Title

Concordance 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], Yang Joon Kim[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) by providing spatial context for cell type identification and analysis. Imaging-based spatial technologies such as Multiplexed Error-Robust Fluorescence In Situ Hybridization (MERFISH) can achieve single-cell resolution, directly mapping single cell identities to spatial positions. MERFISH produces an intrinsically different data type than scRNA-seq and a technical comparison between the two modalities is necessary to ascertain how to best integrate them. We performed MERFISH on mouse liver and kidney and compared the resulting bulk and single-cell RNA statistics with those from the Tabula Muris Senis cell atlas as well as from two Visium datasets. MERFISH quantitatively reproduced the bulk RNA-seq and scRNA-seq results with improvements in overall dropout rates and sensitivity. Finally, we found that MERFISH independently resolved distinct cell types and spatial structure in both liver and kidney. Computational integration with the Tabula Muris Senis atlas did not enhance these results. We conclude that compared to scRNA-seq, MERFISH provides a quantitatively comparable method for measuring single-cell gene expression and can robustly identify cell types 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 (K. H. 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 (Moses and Pachter 2022).

A microscopy-based method, MERFISH 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.

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 (K. H. 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). Evaluating its potential for further biomedical research necessitates a technical analysis in other types of tissue. As MERFISH relies on a limited probe panel of marker genes, it is crucial to ascertain if MERFISH measurements are independently sufficient and informative to conduct single-cell bioinformatic analysis such as cell type clustering and identification, or if they need to be augmented 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 2020) as well as with data from two Visium spatial transcriptomics datasets (Dixon et al. 2022; Guilliams et al. 2022), 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, with MERFISH exhibiting superior dropout rates and sensitivity, we examined MERFISH’s ability to identify different cell types. Intriguingly, MERFISH was able to distinguish between different cell types in liver and kidney tissue with standard single-cell bioinformatics analysis, even with only a 307-gene panel containing marker genes split between three organs. Furthermore, we were

able to resolve clear structure in the spatial distribution of these cell types, for example in spatial patterning of pericentral and periportal hepatocytes and endothelial cells in mouse liver as well as of podocytes in mouse kidney.

We then integrated our MERFISH measurements with Tabula Muris Senis using a couple of computational methods (Lopez et al. 2018; Korsunsky et al. 2019; Kang et al. 2021; 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 enhance cell type identification, indicating that MERFISH data alone can sufficiently resolve distinct cell types. While most cell types were similarly labeled by integration or manual annotation, some cell types like podocytes were consistently misclassified by the integration methods. We ascertained that these errors stemmed from intrinsic differences in RNA statistics between MERFISH and scRNA-seq, rather than from the integration methods themselves.

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 (K. H. 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 (SI 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, flowing in different subsets of fluorescently labeled readout probes 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” (SI Fig 1A). These barcodes and their intracellular positions are combined with cell nucleus and boundary staining, allowing for segmentation and measurement of gene expression at single cell resolution.

SI 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. Staining with a combination of DAPI and membrane protein antibodies produced a fluorescent readout on cell nuclei and boundaries. Using the MERlin image analysis and Cellpose segmentation packages (Emanuel et al. 2020; Stringer et al. 2021), these raw images were processed to obtain positions of single RNA transcripts and segmented geometries of individual cells, respectively. 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). Fourteen datasets were collected from mouse liver (5), kidney (4), and pancreas (5). Of these, all five of the pancreas samples and two of the kidney samples did not yield enough RNA for downstream analysis (Section S1 and SI Fig 2). Furthermore, one of the kidney samples only yielded enough transcripts for bulk RNA analysis. In addition, only one of the liver samples was stained for DAPI and cell boundary antibodies that allowed for single cell segmentation. The other four were used for bulk RNA analysis only. Thus, the bulk RNA analysis in this paper utilized five liver and two kidney samples, and the single-cell RNA analysis utilized one liver and one kidney sample.

Example workflow

Figure 1 shows an example of the data acquisition and image analysis procedure for a ~1cm^2 mouse kidney tissue sample (Fig 1A). 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. Each FOV is first imaged for DAPI and cell boundary antibody stains as shown in Figures 1B and 1C. Then, multiple rounds of hybridization begin and the various smFISH signals are acquired as shown in Figure 1D 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. 1E, colored points; see Methods and Materials for details). The cell boundary antibody stain and DAPI channels can be used to segment and produce single cell boundaries (Fig. 1E, black). These result in processed images containing the single cell boundaries and decoded transcript positions, which are then used for downstream bioinformatic analysis.

Sample quality and RNA integrity are critical to successful MERFISH analysis. We computed the average density of detected RNA transcripts for each tissue 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 2), and we discarded measurements with RIN score lower than 4. All five liver datasets passed this margin, as well as two kidney datasets.

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 2A and 2B 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 = 0.99 and R = 0.95 for liver and kidney, respectively. MERFISH replicates are 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 2C and 2D show the total log-transformed counts per gene in mouse liver and kidney between the two types of technologies, with some marker genes highlighted for illustrative purposes. 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 = 0.61 and R = 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.

As an additional point of comparison, we examined bulk MERFISH results together with another spatial transcriptomics technology, the sequencing-based Visium assay (10x Genomics). We used two publicly available datasets for the mouse liver (Guilliams et al. 2022) and kidney (Dixon et al. 2022), respectively, and conducted the same correlation analysis as before. Since Visium is not a single-cell technology, we reasoned that this bulk analysis would be the best way to compare the two. For each Visium dataset, we created a pseudo-bulk measurement by simply summing up all of the detected RNA in the sample and then compared these data with the bulk MERFISH results.

Figure 2E and 2F show the results of this pseudo-bulk comparison for the mouse liver and kidney, respectively, with some marker genes highlighted. Like before, each point represents the RNA count for a single gene, averaged across different experimental samples for the corresponding technology. In both tissues bulk MERFISH results correlated well with Visium (R = 0.65 and R = 0.69 for liver and kidney, respectively). Thus, at the bulk level, we conclude that MERFISH quantitatively agrees with both RNA-seq and Visium technologies.

Analysis and quality control of single-cell MERFISH results

After image acquisition we extracted 2D segmented cell boundaries and decoded RNA transcripts positions as shown for sample zoomed-in views at a single z-position in Fig. 3A (liver) and Fig. 3E (kidney). Using the segmented boundaries in the median z-position (i.e. 4th of 7), we assigned transcripts to cells by treating each cell's 2D boundary as a cookie cutter and associating every transcript in the whole 3D volume that lay within the boundary's xy coordinates with that cell. Segmentation quality was variable, with some cells 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 (Fig. 3A and 3E, white arrows). Not all detected RNA transcripts lay inside segmented boundaries – the percentage of transcripts assignable to cells was around 70% for both the liver and kidney. The remaining 30% likely consisted of both extracellular RNA as well as RNA near the peripheries of cells that lay slightly outside the segmented boundaries.

We developed a quality control protocol to filter out these low quality cells. From the data, we generated distributions of single-cell 2D areas and total counts per cell (Fig. 3B, C, F, and G). We calculated a per-cell alignment metric with the DAPI signal which we designate the *average*

DAPI score. This score is the average DAPI intensity within the xy borders of each segmented cell in the median (i.e. fourth) z-slice (Fig. 3D and 3H). A well segmented cell would intuitively have a high average DAPI score due to good alignment between the segmentation and the DAPI signal whereas a poorly segmented cell would have a lower score.

Quality control filters were established as follows: segmented cells that were unreasonably small or large were discarded by establishing low-area and high-area cutoffs of 200 μm^2 and 3000 μm^2 in the distribution of cell areas (Fig. 3B and 3F). A minimum cutoff of total RNA transcript count per cell of 80 was established to filter out cells with sparse statistics (Fig. 3C and 3G). Both the mouse liver and kidney datasets exhibited bimodality in the median average DAPI score per cell (Fig. 3D and 3H), and a cutoff of 400 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. 3A and 3E, white arrows).

This quality control procedure resulted in 34,217 liver cells and 126,547 kidney cells being retained out of 83,410 and 212,090 originally segmented cells, respectively. This corresponds to yields of 41% and 60% for the liver and kidney dataset, respectively.

Comparison of single-cell MERFISH results with scRNA-seq

Using 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 2020). First we examined the distribution of total RNA transcript counts per cell shown in Figures 4A and 4B for mouse liver and kidney. For a proper comparison the Tabula Muris Senis data were reduced to a subset containing the same 307 genes as the MERFISH gene panel. In the liver, the MERFISH distribution quantitatively agreed with scRNA-seq (Fig 4A). In the kidney, the distributions agreed up to total counts of approximately 1000 transcripts/cell, after which the MERFISH distribution drops to zero while the scRNA-seq distribution continues.

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 3). 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 3). This effect was more pronounced in the kidney than in the liver, reflecting the increased discrepancy in overall RNA count distributions in the kidney (Fig 4B).

Next, we examined the dropout rate between the technologies, shown in Fig 4C and 4D. 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 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 both the liver and kidney, while both technologies possessed cells with high dropout rates (> 0.8), MERFISH possessed a larger fraction of cells with lower dropout rates (between 0.7 and 0.8) compared to scRNA-seq. Thus, we conclude that MERFISH results in systematically lower dropout rates than scRNA-seq.

We then investigated whether either technology was more sensitive than the other (Fig 4E and 4F). While the previous analysis of dropout rates quantified the number of genes detected per cell, we now quantified the number of cells in which each gene was detected. To do so, we calculated the fraction of cells that had non-zero 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 both the liver (Fig 4E) and the kidney (Fig 4F) MERFISH systematically detected genes in higher proportions of cells than scRNA-seq. Thus, for this gene panel, MERFISH is more sensitive than scRNA-seq.

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 an effective control – substantial detection of these pancreas genes in the MERFISH data would indicate a nontrivial false positive rate. Figures 4G and 4H show the mean transcript count of six pancreas marker genes (*Ace2, Chga, Cldn3, Bambi, Hhex, Pcsk2*) among cells with non-zero 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 effective false positive rates were low in MERFISH and scRNA-seq.

To investigate the statistical properties of MERFISH data more deeply, we compared the mean RNA counts for each gene between MERFISH and scRNA-seq, across cells that possessed nonzero counts for that gene (Fig 5A and 5B for liver and kidney, respectively). We only considered genes that had at least 50 cells with nonzero counts in each dataset, to filter out genes with ambiguous means due to low sample size. In the liver (Fig 5A) there was a decent concordance and correlation of the means between MERFISH and scRNA-seq around the x=y line (R = 0.61), while in the kidney (Fig 5B) the correlation was slightly worse (R = 0.48). Furthermore, in the kidney there was less concordance – the scRNA-seq mean RNA counts were systematically higher across genes. We hypothesized that this was due to a combination of the overcrowding effect mentioned earlier, which would reduce the number of cells in MERFISH detected with high RNA counts, with the segmentation procedure performing worse in

the kidney and potentially failing to account for RNA transcripts located further away from cell nuclei.

scRNA-seq RNA counts are well modeled by the negative binomial distribution (Grün, Kester, and van Oudenaarden 2014; Kharchenko, Silberstein, and Scadden 2014). While this has also been demonstrated to hold for smFISH (So et al. 2011), similar analysis is more sparse for MERFISH, although recent results suggest that MERFISH statistics are also overdispersed and captured with the negative binomial distribution (Zhao et al. 2022). We investigated this by examining the mean-variance relationship of RNA counts for both MERFISH and scRNA-seq for the genes shown in Fig 5A and 5B, again only for cells with nonzero counts. We opted to only examine cells with nonzero counts since MERFISH and scRNA-seq possessed different dropout rates (Fig 4C and 4D). Figure 5C and 5D show this relationship for the liver and kidney, respectively. As a reference, the case where RNA counts follow a Poisson distribution (i.e. the variance equals the mean) is shown in the black dashed line. For both the scRNA-seq (orange dots) and MERFISH (blue dots) data, the variance is smaller than the mean for genes with low expression and higher than the mean for genes with high expression. Thus, both scRNA-seq and MERFISH possess overdispersed statistics for abundant genes. While this is expected for scRNA-seq, these results are novel in the case of MERFISH.

This overdispersion is well captured with the negative binomial distribution – the insets in Fig 5C and 5D show the full distribution of RNA counts for the liver and kidney marker genes *Hmgcs2* and *Kcnj1*, respectively (blue bars), along with fits to the negative binomial distribution (red lines). Additionally, the mean-variance relationship for both MERFISH and scRNA-seq are remarkably similar in both liver and kidney. Together, the data indicate that for our gene panel, MERFISH and scRNA-seq RNA counts exhibit quantitatively similar RNA count statistics that are well modeled by a negative binomial distribution, for cells possessing nonzero counts.

Single cell and spatial analysis of MERFISH results

We explored MERFISH's ability to resolve distinct liver cell types on its own using the segmented single-cell results. 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).

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. The three right plots in Fig 6A show expression levels of some marker genes for periportal hepatocytes (*cyp2f2)*, endothelial cells (*ptprb)*, and Kupffer (*clec4f*) cells.

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, brown and pink). We were also able to distinguish between periportal and pericentral endothelial cells (Fig 6A, gray and olive). Finally,

we detected a very small number of bile duct epithelial cells (Fig 6A, cyan; 60 cells out of 34,217 total).

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 S3; 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 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 Materials). Altogether 2321 liver cells from scRNA-seq were used. 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 S4 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. 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). Additionally, bile duct epithelial cells were not detected in the scRNA-seq reference and thus were not predicted by scANVI. An independent integration using Harmony and Symphony (Korsunsky et al. 2019; Kang et al. 2021) yielded similar results (Section S5 and SI Fig 6).

To directly compare the manual annotations with the predicted annotations from scANVI, we calculated the confusion matrix of cell annotations between the two methods. Each entry i, j in the matrix was defined as the fraction of cells manually annotated with label i that were predicted with label j in scANVI. Thus, perfect agreement would result in the diagonal elements possessing values of 1 and all other off-diagonal elements possessing values of zero. For the

liver data, there was generally high agreement between the two methods (Fig 6F). Because bile duct epithelial cells were not present in the scRNA-seq reference, there was obviously no agreement between scANVI and the manual MERFISH annotations. In addition, the manual MERFISH annotations did not label any “other” endothelial cells or hepatocytes, so those rows are blank.

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 we could resolve distinct clusters with the leiden algorithm. Clusters were annotated 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. 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 fibroblasts via manual annotation; however, even in the scRNA-seq reference they were extremely rare (< 1% of cells).

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); 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), a cell type that only exists in glomeruli (Brunskill et al. 2011), multicellular structures 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. Because the reference scRNA-seq kidney data were more comprehensive than the liver data, we used only three-month-old mouse data. 2327 kidney scRNA-seq cells were used in total. This resulted in the joint UMAP plot shown in SI Fig 5C.

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 the medulla (Fig 7D, olive).

A comparison of the relative cell type frequencies between the scRNA-seq reference, MERFISH results, and scANVI integration is shown in Fig 7E. The kidney sample here covered the entire organ unlike the liver sample in Figure 6; therefore the agreement between MERFISH and scRNA-seq was much more quantitatively apparent (Fig 7E, blue and green). The major differences were that MERFISH detected more endothelial cells, Loop of Henle epithelial cells, and collecting duct epithelial cells than scRNA-seq, as well as fewer podocytes. The scANVI

integration produced many more discrepancies (Fig 7E, orange), such as an abnormally high podocyte and pericyte count compared to scRNA-seq.

Examining the confusion matrix clarified these aberrations (Fig 7F). Although most cell type annotations agreed well between the manual and scANVI annotations, both pericytes and podocytes were poorly predicted by scANVI compared to the manual annotation. In particular, more podocytes were predicted by scANVI to be endothelial cells compared to those podocytes that were correctly labeled. Integration with Harmony and Symphony reproduced the scVI and scANVI results, including these discrepancies (Section S5 and SI Fig 6).

Section S6 and SI Fig 7 explore the basis of this mislabeling in further detail. To summarize, for certain cell types (e.g. podocytes) the MERFISH data are systematically different enough from the same cell types in the scRNA-seq reference that any integration method likely would fail. In other words, the integration performs well for matching cells with similar gene expression profiles together, but the two dataset types themselves are fundamentally different enough that the integration is not well posed. We refer the reader to Section S6 for a more detailed description of this investigation and conclude that the MERFISH-alone manual annotation is more reliable than computational integration for these cell types.

Discussion

MERFISH and other evolving spatial transcriptomic technologies offer a paradigm shift in transcriptomic analysis by combining subcellular spatial resolution with single cell segmentation ability. 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. A technical comparison with existing RNA sequencing data is necessary to fully understand the similarities and differences between the two modalities. An important question is if MERFISH independently enables cell type identification or if it 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 and clearing 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. Combining cell segmentation based on membrane antibody and DAPI stains, we obtained bulk and single-cell RNA counts (Fig 1) that were then compared with the pre-existing Tabula Muris bulk and single cell RNA datasets.

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 2A and 2B). Bulk comparisons with Tabula Muris Senis and publicly available Visium spatial transcriptomics data indicated relatively good agreement (Fig 2C-F). MERFISH typically produced bulk RNA

counts that were several orders of magnitude higher than RNA sequencing, 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 3). From this perspective, imaging-based transcriptomics methods such as MERFISH allow for easily interpretable quality control as these metrics can be ultimately derived from raw images rather than from downstream quantities. The predominant factor influencing signal quality was cell segmentation. After quality control, about half of detected cells were filtered out from a combination of low transcript count, extremal segmented cell areas, and/or poor alignment with DAPI nuclear fluorescence. 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. Recent methods that boost cell segmentation quality by using information from RNA transcript positions (Petukhov et al. 2021) or prior knowledge of cell type-specific gene expression (Littman et al. 2021) may further refine segmentation quality and allow for improved construction of single cell RNA count matrices from subcellular MERFISH signals. Such advances are necessary, as not all cell types may be equally affected by quality control – for example, more oddly shaped cells such as fibroblasts may likely be more difficult to segment and thus could systematically be excluded during quality control. Improvements in protein antibody staining could also boost segmentation quality in this regard and enable more robust detection of diverse cell types.

The filtered single cell MERFISH results were then compared with the Tabula Muris Senis scRNA-seq database (Fig 4). We examined summary statistics in mouse liver and kidney, starting with the total RNA counts per cell (Fig 4A and Fig 4B). For the most part, MERFISH quantitatively reproduced the statistics from scRNA-seq. However, for the 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 4B).

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. Our experience recommends that future MERFISH experiments avoid highly expressing genes in the gene panel to ameliorate 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 3). We expect the dynamic range of MERFISH measurements in mouse liver and kidney to be limited to cells that have under ~1000 labeled transcripts overall.

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 zero counts per single cell (Fig 4C and Fig 4D). The distribution of dropout rates was shifted to lower values for both the liver and kidney compared to scRNA-seq, suggesting that MERFISH was more sensitive than scRNA-seq in those tissues. We then quantified per-gene sensitivity by examining the fraction of cells that possessed non-zero counts for each gene in the 307-gene panel, for both MERFISH and scRNA-seq (Fig 4E and Fig 4F). In both the liver and kidney, this fraction was systematically larger for MERFISH than for scRNA-seq, again indicating MERFISH's increased sensitivity. 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 an effective 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 4G and 4H), indicating that MERFISH possessed low false positive detection rates at a similar level to scRNA-seq. Overall, we conclude that MERFISH possesses quantitatively similar overall statistical behavior to scRNA-seq, with minor discrepancies such as the dynamic range issue discussed above.

To more deeply investigate the single-cell statistics of MERFISH, we examined the concordance and correlation of mean RNA counts among cells with nonzero counts between MERFISH and scRNA-seq (Fig 5A and 5B). In the liver the concordance and correlation were decently high, while in the kidney the correlation was weaker and scRNA-seq produced systematically higher counts (likely due to a combination of cell segmentation flaws and the overcrowding effect mentioned earlier). Nevertheless, this quantitative correspondence prompted us to investigate if MERFISH RNA counts could be modeled with similar statistics as scRNA-seq. Examining the mean-variance relationship of the genes studied in Fig 5A and 5B indicated that among cells with nonzero counts, MERFISH and scRNA-seq indeed possessed similar statistics (Fig 5C and 5D). Furthermore, like scRNA-seq, MERFISH RNA count distributions were well modeled by a negative binomial distribution (Fig 5C and 5D, inset). Such analysis sheds further light on questions of noise in transcriptomic measurements (Grün, Kester, and van Oudenaarden 2014), as the statistical agreement between MERFISh and scRNA-seq suggest that distributions such as the negative binomial actually capture true biological variability in cellular transcription rather than technical noise due to measurement error.

Finally, we investigated MERFISH's ability to resolve distinct cell types in mouse liver and kidney and if 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). 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 these cells mainly on the basis of their spatial co-localization with the pericentral and periportal hepatocytes (Fig 6B, inset; Section S3; 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 4C), 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 as 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 a few artifacts. Direct comparison of cell type annotations between the manual method and scANVI indicated quantitative agreement for the liver (Fig 6F) and semi-quantitative agreement in the kidney (Fig 7F) with some exceptions for cell types such as podocytes. These results and discrepancies were reproduced by integration using Harmony and Symphony (SI Fig 6), suggesting that the underlying cause of the annotation disagreement lay in the data rather than the integration method. Indeed, a deeper investigation into cell type RNA profile similarities and differences between MERFISH and scRNA-seq indicated that this was indeed the case (SI Fig 7). Thus, while computational integration works for the most part, we caution researchers to also manually annotate MERFISH data and check for potential discrepancies between annotation methods.

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. 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 (Fig 4 C-F), our results anchor MERFISH as a technique that offers equivalent, if not better, overall biological signal compared to scRNA-seq. MERFISH and other spatial technologies provide a more complete representation of tissue cell types as fragile and rare cells may be lost in conventional tissue dissociation protocols.

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 similar signal with varying quality depending on cell type. Although here we opted to use scVI and scANVI as well as Harmony and Symphony, many other computational methods exist, such as Seurat (Stuart et al. 2019), 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, especially in light of the discovered systematic differences between MERFISH and scRNA-seq (SI Fig 7).

Substantial care should be taken regarding the design and interpretation of the gene panel. Because of the use of a scRNA-seq reference in creating the list of genes for the panel, the strength of MERFISH is heavily reliant on the quality of the reference and the method of panel construction. For example, the reference scRNA-seq dataset should be of sufficient sequencing depth to avoid creating a gene list based on noisy data. The method of choosing marker genes should also be robust. Although here we utilized simple differential expression analysis for picking marker genes, several recent software packages have been released that enable more nuanced approaches for marker gene panel creation (Aevermann et al. 2021; Missarova et al. 2021; X. Chen, Chen, and Thomson 2022).

Importantly, because of the bespoke nature of these panels, MERFISH data are best examined in comparison to their parent scRNA-seq references. Comparing with other scRNA-seq datasets should be taken with caution due to the strong batch effects between various scRNA-seq datasets. In addition, MERFISH data will only be informative for the genes existing in the panel. Here, we utilized a panel consisting only of marker genes, so investigations of other biological effects such as e.g. cell state or metabolic activity are not possible. For future MERFISH experiments interested in such properties, relevant genes for study should be added to the list of marker genes already present in the panel. Because of the limited size of MERFISH gene panels, there thus will always be a constraint on the number of genes available for study, and efficient panel design is of paramount importance. For example, a follow-up study for the experiments here could utilize a single panel for each tissue of interest, freeing up space on the panels for non-marker genes relevant for effects such as cell cycle or state.

In light of these findings, we envision MERFISH primarily as a targeted approach for following up on pre-existing scRNA-seq studies. 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 C57Black6 females 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 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:
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 the Vizgen Cell boundary Staining Kit (Catalog number: 10400009), a protein-based staining reagent consisting of a Blocking Buffer Premix (PN: 20300012), a Primary Staining Mix (PN: 20300010, 0.5mg/ml) and a Secondary Staining Mix (PN: 20300011, 0.33mg/ml) that can label 3 plasma membrane proteins in cell. To perform the cell boundary staining, samples are first incubated with Blocking solution supplemented with Rnase inhibitor for 1 hour at room temperature. Then the sample is incubated with Primary Staining Solution diluted in Blocking Buffer (1:100 dilution) at room temperature for 1 hour. The sample is washed 3 times with 1X PBS, 5 minutes each time, and then incubated with Secondary Staining Solution diluted in Blocking Buffer (1:33 dilution) at room temperature for 1 hour. After washing with 1XPBS 3 times, 5 minutes each time, the sample is post fixed with 4% PFA at room temperature for 15 minutes 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.

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 with the Cellpose software package (Stringer et al. 2021). The decoded RNA molecules are partitioned into individual cells to generate single-cell RNA count matrices. To do so, the xy segmentation from the median z-slice was used (the 4th position out of 7). After projecting this segmentation up and down through all z-positions, all detected transcripts in all z-positions that lay within the xy boundaries of a cell were assigned to that cell. 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 2020). To remove the effects of age, we subsetted both reference datasets to mice that were three months old.

Comparison with publicly available Visium dataset:
Comparisons of pseudo-bulk counts between Visium and MERFISH results were done with publicly available Visium datasets for mouse liver and kidney (Dixon et al. 2022; Guilliams et al. 2022). The pseudo-bulk counts computed per each dataset were averaged over replicates for five Visium datasets for mouse liver. For the mouse kidney Visium dataset, we used one sham dataset that Dixon et al. 2022 used in their study as a control (healthy mouse).

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, extremal segmented areas, and low average DAPI scores as mentioned in the main text (Fig 3). The comparative analyses in Figure 4 and 5 used the raw counts that resulted from the preprocessing. For the single-cell analysis in Figures 6 and 7, the data were further normalized to a total count of 10,000 transcripts per cell, log-transformed, and scaled such that each gene possessed zero mean and unit variance across cells.

For cell type identification and annotation 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.

Integration of MERFISH data with scRNA-seq reference using scVI and scANVI:
For integrated analysis with scVI and scANVI, 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 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.

Integration of MERFISH data with scRNA-seq reference using Harmony and Symphony:
To integrate MERFISH data with Tabula Muris Senis using Harmony and Symphony, we subsetted the scRNA-seq first using the same method mentioned above ("Integration of MERFISH data with scRNA-seq reference using scVI and scANVI). We then used Harmony and Symphony on the subsetted scRNA-seq to build a reference using the function *symphony::buildReference* with the following settings: 100 Harmony clusters, 100 variable genes to choose per group, and 20 principal components. The cell type annotations of MERFISH data were then predicted using the functions *mapQuery* with default settings and *knnPredict* with k=5 neighbors.

Data and Code 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, FISH probe target region sequences, 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 Neuhoefer 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. George Emanuel and Jiang He own equity in Vizgen.

References

Aevermann, Brian, Yun Zhang, Mark Novotny, Mohamed Keshk, Trygve Bakken, Jeremy Miller, Rebecca Hodge, Boudewijn Lelieveldt, Ed Lein, and Richard H. Scheuermann. 2021. "A Machine Learning Method for the Discovery of Minimum Marker Gene Combinations for Cell Type Identification from Single-Cell RNA Sequencing." *Genome Research* 31 (10): 1767–80. https://doi.org/10.1101/gr.275569.121.

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.

Chen, Xiaoqiao, Sisi Chen, and Matt Thomson. 2022. "Minimal Gene Set Discovery in Single-Cell MRNA-Seq Datasets with ActiveSVM." *Nature Computational Science* 2 (6): 387–98. https://doi.org/10.1038/s43588-022-00263-8.

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

Dixon, Eryn E., Haojia Wu, Yoshiharu Muto, Parker C. Wilson, and Benjamin D. Humphreys. 2022. "Spatially Resolved Transcriptomic Analysis of Acute Kidney Injury in a Female Murine Model." *Journal of the American Society of Nephrology* 33 (2): 279–89. https://doi.org/10.1681/ASN.2021081150.

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.

Grün, Dominic, Lennart Kester, and Alexander van Oudenaarden. 2014. "Validation of Noise Models for Single-Cell Transcriptomics." *Nature Methods* 11 (6): 637–40. https://doi.org/10.1038/nmeth.2930.

Guilliams, Martin, Johnny Bonnardel, Birthe Haest, Bart Vanderborght, Camille Wagner, Anneleen Remmerie, Anna Bujko, et al. 2022. "Spatial Proteogenomics Reveals Distinct and Evolutionarily Conserved Hepatic Macrophage Niches." *Cell* 185 (2): 379-396.e38. https://doi.org/10.1016/j.cell.2021.12.018.

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.

Kang, Joyce B., Aparna Nathan, Kathryn Weinand, Fan Zhang, Nghia Millard, Laurie Rumker, D. Branch Moody, Ilya Korsunsky, and Soumya Raychaudhuri. 2021. “Efficient and Precise Single-Cell Reference Atlas Mapping with Symphony.” *Nature Communications* 12 (1): 5890. https://doi.org/10.1038/s41467-021-25957-x.

Kharchenko, Peter V., Lev Silberstein, and David T. Scadden. 2014. “Bayesian Approach to Single-Cell Differential Expression Analysis.” *Nature Methods* 11 (7): 740–42. https://doi.org/10.1038/nmeth.2967.

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.

Korsunsky, Ilya, Nghia Millard, Jean Fan, Kamil Slowikowski, Fan Zhang, Kevin Wei, Yuriy Baglaenko, Michael Brenner, Po-ru Loh, and Soumya Raychaudhuri. 2019. “Fast, Sensitive and Accurate Integration of Single-Cell Data with Harmony.” *Nature Methods* 16 (12): 1289–96. https://doi.org/10.1038/s41592-019-0619-0.

Littman, Russell, Zachary Hemminger, Robert Foreman, Douglas Arneson, Guanglin Zhang, Fernando Gómez-Pinilla, Xia Yang, and Roy Wollman. 2021. “Joint Cell Segmentation and Cell Type Annotation for Spatial Transcriptomics.” *Molecular Systems Biology* 17 (6): e10108. https://doi.org/10.15252/msb.202010108.

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.

Missarova, Alsu, Jaison Jain, Andrew Butler, Shila Ghazanfar, Tim Stuart, Maigan Brusko, Clive Wasserfall, et al. 2021. “GeneBasis: An Iterative Approach for Unsupervised Selection of Targeted Gene Panels from ScRNA-Seq.” *Genome Biology* 22 (1): 333. https://doi.org/10.1186/s13059-021-02548-z.

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.

Moses, Lambda, and Lior Pachter. 2022. “Museum of Spatial Transcriptomics.” *Nature Methods*, March, 1–13. https://doi.org/10.1038/s41592-022-01409-2.

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. 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.

So, Lok-hang, Anandamohan Ghosh, Chenghang Zong, Leonardo A. Sepúlveda, Ronen Segev, and Ido Golding. 2011. “General Properties of Transcriptional Time Series in Escherichia Coli.” *Nature Genetics* 43 (6): 554–60. https://doi.org/10.1038/ng.821.

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.

Stringer, Carsen, Tim Wang, Michalis Michaelos, and Marius Pachitariu. 2021. “Cellpose: A Generalist Algorithm for Cellular Segmentation.” *Nature Methods* 18 (1): 100–106. https://doi.org/10.1038/s41592-020-01018-x.

Stuart, Tim, Andrew Butler, Paul Hoffman, Christoph Hafemeister, Efthymia Papalexi, William M. Mauck, Yuhan Hao, Marlon Stoeckius, Peter Smibert, and Rahul Satija. 2019. “Comprehensive Integration of Single-Cell Data.” *Cell* 177 (7): 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

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 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.

Zhao, Peiyao, Jiaqiang Zhu, Ying Ma, and Xiang Zhou. 2022. “Modeling Zero Inflation Is Not Necessary for Spatial Transcriptomics.” *Genome Biology* 23 (1): 118. https://doi.org/10.1186/s13059-022-02684-0.

Figure Legends

Figure 1 - **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.

Figure 2 - **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 of bulk RNA counts per gene in mouse (C) liver and (D) kidney between MERFISH and RNA-seq. (E, F) Comparison of bulk MERFISH RNA counts per gene with pseudobulk RNA counts from Visium in mouse (E) liver and (F) kidney. In C-F, counts per gene were averaged across replicates for each technology.

Figure 3 - **Quality control of single cell MERFISH data**. (A) Sample cropped image of mouse liver tissue with segmented cell boundaries (white, magenta), decoded RNA transcripts (gray), DAPI nuclear stain (blue), and cell boundary antibody stain (green). White boundaries indicate cells that have passed the filtering stage, whereas magenta boundaries indicate cells that have been thrown out. (B, C, D) Histograms of (B) cell areas, (C) RNA transcript count per cell, and (D) 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 areas, (G) RNA transcript count per cell, and (H) average DAPI per cell for the image in (E). White arrows in A and E indicate examples of poor cell segmentation.

Figure 4 - **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.

Figure 5 - **Statistical analysis of single-cell MERFISH and Tabula Muris Senis RNA count distributions.** (A, B) Comparison of mean RNA counts in MERFISH vs. scRNA-seq in mouse (A) liver and (B) kidney across cells with nonzero counts. Each dot represents a gene and only genes possessing at least 50 cells with nonzero counts are shown for each tissue. The dotted black line shows the x=y line, indicating one-to-one concordance between MERFISH and scRNA-seq. (C, D) Mean-variance relationship for RNA counts in MERFISH (blue) and scRNA-seq (orange) for mouse (A) liver and (B) kidney, across cells with nonzero counts. Each dot represents a gene and only genes possessing at least 50 cells with nonzero counts are shown for each tissue. The dotted line indicates the x=y line, which represents the Poisson scenario where the mean equals the variance. (C, inset) Distribution of MERFISH RNA counts across cells with nonzero counts for *Hmgcs2*, a liver hepatocyte marker gene. Red line indicates the best fit to a negative binomial distribution. (D, inset) Distribution of MERFISH RNA counts across cells with nonzero counts for *Kcnj1*, a kidney Loop of Henle cell marker gene. Red line indicates the best fit to a negative binomial distribution. Error bars in A and B represent standard error of the mean.

Figure 6 - Single-cell and spatial analysis of MERFISH liver sample. (A) UMAP plots of MERFISH data colored by manually annotated clusters or by normalized, log-transformed, and scaled 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 (gray and brown, respectively) and pericentral endothelial cells and hepatocytes (olive and pink, 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. (F) Confusion matrix of MERFISH cell type annotations between manual method and scANVI predictions. The rows for other endothelial cells and hepatocytes ("o-EC" and "o-hep") are blank because none of the manual annotations were in these groups. Additionally, the row for bile duct epithelial cells is noisy because none of the scRNA-seq reference cells possessed this annotation. 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', 'BD-EC': 'bile duct epithelial cell'.

Figure 7 - Single-cell and spatial analysis of MERFISH kidney sample. (A) UMAP plots of MERFISH data colored by manually annotated clusters or by normalized, log-transformed, and scaled 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, black arrow). (C) DAPI stain 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. (F) Confusion matrix of MERFISH cell type annotations between manual method and scANVI predictions. 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'.

Figures

Figure 1

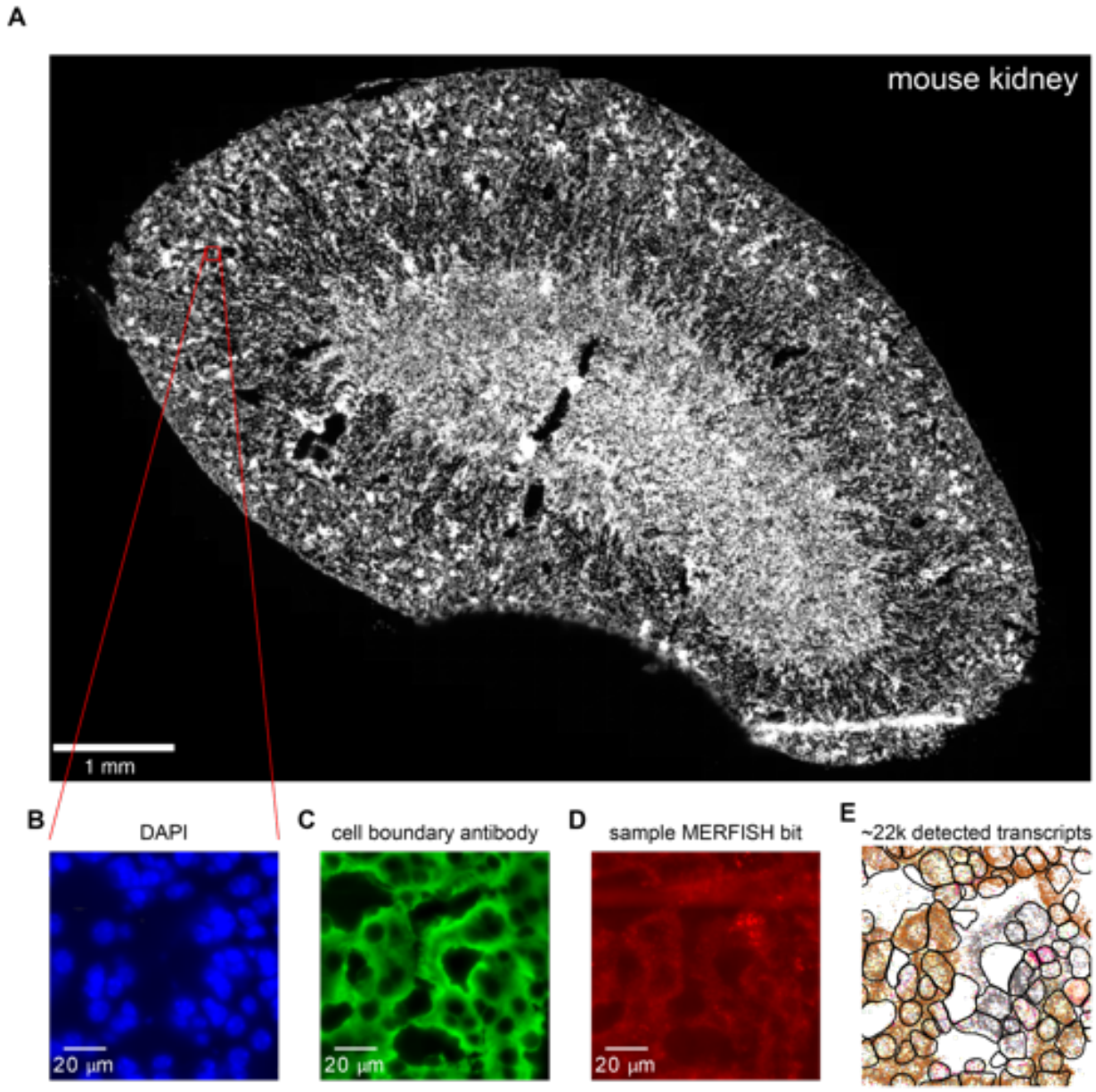


Figure 2

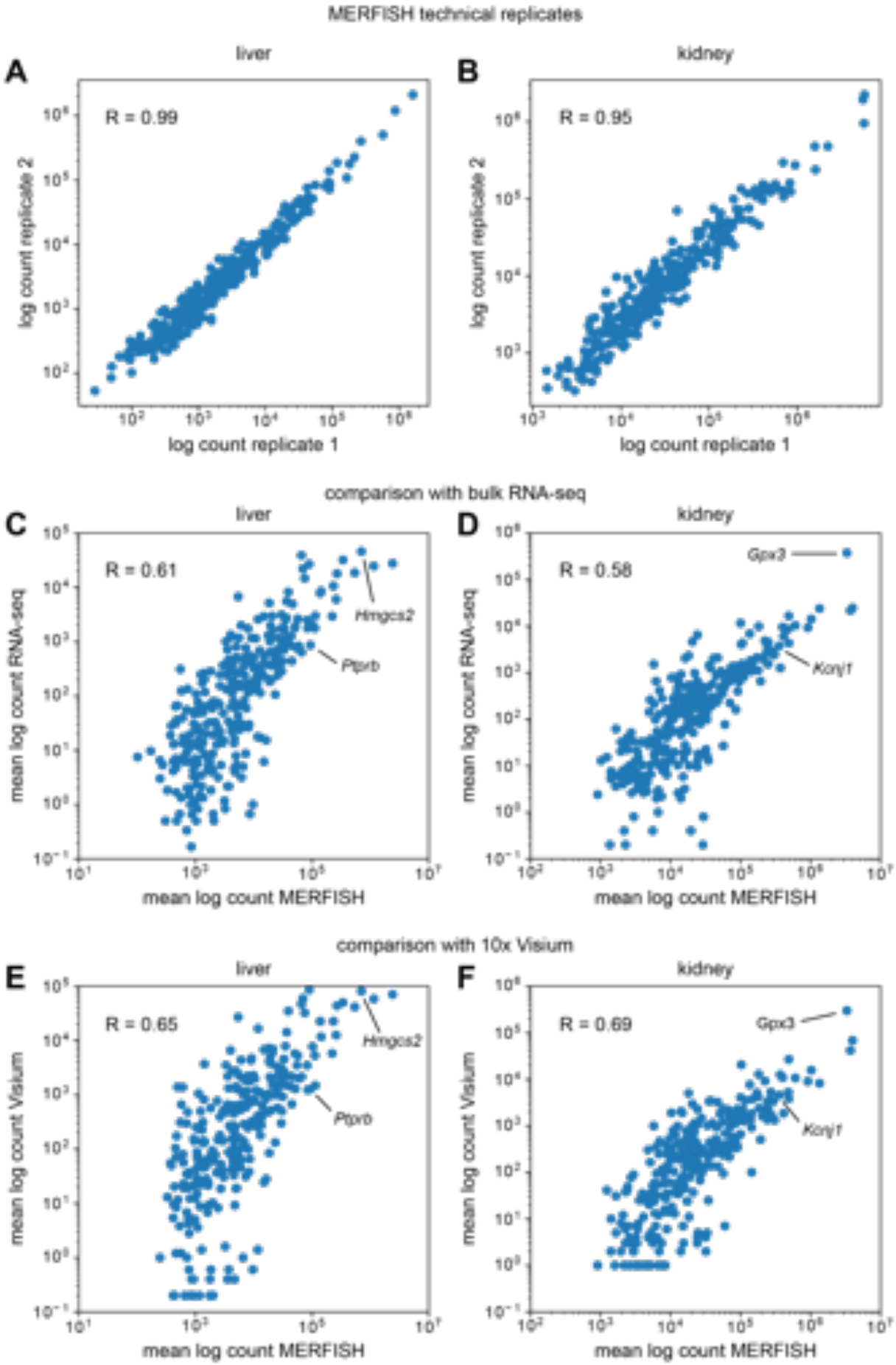

MERFISH technical replicates
A
liver
R = 0.99
log count replicate 2
log count replicate 1
B
kidney
R = 0.95
log count replicate 2
log count replicate 1
comparison with bulk RNA-seq
C
liver
R = 0.61
Hmgcs2
Ptprb
mean log count RNA-seq
mean log count MERFISH
D
kidney
R = 0.58
Gpx3
Kcnj1
mean log count RNA-seq
mean log count MERFISH
comparison with 10x Visium
E
liver
R = 0.65
Hmgcs2
Ptprb
mean log count Visium
mean log count MERFISH
F
kidney
R = 0.69
Gpx3
Kcnj1
mean log count Visium
mean log count MERFISH

Figure 3

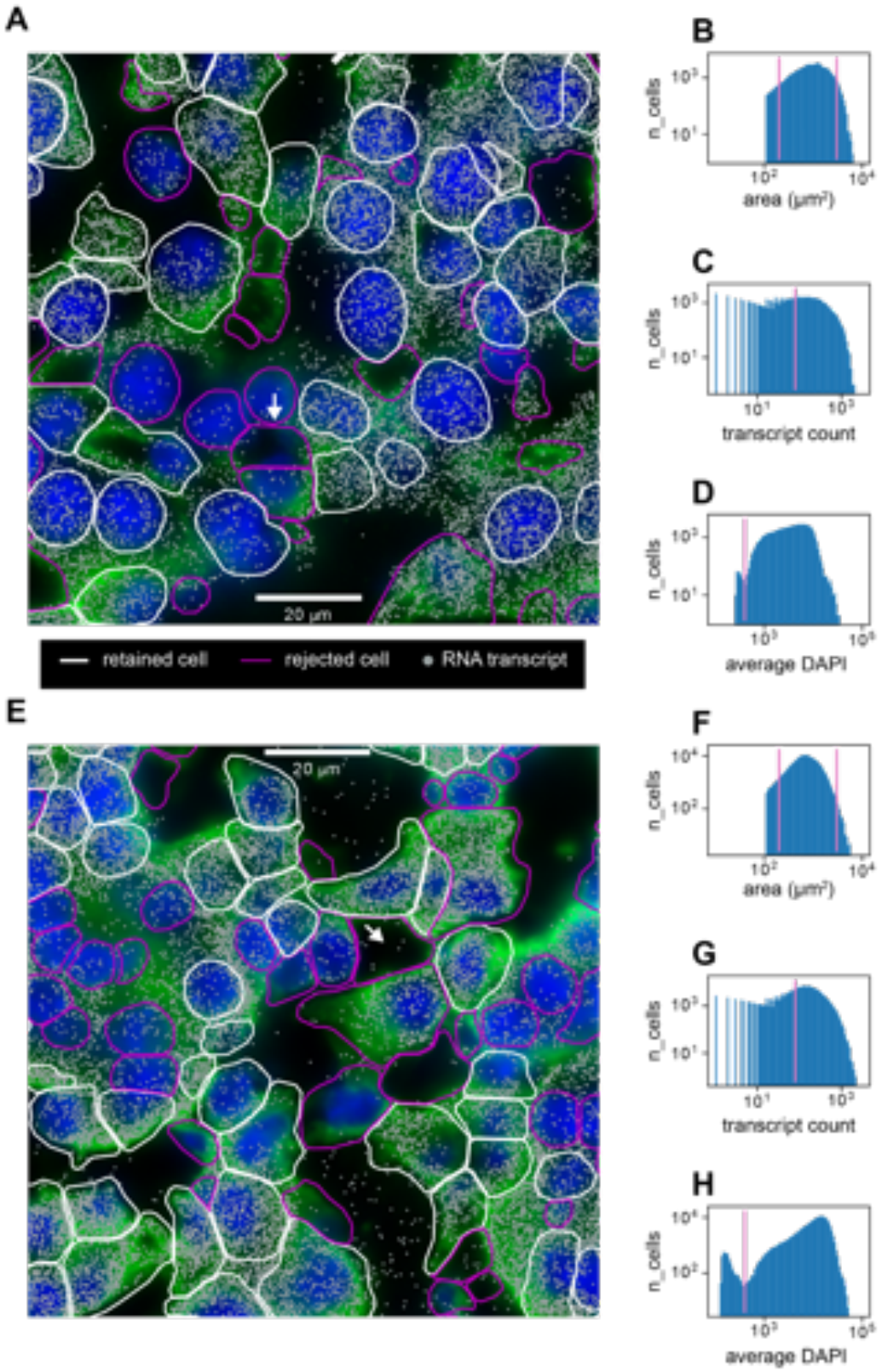

A
B
C
D
E
F
G
H
20 µm
retained cell
rejected cell
RNA transcript
n_cells
area (µm²)
transcript count
average DAPI

Figure 4

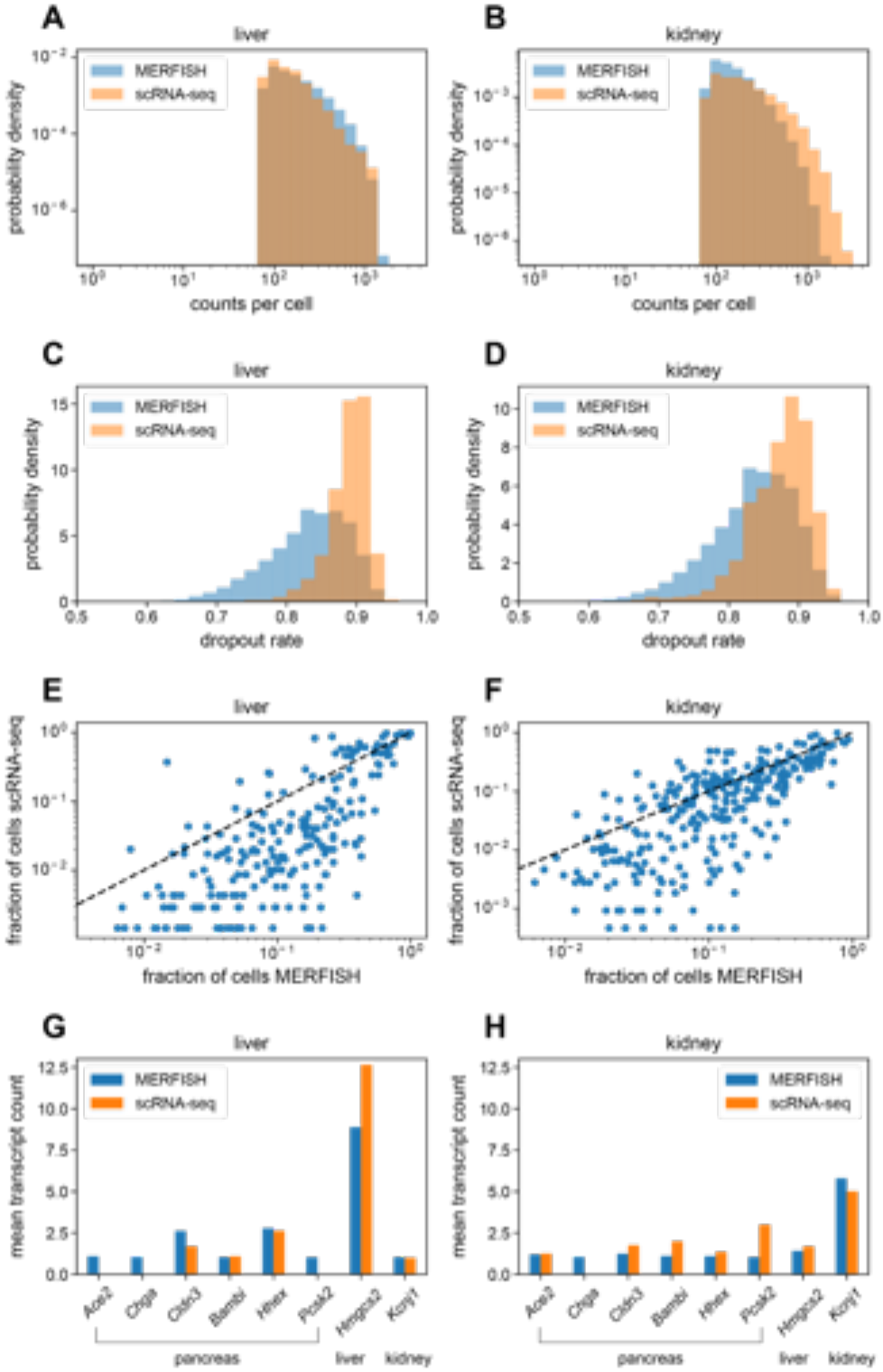

A
liver
MERFISH
scRNA-seq
probability density
counts per cell
B
kidney
MERFISH
scRNA-seq
probability density
counts per cell
C
liver
MERFISH
scRNA-seq
probability density
dropout rate
D
kidney
MERFISH
scRNA-seq
probability density
dropout rate
E
liver
fraction of cells scRNA-seq
fraction of cells MERFISH
F
kidney
fraction of cells scRNA-seq
fraction of cells MERFISH
G
liver
MERFISH
scRNA-seq
mean transcript count
Ace2
Chgb
Cldn3
Bambi
Hhex
Pcsk2
Hmgcs2
Kcnj1
pancreas
liver
kidney
H
kidney
MERFISH
scRNA-seq
mean transcript count
Ace2
Chgb
Cldn3
Bambi
Hhex
Pcsk2
Hmgcs2
Kcnj1
pancreas
liver
kidney

Figure 5

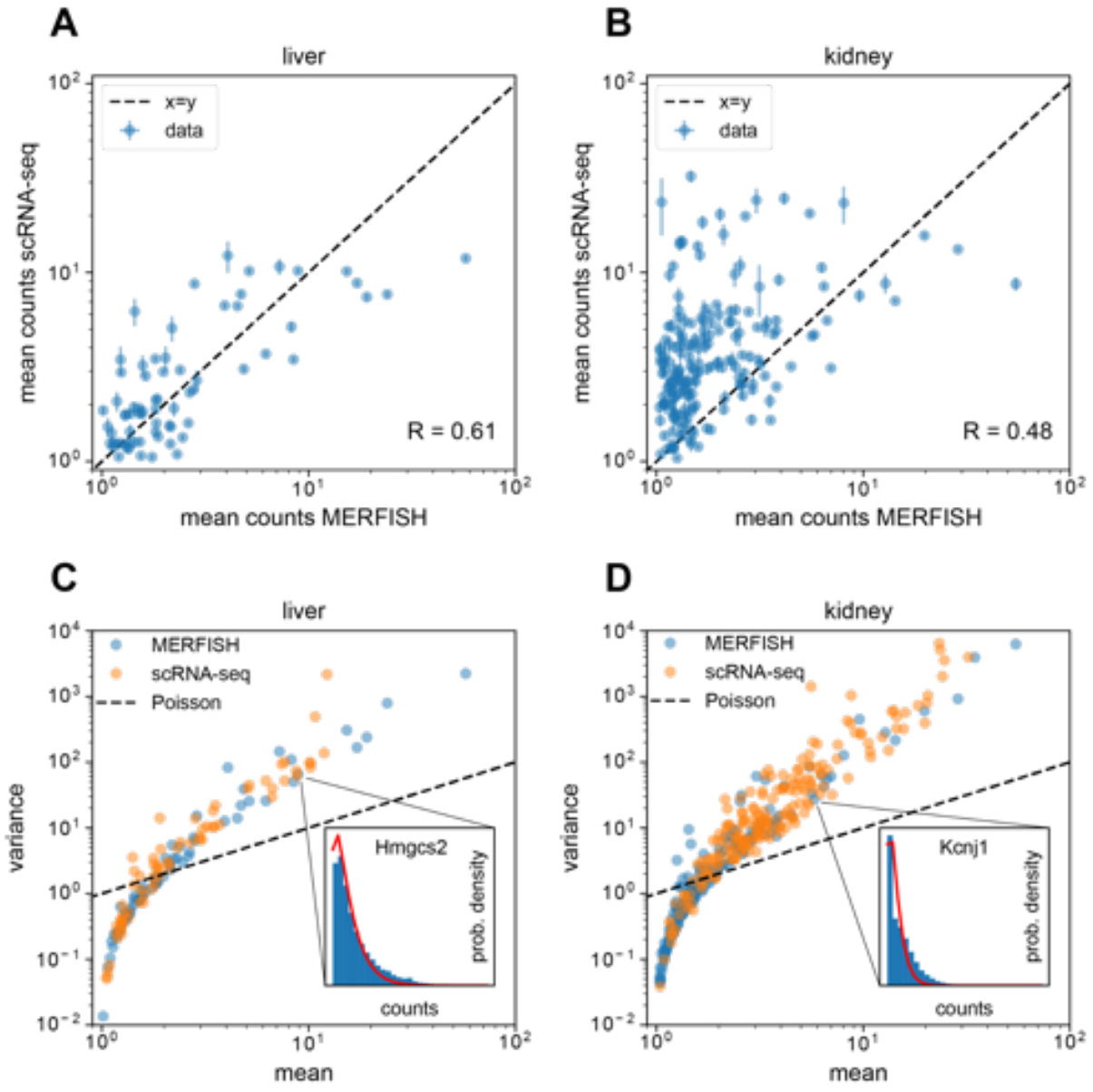

A
liver
x=y
data
R = 0.61
mean counts scRNA-seq
mean counts MERFISH
B
kidney
x=y
data
R = 0.48
mean counts scRNA-seq
mean counts MERFISH
C
liver
MERFISH
scRNA-seq
Poisson
Hmgcs2
prob. density
counts
variance
mean
D
kidney
MERFISH
scRNA-seq
Poisson
Kcnj1
prob. density
counts
variance
mean

Figure 6

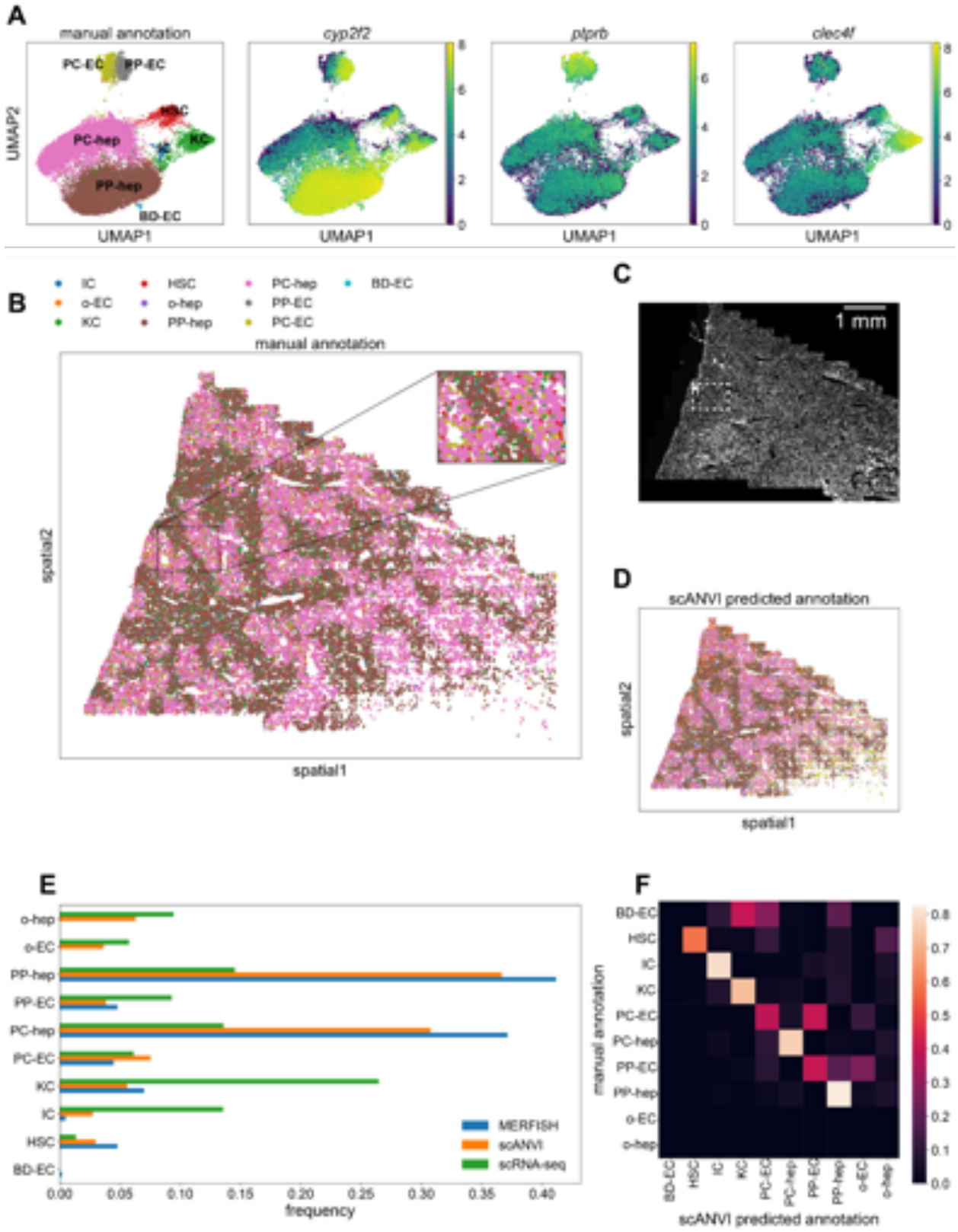
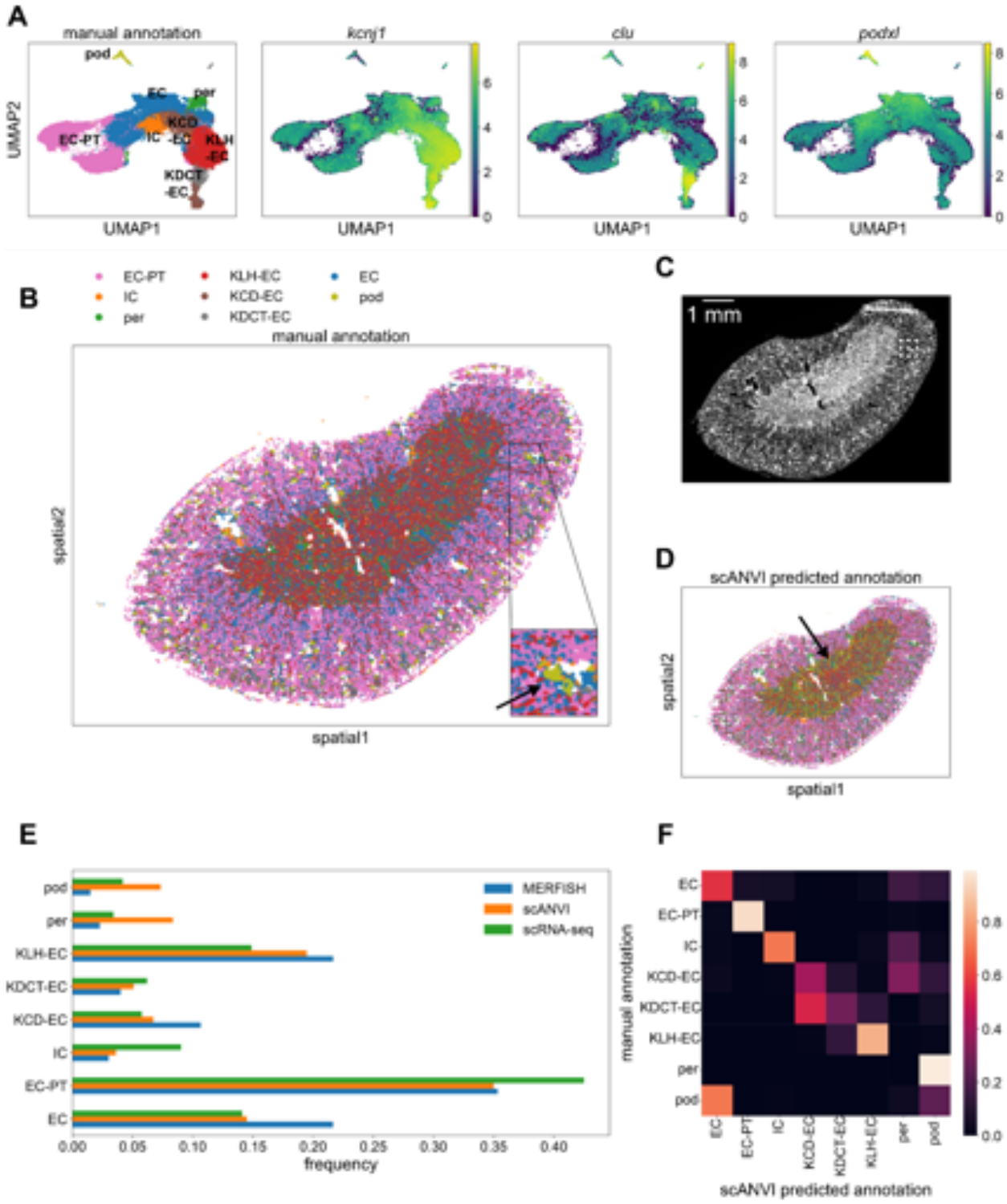
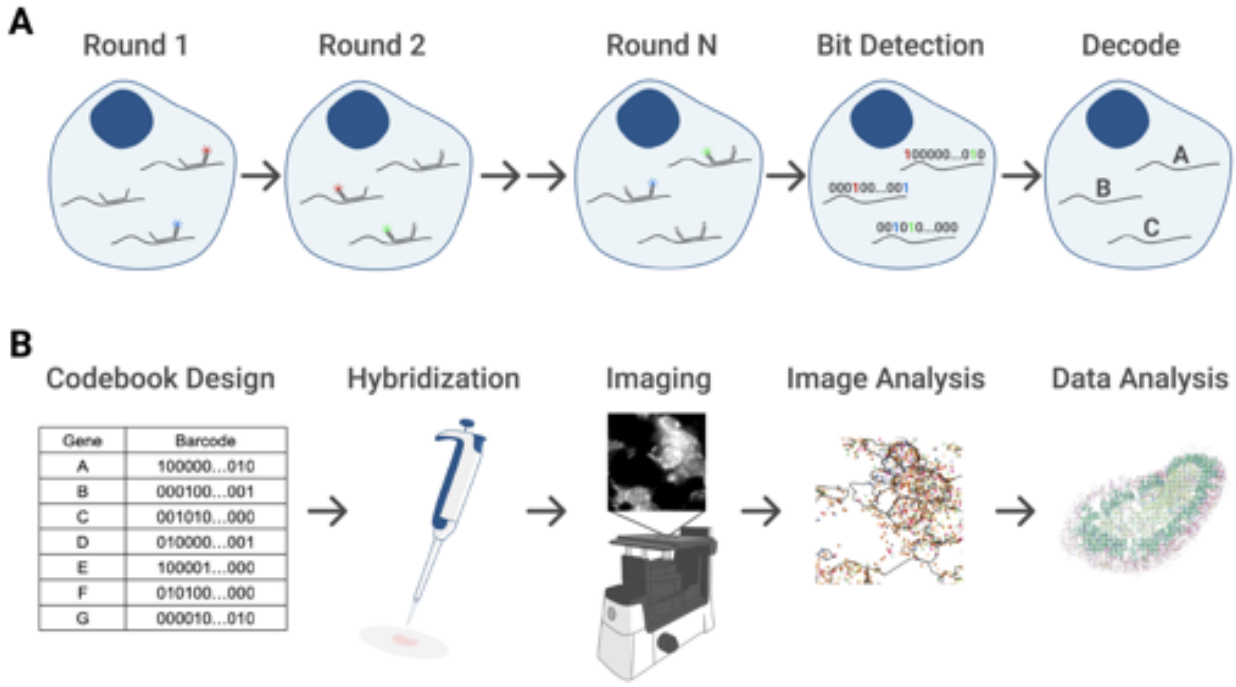

A
manual annotation
cyp2f2
ptprb
clec4f
PC-EC
PP-EC
HSC
PC-hep
KC
IC
PP-hep
BD-EC
UMAP2
UMAP1
B
IC
o-EC
KC
HSC
o-hep
PP-hep
PC-hep
PP-EC
PC-EC
BD-EC
manual annotation
spatial2
spatial1
C
1 mm
D
scANVI predicted annotation
spatial2
spatial1
E
o-hep
o-EC
PP-hep
PP-EC
PC-hep
PC-EC
KC
IC
HSC
BD-EC
MERFISH
scANVI
scRNA-seq
frequency
F
manual annotation
scANVI predicted annotation

Figure 7


A
manual annotation
kcnj1
clu
podxl
pod
EC
per
KCD -EC
EC-PT
IC
KLH -EC
KDCT -EC
UMAP2
UMAP1
B
EC-PT
IC
per
KLH-EC
KCD-EC
KDCT-EC
EC
pod
manual annotation
spatial2
spatial1
C
1 mm
D
scANVI predicted annotation
spatial2
spatial1
E
MERFISH
scANVI
scRNA-seq
pod
per
KLH-EC
KDCT-EC
KCD-EC
IC
EC-PT
EC
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
frequency
F
manual annotation
EC
EC-PT
IC
KCD-EC
KDCT-EC
KLH-EC
per
pod
scANVI predicted annotation
0.0
0.2
0.4
0.6
0.8

Supplementary Text

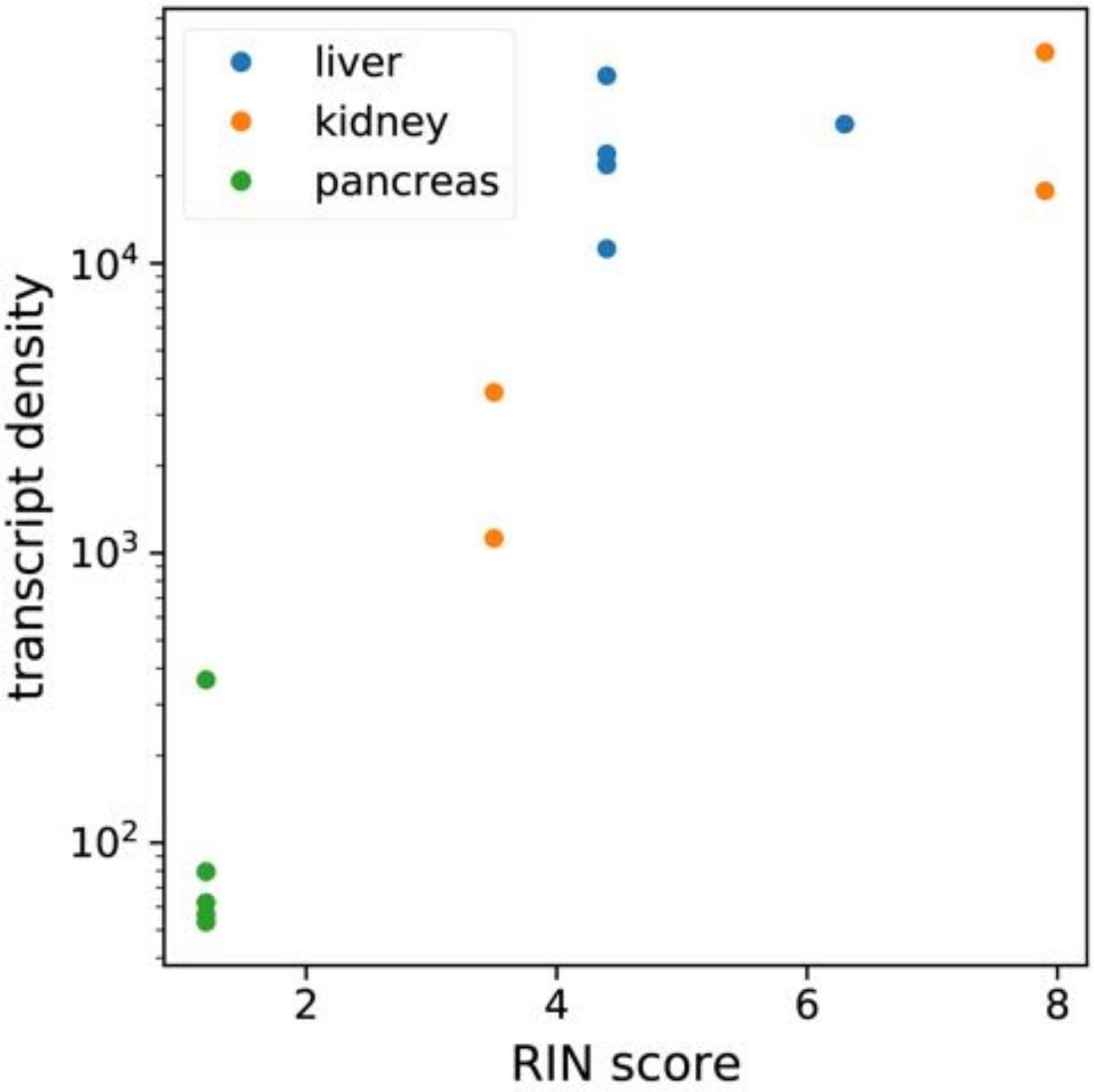


SI Figure 1: Overview of experimental setup. (A) Barcoding scheme across N rounds of multicolor imaging. Each detected fluorescent spot is registered as a "1" in its corresponding bit position. Then, RNA species are identified by matching the barcodes to a preselected gene codebook. (B) Experimental 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.

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 2 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.



SI Figure 2 - Investigation of RNA tissue quality on MERFISH measurements. Average transcript density is shown as a function of tissue RIN score for each tissue sample.

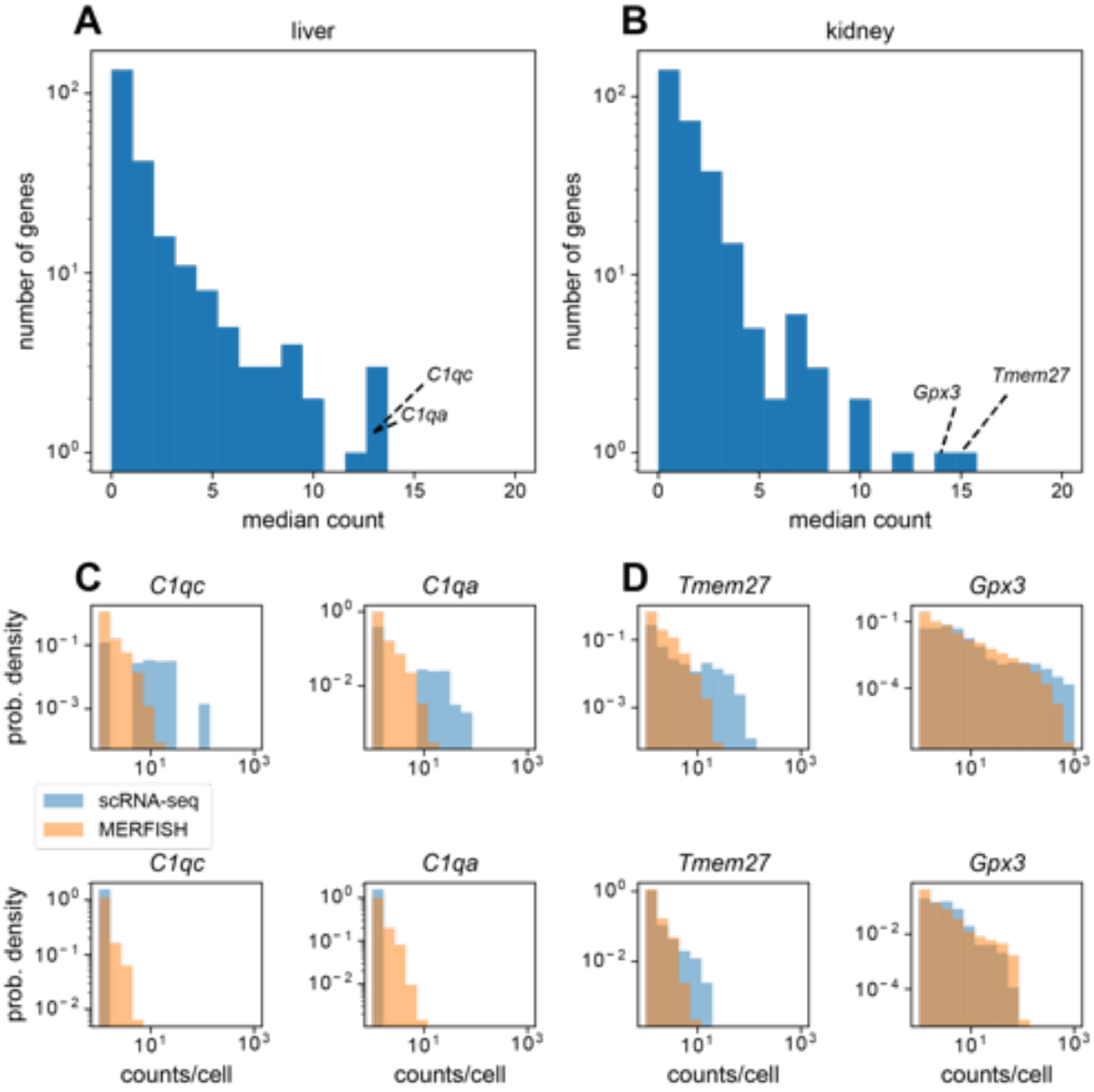
S2 Investigation of molecular crowding in MERFISH measurements

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 3A and SI Fig 3B 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 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 3C and SI Fig 3D 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 more similar in cells with under 100 total transcripts, particularly for the kidney.

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.



SI Figure 3 - 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).

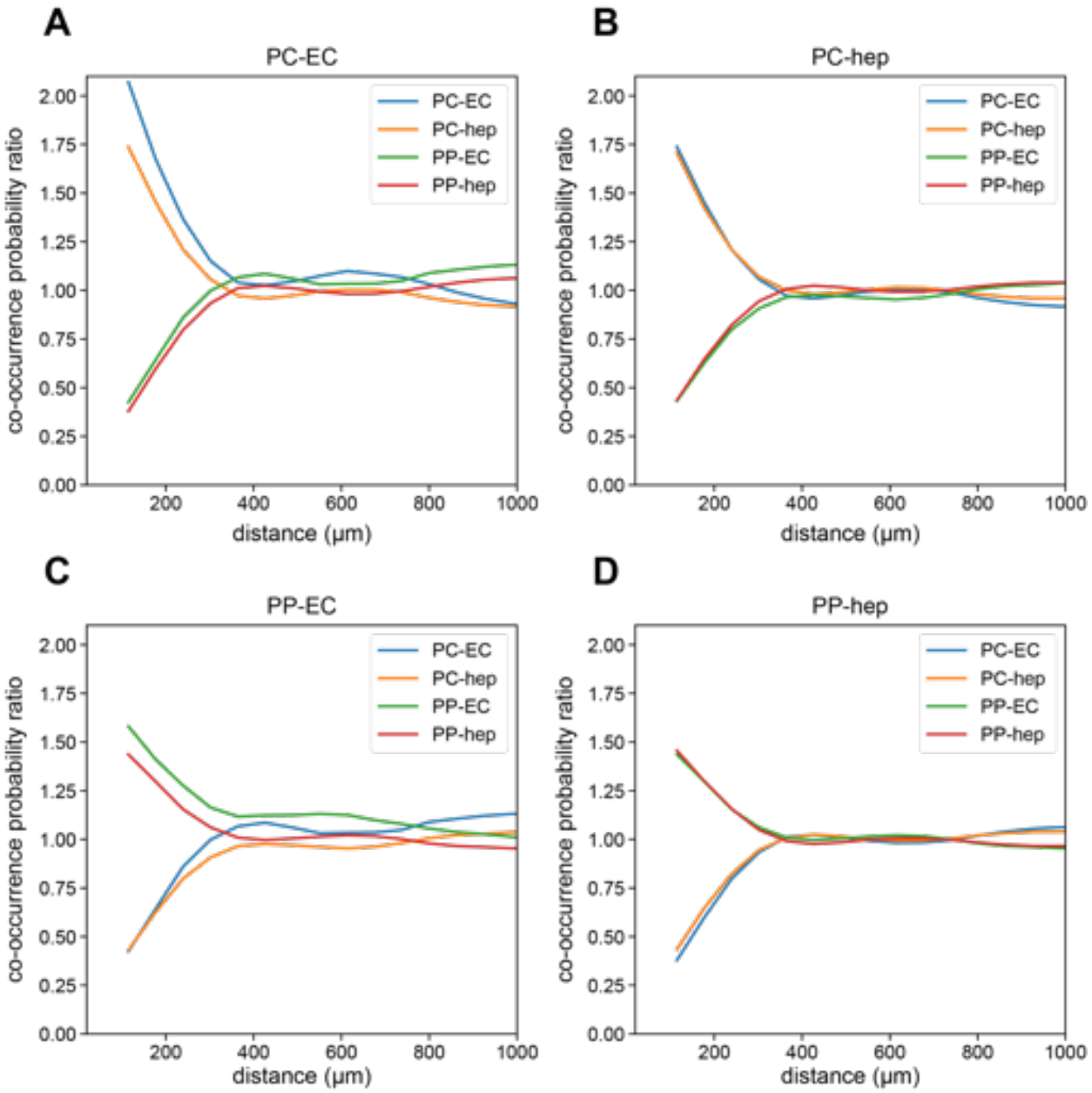
S3 Quantifying co-localization of periportal and pericentral endothelial cells and hepatocytes in mouse liver

To quantify the co-localization of periportal and pericentral endothelial cells and hepatocytes (Fig 6B inset), we computed the co-occurrence probability likelihood 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 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.

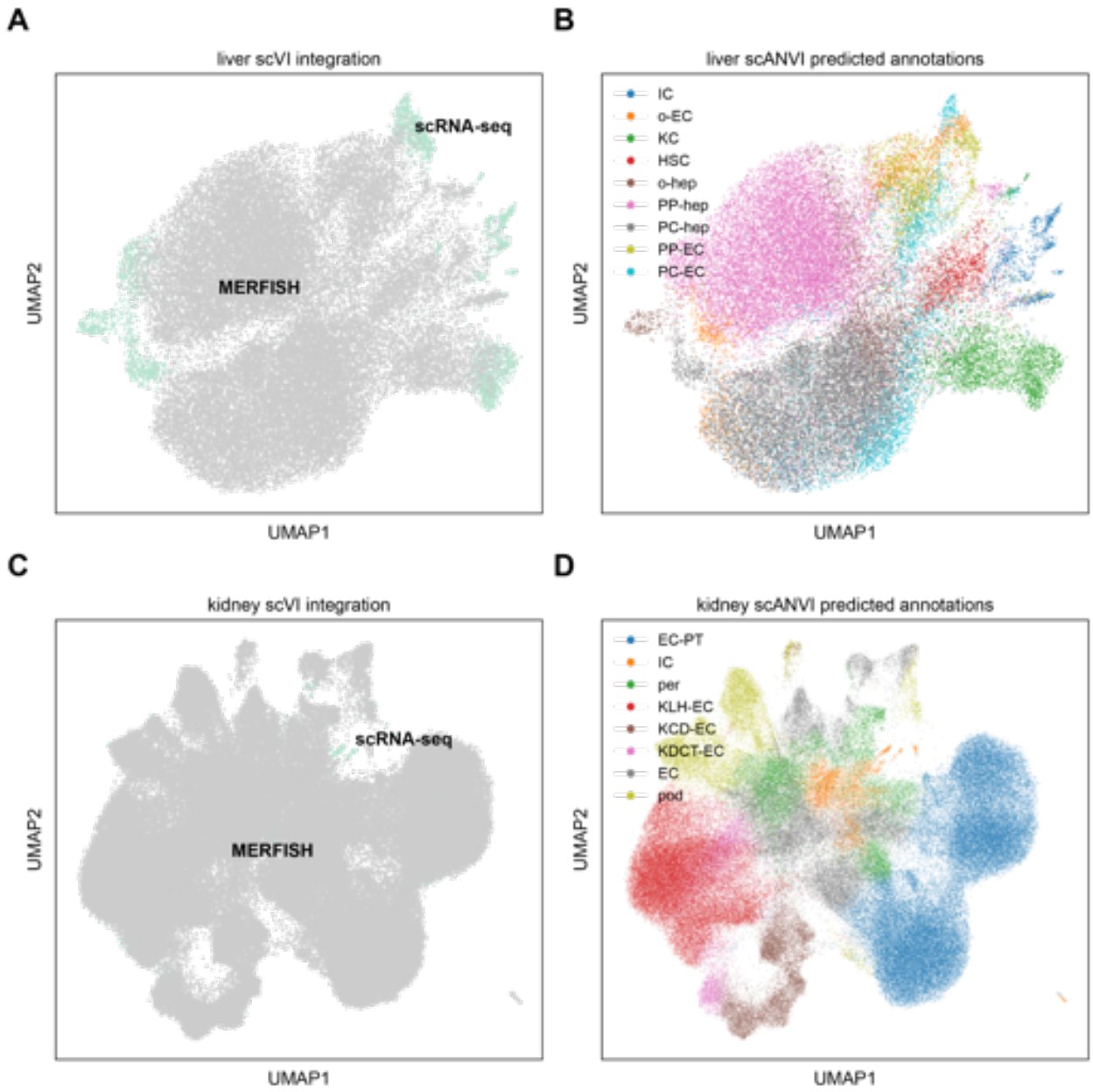


SI Figure 4 - Co-occurrence probability analysis of periportal and pericentral hepatocytes and endothelial cells. (A-D) Co-occurrence probability ratio of (A) pericentral endothelial cells, (B) pericentral hepatocytes, (C) periportal endothelial cells, and (D) periportal hepatocytes with other periportal 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'.

S4 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.

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.

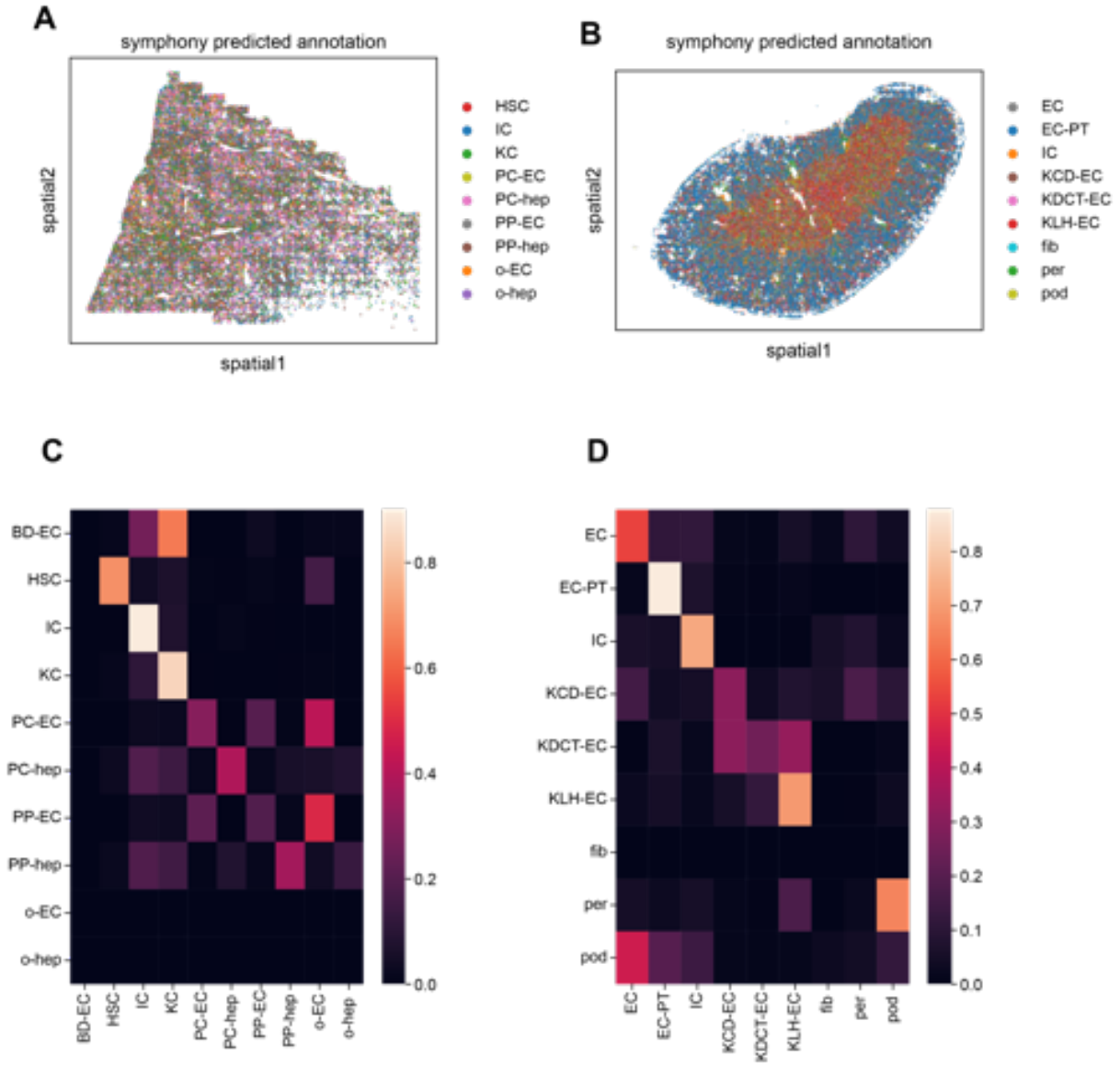


SI Figure 5 - Integration of MERFISH mouse liver and kidney data with Tabula Muris Senis scRNA-seq using scVI and scANVI. (A) UMAP plot of MERFISH liver data (gray) with Tabula Muris Senis (teal) using scVI. (B) UMAP plot of liver integration in (A) with cell type annotations predicted using scANVI. (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

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

S5 Integration of MERFISH data using Harmony and Symphony

To compare with our scVI and scANVI integration, we used Harmony (Korsunsky et al. 2019) and Symphony (Kang et al. 2021) to integrate our MERFISH data with the Tabula Muris Senis scRNA-seq reference and generate predicted cell type annotations for the MERFISH data. SI Figure 6A and 6B show the resulting spatial plots of mouse liver and kidney, colored by the Symphony predicted cell type annotation. Qualitatively, the results are similar to the scANVI results (Fig 6D and 7D), reproducing the structure of the liver (Fig 6D) as well as the artifactual population of podocytes in the kidney medulla (Fig 7D). SI Figure 6C and 6D show the confusion matrices of cell type annotations between the manual method and Symphony, for the mouse liver and kidney. Like with scANVI (Fig 6F and 7F), Symphony generally is consistent with manual annotation, as can be seen by the higher values along the main diagonal. However, Symphony produces slightly more off-diagonal values than scANVI and also misclassifies podocytes and pericytes (SI Fig 6D) just like scANVI. Thus, we conclude that Symphony integration produces generally similar results to scANVI.



SI Figure 6 - Integration of MERFISH mouse liver and kidney data with Tabula Muris Senis scRNA-seq using Harmony and Symphony. (A) Spatial plot of MERFISH mouse liver dataset, colored by cell type annotations predicted by Symphony. (B) Spatial plot of MERFISH mouse kidney dataset, colored by cell type annotations predicted by Symphony. (C) Confusion matrix of MERFISH cell type annotations between manual method and Symphony predictions for mouse liver. The rows for other endothelial cells and hepatocytes (“o-EC” and “o-hep”) are blank because none of the manual annotations were in these groups. Additionally, the row for bile duct epithelial cells (“BD-EC”) is noisy because none of the scRNA-seq reference cells possessed this annotation. (D) Confusion matrix of MERFISH cell type annotations between manual method and Symphony predictions for mouse kidney.

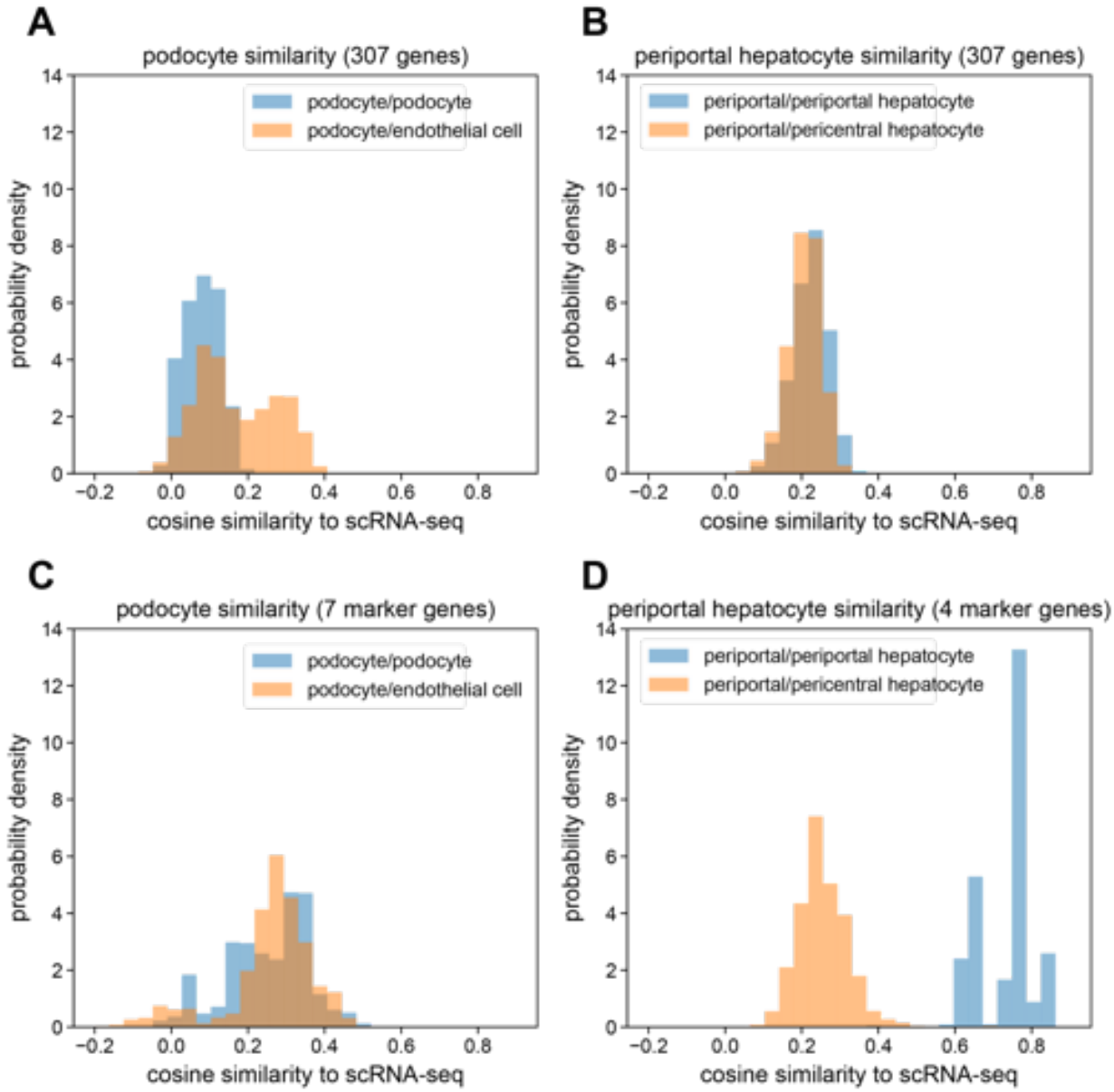
S6 Investigation of systematic differences between MERFISH and scRNA-seq data

The discrepancies produced by both scANVI and Symphony integration for cell types such as podocytes led us to investigate if the integration methods were at fault or if the issue lay within the data itself. To do so, we strove to quantify the similarity in RNA count profiles between different cell types to see if systematic differences between MERFISH and scRNA-seq were a possible source for the integration errors.

We reasoned that perhaps MERFISH RNA count profiles for a certain cell type were more similar to a different cell type in the scRNA-seq reference than for the same cell type. For example, in our kidney data, both integration methods consistently misclassified podocytes as endothelial cells (Fig 7F and SI Fig 7D). To quantify the source of this error, we used the cosine similarity to compare the similarity of MERFISH podocytes with scRNA-seq podocytes and endothelial cells. Specifically, for each MERFISH podocyte we calculated a mean cosine similarity value, defined thusly. For a given cell type pair (e.g. MERFISH podocyte vs. scRNA-seq podocyte), we calculated the cosine similarity between each individual pair of MERFISH/scRNA-seq cells. We then averaged across scRNA-seq cells, generating a mean cosine similarity value for each MERFISH cell and a distribution of mean cosine similarity values across MERFISH cells.

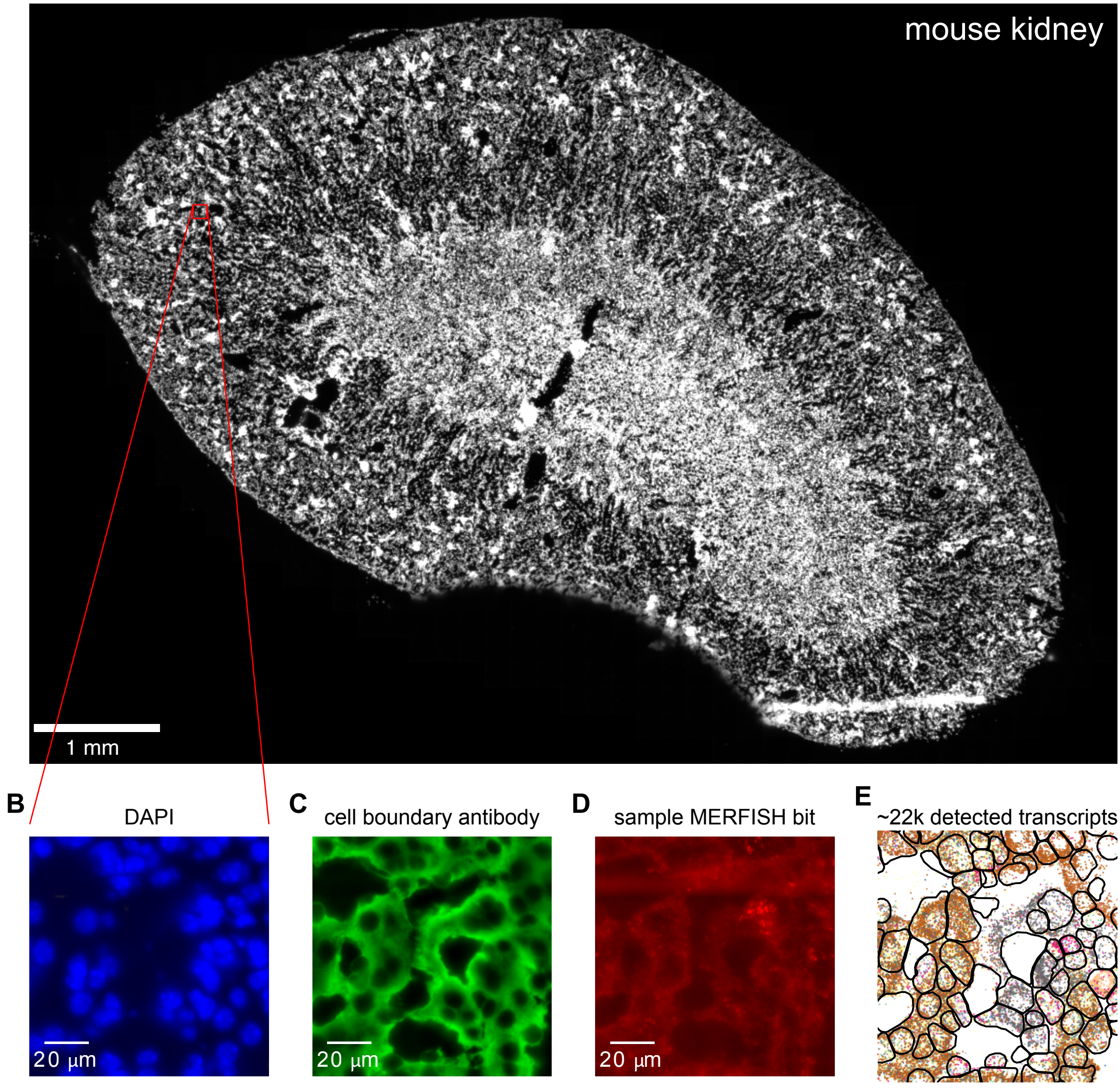
SI Figure 7A shows the distribution of mean cosine similarity values between MERFISH kidney podocytes and scRNA-seq kidney podocytes (blue) or scRNA-seq kidney endothelial cells (orange), calculated by using the cosine similarity across all 307 genes in the panel. Surprisingly, the distribution for kidney endothelial cells possesses larger values than for podocytes, suggesting that MERFISH podocytes may actually have more similar gene expression with scRNA-seq kidney endothelial cells. Even if we compute the mean cosine similarity using just the podocyte marker genes in the panel (*Wt1, Actn4, Synpo, Dag1, Foxc1, Podxl, Mme*), the MERFISH podocytes appear just as similar to scRNA-seq kidney endothelial cells as to scRNA-seq podocytes (SI Fig 7C). We note that this may stem from technical errors such as mis-segmentation of MERFISH podocytes or inaccurate annotation in the reference; nevertheless, we conclude that this is the main reason for the inaccurate cell type annotation prediction from integration.

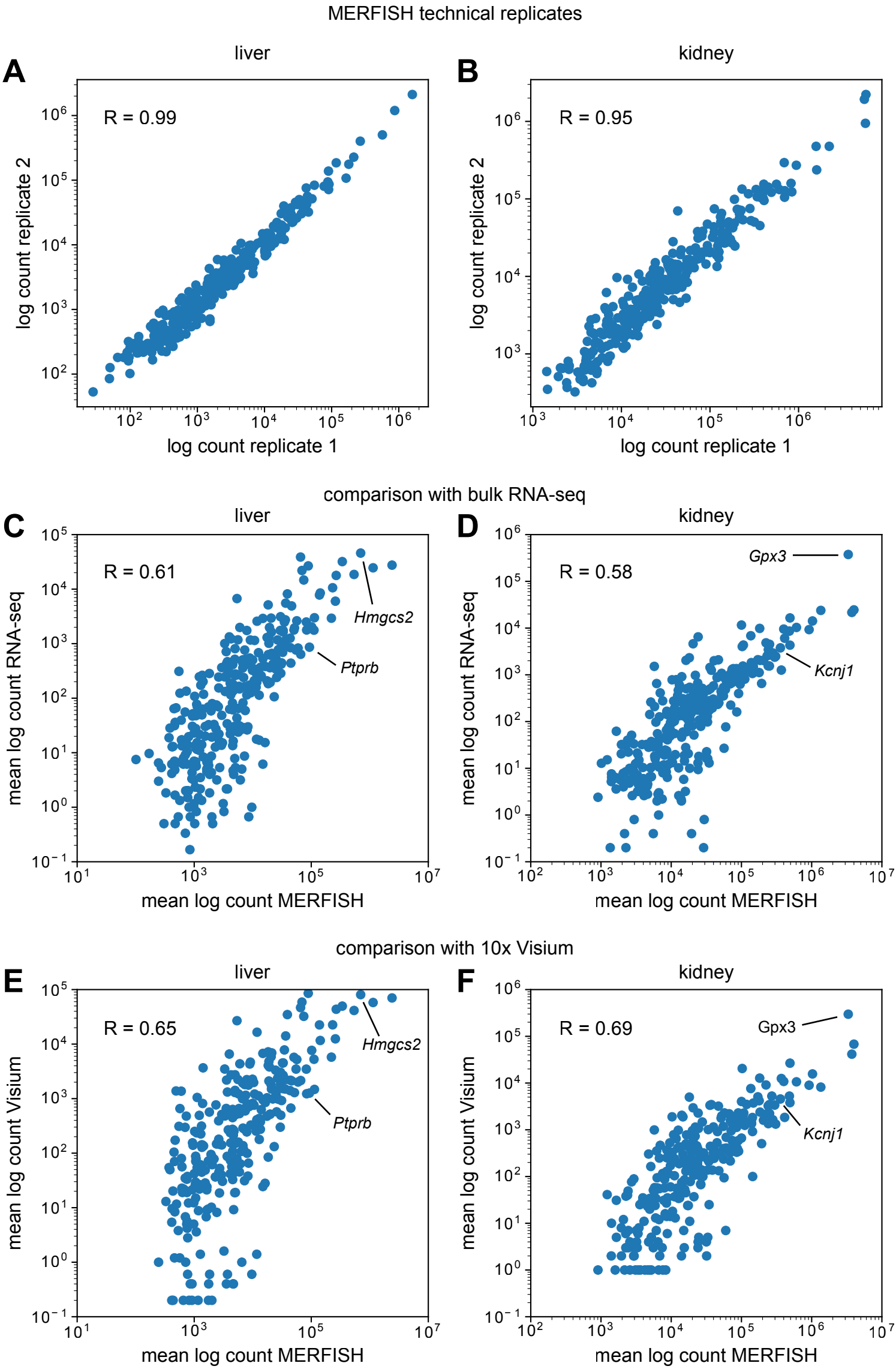
As a control, we then compared mean cosine similarity values between MERFISH liver periportal hepatocytes and scRNA-seq liver periportal hepatocytes or pericentral hepatocytes (SI Fig 7B). These were cells that were very accurately classified with the integration methods (Fig 6F and SI Fig 6C).

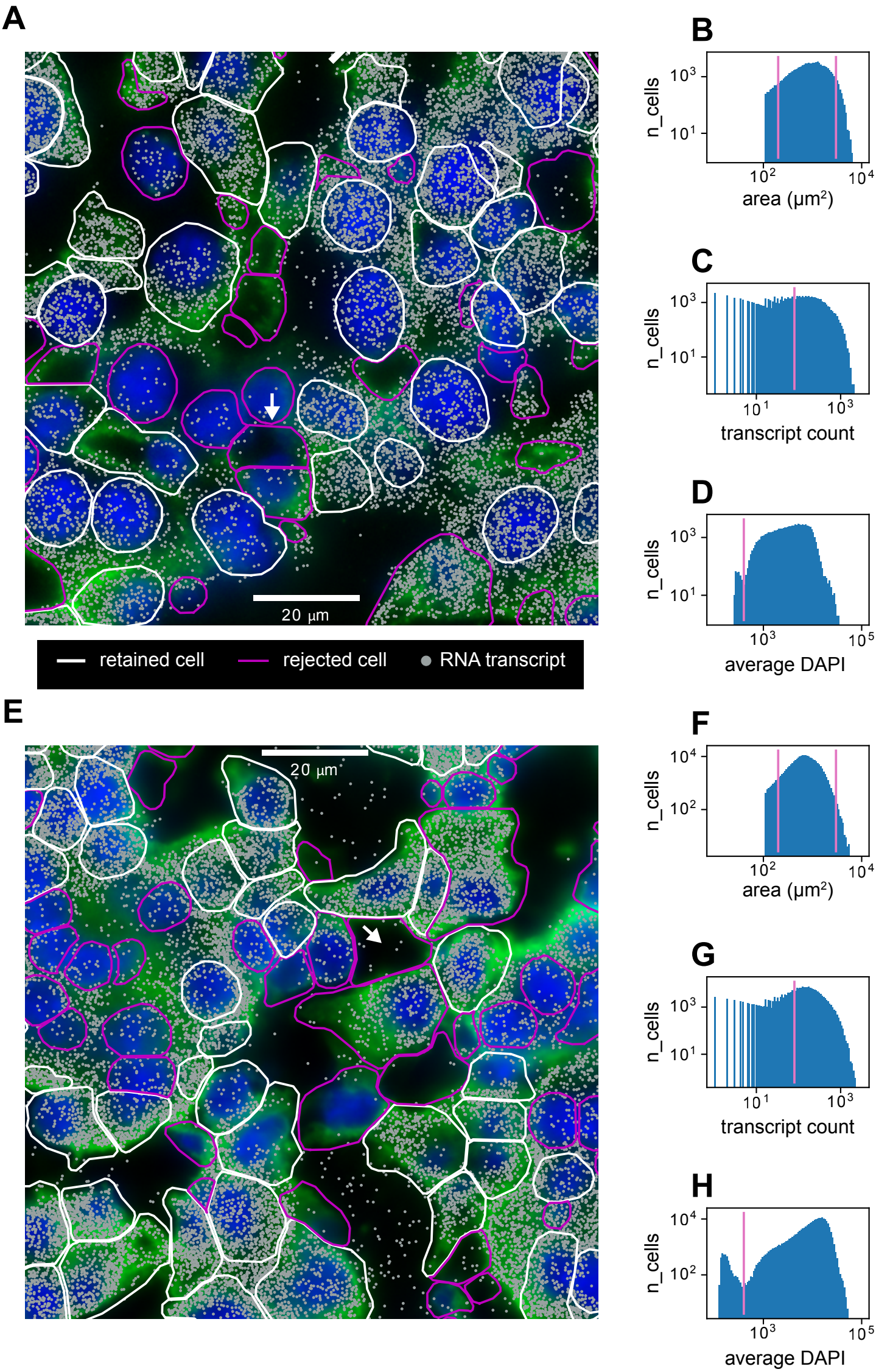


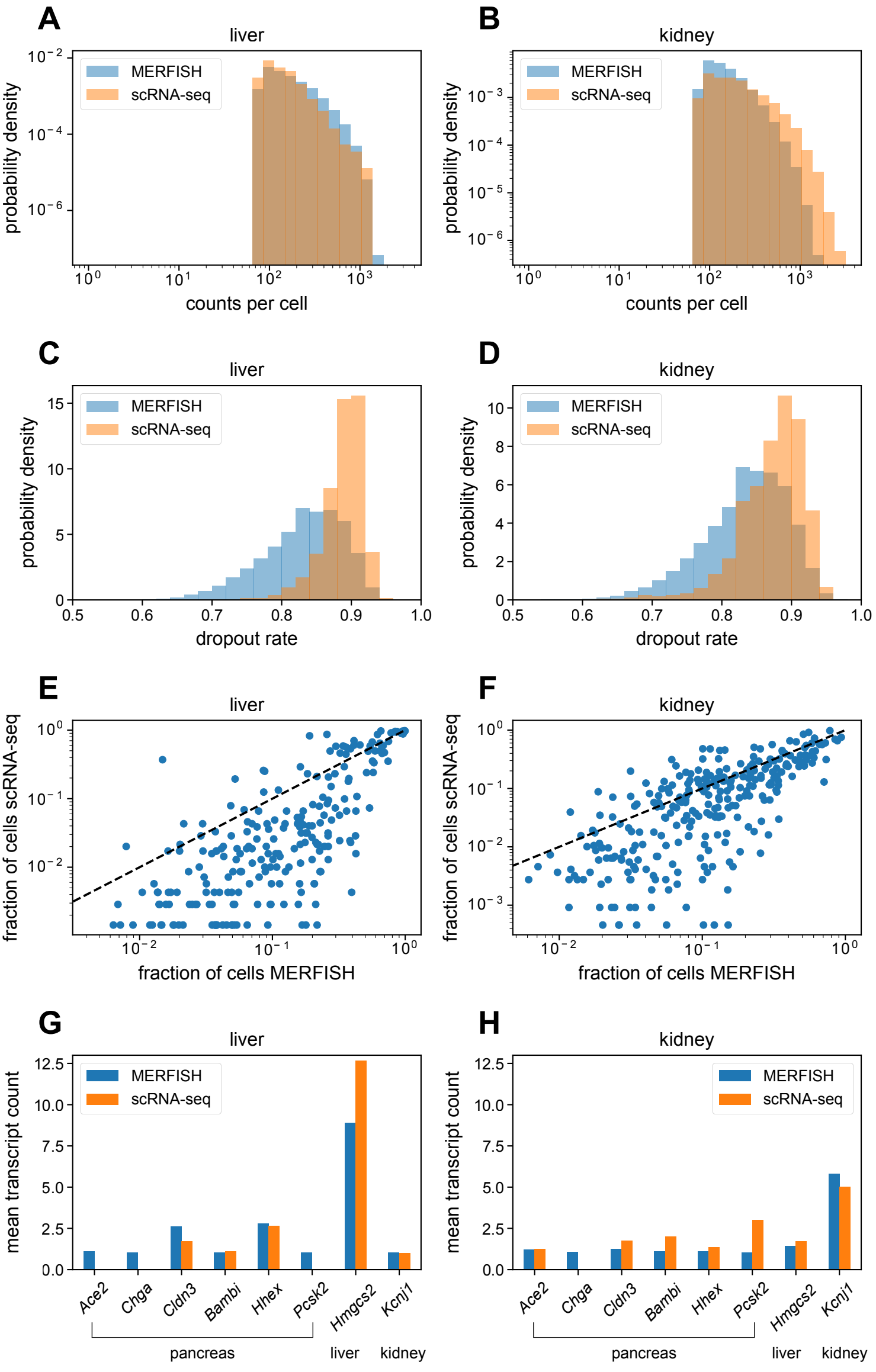
SI Figure 7 - Investigation of systematic differences in RNA expression profiles between MERFISH and scRNA-seq. (A) Distribution of mean cosine similarity values using the full 307-gene panel between manually annotated MERFISH kidney podocytes and cells in the scRNA-seq kidney reference labeled as podocyte (blue) or endothelial cell (orange). (B) Distribution of mean cosine similarity values using the full 307-gene panel between manually annotated MERFISH liver periportal hepatocytes and cells in the scRNA-seq liver reference labeled as periportal hepatocyte (blue) or pericentral hepatocyte (orange). (C) Same as (A) but only using seven podocyte marker genes (*Wt1, Actn4, Synpo, Dag1, Foxc1, Podxl, Mme*). (D) Same as (B) but only using four periportal hepatocyte marker genes (*Cyp2f2, Pck1, Hal, Cdh1*). Mean cosine similarity values in A-D were obtained by calculating the cosine similarity between each

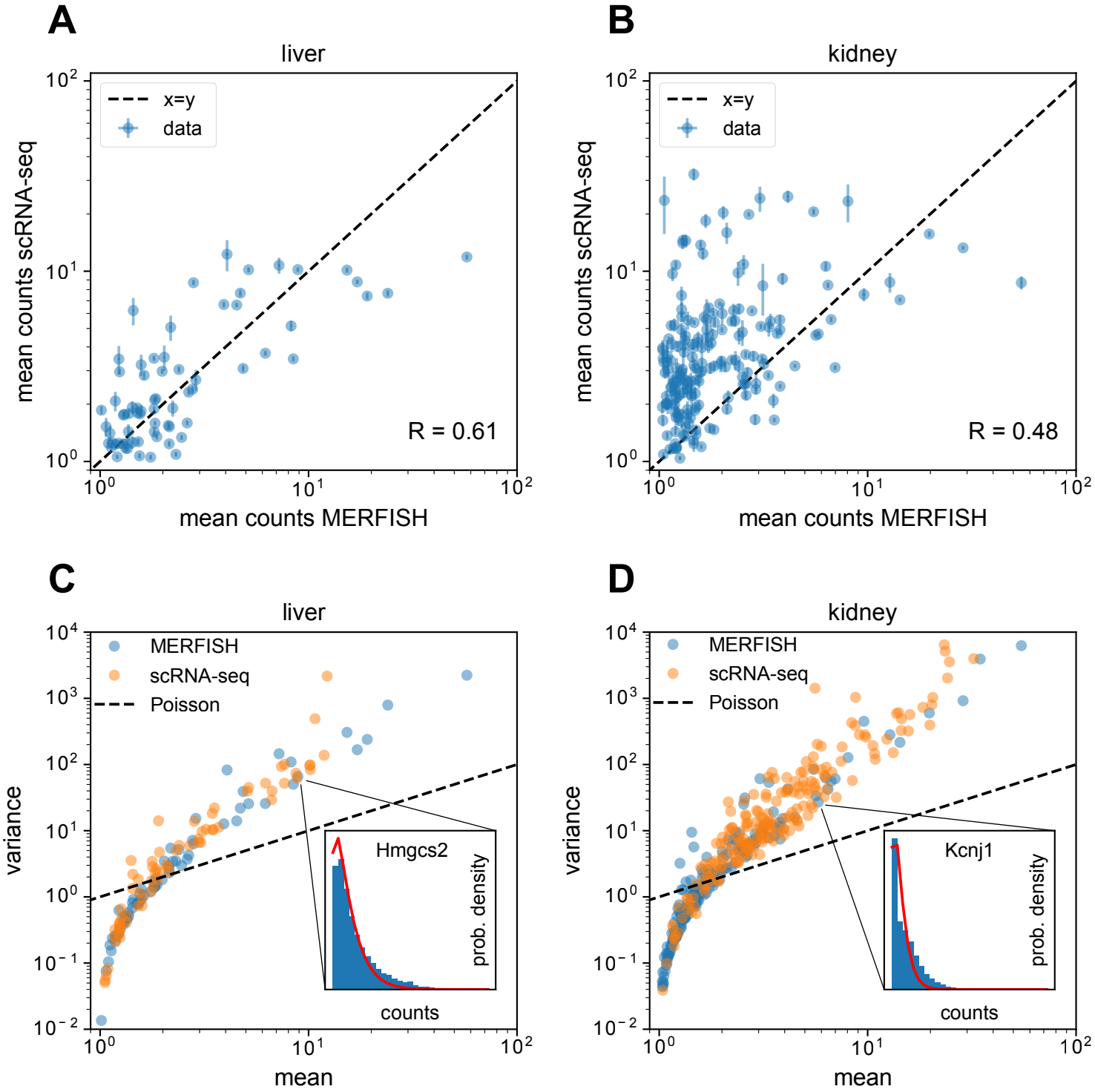
individual cell in a pair (e.g. MERFISH podocyte vs. scRNA-seq podocyte) and then averaging across the scRNA-seq cells, yielding a mean cosine similarity value for each MERFISH cell and a distribution of mean cosine similarity values across MERFISH cells.

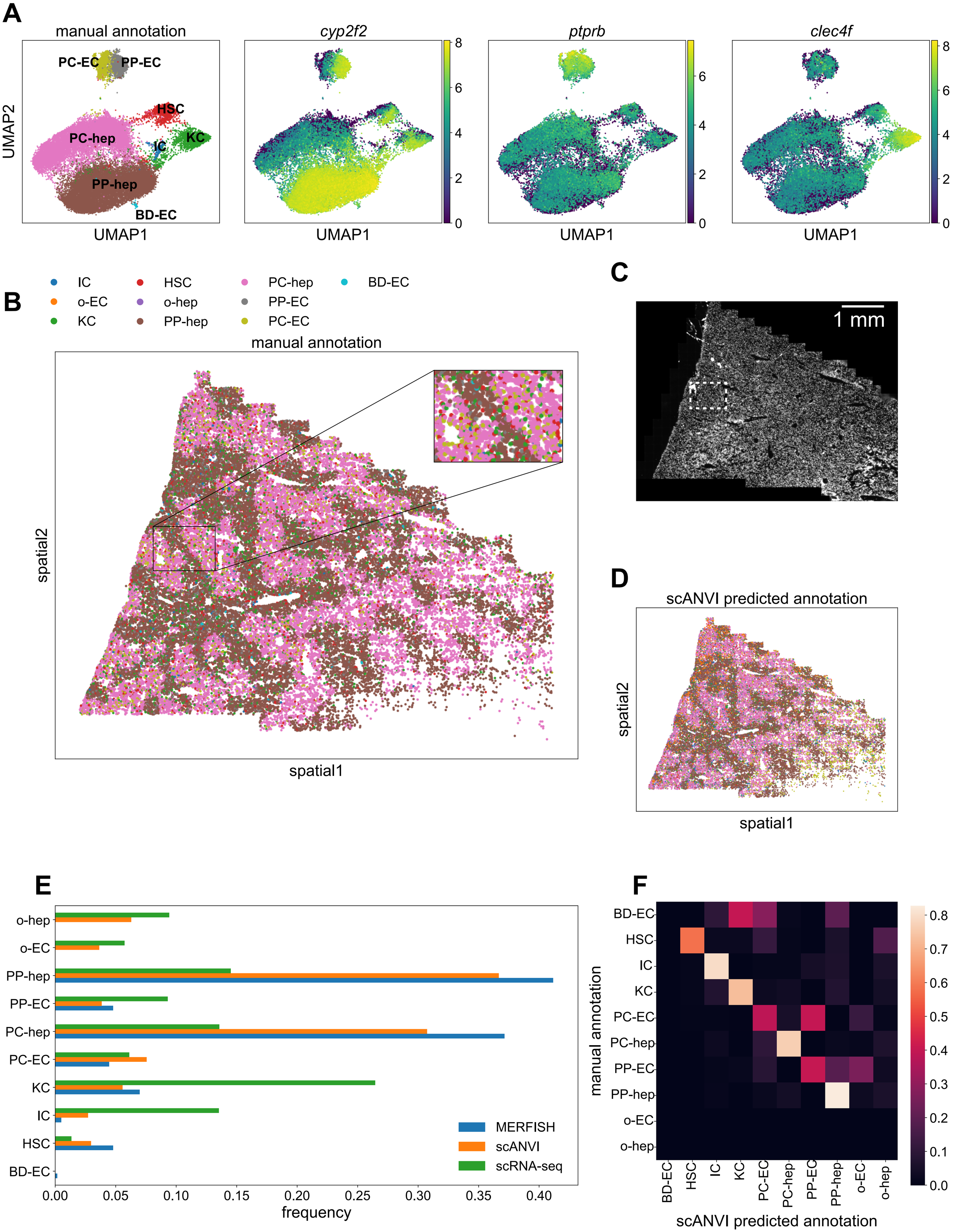

A
mouse kidney
1 mm
B
DAPI
20 µm
C
cell boundary antibody
20 µm
D
sample MERFISH bit
20 µm
E
~22k detected transcripts

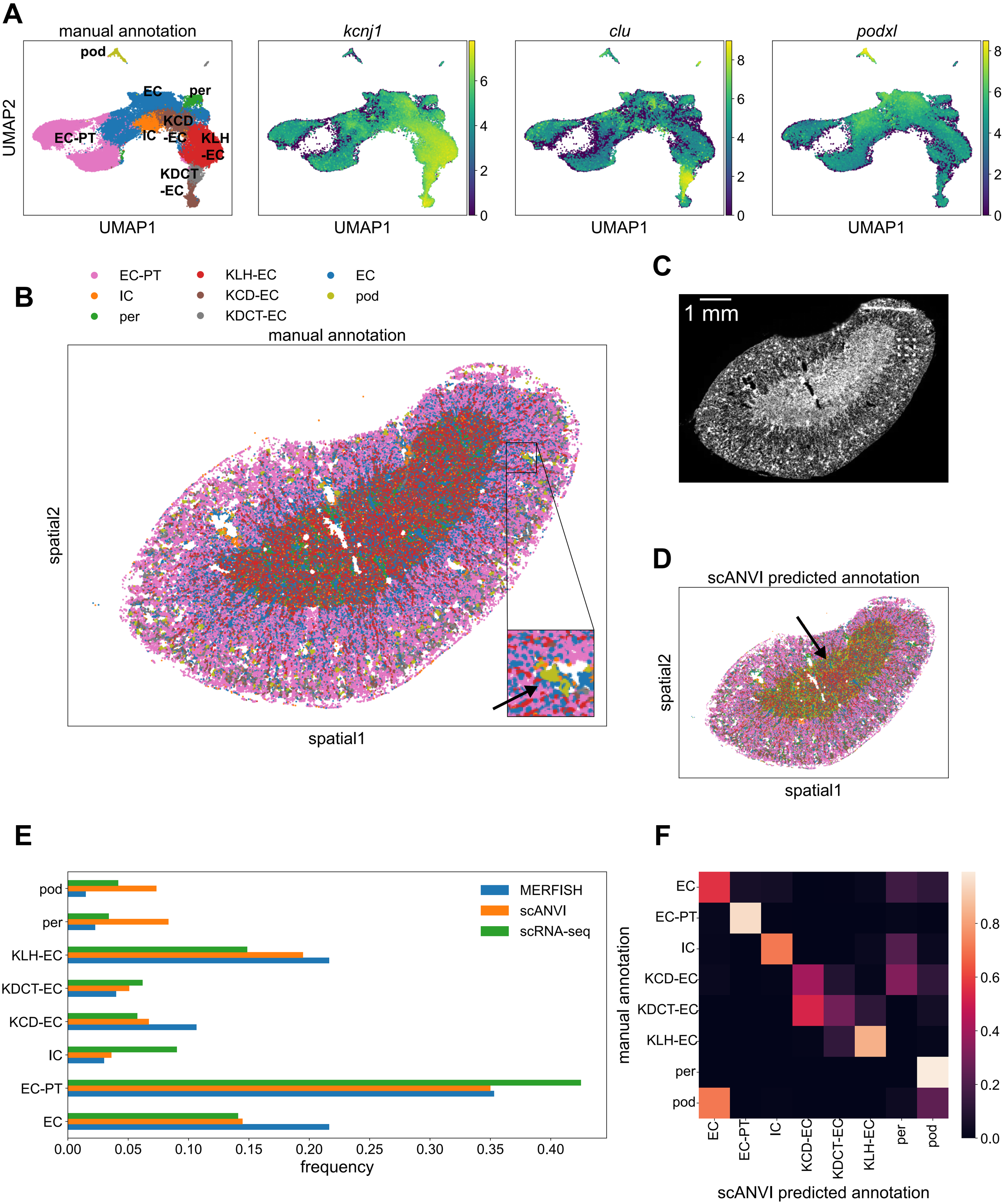

MERFISH technical replicates
A
liver
R = 0.99
log count replicate 2
log count replicate 1
B
kidney
R = 0.95
log count replicate 2
log count replicate 1
comparison with bulk RNA-seq
C
liver
R = 0.61
Hmgcs2
Ptprb
mean log count RNA-seq
mean log count MERFISH
D
kidney
R = 0.58
Gpx3
Kcnj1
mean log count RNA-seq
mean log count MERFISH
comparison with 10x Visium
E
liver
R = 0.65
Hmgcs2
Ptprb
mean log count Visium
mean log count MERFISH
F
kidney
R = 0.69
Gpx3
Kcnj1
mean log count Visium
mean log count MERFISH


A
20 µm
retained cell
rejected cell
RNA transcript
B
n_cells
area (µm²)
C
n_cells
transcript count
D
n_cells
average DAPI
E
20 µm
F
n_cells
area (µm²)
G
n_cells
transcript count
H
n_cells
average DAPI

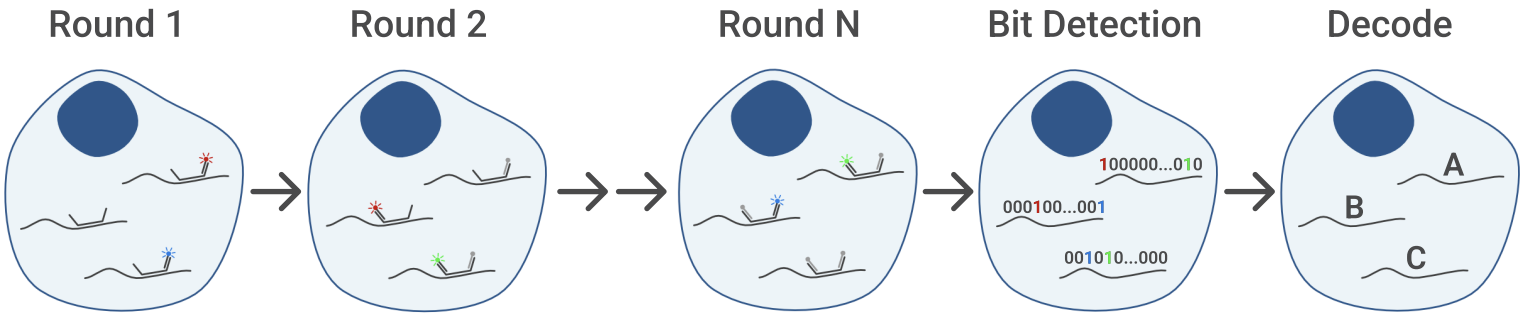

A
liver
MERFISH
scRNA-seq
probability density
counts per cell
B
kidney
MERFISH
scRNA-seq
probability density
counts per cell
C
liver
MERFISH
scRNA-seq
probability density
dropout rate
D
kidney
MERFISH
scRNA-seq
probability density
dropout rate
E
liver
fraction of cells scRNA-seq
fraction of cells MERFISH
F
kidney
fraction of cells scRNA-seq
fraction of cells MERFISH
G
liver
MERFISH
scRNA-seq
mean transcript count
Ace2
Chga
Cldn3
Bambi
Hlhex
Pcsk2
Hmgcs2
Kcnj1
pancreas
liver
kidney
H
kidney
MERFISH
scRNA-seq
mean transcript count
Ace2
Chga
Cldn3
Bambi
Hlhex
Pcsk2
Hmgcs2
Kcnj1
pancreas
liver
kidney


A
liver
x=y
data
R = 0.61
mean counts scRNA-seq
mean counts MERFISH
B
kidney
x=y
data
R = 0.48
mean counts scRNA-seq
mean counts MERFISH
C
liver
MERFISH
scRNA-seq
Poisson
Hmgcs2
prob. density
counts
variance
mean
D
kidney
MERFISH
scRNA-seq
Poisson
Kcnj1
prob. density
counts
variance
mean

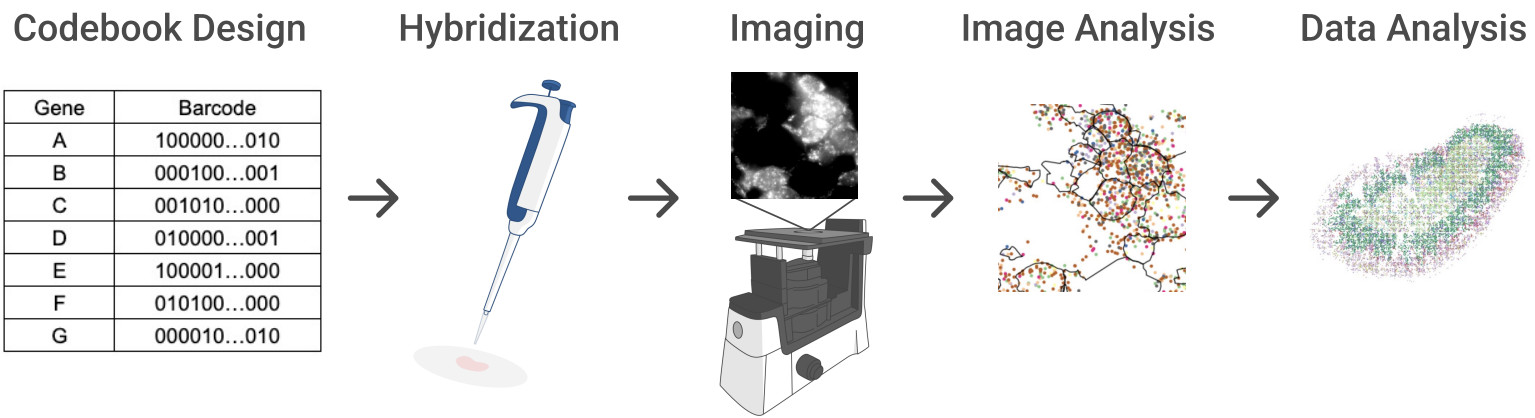

A
manual annotation
PC-EC
PP-EC
HSC
PC-hep
IC
KC
PP-hep
BD-EC
UMAP2
UMAP1
cyp2f2
ptprb
clec4f
B
IC
HSC
PC-hep
BD-EC
o-EC
o-hep
PP-EC
KC
PP-hep
PC-EC
manual annotation
spatial2
spatial1
C
1 mm
D
scANVI predicted annotation
spatial2
spatial1
E
o-hep
o-EC
PP-hep
PP-EC
PC-hep
PC-EC
KC
IC
HSC
BD-EC
MERFISH
scANVI
scRNA-seq
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
frequency
F
manual annotation
scANVI predicted annotation
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8


A
manual annotation
kcnj1
clu
podxl
pod
EC
per
KCD
-EC
EC-PT
IC
KLH
-EC
KDCT
-EC
UMAP2
UMAP1
B
EC-PT
IC
per
KLH-EC
KCD-EC
KDCT-EC
EC
pod
manual annotation
spatial2
spatial1
C
1 mm
D
scANVI predicted annotation
spatial2
spatial1
E
MERFISH
scANVI
scRNA-seq
pod
per
KLH-EC
KDCT-EC
KCD-EC
IC
EC-PT
EC
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
frequency
F
manual annotation
scANVI predicted annotation
EC
EC-PT
IC
KCD-EC
KDCT-EC
KLH-EC
per
pod

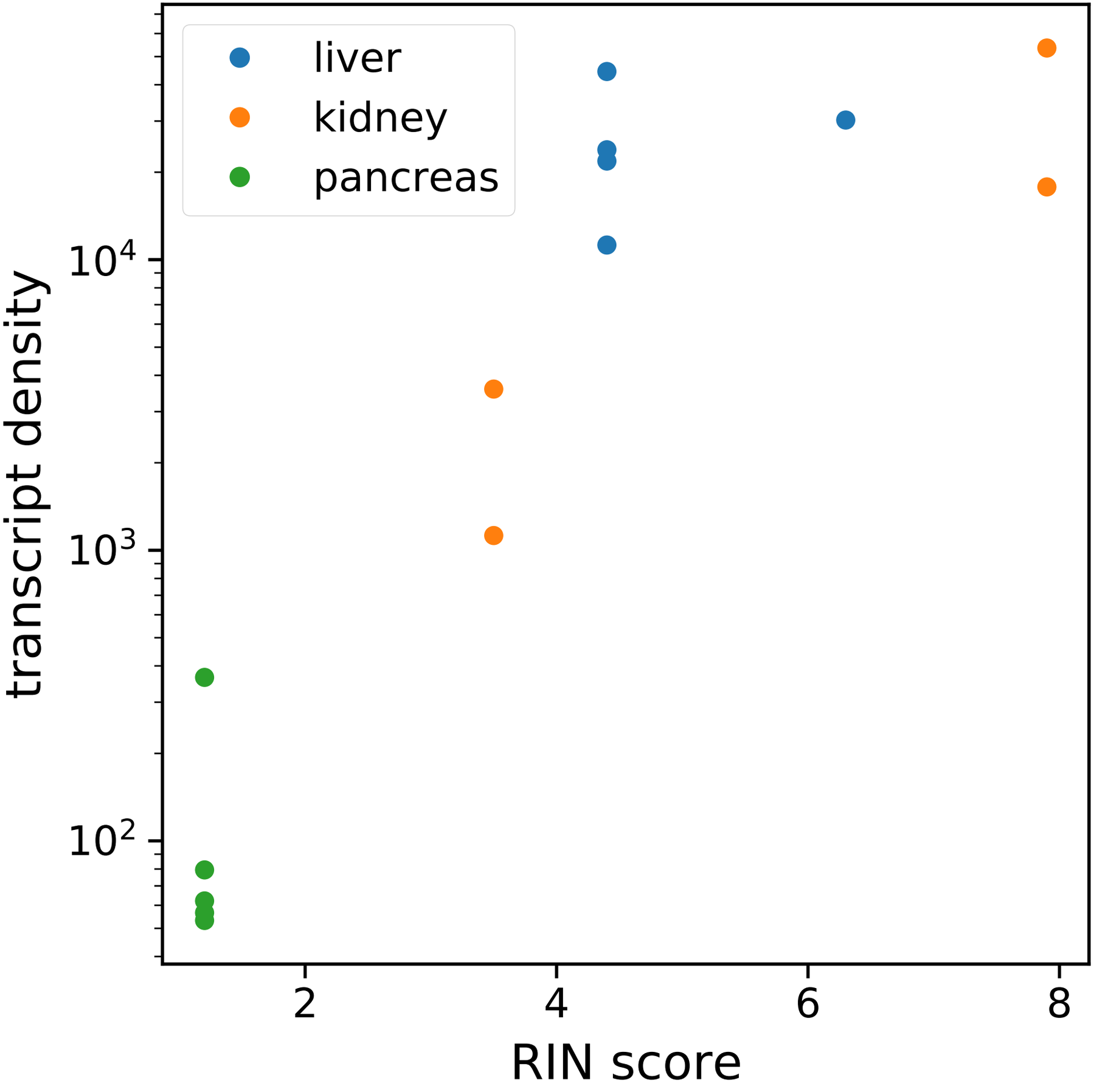
A

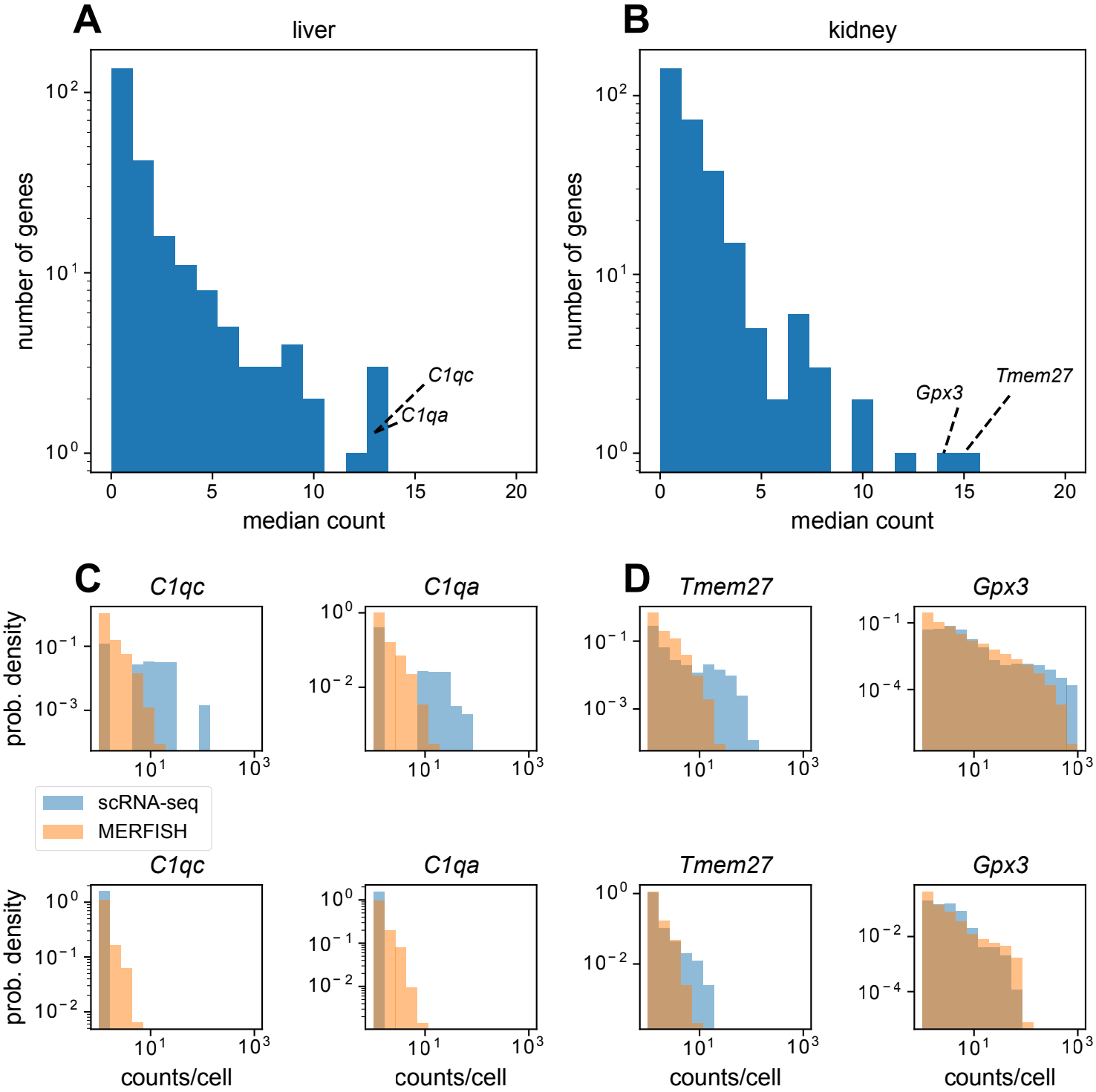


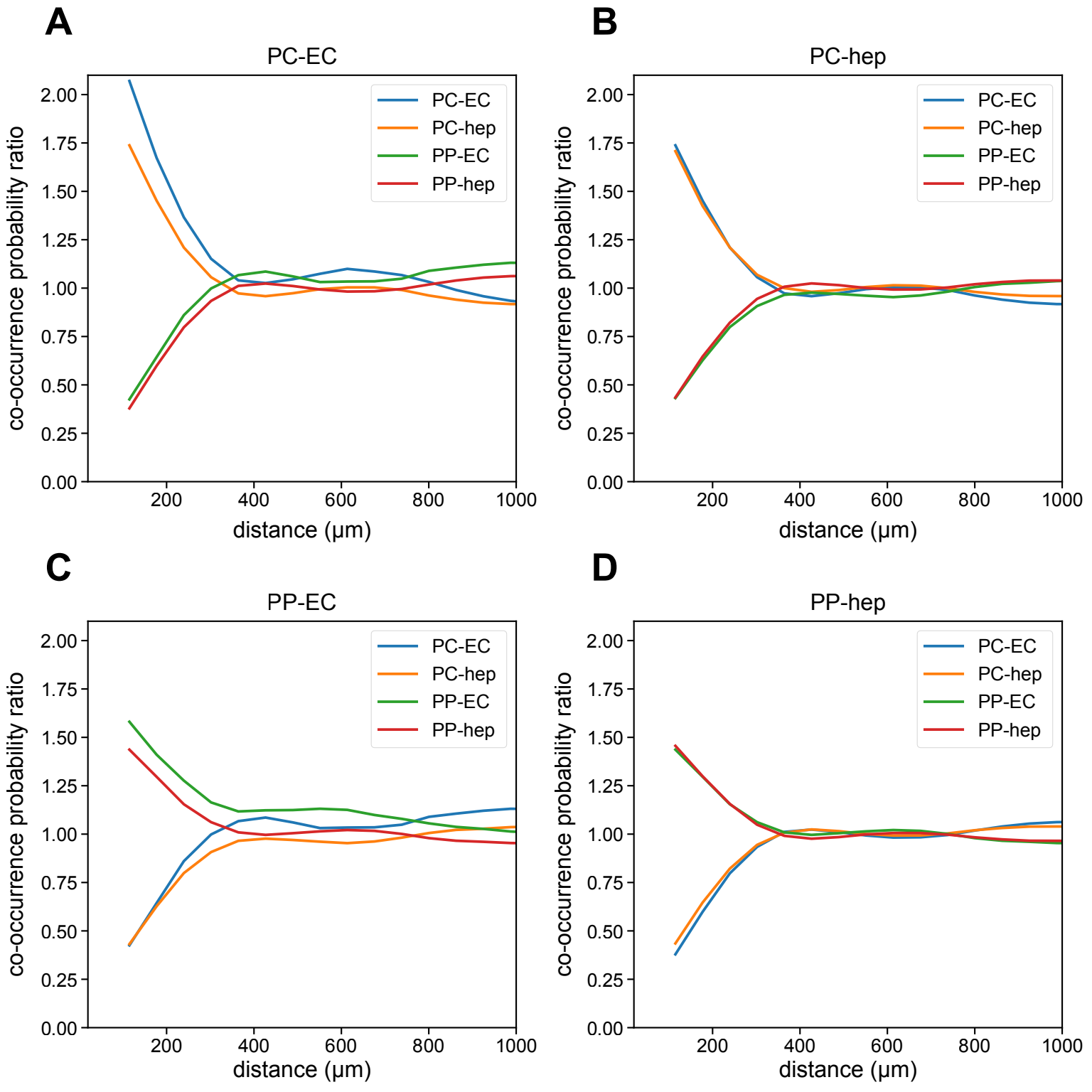
B

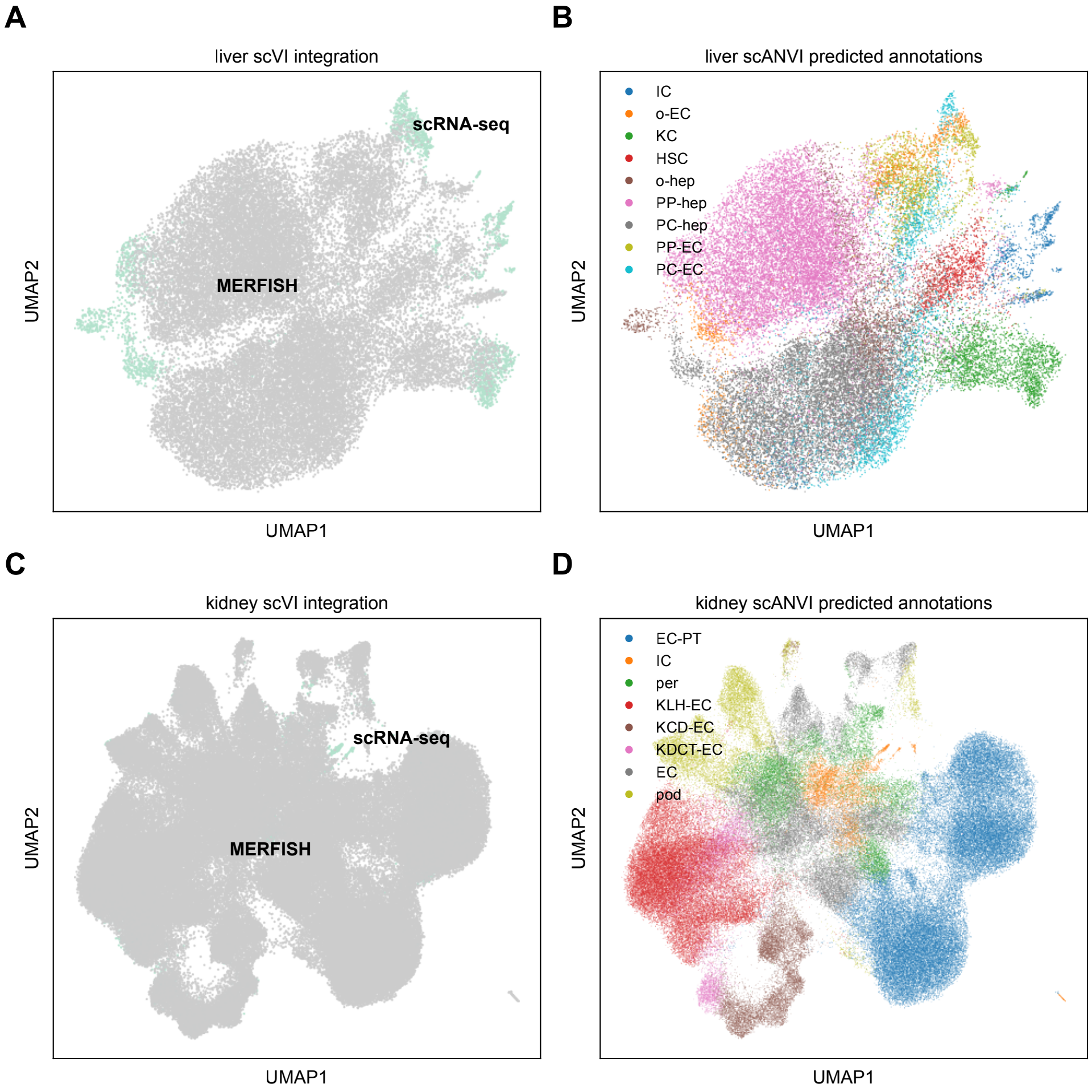


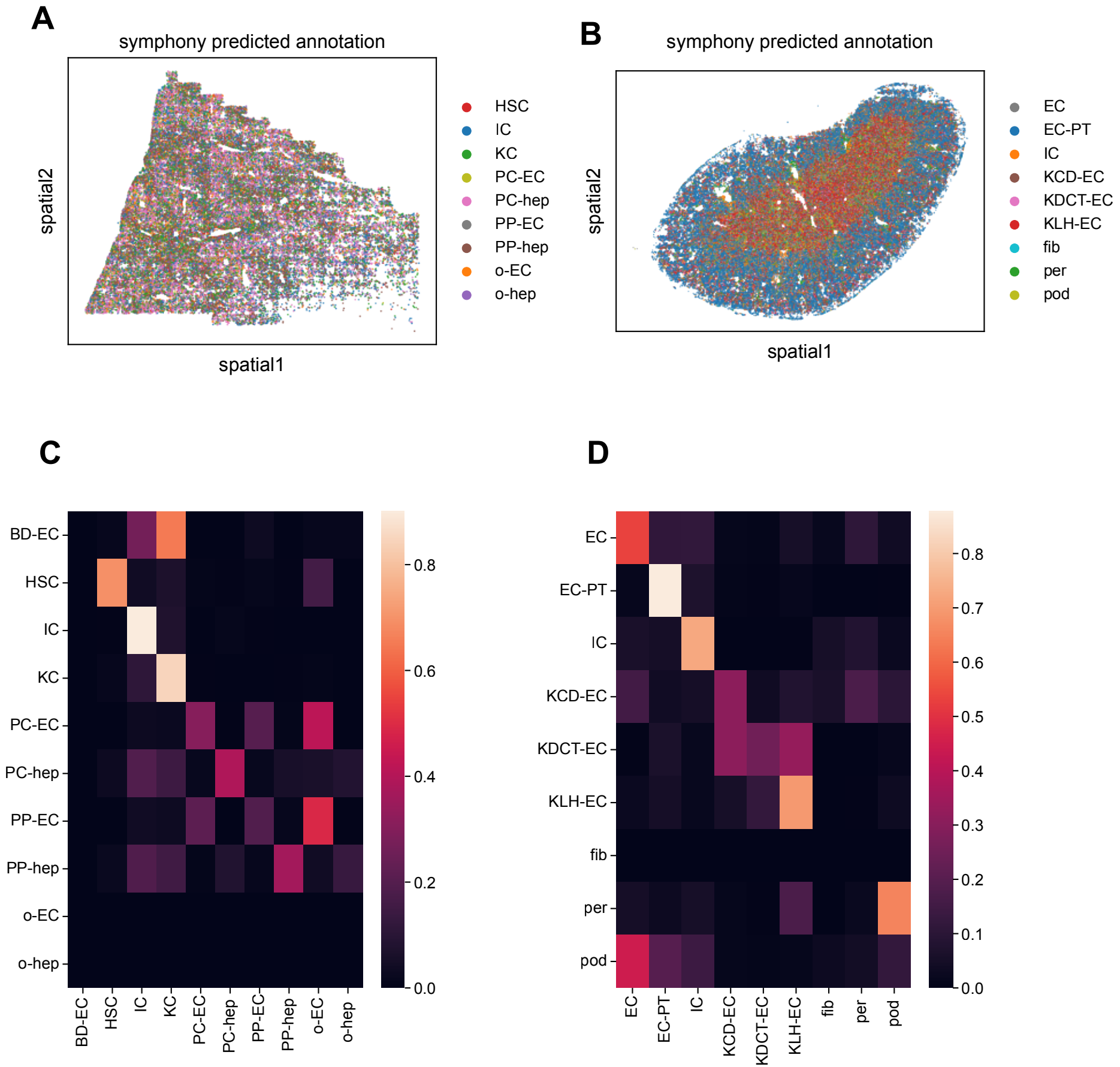
Gene	Barcode
A	100000…010
B	000100…001
C	001010…000
D	010000…001
E	100001…000
F	010100…000
G	000010…010

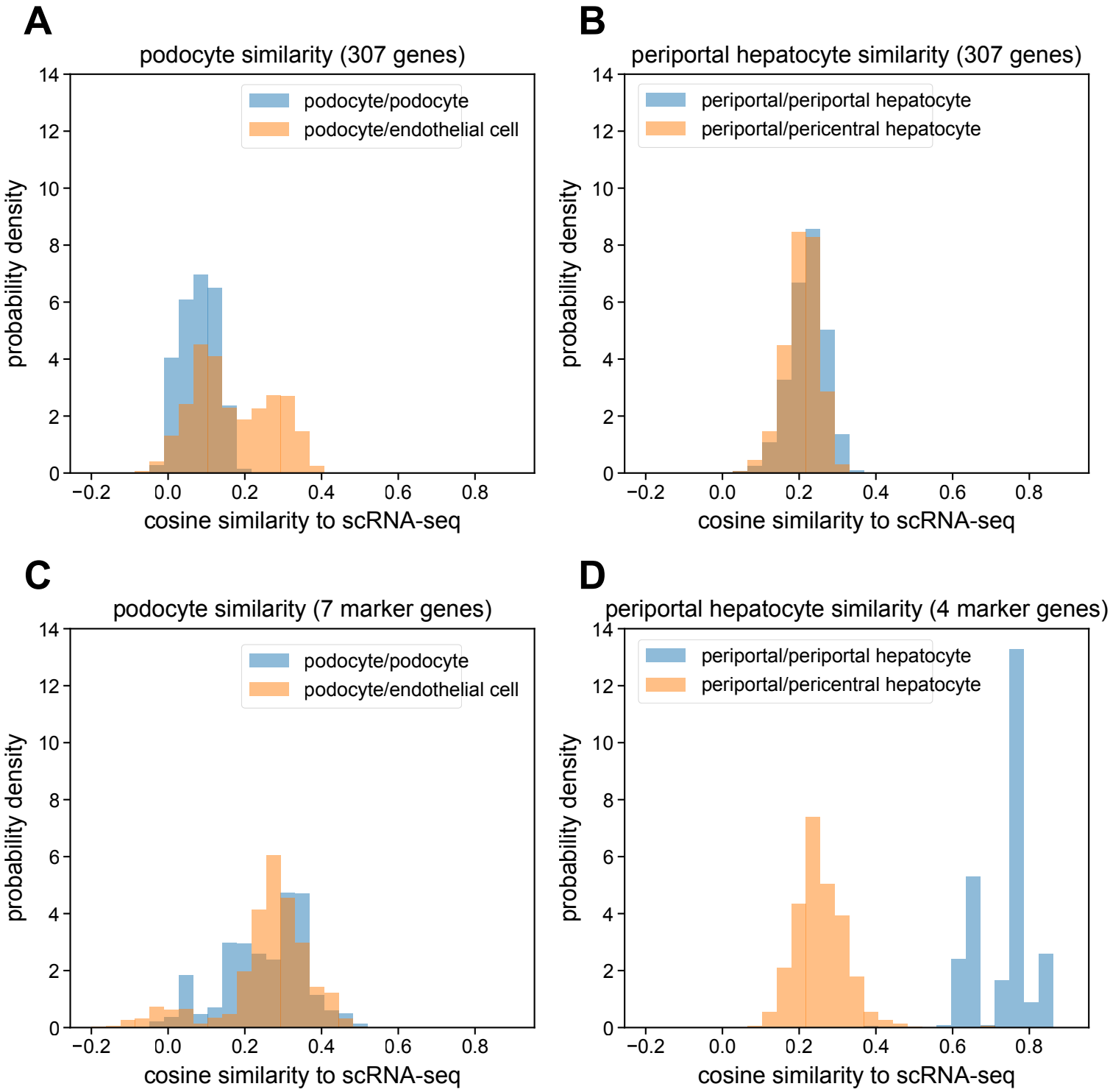

liver
kidney
pancreas
transcript density
10^4
10^3
10^2
2
4
6
8
RIN score


A
liver
number of genes
median count
C1qc
C1qa
B
kidney
number of genes
median count
Gpx3
Tmem27
C
C1qc
C1qa
prob. density
counts/cell
D
Tmem27
Gpx3
scRNA-seq
MERFISH


A
PC-EC
B
PC-hep
C
PP-EC
D
PP-hep
PC-EC
PC-hep
PP-EC
PP-hep
co-occurrence probability ratio
distance (μm)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
200
400
600
800
1000


A
liver scVI integration
scRNA-seq
MERFISH
UMAP2
UMAP1
B
liver scANVI predicted annotations
IC
o-EC
KC
HSC
o-hep
PP-hep
PC-hep
PP-EC
PC-EC
UMAP2
UMAP1
C
kidney scVI integration
scRNA-seq
MERFISH
UMAP2
UMAP1
D
kidney scANVI predicted annotations
EC-PT
IC
per
KLH-EC
KCD-EC
KDCT-EC
EC
pod
UMAP2
UMAP1


A
symphony predicted annotation
spatial2
spatial1
HSC
IC
KC
PC-EC
PC-hep
PP-EC
PP-hep
o-EC
o-hep
B
symphony predicted annotation
spatial2
spatial1
EC
EC-PT
IC
KCD-EC
KDCT-EC
KLH-EC
fib
per
pod
C
BD-EC
HSC
IC
KC
PC-EC
PC-hep
PP-EC
PP-hep
o-EC
o-hep
0.8
0.6
0.4
0.2
0.0
D
EC
EC-PT
IC
KCD-EC
KDCT-EC
KLH-EC
fib
per
pod
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0


A
podocyte similarity (307 genes)
podocyte/podocyte
podocyte/endothelial cell
probability density
cosine similarity to scRNA-seq
B
periportal hepatocyte similarity (307 genes)
periportal/periportal hepatocyte
periportal/pericentral hepatocyte
probability density
cosine similarity to scRNA-seq
C
podocyte similarity (7 marker genes)
podocyte/podocyte
podocyte/endothelial cell
probability density
cosine similarity to scRNA-seq
D
periportal hepatocyte similarity (4 marker genes)
periportal/periportal hepatocyte
periportal/pericentral hepatocyte
probability density
cosine similarity to scRNA-seq