

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies

scoMorphoFISH: A Deep-Learning enabled toolbox for single-cell single-mRNA quantification and correlative (ultra-)morphometry

Florian Siegerist¹, Eleonora Hay^{1,2}, Juan Saydou Dikou¹, Anja Büscher³, Jun Oh⁴, Silvia Ribback⁵, Uwe Zimmermann⁶, Jan Hinrich Bräsen⁷, Olivia Lenoir⁸, Karlhans Endlich¹, Pierre-Louis Tharaux⁸, Nicole Endlich^{1*}

¹Institute for Anatomy and Cell Biology, University Medicine Greifswald, Greifswald, Germany

²Department of Mental and Physical Health and Preventive Medicine, Section of Human Anatomy, University of Campania "Luigi Vanvitelli", Naples, Italy

³Department of Paediatrics II, University Hospital Essen, Essen, Germany

⁴Department of Paediatrics, University Hamburg-Eppendorf, Hamburg, Germany

⁵Department of Pathology, University Medicine Greifswald, Greifswald, Germany

⁶Department of Urology, University Medicine Greifswald, Greifswald, Germany

⁷Nephropathology, Institute of Pathology, Medical School Hannover, Hannover, Germany

⁸Université de Paris, INSERM, U970, PARCC, F-75015 Paris, France

*Address for correspondence:
Prof. Dr. rer. nat. Nicole Endlich
Institute for Anatomy and Cell Biology
University Medicine Greifswald
17487 Greifswald
Germany

Abstract

Increasing the information depth of single kidney biopsies can improve diagnostic precision, personalized medicine and accelerate basic kidney research. Until now, information on mRNA abundance and morphologic analysis has been obtained from different samples, missing out on the spatial context and single-cell correlation of findings. Herein, we present scoMorphoFISH, a modular toolbox to get spatial single-cell single-mRNA expression data optimized for routinely generated kidney biopsies. Deep-Learning was used to virtually dissect tissue sections in tissue compartments and cell types to which single-cell expression data was assigned. Furthermore, we show correlative and spatial single-cell expression quantification with super-resolved podocyte foot process morphometry on the same histological section. In contrast to bulk analysis methods, this approach will help to identify local transcription changes even in less frequent kidney cell types on a spatial single-cell level with single-mRNA resolution. As this method performs well with standard formalin-fixed paraffin-embedded samples and we provide pretrained DL-networks embedded in a comprehensive image analysis workflow, this method can be applied immediately in a variety of settings.

Abbreviations:

DL: Deep Learning

FFPE: formalin-fixed paraffin-embedded

scoMorphoFISH: single-cell correlative Morphometric single-mRNA fluorescence in situ hybridization

smFISH: single-molecule fluorescence in situ hybridization

WF: wide-field microscopy

3D-SIM: Three-dimensional structured illumination microscopy

Introduction

High-precision analysis of kidney biopsies is key to providing diagnosis and targeted therapies for patients. Recently, several methods have been established to improve the analysis depth of formalin-fixed paraffin-embedded (FFPE) kidney biopsies: The filtration barrier can be morphometrically analyzed by 3D structured illumination microscopy (3D-SIM)^{1–3}. The determination of the filtration slit density by PEMP (*podocyte exact morphology measurement procedure*) emerged as a tool that can be combined with co-staining of multiple proteins⁴. However, antibody-based quantification of local protein abundance has limitations as it depends on antibody availability and performance. Additionally, locally secreted factors are typically not captured by immunofluorescence techniques. The use of bulk proteomics and transcriptomics is limited since the expression of frequent cells like proximal tubule cells can mask transcriptional changes in less frequent cell populations. To circumvent this, tissue is manually dissected (e.g. glomeruli from tubulointerstitium)⁵, cell types are enriched by flow cytometry⁶ or single-cell-RNA sequencing is performed⁷. Unfortunately, the contextual and/or morphological information is lost in all approaches due to mechanical dissociation. As biopsy material is typically limited and interpretation in a spatial context required, correlation of multiple techniques on single sections could increase the depth of information.

An antibody-independent way to investigate spatial RNA abundance is *in situ* hybridization (ISH) which has been substantially improved in terms of sensitivity and multiplexing. Recently, diverse methods for single-mRNA visualization and quantification (smFISH) are available^{9–12}. A problem for inter-sample comparability is that smFISH highly depends on preparation-dependent RNA integrity. To rule out this problem, a stable on-slide in-cell reference gene would be required to normalize expression data for different parts of the same biopsy or even over different samples. To assign transcripts to tissue compartments and individual cell types, reliable identification and segmentation of cellular regions of interest (ROIs) is required. Unfortunately, correlative antibody-based cell classification is challenging as *smFISH* requires tissue digestion to liberate fixed mRNAs. Additionally, segmentation tasks are typical bottlenecks in image analysis workflows. To overcome this, Deep-Learning (DL) has been used for segmentation and morphometry of kidney biopsies¹³.

Herein, we present scoMorphoFISH (single-cell **correlative Morphometric** single-mRNA **FISH**), a Deep-Learning-accelerated approach for imaging-based and digital

single-cell single-mRNA quantification. For the first time, we combined spatial single-cell expression data with antibody-based super-resolved podocyte foot process morphometry. We integrate spatial single-cell transcriptomic, (ultra-)morphometric, and classic histology over scales as large as whole FFPE sections down to individual foot processes.

Results

Spatially resolved normalized *single-cell single-mRNA* visualization

To account for RNA-integrity differences, we wanted to identify an on-slide reference gene. Such genes should be constantly expressed and not regulated themselves. We screened the Nephroseq database for the reference genes *ACTB*, *GAPDH*, *POLR2A*, *UBC*, and *PPIB*. In a glomerular disease microarray dataset (Figure 1a), *PPIB* was the most stable gene (Figure 1b). Performance of *PPIB* as a reference gene was exemplarily demonstrated for a dual *smFISH* together with *ACE2* (Figure 1c-d). *PPIB* was homogeneously and strongly expressed (8.1 ± 4.5 transcripts per cell) (Figure 1c). Negative controls did not show a signal (Figure 1e-f).

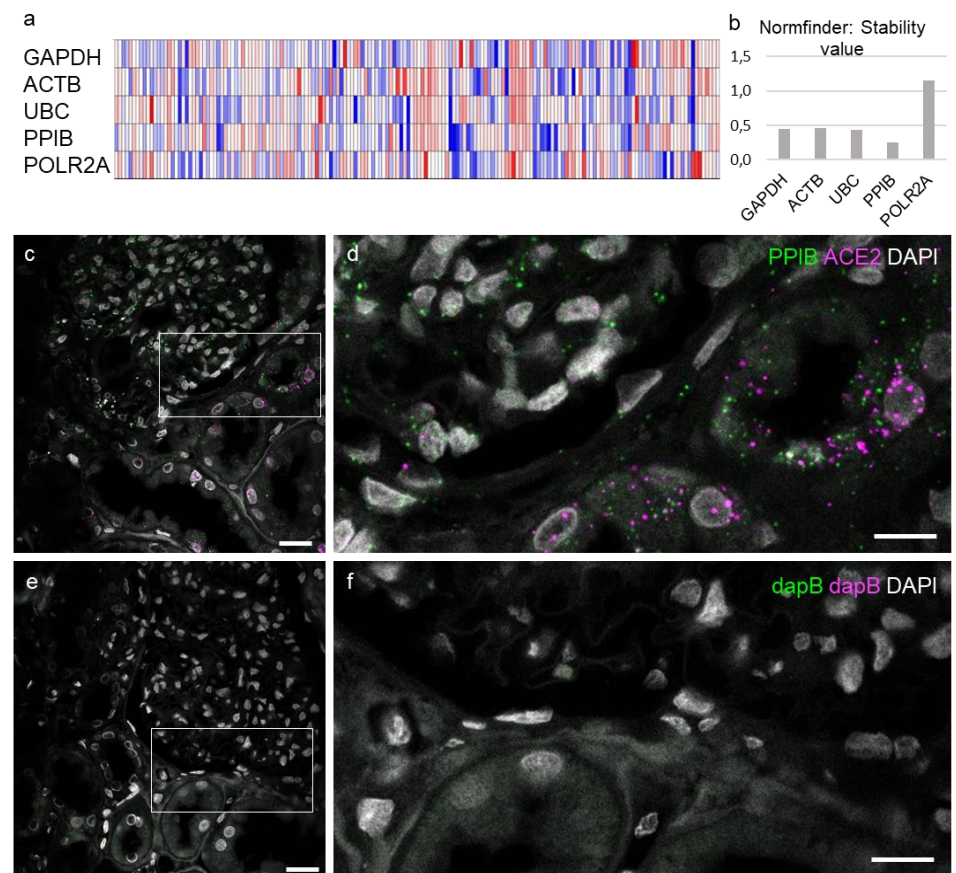


Figure 1: Evaluation of an on-slide reference gene. The Nephroseq database was screened for the five reference genes *GAPDH*, *ACTB*, *UBC*, *PPIB*, *POLR2A*. The heatmap of microarray expression data of the Ju CKD dataset is shown in a. This data was evaluated in terms of stability over samples using the Normfinder algorithm. With a stability value of 0.257, *PPIB* was identified as the most stable gene (b). Images in c and d show *PPIB* and *ACE2* mRNA dual-labelling in FFPE human kidney sections. While *ACE2* was predominantly expressed in tubular cells, *PPIB* is abundant in all cells. Single transcript spots can be clearly distinguished. Negative controls with a probe targeting the bacterial *dapB* gene (e and f) shows specificity of the *smFISH*-signals. Scale bars represent 40 μ m in the overview images and 20 μ m in the magnifications.

Immunofluorescence-based cell classification and Deep Learning-enabled tissue segmentation

To assign transcripts to tissue compartments or cell types, we established immunofluorescence protocols that perform well with smFISH. Glomeruli were labeled with single-step anti-podocin staining (Figure 2a). Less abundant antigens like WT1 were amplified by fluorescent tyramide signal amplification (TSA, Figure 2b).

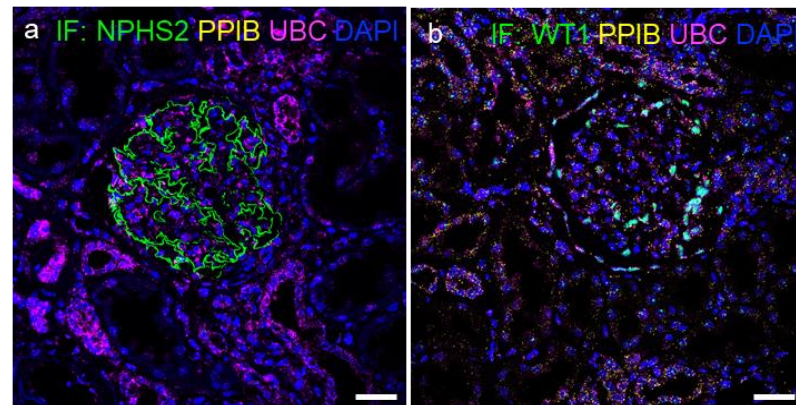


Figure 2 Combination of smFISH and immunofluorescence. The micrograph in a shows combined dual-smFISH with podocin (NPHS2) immunofluorescence. Image b exemplarily shows smFISH combined with WT1 tyramide signal amplified-immunofluorescence. Scale bars represent 40 μ m.

To establish virtual DL-segmentation-based tissue-microdissection, we custom-trained the two DL networks UNet^{14,15} and StarDist with datasets of 200 manually segmented glomeruli and 1033 cell nuclei, respectively (Suppl. Figure 1). As shown in Figure 3, outlines of glomeruli and cell nuclei of raw images were predicted with high reliability and accuracy by the trained networks.

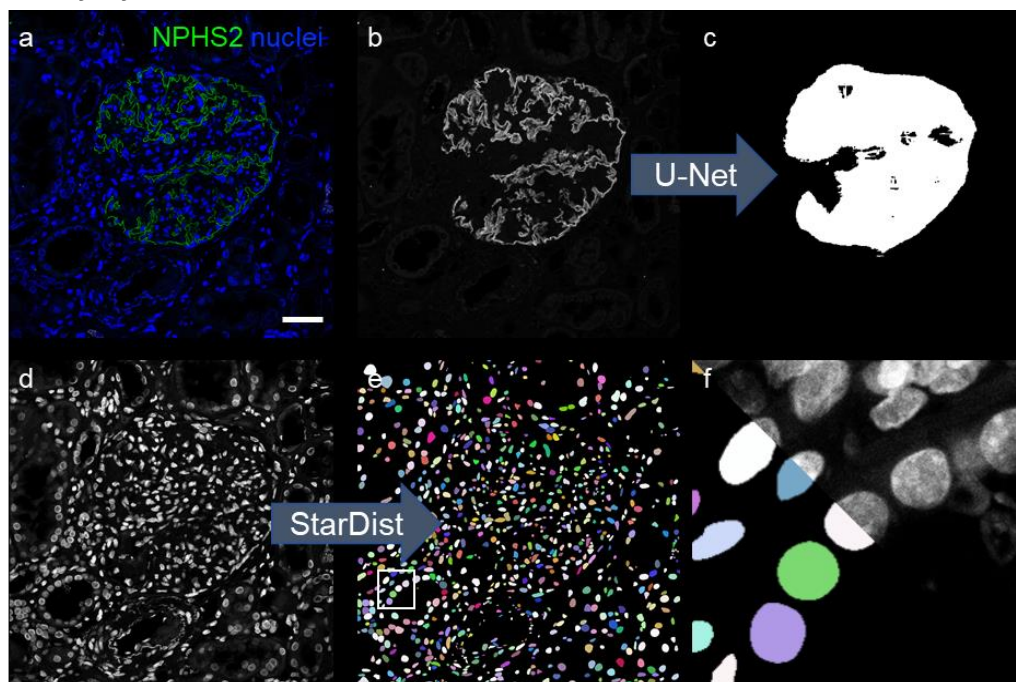
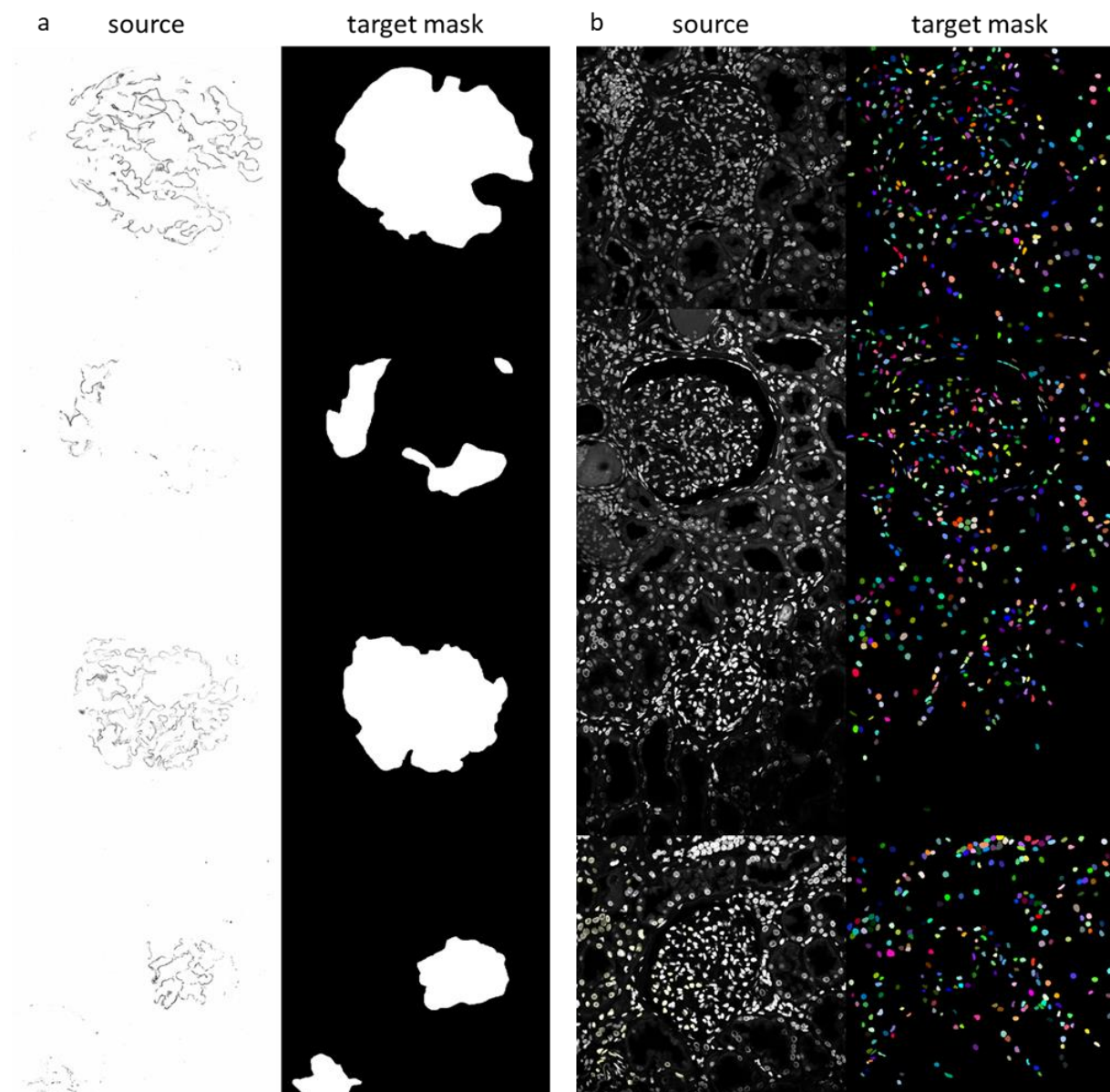


Figure 3 Deep learning-based segmentation of glomerular tufts and cell nuclei: A UNet was trained to segment glomerular outlines of podocin-stained glomeruli as shown Suppl. Fig. 1. As shown in b and d, glomerular outlines were predicted from podocin immunofluorescence micrographs (a, b) which were used as segmentation masks (c). Panel d-f shows how cell nuclei were segmented by the predictions of the trained StarDist network (Suppl. Fig. 2). The scale bar represents 40 μ m

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies

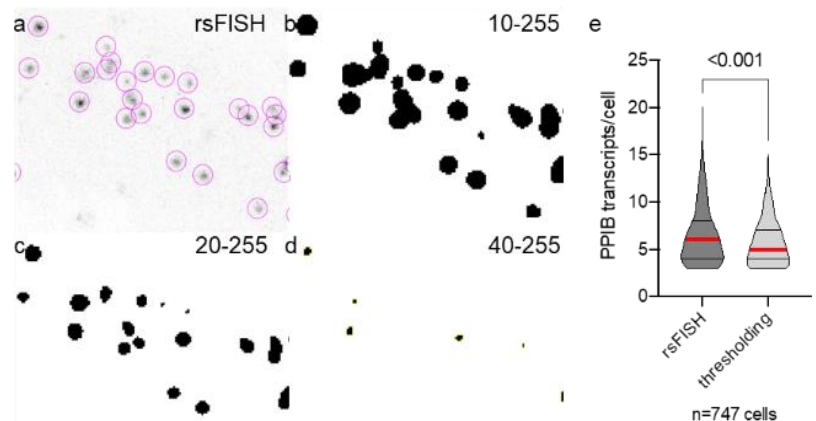


Supplemental Figure 1 Deep learning training datasets for the segmentation of glomeruli in NPHS2-stained kidney section using a UNet (a) or the segmentation of DAPI-labelled cell nuclei (b).

scoMorphoFISH: Quantitative analysis of compartment and cell type-specific mRNA abundance

To spatially map mRNA transcripts, we applied the RS-FISH algorithm that uses radial symmetry to approximate transcript localizations¹⁶. As demonstrated in Suppl. Figure 2, RS-FISH was more precise than classic thresholding-based segmentation. To automatize normalized *smFISH* transcript annotation to DL-segmented cells and tissue compartments,

we established an open-source ImageJ script including the trained DL-networks and RS-FISH. As proof-of-principle, we quantified the spatial abundance of *ACE2* and *WT1*, which showed *ACE2*-clustering in tubulointerstitial cells while *WT1* was highly enriched in the glomerular cell fraction (Figure 4a-d).



Supplemental Figure 2: RS-FISH detection of smFISH spots. In comparison to classic thresholding, rsFISH localizes more transcripts/cell with higher accuracy. Image a shows the input data together with encircled smFISH transcript localizations detected by the RS-FISH algorithm. In contrast to that, b-d show transcript segmentations after conventional thresholding which either lead to confluent transcript spots (b, c) or lose positive signals (d).

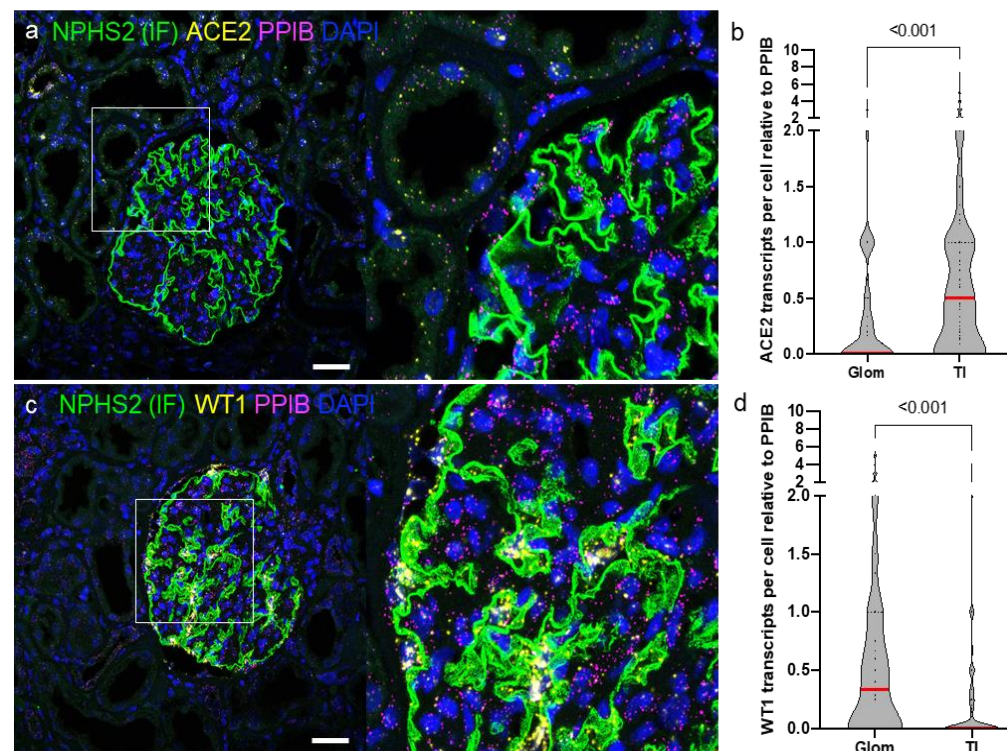


Figure 4: Normalized single-cell expression of *ACE2* and *WT1* of glomerular versus extraglomerular cells. Micrographs in a show dual smFISH for *ACE2* and PPIB combined with podocin (NPHS2) immunofluorescence. Violin plots in b show statistically significant higher expression of *ACE2* in tubulointerstitial (TI) to glomerular (Glom) cells. *ACE2*: n=329 cells *WT1* n=698 cells. $P > 0.001$, two-way Student's t-test for normal-distributed data. Scale bars represent 40 μ m

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies

Irrespective of cellular identity, mean *ACE2* expression was 0.251 normalized transcripts/cell versus 0.07 normalized transcripts/podocyte (Figure 5a-b). Vice versa, *WT1* was 9-fold enriched in podocytes in comparison to all cells (Figure 5c-d). Additionally, we quantified the abundance of *VEGFA* mRNA, a secreted factor in which

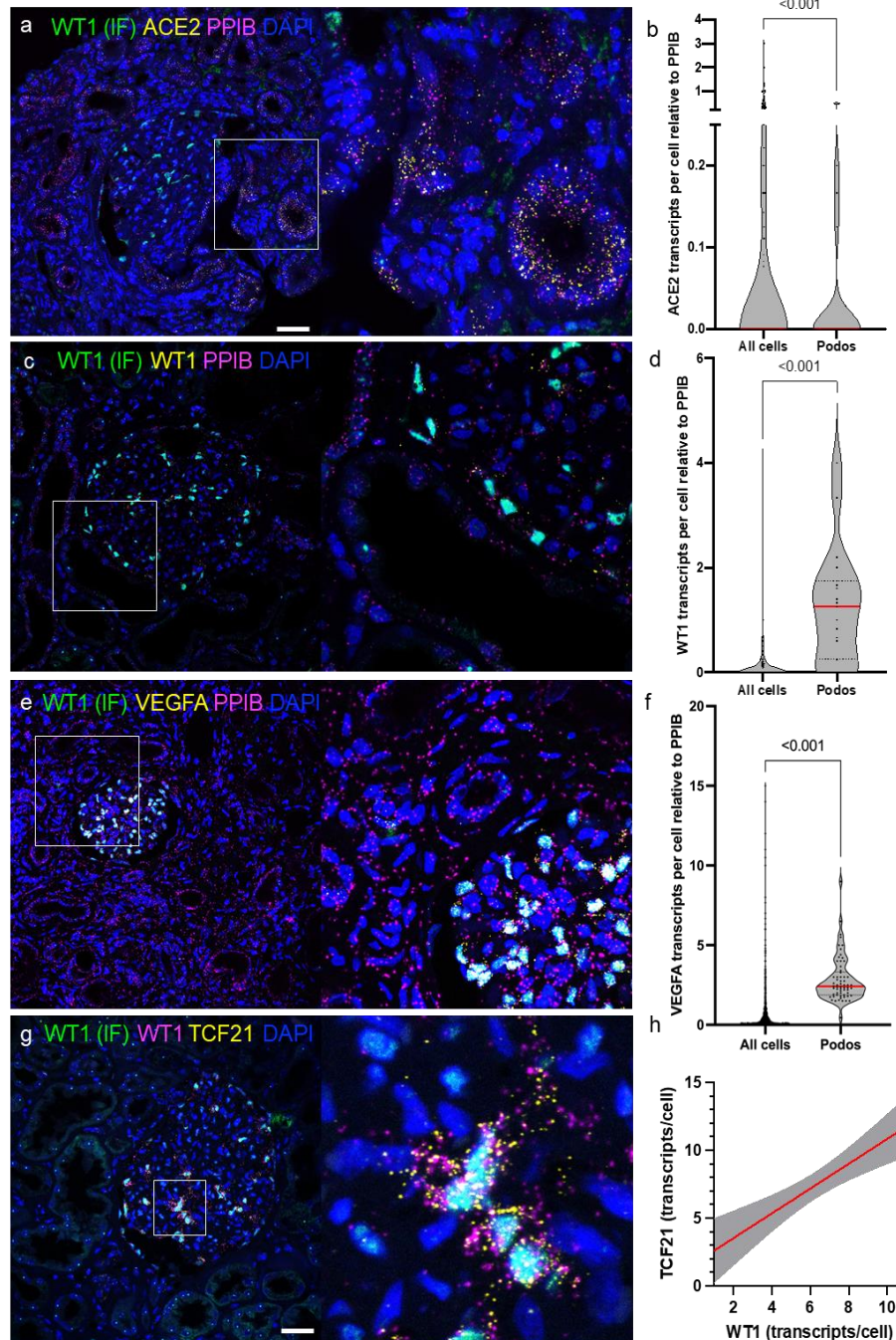


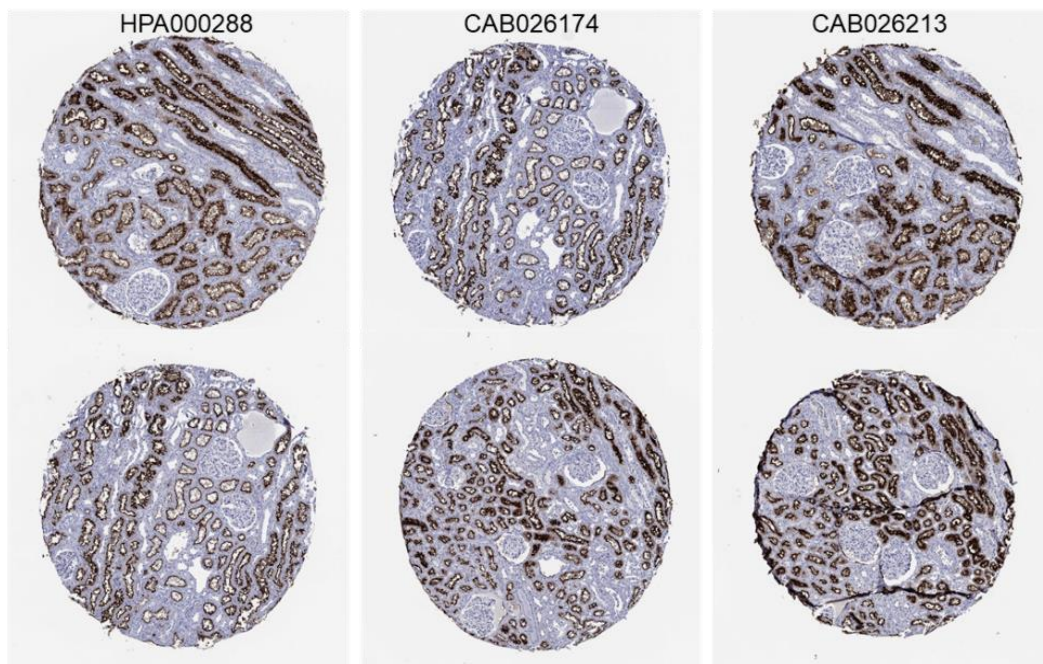
Figure 5 Normalized single-cell expression and co-expression analysis. Normalized single-cell expression of *ACE2* (a-b) *WT1* (c-d) and *VEGFA* (e-f) of the podocyte fraction versus all cells. (*ACE2*: n=324 cells, *WT1*: n=377, *VEGFA*: n=1447 cells). Panel g-h shows single-cell co-expression analysis of *TCF21* and *WT1*. The positive correlation in the plot indicates *TCF21* and *WT1* coexpression ($R^2=0.35$, $p>0.0001$, 95% confidence interval in gray, data from n=40 podocytes of 3 individual glomeruli).

immunostainings are typically not suitable for quantification. Shown in Figure 5e-f are single-cell expression data of *VEGFA*, which strongly clustered in the podocyte

fraction. As shown in Figure 5g-h, a positive correlation of *TCF21* and *WT1* single-cell expression is demonstrated (n=40 podocytes of three glomeruli, $R^2=0.35$, $p>0.0001$).

ACE2 regulation in glomerular disease

SARS-CoV2 has been found in both tubular cells and podocytes¹⁷. This is surprising as podocytes express massively lower levels of the respective entry receptor *ACE2* (Supplemental Figure 3).



Supplemental Figure 3 *ACE2 immunohistochemistry with three different antibodies of the ProteinAtlas database (www.proteinatlas.org). ACE2 is predominantly expressed in tubular*

Since it is not clear whether *ACE2* can be locally upregulated in podocytes, we investigated its expression in different glomerulopathies. As shown in Figure 6, we analyzed single-podocyte *ACE2* expression in 13 different biopsies of patients diagnosed for FSGS (primary and secondary), diabetic nephropathy (DN), membranous nephropathy (MN), Lupus nephritis (LN), IgA nephropathy, ANCA-positive glomerulonephritis as well as Goodpasture syndrome. While in general, *ACE2* expression was very low, one single primary FSGS biopsy showed statistically significantly elevated podocyte *ACE2* mRNA levels (Figure 6b).

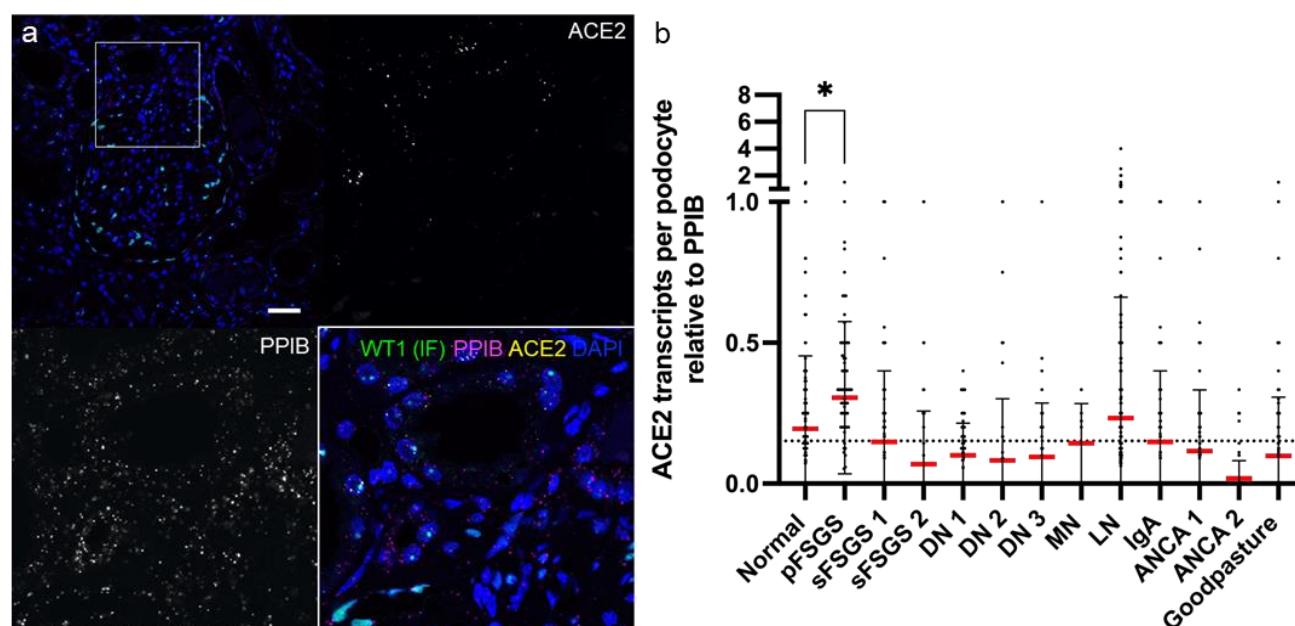


Figure 6: Podocyte ACE2 expression in various biopsies of glomerular diseases. (a) combined smFISH and WT1 immunofluorescence. The horizontal dashed line in b represents the mean expression over all samples. One primary FSGS biopsy shows elevated podocyte-ACE2 expression ($p < 0.05$ in a Kruskal-Wallis test with Dunn's multiple comparison test for non-parametric data).

Correlative super-resolution podocyte foot process morphometry, multiplex smFISH, and standard histology

To correlate single-podocyte mRNA expression with local filtration slit integrity, we imaged *smFISH* and podocin-immunofluorescence stainings with 3D-SIM. For correlation of smFISH and classic histology, sections were retrieved from the mounting medium after imaging, PAS-stained, and again whole-slide scanned. As shown in Supplemental Figure 2a, PAS morphology was sufficient for histopathological assessment, although tissue has been *smFISH*-processed. As shown in Figure 8, side-by-side evaluation of whole-slide widefield, 3D-SIM, and PAS images allowed parallel single-cell expression evaluation, filtration slit structure, and classic morphology. Filtration slits were resolved with a mean resolution of 125 ± 12 nm together (Figure 7c). PEMP filtration slit morphometry did not differ between *smFISH*-processed and native kidney samples (Supplemental Figure 4 c,d). Local mRNA expression and podocyte ultrastructure can be correlated as exemplified for *ACE2* transcript localization in Figure 7c where yellow dots (arrowheads) correspond to single *ACE2* transcripts. Interestingly, besides tubular *ACE2* expression (arrows in Figure 7b), glomerular *ACE2* expression in this glomerulus was present but restricted to podocytes in an area with cuboidal parietal epithelial cells (arrowheads in Figure 7c and e).

Siegerist et al.: Spatial single cell expression analysis in FFPE kidney biopsies

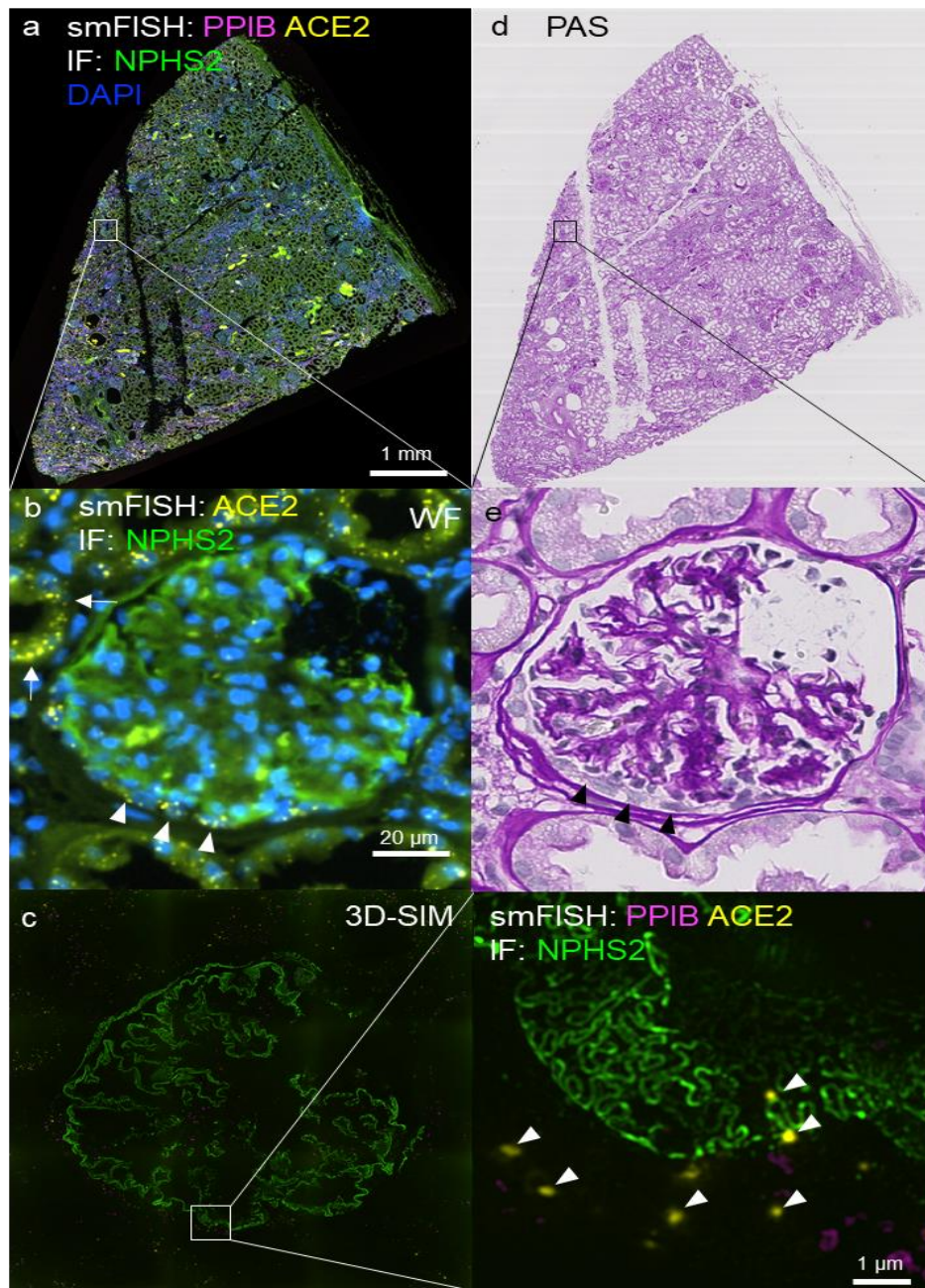
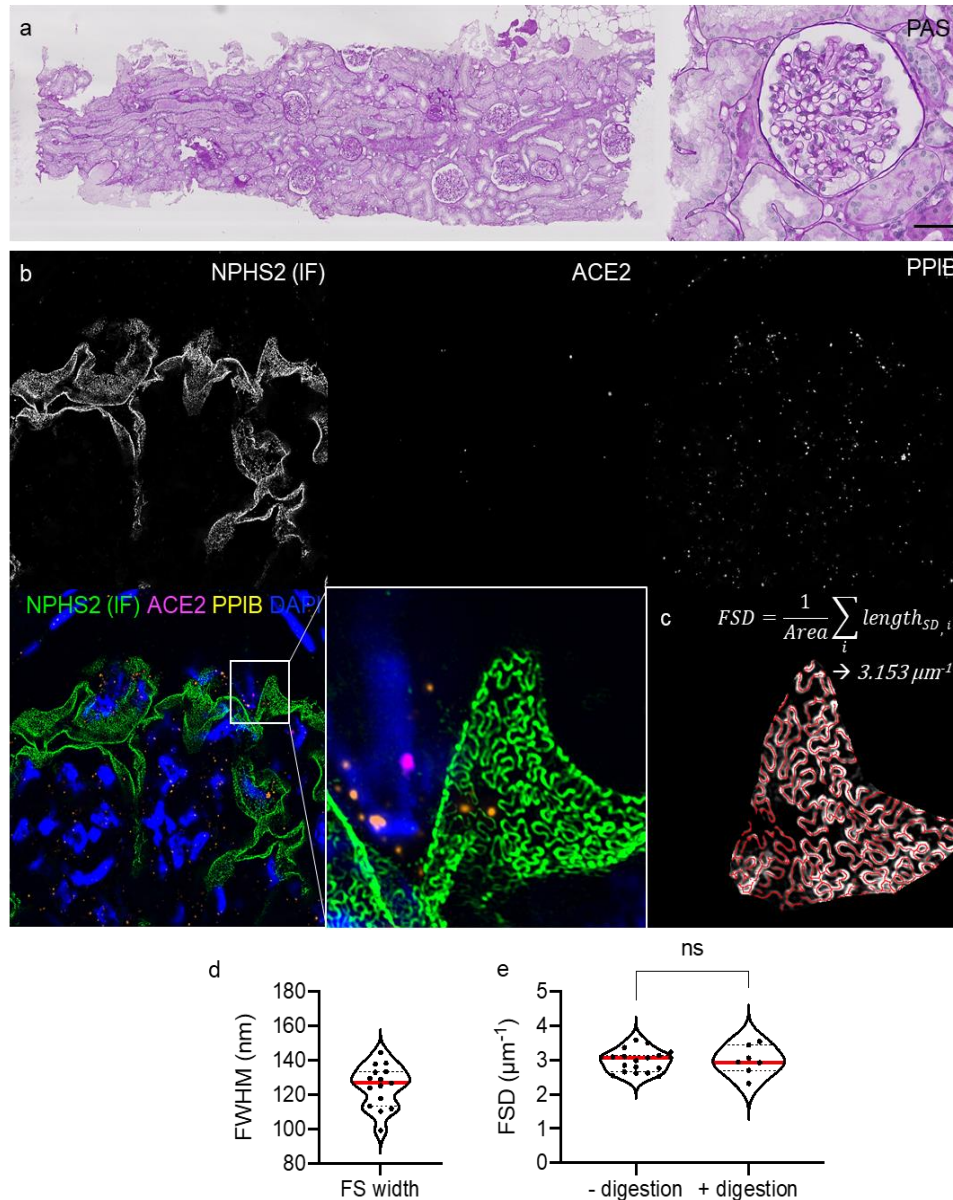


Figure 7: Correlative single-cell transcript quantification, podocyte ultramorphometry and histology. Subsequent processing of the same FFPE section for scMorphoFISH and classic histology enables correlative assessment of local mRNA-expression, podocyte foot process morphology and classic histologic aspects. FFPE sections were whole-slide scanned and aligned. Micrograph a shows a for-channel fluorescence overview over a whole section. Zoomed in in b is a single glomerulus in wide field microscopy with ACE2 expression in tubular cells (arrows) and focally in the glomerulus (arrowheads). This area has been resolved with 3D-SIM, visualizing the NPHS2-positive filtration slit which showed broadened foot processes in areas with ACE2-expression. This data can be correlated with the histologic aspect in d and e that shows parietal basement membrane thickening and cuboidal parietal epithelial cells in this area (arrowheads in e).

A second glomerulus of the same section with normal parietal epithelial cells did not show any ACE2 expression (Supplemental Figure 5). Remnant podocytes in globally

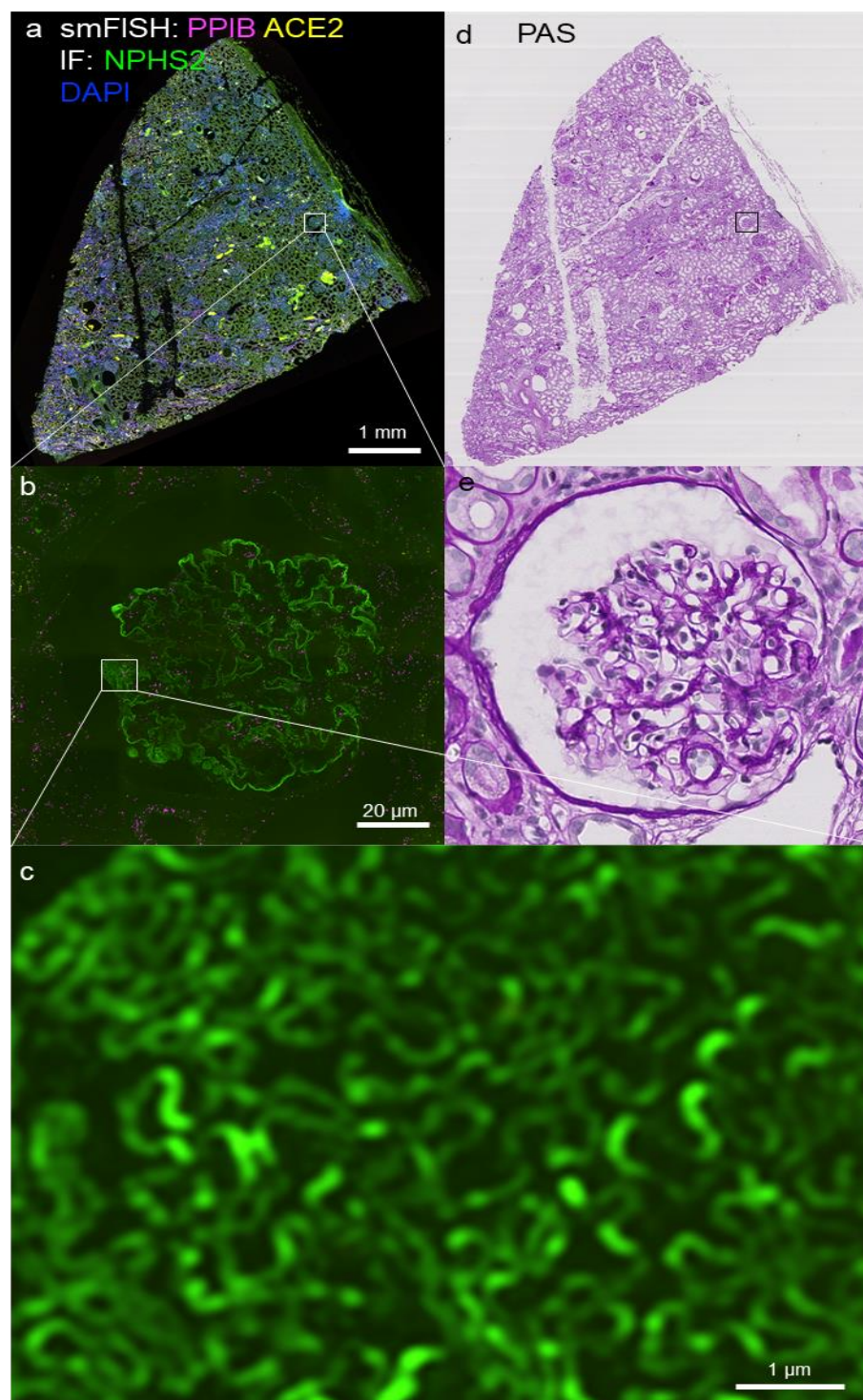
Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies

sclerotic glomeruli showed podocyte foot process effacement but preserved *VEGFA* expression (Supplemental Figure 6).



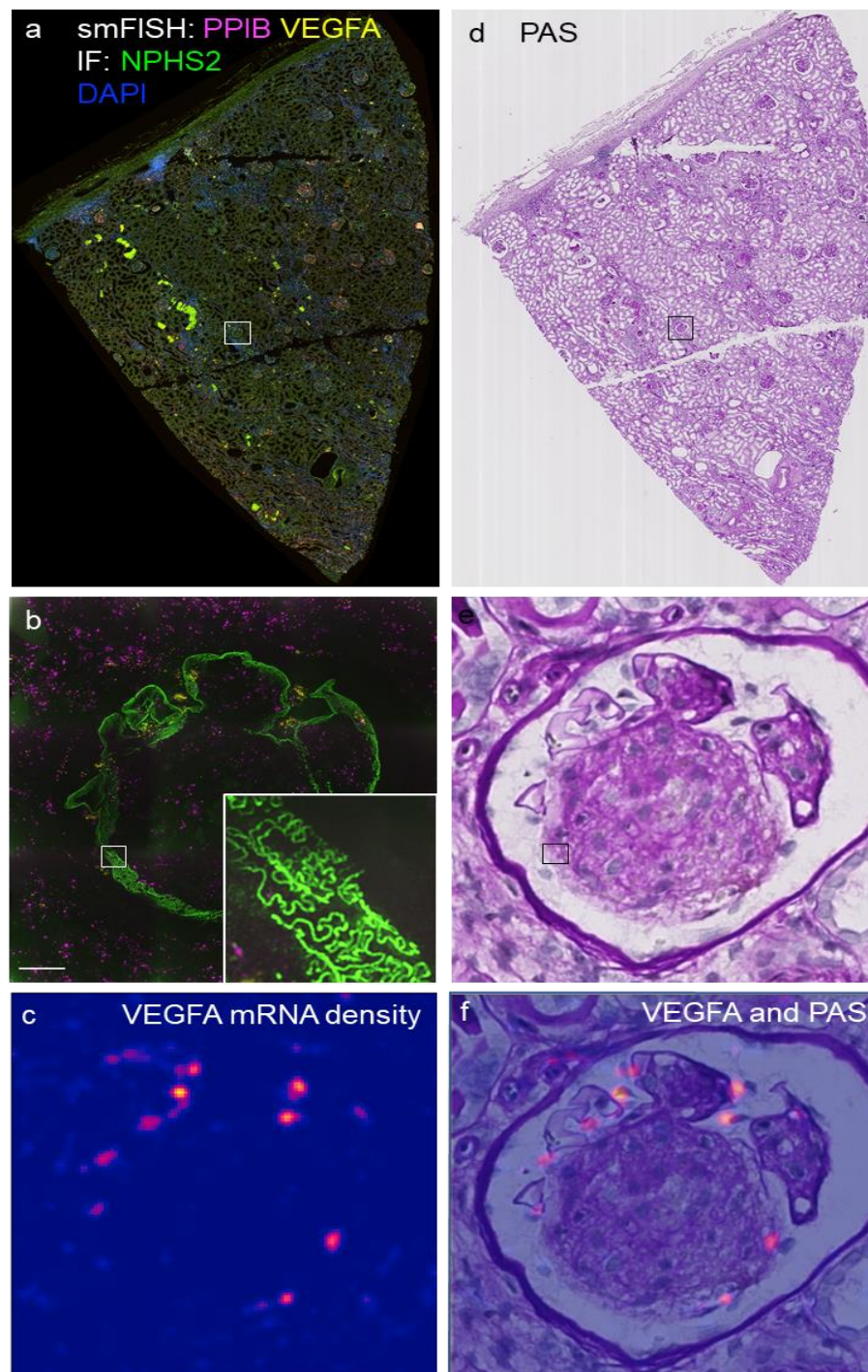
Supplemental Figure 4 Classic PAS histology after tissue digestion showed tissue integrity sufficient for pathohistological assessment (a). Podocyte filtration slits were labelled with a primary-conjugated podocin (NPHS2) antibody (b). Shown in b is the correlative imaging of local podocyte ultrastructure and associated smFISH transcripts. As shown in b, filtration slit density was quantified as the total length of the filtration slit per glomerular capillary area. Optical resolution as determined as the full width at half maximum of the sub-diffraction filtration slit was $125 \text{ nm} \pm 12 \text{ nm}$, sufficient to resolve individual foot processes (d). Digestion required for smFISH had no influence of filtration slit density (e).

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies



Supplemental Figure 5: Correlative single-cell transcript quantification, podocyte ultramorphometry and histology. A second normalized smFISH and correlative NPHS2-stained (a) glomerulus from main Fig.7. with normal podocyte structure (b,c) and no parietal cell changes (d, e).

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies



Supplemental Figure 6: Correlative smFISH for VEGFA and PPIB together with super-resolved NPHS2 IF (a). Micrograph and insert in b show central NPHS2-negativity and lateral foot process effacement. Still, VEGFA-expressing cells were present as shown in the VEGFA mRNA density-plots in c. Correlative PAS histology reveals almost global sclerosis (d, e). Local mRNA density can be directly correlated with PAS histology as demonstrated for VEGFA in f. The scale bar in b represents 20 μ m.

Discussion

Herein, we present *scoMorphoFISH*, a straightforward approach for automated digital spatial *in situ* mRNA quantification, Deep Learning-accelerated antibody-based tissue segmentation combined with correlative podocyte foot process morphometry (PEMP) in routine kidney biopsies. We used *smFISH* to *in situ* visualize single mRNA transcripts inside intact kidney tissue. Besides *smFISH*, other techniques are used to investigate spatial mRNA abundance: Spatial transcriptomics with local reverse transcription and subsequent RNA-sequencing provides a large number of transcripts per area. Unfortunately, it lacks performance with routinely generated clinical samples and is inferior to *smFISH* in terms of sensitivity and spatial resolution¹⁸. The most significant limitation for the performance of *smFISH* is the initial preparation quality which can vary among centers. To overcome this, we have established *PPIB* as an on-slide reference gene to normalize target transcript expression levels across samples or even in different parts of the same sample. We used material derived from five different origins (hereof four centers that contributed archived samples). Sections from one center posed a suboptimal RNA-integrity, which could be due to overt fixation or too high embedding temperature which can be detrimental for in-tissue RNA integrity. We deliberately established *scoMorphoFISH* on these rather challenging archived clinical samples with a higher degree of heterogeneity compared to standardized material generated in a basic-research context. Thus, the functionality of this method is not species-limited and readily usable in rodent models for which we have already shown the performance of PEMP⁴. Due to high sensitivity and specificity, this technique poses excellent functionality in routinely generated and archived formalin-fixed paraffin-embedded (FFPE) material with the potential to correlate spatial transcription and morphometric data with corresponding clinical data and disease history. Therefore, *scoMorphoFISH* enables hypothesis-testing in sample repositories with archived medical records.

Although intense tissue digestion was required for *smFISH*, super-resolved podocyte filtration slit morphology was largely unaffected and filtration slit density was in line with previously published values². Even though we developed *scoMorphoFISH* focusing on glomerular diseases, it can be instantly applied to other tissue compartments and even other organs. Ideally, an antibody can be used to segment respective tissue compartments, which chances we significantly improved when using tyramide signal

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies

amplification. In cases in wherein such antibody is available, antibody-independent *smFISH* expression levels can be used to identify respective cell types.

Typically, tissue segmentation tasks are demanding in heterogeneous sample sets (like kidney biopsies) and therefore traditional bottlenecks in image analysis workflows. DL has great potential to accelerate segmentation tasks and has already been applied to classify glomerulosclerosis¹⁹ or immunofluorescence-based glomerular morphometry¹³. We used two custom-trained DL networks for the virtual microdissection of glomeruli, all cells, and podocytes. The barriers to establishing this method in other settings are low as only commercially available reagents are used, high-quality imaging systems are widely available, data analysis is performed in the open-source image analysis platform FIJI and the source code and pre-trained DL-networks are fully available.

As shown here, correlative assessment of spatial *in situ* single-cell single-mRNA abundance and local podocyte ultramorphology is now possible for the first time on scales as large as entire FFPE sections. We believe that scoMorphoFISH is a valuable addition to the kidney tissue analysis toolbox, can aid researchers in hypothesis testing, and could be a possible step towards the high-precision evaluation of kidney biopsies in clinical settings.

Methods

For comprehensive methods see Supplemental Material and Methods. Archived kidney biopsies from the Departments of Pathology Greifswald, Hannover, Paris and Pediatric Nephrology Essen were used. Customized smFISH was performed with the RNAscope Multiplex Fluorescent V2 Kit (ACDbio). After probe visualization, sections were blocked and antibodies were incubated at 4°C overnight which was either primarily conjugated with Alexa Fluor 488-conjugated dual monoclonal recombinant alpaca anti-rabbit IgG nano secondaries (Chromotek) or detected with the Vectastain Elite ABC-HRP Kit (VectorLaboratories) with TSA with Opal dyes (Akoya Biotech). Sections were mounted in Mowiol for microscopy (Roth, Karlsruhe, Germany). Four-channel image stacks were either acquired on a TCS-SP5 confocal laser scanning system, a Zeiss Elyra PS.1 system (Carl Zeiss Microsystems), or a Nikon N-SIM-E 3D-SIM system. DL-networks were trained within the ZeroCostDL4Mic ecosystem²⁰. ImageJ scripts were developed as IJ1 macros in FIJI²¹. Source codes, as well as example data, are available online: <http://www.github.com/Siegerist>.

Author Contributions

FS established methodology, source codes, performed experiments, trained DL-networks, prepared figures, and wrote the manuscript. EH established methodology and performed experiments. JSD established source codes, generated training data, and trained DL networks. UZ, SR, AB, JO, JHB, OL, PLT contributed kidney biopsies. NE and KE supervised the study, acquired funding, and wrote the manuscript. All authors approved the final version of the manuscript.

Acknowledgments

This work was supported by a starting grant of the Forschungsverbund Molekulare Medizin, Universitätsmedizin Greifswald to FS and by a grant of the Federal Ministry of Education and Research (BMBF, grant 01GM1518B, STOP- FSGS) to NE. This work was generously supported by the Südmeyer fund for kidney and vascular research ('Südmeyer Stiftung für Nieren- und Gefäßforschung') and the Dr Gerhard Büchtemann fund, Hamburg, Germany. This work was also supported by the Fondation pour la Recherche Médicale (FRM, grant "VALID") to PLT. This work highly profited from the ImageJ image-analysis community and the open-source ZeroCostDL4Mic ecosystem. The expert technical assistance of Sindy Schröder, Vedran Drenic, and Marc Niemeyer is highly acknowledged.

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies

Conflicts of interest

NE serves as CEO and holds shares of NIPOKA GmbH. FS holds shares of NIPOKA GmbH, a startup commercializing PEMP. All other authors report no competing interests.

References

1. Zhao Y, Bucur O, Irshad H, et al. Nanoscale imaging of clinical specimens using pathology-optimized expansion microscopy. *Nat Biotechnol.* 2017;35(8):757-764. doi:10.1038/nbt.3892
2. Siegerist F, Ribback S, Dombrowski F, et al. Structured illumination microscopy and automatized image processing as a rapid diagnostic tool for podocyte effacement. *Sci Rep.* 2017;7(1):1-11. doi:10.1038/s41598-017-11553-x
3. Pullman JM, Nylk J, Campbell EC, Gunn-Moore FJ, Prystowsky MB, Dholakia K. Visualization of podocyte substructure with structured illumination microscopy (SIM): a new approach to nephrotic disease. *Biomed Opt Express.* 2016;7(2):302. doi:10.1364/boe.7.000302
4. Tesch F, Siegerist F, Hay E, et al. Super-resolved local recruitment of CLDN5 to filtration slits implicates a direct relationship with podocyte foot process effacement. *J Cell Mol Med.* 2021;25(16):7631-7641. doi:10.1111/jcmm.16519
5. Cohen CD, Frach K, Schlöndorff D, Kretzler M. Quantitative gene expression analysis in renal biopsies: A novel protocol for a high-throughput multicenter application. *Kidney Int.* 2002;61(1):133-140. doi:10.1046/j.1523-1755.2002.00113.x
6. Hatje F, Wedekind U, Sachs W, et al. Tripartite Separation of Glomerular Cell-Types and Proteomes From Reporter-Free Mice. *J Am Soc Nephrol.* Published online 2021:ASN.2020091346. doi:10.1681/asn.2020091346
7. Park J, Shrestha R, Qiu C, et al. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease HHS Public Access. *Science (80-).* 2018;360(6390):758-763. doi:10.1126/science.aar2131
8. Gall JG, Pardue ML. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proc Natl Acad Sci U S A.* 1969;63(2):378-383. doi:10.1073/pnas.63.2.378
9. Cheng YS, Zhuo Y, Hartmann K, et al. Autonomous combinatorial color barcoding for multiplexing single molecule RNA visualization. *bioRxiv.* Published online 2017. doi:10.1101/127373
10. Choi HMT, Schwarzkopf M, Fornace ME, et al. Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Dev.* 2018;145(12):1-10. doi:10.1242/dev.165753
11. Wang F, Flanagan J, Su N, et al. RNAscope: A novel in situ RNA analysis

- platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagnostics*. 2012;14(1):22-29. doi:10.1016/j.jmoldx.2011.08.002
12. Femino AM, Fay FS, Fogarty K, Singer RH. Visualization of single RNA transcripts in situ. *Science* (80-). 1998;280(5363):585-590. doi:10.1126/science.280.5363.585
13. Zimmermann M, Klaus M, Wong MN, et al. Deep learning-based molecular morphometrics for kidney biopsies. *JCI Insight*. 2021;6(7). doi:10.1172/jci.insight.144779
14. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2015;9351:234-241. doi:10.1007/978-3-319-24574-4_28
15. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry. *Nat Methods*. 2019;16(1):67-70. doi:10.1038/s41592-018-0261-2
16. Bahry E, Breimann L, Epstein L, et al. RS-FISH: Precise , interactive and scalable smFISH spot detection using Radial Symmetry. *bioRxiv Prepr*. Published online 2021. doi:10.1101/2021.03.09.434205
17. Puelles VG, Lütgehetmann M, Lindenmeyer MT, Sperhake JP, Wong MN, Allweiss L, Chilla S, Heinemann A, Wanner N, Liu S, Braun F, Lu S, Pfefferle S, Schröder AS, Edler C, Gross O, Glatzel M, Wichmann D, Wiech T, Kluge S, Pueschel K, Aepfelbacher M HT. Multiorgan and Renal Tropism of SARS-CoV-2. *N Engl J Med*. 2020;383(6):59. doi:10.1056/NEJMc2011400
18. Ståhl PL, Salmén F, Vickovic S, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* (80-). 2016;353(6294):78-82. doi:10.1126/science.aaf2403
19. Østergaard M V., Secher T, Christensen M, et al. Therapeutic effects of lisinopril and empagliflozin in a mouse model of hypertension-accelerated diabetic kidney disease. *Am J Physiol Physiol*. 2021;321(2):F149-F161. doi:10.1152/ajprenal.00154.2021
20. Chamier L von, Laine R, Jukkala J, et al. ZeroCostDL4Mic: an open platform to use Deep-Learning in Microscopy. Published online 2020:1-4. doi:10.1101/2020.03.20.000133
21. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: An open-source platform for

Siegerist *et al.*: Spatial single cell expression analysis in FFPE kidney biopsies

- biological-image analysis. *Nat Methods*. 2012;9(7):676-682. doi:10.1038/nmeth.2019
22. Bloice MD, Stocker C, Holzinger A. Augmentor: An Image Augmentation Library for Machine Learning. *arXiv*. Published online 2017:1-5. doi:10.21105/joss.00432
23. Linkert M, Rueden CT, Allan C, et al. Metadata matters: Access to image data in the real world. *J Cell Biol*. 2010;189(5):777-782. doi:10.1083/jcb.201004104
24. Schmidt U, Weigert M, Broaddus C, Myers G. Cell detection with star-convex polygons. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11071 LNCS:265-273. doi:10.1007/978-3-030-00934-2_30

Supplementary information: Comprehensive Methods

Sample preparation

Human kidney biopsies of the Departments of Pathology of Paris and Hannover and of the Department of Pediatric Nephrology Essen were used. Additionally, we used excess healthy kidney tissue of tumor nephrectomies from the Department of Urology of the University Medicine Greifswald. After immersion fixation in 3% PFA overnight at room temperature, kidneys were embedded in paraffin using standard protocols. Care was taken that the temperature did not exceed 60°C. 5 µm FFPE tissue sections were mounted on superfrost slides and air-dried at room temperature.

Multiplex fluorescence in situ hybridization and immunofluorescence staining

For multiplex *smFISH*, the ACDbio RNAscope Multiplex Fluorescent V2 Kit per manufacturer's description with following adaptations: For heat-induced epitope retrieval sections were boiled for 15 min in ACDbio antigen retrieval buffer after deparaffinization in xylene. After protease treatment for 15 min at room temperature, complementary probes were hybridized for two hours at 40°C in the ACDbio HybEZ oven. Sections were stored in 5x SSC (Sigma) overnight. Hybridized probes were detected by complementary amplification probes after HRP-binding signals were detected using Opal 520, 570, and 690 dyes diluted 1:1500 in TSA amplification buffer (ACDbio). After the last detection step, sections were collected in 1x PBS, blocked with 1% normal goat serum, 1% fetal bovine serum, 1% bovine serum albumin, and 0.5% cold fish gelatin for 45 min at room temperature. Whilst blocking, primary affinity-purified rabbit anti podocin antibodies were mixed with Alexa Fluor 488-conjugated secondary dual monoclonal recombinant alpaca anti-rabbit IgG VHH nanobodies (nano secondaries, Chromotek) in blocking solution for a final concentration of 1:150 primary and 1:1000 secondary nanobody. Sections were incubated with the antibody mix at 4°C overnight. Slides were washed in three changes of 1x PBS and nuclei counterstained with 0.1mg/ml DAPI. Slides were rinsed in A.dest and mounted in Mowiol for microscopy (Carl Roth, Karlsruhe, Germany). Until being imaged, sections were stored at 4°C in the dark.

Correlative histology

After being whole-slide imaged by widefield and 3D-SIM on an N-SIM-E 3D-SIM setup (Nikon), slides were immersed in 37°C 1xPBS for 1h. Coverslips were gently removed

and Mowiol mounting medium washed out in 3 changes of 1xPBS. After that, routine PAS staining was performed as described before². Sections were mounted in Eukitt (Carl Roth).

Imaging

To obtain confocal-laser scanning micrographs, a Leica TCS-SP5 system was used. Micrographs were acquired using a 40x 1.2 NA oil immersion objective with a voxel size of 189x189x500 nm (xyz). Stacks over 4 µm were acquired. For super-resolving 3D-structured illumination microscopy, a Zeiss Elyra PS.1 system (Carl Zeiss Microsystems) or a Nikon N-SIM-E was used as described before². Whole slide images of PAS-stained sections were acquired on a Leica SCN400 slidescanner. SCN files were imported and processed in QuPath (v0.3.0).

Deep-Learning

Using the Google Colab-based ZeroCostDL4Mic notebooks, we trained a U-Net, a deep-learning-based neural network. Glomerular tuft outlines determined by NPHS2⁺ glomerular capillaries in stacks of confocal laser scanning micrographs were manually segmented and saved as ROIs in FIJI. The ROIs were then exported as binary 8-bit masks in a corresponding image stack in which glomerular area and background were coded as intensity 255 and 0, respectively. Both image stacks were exported as individual corresponding 512x512 px tiff-files with matching names and uploaded in two separate GoogleDrive-folders as source and template training files. After training, a separate set of files was used for quality control purposes. The UNet model was trained with 200 epochs on 200 paired image patches (image dimensions: (1024, 1024 px), patch size: (512, 512 px)) and an initial learning rate of 3.0000002e-36, using UNet 2D ZeroCostDL4Mic. Key python packages used include tensorflow (v 0.1.12), Keras (v 2.3.1), numpy (v 1.19.5), cuda (v 11.0.221Build cuda_11.0_bu.TC445_37.28845127_0). The training was accelerated using a Tesla K80 GPU. Following parameters were used: number_of_epochs 200; patch_size 512x512; batch_size 4; number_of_steps 23; percentage_validation 10; initial_learning_rate 3.0000002e-36. The readily trained network was exported and saved for subsequent predictions. Predictions were performed in the DeepImageJ plugin by installing the readily trained network. Predicted glomerular regions of interest (ROIs) are saved by the script to the respective source folder and called later by the macro script to define intra- and extraglomerular cells and transcripts.

The StarDist 2D model was trained from scratch for 400 epochs on 4 paired image patches (image dimensions: (2048, 2048 px), patch size: (2048, 2048 px)) with a batch size of 2 and a mae loss function, using the StarDist 2D ZeroCostDL4Mic notebook (v 1.12)²⁰. Key python packages used include tensorflow (v 0.1.12), Keras (v 2.3.1), csbdeep (v 0.6.1), numpy (v 1.19.5), cuda (v 10.1.243). The training was accelerated using a Tesla P100GPU. The dataset was augmented by a factor of 10 using Augmentor²². Following parameters were used: number_of_epochs 400; patch_size 2048x2048; batch_size 2; number_of_steps 30.0; percentage_validation 10; n_rays 32; grid_parameter 2; initial_learning_rate 0.0003. The trained StarDist model was imported to FIJI and is automatically called by the scoMorphoFISH script.

Script development

ImageJ macros were developed in the IJ1 macro language in FIJI²¹. The script requires several different pre-installed plugins: BioFormats²³ to import C-LSM data, StarDist which requires the CSBDeep, and StarDist update sites in FIJI²⁴, rsFISH¹⁶, deepImageJ, and Read and Write Excel. Data in the multichannel tiffs should be ordered with the immunofluorescence channel in C1, *smFISH* in C2/3, and DAPI in C4.

scoMorphoFISH:

The script can perform different tasks. If glomerular outlines are stained by immunofluorescence:

1. **Glomerular vs. tubulointerstitial transcript counter;** For two-channel smFISH + NPHS2 IF + DAPI. The script asks first for the source folder of the multichannel tiff stacks. The script uses the trained UNet to predict the glomerulus segmentation mask from the glomerular staining, takes the outlines of the mask as an ROI, calculates the intra- and extraglomerular area. It then differentially segments intra- and extraglomerular cells and counts transcript in the single-cell ROIs. Output is two-channel intra-and extraglomerular single-cell transcripts, total transcripts intra- and extracellular, intra- and extraglomerular area.

If no glomerular staining is present. The script will ask if glomerular outlines should be segmented manually. If nuclear IF is present, immunofluorescence positive cells are segmented using the trained StarDist network, and expression is differentially measured IF positive and negative cells. This part is established for podocytes but works with every strong nuclear marker.

2. **Podocyte transcript counter:** For two-channel smFISH + WT1 IF + DAPI. The script asks first for the source folder of the multichannel tiff stacks and then manually creates the glomerular outline masks. Podocytes are identified by: intraglomerular position, WT1, and DAPI positivity. Podocytes are segmented with the same model as all nuclei. Podocyte morphometric parameters: Total number/glomerular cross-section, Feret's diameter are exported for estimated glomerular morphometry (podocyte density). Output is intra-, extraglomerular, total, podocyte single-cell dual-channel transcripts.

Output data is exported in one accumulated .xls file per folder analyzed. The script works in batch processing mode and processes all tiff files in a directory.

2D density plots

For the creation of 2D-density plots, the R spatstat package was used: After thresholding and binarization of the smFISH channels, xy-positions of the transcript spots were detected using RS-FISH and exported as csv files from ImageJ. XY positions were loaded to R-studio (version a1.3.1093) and plotted using the density plot function of the SpatStat library.

Data availability

All source codes of scripts used in this manuscript as well as example data is available online: <http://www.github.com/Siegerist>