

# Single-cell profiling of healthy human kidney reveals features of sex-based transcriptional programs and tissue-specific immunity

---

Received: 6 December 2021

---

Accepted: 27 November 2022

---

Published online: 10 December 2022

---

 Check for updates

---

Caitriona M. McEvoy<sup>1,2,3,16</sup>, Julia M. Murphy<sup>1,2,4,16</sup>, Lin Zhang<sup>5,6</sup>, Sergi Clotet-Freixas<sup>1,2</sup>, Jessica A. Mathews<sup>1,2</sup>, James An<sup>1,2,4</sup>, Mehran Karimzadeh<sup>6,7</sup>, Delaram Pouyabahr<sup>8,9</sup>, Shenghui Su<sup>1,2</sup>, Olga Zaslaver<sup>8,9</sup>, Hannes Röst<sup>8,9</sup>, Rangi Arambewela<sup>1,2</sup>, Lewis Y. Liu<sup>1,2,4</sup>, Sally Zhang<sup>10</sup>, Keith A. Lawson<sup>10</sup>, Antonio Finelli<sup>10,11</sup>, Bo Wang<sup>6,7,12,13</sup>, Sonya A. MacParland<sup>1,2,4,13</sup>, Gary D. Bader<sup>8,9,12,14,15</sup>, Ana Konvalinka<sup>1,2,3,11,13,17</sup> ✉ & Sarah Q. Crome<sup>1,2,4,17</sup> ✉

---

Knowledge of the transcriptional programs underpinning the functions of human kidney cell populations at homeostasis is limited. We present a single-cell perspective of healthy human kidney from 19 living donors, with equal contribution from males and females, profiling the transcriptome of 27677 cells to map human kidney at high resolution. Sex-based differences in gene expression within proximal tubular cells were observed, specifically, increased anti-oxidant metallothionein genes in females and aerobic metabolism-related genes in males. Functional differences in metabolism were confirmed in proximal tubular cells, with male cells exhibiting higher oxidative phosphorylation and higher levels of energy precursor metabolites. We identified kidney-specific lymphocyte populations with unique transcriptional profiles indicative of kidney-adapted functions. Significant heterogeneity in myeloid cells was observed, with a *MRC1<sup>+</sup>LYVE1<sup>+</sup>FOLR2<sup>+</sup>CIQC<sup>+</sup>* population representing a predominant population in healthy kidney. This study provides a detailed cellular map of healthy human kidney, and explores the complexity of parenchymal and kidney-resident immune cells.

Kidneys perform complex functions and maintain body homeostasis through a diverse range of specialized parenchymal cells residing in distinct compartments. Within tissues, resident immune populations have specialized functions in surveillance, maintenance of self-tolerance, response to infection and injury, and interface with parenchymal cells to maintain tissue homeostasis<sup>1–3</sup>. There is limited understanding of the network of kidney parenchymal and resident immune cells in human kidney due to lack of access to healthy, fresh tissue, making it challenging to delineate alterations in transcriptional programs and harmful immune responses that occur in native kidney diseases and following kidney transplantation. Much of our knowledge

is based on studies that used kidneys rejected for transplant or tumour-adjacent nephrectomy specimens, where parenchymal populations can have altered molecular programs, and immune populations and their signalling circuits may not be entirely reflective of the steady-state<sup>4,5</sup>. Further, sex-based dichotomy in gene expression within human kidney cell populations has not been thoroughly examined, but is of great significance to acute and chronic kidney disease, ischemia-reperfusion injury (IRI) and progression of diabetic kidney disease, which exhibit a male preponderance<sup>6–8</sup>.

In this study, we define cell populations and transcriptomic signatures of healthy human kidney using the gold standard of tissue

sources – living donor kidney biopsies. We provide a map of healthy human kidney using single-cell RNA sequencing (scRNAseq) of 19 living donors, capturing parenchymal and immune cell transcriptomes reflective of a healthy state. Our map includes common and rare cell types and demonstrates that <1% of all cells in healthy kidney are immune cells. Given known dichotomies between males and females in many kidney diseases and outcomes, we examined underlying baseline transcriptome differences between healthy male and female kidney. We identify sex-specific gene expression profiles among kidney cell populations, and functionally validate differences in proximal tubular (PT) cell metabolism. Finally, as it's known that the immune system responds rapidly to tissue alterations, we examined differences in immune cell subsets and functions in living donor kidney and compared to those observed previously in studies of tumour-affected nephrectomy and deceased donor kidney. We report unique immune populations, including an alternatively-activated myeloid population and kidney-resident lymphocytes, validating markers which differentiate them from their circulating counterparts. Collectively, we demonstrate a unique immune niche in healthy living donor kidney and define molecular programs that distinguish male and female parenchymal PT cell populations at homeostasis.

## Results

### Single-cell map of healthy human kidney

We examined the cellular landscape of human kidney using pre-implantation kidney biopsies from 19 sex-matched living kidney donors (Fig. 1a, b). Our dissociation method was developed to maximize viability while preserving the representation of rare and fragile cell populations, and we employed rigorous quality control. We noted minimal immune cell representation in healthy kidneys (~0.3% of cells captured), and therefore incorporated CD45-enrichment for immune cells in 10/19 biopsy samples (5 female, 5 male) (Fig. 1a). Of 27677 cells in our map, 6899 cells were from CD45-enriched samples, while 20778 cells were from non-CD45-enriched samples. Twenty-three clusters were identified, including several distinct immune cell populations, alongside all anticipated parenchymal populations of the nephron (Fig. 1c, Supplementary Fig. 1a). Clusters were comprised of cells captured from multiple donors, there was no exceptional variability in cell cycle state across clusters, and most clusters had symmetrical distribution of donor sex (Fig. 1b–e, Supplementary Fig. 1).

As anticipated, Proximal Tubular (PT) cells comprised 75% of sequenced cells. Sub-clustering revealed 6 distinct clusters (PT1–PT6) (Fig. 1d, Supplementary Fig. 2a), with some heterogeneity between individuals, methods of sample preparation, and sexes noted (Supplementary Fig. 2b). PT segment-specific separation is evident; PT1, 4, and 6 are enriched for PT segment 1 (S1) marker *SLC5A2* and S1/2-abundant genes (*SLC7A7*, *ANK2*, *SLC44A*, *SLC6A19*, *SLC22A8*), while PT2 shows increased expression of S3-abundant genes (*DCXR*, *AGXT*, *SLC22A7*, *SLC7A13*) (Fig. 1e)<sup>9,10</sup>. PT3 highly expresses dissociation stress-associated genes<sup>3</sup>, together with general (*LRP2*, *CUBN*) and segment-specific PT genes, indicating cell contributions from all PT segments (Fig. 1e, Supplementary Fig. 2c). PT5 (*VIM*<sup>+</sup>*SI00A6*<sup>+</sup>*VCAMI*<sup>+</sup>*DCDC2*<sup>+</sup>*ANXA4*<sup>+</sup>) displays similarity to a putative regenerative PT population – termed ‘scattered tubular cells’ (STC)<sup>11,12</sup>. These genes also characterize a population which expands following IRI and is postulated to reflect failed PT repair, though expression was also observed in healthy kidney<sup>13,14</sup>. Some STC-associated genes were exclusively expressed by PT5 or PT3, while others were expressed by both populations (Supplementary Fig. 2d–f). This transcriptional overlap between the regenerative STC-like PT5 and stressed PT3 cells may indicate attempted initiation of repair in PT3 cells. Transcription factor analysis (Supplementary Fig. 2g) of PT5 genes revealed potential upstream regulators directing cell differentiation and migration (*SNAI2*, *ZNF217*), and epithelial phenotype maintenance (*ELF3*), alongside *NFE2L2*, a key regulator of antioxidant and cytoprotective

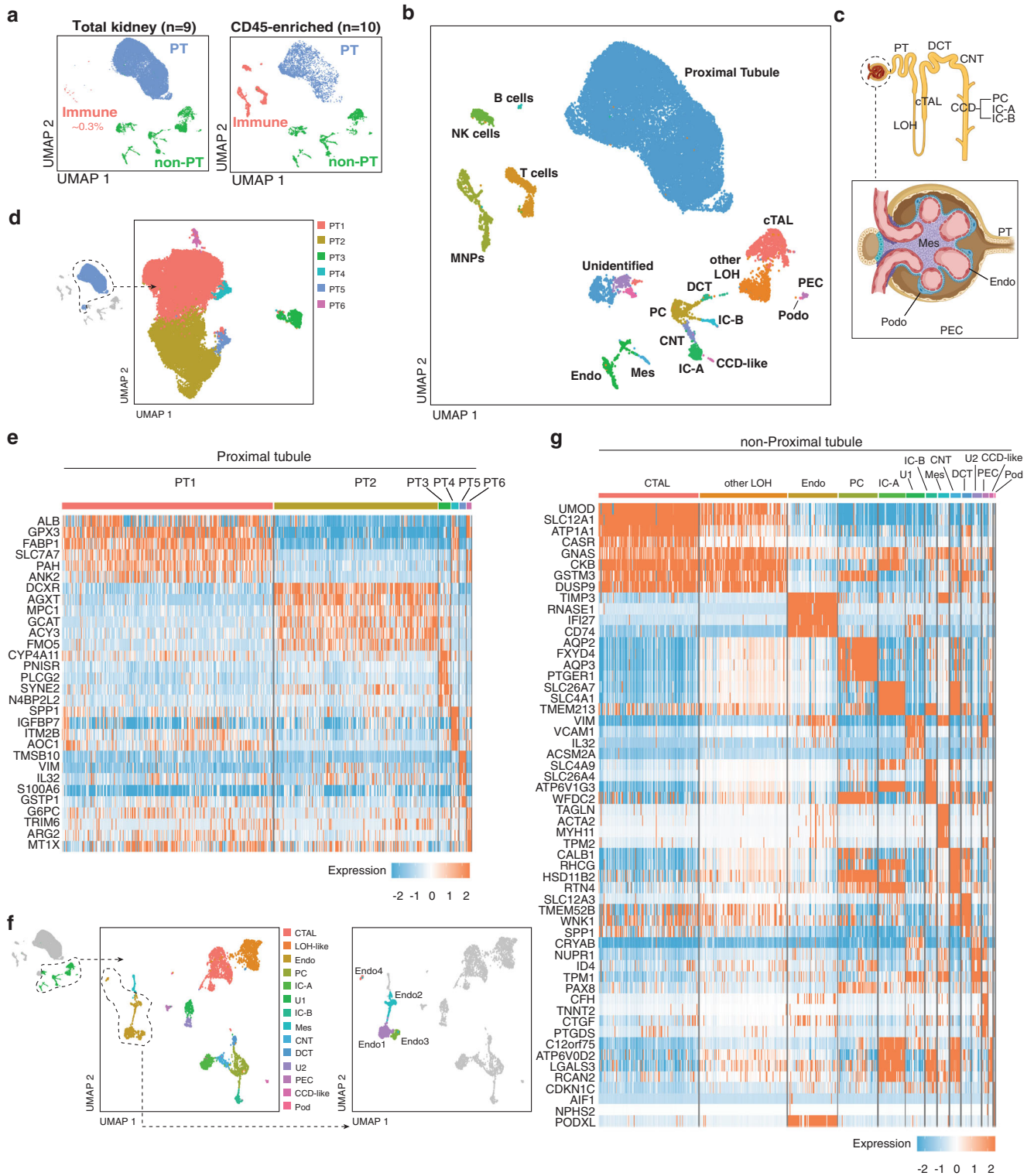
genes<sup>15</sup>. Predicted upstream regulators for PT3 (*EGRI*, *FOS*, and *JUN*) are associated with oxidative stress and fibrogenesis. Predicted regulator *ATF3* (protective in renal IRI<sup>16</sup>) supports potential reparative processes in this cluster (Supplementary Fig. 2g).

*VCAMI*<sup>+</sup> PT cells were recently described in a dataset derived from non-tumour portions of nephrectomy samples by single nucleus RNA sequencing (snRNAseq)<sup>14</sup>. We explored the similarities between *VCAMI*<sup>+</sup> PT cells in our dataset and those described by Muto et al. in an unbiased manner. We annotated PT cells from our dataset based on their overall transcriptional correlation to PT clusters in Muto et al.<sup>14,17</sup>. PT5 cells from our dataset were the most highly correlated with the ‘PT\_VCAMI’ cluster from Muto et al. (Supplementary Fig. 3a), and cells from our dataset annotated by label transfer as ‘PT\_VCAMI’ clearly derive from the PT5 cluster (Supplementary Fig. 3b). Next, we integrated both datasets into one data object (Supplementary Fig. 3c). Although clustering fails to completely overcome technical differences associated with different sequencing modalities (Supplementary Fig. 3d), the conformation of cells in dimensionality reduction again confirms that PT\_VCAMI cells co-localize with our PT5 cells. Additionally, we note greater diversity in the PT\_VCAMI population, as some cells from this annotation also co-localized with PT3 cells and U1/2, supporting the likely overlap in functions among injured/stressed PT cells and STCs (Supplementary Fig. 3e). Finally, correlation analysis between PT5 cluster defining genes and those of PT\_VCAMI identified a significant positive correlation (Spearman's rho = 0.524, *p* = 8.542e-11) (Supplementary Fig. 3f).

Fourteen non-PT parenchymal cell populations were identified<sup>9</sup> (Fig. 1f, g) including rare but important glomerular populations such as podocytes, mesangial cells, and parietal epithelial cells. We detected notable heterogeneity in CTAL and endothelial populations. Two CTAL subpopulations expressing *CLDN10* and *CLDN16*, respectively, identify cells with differing paracellular cation-resorption preferences in *CLDN10*-dominant (Na<sup>+</sup>) versus *CLDN16*-dominant (Ca<sup>2+</sup>, Mg<sup>2+</sup>) tight junctions (Supplementary Fig. 4a–e)<sup>18</sup>. Among endothelial subpopulations (Endo1–4) (Fig. 1f), we identified two populations (Endo1, Endo3) of peritubular capillary cells (*PLVAP*<sup>+</sup>*TMEM88*<sup>+</sup>*DNASE1L3*<sup>+</sup>) (Supplementary Fig. 4a, f, g). Endo1 expressed *ESMI* – required for VEGF-related maintenance of the peritubular capillary network<sup>19</sup>, while Endo3 expressed motility and angiogenesis markers *MARCKS*, *CLU*, *ACKR1*, and *SEMA3D* (Supplementary Fig. 4f). Endo2 (*SOX17*<sup>+</sup>*SERPINE2*<sup>+</sup>*CLDN5*<sup>+</sup>*CXCL12*<sup>+</sup>) represents afferent arterioles and vasa recta, exhibiting reduced *KDR* expression and increased expression of extracellular matrix-encoding genes (Supplementary Fig. 4g). Endo4 expresses the glomerular microvascular endothelial cell markers *EDH3*, *SOST*, and *TBX3*, a transcriptional regulator critical to fenestrated glomerular endothelial development (Supplementary Fig. 4f)<sup>20</sup>. Our description of diversity in intra-renal endothelial populations may have implications for those developing endothelial-targeting therapies for transplant<sup>21</sup>.

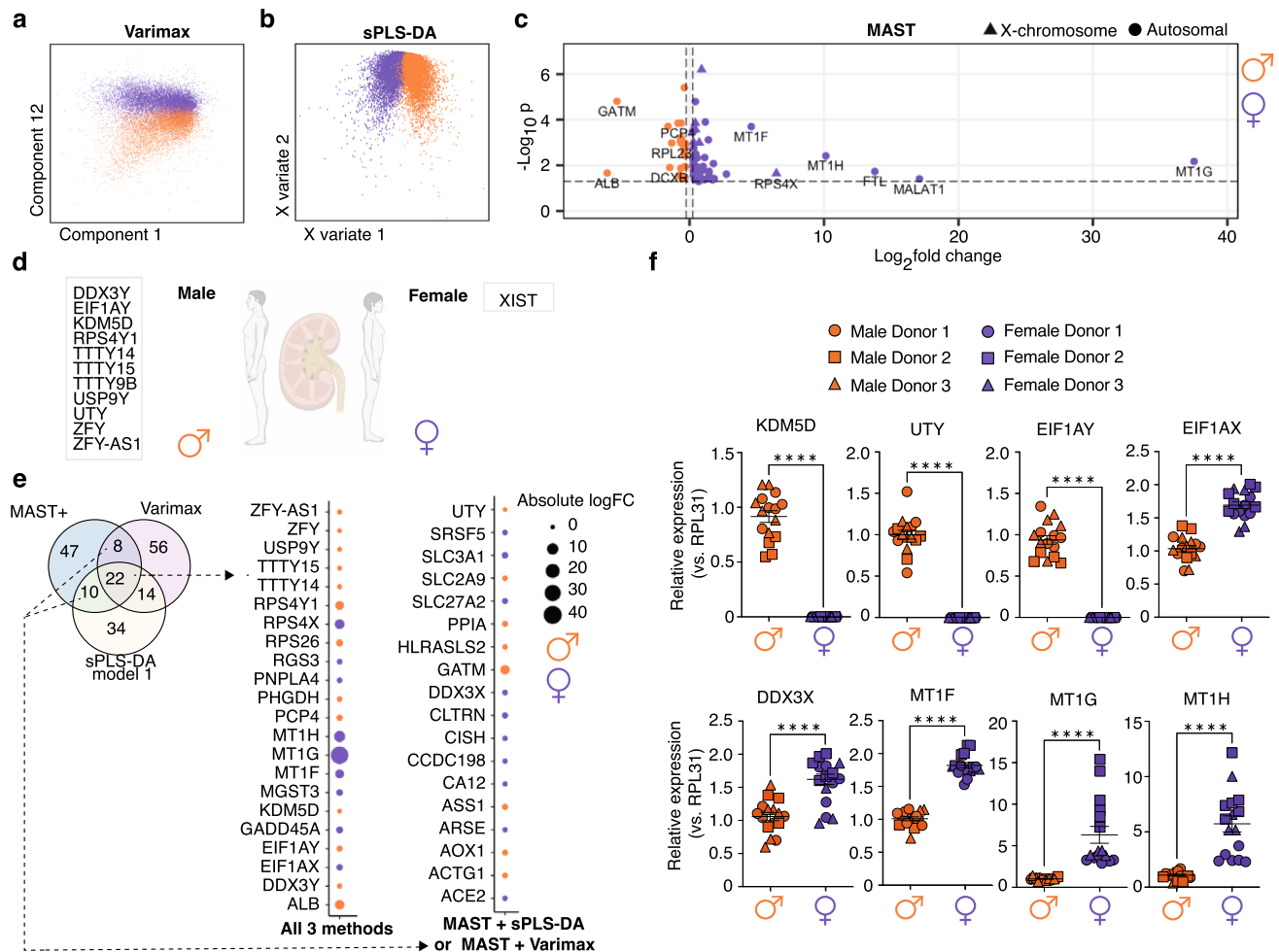
### Identification of sex-based transcriptomic differences in proximal tubular cells

Leveraging the sex-balanced large sample size, we examined differences in gene expression in healthy human kidney between males and females. Using varimax-rotated principal component analysis, we examined individual kidney populations for separation due to donor sex, and observed a clear separation for the PT population (Fig. 2a, Supplementary Fig. 5a). Such separation was not evident in other cell populations, perhaps reflecting insufficient power with fewer cells. Consequently, subsequent analyses focused on PT cells. Using machine learning, we identified the most discriminant subset of genes in our dataset that could correctly classify cell sex. Model-1 (80 genes) correctly classified cell sex with an area under the curve (AUC) of 0.98 (training dataset), and an accuracy of 84% (validation dataset) (Fig. 2b, Supplementary Fig. 5b–d). As X- and Y-linked genes potentially drive sex-biased effects<sup>22</sup>, we removed all sex chromosome-encoded genes



**Fig. 1 | Identification and annotation of kidney parenchymal cells. a** Different cell type proportions were captured by sequencing total kidney homogenate and CD45-enriched samples to create the total combined dataset. **b** UMAP clustering of total combined dataset with cell type annotations. **c** Graphical depiction of location of nephron cell types captured within the data. **d** UMAP plot of compartment-specific analysis of 20772 proximal tubular cells, comprising 6 clusters. **e** Heat map showing the expression levels of cluster marker genes. **f** UMAP plot of compartment-specific analysis of 4436 non-proximal tubular parenchymal cells,

with 14 cell populations represented, including four distinct endothelial clusters. **g** Heat map showing the expression levels of cell type marker genes across the 14 non-PT cell populations. PT Proximal tubule, DCT Distal convoluted tubule, CNT Connecting tubule, LOH Loop of Henle, cTAL Cortical thick ascending limb, CCD Cortical collecting duct, PC Principal cell, IC Intercalated cell, Mes Mesangial, Podo Podocyte, Endo Endothelial, PEC Parietal epithelial cell, NK Natural Killer, MNP Mononuclear phagocyte.



**Fig. 2 | Identifying genes differentially expressed between male and female proximal tubular cells.** 2-Dimensional plots of (a) Varimax-rotated PCA and (b) sPLS-DA showing separation of male and female cells, and (c) volcano plot showing differential expression of genes between sexes from MAST analysis with sample random effect. **d** Genes expressed exclusively by all samples of one sex and none of the opposite sex, which were added to the MAST results for comparison across methods in **e** (termed MAST +). **e** Venn diagram depicting genes identified through each analysis, with bubble plots highlighting genes identified by all three methods or by MAST plus one additional method. The size of the circle is proportional to absolute logFC and the colour indicates whether the gene was higher in male (orange) or female (dark purple) PT cells. Source data are provided in Supplementary Data 1. **f** Differences in gene expression of *KDM5D* ( $p < 0.0001$ ,  $t = 17.32$ ,

$df=30$ ), *UTY* ( $p < 0.0001$ ,  $t = 18.75$ ,  $df=30$ ), *EIF1AY* ( $p < 0.0001$ ,  $t = 18.04$ ,  $df=30$ ), *EIF1AX* ( $p < 0.0001$ ,  $t = 9.077$ ,  $df=29$ ), *DDX3X* ( $p < 0.0001$ ,  $t = 5.619$ ,  $df=29$ ), *MT1F* ( $p < 0.0001$ ,  $t = 16.04$ ,  $df=30$ ), *MT1G* ( $p < 0.0001$ ,  $u = 0$ ), and *MT1H* ( $p < 0.0001$ ,  $t = 6.286$ ,  $df=30$ ) were determined in primary male and female PT cells, and normalized to *RPL31* ( $n = 3$  donors/sex;  $n = 16$  replicates/sex). Source data are provided as a source data file. Group-to-group differences were assessed using two-tailed unpaired t-tests for variables following a normal distribution (*KDM5D*, *UTY*, *EIF1AY*, *EIF1AX*, *DDX3X*, *MT1F*, *MT1H*), and Mann-Whitney tests for variables with a non-parametric distribution (*MT1G*). Data are presented as mean values  $\pm$  SEM. \*\*\*\* $p < 0.0001$ . Orange = males, dark purple = females. Circle, square, and triangle symbols indicate biologically independent donors for each sex.

and derived Model-2 (15 genes), which correctly classified cell sex in the training dataset (AUC 0.85), but had reduced accuracy (68%) in the validation set (Supplementary Fig. 5d–f). Using an independent single-cell kidney dataset for validation<sup>23</sup>, our gene signatures accurately classified cell sex in 79% (Model-1) and 66% (Model-2) of cells (Supplementary Fig. 5d). Next, we identified genes with significant differential expression between males and females ( $n = 75$  genes,  $p$  value  $< 0.05$ ,  $\log_{2}FC > 0.25$ ) (Fig. 2c). As our conservative analysis excluded genes expressed uniquely by one sex (e.g. Y-chromosome-encoded genes), these genes ( $n = 12$ ) were added for downstream analyses (Fig. 2d). Results from our three analyses were compared (Supplementary Data 1). In agreement with previous studies<sup>24,25</sup>, the majority of the sex-biased genes uncovered are located in autosomes, rather than in sex chromosomes. Several sex-biased genes are consistent with previous reports of genes upregulated in murine male (*NAT8*, *FKBP5*, *KDM5D*, *DGKG*) and female (*MGST3*, *SLC3A1*, *CYP4A11*, *RPS29*) PT cells, respectively<sup>24–26</sup>.

Twenty-two genes featured in all three analyses (Fig. 2e), including 9 Y- and 3 X-chromosome encoded genes. An additional 18 genes featured in differential expression analysis (MAST+) and one other analysis (Fig. 2e). The X-chromosome genes reported are known to escape X-chromosome inactivation, explaining their higher expression in females<sup>22</sup>. Many of the autosomal-encoded genes or their family members are associated with primary sex determination (*SRSF5*<sup>27</sup>, *GATM*<sup>28</sup>, *GADD45A*), sex-biased expression (*CISH*, *SRSF5*, *ACTG1*, *GATM*, *AOX1*), or sex-specific effects (*SLC2A9*<sup>29</sup>). Intriguingly, many of the genes have established links with kidney disease, including *SLC27A2* (diabetic kidney disease)<sup>30</sup>, *SLC3A1* (cystinuria), and *GATM*<sup>31</sup>; while others are associated with hypoxia (*PHGD*, *CA12*), inflammation (*PPIA*), and genotoxic stress (*ASS1*). Metallothionein gene family members (*MT1F*, *MT1G*, *MT1H*), which encode cysteine-rich antioxidant proteins<sup>32</sup>, were notably higher in females (Fig. 2c, e). Additional differentially expressed genes also relate to cysteine-glutathione availability and metabolism, including *SLC3A1*<sup>33</sup>, *MGST3*, and *HRASLS2*.

Gene expression signatures derived by scRNAseq from tissues may be affected by the stress induced by tissue dissociation. Notably, male tissues are more susceptible to ischemic damage<sup>7,34</sup>. To ensure dissociation stress did not differentially impact male versus female PT cells, we calculated a gene module score using genes reported to be overexpressed in warm collagenase dissociation<sup>35</sup>, and found no difference in gene score between male and female PT cells, suggesting that the gene expression differences identified exist independently of susceptibility to dissociation stress (Supplementary Fig. 6a). To further verify that genes identified as having sex-biased expression did not relate to stress-induced artifact, we performed an unbiased analysis by integrating data from a single biopsy processed for both scRNAseq (50% of tissue) and snRNAseq (the other 50% of tissue) (Supplementary Fig. 6b, c), and identified many differences in PT cell gene expression between the two sequencing techniques (Supplementary Data 2, Supplementary Fig. 6d). From our gene signature differentiating male from female PT cells, we found 20/40 sex-biased genes overexpressed in scRNAseq versus snRNAseq (using at least 2 methods) and 8/40 overexpressed in snRNAseq versus scRNAseq (Supplementary Fig. 6e), likely due to cytosolic versus nuclear localization of the transcripts of interest leading to differential representation in transcript abundance between techniques<sup>36</sup>. This is particularly notable for metallothionein genes which were captured exclusively in scRNAseq data and relevant to female oxidative stress responses as noted above (Supplementary Fig. 6d). We further assessed sex-biased expression within the Muto et al. dataset, which is comprised of 3 male and 2 female samples<sup>14</sup>, but is generated using snRNAseq which relies on direct lysis of intact tissues. We found a significant positive correlation in DEGs differentiating male and female PT cells between these snRNAseq and scRNAseq datasets ( $\rho = 0.7318$ ,  $p = 0.004129$ ) (Supplementary Fig. 6f).

We next aimed to validate sex-biased gene expression profiles using commercially-available human primary PT cells from 3 male and 3 female independent healthy donors (Supplementary Table 1, age range of donors: 50-59 years old). As expected, Y-linked genes *KDMSD*, *UTY*, and *EIFIAY* were exclusively expressed in male PT cells (Fig. 2f). We also studied the X-linked genes *EIFIAX* and *DDX3X*. While proposed as 'X-inactivation escapees', the extent of X-inactivation can be highly variable across genes, tissues, and individuals<sup>37</sup>. In agreement with our scRNAseq findings, primary female PT cells displayed increased transcript levels of *EIFIAX* and *DDX3X*, compared to male cells (Fig. 2f). Female sex is linked to lower oxidative stress markers in the kidney *in vivo*<sup>6</sup> but whether the sex of PT cells is a major contributor to this effect is unknown. Gene expression of *MTIF*, *MTIG*, *MTIH* was significantly increased in primary female PT cells, compared to male cells, as identified by scRNAseq and validated with qPCR in these independent donors (Fig. 2c, e, f).

Next, we investigated the biological processes enriched among the genes showing sex-biased expression in PT cells. Pathway analysis (Fig. 3a, Supplementary Data 3) revealed processes related to amino acid metabolism, PT transport, and regulation of the inflammatory response as increased in females. Among the pathways increased in males, processes related to mitochondrial aerobic metabolism ('oxidative phosphorylation', 'tricarboxylic acid (TCA) cycle' and 'electron transport chain') predominated. Two additional metabolic processes, namely 'generation of precursor metabolites' and 'nucleoside triphosphate metabolism', were also enriched in males. To validate these observations, we studied functional differences in mitochondrial metabolism and precursor metabolite generation in male and female PT cells. We exposed primary male and female PT cells to minimal media containing glucose and glutamine, which serve as mitochondrial substrates. We then measured their oxygen consumption rate (OCR), as a marker of mitochondrial respiration<sup>38</sup>. Supporting our pathway analysis, male PT cells showed a significant increase in OCR at baseline and after metabolic stress, compared to female PT cells

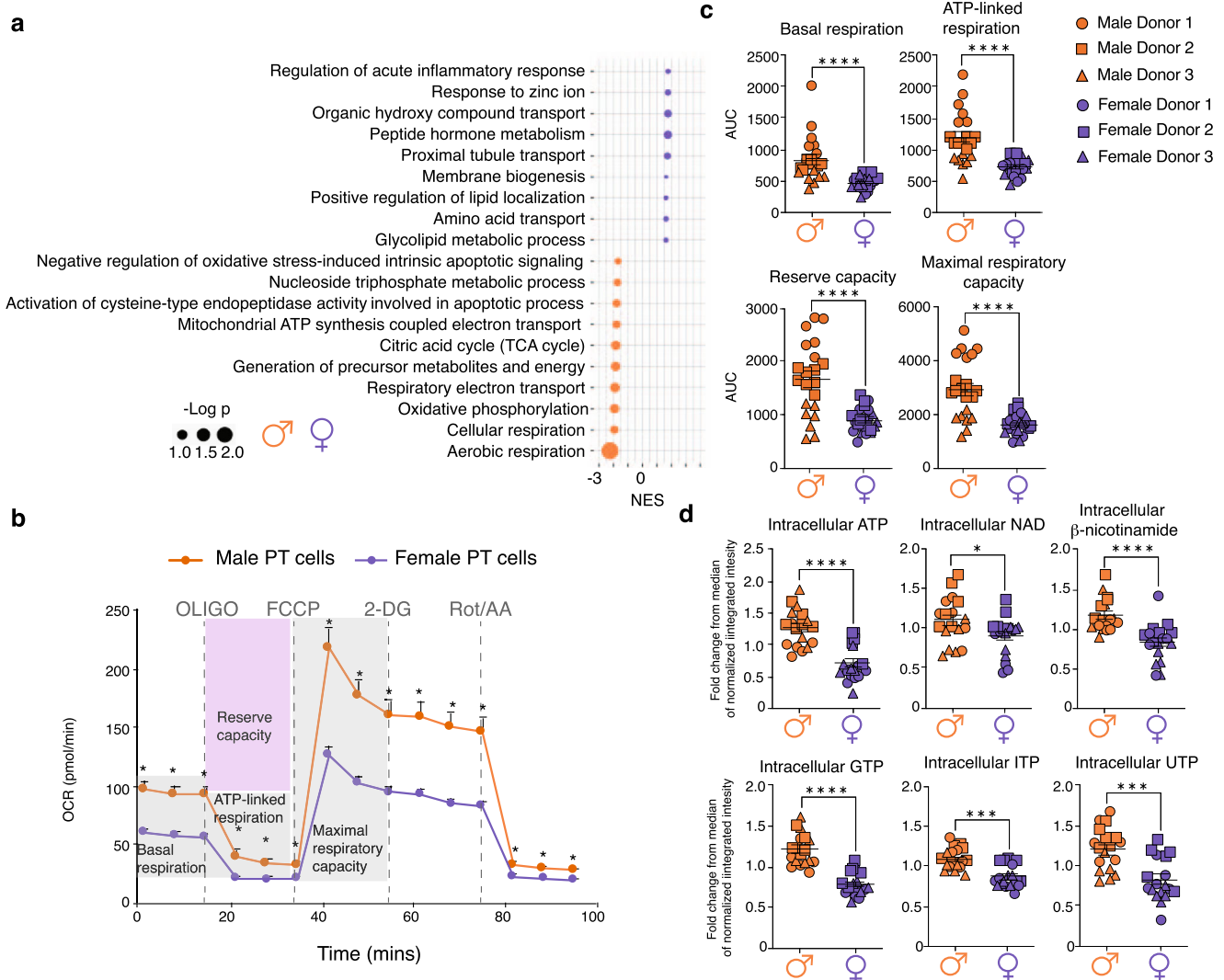
(Fig. 3b). By calculating the corresponding areas under the OCR curves, we determined that male PT cells had a significantly higher basal respiration, ATP-linked respiration, maximal respiratory capacity, and reserve capacity than female cells (Fig. 3c). Together with mitochondrial respiration, glycolysis is a major mechanism of glucose-derived energy production<sup>39</sup>. Thus, a parallel increase in glycolysis and aerobic respiration is often indicative of a higher energy state<sup>40</sup>. In addition to increased OCR, our male PT cells displayed an increase in glycolytic capacity (Supplementary Fig. 7), suggesting that they are energetically more active than female PT cells. Mitochondrial respiration results in the generation of two key energy precursors – NAD and ATP<sup>41</sup>. In line with increased aerobic metabolism, male PT cells exhibited a significant increase in the intracellular levels of NAD,  $\beta$ -nicotinamide mononucleotide (NAD precursor), ATP, and three additional nucleoside triphosphate metabolites – GTP, ITP, and UTP (Fig. 3d). In summary, these data support distinct metabolic function and gene expression in male and female PT cells.

### Immune landscape of healthy human kidney

Despite the relative paucity of immune cells in healthy human kidney, we examined kidney-resident immune cells to delineate their steady-state phenotypes and functions. Sub-clustering of immune cells yielded 12 clusters (Fig. 4a). T cells (*CD3E*<sup>+</sup>), Natural Killer (NK) cells (*NKG7*<sup>+</sup>*CD3E*<sup>+</sup>), and a small B cell population (*CD79A*<sup>+</sup>) mainly expressing the immunoglobulin chain *IGHM* were identified (Fig. 4b, Supplementary Fig. 8a). Plasma cells (*CD38*<sup>+</sup>*XBPT*<sup>+</sup>) were scarce in healthy kidney tissue (Supplementary Fig. 8b). Myeloid clusters (*CD68*<sup>+</sup>) (Fig. 4b) displayed enrichment of phagocyte-related pathways including "receptor-mediated endocytosis", "regulation of TLR signaling", and "antigen processing and presentation via MHC class II" (Supplementary Fig. 8c).

T cell cluster T1 expressed CD4<sup>+</sup> T helper (Th) cell genes (*IL7R*<sup>+</sup>*CD40LG*<sup>+</sup>*LTB*<sup>+</sup>) and enrichment of "T-helper cell differentiation" and "Interleukin-7-mediated signaling" pathways (Fig. 4c, Supplementary Fig. 8c). T1 also included *CCR7*<sup>+</sup>*SELL*<sup>+</sup> cells, suggesting central memory T cell identity (Supplementary Fig. 9)<sup>42</sup>. T2 demonstrates expression of a cytotoxic program (*GZMA*, *GZMB*, *GZMH*, *GZML*, *PRFI*) alongside NK receptor genes (*KLRD1*, *KLRG1*), consistent with effector memory T cell or NKT cell identity (Fig. 4c). T2 also contained some gamma-delta ( $\gamma\delta$ ) T cells, marked by co-expression of TCR chain components *TRDV2* and *TRDC* (Fig. 4b). T3 had sparse expression of resident memory T cell (Trm) markers ( *CXCR6*, *ITGAI*), while T4 was marked by high *GZMK* expression, a marker of circulating age-associated memory T cells (Fig. 4b, c)<sup>43</sup>. FOXP3<sup>+</sup>CD4<sup>+</sup> regulatory T cells were notably absent from scRNAseq and flow cytometry analyses (Supplementary Fig. 10a), while typically observed in kidney pathologies<sup>44,45</sup>, indicating they are likely recruited during inflammation. NK cell cluster NK1 displayed a cytotoxic gene program and broad *FCGR3A*(*CD16*) expression. Flow cytometry confirmed ~95% of renal NK cells are CD56<sup>dim</sup>CD16<sup>+</sup> (Supplementary Fig. 11a). Low abundance of ILC2s, ILC3s and CD56<sup>bright</sup> NK cells was suggested by a predictive classifier and confirmed by flow cytometry (Supplementary Fig. 11b, c).

As we noted differences in our lymphocytes signatures to those reported using other tissue sources, we directly compared lymphocytes in living donor kidney with tumor-affected renal tissue using scRNAseq (Supplementary Fig. 12a). We confirmed the presence of many similar immune populations across tissue sources, yet also observed differences in abundance and transcriptional signatures. When T cell clusters were compared between these different tissue sources, alterations in checkpoint molecule expression (i.e. *TIGIT*, *CTLA4*, *PDCDI*) were noted alongside *NR4A1* (encoding Nur77, which is induced upon antigen receptor signalling)<sup>46</sup> (Supplementary Fig. 12b). Several of these differences were also observed at the protein level by flow cytometry

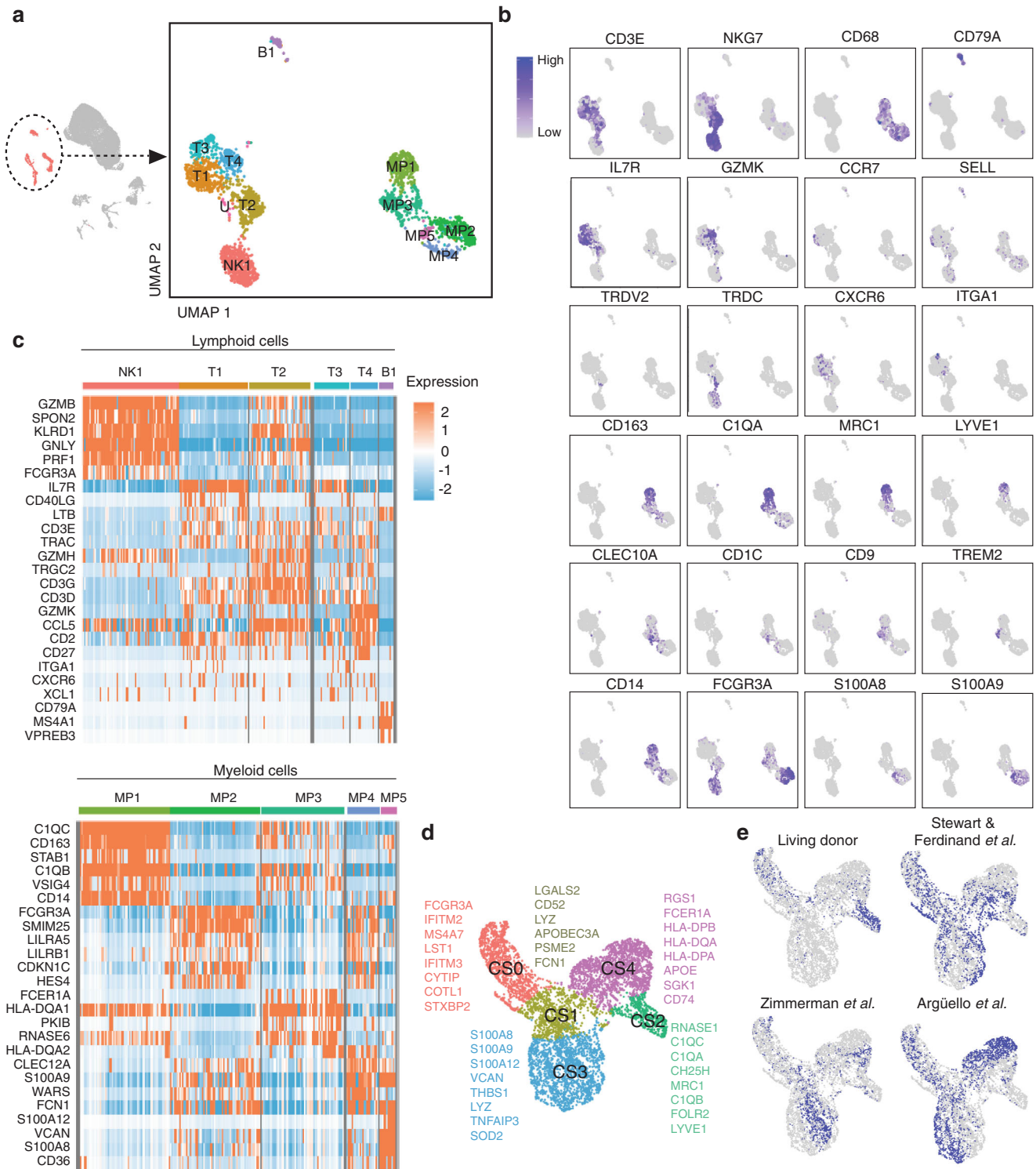


**Fig. 3 | Sex differences in mitochondrial respiration and energy precursor metabolism of proximal tubular cells.** **a** Depiction of selected significant (FDR < 0.25) terms identified by GSEA analysis as being enriched in males and females, respectively. Source data are provided in Supplementary Data 3. *P* values were calculated by Wilcoxon rank sum test without multiple comparison adjustment. **b** Oxygen consumption rate (OCR) was monitored to assess the mitochondrial respiration of male and female PT cells at baseline and after metabolic stress ( $n = 3$  donors/sex;  $n = 21$  replicates/male sex, and 23 replicates/female sex). To induce metabolic stress, the following sequence of drugs was injected: 1  $\mu$ M oligomycin, 0.3  $\mu$ M FCCP, 100 mM 2-DG, 1 mM Rot/AA. The OCR was monitored in male and female PT cells ( $n = 3$  donors/sex;  $n = 21$  replicates/male sex, and 23 replicates/female sex). Data are presented as mean values  $\pm$  SEM. **c** The basal OCR ( $p < 0.0001$ ,  $u = 48$ ), ATP-linked respiration ( $p < 0.0001$ ,  $t = 5.223$ ,  $df = 42$ ), reserve capacity ( $p < 0.0001$ ,  $t = 5.018$ ,  $df = 42$ ) and maximal respiratory capacity ( $p < 0.0001$ ,  $t = 5.281$ ,  $df = 42$ ) of male and female PT cells were calculated from the OCR curves in **b**. Group-to-group differences were assessed using two-tailed unpaired T tests for variables following a normal distribution (ATP-linked

respiration, reserve capacity, maximal respiratory capacity), and Mann-Whitney tests for variables with a non-parametric distribution (Basal respiration). Data are presented as mean values  $\pm$  SEM. **d** In a separate experiment, the intracellular levels of ATP ( $p < 0.0001$ ,  $t = 5.959$ ,  $df = 34$ ), NAD ( $p = 0.029$ ,  $u = 93$ ),  $\beta$ -nicotinamide mononucleotide ( $p < 0.0001$ ,  $t = 4.575$ ,  $df = 34$ ), GTP ( $p < 0.0001$ ,  $t = 7.45$ ,  $df = 34$ ), ITP ( $p = 0.0001$ ,  $u = 46$ ), and UTP ( $p = 0.0001$ ,  $t = 4.316$ ,  $df = 34$ ) were determined in male and female PT cells ( $n = 3$  donors/sex;  $n = 6$  replicates/donor). Data are presented as mean values  $\pm$  SEM. Source data for (**b–d**) are provided as a source data file. Group-to-group differences were assessed using two-tailed unpaired T tests for variables following a normal distribution (ATP,  $\beta$ -nicotinamide, GTP, UTP), and Mann-Whitney tests for variables with a non-parametric distribution (NAD, ITP). \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ . PT proximal tubule, AUC area under the curve, OCR oxygen consumption rate, FCCP *p*-trifluoromethoxy carbonyl cyanide phenyl hydrazone, 2-DG 2-deoxyglucose, Rot rotenone, AA antimycin A,  $df$  degrees of freedom. Orange = males, dark purple = females. Circle, square, and triangle symbols indicate biologically independent donors for each sex.

(Supplementary Fig. 12c, d). Similarly, NK cell clusters exhibited differential expression of *NR4A1*, *AREG*, and *TIGIT* (Supplementary Fig. 12e), with validation of significant differences in *TIGIT* and *AREG* expression by flow cytometry (Supplementary Fig. 12f, g). We also observed high donor heterogeneity in immune infiltration and generally a greater proportion of immune cells in nephrectomy specimens, corroborating that the immune niche in tumor-affected nephrectomy samples is altered from healthy kidney (Supplementary Fig. 12h).

Mononuclear phagocytes (MP) acquire tissue-adapted phenotypes and functions<sup>47</sup>. Definitively attributing macrophage or DC identity to myeloid populations based on gene expression alone is particularly challenging within the kidney due to a lack of consensus on lineage defining markers<sup>48</sup> and here they are annotated more generally as five MP populations. Cluster MP1 highly expressed complementary components (*CIQA*, *CIQB*, *CIQC*) and markers of alternative macrophage activation or anti-inflammatory function (*CDI63*, *LYVE1*, *STAB1*, *MRC1*, *VSIG4*, *FOLR2*) (Fig. 4b, c, Supplementary Fig. 13).



**Fig. 4 | Identification and annotation of kidney immune cells.** **a** Compartment-specific analysis of 2491 immune cells comprising 12 clusters and **b** cell type markers used for cluster annotations. **c** Heatmap of cell-type defining and highly expressed genes by each cluster separated by lymphoid and myeloid lineage.

**d** UMAP plot showing the living donor myeloid cell data clustered together with Stewart and Ferdinand et al., Zimmerman et al., and Argüello et al. to define five cell states across datasets and their respective cluster markers. **e** UMAP plots highlighting the distribution of dataset membership across the cell states.

Efferocytosis receptor *MERTK* expression supports homeostasis or repair functions (Supplementary Fig. 13). MP3 contained cells expressing cDC2 markers (*CLEC10A*, *CD1C*), alongside a subgroup of cells co-expressing lipid-associated genes (*CD9*, *TREM2*, *APOE*, *APOC1*) (Fig. 4b). Similar populations have been identified as kidney-resident macrophages and are expanded in fibrotic tissues<sup>49</sup>. MP2 and MP4 (*FCGR3A*<sup>+</sup>*SIGLEC10*<sup>+</sup>*FCN1*<sup>+</sup>) resemble *CD16*<sup>+</sup> non-classical monocytes

(Fig. 4b, c, Supplementary Fig. 13). MP4 had elevated expression of *IL1B*, MHC Class-II genes, and *CX3CR1* while MP2 had higher expression of *CXCR4* and *FPRI* (Supplementary Fig. 8d). MP5 expressed markers of classical *CD14*<sup>+</sup> monocytes (*S100A8*, *S100A9*, *CD14*, *VCAN*), yet was predominantly derived from one individual with elevated hemoglobin transcripts, indicative of increased circulating cells in this sample (Fig. 4b, c, Supplementary Fig. 8e). Flow cytometry confirmed greater

abundance of CD16<sup>+</sup> cells in kidney relative to blood, as well as low proportions of CD14<sup>+</sup>CD16<sup>-</sup> MPs resembling MP5 and the presence of MRC1<sup>+</sup>HLA-DR<sup>+</sup> MPs in kidney that align with MP1 (Supplementary Fig. 11e, f).

### Identification of a distinct resident macrophage population in healthy kidney

Due to unique aspects of our study, including short ischemic times to which resident macrophages are especially sensitive<sup>50,51</sup>, and use of flushed living donor-derived kidney tissue, we examined shared and unique MP populations in healthy kidney compared to those reported previously in scRNAseq studies of kidney tissue from other sources. CD68<sup>+</sup> cells from three prior studies<sup>3,52,53</sup> were compared and classified to match cluster identities of our study. MPs from these studies most resembled MP5 (classical CD14<sup>+</sup> monocyte-like), the lowest abundance MP cluster in living donor samples (Supplementary Fig. 14a). MP3 (DC-like and lipid-associated MPs) as well as MP2 and MP4 (CD16<sup>+</sup> non-classical monocyte-like) were shared across datasets. Strikingly, few cells from these studies corresponded to MP1 (resident macrophages) – the largest MP population in living donor kidney. Next, CD68<sup>+</sup> cells from these prior studies<sup>3,52,53</sup> and our study were merged, identifying five myeloid cell states (CS) across all studies (Fig. 4d). Based on transcriptomic profiles, CS2 and CS4 include resident macrophages and antigen-presenting cells, CS0 is consistent with non-classical CD16<sup>+</sup> monocytes, CS3 represents classical CD14<sup>+</sup> monocytes and CS1 may represent a transition state, supported by trajectory analysis (Supplementary Fig. 14b, c). CS2, which was almost entirely comprised of living donor kidney cells (Supplementary Fig. 14b), is defined by expression of genes associated with alternatively activated macrophages (*CIQA/B/C/RNASE1<sup>+</sup>CD163<sup>+</sup>LYVE1<sup>+</sup>FOLR2<sup>+</sup>*), in contrast to all other CS which expressed markers associated with monocytes and classically activated macrophages (S100 family members, *FCNI*, *LYZ*, and pro-inflammatory *SOD2*) (Fig. 4d, Supplementary Fig. 14b). This cell state aligns with descriptions of fetal-derived tissue macrophages<sup>54</sup>. CS2 constitutes the predominant MP population in healthy kidney (MP1), while CS3 and CS4 abundance is limited (Fig. 4e).

### Kidney-resident lymphocytes are antigen-experienced with distinct gene expression

Due to unexpected heterogeneity and previously unrecognized transcriptional profiles in kidney lymphocyte populations (Fig. 5a–c), we directly compared lymphocyte proportions, signatures, and phenotypes to those in healthy donor blood. Increased proportions of NK (CD3<sup>+</sup>CD56<sup>+</sup>) and NKT cells (CD3<sup>+</sup>CD56<sup>+</sup>) were noted in kidney, while T cell (CD3<sup>+</sup>CD56<sup>-</sup>) abundance was unchanged (Fig. 5a). CD8<sup>+</sup> T cells were present in higher proportions than CD4<sup>+</sup> T cells in kidney and the presence of  $\gamma\delta$ T cells was validated by flow cytometry (Fig. 5b, Supplementary Fig. 10a).

To identify specific markers and transcriptional profiles of kidney-resident lymphocytes, we integrated our dataset with publicly available PBMC scRNAseq datasets<sup>55,56</sup> (Supplementary Fig. 15a–j, Supplementary Data 4) and validated differences by flow cytometry. Unsurprisingly, blood lymphocytes exhibited higher expression of naïve T cell genes (*CCR7*, *SELL*, *LEF1*, *TCF7*). In contrast, the tissue residency-associated transcription factor *PRDMI* (BLIMP-1)<sup>57</sup> was upregulated in kidney lymphocytes, as was *CD69*, which marks tissue-resident memory T cells (Trms) in several organs and prevents tissue egress via *SIPRI* antagonism<sup>58</sup> (Supplementary Fig. 15j). Antigen-experienced T cells upregulate *CD45RO* and can become Trm<sup>59</sup>. 60–98% of kidney CD4<sup>+</sup> and CD8<sup>+</sup> T cells were CD45RO<sup>+</sup> in contrast to low proportions of Trms in blood (Fig. 5c). NK cells with memory functions may also express *CD45RO*<sup>60</sup>; however, this was not observed in renal NK cells (Fig. 5c). Flow cytometry confirmed elevated *CD69* on T cells and NK cells, with *CD69<sup>+</sup>CD103<sup>+</sup>* co-expression by CD8<sup>+</sup> T cells, consistent with a Trm phenotype (Supplementary Fig. 10c). Further

characterization of memory CD4<sup>+</sup> T helper (Th) cell subsets revealed enrichment of Th1/17 cells with reduced Th2 marker expression (Fig. 5d, Supplementary Fig. 10b).

We also sought to validate Granzyme K production in kidney lymphocytes, as cluster T4 was marked by high *GZMK* expression. In agreement with scRNAseq findings, Granzyme K was detected in 21% of kidney T cells (Fig. 5e), with minimal co-expression with Granzyme B, indicating that Granzyme K<sup>+</sup> T cells form a distinct subset of renal T cells (Fig. 5e). Most Granzyme K<sup>+</sup> T cells also did not have detectable perforin expression (Fig. 5e). This supports an extracellular function for Granzyme K produced by these T cells, rather than the canonical cytolytic function of granzymes dependent on intracellular delivery via perforin.

Kidney lymphocytes were distinguished from circulating lymphocytes by elevated expression of chemokine receptors (*CXCR4*, *CXCR6*), integrin components (*ITGB1*, *ITGA4*), and inhibitory NK receptors (*KLRD1*, *KLRG1*) (Fig. 5f, Supplementary Fig. 15j). Flow cytometry confirmed VLA-4 integrin components  $\alpha 4$  (CD49d) and  $\beta 1$  (CD29) were highly expressed in renal T cells suggesting VLA-4 contributes to their residency or function (Fig. 5g). This is consistent with expression of VLA-4 ligands fibronectin and VCAM-1 in kidney<sup>61</sup>. Kidney NK cells have higher levels of *CD69* compared to circulating NK cells, while no difference in *CD29* or *CD49d* was detected (Fig. 5h). Finally, *CXCR6* protein expression was elevated on kidney T and NK cells, while *CXCR4* was not, despite high gene expression (Fig. 5g–i). Notably, renal myeloid cells expressed *CXCL16*, the chemokine ligand for *CXCR6*, indicating participation in lymphocyte recruitment, a finding which was supported by significant aggregate rank scores using cell-cell communication inference (Supplementary Fig. 10d, Supplementary Data 5, 6).

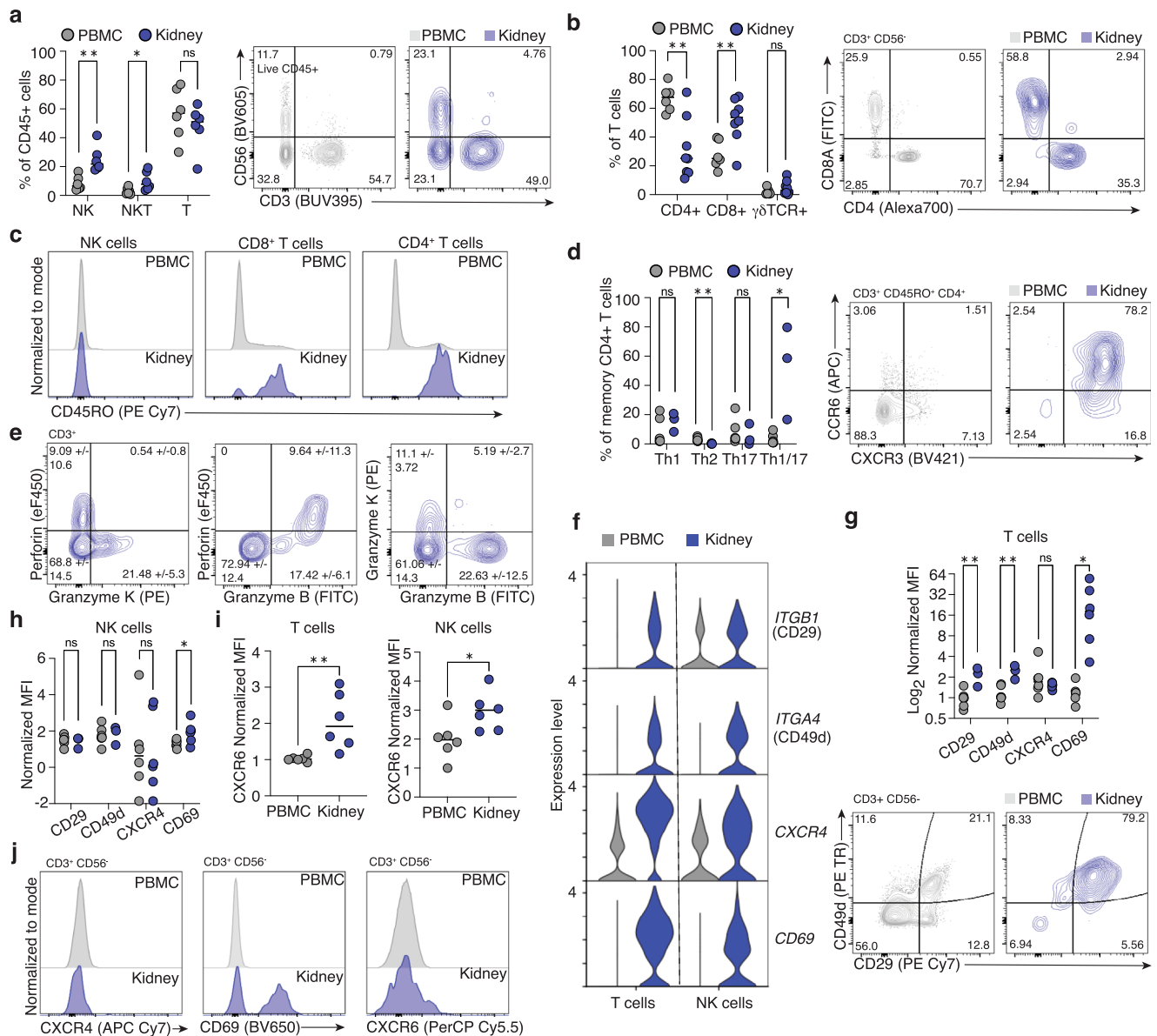
Other differentially expressed genes suggest tissue-adapted function of kidney lymphocytes. *AREG*, encoding the growth factor amphiregulin, was highly expressed by NK1 and validated by flow cytometry (Supplementary Fig. 11d, Supplementary Fig. 12f, g), suggesting tissue-reparative functions. The prostaglandin E2 (PGE2) receptor *PTGER4* and prostaglandin D synthase *PTGDS* were upregulated (Supplementary Fig. 15j), indicating that kidney lymphocytes synthesize and recognize prostaglandins, known mediators of kidney function<sup>62</sup>. PGE2 promotes Th17 and Th1/17 cell development and function, perhaps explaining high proportions of Th1/17 cells observed in healthy human kidney (Fig. 5d)<sup>63</sup>. Collectively, these studies capture the heterogeneity of myeloid and lymphoid populations within healthy human kidney and provide an important reference of immune cell phenotypes and functions at steady state.

## Discussion

We present a scRNAseq study of healthy human kidney using biopsies from living donors. Our high-resolution map detailing the diversity in healthy kidney PT, endothelial, epithelial, and immune subpopulations will inform future studies addressing underlying mechanisms of kidney pathologies, including chronic kidney disease, fibrosis, ischemia-reperfusion injury, renal cancer and allograft rejection. In particular, the non-inflamed nature of this tissue will be particularly relevant as a baseline for studies of immune-mediated kidney disease.

The sex-balanced design in the present study enabled the examination of sex-based dichotomy in gene expression among human kidney cell populations. Prior studies were constrained by small sample size and use of animal models, or instead used bulk transcriptional analysis where sex-specific signatures of individual kidney cell populations cannot be resolved<sup>22,26,64</sup>. Our study is aligned with the conclusion of scRNAseq studies in mouse by Ransick et al.<sup>26</sup> that PT cells are sexually dimorphic. However, the overlap in sexually dimorphic PT genes between human and mouse is small, perhaps due to distinct orthologues in mouse, small number of samples sequenced, or true biological differences between human and mouse.





**Fig. 5 | Characterization of kidney-resident T and NK cells.** **a** NK cells ( $p = 0.0025$ ,  $t = 3.998$ ,  $df = 10$ ) and NKT cells ( $p = 0.0327$ ,  $t = 2.476$ ,  $df = 10$ ) are proportionally enriched in kidney relative to blood, while T cell ( $p = 0.379$ ,  $t = 0.918$ ,  $df = 10$ ) abundance is unchanged. ( $n = 6$ ) **(b)** Within the kidney T cell population, there is an enrichment of CD8<sup>+</sup> T cells ( $p = 0.0060$ ,  $t = 3.327$ ,  $df = 12$ ) and a reduction in CD4<sup>+</sup> T cell abundance ( $p = 0.0025$ ,  $t = 3.815$ ,  $df = 12$ ) with no change in TCR $\gamma\delta$ <sup>+</sup> T cells ( $p = 0.2158$ ,  $u = 14$ ) relative to blood. ( $n = 6$ ) **(c)** Kidney T cells are predominantly antigen-experienced, marked by expression of CD45RO. **(d)** Within kidney memory CD4<sup>+</sup> T cells, there is an enrichment in the Th1/17 subpopulation (CXCR3<sup>+</sup>CCR6<sup>+</sup>) ( $p = 0.0238$ ,  $u = 0$ ) and a reduction in Th2 subpopulation (CRTh2<sup>+</sup>) abundance ( $p = 0.0098$ ,  $t = 3.513$ ,  $df = 7$ ) relative to blood while Th1 (CXCR3<sup>+</sup>) ( $p = 0.3810$ ,  $u = 5$ ) and Th17 (CCR6<sup>+</sup>) ( $p = 0.5476$ ,  $u = 6$ ) proportions were unchanged. ( $n = 3$ ) **(e)** T cells expressing Granzyme K do not co-express perforin, indicating that they are a distinct T cell subset from Granzyme B<sup>+</sup> Perforin<sup>+</sup> cytotoxic T cells. **(f)** Violin plots showing differential gene expression of select markers in kidney T cells and NK cells relative to blood. **(g)** Surface levels of

CD29 ( $p = 0.0061$ ,  $t = 3.869$ ,  $df = 7$ ), CD49d ( $p = 0.0027$ ,  $t = 4.519$ ,  $df = 7$ ) and CD69 ( $p = 0.0203$ ,  $t = 2.756$ ,  $df = 10$ ) were higher on kidney T cells relative to blood as measured by flow cytometry, while CXCR4 ( $p = 0.5887$ ,  $u = 14$ ) was not. ( $n = 6$ ) **(h)** Surface CD69 ( $p = 0.0427$ ,  $t = 2.321$ ,  $df = 10$ ) was higher on kidney NK cells relative to blood while CD29 ( $p = 0.6899$ ,  $t = 0.4159$ ,  $df = 7$ ), CD49d ( $p = 0.9040$ ,  $t = 0.1250$ ,  $df = 7$ ), and CXCR4 ( $p = 0.9326$ ,  $t = 0.0868$ ,  $df = 10$ ) were not. ( $n = 6$ ) **(i)** CXCR6 abundance was higher at the protein level on both T cells ( $p = 0.0086$ ,  $t = 3.258$ ,  $df = 10$ ) ( $n = 6$ ) and NK cells ( $p = 0.0364$ ,  $t = 2.414$ ,  $df = 10$ ) ( $n = 6$ ) relative to blood. **(j)** Histograms showing no difference in CXCR4, increased CD69 and increased CXCR6 protein abundance in kidney T cells relative to blood. Group-to-group differences were assessed using two-tailed unpaired T tests for variables following a normal distribution, and Mann-Whitney tests for variables with a non-parametric distribution. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ . PBMC Peripheral blood mononuclear cells, NK Natural Killer cell, NKT Natural Killer T cells. Gray = PBMCs, Blue = Kidney. Source data for all panels provided as a source data file.

We report striking sex-based transcriptional differences in PT cells, suggesting higher baseline metabolic activity in males, and enhanced expression of antioxidant genes in females. We validated these sex-based observations at the level of gene expression, metabolite generation, and metabolic function. Increased oxidative stress is reported in males<sup>65</sup>, while female sex hormones augment antioxidant

gene transcription<sup>66</sup>. Metallothionein genes (*MT1F*, *MT1G*, *MT1H*), which are potent endogenous antioxidants<sup>67</sup>, were increased in female PT cells. Metallothionein depletion exacerbates diabetic and hypoxia-induced kidney injury<sup>68,69</sup>, whereas augmented expression is protective<sup>70</sup>. Several sex-altered genes further relate to cysteine-glutathione metabolism. Glutathione is critical to cellular antioxidant

defences<sup>71</sup> and glutathione metabolism exhibits sexual dimorphism<sup>25,72</sup>. Identification of these sex-based differences in PT gene expression (enabled by use of scRNAseq which can capture transcripts localized to the mitochondria and cytosol), may provide insights into the well-recognized, but previously unexplained sexual dimorphism observed in most kidney diseases. In particular, why females may be less susceptible to metabolism-related kidney injury<sup>6–8,73,74</sup>.

Our study provides a steady-state map of the kidney immune niche and identifies differences in abundance and phenotype of myeloid and lymphoid populations compared to other kidney tissue sources and peripheral blood. Kidney T cells are predominantly Trms and exhibit unique phenotypes previously unreported in kidney, including Granzyme K<sup>+</sup> T cells. The function of Granzyme K<sup>+</sup> T cells in humans is poorly characterized, and here we show that Granzyme K<sup>+</sup> T cells are a distinct subset separate from Granzyme B<sup>+</sup>Perforin<sup>+</sup> T cells in the kidney. The lack of perforin co-expression suggests that Granzyme K produced by renal T cells may have extracellular targets, such as inducing endothelial cell activation<sup>75</sup>, promoting sensitivity to LPS-induced inflammation<sup>76</sup>, and regulating angiogenesis<sup>77</sup>.

Renal CD4<sup>+</sup> memory Th cells are skewed towards a Th1/17 phenotype, which may be relevant to Th17-related kidney diseases including glomerulonephritis, lupus nephritis, and rejection<sup>78,79</sup>. Renal abundance of CD56<sup>+</sup>CD16<sup>+</sup> NK cells with high expression of amphiregulin compared to circulating NK cells suggests non-canonical tissue-adapted functions. We demonstrate an enrichment of a resident macrophage population, akin to fetal-derived tissue macrophages, with little-to-no presence in prior datasets from discarded deceased donor or tumor nephrectomy specimens, suggesting altered kidney environments impact this myeloid population. Indeed, sensitivity of self-renewing resident macrophage populations to extended ischemic injury and inflammation is reported<sup>51</sup>. Additional comparison of lymphocyte populations in tumor-affected versus living donor renal tissue revealed alterations in tumor-affected tissue relative to the steady-state immune niche in healthy living donor kidney. Increased B and T cell proportions, increased expression of activation and exhaustion-associated molecules by lymphocytes, in addition to a trend for increased immune infiltration in tumour-adjacent nephrectomy specimens was observed (Supplementary Fig. 12), in agreement with prior reports that tumour-affected kidneys can have altered immune infiltrates<sup>5,80</sup>. These findings may have implications on the future studies exploring alterations in immune cells in unaffected kidney tissue of renal cancer patients, for development of immunotherapies.

There are some limitations to using scRNAseq for tissue studies. One such limitation is that tissues must undergo dissociation into single cells, which can result in cell stress and damage, as well as loss of their spatial relationships. While we took care to minimize cell damage in our kidney dissociation protocol, including establishing a very short 20 minute dissociation time and minimizing additional processing, warm dissociation can impact cells, in particular, epithelial cells. This often results in proportional increases in mitochondrial transcripts, as damaged cell membranes are leaky to cytosolic transcripts. Sequencing chemistry can further impact this issue. For example, the proportion of mitochondrial reads captured by 10X Genomics 3' v3 kits is 2-3 times greater than that captured by v2 kits<sup>81,82</sup>. Other groups have addressed this by using higher mitochondrial thresholds for tissues sequenced with 10X Genomics 3' v3 kits as compared with 10X Genomics 3' v2 kits within the same study<sup>82</sup>. 10X Genomics 3' v3 provides higher mRNA detection sensitivity which results in increased power to detect differentially expressed genes, and is superior to v2 chemistry for profiling immune cells<sup>83</sup>. Processing tissues with techniques such as microdissection can yield a lower fraction of mitochondrial reads than droplet-based scRNAseq<sup>84</sup> but does not provide high throughput capacity to profile a broad range of cells.

Nevertheless, the proportion of mitochondrial reads in our dataset is in line with other published kidney scRNAseq datasets, which retained cells with a maximum of 25-80% mitochondrial reads<sup>3,82,85,86</sup>. The lack of batch effects noted across our samples obviated the need for batch correction tools. Furthermore, the lack of bias in dissociation artifacts across different sample groups (e.g. male and female PT cells), as well as our ability to validate observations in external datasets and utilizing orthogonal methods, leads us to conclude that transcriptional variability due to dissociation did not greatly impact our data. Several groups have chosen to use snRNA sequencing to avoid the issue of dissociation artifacts; however, it is known that certain cell types such as lymphocytes are not well suited to capture by snRNAseq due to their low nuclear RNA content. In addition, we demonstrate cytosolic epithelial cell transcripts exhibit sex-biased expression, a finding detectable using scRNAseq, which is not restricted to nuclear transcripts. It is likely, however, that dissociation stress would play a greater role in biologically stressed or damaged tissues, such as in various inflammatory contexts including autoimmunity and transplant rejection. scRNAseq and snRNAseq each offer slightly distinct yet also overlapping information, suggesting that combinations of these methods, together with spatial profiling, would be ideal to dissect kidney biology further.

Collectively, our study of healthy human kidney provides an important reference point for understanding the cellular basis of kidney disease development, represents a 'normal' target for stem cell-derived kidney organoids, and expands our understanding of the complexity of sex-based gene expression and kidney-resident immune populations.

## Methods

### Experimental model and subject details

**Human specimens.** Kidney tissue from tumour-unaffected nephrectomy specimens was used for initial method optimization. Pre-implantation core biopsies were obtained from living donor kidneys after organ retrieval and flushing. 20 living donor kidney samples (10 male donors and 10 female donors) were processed for sequencing. One sample was subsequently removed due to quality control reasons, as described below. Additional living donor kidney samples were used in flow cytometry experiments for method optimization and immunophenotyping. All experiments were conducted with institutional ethics approval from University Health Network (CAPCR: 18-5914, Living donor; CAPCR: 18-5489.0, Tumour nephrectomy). Patient demographic information for sequenced samples and samples used for flow cytometry are summarized in Supplementary Table 2. All patients provided informed written consent for inclusion in this study. Compensation was not provided for participation in this study.

**Murine specimens.** Kidneys from female 8 week old C57BL/6 mice were used for digestion optimization experiments. Mice were ordered directly from Jackson Laboratories under Animal Use Protocol 6156 approved by the Toronto General Hospital Research Institute Animal Care Committee.

### Experimental method details

**Tissue digestion and CD45-enrichment.** All living donor samples used for sequencing were processed within one hour of organ retrieval. Briefly, biopsies were collected in RPMI 1640 (Gibco, cat # 11875119) on ice, and mechanically dissociated with a blade before enzymatic digestion at 37 °C with 0.1 mg/ml DNase I (STEMCELL, cat # 07470), 3300 CDA units/ml Collagenase MA (VitaCyte, cat # 001-2030) and 1430 NP units/ml BP neutral protease (VitaCyte, cat # 003-1000) for 20 minutes at 37 °C with intermittent agitation in a dissociation protocol optimized to maximize viability and to preserve representation of rare and fragile cell populations (Supplementary Fig. 15). Cell suspensions were filtered through 35 µm cell strainer snap-cap FACS

tubes (Falcon, cat# 352235) and a plunger from a 1 ml syringe was used to gently mash remaining tissue in the strainer before rinsing strainer lid with 1:1 volume of FBS (HyClone, cat # SH3039603PM) on ice. A low frequency (<1%) of immune cells in the single cell suspension from a kidney biopsy core (Fig. 1a) necessitated immune enrichment in 10 samples (5 males, 5 females) using magnetic EasySep Human CD45 depletion kit II (STEMCELL, cat# 17898), as per the manufacturer's modified instructions for positive selection of CD45-expressing cells.

**Single-cell RNA sequencing.** Samples were prepared according to 10X Genomics Single Cell 3' v3 Reagent kit user guide<sup>87</sup>. The pilot sequencing sample from tumour-unaffecting nephrectomy tissue was sequenced using 10X Genomics Single Cell 5' v2 Reagents. Samples were washed twice in PBS (Life Technologies) plus 0.04% BSA, and viability was determined by a hemocytometer (Thermo Fisher) via Trypan Blue staining. Following counting, the appropriate volume for each sample was calculated for a target capture of 9,000 cells. For CD45-enriched samples, all cells were sequenced. Samples that were too low in cell concentration as defined by the user guide were washed, re-suspended in a reduced volume and counted again using a haemocytometer prior to loading onto the 10x single cell B chip. After droplet generation, samples were transferred onto a pre-chilled 96-well plate (Eppendorf), heat sealed and incubated overnight in a Veriti 96-well thermos cycler (Thermo Fisher). The next day, sample cDNA was recovered using Recovery Agent provided by 10x and subsequently cleaned up using a Silane DynaBead (Thermo Fisher) mix as outlined by the user guide. Purified cDNA was amplified for 11 cycles before being cleaned up using SPRIselect beads (Beckman). Samples were diluted 4:1 (elution buffer (Qiagen):cDNA) and run on a Bioanalyzer (Agilent Technologies) to determine cDNA concentration. cDNA libraries were prepared as outlined by the Single Cell 3' Reagent Kits v3 user guide with modifications to the PCR cycles based on the calculated cDNA concentration.

The molarity of each library was calculated based on library size as measured bioanalyzer (Agilent Technologies) and qPCR amplification data (Roche). Samples were pooled and normalized to 1.5 nM. Library pool was denatured using 0.2 N NaOH (Sigma) for 8 minutes at room temperature, neutralized with 400 mM Tris-HCl (Sigma). Library pool at a final concentration of 300pM were loaded to sequence on Nova-seq 6000 (Illumina). Samples were sequenced with the following run parameters: Read 1-28 cycles, Read 2- 90, index 1-10 cycles, index 2-10 cycles. Across samples, cells were sequenced to a target depth of 40,000 reads per cell. Mapping and quantification were performed using the 10X Genomics Cell Ranger pipeline version 3.1.0. Cell metric summaries for each sample in Supplementary Data 7.

**Single-nucleus RNA sequencing.** A pilot single-nucleus RNA sequencing experiment was undertaken to compare single-cell versus single-nuclear results from a matched sample. The biopsy from a 58-year-old male was collected fresh and divided into 8 segments, evenly distributed to be processed fresh for single cell RNA sequencing as above, and the remainder was flash frozen in liquid nitrogen. The sample was later retrieved from liquid nitrogen and processed on dry ice according to the protocol in<sup>88</sup> with a lysis buffer containing: 0.32 mM sucrose (BioShop SUC507.1), 5 mM CaCl<sub>2</sub> (VWR, 97062-820), 3 mM MgCl<sub>2</sub> (Thermo Fisher AM9530G), 20 mM Tris-HCl pH 7.5 (Thermo Fisher, 15567027), 0.1% TritonX-100 (Sigma Aldrich T8787-50ML), 0.1 mM EDTA pH 8.0 (Thermo Fisher AM9260G), 40 U/ml Protector RNase inhibitor (Sigma Aldrich 3335399001) in UltraPure DNase/RNase-free water (Thermo Fisher 10977015). The nuclei were captured and sequenced using 10X Genomics Single Cell 3' v3 Reagents as above.

**Data quality control, clustering, differential expression, pathway analysis and cell-cell interaction inference.** Original study recruitment included samples from 20 donors, however, data from one male

donor was poor quality and was excluded from downstream analysis. Thus, our final dataset consisted of 19 donors (10 female, 9 male), with 10 CD45-enriched samples (5 female, 5 male) and 9 samples not enriched for CD45<sup>+</sup> cells referred to as "total kidney" (5 female, 4 males). Kidney donor information is summarized in Supplementary Table 2. To preserve representation of rare cell types with uniquely expressed genes, we retained genes expressed in a minimum of 1 cell in the individual datasets.

Ambient RNA contamination was corrected using the AutoEst function in SoupX<sup>89</sup> (Supplementary Fig. 17). DoubletFinder<sup>90</sup> was used to identify and remove cells most likely to be doublets, rather than implementation of a maximum gene or feature threshold. For total samples, a high doublet rate threshold of 7.5% was applied (as utilized in comparable studies<sup>85</sup>), while for CD45-enriched samples, the doublet rate was calculated as 0.8% per 1000 cells captured, as per 10X Genomics estimated doublet rates<sup>87</sup>. The individual datasets were then merged. Upon merging all of the individual datasets, the cells clustered according to cell type rather than donor/batch, and importantly, no batch correction of the data was required.

Cell type-specific thresholds were set to remove low-quality cells. For immune cells (clusters expressing *PTPRC*), all cells with >10% of UMIs mapped to mitochondrial genes were removed, along with cells that had low transcript abundance (<1000) or gene diversity (<200 unique genes). Separately, prior to removing cells with low transcripts/features, data was mined for the presence of granulocyte lineage cells such as neutrophils which are often removed by typical QC thresholds due to high RNase activity and low gene content, however very few neutrophils (>20) were identified by marker expression in the raw data across all samples. For parenchymal cells, all cells with >40% of mitochondrial-mapped UMIs were removed; this high threshold was imposed due to known high mitochondrial content of proximal tubular cells<sup>91</sup>. Additionally, parenchymal cells with low transcript abundance (<1000) and low gene diversity (<750 unique genes) were removed. Cells expressing hemoglobin genes (*HBB*, *HBA1/2*) ( $n=160$ ) were removed. Following normalization (SCTransform<sup>92</sup>) and feature selection (M3Drop/DANB<sup>93</sup>), principal component analysis was used for dimensionality reduction (RunPCA) and cells were clustered using the Louvain algorithm with 30 principal components (FindNeighbors and FindClusters) (Seurat<sup>94</sup>). Clusters were visualized using UMAP algorithm<sup>95</sup>.

The dataset was divided into 3 broad subgroups identified as being Immune (*PTPRC*<sup>+</sup>) or Parenchymal (Proximal Tubular (expressing *CUBN*, *HNF4A*, *SLC34A1*, *LRP2*, *SLC17A1*) or non-Proximal Tubular) in origin. These subgroups were re-clustered and further annotated using a curated marker list (Supplementary Data 8). Cluster-defining genes were identified by Seurat's FindMarkers<sup>94</sup>. Full lists of both median and average levels of gene expression per cluster and per cell type are provided in Supplemental Data 9–12, respectively.

Ranked gene lists were generated using Wilcoxon rank sum testing from the presto package (wilcoxau function) (<https://github.com/immunogenomics/presto>) were used as input for pathway analysis using GSEA<sup>96</sup>. Reference gene sets were acquired from the Bader lab repository ([http://download.baderlab.org/EM\\_Genesets/](http://download.baderlab.org/EM_Genesets/)) – Geneset used: (Human\_GOBP\_AllPathways\_no\_GO\_ia\_January\_13\_2021\_symbol.gmt.txt). To identify pathways enriched in immune cell clusters, the ranked gene lists were generated for each cluster comparing that cluster versus all other clusters.

Cell-cell communication was inferred from the sequencing data using LIANA which generates a consensus ranking across several methods<sup>97</sup>. The OmniPath interaction database was used<sup>98</sup> with the following methods for inferring interactions implemented through the package: SingleCellSignalR<sup>99</sup>, iTALK<sup>100</sup>, NATMI<sup>101</sup>, Connectome<sup>102</sup>, CellChat<sup>103</sup> and CellPhoneDB<sup>104</sup>. Results are summarized in Supplementary Data 5. Separately, SingleCellSignalR, NATMI, iTALK and

Connectome methods were used to generate a consensus score using the CellPhoneDB database to infer interactions inclusive of multimeric complexes as accounted for in the CellPhoneDB interaction database, summarized in Supplementary Data 6.

**Identification of innate lymphoid cells.** A predictive tool for cell type classification (scPred<sup>105</sup>) was trained on single-cell data generated from flow cytometry-sorted ILCs<sup>106</sup> and T cells<sup>107</sup>. Using this classifier, some cells present within our dataset were putatively identified as ILCs.

**Transcription factor analysis.** Top cluster-defining genes for PT5 and PT3, respectively were uploaded to CHEA3<sup>108</sup> (<https://maayanlab.cloud/chea3/>), and the top 10 predicted upstream regulators were identified.

**Comparison of proximal tubular cells to external datasets.** The dataset associated with Muto et al.<sup>14</sup>, derived from non-tumour kidney cortex samples was compared with our dataset of healthy human kidney. The author-defined annotations, and clustering from Muto et al. were used as a reference from which to annotate the proximal tubular cells from our dataset. Our living donor proximal tubular clusters (PT1-PT6) were annotated using these reference annotations using SingleR, which assigns identities to cells from query datasets on the basis of correlation of transcriptional signatures with a reference dataset<sup>17</sup>. The annotation in SingleR is performed for each single cell independently, rather than on a percluster basis.

For correlation analysis, differentially expressed genes defining male PT cells versus female PT cells, or PT5 and PT\_VCAM1 relative to other clusters within each dataset were identified using Seurat's FindMarkers function. Genes overlapping between the two marker lists were retained and plotted by LogFC. Spearman correlations were calculated using the rcorr function in the Hmisc package.

**Gene signature scoring.** The top 40 genes identified as overexpressed in warm collagenase dissociation<sup>35</sup> were used as an input signature and the gene module was scored for male versus female proximal tubule cells using the UCell package<sup>109</sup>.

**Comparison of kidney immune cells to PBMCs.** To identify differences in gene expression between T cells and NK cells from peripheral blood versus kidney, PBMC data (GSE148665)<sup>55</sup> was integrated with the immune only kidney data using Harmony<sup>110</sup>. A second independent PBMC dataset<sup>56</sup>, was separately integrated with the kidney data for dataset-independent validation. NK cells and T cells (clusters expressing *NKG7* and/or *CD3E*) were compared using Seurat's FindAllMarkers function. Violin plots and volcano plots were created using Seurat and EnhancedVolcano<sup>111</sup>.

**Comparison of myeloid cells.** To identify differences in myeloid cell populations in living kidney donors compared to publicly available human kidney single-cell RNA sequencing datasets from tumour nephrectomy or deceased donor tissue sources, CD68-expressing clusters from Stewart & Ferdinand et al.<sup>3</sup>, Zimmerman et al.<sup>52</sup>, and Argüello et al.<sup>53</sup> were scored using a random forest classifier (SingleCellNet<sup>112</sup>) to identify cells from the published datasets corresponding to the five myeloid clusters in the living donor data. Separately, all myeloid cells from this data and the three published studies were integrated and clustered to identify cell states using OCAT<sup>113</sup>. The datasets were also integrated and batch corrected using Seurat v3 integration (FindIntegrationAnchors and IntegrateData functions). The cell state identities from OCAT were mapped onto the integrated object and marker genes of cell states were identified using Seurat's FindAllMarkers function. Lineage analysis by pseudotime inference was applied to the OCAT-identified clustering of the combined

myeloid populations using slingshot<sup>114</sup>, without indicating any clusters as either start or endpoints.

**Sex differences analysis.** Principal component analysis (PCA) followed by Varimax rotation was performed on all major parenchymal and immune populations. Varimax-rotated principal components 2:25 were serially plotted against component 1, to identify whether a separation on the basis of sex was evident. If seen, the top 100 genes (50 from each end of the gene loading list) associated with the Varimax-rotated principal component were retained for further analysis.

Sex differences in proximal tubular cells were identified using sparse partial least squares discriminant analysis (sPLS-DA) in mixOmics<sup>115</sup>. Using the tuning function (tune.splsda), the optimal values for sparsity parameters were determined to be 1 component with 80 variables (genes). To test the classifier, the data were separated into a training dataset (2/3 of cells sampled) and a query dataset (remaining 1/3). Next, our 80-gene signature was applied to an external dataset (Liao et al.<sup>23</sup>) for validation. Here, the entire living donor dataset was used as the training dataset and the external dataset was used as the query dataset. To determine the contribution of sex chromosome-encoded genes to the model, all X- and Y-chromosome-encoded genes were removed from the datasets prior to analysis, where the tuned parameters identified the optimal model to include 1 component with 15 variables. This 15-gene signature was also validated in the Liao et al. dataset. Hierarchical structure, zero inflation, and pseudoreplication bias in single-cell data pose specific challenges for differential expression analyses<sup>116–118</sup>. To circumvent these limitations, we implemented a mixed effects model using MAST<sup>118,119</sup>. For differential expression testing between male and female proximal tubule cells, the dataset was filtered to include only genes which were expressed in each sample (9792 genes). Differential expression testing was conducted using MAST with a random effect for sample (lmm-cellular detection rate + donor sex + (1| sampleID)). As this approach excluded genes expressed exclusively by one sex (e.g. Y chromosome encoded genes, and XIST), such genes were added to MAST differentially expressed genes (MAST +) for comparison with the results of the other methods (Varimax, sPLS-DA).

All significant genes returned using MAST analysis were subjected to enrichment analysis (GSEA<sup>96,120</sup>) using reference gene sets acquired from the Bader lab repository: ([http://download.baderlab.org/EM\\_Genesets/](http://download.baderlab.org/EM_Genesets/)); Geneset used: (Human\_GOBP\_AllPathways\_no\_GO\_ia\_January\_13\_2021\_symbol.gmt.txt).

**Cryopreservation.** Cells from additional (non-sequenced) fresh living donor biopsies or cells remaining following 10X cell capture for sequencing were resuspended in 90% human serum (Sigma, cat# H4522) and 10% DMSO for cryopreservation and cooled to  $-80^{\circ}\text{C}$  in a Mr.Frosty (Sigma, cat #C1562), then transferred to liquid nitrogen for long term storage.

**Flow cytometry.** After fresh tissue digestion, cells were washed in PBS + 2% FCS before staining. Cryopreserved cells were thawed and washed twice in PBS + 2% FCS. Cells were incubated at  $4^{\circ}\text{C}$  for 15 min with an Fc receptor blocker (BioLegend TruStain FcX, cat # 422302) according to manufacturer instructions before cocktails of surface antibodies were added for 30 min at  $4^{\circ}\text{C}$ . If intracellular targets/transcription factors were included in the panel, cells were resuspended in FOXP3 transcription factor fix perm buffer (eBio, cat # 00-5523-00) and stained with intracellular antibodies in 1X permeabilization buffer (eBio, cat # 00-8333-56). If no intracellular targets were included in the staining panel, cells were fixed in 2% PFA (Thermo Scientific, cat # J19443) after surface staining.

Cells were stained with the following surface antibodies: Anti-human CD8a FITC (1:100, clone RPA-T8, BioLegend, cat # 301050),

Anti-human TCRgd FITC (1:100, clone B1, BioLegend, cat # 331208), Anti-human CD3 FITC (1:100, clone UCHT1, BioLegend, cat # 300440), Anti-human CD8a PerCP (1:50, clone RPA-T8, BioLegend, cat # 301030), Anti-human CXCR6 PerCP Cy5.5 (1:50, clone K041E5, BioLegend, cat # 356010), Anti-human CCR8 PE (1:100, clone L263G8, BioLegend, cat # 360604), Anti-human CD127 PE (1:50, clone hIL-7R-M2L, BD Biosciences, cat # 557938), Anti-human CD15 PE (1:100, clone W6D3, BD Biosciences, cat # 562371), Anti-human CD163 PE (1:50, clone GHI/61, BioLegend, cat # 333606), Anti-human CD49d PE Dazzle 594 (1:100, clone 9F10, BioLegend, cat # 304325), Anti-human CRTh2 PE Dazzle 594 (1:50, clone BM16, BioLegend, cat # 350126), Anti-human CD31 PE Dazzle 594 (1:100, clone WM59, BioLegend, cat # 303130), Anti-human CD16 PE Dazzle 594 (1:100, clone 3G8, BioLegend, cat # 302054), Anti-human CD45 PE-CF594 (1:100, clone HI30, BD Biosciences, cat # 562279), Anti-human CD29 PE Cy7 (1:100, clone TS2/16, BioLegend, cat # 303025), Anti-human CD45RO PE Cy7 (1:50, clone UCHL1, BD Biosciences, cat # 560608), Anti-human MerTK PE Cy7 (1:50, clone 590HIIGIE3, BioLegend, cat # 367610), Anti-human TIGIT PE Cy 7 (1:50, clone MBSA43, Invitrogen, cat # 25-9500-42), Anti-human CD94 APC (1:100, clone HP-3D9, eBioscience, cat # 17-5094-42), Anti-human CCR6 APC (1:25, clone G034E3, BioLegend, cat # 353416), Anti-human CD206 APC (1:50, clone 15-2, BioLegend, cat # 321110), Anti-human CD4 Alexa700 (1:50, clone RPA-T4, eBioscience, cat # 56-0049-42), Anti-human CD127 Alexa700 (1:50, clone eBioRDR5, eBioscience, cat # 56-1278-12), Anti-human CXCR4 APC Cy7 (1:50, clone 12G5, BioLegend, cat # 306528), Anti-human CTLA-4 APC Cy7 (1:25, clone BNI3, BioLegend, cat # 369634), Anti-human CD56 APC Cy7 (1:50, clone HCD56, BioLegend, cat # 318332), Anti-human CD45 APC Cy7 (1:100, clone HI30, BioLegend, cat # 304014), Anti-human CD14 APC eF780 (1:100, clone 61D3, eBioscience, cat # 47-0149-42), Anti-human CXCR3 BV421 (1:50, clone G025H7, BioLegend, cat # 353716), Anti-human CD13 BV421 (1:50, clone WM15, BioLegend, cat # 301716), Anti-human TCRgd BV510 (1:100, clone B1, BioLegend, cat # 331220), Anti-human TCRab BV510 (1:100, clone IP26, BioLegend, cat # 306734), Anti-human CD5 BV510 (1:100, clone L17F12, BioLegend, cat # 364018), Anti-human FcER1 BV510 (1:100, clone AER-37, BioLegend, cat # 334626), Anti-human CD303 BV510 (1:100, clone 201A, BioLegend, cat # 354232), Anti-human CD123 BV510 (1:100, clone 6H6, BioLegend, cat # 306022), Anti-human CD34 BV510 (1:100, clone 581, BioLegend, cat # 343528), Anti-human CD20 BV510 (1:100, clone 2H7, BioLegend, cat # 302340), Anti-human CD3 BV510 (1:100, clone OKT3, BioLegend, cat # 317332), Anti-human CD14 BV510 (1:100, clone MSE2, BioLegend, cat # 301842), Anti-human CD19 BV510 (1:100, clone HIB19, BioLegend, cat # 302242), Anti-human CD4 BV510 (1:100, clone RPA-T4, BioLegend, cat # 300546), Anti-human CD56 BV605 (1:50, clone HCD56, BioLegend, cat # 318334), Anti-human CD69 BV650 (1:100, clone FN50, BioLegend, cat # 310934), Anti-human CD8a BV650 (1:50, clone RPA-T8, BioLegend, cat # 301042), Anti-human CD326 BV650 (1:100, clone 9C4, BioLegend, cat # 324226), Anti-human CD107a BV750 (1:50, clone H4A3, BioLegend, cat # 328638), Anti-human CD103 BV711 (1:100, clone Ber-ACT8, BioLegend, cat # 350222), Anti-human CD10 BV711 (1:100, clone HI10a, BioLegend, cat # 312226), Anti-human CD45 BV711 (1:100, clone HI30, BioLegend, cat # 304050), Anti-human CD3 BV785 (1:100, clone OKT3, BioLegend, cat # 317330), Anti-human HLA-DR BV785 (1:50, clone L243, BioLegend, cat # 307642), Anti-human PD-1 BV785 (1:50, clone EH12.2H7, BioLegend, cat # 329930), Anti-human CD45 BUV395 (1:100, clone HI30, BD Biosciences, cat # 563792), Anti-human CD16 BUV395 (1:100, clone 3G8, BD Biosciences, cat # 563785), Anti-human CD3 BUV395 (1:100, clone UCHT1, BD Biosciences, cat # 563546), Anti-human CD69 BUV496 (1:50, clone FN50, BD Biosciences, cat # 750214), Anti-human CD16 BUV737 (1:100, clone 3G8, BD Biosciences, cat # 564434). The following antibodies were used for intracellular staining: Anti-human TBET FITC (1:50, clone 4B10, BioLegend, cat # 644812), Anti-human Granzyme B FITC (1:100, clone QA16A02, BioLegend, cat # 372206), Anti-human Granzyme K PE (1:25, clone

GM26E7, BioLegend, cat # 370512), Anti-human FOXP3 PE CF594 (1:25, clone 236 A/E7, BD Biosciences, cat # 563955), Anti-human GATA3 PE CF594 (1:25, clone L50-823, BD Bioscience, cat # 563510), Anti-human Amphiregulin PE Cy 7 (1:25, clone AREG559, Invitrogen, cat # 25-5370-42), Anti-mouse Nur77 APC (1:25, clone REA704, Miltenyi, cat # 130-111-231), Anti-human EOMES APC eF780 (1:25, clone WD1928, eBioscience, cat # 47-4877-42), Anti-human RORgt BV650 (1:50, clone Q21-559, BD Biosciences, cat # 563424), Anti-human Perforin eF450 (1:100, clone dG9, Invitrogen, cat # 48-9994-42). Cells were analyzed on a BD LSR Fortessa flow cytometer. Data were plotted using FlowJo v10.7.1 (TreeStar) and Prism (Graphpad, v9). Due to the low numbers of immune cells detected in healthy kidney, samples from both males and females were analysed together and these data, as displayed, are aggregated for sex. Donor sex and age information is reported in Supplementary Table 2.

**PT cell culture.** Commercially available human primary PTs from 6 donors (3 males and 3 females), Lonza Walkersville Inc) were expanded at passage 4, and studied at passage 5. The main donor characteristics are summarized in Supplementary Table 1. Cells were grown in custom-made Dulbecco's modified Eagle's medium (DMEM) containing 5.55 mM D-glucose, 4mM L-glutamine, and 1 mM sodium pyruvate, and supplemented with 10 ng/mL human EGF, 0.05 M hydrocortisone, 1x of Transferrin/Insulin/Selenium (Invitrogen), 10% v/v dialyzed fetal bovine serum (FBS), 50 g/mL streptomycin, and 50units/mL penicillin, as previously<sup>121,122</sup>. Cells were serum-starved for 24 h prior to collection for gene expression, metabolite measurements, and assessment of metabolic function. For gene expression experiments, cells were washed with PBS, harvested with trypsin, and snap-frozen at  $-80^{\circ}\text{C}$  until further analysis.

**Assessment of metabolic function in human primary PT cells.** Mitochondrial respiration was separately assessed in male and female PTECs by measuring their oxygen consumption rate (OCR) in a Seahorse XFe96 analyzer (Agilent). Glycolysis was also assessed by monitoring the extracellular acidification rate (ECAR). Upon 80-90% confluence, cells were detached with 0.25% trypsin (5 min,  $37^{\circ}\text{C}$ ), counted and seeded in a Seahorse XFe96 Cell Culture Microplate at a density of 15,000 cells/well in 100  $\mu\text{L}$  of DMEM complete media. After adhering for 6 h, PT cells were exposed to serum starvation conditions for 24 h. One hour prior to the metabolic function assay, cells were washed with phenol-free basal media (Agilent) and exposed to 150  $\mu\text{L}$  of assay media, which included 2 mM glutamine and 5.55 mM glucose. During the assay, OCR and ECAR were recorded at baseline and after metabolic stress. To induce metabolic stress, 25  $\mu\text{L}$  of oligomycin, p-trifluoromethoxy carbonyl cyanide phenyl hydrazone (FCCP), 2-deoxyglucose (2-DG), and Rotenone + Antimycin A (Rot + AA) were sequentially injected into the microplate wells. After optimization, the following working concentrations were established for each drug: oligomycin: 1  $\mu\text{M}$ ; FCCP: 0.3  $\mu\text{M}$ , 2-DG: 100 mM; Rot: 1  $\mu\text{M}$ ; AA: 1  $\mu\text{M}$ . Basal respiration, ATP-linked respiration, maximal respiratory capacity, and reserve capacity were assessed by calculating the area under the curve (AUC) from OCR curves (Fig. 3b, c). Basal glycolysis, maximal glycolytic capacity, and glycolytic reserve were determined by calculating the AUC from ECAR curves (Supplementary Fig. 7).

#### Cell metabolite measurements

**Sample preparation.** Male and female primary PTs were grown on 6-well plates and subjected to starvation as described above. The levels of intracellular metabolites for each sex were then determined using liquid chromatography-mass spectrometry. After collecting the supernatant, 1 mL of extraction solvent (80:20 mixture of methanol:water) was added into each well, in order to extract intracellular metabolites. Plates were placed on dry ice. The adherent material was then triturated, collected into Eppendorf tubes, and stored at  $-80^{\circ}\text{C}$ .

Cell lysate collection was followed by 3 freeze-thawing cycles in dry ice (to shift sample temperature between  $-80^{\circ}\text{C}$  and  $-20^{\circ}\text{C}$ ). The insoluble material from each sample was then precipitated by centrifugation at full speed for 5 min. The resulting pellet was dried at room temperature and used for total RNA quantification using the Quant-iT Ribogreen assay (Invitrogen). In turn, the metabolite extract was dried under high purity nitrogen gas (turbovap) and resuspended with appropriate volume of buffer (0.5  $\mu\text{L}$  of LC-MS grade water to 1  $\mu\text{g}$  of RNA) based on total RNA levels. The appropriate volumes of heavy-labelled ( $^{13}\text{C}/^{15}\text{N}$ ) reference metabolites were spiked into each reconstituted sample for quantitation. The heavy-labelled metabolites used as internal reference standards were acquired in as a metabolite extract from yeast that had been 99% labelled with  $^{13}\text{C}$ -glucose and  $^{15}\text{N}$ -ammonia. To determine background metabolite signals, a mock plate without cells and equal volume of media was processed in parallel to the study plates.

**Liquid chromatography-mass spectrometry (LC-MS).** Cellular metabolites were measured by injecting 2  $\mu\text{L}$  of sample in full scan MS1 mode using an Agilent 6550 qToF mass spectrometer coupled to an Agilent 1290 binary pump UPLC system. Most polar metabolite analytes presented here were measured using an Agilent ZORBAX ExtendC18 1.8  $\mu\text{m}$ , 2.1 mm $\times$ 150 mm reverse phase chromatography using tributylamine as an ion pairing agent as previously described<sup>123</sup>. The Agilent 6550 qToF was fitted with a dual AJS ESI source and an iFunnel with a gas temperature set to  $150^{\circ}\text{C}$  at 14 L/min and 45 psig. Sheath gas temperature was set to  $325^{\circ}\text{C}$  at 12 L/min. Capillary and nozzle voltages were set to 2000V. Funnel conditions were changed from default to  $-30\text{V}$  DC, high pressure funnel drop  $-100\text{V}$  and RF voltage of 110 V, low pressure funnel drop  $-50\text{V}$  and RF voltage of 60 V. Metabolite annotation in full scan data was achieved by matching exact mass and retention time to an in-house database. The retention time and exact mass database were prepared by analyzing a collection of neat standards using the chromatographic method described above and confirming retention times by MS/MS fragmentation of neat standards.

**Metabolite data analysis.** Metabolite raw data was extracted directly from .d folders and integrated in profile mode using an R-based software package developed by the Rosebrock Lab; ChromXtractorPro (personal correspondence K. Lavery and A. Rosebrock, adam.rosebrock@stonybrook.edu). The metabolites whose intensity in all the study samples fell at or below their intensity in the blank (consisting of resuspension buffer only) were excluded from further analyses. Next, the integrated light (L) intensity of each metabolite was normalized to the intensity of its internal heavy (H) standard. The L/H ratio minimized the potential stochastic variation in the signal produced by the instrument due to changes in humidity and/or temperature, enabling the relative quantitation and comparative analysis of each metabolite. The analysis enabled the detection of 158 intracellular metabolites<sup>124</sup>. Data corresponding to the intracellular levels of NAD,  $\beta$ -nicotinamide mononucleotide, ATP, GTP, ITP, UTP were interrogated.

### Gene expression validation studies

RNA was extracted from the cell pellets of human primary male and female PT cells using the RNAeasy Mini Kit (Qiagen). After quantifying RNA concentration in a Nanodrop instrument (Thermo), 300–700 ng of RNA were retrotranscribed to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Male and female PTs had each been grown and serum-starved as above. In these cells, gene levels of *KDMSD*, *UTY*, *EIPIAY*, *EIPIAX*, *DDX3X*, *MTIF*, *MTIG*, and *MTIH* were measured by real-time quantitative PCR using a Power SYBR<sup>®</sup> Green PCR Master Mix reagent (Applied Biosystems) and normalized to RPL31. The fluorescent signal was measured in a LightCycler<sup>®</sup> 480

Instrument II (Roche). All primer sequences employed in this study are summarized in Supplementary Table 3.

### Quantification and statistical analysis

Statistical tests were conducted within R and using GraphPad Prism 9.1.0 software. For all comparisons, normality was determined using a Shapiro-Wilk test. Group-to-group differences were assessed using two-tailed unpaired T tests for variables following a normal distribution, and Mann-Whitney tests for variables with a non-parametric distribution. All p values below 0.05 were considered significant. Significance level for each test is indicated in the figures. For each experiment, n is reported in the figure legends and represents the number of samples. Where appropriate, disaggregated numbers for each sex are also reported.

### Figure preparation

All figures were prepared using Adobe Illustrator 2020. Kidney structure (Fig. 1c), schematic in 2b and Supplementary Fig. 16a were constructed using Biorender: <https://biorender.com> (licence to use provided).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The raw sequencing data generated in this study have been submitted to NCBI Gene Expression Omnibus under the accession number [GSE202109](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE202109). The final dataset (containing processed files) is also available online<sup>125</sup> via: <https://cells.ucsc.edu/?ds=living-donor-kidney>. Source data and supplementary information are provided with this paper. Previously published sequencing data were accessed as follows: Stewart and Ferdinand et al.<sup>3</sup>. (<https://data.humancellatlas.org/explore/projects/abe1a013-af7a-45ed-8c26-f3793c24a1f4>); Muto et al.<sup>14</sup>, [GSE151302](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE151302); Zimmerman et al.<sup>52</sup>, [GSE128993](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128993); Argüello et al.<sup>53</sup>, [GSE159913](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE159913); Wang et al.<sup>55</sup>, [GSE148665](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148665); 10X Genomics Datasets ([https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.0.2/5k\\_pbmc\\_v3](https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.0.2/5k_pbmc_v3))<sup>56</sup> and ([https://support.10xgenomics.com/single-cell-gene-expression/datasets/2.1.0/t\\_4k](https://support.10xgenomics.com/single-cell-gene-expression/datasets/2.1.0/t_4k))<sup>107</sup>; and Bernink et al.<sup>106</sup>, [GSE114396](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114396). Additional public data repositories used for our analysis include CHEA3<sup>108</sup> (<https://maayanlab.cloud/chea3/>) and the Bader lab repository of genesets: ([http://download.baderlab.org/EM\\_Genesets/](http://download.baderlab.org/EM_Genesets/)). Source data are provided with this paper.

### Code availability

Scripts used for processing the raw data are deposited on Zenodo (<https://doi.org/10.5281/zenodo.6633564>).

### References

1. Kurts, C., Panzer, U., Anders, H. J. & Rees, A. J. The immune system and kidney disease: basic concepts and clinical implications. *Nat. Rev. Immunol.* **13**, 738–753 (2013).
2. Rogers, N. M., Ferenbach, D. A., Isenberg, J. S., Thomson, A. W. & Hughes, J. Dendritic cells and macrophages in the kidney: a spectrum of good and evil. *Nat. Rev. Nephrol.* **10**, 625–643 (2014).
3. Stewart, B. J. et al. Spatiotemporal immune zonation of the human kidney. *Science* **365**, 1461–1466 (2019).
4. Saat, T. C. et al. A comparison of inflammatory, cytoprotective and injury gene expression profiles in kidneys from brain death and cardiac death donors. *Transplantation* **98**, 15–21 (2014).
5. Damman, J. et al. Hypoxia and Complement-and-Coagulation Pathways in the Deceased Organ Donor as the Major Target for Intervention to Improve Renal Allograft Outcome. *Transplantation* **99**, 1293–1300 (2015).

6. Clotet, S. et al. Stable Isotope Labeling with Amino Acids (SILAC)-Based Proteomics of Primary Human Kidney Cells Reveals a Novel Link between Male Sex Hormones and Impaired Energy Metabolism in Diabetic Kidney Disease. *Mol. Cell Proteom.* **16**, 368–385 (2017).
7. Aufhauser, D. D. Jr. et al. Improved renal ischemia tolerance in females influences kidney transplantation outcomes. *J. Clin. Invest.* **126**, 1968–1977 (2016).
8. Neugarten, J. & Golestaneh, L. Female sex reduces the risk of hospital-associated acute kidney injury: a meta-analysis. *BMC Nephrol.* **19**, 314 (2018).
9. Clark, J. Z. et al. Representation and relative abundance of cell-type selective markers in whole-kidney RNA-Seq data. *Kidney Int* **95**, 787–796 (2019).
10. Wu, H. et al. Comparative Analysis and Refinement of Human PSC-Derived Kidney Organoid Differentiation with Single-Cell Transcriptomics. *Cell Stem Cell* **23**, 869–881.e868 (2018).
11. Smeets, B. et al. Proximal tubular cells contain a phenotypically distinct, scattered cell population involved in tubular regeneration. *J. Pathol.* **229**, 645–659 (2013).
12. Berger, K. et al. Origin of regenerating tubular cells after acute kidney injury. *Proc. Natl Acad. Sci. USA* **111**, 1533–1538 (2014).
13. Kirita, Y., Wu, H., Uchimura, K., Wilson, P. C. & Humphreys, B. D. Cell profiling of mouse acute kidney injury reveals conserved cellular responses to injury. *Proc. Natl Acad. Sci. USA* **117**, 15874–15883 (2020).
14. Muto, Y. et al. Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nat. Commun.* **12**, 2190 (2021).
15. Ahmed, S. M., Luo, L., Namani, A., Wang, X. J. & Tang, X. Nrf2 signaling pathway: Pivotal roles in inflammation. *Biochim Biophys. Acta Mol. Basis Dis.* **1863**, 585–597 (2017).
16. Yoshida, T. et al. ATF3 protects against renal ischemia-reperfusion injury. *J. Am. Soc. Nephrol.* **19**, 217–224 (2008).
17. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
18. Milatz, S. et al. Mosaic expression of claudins in thick ascending limbs of Henle results in spatial separation of paracellular Na<sup>+</sup> and Mg<sup>2+</sup> transport. *Proc. Natl Acad. Sci. USA* **114**, E219–E227 (2017).
19. Rocha, S. F. et al. Esm1 modulates endothelial tip cell behavior and vascular permeability by enhancing VEGF bioavailability. *Circ. Res* **115**, 581–590 (2014).
20. Barry, D. M. et al. Molecular determinants of nephron vascular specialization in the kidney. *Nat. Commun.* **10**, 5705 (2019).
21. Liburd, S. T., Jr., Shi, A. A., Pober, J. S. & Tietjen, G. T. Wanted: An endothelial cell targeting atlas for nanotherapeutic delivery in allograft organs. *Am. J. Transplant.* **22**, 1754–1759 (2022).
22. Oliva, M. et al. The impact of sex on gene expression across human tissues. *Science* **369**, p.eaba3066 (2020).
23. Liao, J. et al. Single-cell RNA sequencing of human kidney. *Sci. Data* **7**, 4 (2020).
24. Wu, H., Lai, C. F., Chang-Panesso, M. & Humphreys, B. D. Proximal Tubule Translational Profiling during Kidney Fibrosis Reveals Proinflammatory and Long Noncoding RNA Expression Patterns with Sexual Dimorphism. *J. Am. Soc. Nephrol.* **31**, 23–38 (2020).
25. Vinas, J. L. et al. Sex diversity in proximal tubule and endothelial gene expression in mice with ischemic acute kidney injury. *Clin. Sci. (Lond.)* **134**, 1887–1909 (2020).
26. Ransick, A. et al. Single-Cell Profiling Reveals Sex, Lineage, and Regional Diversity in the Mouse Kidney. *Dev. Cell* **51**, 399–413 e397 (2019).
27. Planells, B., Gomez-Redondo, I., Pericuesta, E., Lonergan, P. & Gutierrez-Adan, A. Differential isoform expression and alternative splicing in sex determination in mice. *BMC Genomics* **20**, 202 (2019).
28. Clement, T. M., Anway, M. D., Uzumcu, M. & Skinner, M. K. Regulation of the gonadal transcriptome during sex determination and testis morphogenesis: comparative candidate genes. *Reproduction* **134**, 455–472 (2007).
29. Doring, A. et al. SLC2A9 influences uric acid concentrations with pronounced sex-specific effects. *Nat. Genet* **40**, 430–436 (2008).
30. Khan, S. et al. Fatty acid transport protein-2 regulates glycemic control and diabetic kidney disease progression. *JCI Insight* **5**, (2020).
31. Reichold, M. et al. Glycine Amidinotransferase (GATM), Renal Fanconi Syndrome, and Kidney Failure. *J. Am. Soc. Nephrol.* **29**, 1849–1858 (2018).
32. Ruttkay-Nedecky, B. et al. The role of metallothionein in oxidative stress. *Int J. Mol. Sci.* **14**, 6044–6066 (2013).
33. Woodard, L. E. et al. Metabolic consequences of cystinuria. *BMC Nephrol.* **20**, 227 (2019).
34. Park, K. M., Kim, J. I., Ahn, Y., Bonventre, A. J. & Bonventre, J. V. Testosterone is responsible for enhanced susceptibility of males to ischemic renal injury. *J. Biol. Chem.* **279**, 52282–52292 (2004).
35. O’Flanagan, C. H. et al. Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. *Genome Biol.* **20**, 210 (2019).
36. Lake, B. B. et al. A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Sci. Rep.* **7**, 6031 (2017).
37. Tukiainen, T. et al. Landscape of X chromosome inactivation across human tissues. *Nature* **550**, 244–248 (2017).
38. Smolina, N., Bruton, J., Kostareva, A. & Sejersen, T. Assaying Mitochondrial Respiration as an Indicator of Cellular Metabolism and Fitness. *Methods Mol. Biol.* **1601**, 79–87 (2017).
39. Chaudhry, R. & Varacallo, M. Biochemistry, Glycolysis. in *StatPearls* (Treasure Island (FL), 2021).
40. Jose, C., Bellance, N. & Rossignol, R. Choosing between glycolysis and oxidative phosphorylation: a tumor’s dilemma? *Biochim Biophys. Acta* **1807**, 552–561 (2011).
41. Canto, C., Menzies, K. J. & Auwerx, J. NAD(+) Metabolism and the Control of Energy Homeostasis: A Balancing Act between Mitochondria and the Nucleus. *Cell Metab.* **22**, 31–53 (2015).
42. Park, C. O. & Kupper, T. S. The emerging role of resident memory T cells in protective immunity and inflammatory disease. *Nat. Med.* **21**, 688–697 (2015).
43. Mogilenko, D. A. et al. Comprehensive Profiling of an Aging Immune System Reveals Clonal GZMK(+) CD8(+) T Cells as Conserved Hallmark of Inflammaging. *Immunity* **54**, 99–115.e112 (2021).
44. Nehar-Belaid, D. et al. Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* **21**, 1094–1106 (2020).
45. Sakai, R. et al. Kidney GATA3(+) regulatory T cells play roles in the convalescence stage after antibody-mediated renal injury. *Cell Mol. Immunol.* **18**, 1249–1261 (2021).
46. Ashouri, J. F. & Weiss, A. Endogenous Nur77 Is a Specific Indicator of Antigen Receptor Signaling in Human T and B Cells. *J. Immunol.* **198**, 657–668 (2017).
47. Epelman, S., Lavine, K. J. & Randolph, G. J. Origin and functions of tissue macrophages. *Immunity* **41**, 21–35 (2014).
48. Gottschalk, C. & Kurts, C. The Debate about Dendritic Cells and Macrophages in the Kidney. *Front Immunol.* **6**, 435 (2015).
49. Ramachandran, P. et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512–518 (2019).
50. Liu, F. et al. Distinct fate, dynamics and niches of renal macrophages of bone marrow or embryonic origins. *Nat. Commun.* **11**, 2280 (2020).

51. Yue, S. et al. Prolonged Ischemia Triggers Necrotic Depletion of Tissue-Resident Macrophages To Facilitate Inflammatory Immune Activation in Liver Ischemia Reperfusion Injury. *J. Immunol.* **198**, 3588–3595 (2017).
52. Zimmerman, K. A. et al. Single-Cell RNA Sequencing Identifies Candidate Renal Resident Macrophage Gene Expression Signatures across Species. *J. Am. Soc. Nephrol.* **30**, 767–781 (2019).
53. Arguello, R. J. et al. SCENITH: A Flow Cytometry-Based Method to Functionally Profile Energy Metabolism with Single-Cell Resolution. *Cell Metab.* **32**, 1063–1075 e1067 (2020).
54. Dick, S. A. et al. Three tissue resident macrophage subsets coexist across organs with conserved origins and life cycles. *Sci. Immunol.* **7**, eabf7777 (2022).
55. Wang, X. et al. BREM-SC: a bayesian random effects mixture model for joint clustering single cell multi-omics data. *Nucleic Acids Res* **48**, 5814–5824 (2020).
56. 10X Genomics, [https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.0.2/5k\\_pbmc\\_v3](https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.0.2/5k_pbmc_v3). (2019).
57. Mackay, L. K. et al. Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science* **352**, 459–463 (2016).
58. Cibrian, D. & Sanchez-Madrid, F. CD69: from activation marker to metabolic gatekeeper. *Eur. J. Immunol.* **47**, 946–953 (2017).
59. Clark, R. A. Resident memory T cells in human health and disease. *Sci. Transl. Med* **7**, 269rv261 (2015).
60. Fu, X. et al. Human natural killer cells expressing the memory-associated marker CD45RO from tuberculous pleurisy respond more strongly and rapidly than CD45RO- natural killer cells following stimulation with interleukin-12. *Immunology* **134**, 41–49 (2011).
61. Chakravorty, S. J., Howie, A. J., Cockwell, P., Adu, D. & Savage, C. O. T lymphocyte adhesion mechanisms within inflamed human kidney: studies with a Stamper-Woodruff assay. *Am. J. Pathol.* **154**, 503–514 (1999).
62. Kim, G. H. Renal effects of prostaglandins and cyclooxygenase-2 inhibitors. *Electrolyte Blood Press* **6**, 35–41 (2008).
63. Boniface, K. et al. Prostaglandin E2 regulates Th17 cell differentiation and function through cyclic AMP and EP2/EP4 receptor signaling. *J. Exp. Med* **206**, 535–548 (2009).
64. Si, H. et al. Human and murine kidneys show gender- and species-specific gene expression differences in response to injury. *PLoS One* **4**, e4802 (2009).
65. Tower, J., Pomatto, L. C. D. & Davies, K. J. A. Sex differences in the response to oxidative and proteolytic stress. *Redox Biol.* **31**, 101488 (2020).
66. Pan, J. S. & Sheikh-Hamad, D. Mitochondrial dysfunction in acute kidney injury and sex-specific implications. *Med. Res. Arch.* **7**, (2019).
67. Schanz, M. et al. Renal effects of metallothionein induction by zinc in vitro and in vivo. *BMC Nephrol.* **18**, 91 (2017).
68. Tachibana, H. et al. Metallothionein deficiency exacerbates diabetic nephropathy in streptozotocin-induced diabetic mice. *Am. J. Physiol. Ren. Physiol.* **306**, F105–F115 (2014).
69. Wu, H. et al. Metallothionein deletion exacerbates intermittent hypoxia-induced renal injury in mice. *Toxicol. Lett.* **232**, 340–348 (2015).
70. Park, L. et al. Tat-enhanced delivery of metallothionein can partially prevent the development of diabetes. *Free Radic. Biol. Med* **51**, 1666–1674 (2011).
71. Wu, G., Fang, Y. Z., Yang, S., Lupton, J. R. & Turner, N. D. Glutathione metabolism and its implications for health. *J. Nutr.* **134**, 489–492 (2004).
72. Wang, L., Ahn, Y. J. & Asmis, R. Sexual dimorphism in glutathione metabolism and glutathione-dependent responses. *Redox Biol.* **31**, 101410 (2020).
73. Saran, R. et al. US Renal Data System 2016 Annual Data Report: Epidemiology of Kidney Disease in the United States. *Am. J. Kidney Dis.: Off. J. Natl Kidney Found.* **69**, A7–A8 (2017).
74. Clotet-Freixas, S. et al. Sex dimorphism in ANGII-mediated crosstalk between ACE2 and ACE in diabetic nephropathy. *Lab Invest* **98**, 1237–1249 (2018).
75. Sharma, M. et al. Extracellular granzyme K mediates endothelial activation through the cleavage of protease-activated receptor-1. *FEBS J.* **283**, 1734–1747 (2016).
76. Wensink, A. C. et al. Granzyme K synergistically potentiates LPS-induced cytokine responses in human monocytes. *Proc. Natl Acad. Sci. USA* **111**, 5974–5979 (2014).
77. Li, S. et al. Extracellular Granzyme K Modulates Angiogenesis by Regulating Soluble VEGFR1 Release From Endothelial Cells. *Front Oncol.* **11**, 681967 (2021).
78. Turner, J. E., Paust, H. J., Steinmetz, O. M. & Panzer, U. The Th17 immune response in renal inflammation. *Kidney Int* **77**, 1070–1075 (2010).
79. Chung, B. H., Yang, C. W. & Cho, M. L. Clinical significance of Th17 cells in kidney transplantation. *Korean J. Intern Med.* **33**, 860–866 (2018).
80. Aran, D. et al. Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nat. Commun.* **8**, 1077 (2017).
81. 10XGenomics. What fraction of mRNA transcripts are captured per cell?, Vol. 2022 (<https://kb.10xgenomics.com/hc/en-us/articles/360001539051-What-fraction-of-mRNA-transcripts-are-captured-per-cell->, 2022).
82. Menon, R., et al. Single cell transcriptomics identifies focal segmental glomerulosclerosis remission endothelial biomarker. *JCI Insight* **5**, (2020).
83. Yamawaki, T. M. et al. Systematic comparison of high-throughput single-cell RNA-seq methods for immune cell profiling. *BMC Genomics* **22**, 66 (2021).
84. Chen, L., Chou, C. L. & Knepper, M. A. A Comprehensive Map of mRNAs and Their Isoforms across All 14 Renal Tubule Segments of Mouse. *J. Am. Soc. Nephrol.* (2021).
85. Park, J. et al. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763 (2018).
86. Zhang, Y., et al. Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc. Natl. Acad. Sci. USA* **118**, (2021).
87. 10XGenomics. [https://assets.ctfassets.net/an68im79xiti/4tjk4KvXzTWgTs8f3tvUjq/2259891d68c53693e753e1b45e42de2d/CG000183\\_ChromiumSingleCell3\\_v3\\_UG\\_Rev\\_C.pdf](https://assets.ctfassets.net/an68im79xiti/4tjk4KvXzTWgTs8f3tvUjq/2259891d68c53693e753e1b45e42de2d/CG000183_ChromiumSingleCell3_v3_UG_Rev_C.pdf). (2021).
88. Slyper, M. et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat. Med.* **26**, 792–802 (2020).
89. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, (2020).
90. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**, 329–337.e324 (2019).
91. Pagliarini, D. J. et al. A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134**, 112–123 (2008).
92. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
93. Andrews, T. S. & Hemberg, M. M3Drop: dropout-based feature selection for scRNASeq. *Bioinformatics* **35**, 2865–2867 (2019).
94. Stuart, T. et al. Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e1821 (2019).
95. Becht, E., et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* (2018).



96. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
97. Dimitrov, D., et al. Cell-cell Communication Inference from Single-cell RNA-Seq Data: a Comparison of Methods and Resources. *Research Square* (2021).
98. Turei, D., Korcsmaros, T. & Saez-Rodriguez, J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat. Methods* **13**, 966–967 (2016).
99. Cabello-Aguilar, S. et al. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res* **48**, e55 (2020).
100. Wang, Y. et al. iTALK: an R Package to Characterize and Illustrate Intercellular Communication. *BioRxiv* (2019).
101. Hou, R., Denisenko, E., Ong, H. T., Ramilowski, J. A. & Forrest, A. R. R. Predicting cell-to-cell communication networks using NATMI. *Nat. Commun.* **11**, 5011 (2020).
102. Raredon, M. S. B. et al. Connectome: computation and visualization of cell-cell signaling topologies in single-cell systems data. *bioRxiv* (2021).
103. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
104. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
105. Alquicira-Hernandez, J., Sathe, A., Ji, H. P., Nguyen, Q. & Powell, J. E. scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome Biol.* **20**, 264 (2019).
106. Bernink, J. H. et al. c-Kit-positive ILC2s exhibit an ILC3-like signature that may contribute to IL-17-mediated pathologies. *Nat. Immunol.* **20**, 992–1003 (2019).
107. 10XGenomics. [https://support.10xgenomics.com/single-cell-gene-expression/datasets/2.1.0/t\\_4k](https://support.10xgenomics.com/single-cell-gene-expression/datasets/2.1.0/t_4k). (2017).
108. Keenan, A. B. et al. ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res* **47**, W212–W224 (2019).
109. Andreatta, M. & Carmona, S. J. UCell: Robust and scalable single-cell gene signature scoring. *Comput Struct. Biotechnol. J.* **19**, 3796–3798 (2021).
110. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
111. Blighe K, R. S., Lewis M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. (<https://github.com/kevinblighe/EnhancedVolcano>., 2020).
112. Tan, Y. & Cahan, P. SingleCellNet: A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Across Species. *Cell Syst.* **9**, 207–213.e202 (2019).
113. Wang, C. X., Zhang, L. & Wang, B. One Cell At a Time (OCAT): a unified framework to integrate and analyze single-cell RNA-seq data. *Genome Biol.* **23**, 102 (2022).
114. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
115. Rohart, F., Gautier, B., Singh, A. & Le Cao, K. A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol.* **13**, e1005752 (2017).
116. Sonesson, C. & Robinson, M. D. Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* **15**, 255–261 (2018).
117. Wang, T., Li, B., Nelson, C. E. & Nabavi, S. Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinforma.* **20**, 40 (2019).
118. Zimmerman, K. D., Espeland, M. A. & Langefeld, C. D. A practical solution to pseudoreplication bias in single-cell studies. *Nat. Commun.* **12**, 738 (2021).
119. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
120. Reimand, J. et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and Enrichment-Map. *Nat. Protoc.* **14**, 482–517 (2019).
121. Konvalinka, A. et al. Determination of an angiotensin II-regulated proteome in primary human kidney cells by stable isotope labeling of amino acids in cell culture (SILAC). *J. Biol. Chem.* **288**, 24834–24847 (2013).
122. Clotet-Freixas, S. et al. Extracellular Matrix Injury of Kidney Allografts in Antibody-Mediated Rejection: A Proteomics Study. *J. Am. Soc. Nephrol.* **31**, 2705–2724 (2020).
123. Wan, L. C. et al. Reconstitution and characterization of eukaryotic N6-threonylcarbamoylation of tRNA using a minimal enzyme system. *Nucleic Acids Res* **41**, 6332–6346 (2013).
124. Clotet-Freixas, S. et al. Cell Sex and Sex Hormones Modulate Kidney Glucose and Glutamine Metabolism in Health and Diabetes. *BioRxiv* (2021).
125. Speir, M. L. et al. UCSC Cell Browser: Visualize Your Single-Cell Data. *Bioinformatics* (2021).

## Acknowledgements

First and foremost, we would like to thank the kidney transplant patients who made this work possible. We would also like to thank the nurses, physicians and surgeons at Toronto General Hospital and the Ajmera Transplant Centre Biobank for efforts to obtain tissue samples, and acknowledge technical support provided by the Princess Margaret Genomics Centre, particularly Troy Ketela and Julissa Tsao, and the Princess Margaret Cancer Centre flow cytometry core. We thank the team at UCSC Cell Browser (Brittney Wick, Matthew Speir, and Maximilian Haeussler) for their assistance in packaging and hosting the online cell browser. Figures 1c, 2b and Supplementary Fig. 16a were created with BioRender.com. This work was supported by funding from the Canadian Institutes for Health Research (CIHR) grant 168960 and support from the UHN Foundation through seed funding and biobank infrastructure support. CMM was supported by the Menkes Family Fellowship and an Ajmera Transplant Centre fellowship. JMM was supported by Ontario Graduate Scholarships, QE II/Aventis Pasteur Graduate Scholarship and a Peterborough K.M. Hunter Foundation Scholarship. SCF was supported by the Kidney Research Scientist Core Education and National Training (KRESCENT) program grant 2019KP-PDF637713. SQC was supported by the Medicine by Design program (Canada First Research Excellence Fund) and Canada Foundation for Innovation (CFI) grant 38308, and holds a Tier 2 Canada Research Chair in Tissue-Specific Immune Tolerance. AK was supported by CFI grant 37205, Kidney Foundation of Canada (KFOC) Biomedical Research grant KFOC160010, and Kidney Research Scientist Core Education and National Training (KRESCENT) program grants CIHR148204, KRES160004, and KRES160005, as well as funding from UHN Foundation (TGTWF 1617-464; TGTWF MKFTR 1718-1268). AF was supported by the Princess Margaret Foundation. GDB was supported by NRRB (U.S. National Institutes of Health, National Center for Research Resources grant numbers P41 GM103504) and by the US National Human Genome Research Institute under award number HG009979.

## Author contributions

C.M.M and J.M.M. contributed equally. A.K. and S.Q.C. jointly supervised the work. C.M.M., J.M.M., A.K., and S.Q.C. designed and implemented the study. A.K. and S.S. established the ATC biobank and the infrastructure required for sample retrieval. S.S., C.M.M., J.M.M., S.Z., K.A.L., A.F. and A.K. liaised with patient care teams to obtain tissues for study. C.M.M., J.M.M., J.A.M., S.C.F., J.A., A.K. and S.Q.C. developed tissue dissociation protocols with input from L.L. and S.A.M. and performed experiments. S.C.F., O.Z., H.R., R.A., contributed to PT validation

experiments. C.M.M., J.M.M. and L.Z. implemented the bioinformatic pipeline with advice and assistance from M.K., D.P., B.W., G.D.B., A.K. and S.Q.C. C.M.M., J.M.M., S.A.M., A.K. and S.Q.C. provided tissue and immune compartment expertise and annotated cell types. B.W., S.A.M., G.D.B., A.K., and S.Q.C. supervised the work. C.M.M., J.M.M., A.K., and S.Q.C. wrote the manuscript, which all authors reviewed and edited.

### Competing interests

GDB is an advisor for Deep Genomics and is on the Scientific Advisory Board of Adela Bio. The remaining authors have no competing interests to declare.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-35297-z>.

**Correspondence** and requests for materials should be addressed to Ana Konvalinka or Sarah Q. Crome.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2022

<sup>1</sup>Toronto General Hospital Research Institute, University Health Network, Toronto, ON, Canada. <sup>2</sup>Ajmera Transplant Centre, University Health Network, Toronto, ON, Canada. <sup>3</sup>Department of Medicine, Division of Nephrology, University Health Network, Toronto, ON, Canada. <sup>4</sup>Department of Immunology, University of Toronto, Toronto, ON, Canada. <sup>5</sup>Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada. <sup>6</sup>Peter Munk Cardiac Centre, University Health Network, Toronto, ON, Canada. <sup>7</sup>Vector Institute, Toronto, ON, Canada. <sup>8</sup>Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada. <sup>9</sup>The Donnelly Centre, University of Toronto, Toronto, ON, Canada. <sup>10</sup>Division of Urology, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada. <sup>11</sup>Institute of Medical Science, University of Toronto, Toronto, ON, Canada. <sup>12</sup>Department of Computer Science, University of Toronto, Toronto, ON, Canada. <sup>13</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada. <sup>14</sup>Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada. <sup>15</sup>The Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON, Canada. <sup>16</sup>These authors contributed equally: Caitriona M. McEvoy, Julia M. Murphy. <sup>17</sup>These authors jointly supervised this work: Ana Konvalinka, Sarah Q. Crome. ✉ e-mail: [Ana.Konvalinka@uhn.ca](mailto:Ana.Konvalinka@uhn.ca); [sarah.crome@utoronto.ca](mailto:sarah.crome@utoronto.ca)