A multiscale atlas of the molecular and cellular architecture of the human kidney

Elizabeth K. Neumann,[1,2] Nathan Heath Patterson,[1,2] Leonoor E.M. Tideman,[3] Lukasz G. Migas,[3] Madeline E. Colley,[1,2] Melissa A. Farrow,[1,2] Jamie L. Allen,[1,2] Emilio S. Rivera,[1,2] Carrie E. Romer,[1,2] Haichun Yang,[4] Maya Brewer,[4] Kavya Sharman,[2,5] Raymond C. Harris,[4,6,7] Agnes B. Fogo,[4,6,7] Danielle B. Gutierrez,[1,2] Mark P. de Caestecker, [4] Richard M. Caprioli, [1,2,8] Raf Van de Plas,[1-3] Jeffrey M. Spraggins[1,2,8,9]

[1]Department of Biochemistry, Vanderbilt University, Nashville, TN, USA 37232.
[2]Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, USA 37232.
[3]Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands.
[4]Division of Nephrology and Hypertension, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN USA 37232.
[5]Program in Chemical & Physical Biology, Vanderbilt University Medical Center, Nashville, TN, USA 37232.
[6]Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN USA 37232.
[7]Departments of Medicine and Pediatrics, Vanderbilt University Medical Center, Nashville, TN, USA 37232.
[8]Department of Chemistry, Vanderbilt University, Nashville, TN, USA 37232.
[9]Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN, USA 37232.
*Correspondence Author: Jeff.Spraggins@Vanderbilt.edu

Summary

The need for deep molecular and spatial characterization of tissues is growing rapidly to study complex diseases and disorders, particularly in the context of relationships between the cellular and molecular architecture from single cells to functional and anatomical regions of tissue. The human kidney is composed of over 26 cell types that actively coordinate to form higher-order structures, such as the nephron. It is not yet understood how these structures vary throughout a single organ or among organs within the human population. Towards this end, we have developed an extensive molecular and cellular atlas of the human kidney consisting of over 3 million cells comprising 75,000 functional tissue units from 13 patients. This atlas was developed using a spatially registered and integrated suite of technologies consisting of imaging mass spectrometry, histological staining, autofluorescence microscopy, and multiplexed immunofluorescence to comprehensively probe centimeter-sized areas of tissue. Through this approach, the cellular organization and metabolomic profiles of glomeruli, proximal tubules, distal tubules, and collecting ducts were determined as well as their intra- and inter-patient variance. Relating these lipid profiles to canonical cell types allowed us to hypothesize the functional role of specific phospholipids that have not previously been described, particularly how they relate to differences in BMI and sex. These integrated data provide a valuable reference for kidney researchers and are publicly available through the NIH Human Biomolecular Atlas Program (https://portal.hubmapconsortium.org/).

Introduction

While the roles of specific cell types and functional tissue units (FTUs) are generally well known, the molecular and numerical heterogeneity throughout a single organ or between organs within the human population has yet to be determined, particularly as a function of patient demographic and anthropometric indicators such as sex and obesity, respectively. Exploring the underlying similarities and

differences between these patient factors is key for extrapolating the drivers of functional efficiency, transition to disease, and disease severity[1-3]. Robust atlases enable these types of questions to be answered and notable atlas efforts include the Allen Brain Atlas[4] and Human Cell Atlas[5,6]. Recently, the National Institutes of Health and private organizations have funded atlas efforts, such as the Human Biomolecular Atlas Program (HuBMAP)[7], BRAIN Initiative[8], Kidney Precision Medicine Project[9], and Human Tumor Atlas Network[10], targeting different healthy as well as diseased organs. Collectively, these initiatives constitute hundreds of scientists from across the world. Here, as part of HuBMAP, we explore the human kidney to investigate how it varies on both the molecular and cellular level within and among different patients, particularly as it relates to obesity and sex.

The kidney is responsible for waste management, electrolyte balance, blood pressure control, and red blood cell production among other functions[11-13]. To execute these myriad functions, this complex organ consists of at least 26 distinct cell types and over a million nephrons[14,15]. A nephron is a functional unit that consists of a glomerulus that is linked to distal and proximal tubular segments, capillaries, lymphatics, and peritubular interstitial spaces[16]. This dynamic network of cell types and multicellular functional tissue units not only varies throughout the kidney itself, but also from one human to another[17]. For instance, the physiological differences in renal function and disease development have been well studied between men and women, including blood pressure regulation[18], diabetes[19], chronic kidney disease[20], and proximal tubule composition[21]. While many studies attribute these differences to sex hormones or size differences[22], none have profiled the spatial or molecular landscape of the kidney with the level of coverage provided in this work. The collection of spatial and molecular profiles across the tissue allows for the creation of a tissue atlas, which can be explored and leveraged to determine how differences between groups manifest physiologically [23,24].

The atlas presented here consists of expansive, integrated, and spatially resolved molecular and cellular information for 13 human patients at unprecedented specificity and scale; we visualize single cells and functional tissue units across centimeters of tissue. This atlas incorporates spatially resolved data from mass spectrometry, highly multiplexed immunofluorescence, histological stains, and autofluorescence microscopy. Each component of the atlas is publicly available (https://portal.hubmapconsortium.org/). Our atlas spans multiple scales of spatial granularity, from large tissue areas to single cells and from averaged populations to individual patients, serving as an exceptionally rigorous resource for exploring the biomolecular landscape that can inform health and disease.

Results and Discussion

Overview of atlas data and methods

There are several key considerations before constructing an atlas derived from data collected over a long time period (months to years): reproducibility, sampling area, sample number, and chemical coverage (Figures 1A and S1)[25]. To begin, remnant tissue with warm and cold ischemia times shorter than 5 min and 30 min, respectively, is collected from unilateral nephrectomies (Figures 1B and S1A-C) where approximately half of the tissue is formalin-fixed and paraffin-embedded (FFPE) for histological (Figure S1D) and pathological assessment (Table S1-3). The other half is fresh frozen (Figure S1E) for subsequent processing for different analytical approaches (Figures 1C and S1G-J), allowing both cellular and molecular profiling of major kidney functional tissue units. Using this workflow, we have acquired 13 datasets from donors with age ranging from 20 to 77 years old, a body mass index (BMI) between 22 and 45.5, and which include both sexes (Table S2). It is estimated that 40% of biomarker related studies

in the nephrology field are misleading because of the exclusion of women[26]. To combat this, we have selected tissues from donors that cover a wider range of patient demographics and clinical factors. This enables us to explore the molecular and cellular heterogeneity of the kidney as it relates to BMI and sex. In brief, datasets consist of autofluorescence (AF) microscopy (Figures S2-3), matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI IMS) in both positive (Figures S4-S31) and negative ion modes (Figure S32-S59), co-detection by indexing (CODEX) multiplexed immunofluorescence (Figures S60-S66, Table S4), and PAS staining (Figure S67). Each modality provides a different, complementary view of the tissue architecture. For instance, MALDI-IMS provides untargeted, spatial information for hundreds of metabolites and lipids nearing cellular resolution[27]. Autofluorescence (AF) microscopy is a label-free microscopy method, which is used as the basis for computational segmentation of the tissue section into known FTUs[28]. The molecular markers corresponding to specific FTUs can subsequently be extracted from the multimodal combination of IMS-supplied mass spectra and AF-supplied FTU masks generated using an interpretable machine learning workflow (Figure 1D.1). Our workflow consists of classifying pixels into FTUs on the basis of their mass spectra using a decision-tree-based gradient boosting model and then quantitatively estimating the importance of each molecular species for recognizing a given functional unit using Shapley additive explanations (SHAP)[29,30]. By bridging the observations from the different imaging modalities and mining their cross-modal relationships in an automated way, it becomes possible to estimate unit-specific molecular markers and empirically determine metabolomic profiles for the segmented structures. To ground these broad label-free modalities and their observations to the body of protein literature, we also employ more targeted modalities such as CODEX multiplexed immunofluorescence microscopy to spatially label and report canonically established cell types and cell states down to sub-cellular resolution. The spatial resolution of these multiplexed fluorescence approaches is 250 nm (diffraction limited) for >60 protein targets, enabling cellular neighborhood analysis (Figure 1D.2)[31-33]. Furthermore, AF microscopy is not only used to automatically obtain FTU masks, but also to establish a common spatial coordinate framework across all modalities and tissue sections. For this reason, AF microscopy is collected on all tissue sections and enables registration and integration of MALDI IMS, multiplexed immunofluorescence, and histological stains[34,35]. Further details are found in the supplemental methods and published elsewhere as a protocol[36] and in the references cited above.

Overview of atlas composition from single cells to functional tissue units

Our atlas provides molecular and cellular profiles across the population of patients and within individual patients, allowing us to explore the diversity within and among human samples (Figure 2). Overall, we detected an average of 222±19 lipids (Figure 2A, Table S5-6) and 3255±754 cells/mm^2 (Table S7) within these samples with extremes of 2100-4336 cells/mm^2 and 192-258 lipid features, respectively. Lipids vary in their relative abundance within individual patients, although the molecular markers associated with different FTUs are largely conserved (Figure 2A). These patterns seem to report genuine biologically relevant connections rather than patient-specific observations. For instance, PC(16:0_18:1) has higher discriminative importance, as measured by its SHAP score, for collecting ducts and its relative signal intensity is generally conserved from patient to patient, while lipids such as PA(36:2) have a high intensity within FTU for VAN0029 (62-year-old, white male, BMI 34.9) compared to the average. The relative composition of cell types is different from patient to patient, especially for cells expressing nestin and ß-catenin (Figure 2B, Table S8-9). Beyond individual cells, we segmented each tissue into discrete functional units (Table S10), with a total of 863 glomeruli, 8008 collecting ducts, 10662 distal tubules, and 55569 proximal tubules segmented within all images (Figure 2C, Figure S68).

To explore how the cellular architecture varies from patient to patient (Figure 2D-E), we calculated the cellular neighborhood composition within a 25 µm neighborhood surrounding each cell for all patients (Figure 2D, Table S11) and individual patients (Figure 2E, Table S12-S19). In brief, these analyses serve a semi-quantitative form of spatial phenotyping with a panel consisting of both classic and less commonly used protein markers[37]. Using these neighborhoods analyses, we determined overall trends with cell types that were correlated (e.g. PARP1 and CD31 expressing cells), uncorrelated (e.g. nestin and ß-catenin expressing cells), and anticorrelated (e.g. renin and KDR expressing cells), using these neighborhoods analyses, although the exact Spearman correlation coefficients varied from patient to patient. Spearman correlation values above 0.7 (highly correlated) are particularly interesting since these suggest conservation of the cellular architecture across the millions of cellular neighborhoods calculated within this dataset. Because cells have unique, but interdependent functions, cells that are often in neighborhoods with one another may interact more regularly. While some of these trends are known, this atlas provides a broad overview of cellular neighborhoods across large tissue areas from 13 patients. Patient-to-patient differences in correlation coefficients likely have functional relevance and may serve as a heuristic for neighborhood reorganization as a function of organ region, age, or pathological state.

Spatially delineating masks for each FTU type, in combination with the interpretable machine learning based detection of FTU-specific markers, enables us to determine molecular marker profiles for (Figures S69-S73 and Table S20), distal tubules (Figures S74-S78 and Table S21), proximal tubules (Figure S79-S83 and Table S22), and collecting ducts (Figures S84-S88 and Table S23). This exploration comprises most segments within the nephron, and its visualizations covers both individual patients and patient population composites. Moreover, we calculate the SHAP score for each lipid to determine which molecules are characteristic of each tissue feature (Figures S89-S90). Our interpretable machine learning computation provides us with molecular features that characterize the functional unit most appropriately[29]. Many molecular features appear to have approximately long normal distributions, such as PA(O-38:1) within glomeruli (Figure S73) and SM(d16:1_20:0) within proximal tubules (Figure S83). The unimodal distributions potentially indicate a single population distribution exists. Conversely, several features, such as PC(16:0_16:0) (Figure S73&S78) and PE(O-38:5) (Figure S78), appear to be bimodal and PE-Cer(40:1) is more complex (Figure S78). This perhaps indicates different physiological states of the tissue units and lends to the diversity and dynamic nature of phospholipids. For instance, cellular membranes with a higher accumulation of PC(16:0_16:0) may be in a more stable, less physiologically active state compared to a lower relative composition.[38]

Spatial metabolomic differences as it relates to obesity and sex

While we are developing an atlas of normal human tissue, characteristics, such as obesity and sex, significantly contribute to the molecular and cellular composition of each patient and should not be ignored. Thus, we can also more granularly analyze this atlas to determine molecular features related to BMI (Figure 3A, Figure S91, and Table S24) and sex (Figure 3B, Figure S92, and Table S25) using the same computational approaches previously described. Stark differences can be observed when subdividing the patients with BMIs above or below 30 as a proxy for obese and non-obese, respectively. For example, we have determined that PAz-PC, PS(26:0), and PE(O-38:5) are some of the top markers with the most interesting distributions for differentiating obese from non-obese samples (Figure 3A.1). Furthermore, cellular neighborhoods involving calbindin or aquaporin 2 (among others) are potentially disrupted between obese and non-obese individuals (Figure 3A.2). Akin to the cellular neighborhood analysis, we calculated Spearman correlation coefficients for metabolites within a one-pixel radius (10

µm) of one another and we determined that there are substantial differences in the molecular neighborhood correlations between obese and non-obese patients (Figure 3A.3). While each of these analyses tells us something different about the system, one of the most notable findings involves PAz-PC which is an oxidized lipid more abundant within the organs of obese patients, which was computationally determined to be the largest discriminator between obese and non-obese patients (Figure 3A.1, Figure S91), and it is known to have higher in abundance in kidneys undergoing renal failure[39]. While it is not understood if the oxidized phosphatidylcholines are produced as a result of or to combat a disease state, their higher abundance in obese patients indicates poorer kidney health. Other lipids, such as PE(O-38:5), show large disparities between obese and non-obese patients based on the violin plots, providing further targets to explore. For instance, PE(O-38:5) may promote the reabsorption of metabolites and proteins from urine as high abundances of this lipid induces negative membrane curvatures that permit endocytosis[40]. Kidneys compensate for higher body mass by undergoing hyperfiltration, eventually leading to disease[41]. This may explain the higher abundance of PE(O-38:5) within obese patients. The functions of most discrete lipids are elusive, and this serves as preliminary evidence to support the essential roles that lipids may possess.

The molecular information can be further supplemented by cellular neighborhood analysis with CODEX analysis of glomerular markers (Figure S93) and tubular markers (Figure S94). Most notably, correlation coefficients decrease for calbindin-expressing cells and MARCKS expressing cells within obese patients compared to their non-obese counterparts. Differences within calbindin expression may indicate reduced absorption of calcium[42], which may be dysregulated in obesity[43]. Differential MARCKS expression may indicate dysregulated sodium reabsorption[44] which is also affected by obesity[45]. Changes within these correlation values perhaps indicate that there is a change in the relative neighborhood that may be required to enable hyperfiltration among other responses. Additional comorbidities or patient information were explored for their effect on cellular neighborhoods (Table S26-S44), and changes in the correlation coefficients could be observed, perhaps indicating restructured cellular neighborhoods between obese and non-obese individuals. While additional patients are required to fully characterize how cellular neighborhoods change as a function of demographic or disease, these initial results support further investigation of cellular neighborhoods within these different tissue parameters, particularly with additional molecular context. This may help provide context to disease progression and the inclusion of additional comorbidities stands to better stratify this disordered state.

Beyond BMI, we explored the molecular differences between male and female patients (Figure S92, and Table S25). Overall, PE(O-38:5) and PS(26:0) accounted for the largest difference between the sexes based on the SHAP analysis. Further inspection of the violin distribution plots shows that both lipids displayed bimodal-like distributions in male patients (Figure 3B.1). Other lipids that comprise male and female differences have complex distributions and demonstrate the diversity and complexity of phospholipids. Additionally, cellular neighborhoods for seven of the 23 marker panel are strongly correlated in men, but only weakly correlated in women (Figure 3B.2). These differences perhaps are a result of multiple neighborhood compositions existing in women with less neighborhood variation in men. This may affect the molecular distributions. Additionally, we determined that there are stronger molecular neighborhood correlation coefficients calculated for men than for women (Figure 3B.3). Through these analyses, we can parse lipids that are highly correlated within a 10 µm neighborhood, such as PE(18:2_20:0) and PE(18:1_18:1), which may indicate that they share either functionally relevant roles in membrane fluidity[46,47] enabling metabolite transport[48-50] or other cellular functions[51,52] as they are known to induce a negative membrane curvature[53]. However, additional validation studies are

required for understanding the exact function (Figure 3B.3). The Spearman correlation coefficients of molecules detected within male and female patients are generally similar, indicating that many phospholipids are likely conserved between the sexes, at least within the context of 10 µm spatial neighborhoods. There are some exceptions to this trend, such as PS(26:0), which has a mostly uncorrelated relationship to other lipids within kidneys from female patients but a higher correlation coefficient with other lipids in kidneys from male patients (Figure 3B.3). This trend may result in phenotypic or disease severity differences between men and women and would be a key candidate for additional research. Additional lipids that show this distribution include CerP(36:2;2O).

Nephron tubular segment composite profiles based on obesity and sex

While metabolomic profiles for the entire tissue provide broad trends, each tissue can be functionally reduced into different segments within a nephron. Metabolomic differences as a function of obesity or sex may differ when examining different functional regions within the kidney and are difficult to examine with bulk approaches. Thus, we extended these analyses to determine differences related to obesity (Figure 4) or sex (Figure 5) within the FTUs of the nephron, such as glomeruli (Figure S95-96 and Table S45-46), proximal tubules (Figure S97-98 and Table S47-48), distal tubules (Figure S99-100 and Table S49-50; and collecting ducts (Figure S101-102 and Table S51-52) as opposed to the whole sampled region. The differences between FTUs of obese and non-obese patients are starker than those for sex. For instance, the lipids that differentiate obese and non-obese patients are conserved among many of the nephron segments: PAz-PC, PC(26:0), PE(O-38:5), and PC(16:1_16:0) (Figure 4.2). As such, these lipids may be a renal response to an obese phenotype and serve as putative therapeutic or diagnostic markers. As an example, obesity represents a chronically inflamed state and under these conditions, coagulation factors are aberrantly expressed or regulated, promoting increased prothrombotic risk[54,55]. The oxidized phospholipid PAz-PC could be an important molecule in the feedback loop that exists between obesity and risk factors for cardiovascular disease as PAz-PC inhibits the tissue factor pathway inhibitor (TFPI)[56,57]. TFPI inhibits the tissue factor dependent, or extrinsic coagulation, pathway by directly and indirectly perturbing the activation of the serine protease Factor Xa that interacts with fibrin, a marker of obesity[58] and linked to cardiovascular disease[59]. Further, PE(O-38:5) may enable hyperfiltration across the entire nephron, as discussed above. While putative, this is the type of insight that can be gleaned from the atlas information, enabling future functional studies aimed at understanding the dynamic role of lipids in physiological function. The molecular neighborhoods are also quite different between obese and non-obese patients indicating that the lipidome is affected, with likely phenotypic effects. Moreover, the cellular neighborhoods of obese patients have lower Spearman correlation coefficients compared to their non-obese counterparts, indicating differential neighborhoods between the two states. This is exemplified by cells expressing calbindin (Figure 4D.1 and E.1; other proteins show this trend as well), which are moderately correlated in non-obese patients and weakly negatively correlated in obese patients.

Further, molecular differences between male and female patients may be related to hormones, which have a known effect on renal function[19,60-66]. However, different lipid distributions exist for each functional unit perhaps indicating a differential response to these hormones[67-69], although the relationship is complicated[70] (Figure 5B-E.2). Largely, each FTU has a different set of top markers. Despite these differences, molecular neighborhoods in men and women are largely similar (Figure 5.3), indicating that a level of global conservation and high correlations (e.g. PA(36:1) and PS(18:1_18:0), Figure 4B.3 or PE(16:0_18:2) and PE(18:2_18:2), Figure 5C.3) between two lipids may indicate a

functional codependence. Additionally, lipids that do not differ significantly between male and female functional units may serve a conserved, housekeeping role. When considering the cellular neighborhoods of markers known to colocalize within a functional unit (Figure 5.1), neighborhoods including cytokeratin 7 within both distal tubules (Figure 5D.1) and collecting ducts (Figure 5E.1) show the largest difference between the sexes. Cytokeratin 7 is a structural polypeptide that is involved in epithelial differentiation and repair within the kidney[71], perhaps lending to a hypothesis for differential response to disease severity.

Conclusion

Here we have constructed a multimodal, high-dimensional atlas of the human kidney, including molecular, cellular, and histological imaging data, as well as pathophysiological information for each patient. As part of the NIH HuBMAP consortium, the data is freely accessible to the greater community (https://portal.hubmapconsortium.org/). As an initial demonstration, we show data comparing the molecular and cellular neighborhood differences between donors with BMIs below or above 30 as a proxy for obesity, and differences between men and women. Additional insight into the cellular and molecular differences between patients from different demographics may help explain the functional ramifications resulting in differential severity in diseases, such as chronic kidney disease, cancer, and diabetes[72-78]. Lipids, particularly in the context of functional tissue units and cellular neighborhoods, are chronically understudied compared to their transcript and protein counterparts, despite their clear involvement in different diseases and disorders. Many large studies assume that genetic expression is sufficient for understanding disease states. While likely true in many if not most cases, there is a known discordance in gene expression and metabolic production, indicating that further development in this area is required for understanding complex physiology with outcomes tied to personalized medicine and improved therapeutics. As a step towards reducing this discrepancy, our atlas serves to catalog the cellular and metabolite differences on a multiscale level. Ultimately, multidimensional and multiscale atlas efforts, such as the one discussed here, provide several key findings for a defined, targeted validation, and serve as a resource for others to use to corroborate orthogonal studies.

Acknowledgements

Support was provided by the NIH Common Fund and National Institute of Diabetes and Digestive and Kidney Diseases (U54DK120058 awarded to J.M.S. and R.M.C.), the National Science Foundation Major Research Instrument Program (CBET–1828299 awarded to J.M.S. and R.M.C.). E.K.N. is supported by a National Institute of Environmental Health Sciences training grant (T32ES007028) and National Institutes of Diabetes and Digestive and Kidney Diseases (1F32DK128887-01A1 awarded to E.K.N). The Cooperative Human Tissue Network is supported by the NIH National Cancer Institute (5 UM1 CA183727-08). R.V. and L.G.M. were furthermore supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Dutch Research Council), ZonMw, FLAG-ERA, and the European Commission through the FLAG-ERA III JTC project SMART BRAIN (NWO 680-91-319 under NWO-domain Exacte en Natuurwetensch., together with NWO-domain Sociale en Geesteswetensch.), which is a Partnering Project to the EU's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 (Human Brain Project SGA2). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Nederlandse Organisatie voor Wetenschappelijk Onderzoek, ZonMw, FLAG-ERA, or the European Commission.

Author Contributions

EKN Designed experiments, performed autofluorescence microscopy, MALDI mass spectrometry, CODEX multiplexed immunofluorescence experiments, and histological staining. Performed cellular and molecular neighborhood analysis. Performed data interpretation in consult with other authors. Conceptualized and wrote the manuscript with edits from other authors. Organized and oversaw the project. **NHP** Performed image registration, CODEX image preprocessing, and segmentation of autofluorescence microscopy. Managed and uploaded data to HuBMAP data repositories in addition to coordinating data transfer between other coauthors. Provided key insight into project development and key input to manuscript preparation. **LEMT** Performed SHAP analyses and provided data to construct each IMS data figure. Provided key input into manuscript writing. **LGM** Performed both intra and inter MALDI IMS data preprocessing including peak alignment, calibration, normalization, and extraction. Provided input into manuscript writing. **MEC** Performed lipidomics sample preparation and generated lipid library. **MAF** Provided key input into data interpretation, analysis, and manuscript preparation. **JA** Performed experimental sample preparation as well as managed tissue repository. **ESR** Aided in lipidomics experiments. **CR** Helped with CODEX multiplexed immunofluorescence sample preparation. **HY** Performed pathological assessment in conjunction with AFB. **MB** Performed tissue embedding and freezing. **KS** Assisted in AF segmentation workflows. RH Provided key insight into project development and manuscript preparation. **ABF** Supervised HY, performed pathological assessment, provided manual segmentation masks for training of automatic segmentation, provided insight into the CODEX antibody panel, and aided in data interpretation. **DBG** Supervised JA and CR and participated in manuscript editing and compilation. **MRdC** Supervised MB and managed IRB protocols and sample acquisition. Helped with data interpretation. **RMC** Supervised EKN, NHP, JA, MC, ESR, MF, CR, KS, and DG. Obtained funding and contributed to manuscript development. **RP** Supervised LEMT and LGM. Provided contributions to mass spectrometry data analysis and input in manuscript writing. **JMS** Supervised EKN, NHP, JA, MC, ESR, MF, CR, KS, and DG. Conceived and directed the project and participated in manuscript writing.

References:

1 Gambara, V., Mecca, G., Remuzzi, G. & Bertani, T. Heterogeneous nature of renal lesions in type II diabetes. *Journal of the American Society of Nephrology* **3**, 1458-1466 (1993).

2 Toyota, E. *et al.* Global heterogeneity of glomerular volume distribution in early diabetic nephropathy. *Kidney International* **66**, 855-861 (2004).

3 Najafian, B., Alpers, C. E. & Fogo, A. B. in *Diabetes and the Kidney* Vol. 170 36-47 (Karger Publishers, 2011).

4 Sunkin, S. M. *et al.* Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Research* **41**, D996-D1008, doi:10.1093/nar/gks1042 (2012).

5 Regev, A. *et al.* The Human Cell Atlas. *eLife* **6**, doi:http://dx.doi.org/10.7554/eLife.27041 (2017).

6 Rozenblatt-Rosen, O. *et al.* Building a high-quality Human Cell Atlas. *Nature Biotechnology* **39**, 149-153, doi:10.1038/s41587-020-00812-4 (2021).

7 Hu, B. C. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature* **574**, 187-192, doi:10.1038/s41586-019-1629-x (2019).

8 Ecker, J. R. *et al.* The BRAIN Initiative Cell Census Consortium: Lessons Learned toward Generating a Comprehensive Brain Cell Atlas. *Neuron* **96**, 542-557, doi:https://doi.org/10.1016/j.neuron.2017.10.007 (2017).

9 Ong, E. *et al.* Modelling kidney disease using ontology: insights from the Kidney Precision Medicine Project. *Nature Reviews Nephrology* **16**, 686-696, doi:10.1038/s41581-020-00335-w (2020).
10 Rozenblatt-Rosen, O. *et al.* The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. *Cell* **181**, 236-249, doi:https://doi.org/10.1016/j.cell.2020.03.053 (2020).
11 Gowda, S. *et al.* Markers of renal function tests. *N Am J Med Sci* **2**, 170-173 (2010).
12 Koeppen, B. M. & Stanton, B. A. *Renal Physiology E-Book: Mosby Physiology Monograph Series*. (Elsevier Health Sciences, 2012).
13 Eaton, D. C. & Pooler, J. P. *Vander's renal physiology*. (Mc Graw Hil Medicall, 2009).
14 Al-Awqati, Q. & Oliver, J. A. Stem cells in the kidney. *Kidney International* **61**, 387-395, doi:10.1046/j.1523-1755.2002.00164.x (2002).
15 Puelles, V. G. *et al.* Glomerular number and size variability and risk for kidney disease. *Current Opinion in Nephrology and Hypertension* **20**, 7-15 (2011).
16 Palmer, L. G. & Schnermann, J. Integrated control of Na transport along the nephron. *Clinical Journal of the American Society of Nephrology* **10**, 676-687 (2015).
17 Keesing, F., Holt, R. D. & Ostfeld, R. S. Effects of species diversity on disease risk. *Ecology letters* **9**, 485-498 (2006).
18 Hilliard, L. M., Sampson, A. K., Brown, R. D. & Denton, K. M. The "His and Hers" of the Renin-Angiotensin System. *Current Hypertension Reports* **15**, 71-79, doi:10.1007/s11906-012-0319-y (2013).
19 Maric, C. Sex, diabetes and the kidney. *American Journal of Physiology-Renal Physiology* **296**, F680-F688, doi:10.1152/ajprenal.90505.2008 (2009).
20 Cobo, G. *et al.* Sex and gender differences in chronic kidney disease: progression to end-stage renal disease and haemodialysis. *Clinical Science* **130**, 1147-1163, doi:10.1042/cs20160047 (2016).
21 Hutchens, M. P., Dunlap, J., Hurn, P. D. & Jarnberg, P. O. Renal Ischemia: Does Sex Matter? *Anesthesia & Analgesia* **107** (2008).
22 Neugarten, J., Kasiske, B., Silbiger, S. R. & Nyengaard, J. R. Effects of Sex on Renal Structure. *Nephron* **90**, 139-144, doi:10.1159/000049033 (2002).
23 Lake, B. B. *et al.* An atlas of healthy and injured cell states and niches in the human kidney. *bioRxiv*, 2021.2007.2028.454201, doi:10.1101/2021.07.28.454201 (2021).
24 Martín-Saiz, L. *et al.* High-Resolution Human Kidney Molecular Histology by Imaging Mass Spectrometry of Lipids. *Analytical Chemistry* **93**, 9364-9372, doi:10.1021/acs.analchem.1c00649 (2021).
25 Neumann, E. K. *et al.* Protocol for multimodal analysis of human kidney tissue by imaging mass spectrometry and CODEX multiplexed immunofluorescence. *STAR Protocols* **2**, 100747, doi:https://doi.org/10.1016/j.xpro.2021.100747 (2021).
26 Bairey Merz, C. N. *et al.* Sex and the kidneys: current understanding and research opportunities. *Nature Reviews Nephrology* **15**, 776-783, doi:10.1038/s41581-019-0208-6 (2019).
27 Spraggins, J. M. *et al.* High-Performance Molecular Imaging with MALDI Trapped Ion-Mobility Time-of-Flight (timsTOF) Mass Spectrometry. *Analytical Chemistry* **91**, 14552-14560, doi:10.1021/acs.analchem.9b03612 (2019).
28 Patterson, N. H. *et al.* Autofluorescence microscopy as a label-free tool for renal histology and glomerular segmentation. *bioRxiv* (2021).
29 Tideman, L. E. M. *et al.* Automated biomarker candidate discovery in imaging mass spectrometry data through spatially localized Shapley additive explanations. *Analytica Chimica Acta* **1177**, 338522, doi:https://doi.org/10.1016/j.aca.2021.338522 (2021).

30 Lundberg, S. M. *et al.* From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* **2**, 56-67, doi:10.1038/s42256-019-0138-9 (2020).
31 Neumann, E. K. *et al.* Highly Multiplexed Immunofluorescence of the Human Kidney using Co-Detection by Indexing (CODEX). *bioRxiv*, 2020.2012.2004.412429, doi:10.1101/2020.12.04.412429 (2020).
32 Radtke, A. J. *et al.* IBEX: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proceedings of the National Academy of Sciences* **117**, 33455-33465, doi:10.1073/pnas.2018488117 (2020).
33 Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e915, doi:10.1016/j.cell.2018.07.010 (2018).
34 Patterson, N. H., Tuck, M., Van de Plas, R. & Caprioli, R. M. Advanced Registration and Analysis of MALDI Imaging Mass Spectrometry Measurements through Autofluorescence Microscopy. *Analytical Chemistry* **90**, 12395-12403, doi:10.1021/acs.analchem.8b02884 (2018).
35 Patterson, N. H. *et al.* Next Generation Histology-Directed Imaging Mass Spectrometry Driven by Autofluorescence Microscopy. *Analytical Chemistry* **90**, 12404-12413, doi:10.1021/acs.analchem.8b02885 (2018).
36 Neumann, E. K. *et al.* Protocol for multimodal analysis of human kidney tissue by imaging mass spectrometry and CODEX multiplexed immunofluorescence. *STAR protocols* **2**, 100747 (2021).
37 Neumann, E. K. *et al.* Highly multiplexed immunofluorescence of the human kidney using co-detection by indexing. *Kidney International* **101**, 137-143, doi:https://doi.org/10.1016/j.kint.2021.08.033 (2022).
38 Pennington, E. R. *et al.* Distinct membrane properties are differentially influenced by cardiolipin content and acyl chain composition in biomimetic membranes. *Biochim Biophys Acta Biomembr* **1859**, 257-267, doi:10.1016/j.bbamem.2016.11.012 (2017).
39 Solati, Z., Edel, A. L., Shang, Y., O, K. & Ravandi, A. Oxidized phosphatidylcholines are produced in renal ischemia reperfusion injury. *PLOS ONE* **13**, e0195172, doi:10.1371/journal.pone.0195172 (2018).
40 Jarsch, I. K., Daste, F. & Gallop, J. L. Membrane curvature in cell biology: An integration of molecular mechanisms. *J Cell Biol* **214**, 375-387, doi:10.1083/jcb.201604003 (2016).
41 Kovesdy, C. P., Furth, S. L. & Zoccali, C. Obesity and Kidney Disease: Hidden Consequences of the Epidemic. *American Journal of Nephrology* **45**, 283-291, doi:10.1159/000458467 (2017).
42 Sooy, K., Kohut, J. & Christakos, S. The role of calbindin and 1,25dihydroxyvitamin D3 in the kidney. *Current Opinion in Nephrology and Hypertension* **9** (2000).
43 Nobre, J. L. *et al.* Role of vitamin D in adipose tissue in obese rats programmed by early weaning and post diet calcium. *Molecular Nutrition & Food Research* **60**, 810-822, doi:https://doi.org/10.1002/mnfr.201500735 (2016).
44 Song, C. *et al.* Myristoylated alanine-rich C kinase substrate-like protein-1 regulates epithelial sodium channel activity in renal distal convoluted tubule cells. *American Journal of Physiology-Cell Physiology* **319**, C589-C604, doi:10.1152/ajpcell.00218.2020 (2020).
45 Hall, J. E. The Kidney, Hypertension, and Obesity. *Hypertension* **41**, 625-633, doi:doi:10.1161/01.HYP.0000052314.95497.78 (2003).
46 Dawaliby, R. *et al.* Phosphatidylethanolamine Is a Key Regulator of Membrane Fluidity in Eukaryotic Cells. *J Biol Chem* **291**, 3658-3667, doi:10.1074/jbc.M115.706523 (2016).
47 Colombo, S. *et al.* Phospholipidome of endothelial cells shows a different adaptation response upon oxidative, glycative and lipoxidative stress. *Scientific Reports* **8**, 12365, doi:10.1038/s41598-018-30695-0 (2018).

48 Lee, H. J., Balasubramanian, S. V., Murer, H., Biber, J. & Morris, M. E. Modulation of sulfate renal transport by alterations in cell membrane fluidity. *Journal of Pharmaceutical Sciences* **88**, 976-980, doi:https://doi.org/10.1021/js990114c (1999).
49 Balasubramanian, S. V., Straubinger, R. M. & Morris, M. E. Salicylic Acid Induces Changes in the Physical Properties of Model and Native Kidney Membranes. *Journal of Pharmaceutical Sciences* **86**, 199-204, doi:https://doi.org/10.1021/js9602756 (1997).
50 Hsu, B. Y. L. *et al.* Membrane fluidity and sodium transport by renal membranes from dogs with spontaneous idiopathic Fanconi syndrome. *Metabolism* **41**, 253-259, doi:https://doi.org/10.1016/0026-0495(92)90267-E (1992).
51 Farge, E., Ojcius, D. M., Subtil, A. & Dautry-Varsat, A. Enhancement of endocytosis due to aminophospholipid transport across the plasma membrane of living cells. *The American journal of physiology* **276**, C725-733, doi:10.1152/ajpcell.1999.276.3.C725 (1999).
52 Patel, D. & Witt, S. N. Ethanolamine and Phosphatidylethanolamine: Partners in Health and Disease. *Oxidative Medicine and Cellular Longevity* **2017**, 4829180, doi:10.1155/2017/4829180 (2017).
53 McMahon, H. T. & Boucrot, E. Membrane curvature at a glance. *J Cell Sci* **128**, 1065-1070, doi:10.1242/jcs.114454 (2015).
54 Samad, F. & Ruf, W. Inflammation, obesity, and thrombosis. *Blood, The Journal of the American Society of Hematology* **122**, 3415-3422 (2013).
55 Miszta, A. *et al.* Application of a plasmin generation assay to define pharmacodynamic effects of tranexamic acid in women undergoing cesarean delivery. *Journal of Thrombosis and Haemostasis* **19**, 221-232, doi:https://doi.org/10.1111/jth.15114 (2021).
56 Bochkov, V. *et al.* Pleiotropic effects of oxidized phospholipids. *Free Radical Biology and Medicine* **111**, 6-24 (2017).
57 Salomon, R. G. Structural identification and cardiovascular activities of oxidized phospholipids. *Circulation research* **111**, 930-946 (2012).
58 Kopec, A. K. *et al.* Thrombin promotes diet-induced obesity through fibrin-driven inflammation. *The Journal of clinical investigation* **127**, 3152-3166 (2017).
59 Ząbczyk, M., Natorska, J. & Undas, A. Fibrin clot properties in atherosclerotic vascular disease: from pathophysiology to clinical outcomes. *Journal of Clinical Medicine* **10**, 2999 (2021).
60 Silbiger, S. & Neugarten, J. Gender and human chronic renal disease. *Gender Medicine* **5**, S3-S10, doi:https://doi.org/10.1016/j.genm.2008.03.002 (2008).
61 Valdivielso, J. M., Jacobs-Cachá, C. & Soler, M. J. Sex hormones and their influence on chronic kidney disease. *Current Opinion in Nephrology and Hypertension* **28** (2019).
62 Ahmed, S. B. & Ramesh, S. Sex hormones in women with kidney disease. *Nephrology Dialysis Transplantation* **31**, 1787-1795, doi:10.1093/ndt/gfw084 (2016).
63 Melo Junior, A. F. *et al.* Involvement of sex hormones, oxidative stress, ACE and ACE2 activity in the impairment of renal function and remodelling in SHR. *Life Sciences* **257**, 118138, doi:https://doi.org/10.1016/j.lfs.2020.118138 (2020).
64 Laughlin, G. A., Barrett-Connor, E. & May, S. Sex-specific determinants of serum adiponectin in older adults: the role of endogenous sex hormones. *International Journal of Obesity* **31**, 457-465, doi:10.1038/sj.ijo.0803427 (2007).
65 Lima-Posada, I. & Bobadilla, N. A. Understanding the opposite effects of sex hormones in mediating renal injury. *Nephrology* **26**, 217-226, doi:https://doi.org/10.1111/nep.13806 (2021).
66 Dubey, R. K., Oparil, S., Imthurn, B. & Jackson, E. K. Sex hormones and hypertension. *Cardiovascular Research* **53**, 688-708, doi:10.1016/s0008-6363(01)00527-2 (2002).
67 Link, J. C. & Reue, K. Genetic Basis for Sex Differences in Obesity and Lipid Metabolism. *Annual Review of Nutrition* **37**, 225-245, doi:10.1146/annurev-nutr-071816-064827 (2017).

68 Mittendorfer, B. Sexual Dimorphism in Human Lipid Metabolism. *The Journal of Nutrition* **135**, 681-686, doi:10.1093/jn/135.4.681 (2005).

69 Gravholt, C. H. *et al.* Glucose Metabolism, Lipid Metabolism, and Cardiovascular Risk Factors in Adult Turner's Syndrome: The impact of sex hormone replacement. *Diabetes Care* **21**, 1062-1070, doi:10.2337/diacare.21.7.1062 (1998).

70 Wang, X., Magkos, F. & Mittendorfer, B. Sex Differences in Lipid and Lipoprotein Metabolism: It's Not Just about Sex Hormones. *The Journal of Clinical Endocrinology & Metabolism* **96**, 885-893, doi:10.1210/jc.2010-2061 (2011).

71 Ng, K. L. *et al.* Utility of cytokeratin 7, S100A1 and caveolin-1 as immunohistochemical biomarkers to differentiate chromophobe renal cell carcinoma from renal oncocytoma. *Transl Androl Urol* **8**, S123-S137, doi:10.21037/tau.2018.11.02 (2019).

72 Ganeshan, B. & Miles, K. A. Quantifying tumour heterogeneity with CT. *Cancer Imaging* **13**, 140-149, doi:10.1102/1470-7330.2013.0015 (2013).

73 Eberhard, A. *et al.* Heterogeneity of Angiogenesis and Blood Vessel Maturation in Human Tumors: Implications for Antiangiogenic Tumor Therapies1. *Cancer Research* **60**, 1388-1393 (2000).

74 Tuomi, T. *et al.* The many faces of diabetes: a disease with increasing heterogeneity. *The Lancet* **383**, 1084-1094, doi:https://doi.org/10.1016/S0140-6736(13)62219-9 (2014).

75 Battaglia, M. *et al.* Introducing the Endotype Concept to Address the Challenge of Disease Heterogeneity in Type 1 Diabetes. *Diabetes Care* **43**, 5-12, doi:10.2337/dc19-0880 (2019).

76 Ma, F., Sun, T., Liu, L. & Jing, H. Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. *Future Generation Computer Systems* **111**, 17-26, doi:https://doi.org/10.1016/j.future.2020.04.036 (2020).

77 Abeysekera, R. A., Healy, H. G., Wang, Z., Cameron, A. L. & Hoy, W. E. Heterogeneity in patterns of progression of chronic kidney disease. *Internal Medicine Journal* **51**, 220-228, doi:https://doi.org/10.1111/imj.14770 (2021).

78 De Nicola, L. & Zoccali, C. Chronic kidney disease prevalence in the general population: heterogeneity and concerns. *Nephrology Dialysis Transplantation* **31**, 331-335, doi:10.1093/ndt/gfv427 (2015).

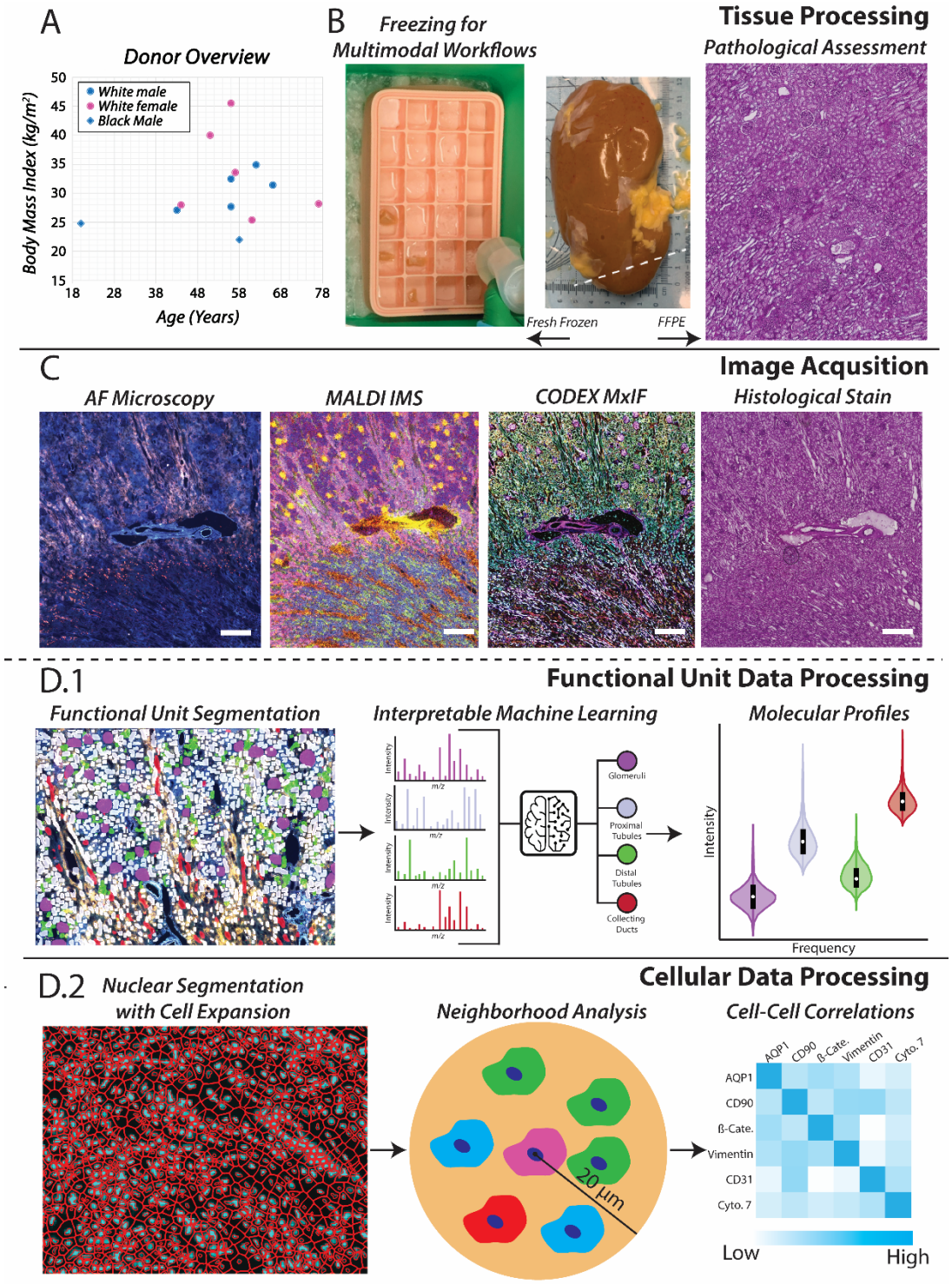

Figure 1. Overview of the experimental and data analysis workflow for atlas creation. A) Graph describing the patient population surveyed in this study. B) Graphical depiction of kidney dissection and processing for both fresh frozen and formalin fixed and paraffin embedding (FFPE) sample preparation. C) Select images from each of the modalities featured within the atlas, including autofluorescence microscopy, MALDI imaging mass spectrometry, co-detection by indexing multiplexed

immunofluorescence microscopy, and histological staining. D) Data analysis pipeline for D.1 functional tissue units and D.2 single cells. Additional information can be found in the methods section of the SI.

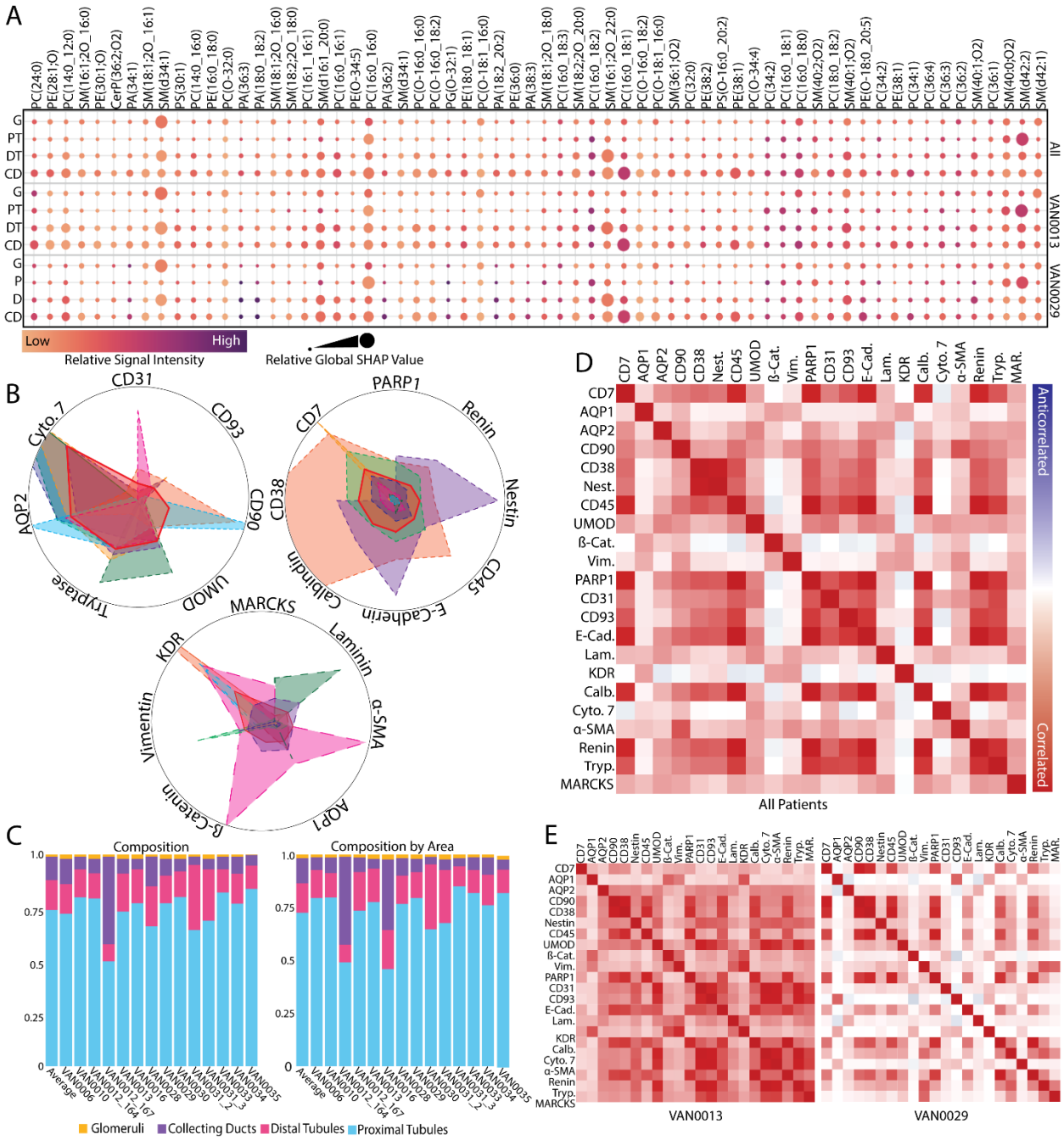

Figure 2. Overview of data represented within the kidney atlas. **A)** Intensity and SHAP scores for a subset of 64 lipids as a composite of all patients and then two extracted patients for glomeruli (G), proximal tubules (PT), distal tubules (DT), and collecting ducts (CD). **B)** Radar plots showing the relative number of cells expressing each protein antigen averaged (solid red line) and on an individual patient level (dashed line), ordered by relative cell number. **C)** Relative composition of different tissue units within each patient holistically (left) and normalized by area (right). **D)** Protein neighborhoods derived for all patients and **E)** a subset of patients that demonstrates how the cellular architecture varies between different tissue.

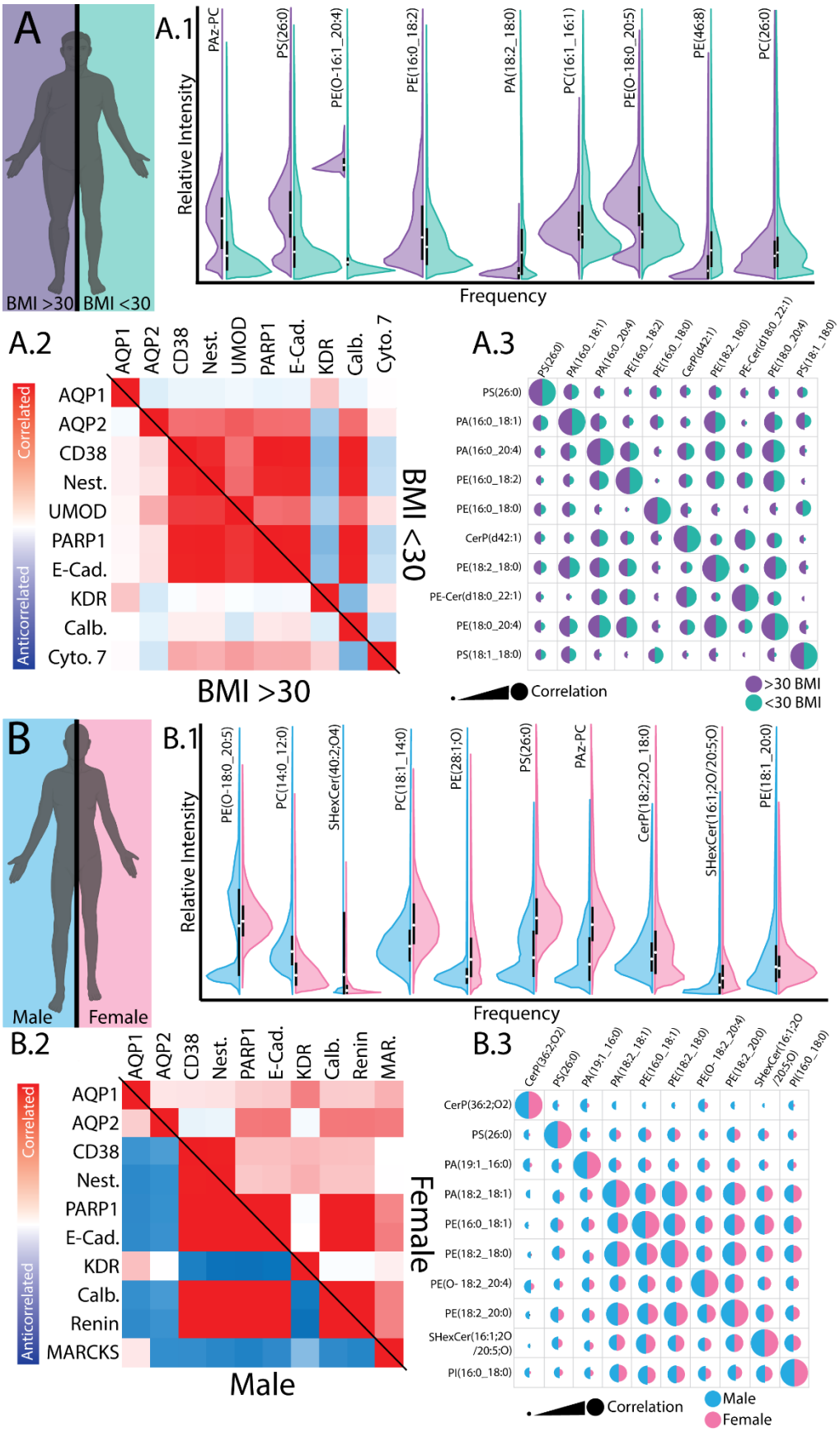
A
BMI >30
BMI <30
A.1
Relative Intensity
PAz-PC
PS(26:0)
PE(O-16:1_20:4)
PE(16:0_18:2)
PA(18:2_18:0)
PC(16:1_16:1)
PE(O-18:0_20:5)
PE(46:8)
PC(26:0)
Frequency
A.2
AQP1
AQP2
CD38
Nest.
UMOD
PARP1
E-Cad.
KDR
Calb.
Cyto. 7
Correlated
Anticorrelated
BMI <30
BMI >30
A.3
PS(26:0)
PA(16:0_18:1)
PA(16:0_20:4)
PE(16:0_18:2)
PE(16:0_18:0)
CerP(d42:1)
PE(18:2_18:0)
PE-Cer(d18:0_22:1)
PE(18:0_20:4)
PS(18:1_18:0)
Correlation
>30 BMI
<30 BMI
B
Male
Female
B.1
Relative Intensity
PE(O-18:0_20:5)
PC(14:0_12:0)
SHexCer(40:2;O4)
PC(18:1_14:0)
PE(28:1;O)
PS(26:0)
PAz-PC
CerP(18:2;2O_18:0)
SHexCer(16:1;2O/20:5;O)
PE(18:1_20:0)
Frequency
B.2
AQP1
AQP2
CD38
Nest.
PARP1
E-Cad.
KDR
Calb.
Renin
MAR.
Correlated
Anticorrelated
Female
Male
B.3
CerP(36:2;O2)
PS(26:0)
PA(19:1_16:0)
PA(18:2_18:1)
PE(16:0_18:1)
PE(18:2_18:0)
PE(O- 18:2_20:4)
PE(18:2_20:0)
SHexCer(16:1;2O /20:5;O)
PI(16:0_18:0)
Correlation
Male
Female

Figure 3: Molecular and cellular profiles of the patients subdivided by **A)** BMI and **B)** sex. **A.1)** Violin plots for nine lipids with the highest SHAP scores that differentiate tissue originating from male (blue) and female (pink) patients. **A.2)** Cellular neighborhoods for ten markers that vary between men and women, where red is highly correlated and blue is highly anticorrelated based on Spearman correlation values and a neighborhood size of 25 µm from the cell center. **A.3)** Molecular neighborhoods for ten lipids detected in negative ion mode where circle radius is proportional to the Spearman correlation coefficient. Pixel neighborhoods were 1 pixel (10 µm) from the pixel of interest. **B.1)** Violin plots for nine lipids with the highest SHAP scores that differentiate tissue originating from patients with BMIs above 30 (purple) and below 30 (green). **B.2)** Cellular neighborhoods for ten markers that vary between obese and non-obese, where red is highly correlated and blue is highly anticorrelated based on Spearman correlation values and a neighborhood size of 25 µm from the cell center. **B.3)** Molecular neighborhoods for ten lipids detected in negative ion mode where circle radius is proportional to the Spearman correlation coefficient. Pixel neighborhoods were 1 pixel (10 µm) from the pixel of interest.

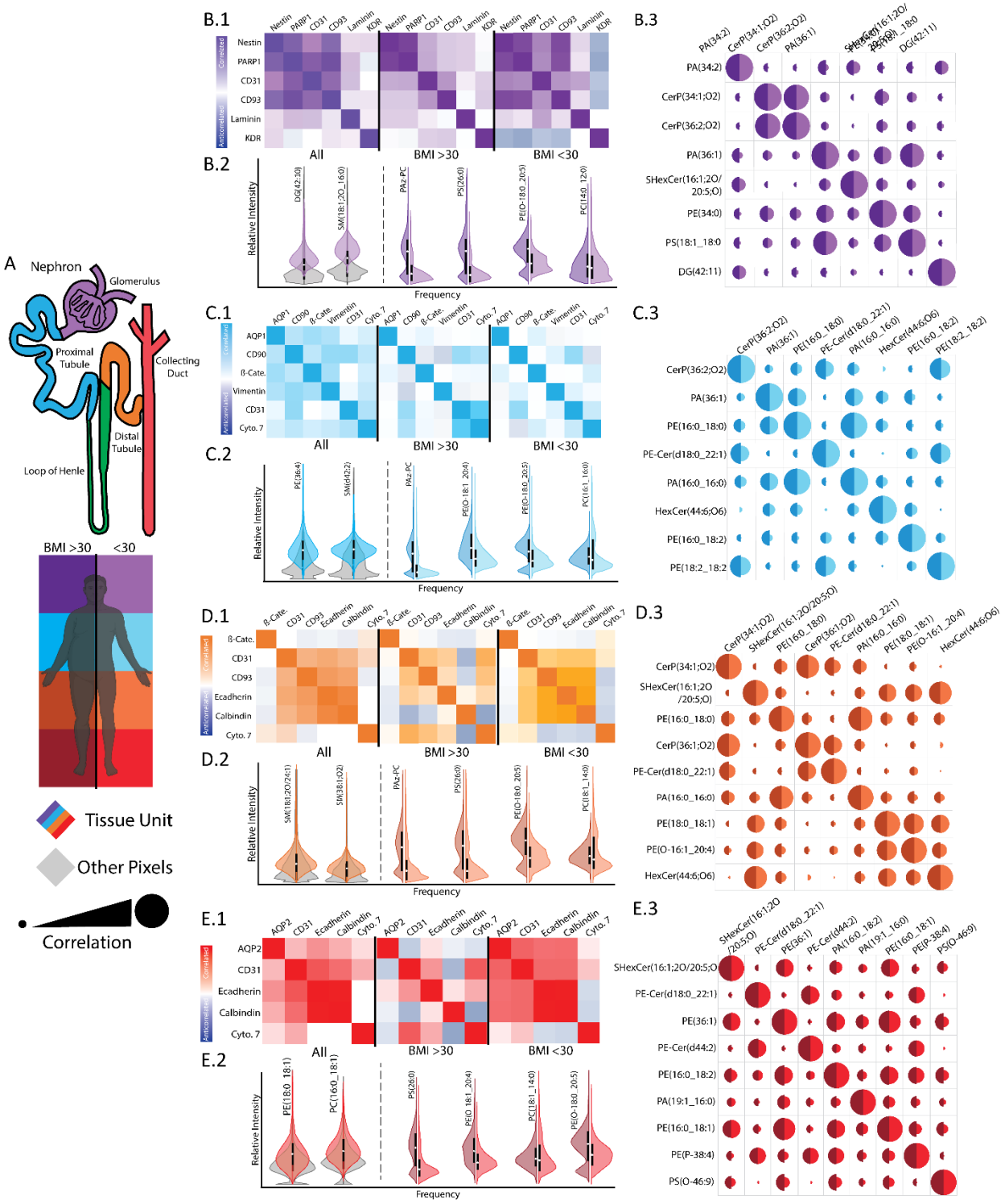

Figure 4: Summary of molecular and cellular differences between FTUs from patients with BMIs above 30 (obese) and below 30 (non-obese). **A)** Color-coded legend for interpretation of the remaining figures associated with **B)** glomeruli, **C)** proximal tubules, **D)** distal tubules, and **E)** collecting ducts. Differences between obese and non-obese patients can be seen in the extrapolated **1)** cellular neighborhoods, **2)** molecular species, and molecular neighborhoods associated with different segments of the nephron.

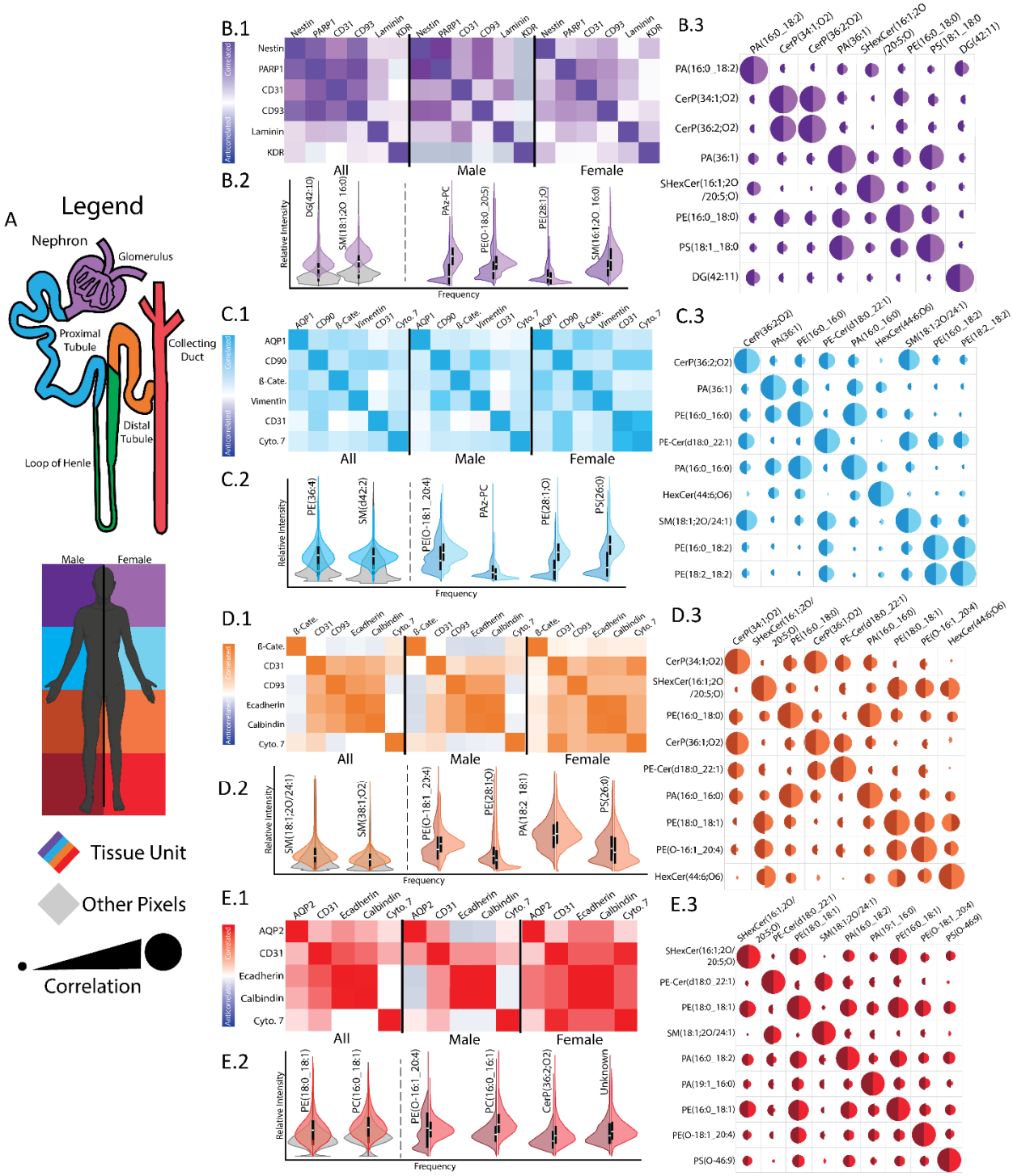

Figure 5: Summary of molecular and cellular differences between FTUs from male and female patients. **A)** Color-coded legend for interpretation of the remaining figures associated with **B)** glomeruli, **C)** proximal tubules, **D)** distal tubules, and **E)** collecting ducts. Differences between the sexes can be seen in the extrapolated **1)** cellular neighborhoods, **2)** molecular species, and **3)** molecular neighborhoods associated with different segments of the nephron.