bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Title

Comparative analysis of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing

Authors

Jonathan Liu[1,*], Vanessa Tran[1,*], Venkata Naga Pranathi Vemuri[1], Ashley Byrne[1,†], Michael Borja[1], Snigdha Agarwal[1], Ruofan Wang[1], Kyle Awayan[1], Abhishek Murti[2], Aris Taychameekiatchai[2], Bruce Wang[2], George Emanuel[3], Jiang He[3], John Haliburton[1,‡], Angela Oliveira Pisco[1], Norma Neff[1,#]

[1] Chan Zuckerberg Biohub, 499 Illinois St, San Francisco, CA 94158
[2] School of Medicine, University of California, San Francisco, 533 Parnassus Ave, San Francisco, CA 94143
[3] Vizgen Inc. 61 Moulton Street, Cambridge, MA, USA, 02138
* equal authorship
[†] Current affiliation: Department of Microchemistry, Proteomics, Lipidomics and Next Generation Sequencing, Genentech, Inc., South San Francisco, CA, USA
[‡] Current affiliation: Prolific Machines, 479 Jessie St, San Francisco, CA 94103
[#] Correspondence to norma.neff@czbiohub.org

Abstract

Spatial transcriptomics extends single cell RNA sequencing (scRNA-seq) technologies by providing spatial context for cell type identification and analysis. In particular, imaging-based spatial technologies such as Multiplexed Error-Robust Fluorescence In Situ Hybridization (MERFISH) can achieve single-cell resolution, allowing for the direct mapping of single cell identities to spatial positions. Nevertheless, because MERFISH produces an intrinsically different data type than scRNA-seq methods, a technical comparison between the two modalities is necessary to ascertain how best to integrate them. Here, we used the Vizgen MERSCOPE platform to perform MERFISH on mouse liver and kidney tissues and compared the resulting bulk and single-cell RNA statistics with those from the existing Tabula Muris Senis cell atlas. We found that MERFISH produced measurements that quantitatively reproduced the bulk RNA-seq and scRNA-seq results, with some minor differences in overall gene dropout rates and single-cell transcript count statistics. Finally, we explored the ability of MERFISH to identify cell types, and found that it could independently resolve distinct cell types and spatial structure in both liver and kidney. Computational integration with the Tabula Muris Senis atlas using scVI and scANVI did not noticeably enhance these results. We conclude that compared to scRNA-seq, MERFISH provides a quantitatively comparable method for measuring single-cell gene expression, and that efficient gene panel design allows for robust identification of cell

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

types with intact spatial information without the need for computational integration with scRNA-seq reference atlases.

Introduction

Named "Method of the Year" by Nature Methods in 2021 (Marx 2021), spatial transcriptomics promises to revolutionize biological investigation by allowing researchers to study cells' transcriptomes in their native spatial context. Recently, a plethora of various spatial transcriptomics techniques have arisen, including platforms such as Slide-Seq (Stickels et al. 2021), 10x Visium (Ståhl et al. 2016), Seq-FISH (Eng et al. 2019), MERFISH (Chen et al. 2015), STARmap (X. Wang et al. 2018), and GeoMX Digital Spatial Profiler (Merritt et al. 2020), to name a few examples. These techniques all involve different fundamental mechanisms, ranging from traditional RNA sequencing performed on spatially barcoded chips to single-molecule fluorescence microscopy.

MERFISH, one of the aforementioned microscopy-based methods, provides a particularly compelling approach to complement pre-existing single-cell RNA sequencing (scRNA-seq) techniques. While it can only study hundreds to a few thousands of genes, MERFISH makes up for this limitation by offering single-molecule capability. This results in subcellular spatial resolutions that are orders of magnitude finer than sequencing-based spatial approaches, which are currently limited to spatial resolutions at the length scale of a single cell or greater. Thus, MERFISH can deeply study a target list of genes of interest with spatial context, drawing upon and complementing pre-existing insights offered by traditional single cell analysis.

However, since MERFISH is still relatively new, it lacks a systematic technical comparison with RNA sequencing technologies, especially at the single-cell level. Additionally, while MERFISH has provided numerous insights in cell culture (Chen et al. 2015; Moffitt, Hao, Wang, et al. 2016; Moffitt, Hao, Bambah-Mukku, et al. 2016; Xia et al. 2019), only recently has MERFISH been successfully used in tissue samples, notably in the mouse brain (Moffitt et al. 2018; X. Wang et al. 2018; G. Wang et al. 2020; Zhang et al. 2020) and fetal liver (Liu et al. 2020; Lu et al. 2021). Thus, its potential for further biomedical research necessitates a technical analysis in other types of tissue. Finally, since MERFISH relies on a limited gene panel typically consisting of informative marker genes, it is crucial to ascertain if MERFISH measurements are independently sufficient to conduct single-cell bioinformatic analysis such as cell type clustering and identification, or if they need to be integrated with reference scRNA-seq atlases using computational methods.

In this paper, we used the Vizgen MERSCOPE platform to conduct a technical comparison between MERFISH and RNA sequencing in mouse liver and kidney tissues. By comparing the insights from MERFISH with traditional bulk and scRNA-seq results from Tabula Muris Senis (Schaum et al. 2020, The Tabula Muris Consortium et al. 2020), we explored the advantages and limitations of MERFISH for bioinformatic analysis. After establishing that the statistics exhibited by both imaging and sequencing modalities were fairly similar, we examined MERFISH's ability to identify different cell types. Intriguingly, MERFISH was able to distinguish between different cell types in liver and tissue with standard single-cell bioinformatics analysis,

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

even with only a 307-gene panel containing marker genes split between three organs. Furthermore, we were able to resolve clear spatial structure in the morphological distribution of these cell types, for example in spatial patterning of pericentral and periportal hepatocytes and endothelial cell subpopulations in mouse liver.

We then integrated our MERFISH measurements with Tabula Muris Senis using scVI (Lopez et al. 2018) and scANVI (Xu et al. 2021) to automatically predict cell type annotations and investigate what additional insights were offered via computational integration. Interestingly, computational integration did not noticeably enhance cell type identification, indicating that MERFISH data alone can sufficiently resolve distinct cell types.

Results

MERFISH utilizes a combination of single-molecule fluorescence in-situ hybridization (smFISH) and combinatorial labeling of RNA transcripts with optical barcoding to achieve a highly multiplexed, single-molecule readout of transcriptional activity in fixed samples (Chen et al. 2015). Here, we briefly describe the essence of the technique as applied in this work.

MERFISH provides targeted labeling of RNA transcripts by using a preselected gene panel, where each transcript is assigned a unique binary barcode (Fig 1A). These barcodes are error robust, allowing for reliable transcript identification even with several hundreds of genes (Moffitt, Hao, Wang, et al. 2016). After hybridizing a sample with encoding probes that effectively imprint the desired barcodes onto each RNA species, the barcode is then detected by sequential rounds of multichannel imaging, with different subsets of fluorescently labeled readout probes flowing in to hybridize with the barcode region of encoding probes. Fluorescent spots are computationally decoded into their respective binary barcodes wherein the presence of a spot indicates a “1” and absence indicates a “0” (Fig 1A). These barcodes and their positions, combined with cell nucleus and boundary staining and segmentation, allow for measurement of gene expression at single cell resolution.

Figure 1B summarizes the experimental workflow. After designing a gene panel barcode scheme, we hybridized encoding probes to a tissue sample and imaged it (see Methods and Materials for detailed overview of protocol). The resulting raw images contained the necessary information to decode RNA transcript counts. In addition, the sample was stained with a combination of DAPI and membrane protein antibodies, producing a fluorescent readout on cell nuclei and boundaries. Using the MERlin image analysis package (Emanuel et al. 2020), we processed these raw images to obtain positions of single RNA transcripts and segmented geometries of individual cells. These data were then post-processed to calculate more parsimonious summary statistics, such as individual cell volumes and single-cell count matrices of each RNA species.

In this work, we investigated a 307-gene panel in mouse liver and kidney. The genes were selected to be differentially expressed cell type marker genes for mouse liver, kidney, or pancreas (see Methods and Materials for more details). In total, 14 datasets were taken with 5, 4, and 5 datasets for mouse liver, kidney, and pancreas, respectively. Of these, all of the pancreas samples and two of the kidney samples did not yield enough RNA for downstream analysis (Section S1 and SI Fig 1). In addition, only one of the liver samples was stained for

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DAPI and cell boundary antibodies that allowed for single cell segmentation – the other four were used for bulk RNA analysis only.

Example workflow

Figure 2 shows an example of the data acquisition and image analysis procedure for a mouse kidney tissue sample. From a ~1cm^2 tissue sample (Fig 2A), zoomed-in images are acquired for various fields of view (FOVs). Each image consists of a z-stack of seven z-positions, with a space of 0.7 microns between each z-position. Within each FOV, the sample is first imaged for the DAPI and cell boundary antibody stains, shown for a zoomed-in snippet in Fig 2B and 2C. Then, the multiple rounds of hybridization begin and the various smFISH signals are acquired, as shown in Fig. 2D for a single bit.

From examining the presence, absence, and colocalization of individual smFISH spots, the digital barcodes of unique RNA species can be decoded using computational image analysis (Fig. 2E, colored points; see Methods and Materials for details). Furthermore, the cell boundary antibody stain and DAPI channels can be used to segment and produce single cell boundaries (Fig. 2E, black). These result in processed images containing the single cell boundaries and decoded transcript positions, which are then used for downstream bioinformatic analysis.

Prior to further investigation, we first examined our measurements for sample quality. To do so, we computed the average density of detected RNA transcripts for each sample and compared this number with the RNA integrity number (RIN), a common metric for measuring RNA quality (Schroeder et al. 2006). The two metrics correlated quite well (Section S1 and SI Fig 1), and we discarded measurements with RIN score lower than 4. All five liver datasets passed this margin, as well as two kidney datasets.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

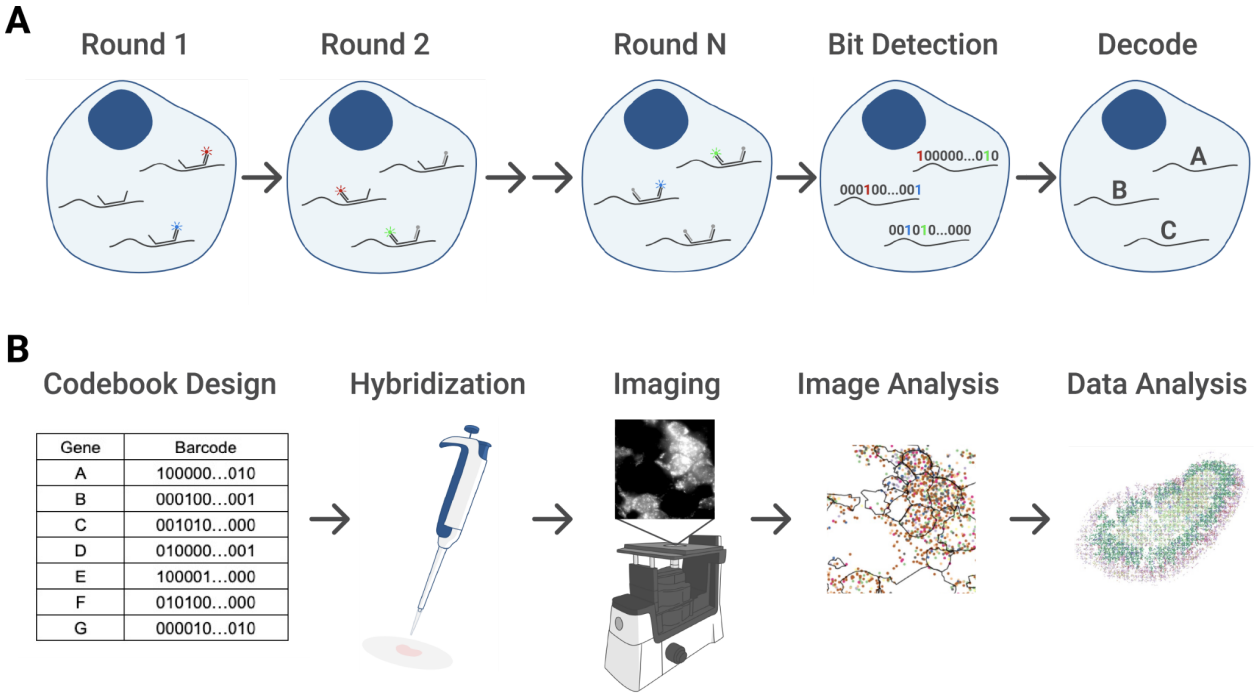


Figure 1: **Overview of experimental setup.** (A) Barcoding scheme across N rounds of multicolor imaging. Each detected fluorescent spot is registered as a "1" in its corresponding bit position. Then, RNA species are identified by matching the barcodes to a preselected gene codebook. (B) Experimental workflow. First, the gene panel and codebook are designed. Then, probes are hybridized onto the sample, which has been mounted and fixed onto a coverslip that is in turn placed in a flow cell connected to an automated fluidic system. The sample is imaged in 5 color channels at multiple fields of view and z-positions using an inverted epi-fluorescence microscope. Images are analyzed to identify RNA molecules and their spatial coordinates as well as cell boundaries. These data result in single-cell count matrices that can then be analyzed with standard bioinformatics tools to achieve goals such as cell type identification.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

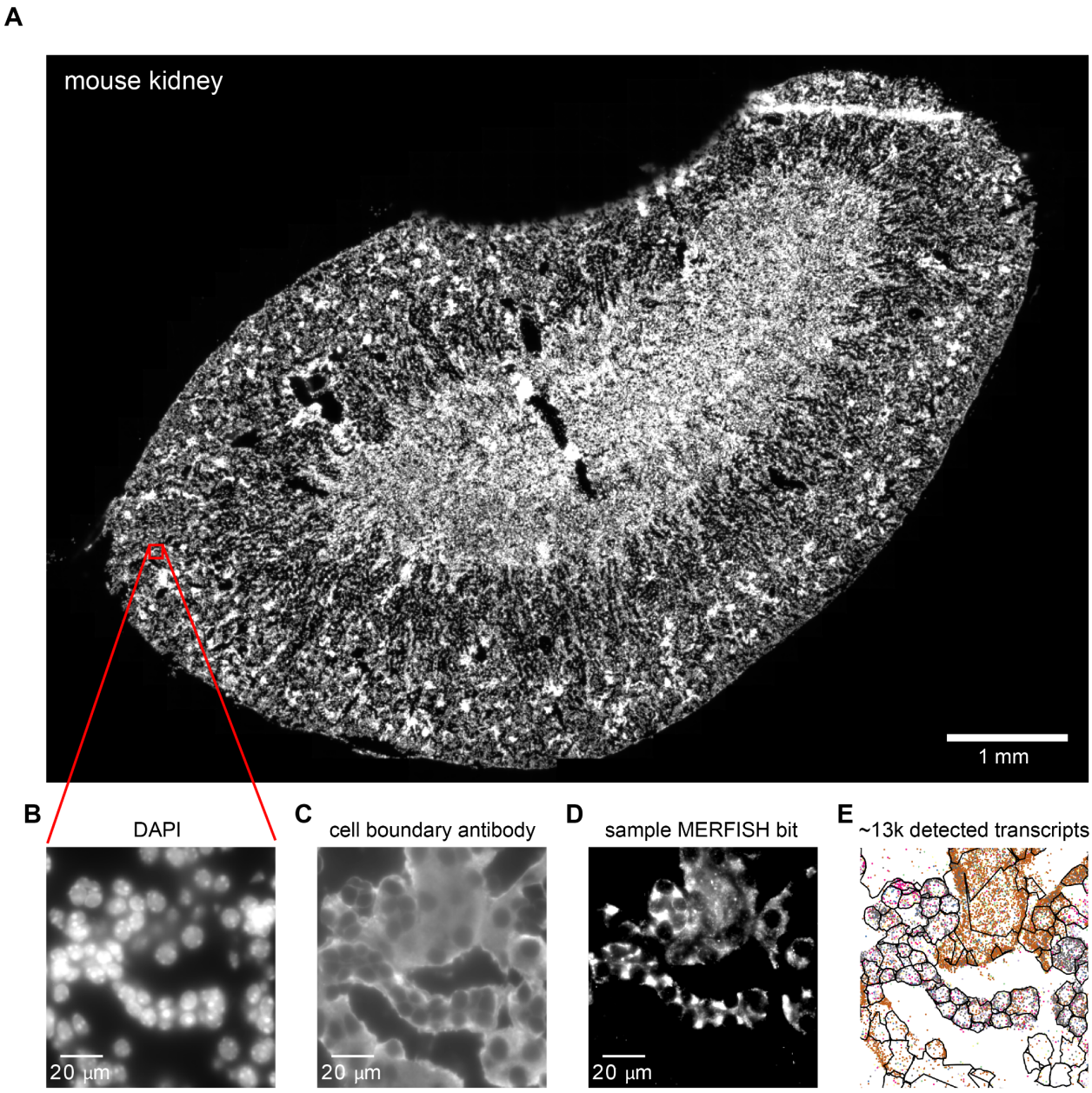


Figure 2 - Sample MERFISH data acquisition workflow. (A) Low-resolution image of DAPI channel for a mouse kidney tissue sample. Red box indicates zoomed-in region displayed in B-E. (B, C) DAPI (B) and (C) cell boundary antibody stain channels for the zoomed-in region. (D) MERFISH signal for a single barcode bit channel in the same zoomed-in region. (E) Positions of decoded mRNA transcripts (colorful dots) and segmented cell boundaries (black) in the same zoomed-in region after running data through image analysis pipeline, with each color representing a unique gene species. Panels B-D show raw images with brightness and contrast levels selected for ease of visualization.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Comparison of MERFISH results with bulk RNA-seq

First, we examined noise in MERFISH measurements by comparing bulk RNA transcript counts between technical replicates. Figure 3A and 3B show total log-transformed counts of detected RNA transcripts for each gene in the panel between two technical replicates taken from the same tissue block in mouse liver and kidney. Both tissues exhibited high correlation, with R^2 = 0.99 and R^2 = 0.95 for liver and kidney, respectively. Thus, results from MERFISH replicates were extremely reproducible.

Then, we compared bulk MERFISH results with bulk RNA-seq data from Tabula Muris Senis (Schaum et al. 2020). Since the mice used in this study were three months old, we only considered those mice from Tabula Muris Senis that were also three months old. Figure 3C and 3D show the total log-transformed counts per gene in mouse liver and kidney, respectively, between the two types of technologies. Each point represents the RNA count for a single gene, averaged across different experimental samples for the corresponding technology. While the correlation was lower than between MERFISH technical replicates, it was still apparent (R^2 = 0.61 and R^2 = 0.58 for liver and kidney, respectively), albeit slightly lower than previous MERFISH studies in cell culture (Xia et al. 2019). Furthermore, MERFISH systematically detected more transcripts than bulk RNA-seq, with fold-change increases of ~10x to over ~1000x.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

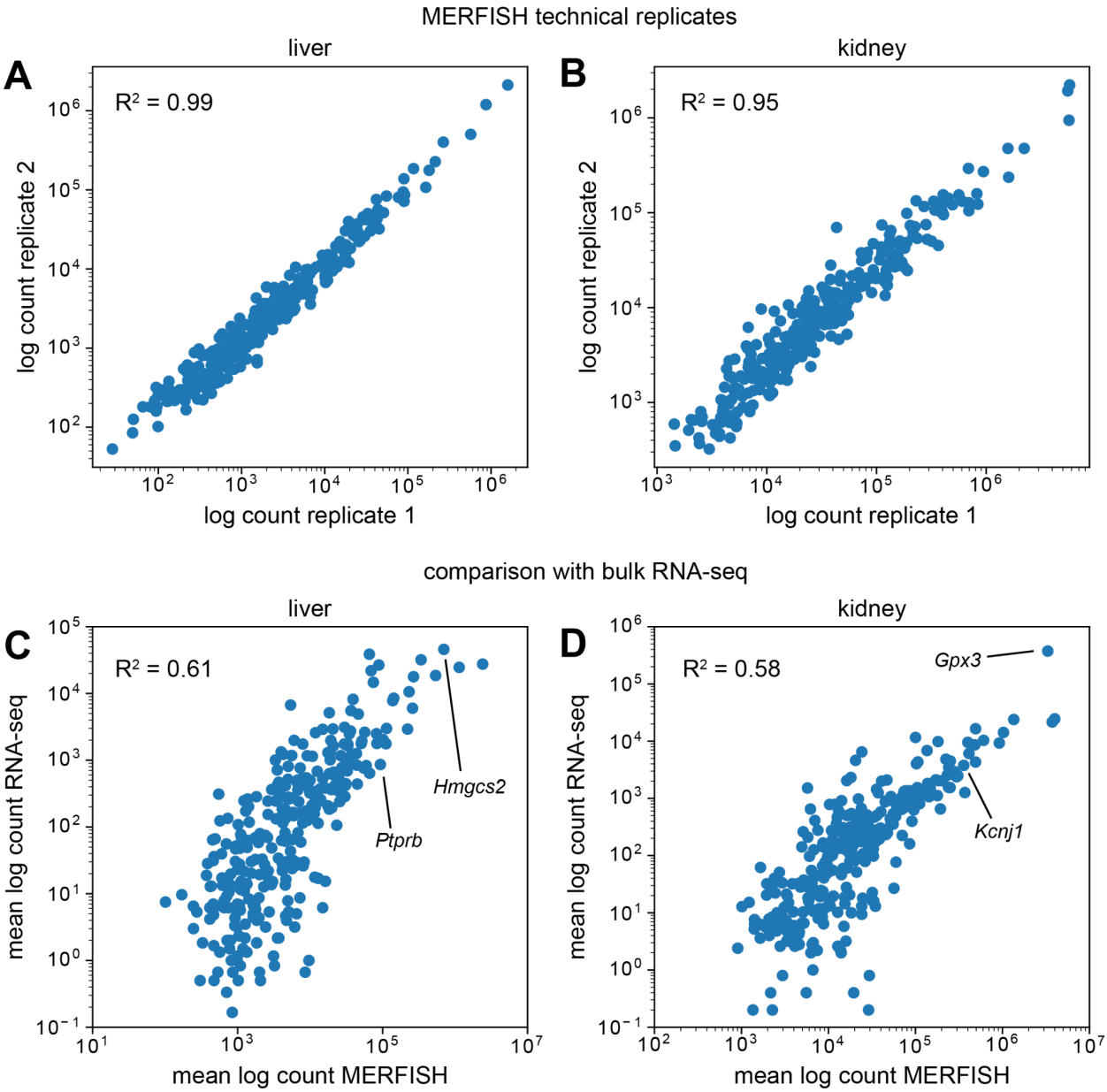


Figure 3 - Comparison of MERFISH data with bulk RNA-seq from Tabula Muris Senis. (A, B) Bulk RNA counts per gene in mouse (A) liver and (B) kidney between MERFISH sample replicates. (C, D) Comparison in bulk RNA counts per gene in mouse (C) liver and (D) kidney between MERFISH and RNA-seq. In C and D, counts per gene were averaged across replicates for each technology.

Analysis and quality control of single-cell MERFISH results

After image analysis, we extracted 3D segmented cell boundaries and decoded RNA transcripts positions, as shown for sample zoomed-in views at a single z-position in Fig. 4A and 4E for mouse liver and kidney, respectively. Segmentation quality was variable, with some cells

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

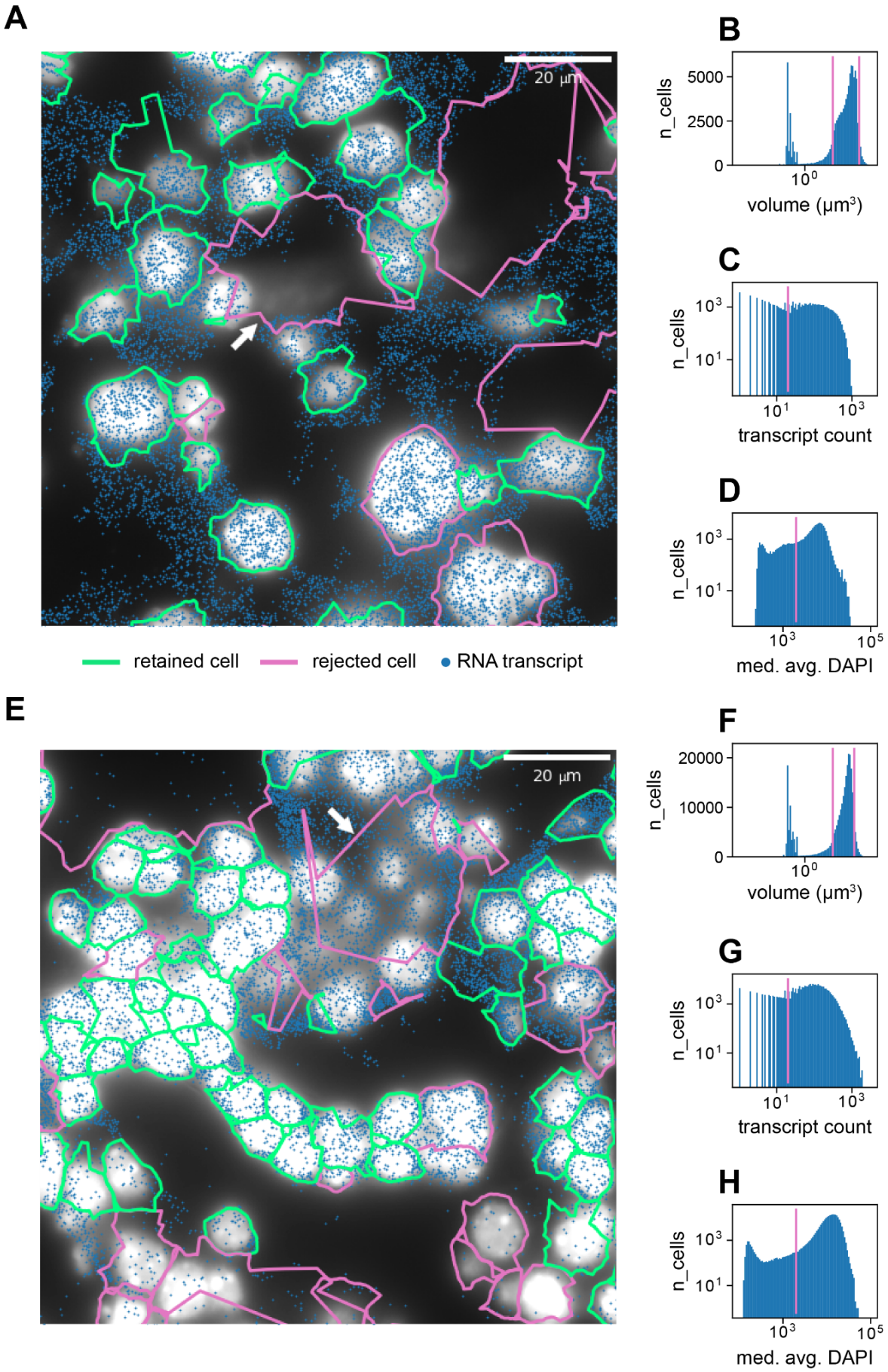
possessing segmented boundaries that appeared quite reliable and others possessing aberrant segmented boundaries. These discrepancies were due to low-quality cell boundary and/or DAPI stain signals, resulting in the image analysis software constructing segmented boundaries in the empty spaces between cells, or grouping multiple cells together (Fig. 4A and 4E, white arrows). In addition, not all detected RNA transcripts lay inside segmented boundaries – the typical percentage of transcripts assignable to cells was around 30-50%.

To filter out these poor quality cells, we developed a quality control protocol. From the data, we generated distributions of single-cell 3D volumes and total counts per cell (Fig. 4B, C, F, and G). In addition, we calculated a per-cell alignment metric with the DAPI signal, which we call the *median average DAPI* score. This score is the median across z-slices of the average DAPI intensity within the xy borders of each segmented cell in each z-slice (Fig 4D and 4H). Intuitively, a well segmented cell would have a high median average DAPI score due to good alignment between the segmentation and the DAPI signal, whereas a poorly segmented cell would have a lower score.

We established quality control filters as follows: segmented cells that were unreasonably small or large were discarded by establishing low-volume and high-volume cutoffs of 50 μm^3 and 2000 μm^3 (liver) and 50 μm^3 and 1000 μm^3 (kidney), respectively, in the distribution of cell volumes (Fig. 4B and 4F). Additionally, a minimum cutoff of total RNA transcript count per cell of 20 was established to filter out cells with sparse statistics (Fig 4C and 4G). Finally, both the mouse liver and kidney datasets exhibited bimodality in the median average DAPI score per cell (Fig. 4D and 4H), and a cutoff of 2000 was established in order to retain cells with a high score. This allowed us to filter out segmented cells that were actually composed of empty space (Fig. 4A, white arrow) or consisted of multiple cells grouped together (Fig. 4E, white arrow).

This quality control procedure resulted in 36,254 liver cells and 210,548 kidney cells being retained out of 98,317 and 508,352 originally segmented cells, respectively. Note that many of the rejected cells were segmentation artifacts with unphysically small volumes (Fig 4B and 4F), so the true original cell count was likely lower.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4 - **Quality control of single cell MERFISH data**. (A) Sample cropped image of mouse liver tissue with segmented cell boundaries (green, magenta), decoded RNA transcripts (blue), and DAPI (grayscale). Green indicates cells that have passed the filtering stage, whereas magenta indicates cells that have been thrown out. (B, C, D) Histograms of (B) cell volumes, (C) RNA transcript count per cell, and (D) median average DAPI per cell for the image in (A). Magenta indicates minimum or maximum cutoff values for filtering criteria. (E) Sample cropped image of mouse kidney tissue. (F, G, H) Histograms of (F) cell volumes, (G) RNA transcript count per cell, and (H) median average DAPI per cell for the image in (E). White arrows in A and E indicate examples of poor cell segmentation.

Comparison of single-cell MERFISH results with scRNA-seq

From the segmented data, we constructed single-cell RNA count matrices in order to compare the single-cell MERFISH results with the three-month Tabula Muris Senis scRNA-seq data (The Tabula Muris Consortium et al. 2020). First, we examined the distribution of total RNA transcript counts per cell, as shown in Fig 5A and 5B for mouse liver and kidney. To allow for a proper comparison, the Tabula Muris Senis data were reduced to a subset containing the same 307 genes as the MERFISH gene panel. As shown in Fig 5A and 5B, the MERFISH distribution quantitatively agrees with scRNA-seq up tol total counts of approximately 1000 transcripts/cell, after which the MERFISH distribution drops to zero while the scRNA-seq distribution continues. This effect was more pronounced in the kidney than in the liver.

We hypothesize that this is due to the fact that the gene panel contained a few highly abundant genes in mouse liver and kidney such as *C1qc* and *Gpx3* respectively (Section S2 and SI Fig 2); as a result, for the minority of cells containing extremely high numbers of transcripts (i.e. over 1000 per cell), the smFISH spots in the MERFISH images would be too dense due to overcrowding of RNA molecules, preventing accurate identification of single RNA transcripts. Examination of single-cell distributions of these abundant genes verified this hypothesis, as the distributions resulting from MERFISH agreed with those from scRNA-seq for cells with low overall transcript counts but not for cells with high overall transcript counts (Section S2 and SI Fig 2).

Next, we examined the dropout rate between the technologies, shown in Fig 5C and 5D. For each cell in a given tissue, this was defined as the fraction of genes with zero counts out of the whole 307-gene panel. For both technologies, the dropout rates were rather high (> 0.7) due to the fact that the gene panel contained marker genes specific for liver, kidney, and pancreas. Nevertheless, the relative difference in the distribution between technologies was informative. In the liver (Fig 5C), the distribution of dropout rates in MERFISH was wider than in scRNA-seq, with more cells possessing lower dropout rates (< 0.8) in MERFISH. In the kidney (Fig 5D), the distribution of dropout rates was slightly shifted to the right (i.e. higher dropout rates) in MERFISH compared to scRNA-seq. This is likely due to the molecular crowding effect described above leading to a high density of the fluorescent signal that exceeds the diffraction limit, which cannot be resolved and hence prevents those transcripts from being counted.

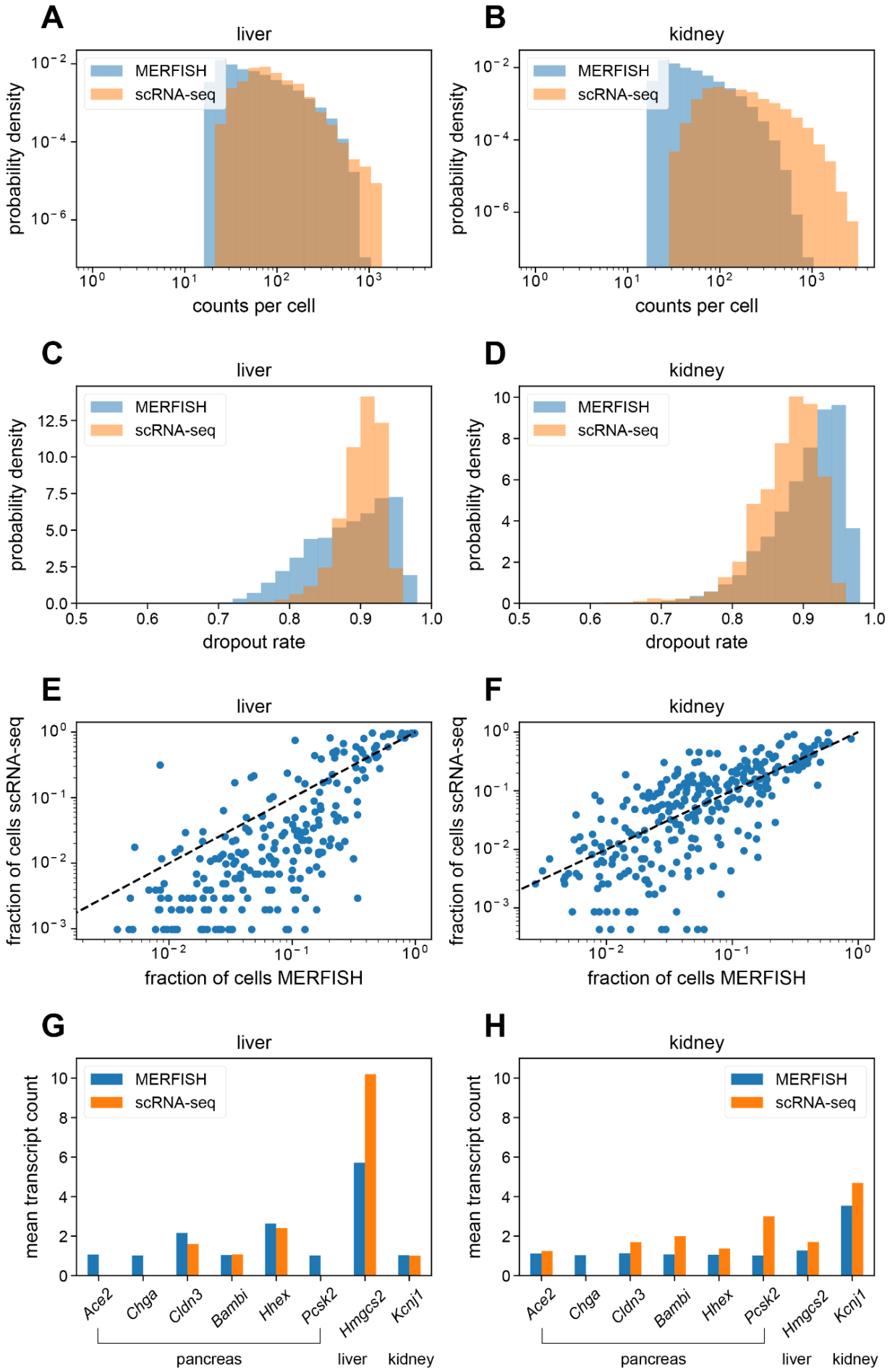
We then investigated whether either technology was more sensitive than the other (Fig 5E and 5F). While the previous analysis of dropout rates quantified the number of genes

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

detected per cell, here we quantified the number of cells in which each gene was detected. To do so, we calculated the fraction of cells that had nonzero counts for each gene. If the two technologies had identical sensitivities, then a scatter plot of this fraction for each gene between technologies would fall on the x=y line. Bias in either direction would indicate that one technology was more sensitive. In the liver (Fig 5E), MERFISH systematically detected genes in higher proportions of cells than scRNA-seq. In contrast, there was no clear systematic bias in the kidney (Fig 5F). Thus, for this gene panel, we concluded that MERFISH was more sensitive than scRNA-seq in the liver and about as sensitive in the kidney.

Finally, we investigated possible detection of false positives (i.e. incorrectly decoded RNA transcripts) in MERFISH. To do so, we reasoned that the pancreas marker genes in the panel could be used as a control – substantial detection of these pancreas genes in the MERFISH data would indicate a nontrivial false positive rate. Figures 5G and 5H show the mean transcript count of six pancreas marker genes (*Ace2, Chga, Cldn3, Bambi, Hhex, Pcsk2*) among cells with nonzero counts in both the MERFISH and Tabula Muris Senis datasets. The standard errors of the mean were also calculated but were too small to visualize. For comparison, a liver and kidney marker gene (*Hmgcs2* and *Kcnj1*, respectively) are shown as well. The values for both MERFISH and scRNA-seq were similar for the pancreas genes with a few exceptions likely due to expected dropouts in scRNA-seq (*Ace2, Chga, Pcsk2*), indicating that both technologies have comparable false positive rates. In addition, the expression levels of *Hmgcs2* were high in liver and low in kidney, and the expression levels of *Kcnj1* were low in liver and high in kidney, as expected. There was a slight discrepancy in the measured expression of the highly expressed *Hmgcs2* transcript in liver between MERFISH and scRNA-seq, probably due to the molecular crowding effect already discussed. Together these data indicated that the false positive rates were low in MERFISH and scRNA-seq.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5 - **Comparison of single-cell MERFISH data with Tabula Muris Senis**. (A, B) Histograms of total RNA transcript count per cell in mouse (A) liver and (B) kidney for the gene panel, for MERFISH (blue) and scRNA-seq (orange). (C, D) Histograms of dropout rates per cell in mouse (C) liver and (D) kidney for the gene panel. (E, F) Comparison between MERFISH and scRNA-seq of per-gene fraction of total cells out of the whole population of each dataset that have nonzero counts for mouse (E) liver and (F) kidney. Here, each dot represents a single gene. (G, H) Mean transcript count of pancreas marker genes in mouse (G) liver and (H) kidney, for MERFISH (blue) and scRNA-seq (orange). A liver marker gene (*Hmgcs2*) and kidney marker gene (*Kcnj1*) are shown for comparison. Standard error of the mean across cells was negligible and too small to visualize.

Single cell and spatial analysis of MERFISH results

Using the segmented single-cell results, we explored MERFISH's ability to resolve distinct liver cell types on its own. First, we visualized a low-dimensional embedding using UMAP (McInnes et al. 2018) and conducted unsupervised clustering using the leiden algorithm (Traag, Waltman, and van Eck 2019). To improve signal quality, prior to UMAP visualization we further filtered out cells that had fewer than 80 total RNA transcript counts (Section S3 and SI Fig 3).

Even with just the 307-gene panel, we observed clear separation of clusters corresponding to various liver cell types in the UMAP plot (Fig 6A, left). By examining the most informative genes representative of each group, we were able to assign cell type identities to the various clusters, with the exception of a cluster of three cells that we were unable to annotate (these cells were thus removed from analysis). The three right plots in Fig 6A show expression levels of some example genes, *cyp2f2*, *ptprb*, and *clec4f*, that are marker genes for periportal hepatocytes, endothelial cells, and Kupffer cells, respectively.

As expected, the bulk of the cells were hepatocytes. Notably, the data were able to resolve between different subpopulations of hepatocytes, with both periportal and pericentral hepatocytes clearly represented in the data (Fig 6A, pink and gray). Additionally, we were also able to distinguish between periportal and pericentral endothelial cells (Fig 6A, olive green and cyan).

Projecting these cell type annotations onto the spatial plot of mouse liver allowed us to observe clear spatial structure of the different cell types (Fig 6B). In particular, the periportal and pericentral hepatocytes and endothelial cells exhibited spatial segregation via prominent maze-like patterning. Focusing on this zonation revealed clear co-localization of periportal hepatocytes and endothelial cells, as well as of pericentral hepatocytes and endothelial cells (Fig 6B, inset; Section S4; SI Fig 4). A DAPI stain of the sample is shown in Fig 6C for reference.

To see if we could further enhance our signal quality, we integrated the MERFISH liver dataset with the annotated Tabula Muris Senis liver cell atlas (The Tabula Muris Consortium et al. 2020) using scVI (Lopez et al. 2018). To do so, we retained mice from the atlas that were three months old, in addition to non-hepatocyte cells from the one-month dataset to increase cell diversity, as the three-month data primarily consisted of hepatocytes (see Methods and

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Materials). We then subsetted these data to the same 307-gene panel as the MERFISH data and then trained an scVI model to combine the two datasets. Details of the analysis are given in Section S5 and the combined UMAP in the joint embedding is shown in SI Fig 5A. Finally, we used scANVI (Xu et al. 2021) to predict cell type labels in the MERFISH dataset from the annotations from Tabula Muris Senis (SI Fig 5B). The resulting spatial plot with projected scANVI cell type annotations (Fig 6D) qualitatively reproduced the spatial results from the MERFISH data alone (Fig 6B).

Comparisons of cell type frequencies between the scRNA-seq reference, MERFISH data, and scANVI integration are shown in Figure 6E. There was a discrepancy between scRNA-seq and MERFISH in the frequencies of hepatocyte subpopulations, likely because the tissue sample used here did not contain the whole liver and thus lacked a globally representative population of hepatocytes. Furthermore, the scRNA-seq reference contained other hepatocytes and endothelial cells that were neither periportal nor pericentral; these cell types were not readily detected in the manual annotation scheme used for the MERFISH data. More notably, MERFISH detected substantially fewer immune and Kupffer cells and more hepatic stellate cells than the scRNA-seq reference, hinting at the fact that *in situ* hybridization technologies could offer a more accurate estimate of cell type proportions since scRNA-seq may overestimate immune cell counts (Wu et al. 2019; Denisenko et al. 2020; Ding et al. 2020; Koenitzer et al. 2020; Slyper et al. 2020).

We then repeated this analysis for the MERFISH kidney data. The left panel in Fig 7A shows a UMAP plot of the MERFISH data alone from a single tissue sample, where again we could resolve distinct clusters with the leiden algorithm. Here, we filtered out cells that had total RNA counts below 150 (Section S6 and SI Fig 6). We then annotated the various clusters as before, by examining the most informative genes in each cluster. The three right plots in Figure 7A show the example marker genes *kcnj1*, *clu,* and *podxl*, which along with other marker genes allowed us to detect Loop of Henle epithelial cells, distal convoluted tubule epithelial cells, and podocytes, respectively. Altogether, while the kidney contains many more known cell types than the liver, we were able to resolve broad categories of cell types, including endothelial cells, various epithelial cell groups, and podocytes. We did not detect immune cells, fibroblasts, and pericytes via manual annotation.

Figure 7B shows the corresponding mouse kidney spatial plot, with cell type labels projected from this annotation process. We observed clear stratification of the kidney into different layers – epithelial proximal tubule cells were localized to the cortex (Fig 7B, pink) while loop of Henle epithelial cells were localized to the medulla (Fig 7B, red). Endothelial cells were more uniformly distributed throughout the kidney sample (Fig 7B, blue). Notably, we also discovered clusters of podocytes (Fig 7B, olive green), a cell type that only exists in glomeruli (Brunskill et al. 2011), clusters of blood vessels that exist dispersed throughout the kidney cortex. A DAPI stain is shown for reference in Figure 7C.

We then again used scVI to integrate the kidney MERFISH data with the Tabula Muris Senis kidney scRNA-seq reference atlas, using a similar subsetting protocol as we did for the liver data (Section S5). Because the reference scRNA-seq kidney data were more comprehensive than the liver data, we used only three-month-old mouse data. This resulted in the joint UMAP plot shown in SI Fig 5C.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Using scANVI, we then predicted cell types in the MERFISH data from the annotations in Tabula Muris Senis (SI Fig 5D). Although the spatial kidney plot with predicted scANVI cell type labels exhibited qualitatively similar spatial structure (Fig 7D) to the spatial results from MERFISH alone (Fig 7B), it did produce some clear artifacts. For example, podocytes only exist in spatial clusters concentrated in glomeruli in the cortex, but the integration predicted the existence of a large podocyte population that lay in a ring around the medulla (Fig 7D, olive green).

Examining the relative cell type frequencies between the scRNA-seq reference, MERFISH results, and scANVI integration shed additional light on this discrepancy (Fig 7E). Because the kidney sample here covered the entire organ, as opposed to the liver sample in Figure 6, the agreement between MERFISH and scRNA-seq was much more quantitatively apparent (Fig 7E, blue and green). The major differences were that MERFISH did not detect pericytes, fibroblasts, or immune cells, and detected more endothelial cells and fewer podocytes than scRNA-seq. The scANVI integration produced many more discrepancies (Fig 7E, orange), such as an abnormally high podocyte count and a low epithelial proximal tubule cell count compared to scRNA-seq. We thus take the MERFISH-alone results to be more reliable than the scANVI integration.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

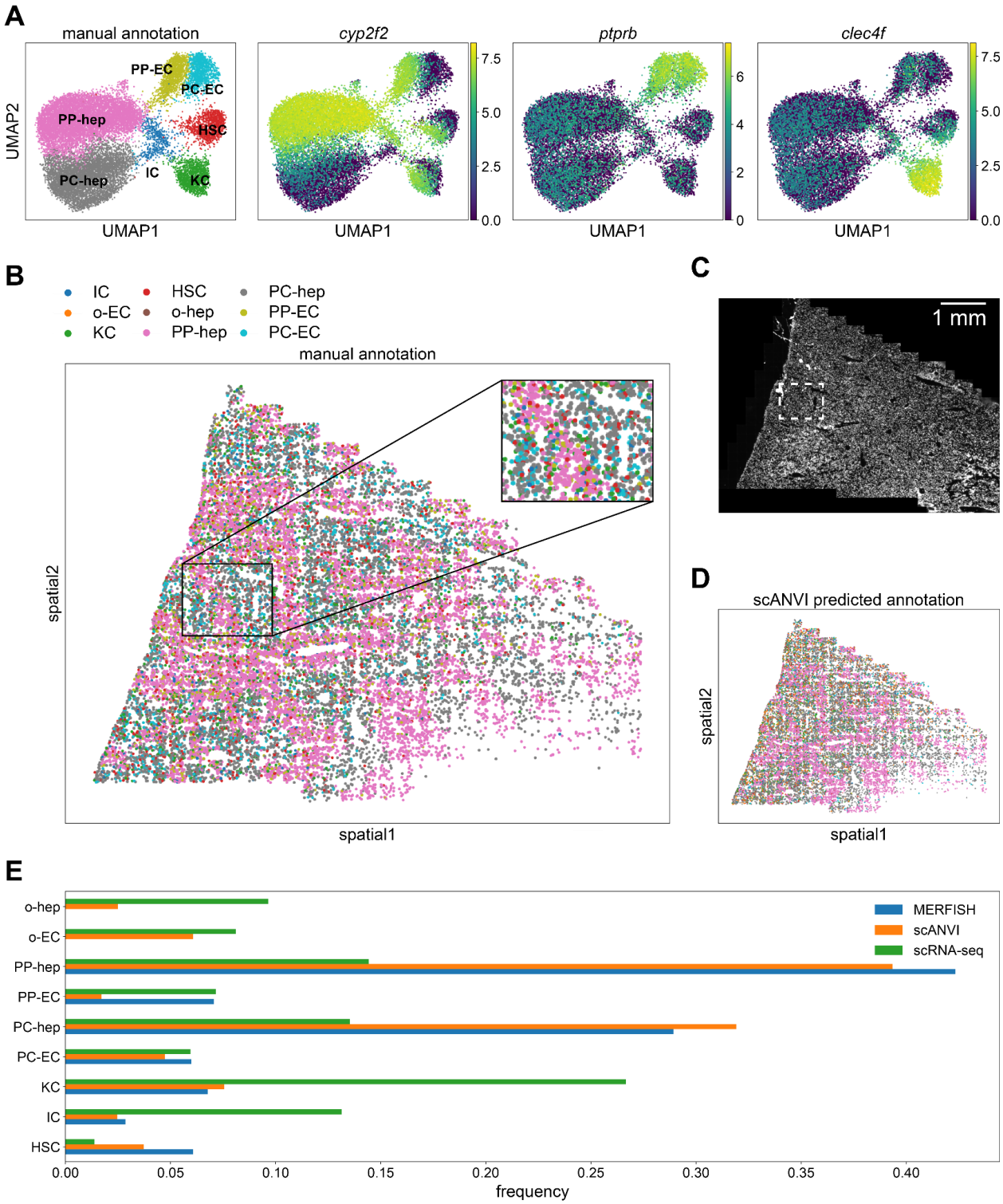


Figure 6 - Single-cell and spatial analysis of MERFISH liver sample. (A) UMAP plots of MERFISH data colored by manually annotated clusters and normalized, log-transformed expression of example marker genes. (B) Spatial plot of MERFISH dataset alone, colored by manually annotated cell types in A. Inset highlights spatial co-localization of periportal endothelial cells and hepatocytes (olive green and pink, respectively) and

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pericentral endothelial cells and hepatocytes (cyan and gray, respectively). (C) DAPI stain of liver sample. White box indicates the same inset region as in panel C. (D) Spatial plot of MERFISH dataset using scANVI predicted cell type labels. Legend is the same as in panel B. (E) Cell type composition for scRNA-seq and MERFISH datasets. Each point in A, B, and D represents a single cell. Cell type abbreviations are as follows: 'IC': 'immune cell', 'o-EC': 'other endothelial cell', 'KC': 'Kupffer cell', 'HSC': 'hepatic stellate cell', 'o-hep': 'other hepatocyte', 'PP-hep': 'periportal hepatocyte', 'PC-hep': 'pericentral hepatocyte', 'PP-EC': 'periportal endothelial cell', 'PC-EC': 'pericentral endothelial cell'.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

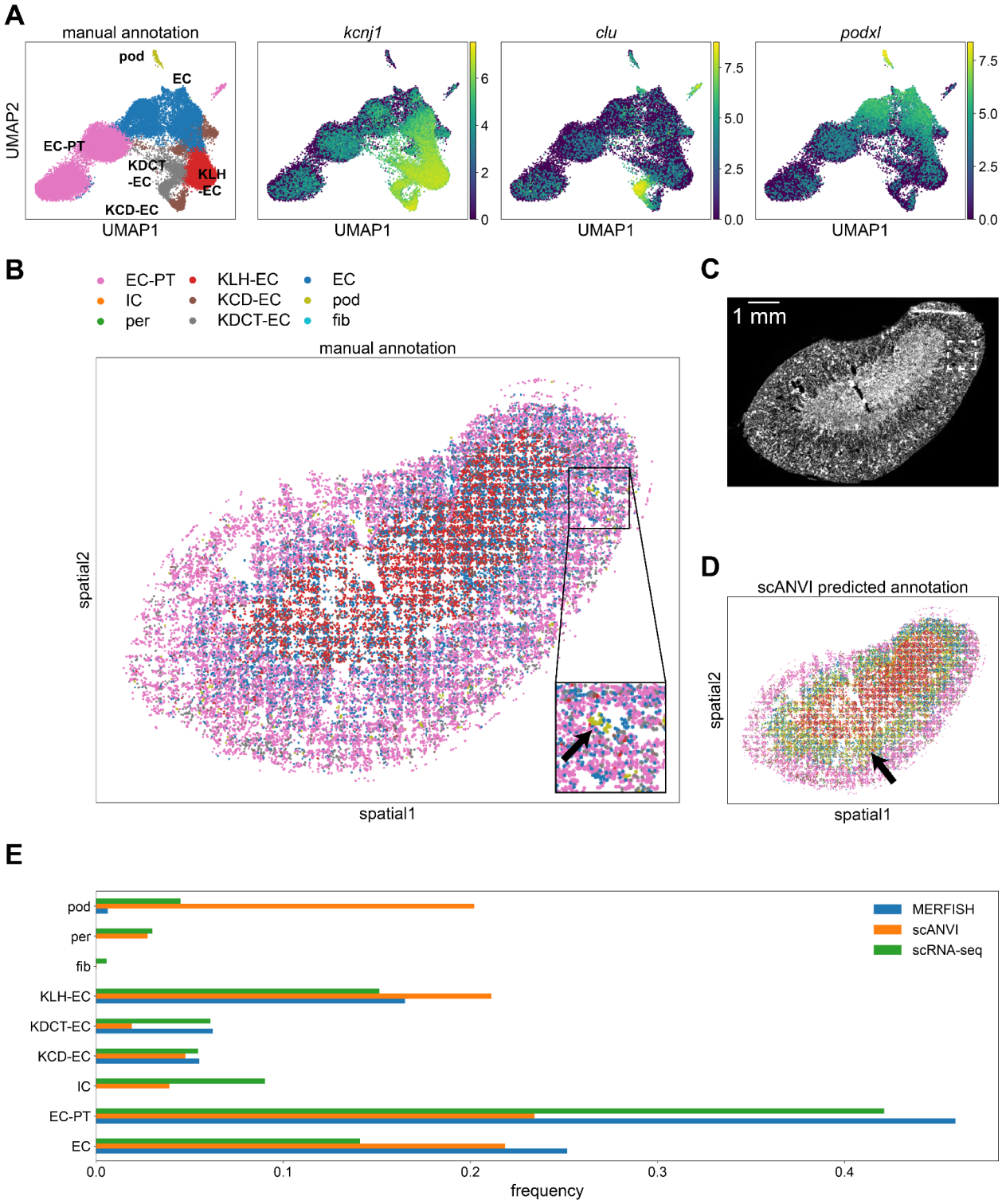


Figure 7 - Single-cell and spatial analysis of MERFISH kidney sample. (A) UMAP plots of MERFISH data colored by manually annotated clusters and normalized, log-transformed expression of example marker genes. (B) Spatial plot of MERFISH dataset alone, colored by manually annotated cell types in A. Inset shows an example of a podocyte cluster in the kidney cortex region (olive green, black arrow). (C) DAPI stain

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of kidney sample. White box indicates the same inset region as in panel C. (D) Spatial plot of MERFISH dataset using scANVI predicted cell type labels. Black arrow indicates falsely predicted podocytes distributed in a ring-like structure around the medulla. Legend is the same as in panel B. (E) Cell type composition for scRNA-seq and MERFISH datasets. Each point in A, B, and D represents a single cell. Cell type abbreviations are as follows: 'EC-PT': 'epithelial cell of proximal tubule', 'IC': 'immune cell', 'per': 'pericyte', 'KLH-EC': 'kidney loop of Henle epithelial cell', 'KCD-EC': 'kidney collecting duct epithelial cell', 'KDCT-EC': 'kidney distal convoluted tubule epithelial cell', 'EC': 'endothelial cell', 'pod': 'podocyte', 'fib': 'fibroblast'.

Discussion

A burgeoning technology, MERFISH offers a paradigm shift in transcriptomic analysis by combining subcellular spatial resolution with single cell segmentation ability. However, while MERFISH promises to extend single-cell transcriptomic analysis to the spatial domain, it is unclear as to what extent MERFISH measurements are quantitatively comparable to those from traditional RNA sequencing technologies. Thus, a technical comparison with existing RNA sequencing data is necessary to fully understand the similarities and differences between the two modalities. In particular, it remains to be seen if MERFISH can independently enable cell type identification, or if instead such a task requires computational integration with more comprehensive reference scRNA-seq databases.

Here, we used MERFISH to measure the RNA profiles of 307 liver, kidney, or pancreas cell type marker genes in mouse liver and kidney. By utilizing a fixation protocol in conjunction with an automated microfluidics setup and fluorescence microscopy, we measured single-molecule positions of the various genes in large (~1cm^2) sections of fresh-frozen tissue (Fig 1). Combined with cell segmentation based on membrane antibody and DAPI stains, we obtained bulk and single-cell RNA counts (Fig 2) that were then compared with the pre-existing Tabula Muris bulk and single cell RNA datasets.

First, we assessed technical noise by comparing RNA counts between technical replicates of mouse liver and kidney samples, and found extremely high correlation between them (Fig 3A and 3B). Bulk comparisons with Tabula Muris Senis indicated relatively good agreement (Fig 3C and 3D). MERFISH typically produced bulk RNA counts that were several orders of magnitude higher than RNA-seq, an important point in its favor for the measurement of sparse genes.

We then investigated the single-cell RNA counts from our MERFISH measurements. We first developed a quality control protocol to filter out poor cells based on metrics involving spatial morphology, transcript counts, and alignment with DAPI nuclear fluorescence (Fig 4). From this perspective, imaging-based transcriptomics methods such as MERFISH allow for easily interpretable quality control, since these metrics can be ultimately derived from raw images rather than downstream quantities. The predominant factor influencing signal quality was cell segmentation – there were a substantial number of cells that were improperly segmented due to poor fluorescence in the cell nucleus or boundary stain, resulting in roughly one-third to one-half

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of detected cells being filtered out in the quality control protocol. While this conservative filtering strengthens downstream signal quality, it also poses the need for improved computational segmentation in the future to fully harness the massive amounts of data contained in the raw images. For example, recent methods that boost cell segmentation quality by using information from RNA transcript positions (Petukhov et al. 2021) may further refine segmentation quality and allow for improved construction of single cell RNA count matrices from subcellular MERFISH signals.

The filtered single cell MERFISH results were then compared with the Tabula Muris Senis scRNA-seq database (Fig 5). We examined summary statistics in mouse liver and kidney, starting with the total RNA counts per cell (Fig 5A and Fig 5B). For the most part, MERFISH quantitatively reproduced the statistics from scRNA-seq. However, for both liver and kidney, MERFISH was unable to resolve a minority of cells that possessed about 1000 or more total RNA counts from the 307-gene panel (Fig 5A and Fig 5B).

We hypothesized that this was due to the presence of highly abundant genes resulting in crowding of RNA molecules, bringing the density of fluorescent spots to a level exceeding that of the diffraction limit of the microscope and thus obscuring the ability to resolve them. Because MERFISH uses a barcode labeling scheme, this would then reduce signal quality in many genes in the panel for these highly expressing cells. For future MERFISH experiments, we thus recommend avoiding highly expressing genes in the gene panel to avoid this issue. Although the gene panel design implemented here included a filtering step based on overall bulk RNA measurements (see Methods and Materials for description), future selection criteria could incorporate total RNA counts from single-cell reference data as well (Section S2 and SI Fig 2). Overall, we expect the dynamic range of MERFISH measurements in mouse liver and kidney to be limited to cells that have under ~1000 labeled transcripts.

We also investigated measurement sensitivity by quantifying the number of genes measured per cell as well as the number of cells that measured each gene. First, we calculated the dropout rate in liver and kidney, defined as the fraction of genes in the panel possessing nonzero counts per single cell (Fig 5C and Fig 5D). The distribution of dropout rates was slightly lower in the liver compared to scRNA-seq, suggesting that MERFISH was more sensitive than scRNA-seq in that tissue. In the kidney, the distribution was slightly higher, likely due to the molecular crowding issue mentioned above playing a larger role in the kidney than in the liver for this gene panel. We then quantified per-gene sensitivity by examining the fraction of cells that possessed nonzero counts for each gene in the 307-gene panel, for both MERFISH and scRNA-seq (Fig 5E and Fig 5F). In the liver, this fraction was systematically larger for MERFISH than for scRNA-seq, again indicating MERFISH's increased sensitivity. In the kidney, there was no substantial difference between technologies. Taken together, these results suggest that *in situ* hybridization technologies like MERFISH can measure a more intrinsically correct view of RNA statistics than a dissociative technology such as scRNA-seq.

Next, we undertook a negative control by investigating several genes in the panel that were marker genes for pancreas cell types. Examining the mean transcript count for these genes revealed similarly low values for both MERFISH and scRNA-seq compared to expression levels for positive marker genes (Fig 5G and 5H), indicating that MERFISH possessed low false positive detection rates at a similar level to scRNA-seq. Overall, we conclude that MERFISH

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

possesses quantitatively similar overall statistical behavior to scRNA-seq, with minor discrepancies such as the dynamic range issue discussed above.

Finally, we investigated MERFISH's ability to resolve distinct cell types in mouse liver and kidney, and whether computational integration of MERFISH data with a scRNA-seq reference atlas could boost this resolving potential. Single cell analysis of the MERFISH data using UMAP and leiden clustering indicated reasonably good separation of cell types in both liver and kidney (Fig 6A and 7A). These unsupervised clustering results were clean enough to enable manual annotation of cell types based on representation of marker genes in the panel. After projecting these single-cell annotations onto a spatial plot, we were able to resolve clear spatial structure in both mouse liver and kidney samples (Fig 6B and 7B). For example, the MERFISH data were able to distinguish between different liver hepatocyte and endothelial cell populations, exemplified by different spatial patterning of periportal and pericentral hepatocytes and endothelial cells (Fig 6B). Notably, although canonical marker genes for pericentral and periportal endothelial cells such as *Wnt2* and *Dll4* were not present in our gene panel, our MERFISH results could still identify them mainly on the basis of their spatial co-localization with the pericentral and periportal hepatocytes (Fig 6B, inset; Section S4; SI Fig 4). Thus, the incorporation of spatial information into analysis can greatly aid in annotation of cell types by comparing the annotations with expected spatial patterning.

Interestingly, computational integration of the MERFISH data with the Tabula Muris Senis scRNA-seq atlas using scVI did not noticeably improve cell type identification ability. The joint UMAP plots for both liver and kidney were fairly noisy, especially for kidney (SI Fig 5A and SI Fig 5C), suggesting that MERFISH and scRNA-seq exhibit systematically different statistics at the individual RNA species level that hinder efficient harmonization. An additional complicating factor is in the throughput of cells – the individual MERFISH datasets used here possessed tens of thousands of cells, whereas the Tabula Muris Senis possessed only a few thousand total cells for each organ. Therefore, future investigations of this sort of computational integration should include a nuanced accounting of the imbalance in throughput between different technological modalities.

After using scANVI to predict cell type labels in the MERFISH data from the annotations in the Tabula Muris Senis dataset, we generated spatial plots of mouse liver and kidney with these automatically generated labels (Fig 6D and 7D). Although the spatial patterning in the liver was qualitatively similar to the results from the manual annotation, the spatial results of the scANVI integration in the kidney produced artifacts, such as the existence of podocytes in a ring around the medulla (Fig 7D), although podocytes actually exist only in clusters within glomeruli in the cortex.

To more quantitatively investigate MERFISH's cell type resolving ability with and without computational integration, we compared cell type proportions between Tabula Muris Senis, MERFISH, and the integrated dataset in both mouse liver and kidney (Fig 6E and Fig 7E). Especially for the kidney, MERFISH resulted in cell type proportions that were generally closer to Tabula Muris Senis than the integrated dataset, further highlighting MERFISH's potential as a standalone technology. Interestingly, MERFISH detected far fewer immune cells than scRNA-seq in both organs (Fig 6E and 7E), supporting the hypothesis that scRNA-seq might overestimate immune cell counts and thus may not produce the most accurate reflection of minority cell types (Wu et al. 2019; Denisenko et al. 2020; Ding et al. 2020; Koenitzer et al.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2020; Slyper et al. 2020). This was further highlighted by the fact that MERFISH detected more hepatic stellate cells in mouse liver compared to scRNA-seq (Fig 6E). In combination with the increased measurement sensitivity discussed above, our results anchor MERFISH as a technique that offers equivalent, if not better, overall biological signal compared to scRNA-seq.

Thus, we conclude that with efficient gene panel design, MERFISH cleanly resolves cell types and spatial structure alone, and that computational integration with scRNA-seq reference datasets provide no substantial gain of signal. Although here we opted to use scVI and scANVI, many other computational methods exist, such as tangram (Biancalani et al. 2021) and giotto (Dries et al. 2021), to name a few examples. Further work should be undertaken to systematically compare these various integration methods for different tissue types and spatial measurement technologies.

In light of these findings, we envision MERFISH primarily as a standalone targeted approach for following up on pre-existing scRNA-seq studies. Due to the limited size of the gene panel in a single MERFISH experiment, smart design of the panel from existing scRNA-seq data is necessary. Here we used a single panel of 307 genes to investigate three separate organs – liver, kidney, and pancreas. Future experiments can likely be improved by limiting a gene panel to a single organ of interest. Because the overall statistics of MERFISH measurements are quantitatively comparable to those from scRNA-seq, and because MERFISH data alone appear sufficient to robustly identify cell types and reproduce spatial structure, there appears to be no urgent need for computational integration between the two modalities, at least in the context of cell type identification in liver and kidney.

Methods and Materials

Mouse tissue sample:

The mice in this study were wild type female and harvested at the age of 3 months. The tissues of interest were fresh-frozen, embedded in optimal cutting temperature (OCT) compound and stored at -80℃ until cryosectioning.

RNA quality measurement:

The Qiagen RNeasy Mini Kit was used to isolate RNA from tissue samples and the RIN score was measured for each using the Agilent Tapestation system.

Design of gene panel:

The MERFISH gene panel consisted of 307 genes, requiring 22 bits per barcode and 8 rounds of 3-color hybridization. The gene panel was designed by selecting the top 10 differentially expressed genes in each cell type according to Tabula Muris Senis cell atlas scRNA-seq data (The Tabula Muris Consortium et al. 2020). Initially the list included 424 genes across 19 kidney, 19 liver and 11 pancreas cell types. Genes with fewer than 30 target regions per transcript or higher than an abundance threshold of 800 FPKM (Fragments Per Kilobase of transcript per Million mapped reads) were removed from the list. The final 307 targetable genes possessed a total abundance of ~9000 FPKM.

Cryosectioning, staining and hybridization:

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Frozen OCT-embedded tissue was sectioned at -15℃ to a thickness of 10 μm and mounted onto a functionalized 20 mm coverslip treated with yellow green (YG) fluorescent microspheres. The mounted tissues were then fixed with 4% PFA (Paraformaldehyde) in 1X PBS (Phosphate Buffered Saline), washed with 1X PBS, and stored in 70% ethanol at 4℃ for at least one day and no more than one month before proceeding. Next, the cell boundary was stained using a primary antibody mix followed by a secondary antibody mix. The antibody stained tissue section was post-fixed with 4% PFA in 1X PBS and washed with 1X PBS. The sample was then treated with 30% formamide in 2X SSC (Saline Sodium Citrate, formamide buffer) before the encoding probe hybridization buffer mix (MERFISH library mix) was applied. The sample was incubated for 36-48 hours in a 37℃ cell culture incubator while submerged in the MERFISH library mix.

Gel embedding and tissue clearing:

After hybridization of encoding probes to mRNA transcripts was complete, the sample was washed twice with a formamide buffer and embedded in an acrylamide/bis-acrylamide gel. Polymerization catalyzed by ammonium persulfate and NNN'Tetramethyl-ethylindiamin (TEMED) took approximately 1.5 hours. Once gel formation was confirmed, the sample was rinsed briefly with 2X SSC then incubated at 37℃ in proteinase K supplemented clearing solution (2% Sodium Dodecyl Sulfate (SDS), 0.5% Triton-X 100 in 2X SSC). After 1 day, the tissue became transparent and was ready for imaging.

Imaging:

The gel embedded and cleared sample was washed repeatedly with 2X SSC to reduce autofluorescence from residual SDS in the gel. Then the first set of fluorescent probes (hybridization buffer A1) was pre-hybridized to the cell boundary. After 15 minutes, the sample was incubated in Wash Buffer from the Vizgen Imaging Reagent Kit for 10 minutes and gently assembled into the flow cell (Bioptechs, FCS2). The reagents (Wash buffer, Imaging Buffer, Rinse Buffer, Extinguishing Buffer) and hybridization buffers (up to 8) were loaded into the fluidic system, which was controlled by a user interface on a desktop computer. After priming the system, the flow cell was inserted into the fluidic path and Imaging Buffer was delivered to the sample.

Using a Nikon Eclipse Ti2 inverted microscope equipped with a 10X objective and 405nm laser channel, a low resolution mosaic was constructed from the resulting DAPI signal. Then the region of interest was selected, generating a text file containing a list of position coordinates. This position list, along with the corresponding fluidics recipe configuration file, was inputted to the automated fluidics and imaging control program. Using a 60X oil immersion objective, 8 rounds of 3-color imaging were performed. The cell boundary and DAPI stains were imaged at 7 focal planes on the z axis for each tiled field of view (FOV). Imaging was followed by incubation in Extinguishing Buffer, Rinse Buffer, Hybridization Buffer (corresponding to the subsequent round of MERFISH readout probes), Wash Buffer, and Imaging Buffer. In each round, fluorescent probes were imaged at 7 focal planes on the z axis using 749 nm, 638 nm and 546 nm laser illumination. In addition, a single image of the fiducial beads was acquired at each FOV using 477 nm illumination. The resulting raw images stack was saved in various .dax files. As a point of reference, the raw images from a single run of MERFISH for a ~1cm^2 tissue sample contain about 1 TB of data.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Image Analysis:

To process the raw image files from the MERFISH experiments, we utilized the MERlin image analysis pipeline (Emanuel et al. 2020). Here we briefly describe the pipeline.

Initially, image stacks obtained from different MERFISH rounds are aligned to correct for microscope stage drift by maximizing their cross-correlation with fiducial bead images. The aligned images are then passed through a high-pass filter to remove background noise, as well as a deconvolution process to clarify the RNA spots in preparation for bit-calling i.e decoding a bit as 1 or 0 based on fluorescence detection (Guo et al. 2019).

Individual RNA molecule barcodes are then decoded from using a pixel-based decoding algorithm that utilizes filters based on area and intensity. Then, an adaptive barcoding scheme is used to correct misidentified barcodes that do not correspond to any barcodes in the codebook. The level of correction can be set to a user-specified final misidentification rate. Here, we set the misidentification rate to 5%.

Then, cells are segmented by utilizing the information from the nuclear DAPI and cell membrane antibody stains. The decoded RNA molecules are partitioned into individual cells to generate single-cell RNA count matrices. Note that the whole pipeline is run for each imaging field-of-view and then tiled over the entire sample imaging area of ~1cm^2.

Comparison with Tabula Muris Senis reference cell atlas:

Comparisons of bulk and single-cell MERFISH results were done with the bulk and droplet-based single-cell RNA-seq results in Tabula Muris Senis (Schaum et al. 2020, The Tabula Muris Consortium et al. 2020). To remove the effects of age, we subsetted both reference datasets to mice that were three months old.

Single-cell bioinformatic analysis:

Single-cell MERFISH and RNA-seq results were analyzed using *scanpy* (Wolf, Angerer, and Theis 2018). All results were preprocessed by filtering out cells with low counts as mentioned in the main text, normalizing to a total count of 10,000 transcripts per cell, and log-transforming.

For single-cell analysis of MERFISH results alone, principal components were first computed using *scanpy*'s *tl.pca()* function with default settings. Then, neighborhood graphs were computed using *scanpy*'s *pp.neighbors()* function. UMAP plots and leiden clustering were calculated using *tl.umap()* and *tl.leiden()*. To annotate the computed clusters, we examined the top differentially expressed genes in each cluster using the *tl.rank_genes_groups()* function and compared with known marker genes for the various cell types.

For integrated analysis, the Tabula Muris Senis scRNA-seq data were subsetted to retain the same 307 genes as the MERFISH data. For mouse liver, the three-month data did not possess many non-hepatocyte cells, so we included non-hepatocyte data from the one-month data to increase cell diversity. For mouse kidney, we used only three-month data. The MERFISH and scRNA-seq datasets were concatenated, and then we trained a joint model using scVI's default settings. The resulting neighborhood graph and UMAP plots were computed on the scVI latent

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

space. For cell type annotation transfer, scANVI was run using 20 epochs and 100 samples per label.

The liver cell type annotations in Tabula Muris Senis were revised after consultation with tissue experts. Furthermore, the kidney cell type annotations were overly specific, so we coarse-grained them into broader categories.

Data Availability

Raw MERFISH data are available for download on AWS s3 at s3://czb-tabula-muris-senis/spatial-transcriptomics/MERFISH-data/. The gene panel codebook, RIN scores, and processed data including decoded MERFISH transcript information, cell boundaries, bulk RNA statistics, and single cell RNA counts with cell type annotations are available for download from figshare at https://figshare.com/projects/MERFISH_mouse_comparison_study/134213. A Github repository containing the code needed to reproduce the figures is available at https://github.com/czbiohub/MERFISH-mouse-comparison.

Acknowledgements

We thank Patrick Neuheofer for providing the frozen OCT embedded mouse tissue samples used in this work, as well as Gabriel Loeb for helpful discussions regarding interpretation of the MERFISH kidney data. We are also grateful to Sandra Schmid, Rafael Gómez-Sjöberg, and Alejandro Granados for thoughtful comments on this manuscript. This work was supported by the Chan Zuckerberg Biohub.

References

Biancalani, Tommaso, Gabriele Scalia, Lorenzo Buffoni, Raghav Avasthi, Ziqing Lu, Aman Sanger, Neriman Tokcan, et al. 2021. “Deep Learning and Alignment of Spatially Resolved Single-Cell Transcriptomes with Tangram.” *Nature Methods*, October, 1–11. https://doi.org/10.1038/s41592-021-01264-7.

Brunskill, Eric W., Kylie Georgas, Bree Rumballe, Melissa H. Little, and S. Steven Potter. 2011. “Defining the Molecular Character of the Developing and Adult Kidney Podocyte.” *PLOS ONE* 6 (9): e24640. https://doi.org/10.1371/journal.pone.0024640.

Chen, Kok Hao, Alistair N. Boettiger, Jeffrey R. Moffitt, Siyuan Wang, and Xiaowei Zhuang. 2015. “Spatially Resolved, Highly Multiplexed RNA Profiling in Single Cells.” *Science* 348 (6233). https://doi.org/10.1126/science.aaa6090.

Denisenko, Elena, Belinda B. Guo, Matthew Jones, Rui Hou, Leanne de Kock, Timo Lassmann, Daniel Poppe, et al. 2020. “Systematic Assessment of Tissue Dissociation and Storage Biases in Single-Cell and Single-Nucleus RNA-Seq Workflows.” *Genome Biology* 21 (1): 130. https://doi.org/10.1186/s13059-020-02048-6.

Ding, Jiarui, Xian Adiconis, Sean K. Simmons, Monika S. Kowalczyk, Cynthia C. Hession, Nemanja D. Marjanovic, Travis K. Hughes, et al. 2020. “Systematic Comparison of Single-Cell and Single-Nucleus RNA-Sequencing Methods.” *Nature Biotechnology* 38

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(6): 737–46. https://doi.org/10.1038/s41587-020-0465-8.

Dries, Ruben, Qian Zhu, Rui Dong, Chee-Huat Linus Eng, Huipeng Li, Kan Liu, Yuntian Fu, et al. 2021. "Giotto: A Toolbox for Integrative Analysis and Visualization of Spatial Expression Data." *Genome Biology* 22 (1): 78. https://doi.org/10.1186/s13059-021-02286-2.

Emanuel, George, seichhorn, Hazen Babcock, leonardosepulveda, and timblosser. 2020. *ZhuangLab/MERlin: MERlin v0.1.6*. Zenodo. https://doi.org/10.5281/zenodo.3758540.

Eng, Chee-Huat Linus, Michael Lawson, Qian Zhu, Ruben Dries, Noushin Koulena, Yodai Takei, Jina Yun, et al. 2019. "Transcriptome-Scale Super-Resolved Imaging in Tissues by RNA SeqFISH+." *Nature* 568 (7751): 235–39. https://doi.org/10.1038/s41586-019-1049-y.

Guo, Min, Yue Li, Yijun Su, Talley Lambert, Damian Dalle Nogare, Mark W. Moyle, Leighton H. Duncan, et al. 2019. "Accelerating Iterative Deconvolution and Multiview Fusion by Orders of Magnitude." https://doi.org/10.1101/647370.

Koenitzer, Jeffrey R., Haojia Wu, Jeffrey J. Atkinson, Steven L. Brody, and Benjamin D. Humphreys. 2020. "Single-Nucleus RNA-Sequencing Profiling of Mouse Lung. Reduced Dissociation Bias and Improved Rare Cell-Type Detection Compared with Single-Cell RNA Sequencing." *American Journal of Respiratory Cell and Molecular Biology* 63 (6): 739–47. https://doi.org/10.1165/rcmb.2020-0095MA.

Liu, Miao, Yanfang Lu, Bing Yang, Yanbo Chen, Jonathan S. D. Radda, Mengwei Hu, Samuel G. Katz, and Siyuan Wang. 2020. "Multiplexed Imaging of Nucleome Architectures in Single Cells of Mammalian Tissue." *Nature Communications* 11 (1): 2907. https://doi.org/10.1038/s41467-020-16732-5.

Lu, Yanfang, Miao Liu, Jennifer Yang, Sherman M. Weissman, Xinghua Pan, Samuel G. Katz, and Siyuan Wang. 2021. "Spatial Transcriptome Profiling by MERFISH Reveals Fetal Liver Hematopoietic Stem Cell Niche Architecture." *Cell Discovery* 7 (1): 1–17. https://doi.org/10.1038/s41421-021-00266-1.

Marx, Vivien. 2021. "Method of the Year: Spatially Resolved Transcriptomics." *Nature Methods* 18 (1): 9–14. https://doi.org/10.1038/s41592-020-01033-y.

McInnes, Leland, John Healy, Nathaniel Saul, and Lukas Großberger. 2018. "UMAP: Uniform Manifold Approximation and Projection." *Journal of Open Source Software* 3 (29): 861. https://doi.org/10.21105/joss.00861.

Merritt, Christopher R., Giang T. Ong, Sarah E. Church, Kristi Barker, Patrick Danaher, Gary Geiss, Margaret Hoang, et al. 2020. "Multiplex Digital Spatial Profiling of Proteins and RNA in Fixed Tissue." *Nature Biotechnology* 38 (5): 586–99. https://doi.org/10.1038/s41587-020-0472-9.

Moffitt, Jeffrey R., Dhananjay Bambah-Mukku, Stephen W. Eichhorn, Eric Vaughn, Karthik Shekhar, Julio D. Perez, Nimrod D. Rubinstein, et al. 2018. "Molecular, Spatial, and Functional Single-Cell Profiling of the Hypothalamic Preoptic Region." *Science* 362 (6416). https://doi.org/10.1126/science.aau5324.

Moffitt, Jeffrey R., Junjie Hao, Dhananjay Bambah-Mukku, Tian Lu, Catherine Dulac, and Xiaowei Zhuang. 2016. "High-Performance Multiplexed Fluorescence in Situ Hybridization in Culture and Tissue with Matrix Imprinting and Clearing." *Proceedings of the National Academy of Sciences* 113 (50): 14456–61. https://doi.org/10.1073/pnas.1617699113.

Moffitt, Jeffrey R., Junjie Hao, Guiping Wang, Kok Hao Chen, Hazen P. Babcock, and Xiaowei Zhuang. 2016. "High-Throughput Single-Cell Gene-Expression Profiling with Multiplexed Error-Robust Fluorescence in Situ Hybridization." *Proceedings of the National Academy of Sciences* 113 (39): 11046–51. https://doi.org/10.1073/pnas.1612826113.

Palla, Giovanni, Hannah Spitzer, Michal Klein, David Fischer, Anna Christina Schaar, Louis Benedikt Kuemmerle, Sergei Rybakov, et al. 2022. "Squidpy: A Scalable Framework for Spatial Omics Analysis." *Nature Methods* 19 (2): 171–78.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://doi.org/10.1038/s41592-021-01358-2.

Petukhov, Viktor, Rosalind J. Xu, Ruslan A. Soldatov, Paolo Cadinu, Konstantin Khodosevich, Jeffrey R. Moffitt, and Peter V. Kharchenko. 2021. "Cell Segmentation in Imaging-Based Spatial Transcriptomics." *Nature Biotechnology*, October, 1–10. https://doi.org/10.1038/s41587-021-01044-w.

Schaum, Nicholas, Benoit Lehallier, Oliver Hahn, Róbert Pálovics, Shayan Hosseinzadeh, Song E. Lee, Rene Sit, et al. 2020. "Ageing Hallmarks Exhibit Organ-Specific Temporal Signatures." *Nature* 583 (7817): 596–602. https://doi.org/10.1038/s41586-020-2499-y.

Schroeder, Andreas, Odilo Mueller, Susanne Stocker, Ruediger Salowsky, Michael Leiber, Marcus Gassmann, Samar Lightfoot, Wolfram Menzel, Martin Granzow, and Thomas Ragg. 2006. "The RIN: An RNA Integrity Number for Assigning Integrity Values to RNA Measurements." *BMC Molecular Biology* 7 (1): 3. https://doi.org/10.1186/1471-2199-7-3.

Slyper, Michal, Caroline B. M. Porter, Orr Ashenberg, Julia Waldman, Eugene Drokhlyansky, Isaac Wakiro, Christopher Smillie, et al. 2020. "A Single-Cell and Single-Nucleus RNA-Seq Toolbox for Fresh and Frozen Human Tumors." *Nature Medicine* 26 (5): 792–802. https://doi.org/10.1038/s41591-020-0844-1.

Ståhl, Patrik L., Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello, et al. 2016. "Visualization and Analysis of Gene Expression in Tissue Sections by Spatial Transcriptomics." *Science* 353 (6294): 78–82. https://doi.org/10.1126/science.aaf2403.

Stickels, Robert R., Evan Murray, Pawan Kumar, Jilong Li, Jamie L. Marshall, Daniela J. Di Bella, Paola Arlotta, Evan Z. Macosko, and Fei Chen. 2021. "Highly Sensitive Spatial Transcriptomics at Near-Cellular Resolution with Slide-SeqV2." *Nature Biotechnology* 39 (3): 313–19. https://doi.org/10.1038/s41587-020-0739-1.

The Tabula Muris Consortium, Nicole Almanzar, Jane Antony, Ankit S. Baghel, Isaac Bakerman, Ishita Bansal, Ben A. Barres, et al. 2020. "A Single-Cell Transcriptomic Atlas Characterizes Ageing Tissues in the Mouse." *Nature* 583 (7817): 590–95. https://doi.org/10.1038/s41586-020-2496-1.

Traag, V. A., L. Waltman, and N. J. van Eck. 2019. "From Louvain to Leiden: Guaranteeing Well-Connected Communities." *Scientific Reports* 9 (1): 5233. https://doi.org/10.1038/s41598-019-41695-z.

Wang, Guiping, Cheen-Euong Ang, Jean Fan, Andrew Wang, Jeffrey R. Moffitt, and Xiaowei Zhuang. 2020. "Spatial Organization of the Transcriptome in Individual Neurons." https://doi.org/10.1101/2020.12.07.414060.

Wang, Xiao, William E. Allen, Matthew A. Wright, Emily L. Sylwestrak, Nikolay Samusik, Sam Vesuna, Kathryn Evans, et al. 2018. "Three-Dimensional Intact-Tissue Sequencing of Single-Cell Transcriptional States." *Science* 361 (6400). https://doi.org/10.1126/science.aat5691.

Wolf, F. Alexander, Philipp Angerer, and Fabian J. Theis. 2018. "SCANPY: Large-Scale Single-Cell Gene Expression Data Analysis." *Genome Biology* 19 (1): 15. https://doi.org/10.1186/s13059-017-1382-0.

Wu, Haojia, Yuhei Kirita, Erinn L. Donnelly, and Benjamin D. Humphreys. 2019. "Advantages of Single-Nucleus over Single-Cell RNA Sequencing of Adult Kidney: Rare Cell Types and Novel Cell States Revealed in Fibrosis." *Journal of the American Society of Nephrology* 30 (1): 23–32. https://doi.org/10.1681/ASN.2018090912.

Xia, Chenglong, Jean Fan, George Emanuel, Junjie Hao, and Xiaowei Zhuang. 2019. "Spatial Transcriptome Profiling by MERFISH Reveals Subcellular RNA Compartmentalization and Cell Cycle-Dependent Gene Expression." *Proceedings of the National Academy of Sciences* 116 (39): 19490–99. https://doi.org/10.1073/pnas.1912459116.

Xu, Chenling, Romain Lopez, Edouard Mehlman, Jeffrey Regier, Michael I. Jordan, and Nir Yosef. 2021. "Probabilistic Harmonization and Annotation of Single-Cell Transcriptomics

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Data with Deep Generative Models." *Molecular Systems Biology* 17 (1): e9620. https://doi.org/10.15252/msb.20209620.

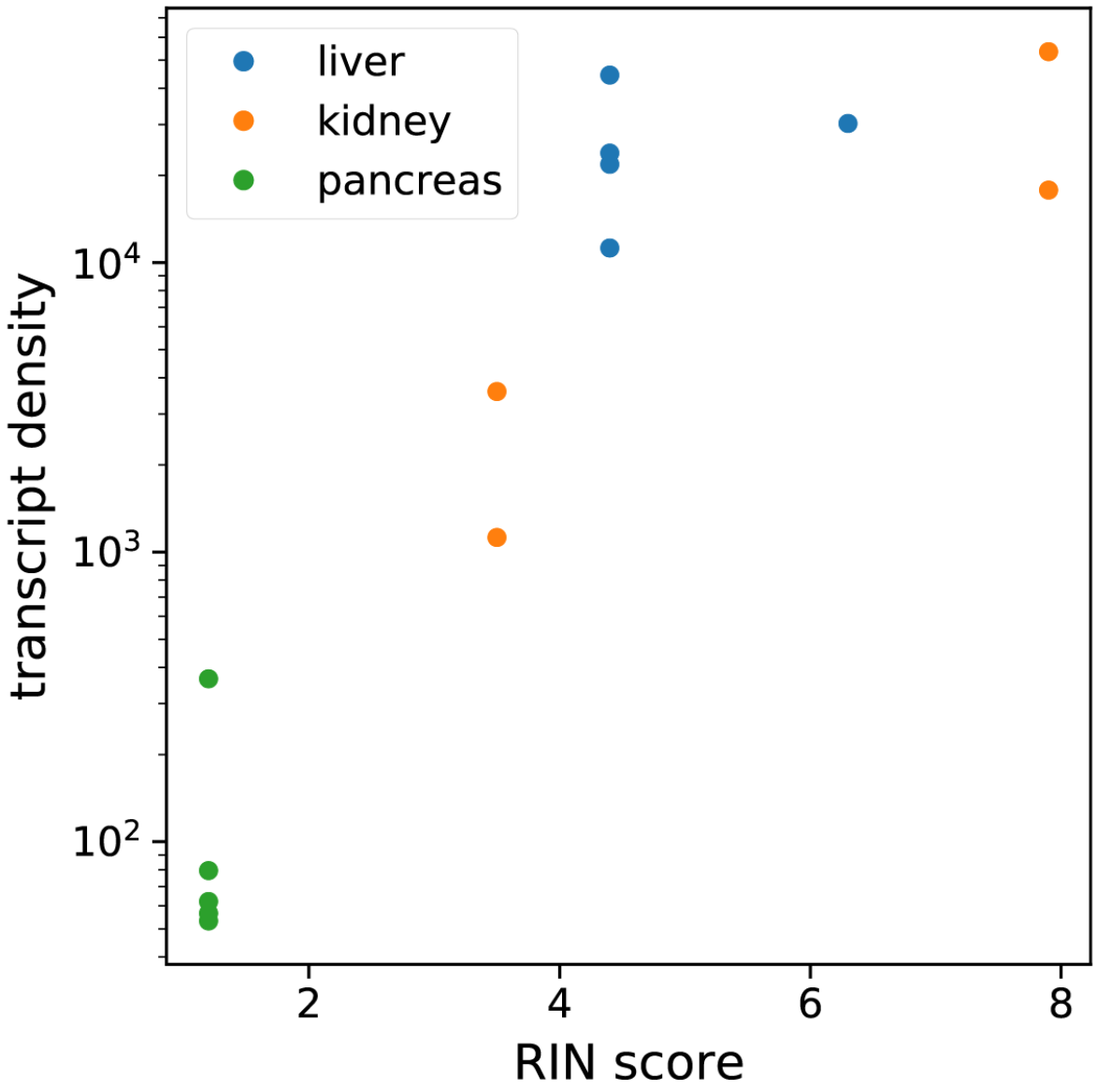
Zhang, Meng, Stephen W. Eichhorn, Brian Zingg, Zizhen Yao, Hongkui Zeng, Hongwei Dong, and Xiaowei Zhuang. 2020. "Molecular, Spatial and Projection Diversity of Neurons in Primary Motor Cortex Revealed by in Situ Single-Cell Transcriptomics." https://doi.org/10.1101/2020.06.04.105700.

Supplementary Text

S1 RNA integrity analysis

To investigate the relationship between MERFISH signal quality and RNA integrity in the tissue, we measured the RNA integrity number (RIN) (Schroeder et al. 2006) in each of our tissue samples (Methods and Materials). As a metric for MERFISH signal quality, we calculated the *transcript density*, defined as the total number of detected RNA transcripts divided by the number of fields of view imaged using the microscope. SI Fig 1 shows the results of this analysis. There was a clear correlation between RIN score and transcript density up until a RIN score of about 5, after which the transcript density plateaued. Thus, we opted to exclude datasets before the plateau and did not consider the kidney and pancreas MERFISH datasets with RIN score below 4 in this paper.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



SI Figure 1 - Investigation of RNA tissue quality on MERFISH measurements. Average transcript density is shown as a function of tissue RIN score for each tissue sample.

S2 Investigation of molecular crowding in MERFISH measurements

To assess the impact of molecular crowding of RNA transcripts obscuring fluorescent signal, we reasoned that such an effect would be most apparent in the most abundant genes in the gene panel. SI Fig 2A and SI Fig 2B show distributions of the median count of each gene in the MERFISH panel, calculated from cells in scRNA-seq that registered nonzero counts for that gene. Although most genes possess median counts fewer than 5, there are some genes with

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

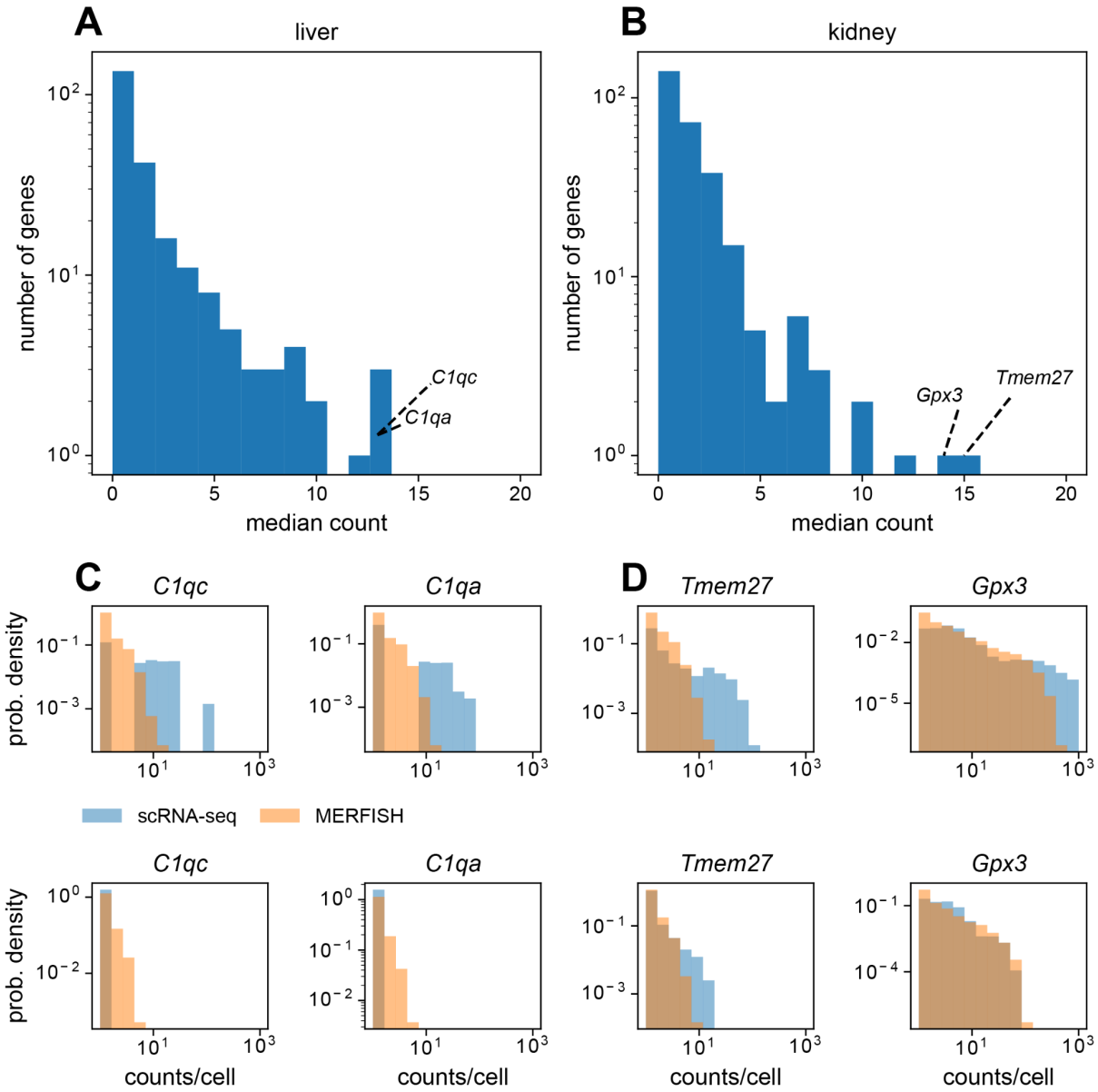
median counts over 10. The top two most abundant genes in the MERFISH panel from scRNA-seq were *C1qc* and *C1qa* in mouse liver, and *Gpx3 and Tmem27* in mouse kidney.

We then hypothesized that molecular crowding should only occur in the regime of high transcript count. Thus, for cells with high overall numbers of transcripts, MERFISH measurements may become unreliable due to the high density of fluorescent spots making reliable single-molecule detection difficult. In contrast, for cells with low overall numbers of transcripts, MERFISH measurements should be more reliable.

To investigate this hypothesis, we examined the distribution of detected counts per cell for these two most abundant genes in mouse liver and kidney, between MERFISH and scRNA-seq. SI Fig 2C and SI Fig 2D show the results, where we split the analysis by overall transcript count. The top row shows these distributions for cells with total transcript count over 100, whereas the bottom row shows these distributions for cells with total transcript count under 100. In both liver and kidney, we notice that the MERFISH and scRNA-seq distributions differ substantially for higher count values in cells with over 100 total transcripts. In contrast, the distributions are much more similar in cells with under 100 total transcripts.

Thus, we concluded that our MERFISH measurements were entering the molecular crowding regime, causing the quantitative agreement in statistics between MERFISH and scRNA-seq to diverge for cells with high overall RNA count number.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



SI Figure 2 - Investigation of molecular crowding in MERFISH measurements. (A, B) Distribution of median counts per gene across cells with nonzero counts in mouse liver (A) and kidney (C). Text annotations highlight the top two most abundant genes in each tissue. (C, D) Distribution of counts per cell for top two most abundant genes in mouse liver (C) and (D) kidney, in cells that registered over 100 total counts (top), or in cells that registered under 100 total counts (bottom).

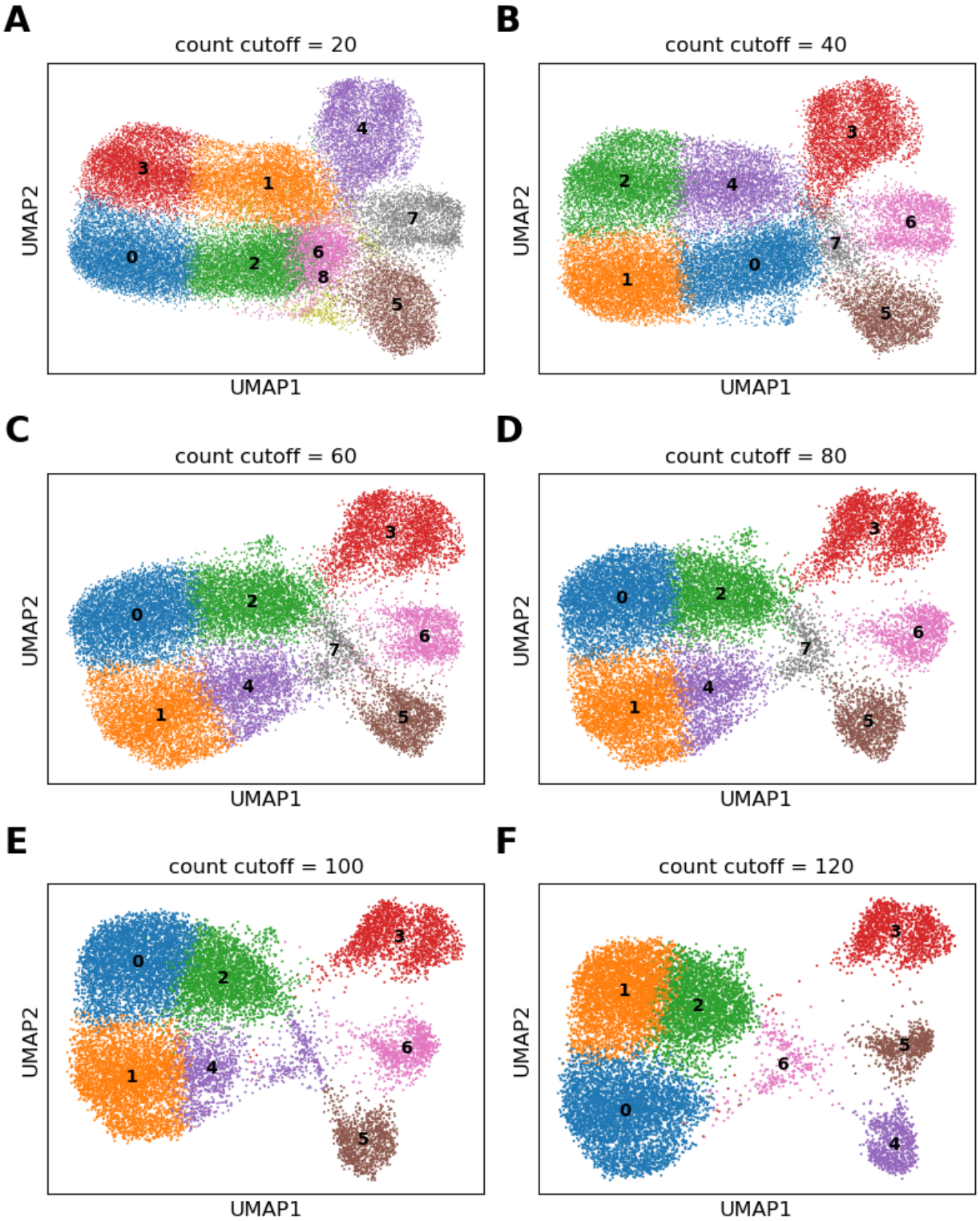
S3 Single cell analysis in mouse liver

Before conducting single-cell analysis of our MERFISH liver results, we further filtered the single-cell RNA count data to strengthen data quality. To do so, we created

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

data subsets with increasing numbers of minimum total RNA counts per cell and examined the resulting UMAP plots with leiden clusters (SI Fig 3). At lower values of this minimum cutoff, the various clusters were not well separated, indicating the presence of low-quality single cell that were introducing noise into the analysis. At a cutoff of 80 total counts per cell, the clusters began to separate more clearly. Finally, with higher cutoff values of 100 and above, there was no substantial increase in clustering quality. Thus we chose a filtering cutoff of 80 total RNA counts per cell for the analysis used in the main text.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



SI Figure 3 - Thresholding of MERFISH mouse liver data for single cell analysis. (A-F) UMAP plot of MERFISH mouse liver dataset used in Fig 6 for various threshold values of total RNA count per cell. For the main text, a total RNA count cutoff of 80 per cell was chosen.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

S4 Quantifying co-localization of periportal and pericentral endothelial cells and hepatocytes in mouse liver

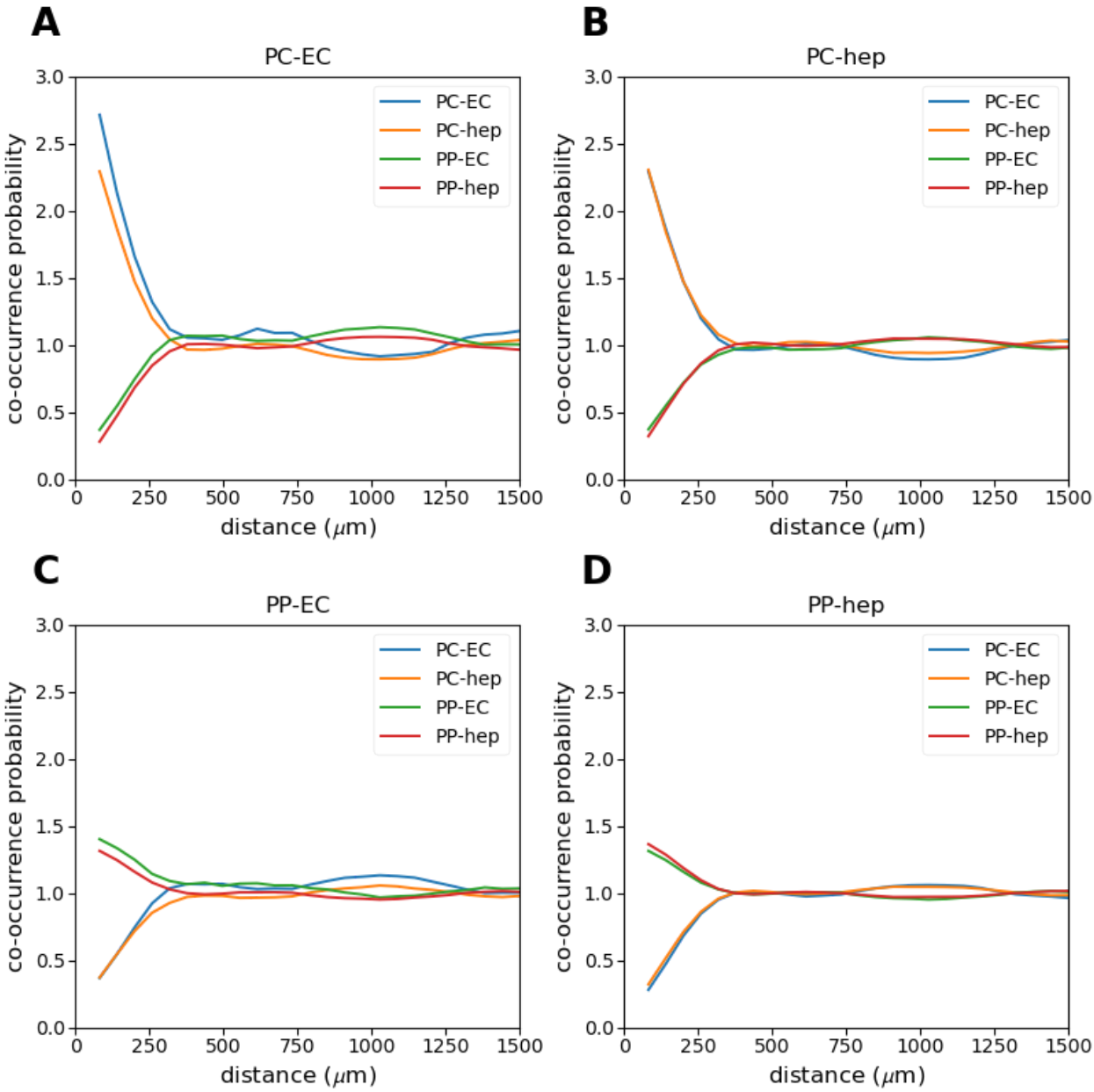
To quantify the co-localization of periportal and pericentral endothelial cells and hepatocytes (Fig 6B inset), we computed the co-occurrence probability of finding each cell type. To do so, we calculated the cluster co-occurrence ratio R provided in the *squidpy* package with the *squidpy.gr.co_occurrence* function (Palla et al. 2022). Briefly, the ratio describes the normalized probability of locating a cell type i in a radius of size d, conditioned on the existence of a cell type j:

$$R(d) = \frac{P(i|j,d)}{P(i,d)}$$

The ratio is computed at varying length scales d. Thus, higher values of R correspond to higher likelihoods of co-localization, and lower values to lower likelihoods of co-localization. In the case of spatial independence of two cell types, the ratio then takes a value of one.

SI Figure 4 shows this co-occurrence probability ratio R for varying length scales d for pairs of periportal and pericentral endothelial cells and hepatocytes. We notice that for small length scales (< 500 microns), pericentral endothelial cells and hepatocytes tend to co-localize with each other with $R > 1$ (SI Fig 4A and 4B). This also holds true for periportal endothelial cells and hepatocytes (SI Fig 4C and 4D). In contrast, periportal and pericentral cells do not co-localize, with $R < 1$. At larger length scales (> 500 microns), the co-occurrence probabilities trend to $R = 1$, indicating no strong co-localization effects.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



SI Figure 4 - Co-occurrence probability analysis of periportal and pericentral hepatocytes and endothelial cells. (A-D) Co-occurrence probability of (A) pericentral endothelial cells, (B) pericentral hepatocytes, (C) periportal endothelial cells, and (D) periportal hepatocytes with other periportal or pericentral endothelial cells or hepatocytes. Cell type abbreviations are as follows: 'PP-hep': 'periportal hepatocyte', 'PC-hep': 'pericentral hepatocyte', 'PP-EC': 'periportal endothelial cell', 'PC-EC': 'pericentral endothelial cell'.

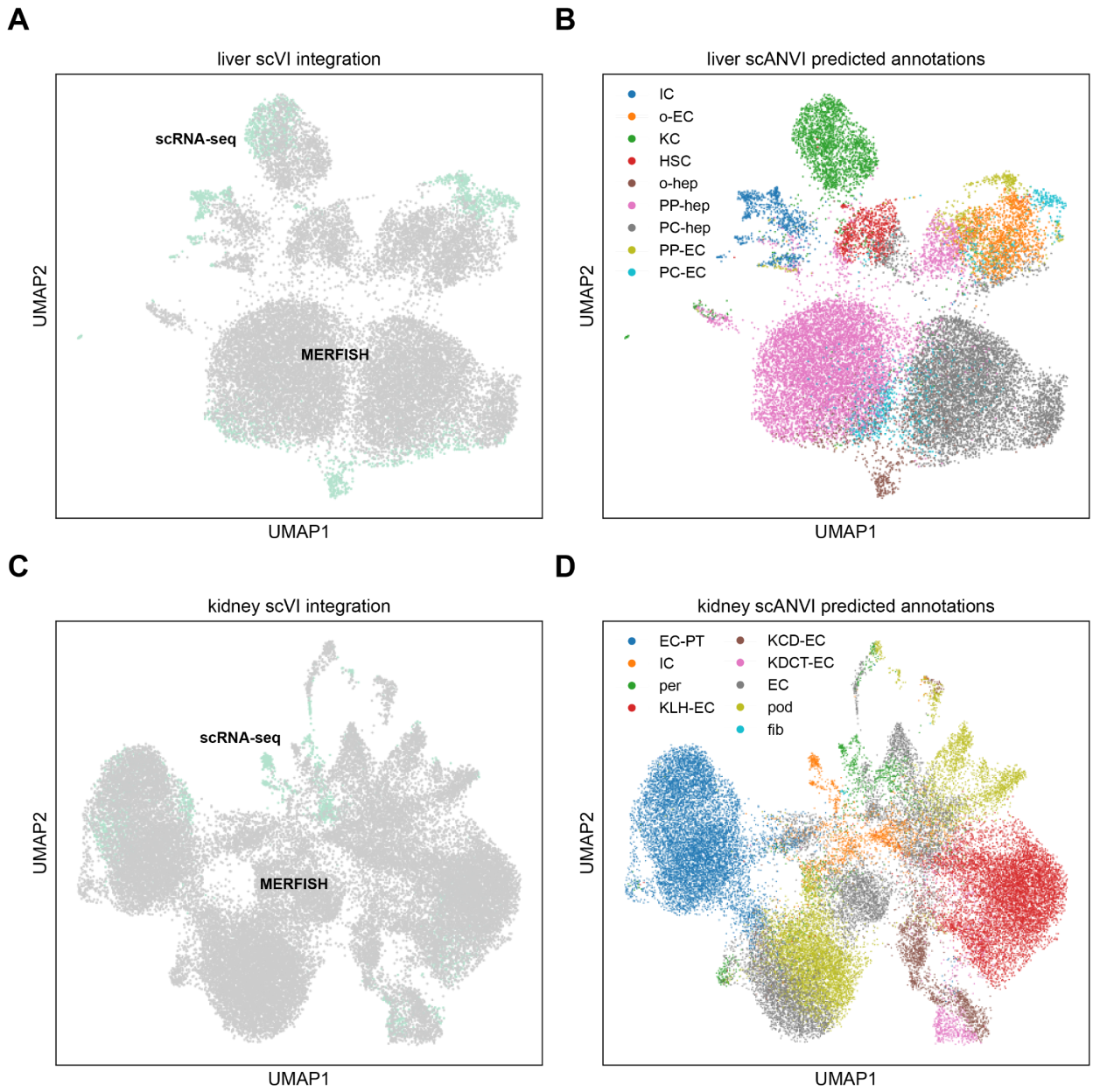
bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

S5 Integration of MERFISH data using scVI and scANVI

SI Figure 5A shows the results of the scVI integration of the MERFISH liver data with Tabula Muris Senis. Surprisingly, the two technologies were not particularly well mixed, suggesting that the underlying statistics of the two modalities are systematically different enough that constructing a good joint embedding may be difficult (SI Fig 5A). SI Figure 5B shows the same UMAP plot with colors according to predicted cell type labels from scANVI. In general, the scANVI joint UMAP plot was of similar quality to the MERFISH-only results. However, some clusters were noisy, particularly the pericentral endothelial cell cluster (SI Fig 5B, cyan), which was highly entangled with the pericentral and periportal hepatocyte clusters (SI Fig 5B, gray, pink).

SI Figure 5C shows the same scVI integration for mouse kidney. Just like the liver data, the kidney data failed to exhibit good mixing of the two modalities. SI Figure 5D shows the same combined integrated kidney UMAP with colors given by cell types predicted using scANVI. Unlike the liver, the integrated kidney combined UMAP plot (SI Fig 5D) was much noisier than the MERFISH counterpart (Fig 7A).

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



SI Figure 5 - Integration of MERFISH mouse liver and kidney data with Tabula Muris Senis scRNA-seq using scVI and scANVI. (A) UMAP plot of MERFISH liver data (gray) with Tabula Muris Senis (teal) using scVI. (B) UMAP plot of liver integration in (A) with cell type annotations predicted using scANVI. (C) UMAP plot of MERFISH kidney data (gray) with Tabula Muris Senis (teal) using scVI. (D) UMAP plot of kidney integration in (C) with cell type annotations predicted using scANVI. Cell type abbreviations in (B) are as follows: 'IC': 'immune cell', 'o-EC': 'other endothelial cell', 'KC': 'Kupffer cell', 'HSC': 'hepatic stellate cell', 'o-hep': 'other hepatocyte', 'PP-hep': 'periportal hepatocyte', 'PC-hep': 'pericentral hepatocyte', 'PP-EC': 'periportal endothelial cell', 'PC-EC': 'pericentral endothelial cell'. Cell type abbreviations in (D) are as follows: 'EC-PT': 'epithelial cell of proximal tubule', 'IC': 'immune cell', 'per': 'pericyte', 'KLH-EC': 'kidney loop of

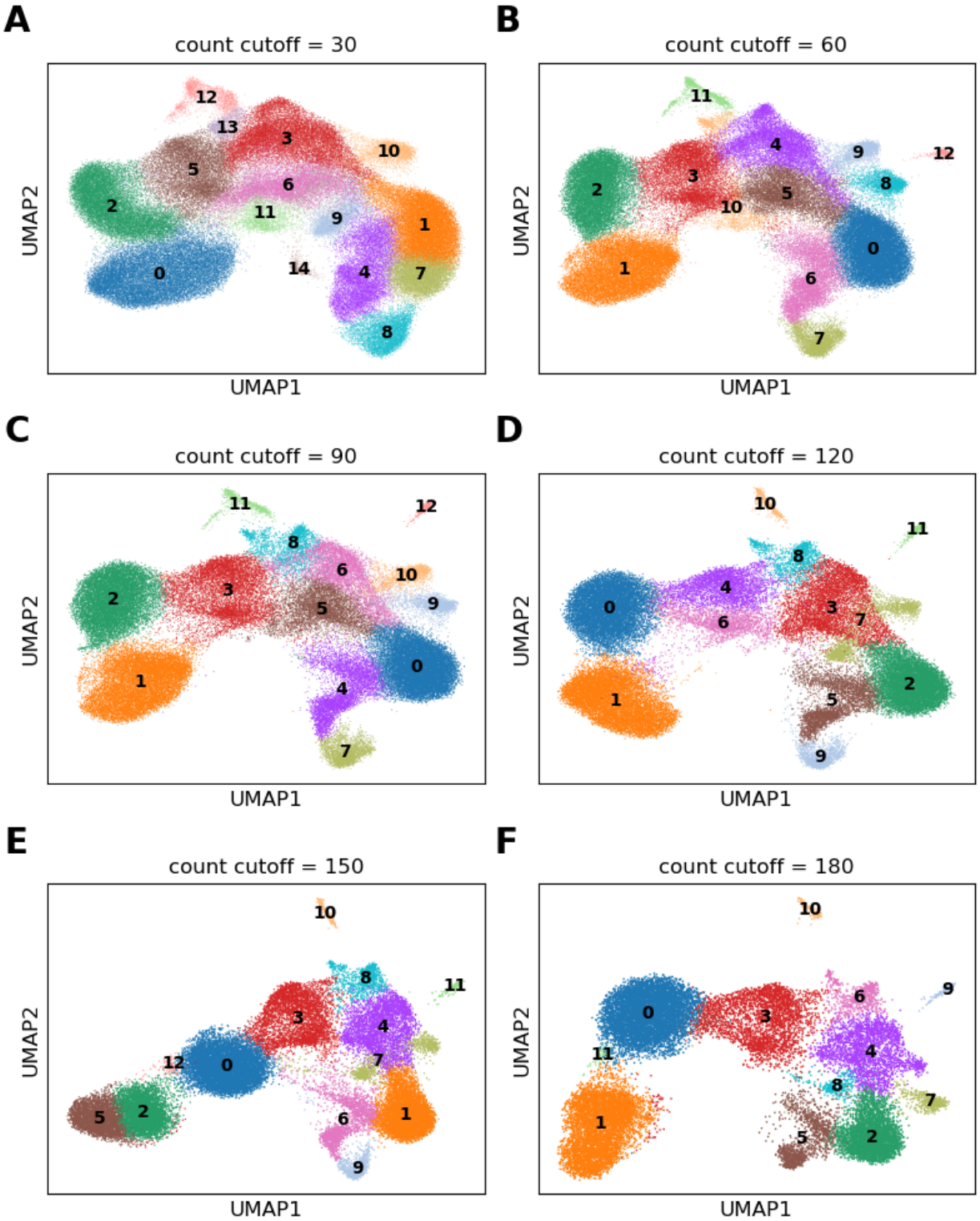
bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Henle epithelial cell', 'KCD-EC': 'kidney collecting duct epithelial cell', 'KDCT-EC': 'kidney distal convoluted tubule epithelial cell', 'EC': 'endothelial cell', 'pod': 'podocyte', 'fib': 'fibroblast'.

S6 Single cell analysis in mouse kidney

Before conducting single-cell analysis of our MERFISH kidney results, we further filtered the single-cell RNA count data to strengthen data quality. To do so, we created data subsets with increasing numbers of minimum total RNA counts per cell and examined the resulting UMAP plots with leiden clusters (SI Fig 4). At lower values of this minimum cutoff, the various clusters were not well separated, indicating the presence of low-quality single cells that were introducing noise into the analysis. At a cutoff of 150 total counts per cell, the clusters began to separate more clearly. Finally, with a higher cutoff value of 180, there was no substantial increase in clustering quality. Thus we chose a filtering cutoff of 150 total RNA counts per cell for the analysis used in the main text.

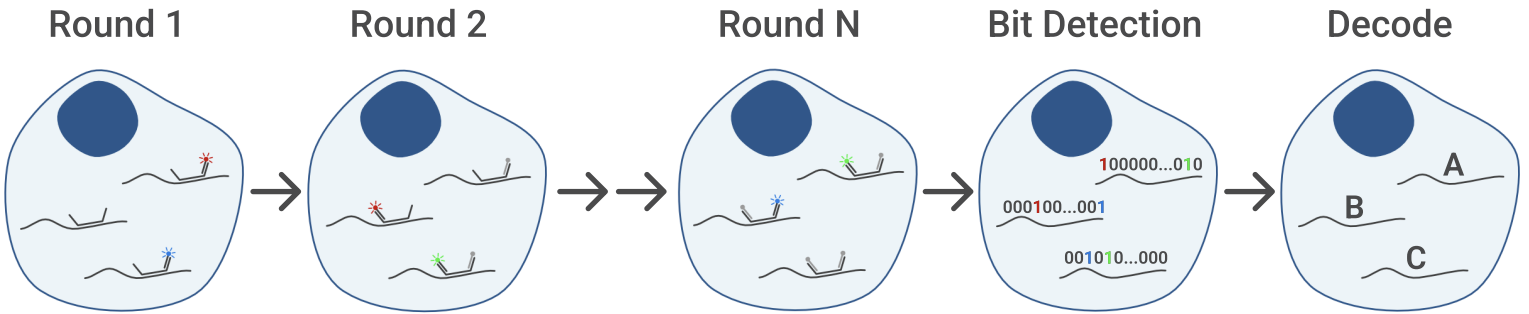
bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



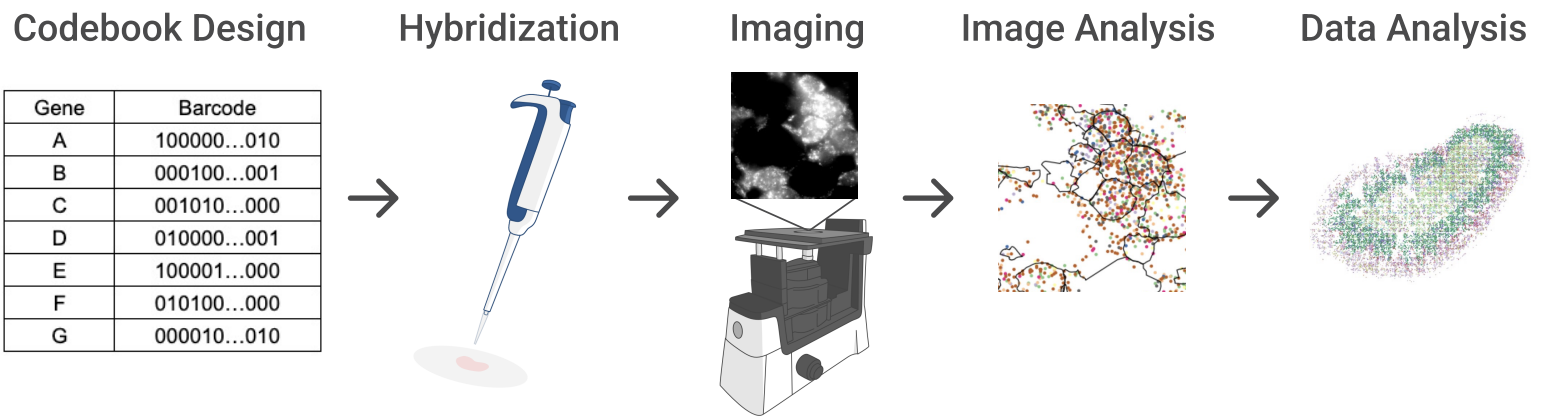
SI Figure 6 - Thresholding of MERFISH mouse kidney data for single cell analysis. (A-F) UMAP plot of MERFISH mouse kidney dataset used in Fig 6 for various threshold values of total RNA count per cell. For the main text, a total RNA count cutoff of 150 per cell was chosen.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

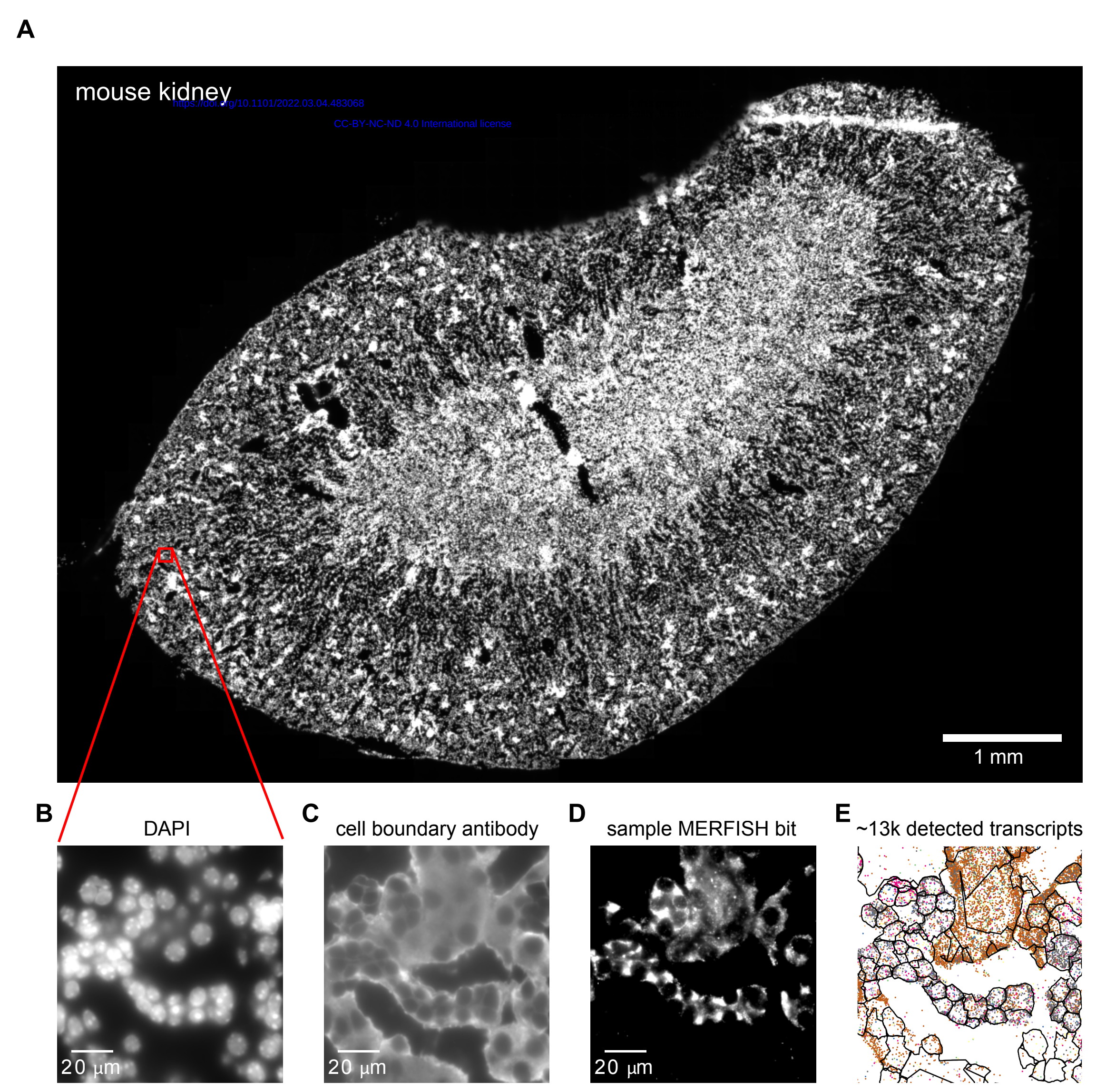
A

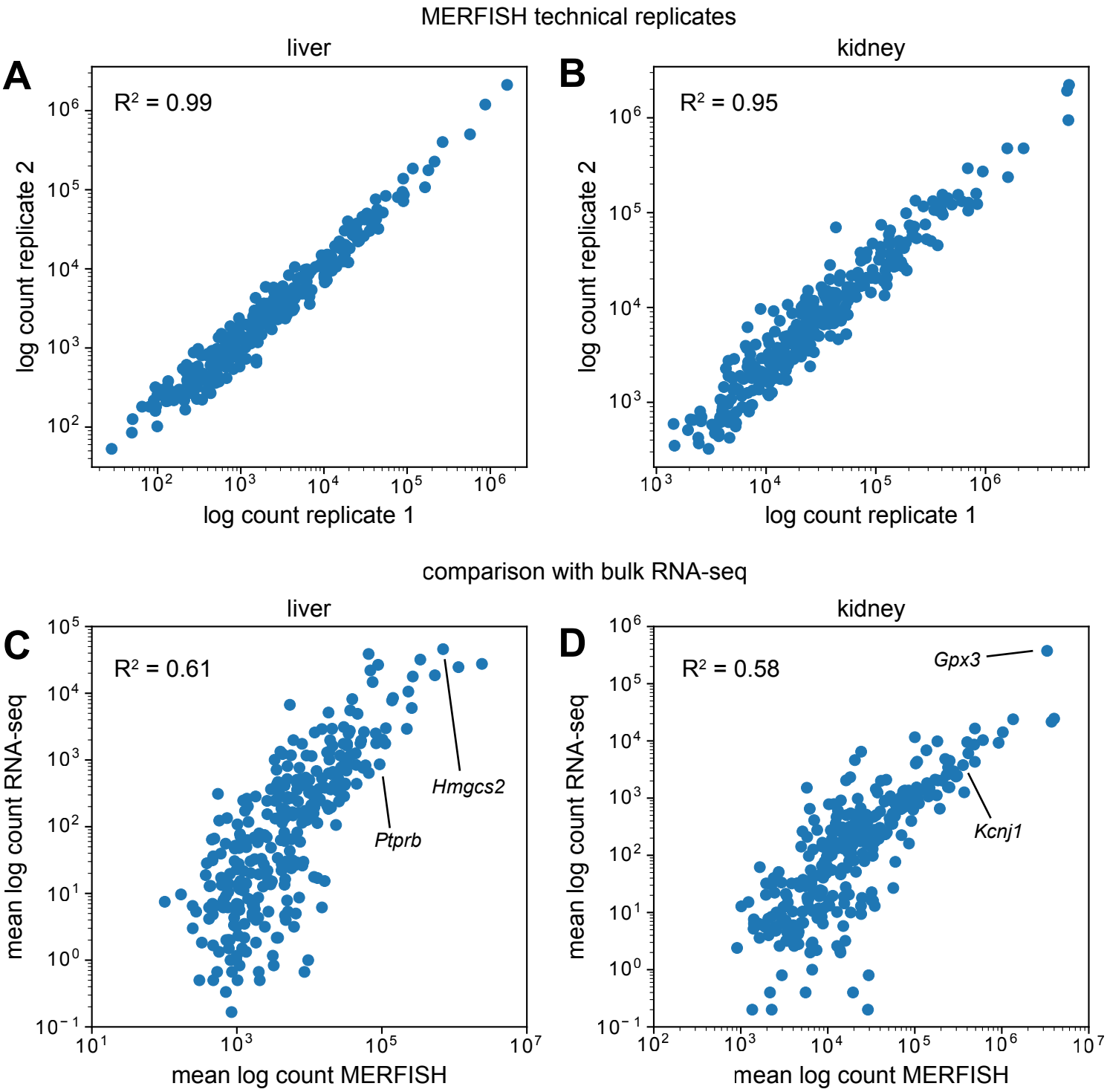


B

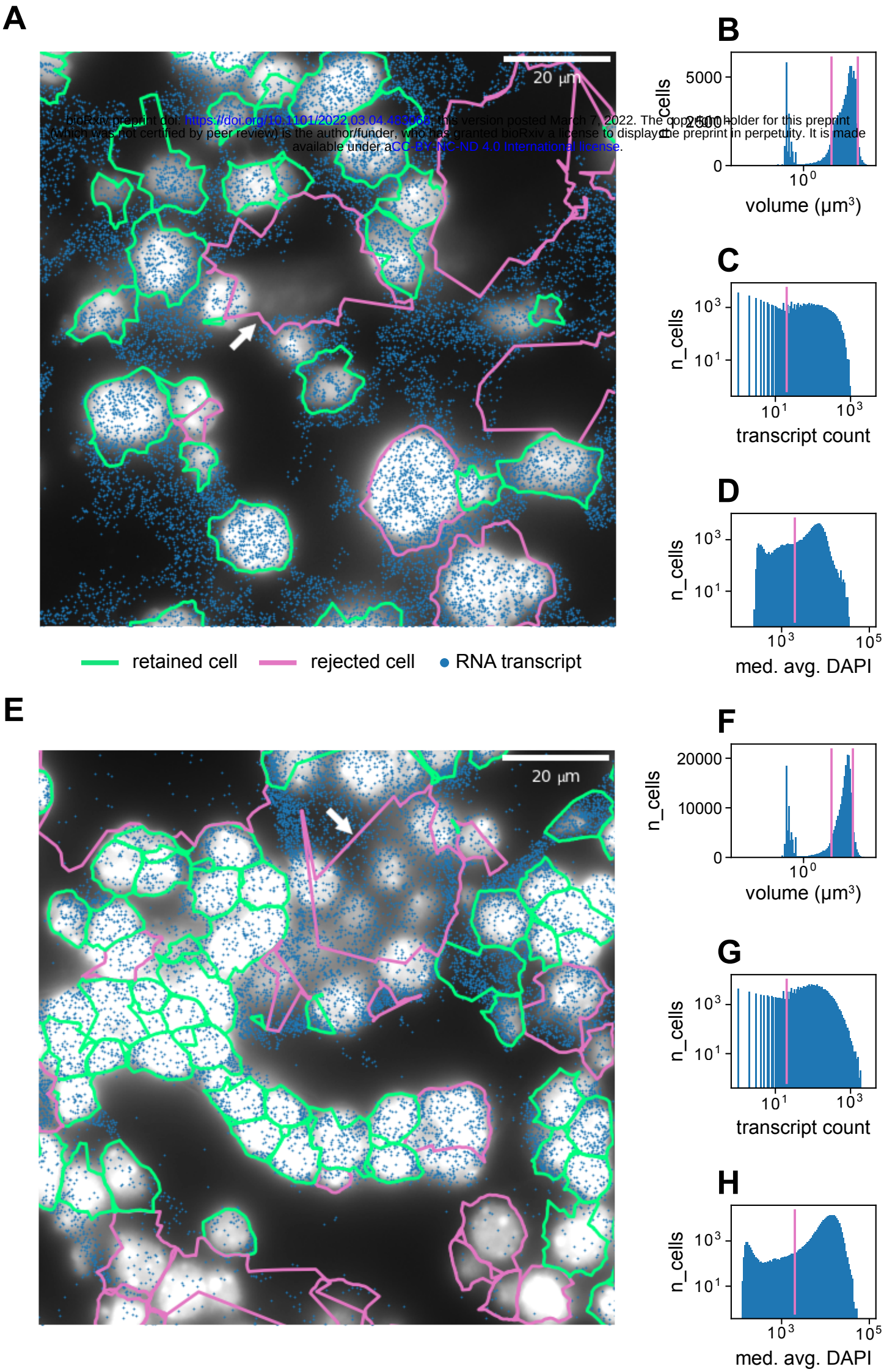


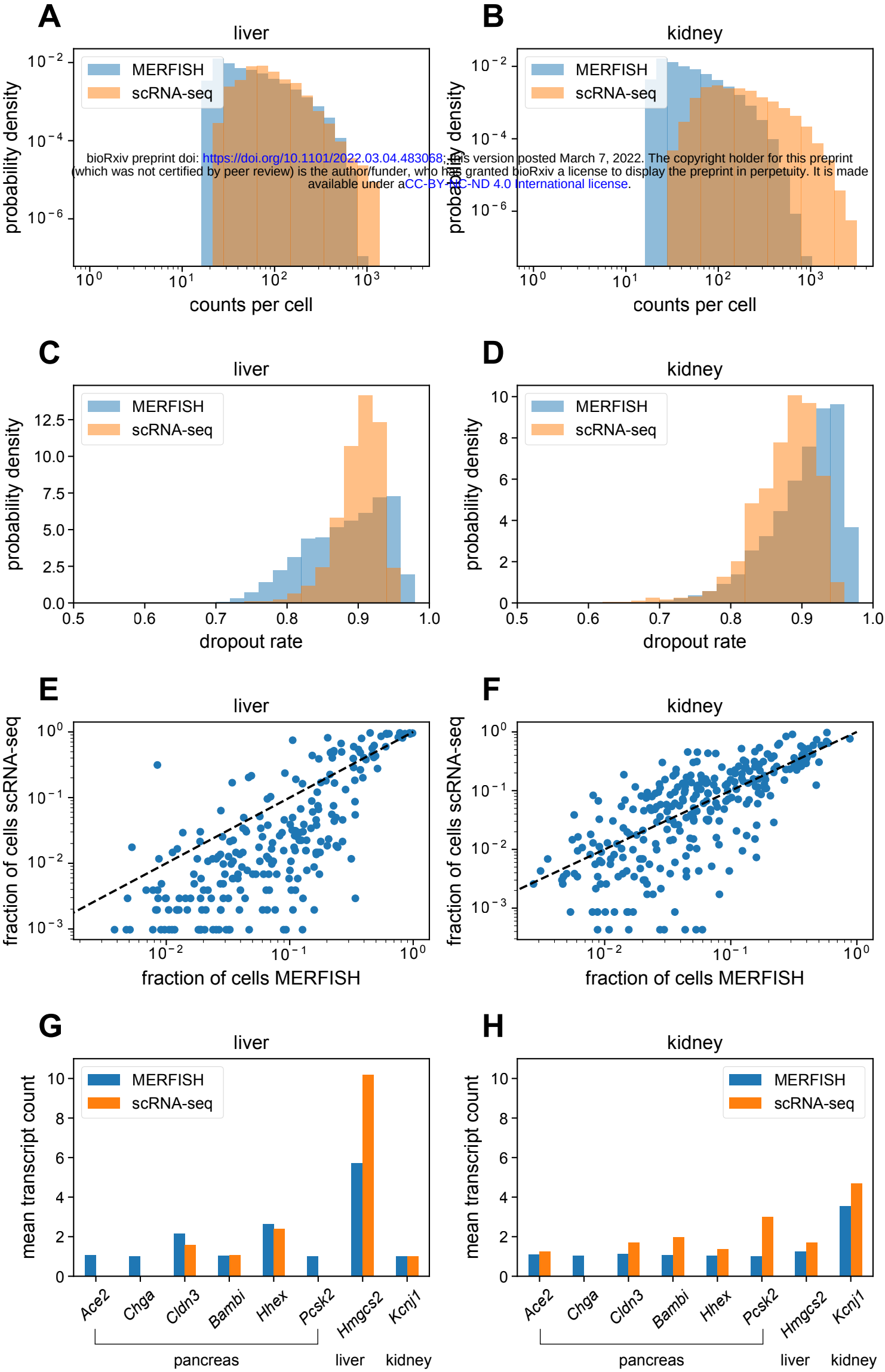
Gene	Barcode
A	100000…010
B	000100…001
C	001010…000
D	010000…001
E	100001…000
F	010100…000
G	000010…010


A
mouse kidney
https://doi.org/10.1101/2022.03.04.483068
CC-BY-NC-ND 4.0 International license
1 mm
B
DAPI
20 µm
C
cell boundary antibody
20 µm
D
sample MERFISH bit
20 µm
E
~13k detected transcripts

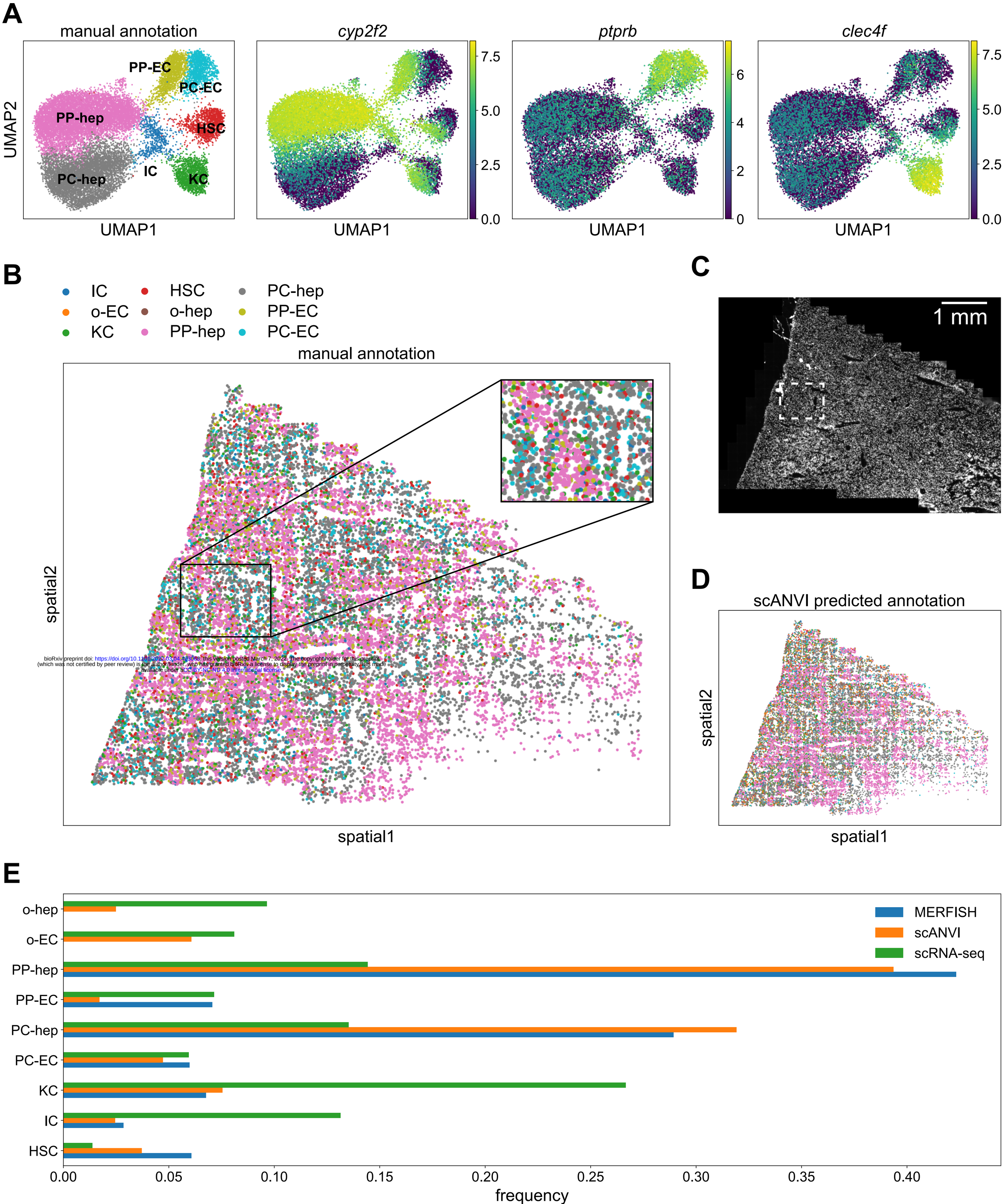

MERFISH technical replicates
liver
kidney
A
B
R² = 0.99
R² = 0.95
log count replicate 2
log count replicate 1
comparison with bulk RNA-seq
C
D
R² = 0.61
R² = 0.58
Hmgcs2
Ptprb
Gpx3
Kcnj1
mean log count RNA-seq
mean log count MERFISH

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483868; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

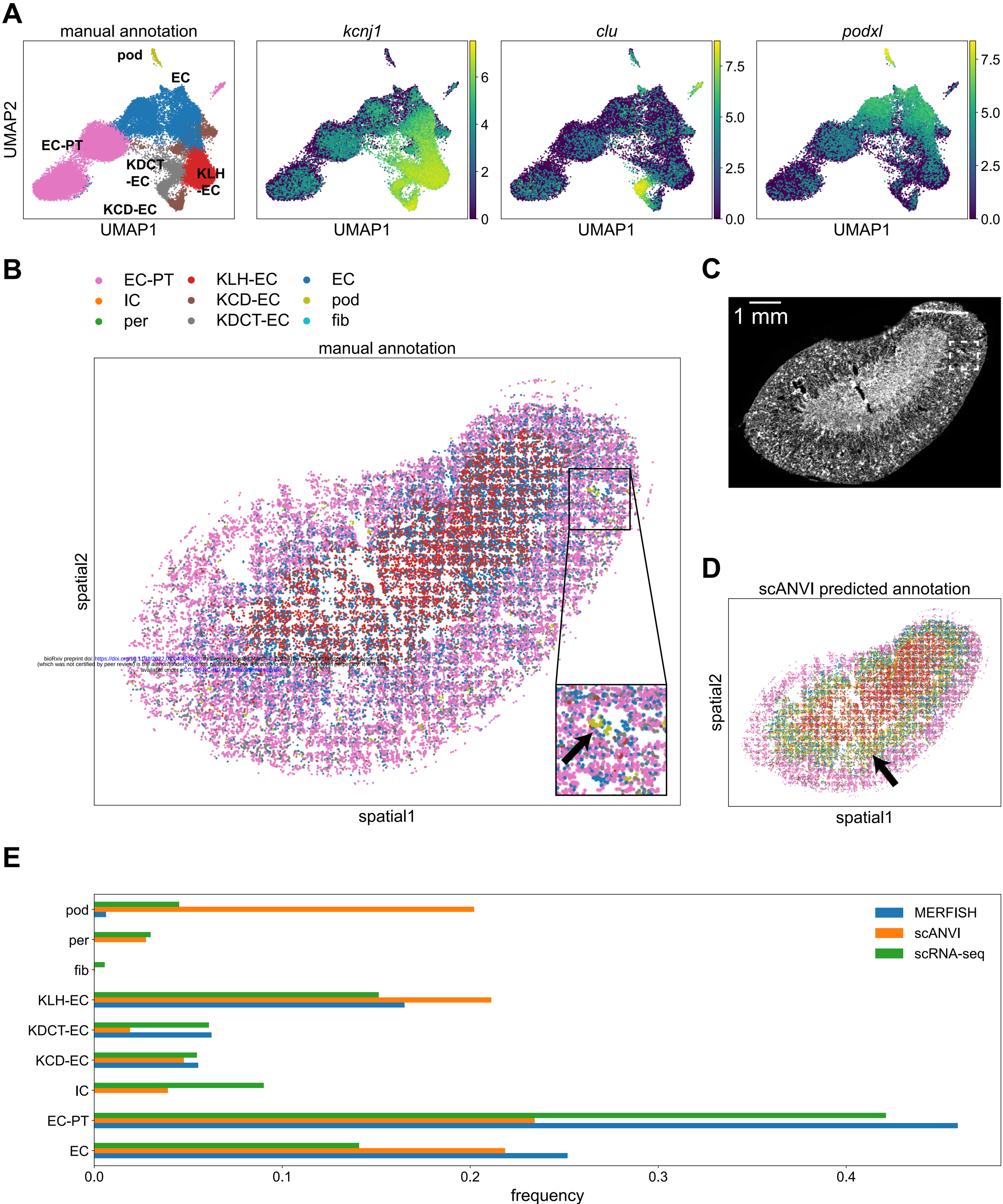




bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2022.03.04.483068; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.