Extended Data

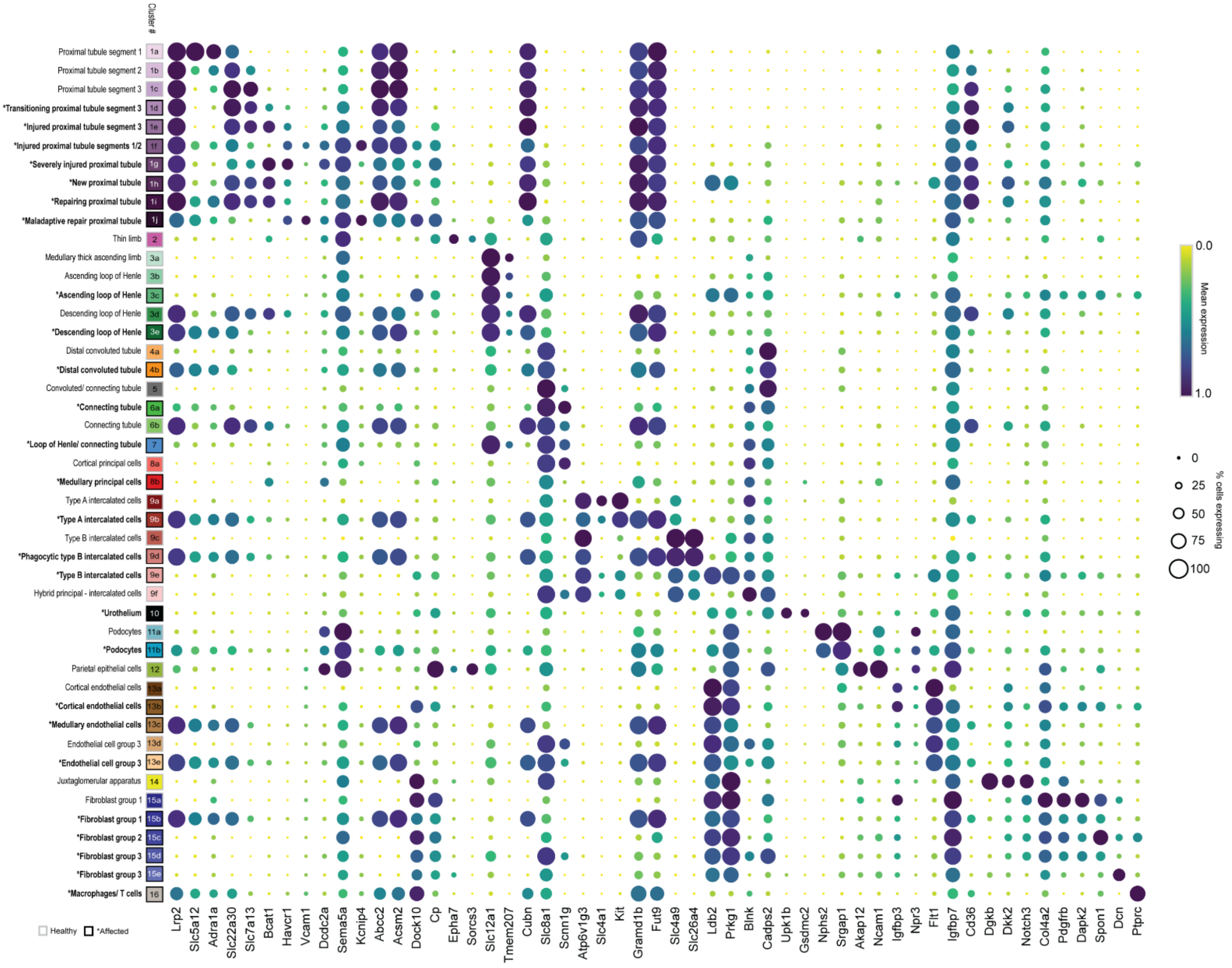
Single-nucleus transcriptomics illuminates sex differences during murine *Escherichia coli* pyelonephritis

Teri N. Hreha,[1]* Abigail L. Manson,[4]* Christina A. Collins,[1] Haojia Wu,[2] Christophe Georgescu,[5] Benjamin D. Humphreys,[2] Ashlee M. Earl,[4†] and David A. Hunstad[1,3†]

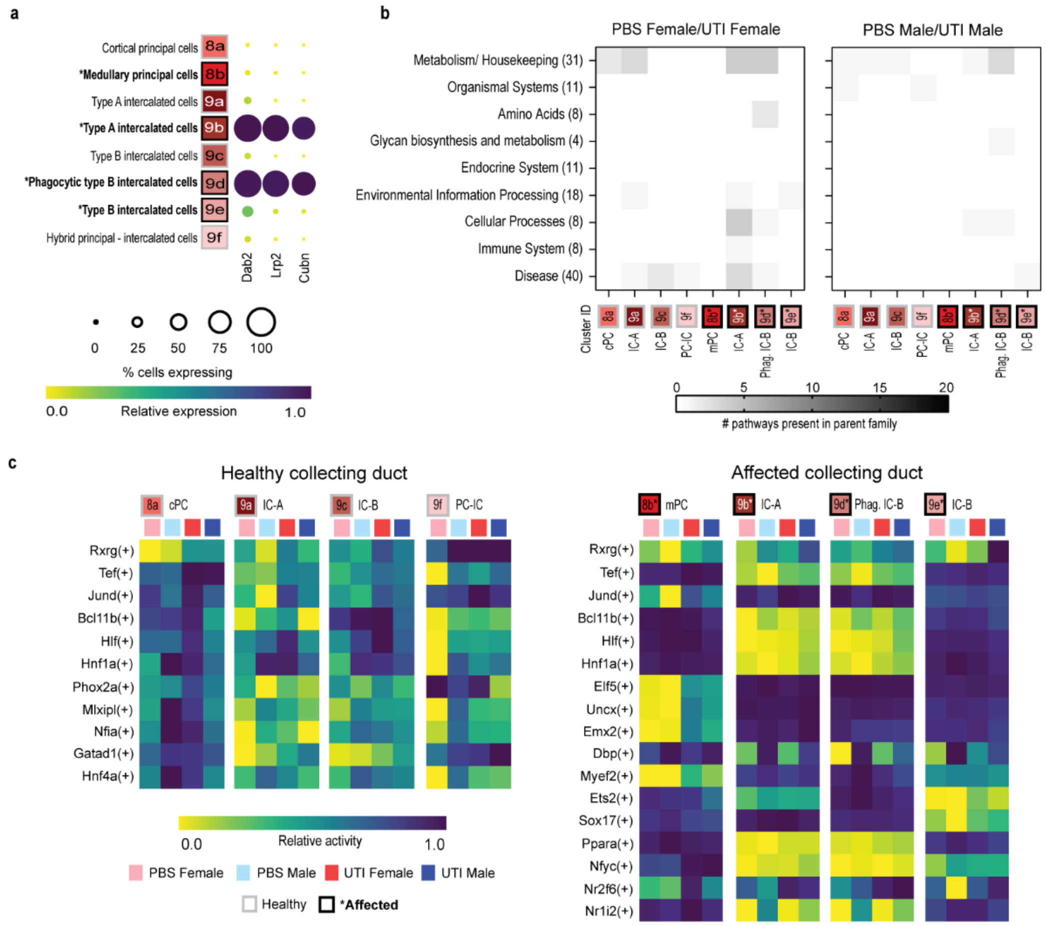
Departments of [1]Pediatrics, [2]Medicine, and [3]Molecular Microbiology, Washington University School of Medicine, St. Louis, MO, USA; [4]Infectious Disease and Microbiome Program and [5]Genomics Platform, Broad Institute of MIT and Harvard, Cambridge, MA, USA

*These authors contributed equally to the work

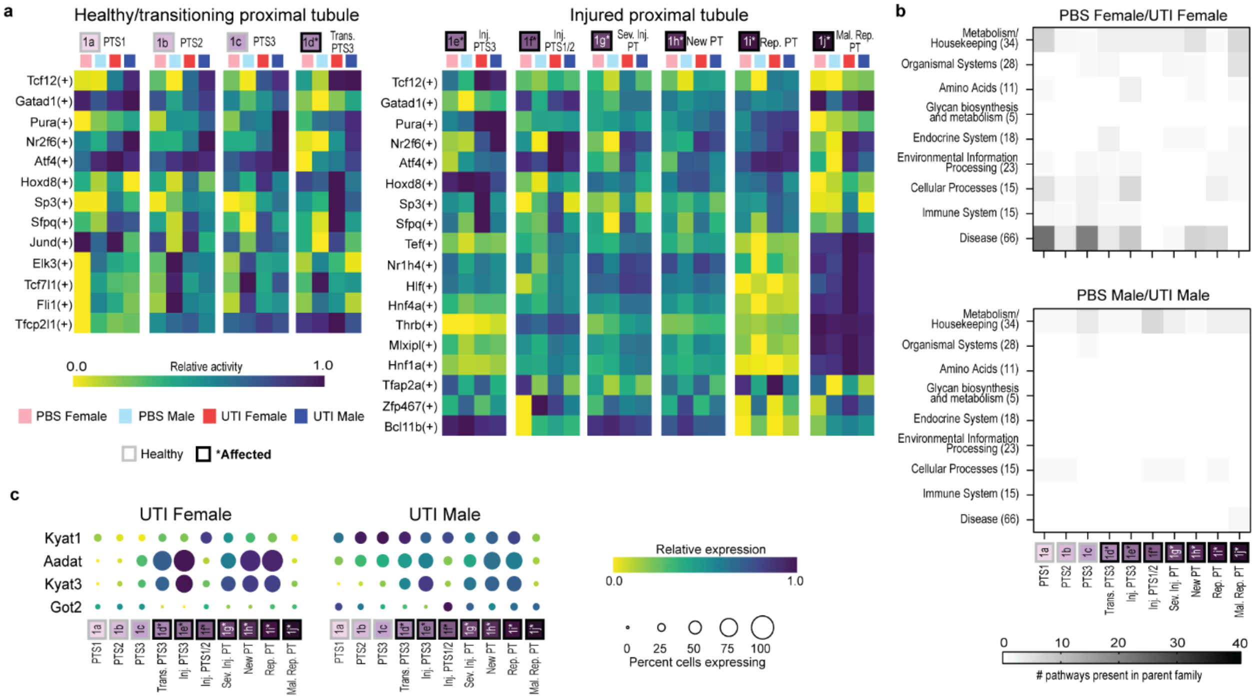
†These authors contributed equally to the work



Extended Data Figure 1. Marker genes for set of 46 cell type clusters, containing healthy clusters (outlined in light grey boxes) and affected clusters (bolded and outlined in black boxes). Color indicates mean expression, and circle size represents the percent of cells expressing this gene.



Extended Data Figure 2. a, Genes involved in phagocytosis are expressed in both affected Type A and a subset of affected Type B intercalated cells. Color indicates mean expression, and circle size represents the percent of cells expressing each gene. **b,** Heatmaps depicting the number of KEGG pathways within each parent family that were significantly enriched within the 200 most upregulated DEGs when comparing PBS mice to their UTI counterparts. **c,** Relative activity for the most downregulated transcription factor regulons in the collecting duct during UTI, as compared to PBS conditions. The left and right heatmaps show grouped healthy and affected cell types, respectively. For each grouping, the superset of the five most downregulated regulons in each cell type are plotted. Healthy clusters are outlined in light grey boxes, and affected clusters are bolded and outlined in black boxes.



Extended Data Figure 3. a, Relative activity for the most downregulated transcription factor regulons in the proximal tubule during UTI, as compared to PBS conditions. The left and right heatmaps show grouped healthy and affected cell types, respectively. For each grouping, the superset of the five most downregulated regulons in each cell type are plotted. Healthy clusters are outlined in light grey boxes, and affected clusters are bolded and outlined in black boxes. **b,** Heatmaps depicting the number of KEGG pathways within each parent family that were significantly enriched within the 200 most upregulated DEGs when comparing PBS mice to their UTI counterparts. **c,** Gene expression plots of kynurenine aminotransferases I (*Kyat1*), II (*Aadat*), III (*Kyat3*), and IV (*Got2*) in the proximal tubule of UTI female and male mice.

Extended Datasets provided online **https://figshare.com/s/f72794257c028472bfa8**

Extended Dataset 1. Differences between conditions in pySCENIC data in Male, Female, and Andro mice.

Extended Dataset 2. KEGG pathways significantly enriched in the 200 most upregulated DEGs in Male, Female, and Andro mice.

Extended Dataset 3. GO pathways significantly enriched in the 200 most upregulated DEGs in each cluster, refined to pathways present in at least 3 healthy or affected clusters.

Extended Dataset 4. KEGG pathways significantly enriched in the 200 most upregulated DEGs comparing UTI and PBS conditions in the same sex, or between male and female PBS-inoculated mice, in the collecting duct.

Extended Dataset 5. Differences between pySCENIC data in the collecting duct, comparing UTI versus PBS conditions, along with activity of the most and least active TFs in UTI by condition.

Extended Dataset 6. KEGG pathways significantly enriched in the 200 most upregulated DEGs comparing UTI and PBS conditions in the same sex, or between male and female PBS-inoculated mice, in the proximal tubule.

Extended Dataset 7. Differences between pySCENIC data in the proximal tubule, comparing UTI versus PBS conditions, along with activity of the most and least active TFs in UTI by condition.

Extended Dataset 8. CellPhone DB output by condition, from cell type clustering with both 46 and 16 clusters.

Extended Dataset 9. Condition-specific principal cell, intercalated cell, and proximal tubule CellPhoneDB interactions.

Extended Dataset 10. Cell cluster identification and gene expression tables for 98-cluster Leiden and 16- and 46-cluster analyses.