bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Deep learning-based segmentation and quantification of podocyte foot process morphology

Linus Butt[1,*], David Unnersjö-Jess[1,2,*], Martin Höhne[1,2], German Sergei[3], Anna Witasp[4], Annika Wernerson[4], Jaakko Patrakka[5], Peter F. Hoyer[6], Hans Blom[7], Bernhard Schermer[1,2], Katarzyna Bozek [3,#], Thomas Benzing[1,2,3,#]

[1] Department II of Internal Medicine and Center for Molecular Medicine Cologne (CMMC), University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany
[2] Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Germany
[3] Center for Molecular Medicine Cologne (CMMC), University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany
[4] Division of Renal Medicine, Department of Clinical Sciences, Intervention and Technology, Karolinska Institute, Stockholm, Sweden.
[5] KI/AZ Integrated CardioMetabolic Center, Department of Laboratory Medicine. Karolinska Institutet at Karolinska University Hospital Huddinge, Stockholm, Sweden
[6] Pediatric Nephrology, Pediatrics II, University of Duisburg-Essen, Essen, Germany
[7] Science for Life Laboratory, Dept. of Applied Physics, Royal Institute of Technology, Solna, Sweden

* equal contribution (order determined by alphabetical order)
[#] shared senior authorship

Correspondence to: linus.butt@uk-koeln.de; david.unnersjoe-jess@uk-koeln.de; k.bozek@uni-koeln.de

Running title: Deep learning-based segmentation and quantification of podocyte foot process morphology

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

The kidneys constantly filter enormous amounts of fluid, with almost complete retention of albumin and other macromolecules in the plasma. Diseases of podocytes at the kidney filtration barrier reduce the glomerular capillary surface area available for filtration and alter the intrinsic permeability of the capillary wall resulting in albuminuria, however, direct quantitative assessment of the underlying morphological changes has not been possible so far. Here we developed a deep learning-based approach for segmentation of foot processes in images acquired with super-resolved stimulated emission depletion (STED) microscopy or confocal microscopy. Our method – Automatic Morphological Analysis of Podocytes (AMAP) – detected 87-95% manually-annotated foot processes and additionally recognized 1.3 - 2.17-fold more. It also robustly determined morphometric parameters, at a Pearson correlation of $r > 0.71$ with a previously published semi-automated approach, across a large set of mouse tissue samples. The artificial intelligence algorithm was applied to a set of human kidney disease conditions allowing comprehensive quantifications of various underlying morphometric parameters. These data confirmed that when podocytes are injured, they take on a more simplified architecture and the slit-diaphragm length is much reduced, resulting in a reduction in the filtration slit area and a loss of the buttress force of podocytes which increases the permeability of the GBM to albumin.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

The capacity of the mammalian kidney to filter vast amounts of fluid (180 L/day in humans) and to almost completely restrict the passage of macromolecules (e.g. albumin) relies on the intricate glomerular filtration barrier. This barrier consists of a fenestrated endothelium, the glomerular basement membrane (GBM) and specialized post-mitotic epithelial cells, called podocytes[1]. Damage to any of the three layers upon glomerular injury results in a pathologically increased filter permeability, with albuminuria being the most prominent clinical symptom of a compromised filtration barrier[2]. Due to the nanoscale dimensions of the filter, electron microscopy is widely used in clinical pathology and research to qualitatively, and in some instances semi-quantitatively, assess morphological alterations that are associated with glomerular injury[3]. Pathological alterations of podocytes, particularly affecting their so-called foot processes (FPs) and the slit diaphragm (SD) as their only cell junction, are frequently seen in virtually all types of glomerular diseases. These alterations are summarized in the term *foot process effacement*, which includes widening and ultimately loss of foot processes together with a progressive shortening of total SD length[4,5]. The advent of super-resolution light microscopy techniques allowed the visualization and subsequent quantification of morphological alterations upon FP effacement[6–11]. Recently, our group has used this approach in the podocin$^{R231Q/A286V}$ mouse model of hereditary focal and segmental glomerulosclerosis (FSGS) to provide a model of kidney ultrafiltration in which the size selectivity of the filtration barrier is dependent on the compression of the GBM[12]. In that study, morphological alterations of FPs, particularly the decrease in SD length, correlated robustly with levels of albuminuria in podocin$^{R231Q/A286V}$ mice. Integrating the morphological data into biophysical models of ultrafiltration, it was proposed that injured podocytes lack the ability to adequately counteract the filtration pressure, which leads to a relaxation of the fiber matrix within the GBM resulting in an increased permeability to albumin. Until now, the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

morphological analyses underlying this model had to be carried out manually or semi-automatically. They required annotation of regions of interest (ROIs) and manual assignment of individual FPs, which is not only time consuming but also investigator-dependent, currently impeding the broad use of FP morphometrics in research and diagnostics.

In this study, we combined the previously established morphological analyses of podocytes with a machine-learning algorithm in order to enable the first completely automated segmentation and quantification of podocyte ultrastructure. New machine learning methods automate bioimage analysis at a human-level accuracy with cancer histopathology as one of its most prominent applications[13,14]. In nephrology, deep learning methods have been proposed for the segmentation of entire renal structures or entire podocytes[15–17]. These methods can facilitate expert kidney biopsy assessment by automatically marking the important morphological elements. However, none of the existing segmentation methods allows the recognition of the subcellular structure of podocytes.

Here we propose Automatic Morphological Analysis of Podocytes (AMAP) – a fully automated method for detection of FPs and the overall SD pattern from high- and super-resolution optical microscopy images of podocytes. Our method is based on deep learning instance segmentation that has been trained and tested on a broad range of disease and imaging conditions, correctly detecting 87-95% of FPs across our datasets and allowing to reproduce the morphometric parameters originally extracted using the ImageJ macro – referred to as macro throughout the manuscript. Importantly, it successfully generalizes to human tissue samples as well as to images acquired using a diffraction-limited confocal microscope, a system commonly available in clinical pathology labs. Our approach opens the way towards systematic study of pathologies occurring at these nanoscale structures in kidney disease.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RESULTS AND DISCUSSION

Method accuracy

Using a previously published STED imaging protocol[6], we generated a set of 209 images representing healthy and diseased mouse tissue at various ages and degrees of podocyte effacement. The images represent FP shapes by staining the SD protein nephrin (see Fig. 1a for an overview). Annotation of these images, including both SD and FP areas (Supplementary Fig. 1), was generated with a network pre-trained on a small set of macro-labeled images. These annotations were then further manually corrected in order to obtain a possibly complete annotation of each image. For improved accuracy, each file was annotated by at least two individuals. Test images were manually selected to reflect the variability of FP effacement present in the training set.

We next trained a segmentation convolutional neural network (CNN) on the training set. The segmentation approach incorporates both semantic and instance segmentation (Fig. 1b-d)[18]. Semantic segmentation classifies pixels into three classes: background, FP, and SD (Fig. 1c). Instance segmentation additionally assigns FP pixels into separate FP instances (Fig. 1d). Briefly, the instance assignment is based on the clustering of CNN-derived numeric pixel representations. These representations are constructed so that pixels belonging to the same instance lie close together in the representation space while a predefined margin separates different instances. Clustering of such pixel representations results in the separation of individual FPs in an image (Supplementary Fig. 2).

We trained the network and inspected its performance on the test-set images assessing both the semantic and instance segmentation accuracy. On average, AMAP assigned the correct class to 73% and 76% FP and SD pixels, respectively. Relative to the number of pixels in the FP and SD class, 23-27% of pixels were misclassified, i.e., incorrectly assigned either to

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

background, FP, or SD. Based on visual inspection of the results, we noted that these errors were sometimes due to the incomplete annotation and uncertainty in the exact FP and SD boundaries (Fig. 1e-f).

The FP pixels were additionally split into separate instances, and we matched the predicted instances with the labels. Predicted instances that overlapped with a labeled FP were counted as true positives, and we quantified their area of overlap with the matched label. In the test images, we detected 87% of all labeled FPs (Fig. 1g). Matched FP instances overlapped on average at 0.74 quantified as Intersection over Union (IoU) of the labeled and predicted instances. Relative to the number of labeled FPs, 28% more FPs were predicted with our method, while 13% were not detected (Fig. 1g). Notably, some of the potential false positive detections were due to missing labels and should not be considered errors.

Morphometric parameter estimation

As further validation of this approach, we asked if such segmentation performance allows to accurately reproduce the morphometric parameters extracted using a previously published macro-based analysis of the podocin$^{R231Q/A286V}$ FSGS mouse model and their control littermates[12]. Podocin$^{R231Q/A286V}$ mice showed quantifiable morphological alterations of individual FPs as well as the overall SD pattern that correlated with disease progression. We ran the segmentation method on the set of 174 nephrin-stained images used in the original publication (Fig. 2a). The imaged samples include control and mutant mice from 0 to 20 weeks of age. Out of the 19,676 labeled FPs in the published dataset, we correctly detected 18,673 (95%, Fig. 2b) with an area overlap of 0.81 IoU on average. However, we detected an additional 25,193 FPs that did not overlap with any of the labeled ones in the published dataset. Visual inspection of these predictions suggests that they predominantly represent FPs

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that were overlooked in the macro analysis rather than errors. Our method failed to detect 1,003 labeled FPs, which represents 5% of all annotated FPs.

Instead of manually verifying the over 40,000 FPs segmented by AMAP, we inspected to what extent the FPs inferred by our method reproduce the morphometric parameters of the labeled FPs in each age and disease stage (see Supplementary Fig. 3 for an overview of quantified FP parameters). Even though the numbers of FPs detected with AMAP vs. macro differed, we found an overall high agreement of the morphometric parameters ($r > 0.71$, $p < 0.0001$, permutation) in area, perimeter and circularity values (Fig. 2c-e). The absolute values for area and perimeter inferred by AMAP were overall lower compared to the ones derived from the macro (Fig. 2f-g). However, comparing the morphological parameters of FPs detected both by AMAP and macro shows a high agreement in absolute values (Supplementary Fig. 4), suggesting that the additional FPs not marked using the macro but detected by AMAP tend to be smaller. Consistent with the findings in the original dataset, quantification of perimeter and circularity by the automated approach allowed the distinction between control and mutant mice, whereas the FP area alone is not sufficient to make this distinction (Fig. 2f-h). These results illustrate that the segmentation accuracy of our method is sufficient to robustly reflect the alterations in FP morphology characteristic of the progression of FSGS.

It has been shown before that FP effacement affects not only the morphology of individual FPs but also the overall SD length[6,12,19]. Additionally, SD length has been mechanistically linked to the occurrence of albuminuria[12]. Therefore, we developed a method to automatically assign regions of interest (ROIs) based on the SD staining pattern and quantify the SD length within the boundaries of the ROIs (Fig. 3a, Supplementary Fig. 3). Overall, AMAP resulted in 20 % lower absolute values for SD length per area compared to the macro (Supplementary Fig. 5a), which is due to the 25 % on average larger ROIs assigned by AMAP (Fig. 3 b,

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig. 5b), whereas the total SD length quantified by AMAP was only 5 % larger (Fig. 3c). Even though there was a difference in absolute values of ROI area and total SD length detected, the respective values correlated robustly between the macro and AMAP (Fig. 3d-e). Consequently, the previously observed decrease of SD length per area in mutant as compared to control mice was reliably detected by our approach (Fig. 3f-g). Notably, the AMAP-derived values also allow us to delineate the decrease in SD length per area with aging (Supplementary Fig. 5c). Taken together, our deep learning approach shows the potential to detect pathological alterations in individual FP morphology as well as in global SD abundance. Furthermore, plotting the mean values of AMAP-derived SD length per area against the levels of albuminuria reveals a similarly robust correlation between SD length and albuminuria as was already shown for the macro-based SD length values (Fig. 3 h).

Performance on human tissue samples

The imaging protocol we have utilized in the above analyses is both lengthy (3-4 days) and requires the use of super-resolved STED microscopy. We have recently published a protocol for fast and straightforward confocal microscopy-based imaging of FPs (referred to as *fast protocol* throughout the manuscript), which reduces the time needed for sample preparation and imaging to only 5 h[9]. We here investigated the applicability of AMAP to images acquired using this fast protocol. To adapt the method to the lower optical resolution of this protocol we generated an additional set of 15 and 40 annotated images of human and mouse tissue, respectively, obtained with the fast protocol. To facilitate the annotation of mouse tissue, we also generated matched images using STED microscopy. We inferred FP and SD segmentations using the network described above, corrected them manually, and used the results as labels of the confocal images (Supplemental Fig. 6a). Images of human tissue were annotated manually (Supplementary Fig. 6b). We added the 55 images to the existing training set and ran another training procedure on the network initially trained on the STED

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

microscopy images. As evaluated by visual inspection, this network retraining resulted in improved segmentation results also for images of lower quality and resolution (Supplementary Fig. 6c).

We tested the adapted segmentation network and the subsequent analysis of morphometric parameters on a set of 43 confocal microscopy images of human tissue. Most of the images, except for the FSGS patient, are part of a previously published set of images[9]. The samples included healthy individuals and a range of patients, including congenital nephrotic syndrome with mutations in the *TRPC6/NPHS2* genes, minimal change disease (MCD), FSGS, and IgA nephropathy.

Based on visual inspection, the adapted segmentation approach offers good accuracy in the spatial resolution-wise more challenging images of human tissue imaged using the fast protocol (Fig. 4a). Moreover, the morphometric parameters of the segmented FPs reflected different degrees of effacement and morphological changes in each patient in agreement with what has been previously published[9] (Fig. 4b-c). In the FSGS mouse model (see Figures 2 and 3), a decrease in SD length is always accompanied by an increase in FP circularity. Interestingly, this correlation was not always observed in the patient samples. For example, the FSGS patient showed a relatively high SD length and a high circularity as compared to the other diagnoses, whereas the opposite was observed for the *TRPC6/NPHS2* patient. This suggests that effacement patterns are complex and that multiple morphometric parameters might be needed to characterize each type of disease comprehensively. We therefore visualized several morphometric parameters (SD length, SD grid index, FP area, FP perimeter, FP circularity (Supplementary Fig. 3)) using umap[20] (Fig. 4d), which illustrates the morphometry-based grouping among the different diagnoses. For example, the FSGS and the MCD patients appear away from one another in the plot. Both of these patients show similar SD lengths (Fig. 4b) but a significantly different circularity (Fig. 4c). Based on visual

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

inspection of the images in Fig. 4a, the general SD density is similar while the morphology of individual FPs differs with the FSGS patient displaying shorter and wider FPs than the MCD patient. The *TRPC6/NPHS2* patient is positioned away from control samples and shows a more extensive spread of data points. This reflects the more focally heterogeneous effacement pattern in this patient as compared to the others. To show the throughput capacity of AMAP we also imaged large field-of-views of 155*155 μm^2, which take around 2-6 minutes to acquire with any high-end confocal microscope. AMAP segmented 6,452 and 176 FPs from the control and *TRPC6/NPHS2* patients, respectively, in < 1 h (Fig. 4 e-f). By comparison, it would take ~ 7 h of hands-on work for a trained user to carry out this analysis with the macro. Moreover, while AMAP detects all recognizable FPs, even very experienced users will only be able to annotate a fraction of all FPs within a reasonable amount of time and effort with the semi-automated macro. While based on a limited number of tissue samples, these results suggest that our fully automated approach for segmentation and morphometric parameter quantification opens up possibilities for a comprehensive, large-scale description of FP effacement across kidney pathologies.

Discussion

In recent years, light microscopy-based imaging of podocyte ultrastructure has gained an enormous momentum. Not only does this technique complement electron microscopy imaging, but also the possibility of efficiently imaging much larger tissue areas enabled the quantitative analysis of morphological changes, especially in FP morphology in disease and disease progression[9,12,19]. Combining quantitative analyses with biophysical modelling of ultrafiltration, our group has previously proposed a model, which mechanistically linked the well-known features of FP effacement with the occurrence of albuminuria[12]. Although powerful, these quantitative analyses still rely on manual assignment of ROIs and a manual

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

step in the segmentation of individual FPs. This is not only time-consuming but also potentially prone to bias.

By taking advantage of the advances in deep learning, we here present a fully automatic segmentation of FPs and the SD. We demonstrate the accuracy of this approach by successfully reproducing previously published results in a mouse model for FSGS[12]. Even though absolute values for each extracted parameter in some cases differ slightly, as compared to our previous analysis using the macro, we show that differences between groups (age, genotype) are highly conserved. Thus, our fully automated approach can readily be applied to quantitatively describe alterations to FP morphology with the same accuracy and substantially higher throughput as our previous approach, while eliminating any potential user bias. Importantly, AMAP confirmed the correlation of SD length and levels of albuminuria, thereby supporting the previously proposed mechanism of albuminuria upon podocyte injury[12].

We further validate that the deep learning network can be adapted to successfully segment the SD and FPs in human patient samples. These images were acquired using a recently published, pathology-optimized, fast, and simple protocol, which utilizes confocal microscopy[9]. Interestingly, we show that multiparameter analysis of FP morphometrics in these patients shows particular groupings of different disease types. This finding indicates that the view of FP effacement as a uniform process might have to be revised and that effacement patterns could differ depending on the underlying disease. We are aware that the presented dataset is limited and that more data points are needed to establish these findings in the future. Still, we suggest that the comprehensive quantitative analysis proposed above, which we named AMAP (Automatic Morphological Analysis of Podocytes) could lead to a better understanding of the FP morphology and its change throughout the disease progression. Depending on the size of the data, the method can be deployed on a local machine or any

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

more advanced computer station equipped with a graphics processing unit (GPU) which substantially speeds up the CNN processing. At this point, the software needs specific coding knowledge in order to run the processing scripts. As the interest in AMAP grows, we will implement it as a user-friendly tool. Apart from saving working hours, incorporating automatically derived morphometric parameters in clinical renal pathology routines could potentially allow for more precise diagnostics due to the additional morphometric data it provides.

In summary, we present AMAP, the first imaging and analysis strategy allowing for non-biased, fully automated quantification of FP morphology at the nanoscale. This investigator bias-free analysis confirms our previous finding that the SD length correlates robustly with levels of albuminuria in a mouse model of hereditary FSGS, thereby supporting the mechanistic link between a simplified podocytes ultrastructure and the occurrence of albuminuria. Importantly, we demonstrate that AMAP is readily applicable to human samples processed with our recently published fast protocol, which no longer requires sophisticated super-resolution microscopy. The combination of AMAP and the fast protocol might in the future allow for multi-scale three-dimensional and quantitative kidney diagnostics using only one sample preparation, imaging, and analysis workflow[15–17].

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

MATERIALS AND METHODS

All methods can be found in the supplement.

ACKNOWLEDGEMENTS

We thank the CECAD Imaging Facility, Cologne, Germany and the Advanced Light Microscopy (ALM) Facility, Solna, Sweden for their support in the acquisition of microscopy data. This work was supported by the Clinical Research Unit (CRU) 329 (KFO 329; A1, A6 and A7) as well as partly by FOR 2743 of the Deutsche Forschungsgemeinschaft (DFG). KB and GS were supported by the North Rhine-Westphalia return program (311-8.03.03.02-147635) and BMBF program Junior Group Consortia in Systems Medicine (01ZX1917B).

DISCLOSURE

All the authors declared no competing interests.

REFERENCES

1. Benzing T, Salant D. Insights into Glomerular Filtration and Albuminuria. *N Engl J Med*. 2021;384(15):1437-1446. doi:10.1056/NEJMra1808786

2. Haraldsson B, Nyström J, Deen WM. Properties of the Glomerular Barrier and Mechanisms of Proteinuria. *Physiol Rev*. 2008;88(2):451-487. doi:10.1152/physrev.00055.2006

3. Amann K, Haas CS. What you should know about the work-up of a renal biopsy. *Nephrol Dial Transplant*. 2006;21(5):1157-1161. doi:10.1093/ndt/gfk037

4. Ichimura K, Miyaki T, Kawasaki Y, Kinoshita M, Kakuta S, Sakai T. Morphological Processes of Foot Process Effacement in Puromycin Aminonucleoside Nephrosis Revealed by FIB/SEM Tomography. *J Am Soc Nephrol*. 2019;30(1):96 LP - 108. doi:10.1681/ASN.2018020139

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5. Suleiman HY, Roth R, Jain S, Heuser JE, Shaw AS, Miner JH. Injury-induced actin cytoskeleton reorganization in podocytes revealed by super-resolution microscopy. *JCI insight*. 2017;2(16):e94137. doi:10.1172/jci.insight.94137

6. Unnersjö-Jess D, Scott L, Blom H, Brismar H. Super-resolution stimulated emission depletion imaging of slit diaphragm proteins in optically cleared kidney tissue. *Kidney Int*. 2016;89(1):243-247. doi:10.1038/ki.2015.308

7. Suleiman H, Zhang L, Roth R, et al. Nanoscale protein architecture of the kidney glomerular basement membrane. *Elife*. 2013;2. doi:10.7554/eLife.01149

8. Unnersjö-Jess D, Scott L, Sevilla SZ, Patrakka J, Blom H, Brismar H. Confocal super-resolution imaging of the glomerular filtration barrier enabled by tissue expansion. *Kidney Int*. 2018;93(4):1008-1013. doi:10.1016/j.kint.2017.09.019

9. Unnersjö-Jess D, Butt L, Höhne M, et al. A fast and simple clearing and swelling protocol for 3D in-situ imaging of the kidney across scales. *Kidney Int*. 2021;99(4):1010-1020. doi:10.1016/j.kint.2020.10.039

10. Pullman JM. New Views of the Glomerulus: Advanced Microscopy for Advanced Diagnosis. *Front Med*. 2019;6:37. doi:10.3389/fmed.2019.00037

11. Siegerist F, Endlich K, Endlich N. Novel Microscopic Techniques for Podocyte Research. *Front Endocrinol (Lausanne)*. 2018;9:379. doi:10.3389/fendo.2018.00379

12. Butt L, Unnersjö-Jess D, Höhne M, et al. A molecular mechanism explaining albuminuria in kidney disease. *Nat Metab*. 2020;2(5):461-474. doi:10.1038/s42255-020-0204-y

13. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-1567. doi:10.1038/s41591-018-0177-5

14. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25(8):1301-1309. doi:10.1038/s41591-019-0508-1

15. Bouteldja N, Klinkhammer BM, Bülow RD, et al. Deep Learning–Based Segmentation and Quantification in Experimental Kidney Histopathology. *J Am Soc Nephrol*. 2021;32(1):52 LP - 68. doi:10.1681/ASN.2020050597

16. Jayapandian CP, Chen Y, Janowczyk AR, et al. Development and evaluation of deep learning–based segmentation of histologic structures in the kidney cortex with multiple histologic stains. *Kidney Int*. 2021;99(1):86-101. doi:https://doi.org/10.1016/j.kint.2020.07.044

17. Zimmermann M, Klaus M, Wong MN, et al. Deep learning-based molecular morphometrics for kidney biopsies. *JCI Insight*. March 2021. doi:10.1172/jci.insight.144779

18. Brabandere B De, Neven D, Gool L Van. Semantic Instance Segmentation with a Discriminative Loss Function. 2017. arXiv:1708.02551

19. 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):11473. doi:10.1038/s41598-017-11553-x

20. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. 2020. arXiv:1802.03426

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

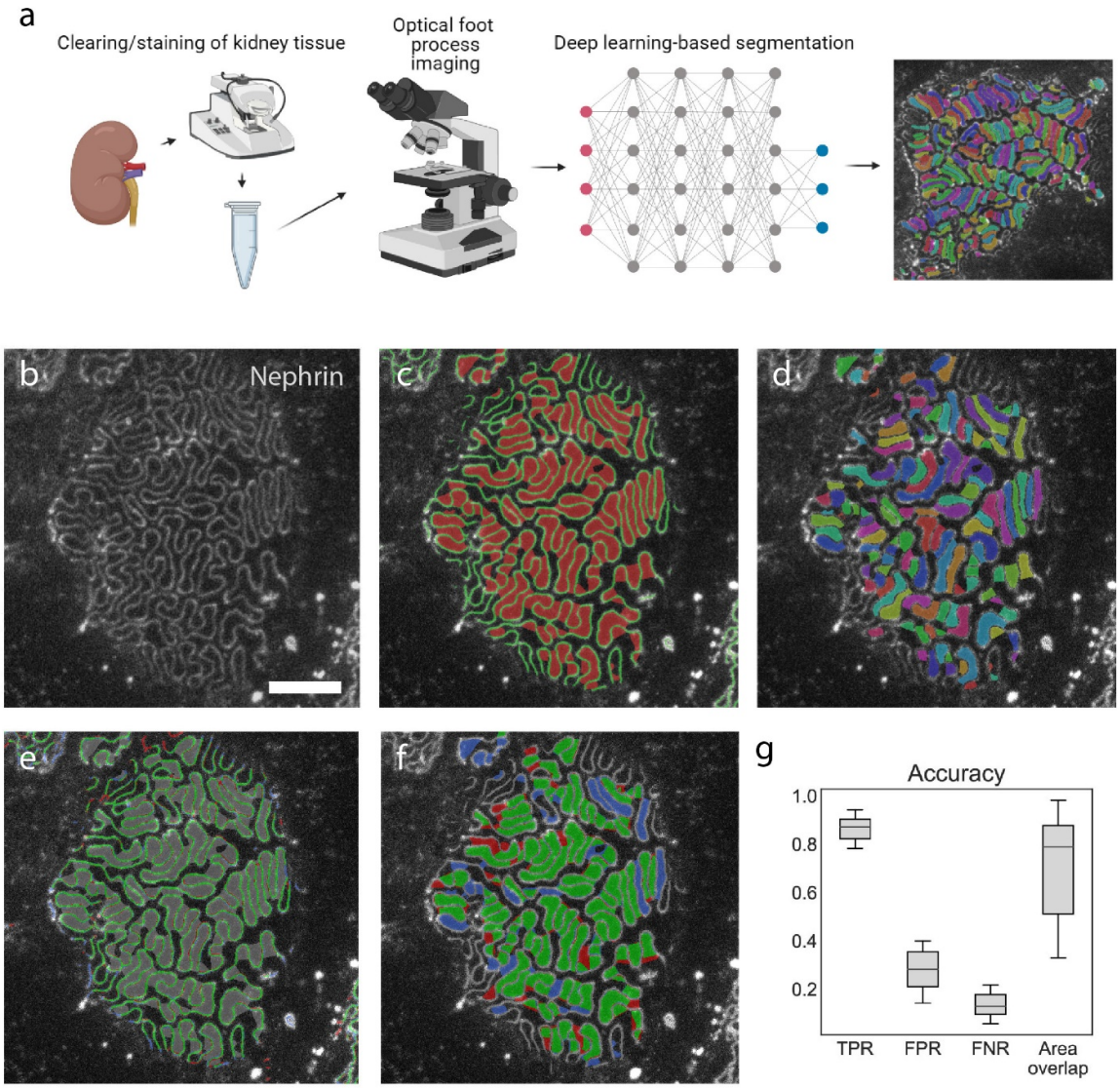
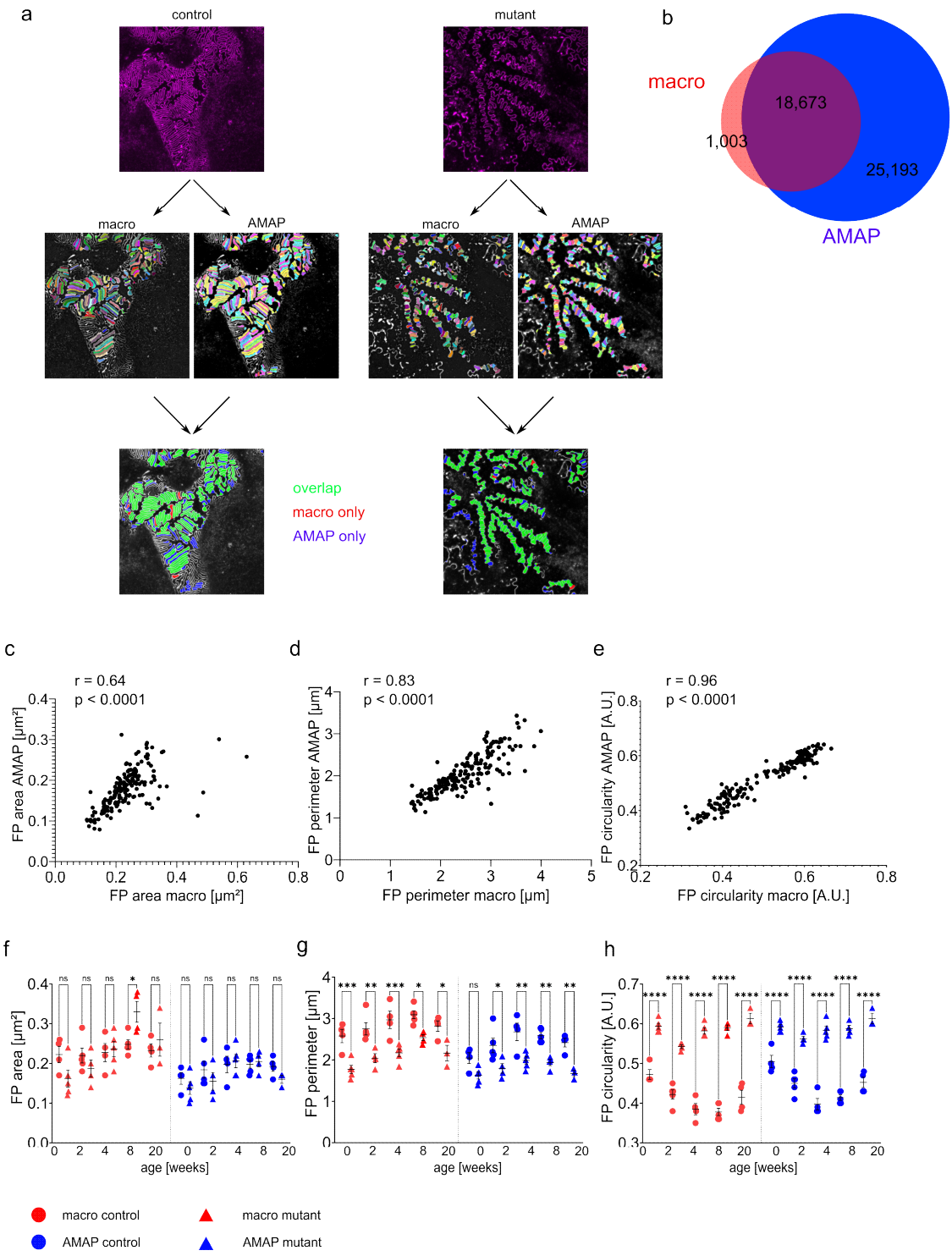


Figure 1. Overview of the imaging and analysis workflow and segmentation. (a) Schematic overview of AMAP. Previously published sample preparation and imaging protocols are used for visualization of FPs. Convolutional neural network is applied to extract FP and SD regions allowing for fully automatized segmentation of FPs in kidney samples from mice and humans. (b) Raw nephrin-stained STED image from the test set. Scale bar 2 µm. (c) Outcome of semantic segmentation. FP pixels are marked in red, SD pixels in green. (d) Outcome of instance segmentation. Separate FP instances are marked with different colors. (e) Accuracy of SD semantic segmentation. Pixels correctly predicted as SD are marked in green, those

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

labeled but not predicted as SD (false negatives) are marked in red, those predicted but not labeled as SD (false positives) are marked in blue. (f) Same as (e) for FP pixels. (g) FP instance prediction accuracy in the entire test set of 209 pictures. True positive rate (TPR), false-positive rate (FPR), false-negative rate (FNR) are quantified relative to the number of manually determined FPs in a given image. Area overlap is quantified as intersection over union (IoU). FPR includes a large proportion of not labeled FPs and does not accurately reflect the actual errors in prediction.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2. Comparison of FP and SD morphology in podocin$^{R231Q/A286V}$ ("mutant") and control mice from a published macro-based dataset with AMAP. (a) Illustration of the macro and AMAP FP segmentation in images of a control and a mutant mouse. Differences and overlap of detection are color-coded as indicated. (b) Venn diagram showing the overlap and differences in FP detection. AMAP detected 18,673 of the total number of 19,676 FPs that were assigned using the macro. 25,193 FPs were additionally assigned by AMAP. (c)-(e) Correlation of FP area (c), FP perimeter (d) and FP circularity (e) between the macro- and AMAP-assigned FPs. Measurements of all three parameters correlate significantly between the approaches. Each dot represents the values originating from one image (n = 174 images). r = Pearson correlation coefficient, p = p-value. (f)-(h) Comparison of the values for FP area (f), FP perimeter (g) and FP circularity (h) in macro-assigned (red) or AMAP-assigned (blue) FPs in control and age-matched mutant mice. With the exception of FP area at 8 weeks and FP perimeter at 0 weeks, the detection of significant differences is comparable between the approaches. Each dot/triangle represents one mouse. Data are presented as mean ± SEM. Sidak's multiple comparison test was performed to determine statistical significance. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

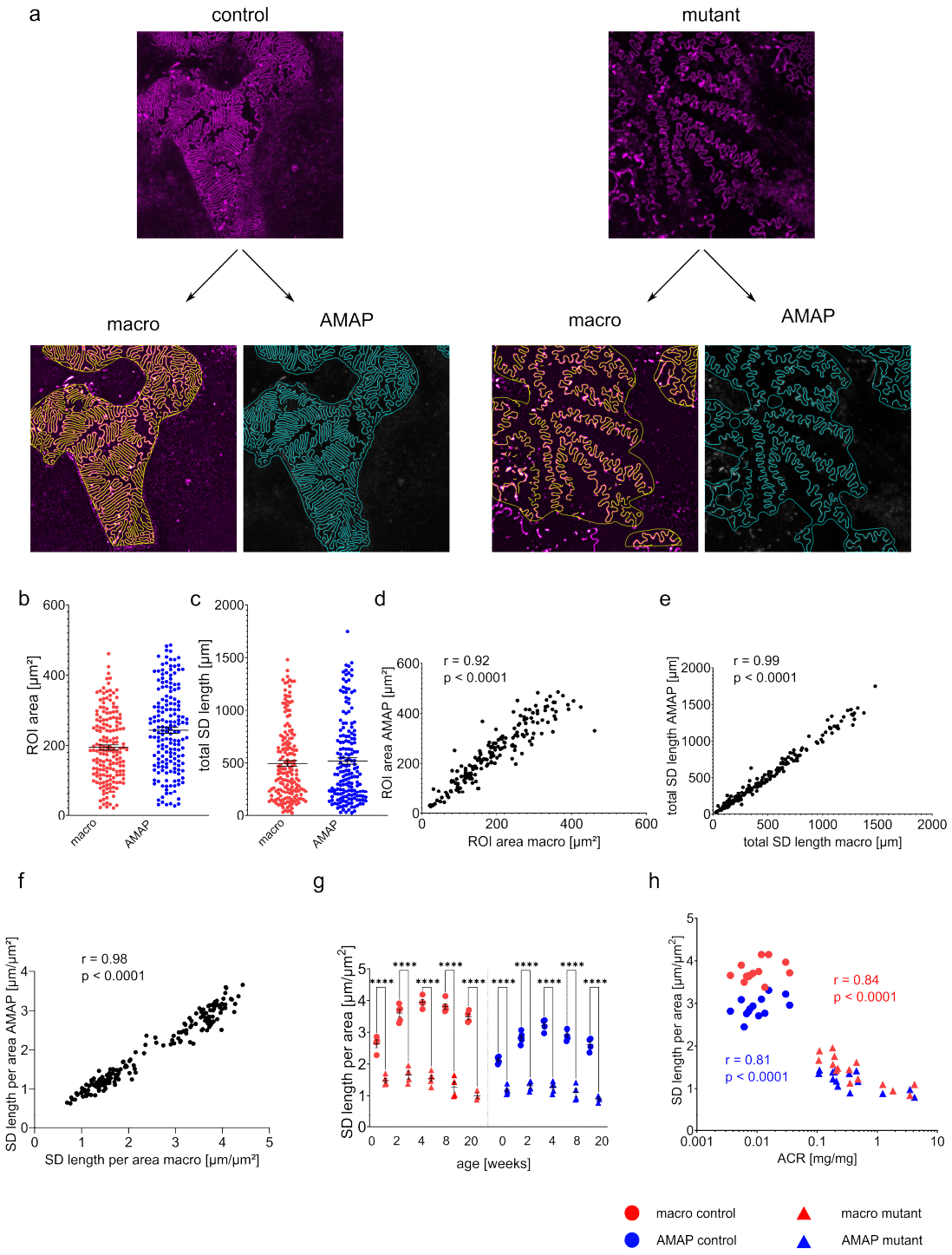


Figure 3. Comparison of SD length in podocin$^{R231Q/A286V}$ ("mutant") and control mice calculated with the macro and AMAP. (a) Illustration of the quantification of the SD length and ROIs with the two approaches in images of a control and a mutant mouse. (b) Area of

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

manually assigned (red) and AMAP-assigned (blue) ROIs. Each dot represents the ROI area of one image (n = 174 images). Horizontal bars represent the median. (c) Total SD length quantified with the macro (red) and AMAP (blue) ROIs. Each dot represents the ROI area of one image (n = 174 images). Horizontal bars represent the median. (d)-(f) Correlation of ROI area (d), total SD length (e) and SD length per area (f) of macro- and AMAP-assigned ROIs. Each dot represents the ROI area of one image (n = 174 images). r = Pearson correlation coefficient, p = p-value. (g) Comparison of the SD per area values in macro- (red) or AMAP-assigned (blue) ROIs in control and age-matched mutant mice. Differences between genotypes are equally detected between the two approaches. Each dot/triangle represents one mouse. Data are presented as mean ± SEM. Sidak's multiple comparison test was performed to determine statistical significance. **** $p < 0.0001$. (h) Scatter plots of the mean SD length per area (macro in red, AMAP in blue) against the urinary albumin creatinine ratio (ACR). Two-tailed Spearman's rank correlation was used to determine statistical significance. r, Spearman's rank coefficient.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

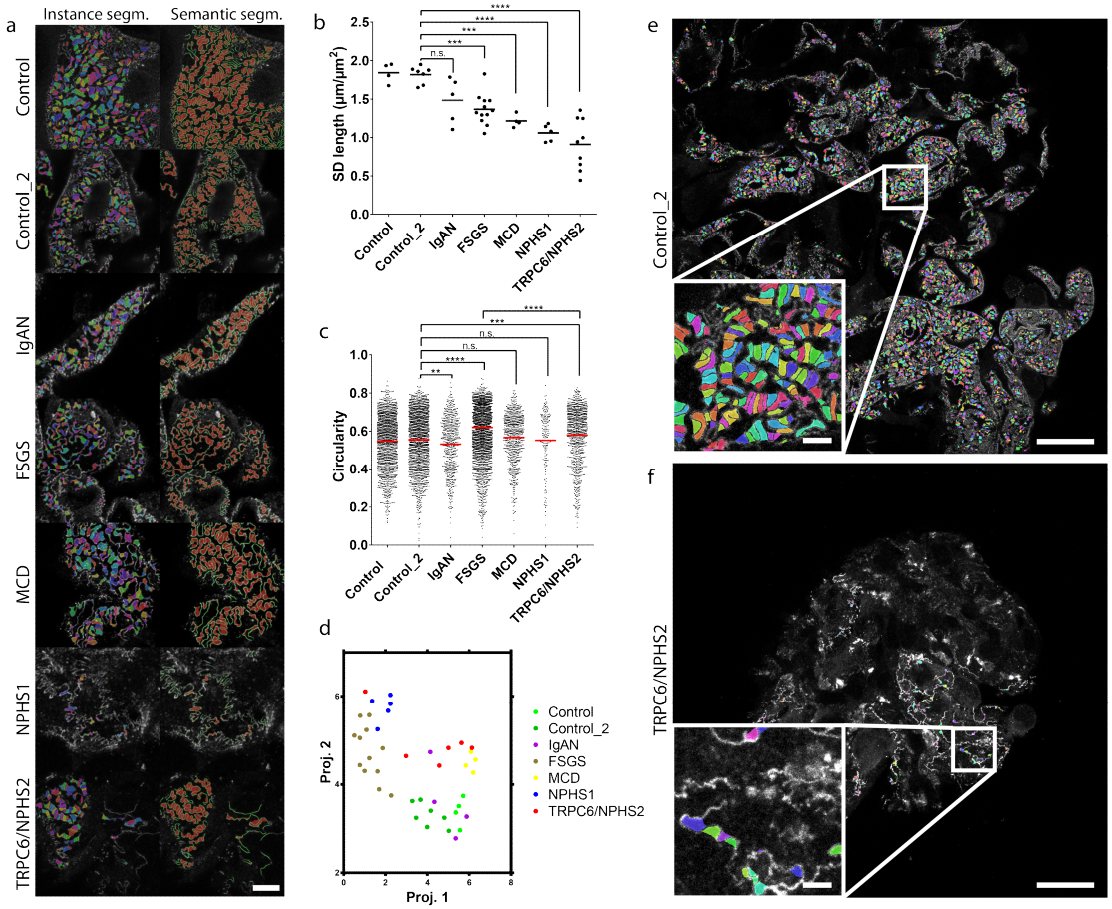


Figure 4. AMAP-based FP morphometry in human patients in images acquired with a confocal microscope using a new 5-hour-sample preparation protocol. (a) Each row represents one patient as indicated with raw images overlaid with the results of instance (left) and semantic (right) segmentation. All samples were stained for nephrin with Alexa-405 or Alexa-488 except for the NPHS1 patient's sample, which was stained for podocin with Alexa-555. Scale bar 5 µm. (b) SD length per area for all patients shows a significant decrease for all patients except for the patient with IgAN. Each dot represents one image (one patient per diagnosis). The black line represents the mean. Tukey's multiple comparison test was performed to determine statistical significance. *** $p < 0.001$, **** $p < 0.0001$. (c) Circularity score of FPs from all patients. Each data point represents one FP. Interestingly, only the FSGS and the TRPC6/NPHS2 patients show significant increases compared to

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.14.448284; this version posted June 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

control, indicating different morphometric fingerprints for different diagnoses. The red line represents the mean. Dunnet's multiple comparison test was performed to determine statistical significance. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. (d) Multiparameter projection (umap) of all patients. (e-f) 155 um field-of views maximum intensity projections of z-stacks of a control patient (a) and a sclerosed glomerulus of the *TRPC6/NPHS2* patient. Insets show zoomed views of the indicated areas. 6,452 FPs are segmented from the image in (e), showing the data throughput capacity of AMAP. Scale bars 20 µm and 2 µm (inset).