955 Main Street, Buffalo, New York, 14203

Tel: 716-829-2265; E-mail: pinakisa@buffalo.edu

Abstract

Background: ~~Type 2~~ Diabetic Nephropathy (~~T2~~DN), in the setting of type 2 diabetes, is the world's leading cause of chronic kidney disease and end-stage kidney disease (ESKD). The increasing prevalence and heterogeneous phenotype of ~~T2~~DN complicate the approach to treating patients. While kidney biopsy is the gold standard for (1) exclusion of non-DN diagnoses and (2) confirming diagnosis of DN, it is imperfect in predicting progression to ESKD. Artificial intelligence (AI) has the potential to improve classification of ~~T2~~DN, predict progression risk, and via integration with urinary proteomic profiles identify novel urinary biomarkers, taken together augmenting and going beyond current pathology practice.

Methods: We studied whole slide images (WSIs) of periodic acid-Schiff stained renal biopsies from 56 DN patients with associated urinary proteomics data ~~for 2038 proteins~~. We identified urinary proteins differentially expressed in patients that developed ESKD within two years of biopsy. Six renal subcompartments were computationally segmented from each WSI using our previously published H-AI-L pipeline. Hand-engineered image features for glomeruli and tubules, and urinary proteins were used as inputs to deep learning frameworks to predict patient's ESKD status two years after biopsy. Differentially measured proteins were spatially registered using available snRNA-seq data from early stage diabetics, and were correlated with digital image features from the same spatial compartment using Spearman rank sum coefficient.

Results: Differentially abundant urinary proteins were most predictive of ESKD, with AUC = 0.95, with weaker performance of tubular and glomerular features, AUC = 0.73 and AUC = 0.66, respectively. Additionally, some proteins were associated in renal parenchymal cell types/compartments correlated with differential digital image features (i.e., EGF with tubular basement membrane contrast, and C7 with glomerular PAS-stain contrast).

Conclusions: By integrating urinary biomarkers with image biomarkers we demonstrate that

urinary proteins can robustly predict progression to ESKD, outperforming tubular and glomerular computational features. Using computational methods like ours can be useful more broadly to investigate links between molecular profiles and image features to improve pathophysiological understanding of the nature of DN progression with translational potential.

Introduction

~~Type 2~~ Diabetic Nephropathy (~~T2~~DN), in the setting of type 2 diabetes, is the leading driver of chronic kidney disease (CKD) and end-stage kidney disease (ESKD) worldwide[1]. The incidence of DN is increasing due to high prevalence of diabetes, with subsequent kidney complications such as proteinuria and a decline in kidney function affect more than 40% of the patients[2-4]. The result is an increased burden of cardiovascular morbidity and mortality[5], but the current therapeutic tools including lifestyle modification, blockers of renin-angiotensin-aldosterone system and SGLT2 inhibitors are insufficient to manage the rising incidence of DN and comorbidities. The heterogeneous phenotype of DN further complicates a comprehensive approach, with a subset of patients have an abrupt decline in kidney function requiring dialysis within a few years of diabetes diagnosis[6].

Kidney biopsy is the gold standard for the diagnosis of kidney disease, and histology aids in predicting the outcome and/or response to therapy[7]. The utility of a renal biopsy in the setting of diabetes, beyond staging of DN, is critical to identify non-diabetic processes (up to 30% of cases) [8,9] with a different etiology and disease trajectory compared with DN [10,11]. For improved classification and staging of DN, the Renal Pathology Society generated a DN classification schema[12]. This system reflects the course of progressive DN well[13,14], but has been hindered by the lack of practical application and implementation[15]. Recent approaches have interrogated molecular profiles with classical grading systems of tissue histology[16-18], but this has not yet been reported for the RPS DN classification schema.

Artificial intelligence (AI) has potential to classify histology of kidney diseases in an automated fashion, which reduces the burden on pathologists and promises to improve reproducibility and robustness[19]. Our previous work demonstrated digital classification of diabetic glomerulopathy in renal biopsies [20,21] where our models achieved high levels of

agreement with routine pathologist classification. To demonstrate the value added by these computational techniques, AI-based supervision for morphological changes related with molecular profiles or patient outcomes but imperceptible to the human eye are a promising next step [22-24]. Herein, we integrated urinary proteomic profiles with AI-based histology images from biopsy-confirmed DN cases. The urinary proteomic model had a greater performance than the AI-based image feature models in predicting early deterioration of kidney function. Several significant proteins and/or relevant pathways were associated with histological changes at the pixel and morphological level. These results suggest that the urinary proteomic dataset could guide better annotation of meaningful image features by AI, which were both within and beyond range of human eye. Based on the results, AI will enable the extraction of histologic data related with encoded molecular implications, as the next revolution in the histology-guided care of DN and generate new hypothesis for discovering novel pathobiology with diagnostic and prognostic implications pertaining to diabetes.

Results

Renal Tissue Multicompartment Segmentation

From each subject, one PAS-stained renal tissue biopsy whole slide image (WSI) was segmented into 6 classes: image background, renal interstitium (excluding perivascular stroma), non-sclerotic glomeruli, globally sclerotic glomeruli, tubules, and arteries/arterioles. The segmentation was first performed by our trained Panoptic segmentation network, and each slide was then manually corrected for any errors or missed segmentations. Examples of corrected network predictions are shown in ***Fig. 1.*** For this study, only cortical regions of the biopsies were used for analyses. Therefore, medullary regions were manually annotated and excluded from further analyses. Panoptic Quality (PQ)[25,26], was measured for segmentation of glomeruli, globally sclerotic glomeruli, and tubules in cortical regions. On average, we measured PQ = 0.76, 0.79, and 0.89 for segmenting glomeruli, globally sclerotic glomeruli, and tubules, respectively.

Renal Tissue Sub-Compartmentalization

Glomerular sub-compartmentalization was done similarly to the method described by Ginley *et al.*[20] in our earlier work. Following the same motivation, tubular sub-compartmentalization was performed to simplify the description of the structure. Each instance of glomerulus and tubule was further segmented into three components: (1) nuclear component, (2) PAS-positive component, and (3) white space component. Examples of the tubular sub-compartmentalization are shown in ***Fig. 2.*** This method of compartmentalization has been shown previously to be useful in classifying patients into stages of DN, by using features extracted from the compartmentalized glomeruli[20]. However, this method is novel for tubules, and in predicting the progression of DN. We assessed the validity of these methods in this work, but did not analyze artery and arteriolar morphometrics.

Renal Tissue Image Feature Quantification

Hand-engineered digital image features were measured from each compartmentalized instance of glomeruli and tubules. The features measured per glomerulus have been previously described by Ginley *et al.*[20] in our earlier work. New features were also defined in addition to our existing set, including component-specific distance metrics, such as quantiles of pixel distance from the glomerular border. In total, there were 315 digital image features measured for each glomerulus. The full feature definitions are listed in ***Supp. Doc. 2.***

Similar digital image features were measured for each renal tubule. From the three components segmented from each tubule (see *Renal Tissue Sub-Compartmentalization*), features were measured on pixel colors, textures, morphology, containment, and inter- and intra-structural distances. Examples of these feature extractions are shown in ***Fig. 3.*** One unique aspect of this set was the features added to measure more specific morphology of the tubules. While glomeruli typically appear circular in histological sections, tubules can have a highly variable shape. To account for this variance, we measured tubular curl[27], major axis length, perimeter, and area. These features can describe the tortuosity and thickness of each tubule. Also, features were defined to quantify the morphology of the tubular basement membrane (TBM). During feature extraction, the TBM was segmented from the rest of the PAS+ area, and color, textural, and size features were measured using the resulting segmentation. Some examples of TBM thickness quantifications are shown in ***Fig. 3.*** The TBM has previously been shown to display renal damage in patients with diabetic kidney disease[28]. In total, 207 digital image features were measured for each segmented tubule. These features are listed in ***Supp. Doc. 2.***

Urinary Proteomics Data Feature Manifold Classification

The feature sets used in this study were initially analyzed using Seurat[29,30]. Each feature set (urinary proteins, glomerular image features, tubular image features) was visualized separately

using Uniform Manifold Approximation and Projection (UMAP) to project the high-dimensional feature-space to a 2-dimensional plot. For patient-wise visualization of image feature sets in UMAP space, each feature was averaged across all object instances (namely, for all the segmented glomeruli or tubules) in a single biopsy. An SVM was trained on each feature set to classify patients on whether they progressed to ESKD within 2 years of biopsy. To account for the imbalance in outcomes, the class weights were adjusted inversely proportional to class frequencies[31]. For aggregated glomerular image features, the SVM achieved Matthews Correlation Coefficient (*MCC*), $MCC = 0.31$. For tubular image features, we achieved $MCC = 0.39$, and for the full urine protein set, $MCC = 0.60$. The simplified feature spaces and SVM hyperplanes are shown in ***Fig. 4***.

Using Seurat, we were also able to identify 45 proteins that were differentially measured ($p < 0.05$) in patients that progressed to ESKD within 2 years. Cluster biomarkers were found using the *FindMarkers* function. Of these proteins, several were measured to be significantly higher in patients progressing to ESKD (***Table 1A***), and others were measured to be significantly lower (***Table 1B***). After reducing the number of features in the urine protein set, we retrained the SVM, and achieved an improved $MCC = 0.72$.

Mapping Urinary Proteomics Features to Pertinent Renal Cells using Single-Nuclei RNA Sequencing

The percentage of cells and average expression for each of the 45 differentially measured urinary proteins were quantified using single-nuclei RNA (snRNA) sequencing data for both non-diabetic and diabetic kidney tissues ($n = 6$). Cell type candidate biomarkers, and protein expression for the differentially measured urinary proteins are shown in ***Fig. 5***. A total of 11 clusters including proximal convoluted tubule, parietal epithelial cell, thick ascending loop, distal convoluted tubule, collecting duct (principal cell, intercalated cell type A and B),

podocyte, endothelial cell, mesangial cell & fibroblast, and immune cells were identified by cell marker genes, and sub-clustering of mesangial cell and fibroblast yielded three subclusters (fibroblast, myofibroblast, and mesangial cell, ***Fig. 5A-B)***. The gene expression of urine proteins increased and decreased in progressed patients is summarized in ***Fig. 5C.*** This result provides insights on which protein types are expressed in renal cell types, since it is hard to precisely localize urinary proteins. Additionally, we quantified the changes in protein expression between non-diabetic and diabetic cases. For example, there was a clear increase in the number of fibroblasts expressing the *C7* gene in diabetic kidneys, and it can be visualized in the UMAP plot with subclustered mesangial cells and fibroblasts (***Fig. 5E***). This relationship was further explored in the study below correlating image and molecular features.

Disease Progression Prediction using AI

AI was implemented to predict disease progression in patients with DN because of its unique ability to detect and model nonlinear relationships[32], which is important for studying nonlinear disease progression such as that of CKD[33].

We used two neural network architectures to predict progression in DN patients. Urinary proteins are patient-level measurements, and therefore these features can be used directly to produce a classification through a series of dense layers, in a Fully Connected Neural Network (FCNN). However, the digital image features measured in this study are measured from multiple instances (namely, multiple segmented glomeruli or tubules) per patient, and must be aggregated to form a single prediction for a patient. We therefore used a Recurrent Neural Network (RNN) to incorporate all glomeruli or tubules to form predictions following recipes discussed in our previous publication[20]. Since some of our data is unlabeled, we trained our networks using a self-training semi-supervised scheme[34]. The results are aggregated from each patient when used in the holdout set during 10-fold cross-validation. We compared the SoftMax

predictions of each network to the patient's ground truth label, and assessed the performance with the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. These results are shown in ***Fig. 6***.

We achieved the best performance when training on differentially measured urinary proteins in an FCNN, with $AUC = 0.95$. This score indicates nearly perfect agreement between the classifier network and the ground truth label. Between the image feature types, those measured for tubules outperformed glomeruli. Tubular image features achieved $AUC = 0.73$, while glomerular image features achieved $AUC = 0.66$. The imbalance between the number of glomeruli and tubules present in a biopsy could explain the difference in network performance, but these biopsies are standard for pathologists' diagnoses. Therefore, tubular features could be considered better indicators of disease progression.

Correlation of Renal Tissue Morphometry and Urinary Proteomics Data

Using Spearman's rank correlation coefficient, we measured the degree of correlation between digital image features and urinary protein measurements, while controlling for confounding variables. We were interested in understanding how a urinary proteome may be reflected in tissue sections and analyzing the significance of these relationships. Each digital image feature across all structures in each patient subject was matched with the subject's urinary proteomic profile, and with their associated molecular pathway scores. The measured correlation coefficients were aggregated into the heatmaps shown in ***Fig. 7***.

The associated heatmap color represents the degree of correlation between an image feature and a urinary protein or pathway, indicated by the colorbar. On average, correlation coefficients with glomerular image features were greater than those with tubules, but *p*-values were lower for tubules, likely due to differences in sample size. We were most interested in the correlations with the largest magnitude, and these molecular-image pairs were ranked accordingly. A full

list of the corresponding image features, along with the corresponding urinary proteins, and pathways, is listed in ***Supp. Doc. 2.***

Discovery via Renal Tissue Image Pixel Parsing

Of the many Spearman rank correlation coefficients measured between image features and urinary proteomics molecular data, we were most interested in those with greatest magnitude. The goal is to identify the image features and proteins/pathways that may have some direct or indirect relationship.

As a select example, when looking at the correlation coefficients calculated between non-sclerotic glomerular image features and urinary proteins, the most highly correlated pair consisted of the protein Complement C7, and the standard deviation of green pixel values in PAS-positive region, with a coefficient value of -0.43 ($p < 0.05$). Along with this image feature, Complement C7 was also more highly correlated with the standard deviations of red and blue pixel values in PAS-positive regions, with coefficient values of -0.39 and -0.37 ($p < 0.05$), respectively. These pixel values are directly related to the perceived brightness in an image, and of the three channels, the green channel has the most significant effect[35]. Furthermore, the standard deviation of the pixel values in an image is related to the image contrast[36], specifically within the PAS-positive region of the segmented glomerulus. Therefore, since this feature and protein have an inverse correlation, as the amount of C7 present in the urine increases, one should expect to see a decrease in this image feature. C7 is a protein expressed in renal fibroblasts, as demonstrated in our snRNA sequencing and urinary proteomics mapping data. C7 is one of the proteins of the membrane attack complex (MAC)[37], which may target the glomerulus based on this correlation. Furthermore, one can link this relationship to outcome, since C7 is also a urinary protein that is differentially elevated in patients developing ESKD within 2 years post-biopsy/urinalysis. Some image feature examples are shown in ***Fig. 8.***

When looking at correlation coefficients between tubular image features and urinary proteins, a few interesting cases arose. One of the most highly correlated pairs consisted of the protein Epidermal Growth Factor (EGF), and the standard deviation of green pixel values in the TBM regions, with a coefficient value of 0.21 ($p < 0.05$). Similar to the glomerular example, EGF was also relatively highly correlated with the standard deviations of the red and blue pixels in the TBM regions. As the amount of EGF present in the urine decreases, one should also expect to see a decrease in these features in the image data. EGF has previously been implicated in the development of AKI and CKD[38], and is expressed in renal tubular cells, as supported by our snRNA-seq results. Additionally, five of the other six most highly correlated feature-protein pairs involve the protein Osteopontin (SPP1). SPP1 is another protein that is found in renal tubular cells, as shown in both our snRNAseq results, and previous studies[39], while the receptor is found in immune cells[40]. These urinary proteins were elucidated by this study for their relations to renal morphology in histological images, but their specific roles in kidney diseases are currently not fully understood[41].

Discussion

The ability to predict disease progression in patients with DN is valuable in identifying at-risk patients. Identifying risk early in disease progression allows the clinicians to make informed decisions about screening and treatments for patients[42]. Using AI, we were able to show that urinary proteins can predict which DN patients would progress to ESKD within two years of urine collection. However, digital image features were proven to be of lesser value than urinary proteins. Between the two sets of digital image features, those measured at the individual tubules level still had higher prognostic power in DN progression than the previously published glomerular based features.

Biopsy is the current gold standard for diagnosis and outcome prediction of kidney disease, and predicting response to therapy. In this work, we examined the digital morphology of glomeruli and tubules as seen in renal biopsies. It is much more manageable for a pathologist to analyze individual glomeruli, since these structures are significantly less present in the biopsy as compared to tubules. The large number of tubules present in a biopsy makes it difficult for pathologists to integrate data over the entire WSI and make an objective prognosis. In this work, we suggest that tubular image features from a standard needle biopsy are more predictive of disease progression than glomerular features. Therefore, we also suggest that computational methods like those we have demonstrated provide an advantage over current diagnosis standards, since our methods can provide a reproducible and comprehensive view of the tubular morphology.

We have also demonstrated the ability to link urinary proteomics data to quantitative image features. Qualitative image analysis has long been the standard for assessing pathology, but our results show that proteomics are robust in assessing progression in DN. Therefore, discovering and investigating links between image features and molecular profiles can aid in further

research to increase our understanding of the nonlinear nature of DN. This methodology is also not limited to DN, or specific to renal pathology. Investigating these links further could lead to discoveries in the area of diabetic kidney disease. Integration of other data types are also applicable in this case, depending on investigative goals and resources.

As mentioned by Ginley *et al.*[20], actual compartmental distinction in both glomeruli and tubules is complex. This sub-compartmentalization does not fully reflect actual components of the glomerulus and tubule. The simplification of each structure into three components based on their colorimetric appearance is a limitation to our performance in quantifying disease progression. Also, we largely analyzed each data source (image and molecular) individually, and there is a potential value in integrating image and molecular data to make prognoses, which can be an area of future investigations. Lastly, our snRNA sequencing data comes from non-diabetics and early diabetic, while the urinary proteomics and histology are sampled from various stages of DN. Therefore, the sequencing data may not be fully reflective of the conclusions we have made. To mitigate this shortcoming, in the future we will focus on generating image, urinary proteomics, and snRNA sequencing data from same subject to tighten the conclusions made in this work.

Methods

Human Samples

Study subjects were from Seoul National University Hospital. Human data collection followed protocols approved by the Institutional Review Board at the Seoul National University (SNU) College of Medicine (H-1812-159-998), Seoul, Korea. All experiments were performed according to federal guidelines and regulations. Individual subject data contain patient demographic and medical history information, including height, weight, age, history of diabetes, stroke, and presence of hypertension. Blood tests were administered to the subjects at the time of biopsy. Serum creatinine was used to measure the estimated glomerular filtration rate (eGFR) at the time of biopsy using the CKD-EPI equation[43]. Patients were re-evaluated and new blood samples were collected at one- and two-years following the initial biopsy and urine collection, and eGFR was recalculated. For each timepoint, it was determined whether serum creatinine had doubled compared with baseline, whether eGFR had reached 50% of the baseline value, and whether the patient had reached end stage kidney disease (ESKD). A full description of patient baseline characteristics can be found in ***Supp. Doc. 1.***

Whole Slide Image Data

Image data for this study consisted of brightfield microscopy whole-slide images (WSIs) of periodic acid-Schiff (PAS)-stained renal biopsies from n = 56 human DN subjects. An additional n = 30 WSIs from human DN patients were added as an additional set, but outcome data was not available for these patients. Whole slide imaging was done using a digital slide scanner (Aperio AT2, Leica Biosystems, Wetzlar, Germany). Pixel resolution of the images is 0.25 μm per pixel. Each biopsy was graded by a pathologist for glomerular lesions, following the Renal Pathology Society DN scoring system[12]. Other features scored included for interstitial fibrosis and tubular atrophy (IFTA), interstitial inflammation, arteriolar hyalinosis,

and arteriosclerosis.

Urine Sample Preparation

For protein digestion, 1-2 ml urine sample from each subject was concentrated to 250 ul using a spin filter with a molecular weight cut-off of 3 kDa (Millipore, Billerica, MA). The protein concentration was measured according to the Bradford assay protocol (Bio-Rad protein assay kit, Bio-Rad, Hercules, CA). For protein digestion, 50 µg of urine proteins was precipitated by adding a 5-fold volume of ice-cold acetone. The precipitated samples were reconstituted in 50 µl of SDT buffer (2% SDS, 0.1 M dithiothreitol in 0.1 M Tris HCl; pH, 8.0). After being heated at 95°C, the denatured proteins were digested by a filter-aided sample preparation (FASP) method as previously described[44] with some modifications. Briefly, protein samples were loaded onto a 30K amicon filter (Millipore, Billerica, MA), and buffer was exchanged with UA solution (8M UREA in 0.1M Tris-HCl; pH, 8.5) via centrifugation. After three buffer exchanges with UA solution, the reduced cysteines were alkylated with 0.05 M iodoacetamide (IAA) in UA solution for 30 min at room temperature in the dark. Thereafter, UA buffer was exchanged for 40 mM ammonium bicarbonate (ABC) twice. The protein samples were digested with trypsin/LysC (enzyme to substrate ratio of 1:100) at 37°C for 16 h. The resulting peptides were collected in new tubes via centrifugation, and an additional elution step was performed using 40 mM ABC and 0.5 M NaCl. All resulting peptides were acidified with 10% trifluoroacetic acid and desalted using a homemade styrene divinylbenzene reversed-phase sulfonate (SDB-RPS)-StageTips[45,46]. The peptides were initially washed with 0.2% TFA, and were then sequentially eluted with 40, 60, and 80% acetonitrile (ACN) containing 1% ammonia. Fractionated peptides were completely dried with a vacuum dryer and stored at -80°C.

Establishment of a Matching Spectral Library

To construct a matching spectral library for matching between runs, pooled urine samples were

digested using the two-step filter-aided sample preparation as described previously[46,47]. Digested peptides were desalted using Oasis HLB solid-phase extraction (SPE). For the in-depth data set, 100 µg of purified peptides were fractionated using an Agilent 1260 bioinert HPLC (Agilent, Santa Clara, CA) equipped with an analytical column (4.6 × 250 mm, 5-µm particle). High-pH reversed-phase peptide fractionation was performed at a flow rate of 0.8 ml/min over a 60-min gradient using solvent A (15 mm ammonium hydroxide in water) and solvent B (15 mM ammonium hydroxide in 90% acetonitrile). A total of 96 fractions was collected each minute and non-contiguously pooled into 24 fractions. The fractions were dried in a vacuum centrifuge and stored at -80°C until liquid chromatography-tandem mass spectrometry (LC-MS/MS).

LC-MS/MS Analysis

LC-MS/MS analysis was performed using Quadrupole Orbitrap mass spectrometers, Q-exactive plus (Thermo Fisher Scientific, Waltham, MA) coupled to an Ultimate 3000 RSLC systems (Dionex, Sunnyvale, CA) with a nano-electrospray source as previously described with some modifications[44,45]. Peptide samples were separated on a two-column setup with a trap column (300 µm I.D. × 0.5 cm, C18 3 µm, 100 Å) and analytical column (50 µm I.D. × 50 cm, C18 1.9 µm, 100 Å). Prior to sample injection, the dried peptide samples were redissolved in solvent A (2% acetonitrile and 0.1% formic acid). After the samples were loaded onto the nano LC, a 90-min gradient from 8% to 30% solvent B (100% acetonitrile and 0.1% formic acid) was applied to all samples. The spray voltage was 2.0 kV in positive ion mode and the temperature of the heated capillary was set to 320°C. Mass spectra were acquired in data-dependent mode using a top 15 method on a Q-Exactive. The Orbitrap analyser scanned precursor ions with a mass range of 300–1650 m/z and resolution of 70,000 at m/z 200. Higher-energy collisional dissociation (HCD) scans were acquired on the Q-Exactive at a resolution

of 17,500. HCD peptide fragments were acquired at a normalized collision energy of 28. The maximum ion injection times for the survey and MS/MS scans were 20 and 120 ms, respectively.

Data Processing for Label-Free Quantification

Mass spectra were processed with MaxQuant (version 1.6.1.0)[48]. MS/MS spectra were searched against the Human Uniprot protein sequence database (December 2014, 88,657 entries) using the Andromeda search engine[49]. Primary searches were performed using a six-ppm precursor ion tolerance for total protein level analysis. The MS/MS ion tolerance was set to 20 ppm. Cysteine carbamido-methylation was set as a fixed modification. *N*-Acetylation of proteins and oxidation of methionine were set as variable modifications. Enzyme specificity was set to full tryptic digestion. Peptides with a minimum length of six amino acids and up to two missed cleavages were considered. The required false discovery rate (FDR) was set to 1% at the peptide, protein, and modification levels. To maximize the number of quantification events across samples, matching between runs was performed using the pooled urine sample as a library.

Human AI Loop (HAIL) Pipeline

To conduct computational segmentation of renal micro-compartments, we used our previously published, publicly available HAIL pipeline[50]. This pipeline allows users to train segmentation CNNs directly from annotations produced in the WSI viewer Aperio ImageScope®. HAIL allows for an iterative process of network training, where the network's own predictions produced after each training iteration can be manually corrected and used to bootstrap further annotation data.

Panoptic Segmentation

Panoptic segmentation[26] was performed on PAS stained WSIs to classify pixels into six

categories, and to resolve separate instances (e.g., individual tubules) of the same class. These six categories include image background, renal interstitium (excluding perivascular stroma), non-sclerotic glomeruli, globally sclerotic glomeruli, renal tubules, and arteries/arterioles. To train the model, 126 WSIs of kidney biopsy from native diabetic, lupus nephritis, and transplant patients were selected due to the presence of minimal cortex (<10%) in the slide, and fully annotated for all six categories. The network architecture was modelled closely after Google's Panoptic-Deeplab model[51]. There are three primary differences between our network model, and DeepLab-Panoptic. First, to reduce the memory overhead, the feature encoding backbone used is a ResNet50[52]. Second, the output stride of our encoder is only 8 as opposed to 16 as used in the DeepLab-Panoptic model. Finally, for the output of our instance decoder branch, rather than predicting a Hough transform, this branch predicts the distance transform of each object. Then, instance segmentation is performed by computing a watershed transform[53] on the predicted output distance transform map. Classification IDs of the predicted objects are settled with the majority voting rule used in Panoptic-DeepLab. The network was trained on cropped image patches of 560 × 560 pixel size and batch size 2 for approximately 3.8M steps. The trained network was then used for prediction where medullary regions were manually annotated and excluded from the analysis. Finally, glomeruli and tubules with artefact were manually removed from the final segmentations to eliminate erroneous quantification. Segmentation accuracy was assessed using the Panoptic Quality (PQ) metric[25], created by the original developers of Panoptic segmentation. This metric joins segmentation quality and recognition quality to assess both semantic and instance segmentation performance. PQ was only measured for segmentation of glomeruli, globally sclerotic glomeruli, and tubules, since these were the classes used in later analyses. We selected 5 random WSIs from patients that progressed to ESKD within 2 years and 5 random WSIs from patients that did not, and these

images were used for performance analysis. Our manually corrected annotations for these slides were further corrected by a pathologist, and used as a ground truth.

Glomerular and Tubular Sub-Compartmentalization

To simplify compartmentalization of the annotated regions, glomerular pixels are assigned to one of three components based on their appearance in the PAS-stained biopsies, according to Ginley *et. al*[20]: (1) nuclear components, (2) PAS-positive components, containing the glomerular basement membrane, mesangium, and Bowman's capsule, and (3) luminal components consisting of capillary lumina and the Bowman's space. For tubular sub-compartmentalization, the three components were similar.

Nuclei were detected in the tissue by a custom coded Panoptic segmentation network. This network was trained on previously annotated data described by Ginley *et. al*[20]. In both glomerular and tubular images, nuclei were segmented by the same network. In glomeruli, PAS+ components were segmented by a several step process. First, the original RGB image was transformed to the hue-saturation-value color space[54], the saturation channel was isolated, and the dark regions were brightened using a gamma transform[55]. For each glomerulus, this gamma was set at a constant of 0.7. Otsu's thresholding method[56] was then applied to segment the PAS+ area, excluding the already segmented nuclei. Once the nuclear and PAS+ components were segmented, the remaining pixels not contained in these components were classified as luminal pixels. PAS+ components of the tubules were segmented in a slightly different fashion. First, the RGB image was transformed to the LAB color space[57]. Second, the lightness color channel (L) was isolated and a constant threshold of 80 was applied to the 8-bit pixels. All pixels, not already classified as nuclei, above this threshold were classified as luminal components. All other pixels not contained in the nuclear or luminal segmentation were classified as PAS+ components.

Feature Engineering

For glomerular images, several digital image feature types are defined according to the methods described by Ginley *et al*[20]. These feature types included color, texture, morphology, containment, interstructural distances, and intrastructural distances. A total of 315 image features were measured for each glomerulus. For tubular images, similar digital features are measured for the same respective structures. However, additional features are defined for the tubular basement membrane, and to measure the tortuosity of the tubules. A total of 207 image features are measured for each tubule. A full list of both digital image feature sets can be found in ***Supp. Doc. 2.***

Image/ Proteomic Feature Manifold Classification

Image/proteomic feature manifolds were studied in low dimensional space using a state-of-the-art software Seurat[29,30]. Seurat is typically used for studying single-cell RNA-sequencing (scRNAseq) data, but can be extended to other high-dimensional data, such as digital image features and urinary proteomics. For image feature analysis, each datapoint is modelled as a computationally segmented individual microcompartment (e.g., tubule, glomerulus), and corresponding feature values are quantified engineered feature data. For urinary proteomics analysis, each datapoint is modelled as a single subject, and corresponding feature values are measured urinary proteomics features. As a first step, the data is normalized by a global-scaling normalization method. This method normalizes each feature by min-max scaling values between 0-1, multiplying by a scale factor, and then log-transforming the result. The data is then scaled to zero mean and unit variance. Principal component analysis was then performed on the scaled data. Non-linear dimensionality reduction using UMAP (Uniform Manifold Approximation and Projection)[58] is performed using the first 20 principal components per datapoint. Class labels (e.g., subjects with two-year ESKD vs the rest) of interest are applied

to these points following dimensionality reduction.

We analyzed differentially measured image or proteomics between various data classes to characterize these classes. Features that are differentially expressed between classes are discovered using the *FindMarkers* function in Seurat, which uses the Wilcoxon rank sum test[59] to determine statistical significance. This function also automatically performs *p*-value adjustment using Bonferroni correction, based on the number of features in the dataset. Features with an adjusted $p < 0.05$ were identified as differentially expressed.

To classify subjects based on two-year ESKD binary outcome measure, we performed binary classification of the UMAP feature data using a support vector classifier[60]. One-dimensional hyperplanes were optimized using a radial basis function kernel, with balanced class weight, to classify the subjects. The performance of the classifier was quantified using Matthews correlation coefficient[61].

Mapping Urinary Proteins to Source Renal Parenchymal Cells using Single-Nuclei RNA Sequencing Data

While fusing urinary proteomics data and digital renal tissue image features is an important goal of our study, it is important to map the urinary proteomics data to pertinent renal cell types to investigate biological relevance of our findings. For this mapping, as a quality control, pooled single-nuclei RNA sequencing (snRNAseq) data of early diabetic (n = 3) and nondiabetic (n = 3) patients from a previous study[62] were used. Genes expressed in ≥3 cells, cells with ≥500 unique molecular identifier counts were used. The R[63] and Seurat[64] software packages as well as pheatmap (version 1.0.12) package embedded in R were used for data analyses and visualization. For the mapping the following functionalities from these software packages were used: *SCTransform* for merging data with adjusting batch-effects; *FindNeighbors* and *FindClusters* for finding clusters; *RunUMAP, FeaturePlot, DotPlot,* and

pheatmap for data visualization; and *Findmarkers* for identifying differentially expressed genes.

Disease Progression Prediction using Urinary Proteins by Semi-Supervised Fully Connected Neural Network

For each patient, 2038 unique urinary proteins were measured at the time of the biopsy, and for the additional (n = 30) set, 2313 urinary proteins. To predict patient progression to ESKD, we designed a fully connected neural network (FCNN). For all the AI networks, a binary outcome for ESKD progression within two years of biopsy was used as ground truth. Due to the small sample size, using the entire protein set caused overfitting of the network. Therefore, only differentially measured proteins are used as input. The network architecture consists of an input layer for protein data, then a series of two dense layers, each with 30 hidden nodes. These dense layers both contain leaky rectified linear unit (ReLU) activation functions, and 50% dropout during training. The last dense layer connects to a two-node prediction layer, with a SoftMax activation function. Cross-entropy was used to compute the loss for this network. The network was trained with an Adam optimizer, with the learning rate set at 0.001.The network was trained using self-supervised semi-supervision[34] with n = 56 labelled cases and n = 30 unlabelled cases. First, the labelled dataset was split into 90% training and 10% testing, and the network was trained using just this training set for 500 iterations with batch size 50. Then, predictions were produced for the additional dataset of n = 30 that did not have any outcome label. The SoftMax output was clipped at a threshold over 0.90, so any network predictions with more than 90% confidence were kept, and assigned the label produced by the network. Next, the newly labelled cases were added to the original set, and a single training step of batch 50 was run. Then, predictions are reproduced for the second dataset of n = 30, and the process is repeated for a total of 100 iterations. Lastly, predictions are run on the 10% holdout labelled set, and compiled

into the final predictions. The network is then reset, and a different 10% holdout is taken for testing. This was repeated for 10 folds and the results reported are compiled from 100 trials of 10-fold cross validation. Performance was calculated using the area under the receiver operator characteristic curve. The generalized network architecture is shown in ***Supp. Doc. 1.***

Disease Progression Prediction using Image features by Semi-Supervised Recurrent Neural Network

To predict DN progression in our cohort using quantified image features from glomerular or tubular compartments, we employed the recurrent neural network (RNN) architecture designed by Ginley *et al*[20]. The goal is to predict an outcome label for a given WSI by incorporating at the network input the ensemble of quantified image features per compartment basis serially with respect to all the glomeruli or tubules in a biopsy. The main modification of the network architecture with that described in Ginley *et al.*[20] is the final prediction layer, where we use two prediction nodes for our binary outcome (ESKD vs no-ESKD in two years after biopsy), with a SoftMax activation. All training parameters, including learning rate, batch size, and training steps were kept the same. Semi-supervised training was coded similarly to the FCNN above. First, 500 training steps of batch size 256 and 10-time steps were completed with 90% of the labelled dataset. Predictions were run on the unlabelled patients, and those with prediction confidence > 75% were kept, assigned the predicted label, and added back to the training set. Then, 100 additional training steps are run with the expanded training set, and predictions on all the unlabelled data are rerun. This process is repeated 5 times for 500 additional training steps. Final predictions are compiled from 10 trials of 10-fold cross-validation. The generalized network architecture is shown in ***Supp. Doc. 1.***

Correlation of Renal Tissue Morphometry and Molecular Data

Our previous experiments investigate how urine proteins and digital image features of renal

structures relate to DN outcome. We next sought to understand how these two orthogonal data modes are related. These relationships were quantified using partial Spearman's rank correlation coefficient, which controlled for several covariates, including sex, age, height, weight, hypertension, number of years with diagnosed diabetes, and history of stroke and ischemic heart disease. For this analysis, we used *partialcorr* command embedded in MATLAB (Mathworks, Natick, MA)[65]. Coefficients were measured for one-to-one relationships between quantified structural image features and protein expression or scores of molecular pathways formed by an ensemble of proteins. We adjusted the *p*-values of the correlation measures using the Benjamini-Hochberg method[66].

For the above correlative study, to quantify the functionality of the measured urinary proteins at pathway level, we developed a scoring method for upregulation and downregulation of molecular pathways, using Ingenuity Pathway Analysis (IPA; Qiagen, Hilden, Germany)[67]. First, the proteins included in each pathway were identified. Then, the raw protein measurements were added by one, and the resulting values were log-transformed. The resulting data were then min-max scaled between 0-1 per protein basis. Following scaling, values were adjusted by their biological expectation of upregulation/downregulation in pertinent pathways, according to the analytical algorithms embedded in IPA. Namely, the scores of the proteins that were expected to be upregulated during the upregulation of a particular pathway were kept the same. Scores of the proteins that were expected to be downregulated during the upregulation of a particular pathway were revised by subtracting their pertinent values from 1. This was done so that all protein values would increase/decrease the same during upregulation/downregulation of a pathway, respectively. Lastly, the scaled values of the proteins for each pathway were summed for each patient, to obtain one score per pathway.

Discovery via Renal Tissue Image Pixel Parsing

Of the most highly correlated pairs (structural image features vs protein expression or pathway scores), we focused our studies on proteins or molecular pathways with biological relevance to renal pathology. This focus limited the proteins of interest to those expressed in renal cell types, and molecular pathways of interest to those involving aspects of renal function. Distributions across patients were investigated for structural image features highly correlated with these biologically relevant molecular candidates. Discovered image features were visualized in MATLAB, via mapping the quantified image features in the image space. Projected maps were qualitatively studied by renal pathology experts, and relationships with compelling trends were used for future hypothesis generation.

Data Availability

Codes are available at [insert github link here].

Author Contributions

NL corrected computational segmentations, conceptualized and performed the quantitative analyses, designed, and conducted the computational methods, completed all statistical analyses, interpreted the results, and wrote the manuscript. DY …. DH …. BG guided NL in conducting the work, reviewed the codes, as well as critically analyzed results. KCM …. AR and JET critically analyzed the results of the correlation study to decipher the biological meaning of the image level features corresponding to molecular markers pertinent from the urinary proteomic data. JZ contributed with the statistical analysis conducted in the study. KYJ generated ground-truth segmentation boundaries for analyzing the performance of the multi-compartment renal tissue segmentation. SSH co-conceived the overall study with PS integrating urinary proteomics data with renal tissue image data, optimized urinary proteomics data generation, as well as spearheaded the database generation with matching renal tissue whole slide image data as well as with corresponding outcome data. PS co-conceived the study with SSH, assisted in manuscript preparation, coordinated with the study team, assisted in study design, supervised the computational implementation, and critically analyzed the results.

Acknowledgement

The project was supported by NIDDK grant R01 DK114485 (PS), a glue grant (PS) from the NIDDK Kidney Precision Medicine Project, a multi-disciplinary small team grant RSG201047.2 (PS) from the State University of New York, a pilot grant (PS) from the University of Buffalo's Clinical and Translational Science Institute (CTSI) grant 3UL1TR00141206 S1, a DiaComp Pilot & Feasibility Project 21AU4180 (PS) with support from NIDDK Diabetic Complications Consortium grants U24 DK076169 and U24 DK115255,

NIH-OD grant U54 HL145608 (PS), and NIDDK grant R01 DK131189 (PS).

References

1 Koye, D. N., Magliano, D. J., Nelson, R. G. & Pavkov, M. E. The Global Epidemiology of Diabetes and Kidney Disease. *Advances in chronic kidney disease* **25**, 121-132, doi:10.1053/j.ackd.2017.10.011 (2018).

2 Doshi, S. M. & Friedman, A. N. Diagnosis and Management of Type 2 Diabetic Kidney Disease. *Clinical journal of the American Society of Nephrology : CJASN* **12**, 1366-1373, doi:10.2215/cjn.11111016 (2017).

3 Adler, A. I. *et al.* Development and progression of nephropathy in type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS 64). *Kidney international* **63**, 225-232, doi:10.1046/j.1523-1755.2003.00712.x (2003).

4 Fiorentino, M. *et al.* Renal biopsy in patients with diabetes: a pooled meta-analysis of 48 studies. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* **32**, 97-110, doi:10.1093/ndt/gfw070 (2017).

5 Tuttle, K. R. *et al.* Diabetic kidney disease: a report from an ADA Consensus Conference. *Diabetes care* **37**, 2864-2883, doi:10.2337/dc14-1296 (2014).

6 Jiang, G. *et al.* Progression of diabetic kidney disease and trajectory of kidney function decline in Chinese patients with Type 2 diabetes. *Kidney international* **95**, 178-187, doi:10.1016/j.kint.2018.08.026 (2019).

7 Luciano, R. L. & Moeckel, G. W. Update on the Native Kidney Biopsy: Core Curriculum 2019. *American journal of kidney diseases : the official journal of the National Kidney Foundation* **73**, 404-415, doi:10.1053/j.ajkd.2018.10.011 (2019).

8 Sharma, S. G. *et al.* The modern spectrum of renal biopsy findings in patients with diabetes. *Clinical journal of the American Society of Nephrology : CJASN* **8**, 1718-1724, doi:10.2215/CJN.02510213 (2013).

9 Mazzucco, G. *et al.* Different patterns of renal damage in type 2 diabetes mellitus: a multicentric study on 393 biopsies. *American journal of kidney diseases : the official journal of the National Kidney Foundation* **39**, 713-720, doi:10.1053/ajkd.2002.31988 (2002).

10 Wong, T. Y. *et al.* Renal outcome in type 2 diabetic patients with or without coexisting nondiabetic nephropathies. *Diabetes care* **25**, 900-905, doi:10.2337/diacare.25.5.900 (2002).

11 Li, L. *et al.* Renal pathological implications in type 2 diabetes mellitus patients with renal involvement. *Journal of diabetes and its complications* **31**, 114-121, doi:10.1016/j.jdiacomp.2016.10.024 (2017).

12 Tervaert, T. W. *et al.* Pathologic classification of diabetic nephropathy. *J Am Soc Nephrol* **21**, 556-563, doi:10.1681/ASN.2010010010 (2010).

13 An, Y. *et al.* Renal histologic changes and the outcome in patients with diabetic nephropathy. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* **30**, 257-266, doi:10.1093/ndt/gfu250 (2015).

14 Mise, K. *et al.* Renal prognosis a long time after renal biopsy on patients with diabetic nephropathy. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* **29**, 109-118, doi:10.1093/ndt/gft349 (2014).

15 Qi, C., Mao, X., Zhang, Z. & Wu, H. Classification and Differential Diagnosis of

Diabetic Nephropathy. *Journal of diabetes research* **2017**, 8637138, doi:10.1155/2017/8637138 (2017).

16 Loupy, A. *et al.* The Banff 2019 Kidney Meeting Report (I): Updates on and clarification of criteria for T cell- and antibody-mediated rejection. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* **20**, 2318-2331, doi:10.1111/ajt.15898 (2020).

17 Sethi, S. & Fervenza, F. C. Standardized classification and reporting of glomerulonephritis. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* **34**, 193-199, doi:10.1093/ndt/gfy220 (2019).

18 Gadegbeku, C. A. *et al.* Design of the Nephrotic Syndrome Study Network (NEPTUNE) to evaluate primary glomerular nephropathy by a multidisciplinary approach. *Kidney international* **83**, 749-756, doi:10.1038/ki.2012.428 (2013).

19 Huo, Y., Deng, R., Liu, Q., Fogo, A. B. & Yang, H. AI applications in renal pathology. *Kidney international*, doi:10.1016/j.kint.2021.01.015 (2021).

20 Ginley, B. *et al.* Computational Segmentation and Classification of Diabetic Glomerulosclerosis. *Journal of the American Society of Nephrology : JASN* **30**, 1953-1967, doi:10.1681/asn.2018121259 (2019).

21 Ginley, B. *et al.* Automated Computational Detection of Interstitial Fibrosis, Tubular Atrophy, and Glomerulosclerosis. *Journal of the American Society of Nephrology : JASN* **32**, 837-850, doi:10.1681/asn.2020050652 (2021).

22 Coudray, N. *et al.* Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature medicine* **24**, 1559-1567, doi:10.1038/s41591-018-0177-5 (2018).

23 Kather, J. N. *et al.* Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nature medicine* **25**, 1054-1056, doi:10.1038/s41591-019-0462-y (2019).

24 Mobadersany, P. *et al.* Predicting cancer outcomes from histology and genomics using convolutional networks. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E2970-E2979, doi:10.1073/pnas.1717139115 (2018).

25 Kirillov, A., He, K., Girshick, R., Rother, C. & Dollár, P. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 9404-9413.

26 Kirillov, A., He, K., Girshick, R. B., Rother, C. & Dollár, P. Panoptic Segmentation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9396-9405 (2019).

27 Wirth, M. A. (Purdue University Cytometry Laboratories, 2004).

28 Tyagi, I., Agrawal, U., Amitabh, V., Jain, A. K. & Saxena, S. Thickness of glomerular and tubular basement membranes in preclinical and clinical stages of diabetic nephropathy. *Indian J Nephrol* **18**, 64-69, doi:10.4103/0971-4065.42336 (2008).

29 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821, doi:10.1016/j.cell.2019.05.031 (2019).

30 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology* **36**, 411-420, doi:10.1038/nbt.4096 (2018).

31 Yi-Min, H. & Shu-Xin, D. in *2005 International Conference on Machine Learning and Cybernetics.* 4365-4369 Vol. 4367.

32 Tu, J. V. Advantages and disadvantages of using artificial neural networks versus

logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology* **49**, 1225-1231, doi:https://doi.org/10.1016/S0895-4356(96)00002-9 (1996).

33 Li, L. *et al.* Longitudinal Progression Trajectory of GFR Among Patients With CKD. *American Journal of Kidney Diseases* **59**, 504-512, doi:https://doi.org/10.1053/j.ajkd.2011.12.009 (2012).

34 Fazakis, N., Karlos, S., Kotsiantis, S. & Sgarbas, K. Self-Trained LMT for Semisupervised Learning. *Comput Intell Neurosci* **2016**, 3057481-3057481, doi:10.1155/2016/3057481 (2016).

35 Bezryadin, S., Bourov, P. & Ilinih, D. Brightness Calculation in Digital Image Processing. *International Symposium on Technologies for Digital Photo Fulfillment* **2007**, 10-15, doi:10.2352/ISSN.2169-4672.2007.1.0.10 (2007).

36 Moulden, B., Kingdom, F. & Gatley, L. The Standard Deviation of Luminance as a Metric for Contrast in Random-Dot Images. *Perception* **19**, 79-101, doi:10.1068/p190079 (1990).

37 Würzner, R. Modulation of complement membrane attack by local C7 synthesis. *Clin Exp Immunol* **121**, 8-10, doi:10.1046/j.1365-2249.2000.01263.x (2000).

38 Zepeda-Orozco, D., Wen, H. M., Hamilton, B. A., Raikwar, N. S. & Thomas, C. P. EGF regulation of proximal tubule cell proliferation and VEGF-A secretion. *Physiol Rep* **5**, e13453, doi:10.14814/phy2.13453 (2017).

39 Kaleta, B. The role of osteopontin in kidney diseases.

40 MacDonald, L. *et al.* COVID-19 and RA share an SPP1 myeloid pathway that drives PD-L1+ neutrophils and CD14+ monocytes. *JCI Insight* **6**, e147413, doi:10.1172/jci.insight.147413 (2021).

41 Isaka, Y. Epidermal growth factor as a prognostic biomarker in chronic kidney diseases. *Ann Transl Med* **4**, S62-S62, doi:10.21037/atm.2016.10.64 (2016).

42 Ahuja, K. & Schaar, M. V. D. Risk-Stratify: Confident Stratification Of Patients Based On Risk. *ArXiv* **abs/1811.00753** (2018).

43 Levey, A. S. *et al.* A new equation to estimate glomerular filtration rate. *Ann Intern Med* **150**, 604-612, doi:10.7326/0003-4819-150-9-200905050-00006 (2009).

44 Grzegorski, S. J. *et al.* Disruption of the kringle 1 domain of prothrombin leads to late onset mortality in zebrafish. *Scientific reports* **10**, 4049-4049, doi:10.1038/s41598-020-60840-7 (2020).

45 Rhee, S. J. *et al.* Comparison of serum protein profiles between major depressive disorder and bipolar disorder. *BMC Psychiatry* **20**, 145-145, doi:10.1186/s12888-020-02540-0 (2020).

46 Han, D., Jin, J., Woo, J., Min, H. & Kim, Y. Proteomic analysis of mouse astrocytes and their secretome by a combination of FASP and StageTip-based, high pH, reversed-phase fractionation. *PROTEOMICS* **14**, 1604-1609, doi:https://doi.org/10.1002/pmic.201300495 (2014).

47 Han, D., Jin, J., Yu, J., Kim, K. & Kim, Y. Integrated approach using multistep enzyme digestion, TiO2 enrichment, and database search for in-depth phosphoproteomic profiling. *PROTEOMICS* **15**, 618-623, doi:https://doi.org/10.1002/pmic.201400102 (2015).

48 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**, 2301-2319, doi:10.1038/nprot.2016.136 (2016).

49 Cox, J. *et al.* Andromeda: A Peptide Search Engine Integrated into the MaxQuant

Environment. *Journal of Proteome Research* **10**, 1794-1805, doi:10.1021/pr101065j (2011).

50 Lutnick, B. *et al.* An integrated iterative annotation technique for easing neural network training in medical image analysis. *Nature Machine Intelligence* **1**, 112-119, doi:10.1038/s42256-019-0018-3 (2019).

51 Cheng, B. *et al.* Panoptic-DeepLab: A Simple, Strong, and Fast Baseline for Bottom-Up Panoptic Segmentation. *arXiv e-prints*, arXiv:1911.10194 (2019).

52 He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. *arXiv e-prints*, arXiv:1512.03385 (2015).

53 Gonzalez, R. C. & Woods, R. E. *Digital Image Processing*. 3 edn, 1-976 (Prentice Hall, 2007).

54 Ganesan, P., Rajini, V., Sathish, B. S. & Shaik, K. B. in *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT).* 101-105.

55 Guan, X., Jian, S., Hongda, P., Zhiguo, Z. & Haibin, G. in *2009 Second International Symposium on Computational Intelligence and Design.* 60-63.

56 Bangare, S., Dubal, A., Bangare, P. & Patil, S. Reviewing Otsu's Method For Image Thresholding. *International Journal of Applied Engineering Research* **10**, 21777-21783, doi:10.37622/IJAER/10.9.2015.21777-21783 (2015).

57 Connolly, C. & Fleiss, T. A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *IEEE Transactions on Image Processing* **6**, 1046-1048, doi:10.1109/83.597279 (1997).

58 McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv e-prints*, arXiv:1802.03426 (2018).

59 Gallant, A. R. *Nonlinear Statistical Models*. (Wiley, 1987).

60 Hastie, T., Tibshirani, R., Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2 edn, 768 (Springer, 2009).

61 McHugh, M. L. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* **22**, 276-282 (2012).

62 Wilson, P. C. *et al.* The single-cell transcriptomic landscape of early human diabetic nephropathy. *Proceedings of the National Academy of Sciences* **116**, 19619, doi:10.1073/pnas.1908706116 (2019).

63 *The Comprehensive R Archive Network (V 4.0.3)*, <http://cran.r-project.org> (

64 *Seurat (V 3.2.3)*, <https://satijalab.org/seurat/> (

65 *MATLAB*, <https://www.mathworks.com/products/matlab.html> (

66 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).

67 *Ingenuity Pathways Analysis (IPA)*, <https://www.nihlibrary.nih.gov/resources/tools/ingenuity-pathways-analysis-ipa> (