi.org/10.1016/j.ddmod.2010.08.002> (2010).

21. Xu, B. J. et al. Proteomic patterns and prediction of glomerulosclerosis and its mechanisms. *J. Am. Soc. Nephrol.* **16**, 2967–2975. <https://doi.org/10.1681/ASN.2005030262> (2005).
22. Zhang, Y. & Kompa, A. R. A practical guide to subtotal nephrectomy in the rat with subsequent methodology for assessing renal and cardiac function. *Nephrology (Carlton)* **19**, 552–561. <https://doi.org/10.1111/nep.12279> (2014).
23. Gillet, L. C. et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* <https://doi.org/10.1074/mcp.O111.016717> (2012).
24. Gessulat, S. et al. Prosit: Proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat. Methods* **16**, 509–518. <https://doi.org/10.1038/s41592-019-0426-7> (2019).
25. Klein, J. et al. Renal fibrosis: Insight from proteomics in animal models and human disease. *Proteomics* **11**, 805–815. <https://doi.org/10.1002/pmic.201000380> (2011).
26. Schaefer, L. Small leucine-rich proteoglycans in kidney disease. *J. Am. Soc. Nephrol.* **22**, 1200–1207. <https://doi.org/10.1681/ASN.2010050570> (2011).
27. Wang, K. et al. Lumican is elevated in the lung in human and experimental acute respiratory distress syndrome and promotes early fibrotic responses to lung injury. *J. Transl. Med.* **20**, 392. <https://doi.org/10.1186/s12967-022-03597-z> (2022).
28. Krishnan, A. et al. Lumican, an extracellular matrix proteoglycan, is a novel requisite for hepatic fibrosis. *Lab. Invest.* **92**, 1712–1725. <https://doi.org/10.1038/labinvest.2012.121> (2012).
29. Feng, S. et al. Identification of lumican and fibromodulin as hub genes associated with accumulation of extracellular matrix in diabetic nephropathy. *Kidney Blood Press. Res.* **46**, 275–285. <https://doi.org/10.1159/000514013> (2021).
30. Vivekanandan-Giri, A. et al. Urine glycoprotein profile reveals novel markers for chronic kidney disease. *Int. J. Proteomics* **2011**, 1–18. <https://doi.org/10.1155/2011/214715> (2011).
31. Rhode, H. et al. Urinary protein-biomarkers reliably indicate very early kidney damage in children with alport syndrome independently of albuminuria and inflammation. *Kidney Int. Rep.* <https://doi.org/10.1016/j.ekir.2023.09.028> (2023).
32. Liu, M. et al. Targeting collagen type III in proteinuric kidney disease: Informing drug potential using the Jaccard-Tanimoto index. *Processes* **8**, 996. <https://doi.org/10.3390/pr8080996> (2020).
33. Hijmans, R. S. et al. Urinary collagen degradation products as early markers of progressive renal fibrosis. *J. Transl. Med.* **15**, 63. <https://doi.org/10.1186/s12967-017-1163-2> (2017).
34. Good, D. M. et al. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Mol. Cell. Proteomics* **9**, 2424–2437. <https://doi.org/10.1074/mcp.M110.001917> (2010).
35. Zoidakis, J. et al. Profilin 1 is a potential biomarker for bladder cancer aggressiveness. *Mol. Cell. Proteomics* <https://doi.org/10.1074/mcp.M111.009449> (2012).
36. Poppelaars, F. & Thurman, J. M. Complement-mediated kidney diseases. *Mol. Immunol.* **128**, 175–187. <https://doi.org/10.1016/j.molimm.2020.10.015> (2020).
37. Drozdziak, M., Drozdziak, M. & Oswald, S. Membrane carriers and transporters in kidney physiology and disease. *Biomedicines* **9**, 426. <https://doi.org/10.3390/biomedicines9040426> (2021).
38. Hayes, J. D., Flanagan, J. U. & Jowsey, I. R. Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* **45**, 51–88. <https://doi.org/10.1146/annurev.pharmtox.45.120403.095857> (2005).
39. Le, T. H. GSTM1 gene, diet, and kidney disease: implication for precision medicine?: Recent advances in hypertension. *Hypertension* **78**, 936–945. <https://doi.org/10.1161/HYPERTENSIONAHA.121.16510> (2021).
40. Zambrano, S. et al. L-carnitine attenuates the development of kidney fibrosis in hypertensive rats by upregulating PPAR- γ . *Am. J. Hypertens.* **27**, 460–470. <https://doi.org/10.1093/ajh/hpt268> (2014).
41. Daza-Arnedo, R. et al. Dipeptidyl peptidase-4 inhibitors and diabetic kidney disease: A narrative review. *Kidney Med.* **3**, 1065–1073. <https://doi.org/10.1016/j.xkme.2021.07.007> (2021).
42. Jiang, W., Roemer, M. E. & Newsham, I. F. The tumor suppressor DAL-1/4.1B modulates protein arginine N-methyltransferase 5 activity in a substrate-specific manner. *Biochem. Biophys. Res. Commun.* **329**, 522–30. <https://doi.org/10.1016/j.bbrc.2005.01.153> (2005).
43. Singh, V. et al. DAL-1/4.1B tumor suppressor interacts with protein arginine N-methyltransferase 3 (PRMT3) and inhibits its ability to methylate substrates in vitro and in vivo. *Oncogene* **23**, 7761–71. <https://doi.org/10.1038/sj.onc.1208057> (2004).
44. Kato, M. & Natarajan, R. Epigenetics and epigenomics in diabetic kidney disease and metabolic memory. *Nat. Rev. Nephrol.* **15**, 327–345. <https://doi.org/10.1038/s41581-019-0135-6> (2019).
45. Kim, J. E. et al. Multisample mass spectrometry-based approach for discovering injury markers in chronic kidney disease. *Mol. Cell. Proteomics* **20**, 100037. <https://doi.org/10.1074/mcp.RA120.002159> (2021).
46. Fernandes, M. & Husi, H. Establishment of an integrative multi-omics expression database CKDdb in the context of chronic kidney disease (CKD). *Sci. Rep.* **7**, 40367. <https://doi.org/10.1038/srep40367> (2017).
47. Johnson, T. S. et al. The role of transglutaminase in the rat subtotal nephrectomy model of renal fibrosis. *J. Clin. Invest.* **99**, 2950–2960. <https://doi.org/10.1172/JCI119490> (1997).
48. Lee, V. W. S. & Harris, D. C. H. Adriamycin nephropathy: a model of focal segmental glomerulosclerosis. *Nephrology (Carlton)* **16**, 30–38. <https://doi.org/10.1111/j.1440-1797.2010.01383.x> (2011).
49. Atkinson, J., Boden, T., Mocho, J.-P. & Johnson, T. Refining the unilateral ureteral obstruction mouse model: No sham, no shame. *Lab. Anim.* **55**, 21–29. <https://doi.org/10.1177/0023677220909401> (2021).
50. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265–275 (1951).
51. Lample, S. et al. Improved workflows for high throughput library preparation using the transposome-based nextera system. *BMC Biotechnol.* **13**, 104. <https://doi.org/10.1186/1472-6750-13-104> (2013).
52. Wu, D. & Smyth, G. K. Camera: A competitive gene set test accounting for inter-gene correlation. *Nucl. Acids Res.* **40**, e133–e133. <https://doi.org/10.1093/nar/gks461> (2012).
53. Furini, G. et al. Proteomic profiling reveals the transglutaminase-2 externalization pathway in kidneys after unilateral ureteric obstruction. *J. Am. Soc. Nephrol.* **29**, 880–905. <https://doi.org/10.1681/ASN.2017050479> (2018).
54. Tyanova, S. et al. The perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740. <https://doi.org/10.1038/nmeth.3901> (2016).
55. Wickham, H. *ggplot2* (Springer International Publishing, 2016). <https://doi.org/10.1007/978-3-319-24277-4>.
56. Lex, A., Gehlenborg, N., Strobelt, H., Vuilleumot, R. & Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992. <https://doi.org/10.1109/TVCG.2014.2346248> (2014).
57. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucl. Acids Res.* **50**, D543–D552. <https://doi.org/10.1093/nar/gkab1038> (2022).
58. Edgar, R., Domrachev, M. & Lash, A. E. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucl. Acids Res.* **30**, 207–210. <https://doi.org/10.1093/nar/30.1.207> (2002).
59. Barrett, T. et al. NCBI GEO: Archive for functional genomics data sets—update. *Nucl. Acids Res.* **41**, D991–D995. <https://doi.org/10.1093/nar/gks1193> (2013).

Author contributions

K.B., E.T., J.A., M.S., T.S.J., E.V. and B.T. conceived and designed the research. E.T., C.C., L.H. and B.A. per-

formed the experiments. K.B., E.T., A.S., D.J.B, J.S., E.V. and B.T. analysed and interpreted the data. K.B., E.T., E.V., T.S.J., and B.T. wrote the paper. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare that they have no conflict of interest, including financial and non-financial.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-93894-6>.

Correspondence and requests for materials should be addressed to K.B., E.A.M.V. or B.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025