nature methods

Article

https://doi.org/10.1038/s41592-026-03237-0

Resolving cell–cell interaction networks and their molecular logic in complex tissues

Received: 26 February 2025

Accepted: 17 August 2026

Published online: 5 October 2026

Check for updates

Lei Tang[1,2,3,8], Kang Tian[1,2,3,8] ✉, Xiaolei Fu[1,2,3,4,8], Yingjia Xu[1,2,5,8], Jincheng Wu[4], Jinsong Zhang[1,2,3], Xi-Wen Wang[1,2,3], Chunxiang Ye[6], Qiong Wu[4], Wei Wu[4] ✉, Changjiang Feng[7] ✉ & Qiangfeng Cliff Zhang[1,2,3] ✉

Cells in complex organisms function through extensive interactions, yet mapping these interaction networks at scale remains challenging. Here we present CCI-seq, a high-throughput method to unbiasedly capture cell–cell interactions across a proximity continuum by combining cell clump combinatorial indexing with single-cell sequencing. CCI-seq identified known interactions and fine-grained cellular organization in mouse kidney and intestine, and uncovered aberrant interactions and disrupted spatial organization in adenomatous polyposis coli knockout (*Apc*-KO) intestines. Applied to human colorectal cancer, it revealed subtype-specific interactions linked to clinical classifications. By leveraging a single-cell RNA sequencing backbone, CCI-seq achieves deep transcriptome coverage, enabling analysis of molecular states arising from interactions. This functional resolution revealed that interacting cells associate with distinct transcriptional programs—driving ion transport in kidney, antigen presentation in *Apc*-KO villi and NF-κB inflammatory activation in colorectal cancer. Thus, CCI-seq provides a scalable platform to unveil large-scale cell–cell interaction networks and dissect their molecular and functional impacts, enhancing our understanding of complex multicellular systems.

Higher-order biological structures in complex organisms are composed of diverse types of cells with distinct architectural and functional organization[1]. Knowledge about the cell–cell interactions that build these structures is crucial for elucidating their functions. These interactions encompass a continuum from direct physical contact to short-range spatial neighbors[2]. While single-cell RNA sequencing (scRNA-seq) reveals cell type composition, it lacks spatial information about cell organization[3]. In contrast, spatial transcriptomics (ST) methods preserve this information and generate a spatially resolved view of gene activity, which has led to a markedly improved understanding of the functions of complex biological structures[3,4]. Nonetheless, even the most advanced ST workflows have limitations (including those achieving subcellular resolution)[5–10], such as a restricted number of detectable genes[11], difficulties in accurately delineating cell boundaries[12–14], and the need for special devices and high costs[15], all of which impede the broader application of these approaches.

[1]State Key Laboratory of Membrane Biology, Beijing Advanced Innovation Center for Structural Biology & Frontier Research Center for Biological Structure, School of Life Sciences, Tsinghua University, Beijing, China. [2]MOE Key Laboratory of Bioinformatics, Center for Life Sciences and Artificial Intelligence, Center for Synthetic and Systems Biology, School of Life Sciences, Tsinghua University, Beijing, China. [3]Tsinghua-Peking Center for Life Sciences, Beijing, China. [4]State Key Laboratory of Molecular Oncology, MOE Key Laboratory of Protein Science, Beijing Advanced Innovation Center for Structural Biology, School of Life Sciences, Tsinghua University, Beijing, China. [5]Peking University-Tsinghua University-National Institute of Biological Sciences Joint Graduate Program, School of Life Sciences, Tsinghua University, Beijing, China. [6]Department of General Surgery, Beijing Chao-Yang Hospital, Capital Medical University, No. 8 Gongtinan Road, Chaoyang District, Beijing, China. [7]Department of Thoracic Surgery, the First Medical Center, Beijing 301 Hospital, Beijing, China. [8]These authors contributed equally: Lei Tang, Kang Tian, Xiaolei Fu, Yingjia Xu.
✉e-mail: tiankang@mail.tsinghua.edu.cn; wwu@mail.tsinghua.edu.cn; fengcj301@163.com; qczhang@tsinghua.edu.cn

https://doi.org/10.1038/s41592-026-03237-0

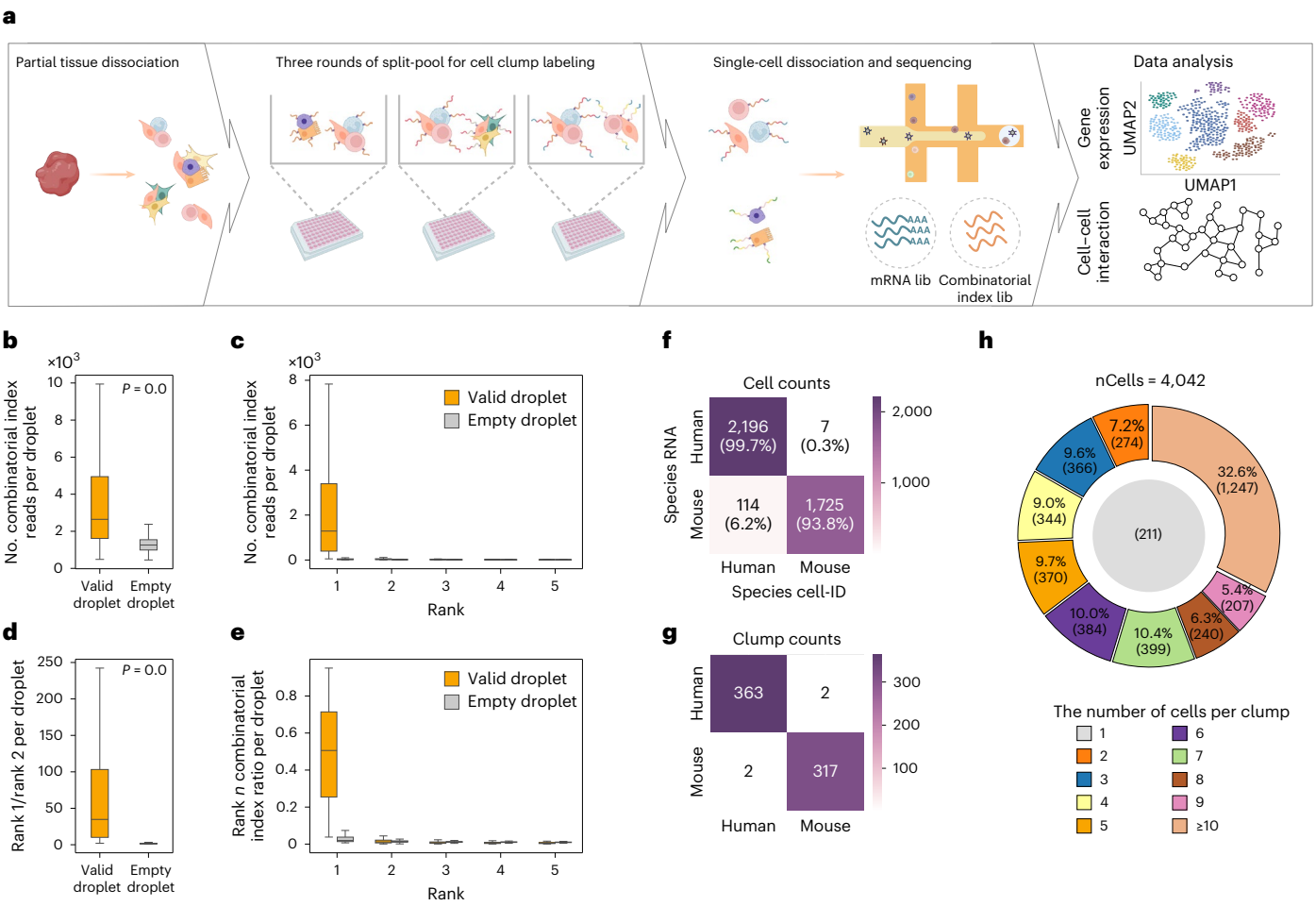


Fig. 1 | CCI-seq achieves high-quality cell labeling and accurate identification of cell–cell interactions. a, Schematic of CCI-seq workflow. **b**–**e**, Boxplot showing the following metrics within each droplet after filtering: the total number of combinatorial indices (**b**), the number of top five indices (**c**), the ratio of rank 1 to rank 2 indices (**d**) and the ratio of top five indices (**e**). Valid droplets are droplets with qualified RNAs, $n = 4{,}042$. Empty droplets indicate droplets without qualified RNAs, $n = 85{,}449$. Boxes show interquartile range. Midline shows the median. Whiskers show 1.5× interquartile range. Significance estimated using a one-sided Student's t-test. **f**, Confusion matrix comparing cell counts between RNA-based annotations (y axis, referred as the ground truth) and species-specific Cell-ID annotations (x axis). **g**, Confusion matrix showing the composition of cell types of 682 cell clumps, two of which contained both human and mouse cells. **h**, Pie plot showing the relative abundance of cells assigned to cell clumps of the indicated sizes ($n = 4{,}042$ cells with high-quality Cell-IDs). Panel **a** created in BioRender; Tang, L. https://biorender.com/y6br4qw (2026).

Over the past decade, technologies that retain information about interacting cells and/or spatial positions during single-cell analysis have been developed as alternatives to ST. For instance, proximity labeling methods[16–23] label neighboring cells of 'bait' cells before scRNA-seq; however, these methods require a pre-selection step that limits the analysis to a single bait cell type. Cell clump deconvolution methods[24–28] use partial dissociation of cells (which retains the tissue microenvironment) followed by scRNA-seq and computational deconvolution to infer constituent cells within a clump; however, these methods are subject to deconvolution inaccuracies. Spatial nuclei indexing technologies[29–31] blur the distinction between ST and the aforementioned methods by tagging single nuclei within a tissue section using spatial oligonucleotide barcoding. While the methods show great promise, they also face the challenge of low nuclei recovery rates, which has led to sparse tissue sampling and a failure to capture data for the large majority of cell–cell interactions in a tissue[31]. Further, these methods are restricted to RNA molecules in the nucleus, yet a large proportion of mRNAs are in the cytoplasm[32].

Here, we introduce CCI-seq, a method that captures large-scale cell–cell interaction networks by integrating partial tissue dissociation and combinatorial cell clump indexing with single-cell sequencing. A key advantage of CCI-seq is its ability to resolve the molecular states associated with these interactions, moving beyond simple network mapping. To demonstrate its utility, we applied CCI-seq across healthy, precancerous and cancerous tissues from both mice and humans. In the mouse kidney, CCI-seq recapitulated known interactions within the collecting duct and glomerulus and revealed that epithelial cells engaging distinct partners were associated with specialized ion transport programs. In the intestine, the method revealed the spatial arrangement of known enterocyte and goblet cell subtypes along the crypt–villus axis. Comparing healthy and precancerous, adenomatous polyposis coli knockout (*Apc*-KO) mouse intestines, our CCI-seq analyses discovered disrupted cellular organization and abnormal interactions, where villus-tip enterocytes engaging T cells were linked to heightened antigen presentation. Finally, in human colorectal cancer (CRC) samples, CCI-seq identified subtype-specific epithelial-immune networks, such as tumor-monocyte interactions characterized by a nuclear factor (NF)-κB inflammatory signature. Together, these findings establish CCI-seq as a scalable approach to simultaneously map cell–cell interaction networks and characterize the molecular states coupled to them, providing deep functional insights into diverse biological systems and clinical applications.

Results

Accurate identification of cell–cell interactions by CCI-seq

To perform CCI-seq, samples are partially dissociated into small cell clumps (<20 cells) to preserve native cell–cell contacts. Microscopic inspection was implemented as a key quality control step to verify cell viability and clump size. Following this, a barcoding cholesterol-modified oligonucleotide (CMO) is applied to anchor the cell membranes of all cells within the clumps[33], followed by three rounds of split-pool barcode ligation to generate a unique combinatorial index to label each cell clump[34]. Finally, the clumps are fully dissociated into single cells for droplet-based sequencing, which captures both the transcriptome and the clump-derived index for each cell (Fig. 1a and Methods). The most abundant index recovered from each cell is defined as its 'Cell-ID'; by design, all cells originating from the same clump share the same Cell-ID and are thus identified as interacting with each other.

We assessed CCI-seq using three-dimensional spheroids of human HEK293T and mouse NIH/3T3 cells (Extended Data Fig. 1a,b and Methods). We first confirmed that the labeling process itself was efficient, with fluorescein amidites-conjugated CMO barcodes evenly labeling both interior and exterior cells of a spheroid (Extended Data Fig. 1c). We then labeled human and mouse spheroids with species-specific barcodes, mixed them, performed combinatorial indexing, and conducted complete dissociation and scRNA-seq. We obtained 3,826 human cells and 3,643 mouse cells of high-quality transcriptomic data, a median of over 11,000 RNA unique molecular identifiers (UMIs) and 3,400 genes per cell, comparable to standard scRNA-seq experiments[35] and indicating that the CCI-seq protocol does not compromise scRNA-seq data acquisition (Extended Data Fig. 1d–f).

Next, we assessed the specificity of the indexing and the accuracy of cell identification. After data filtering (Methods), we obtained 2,203 human cells and 1,839 mouse cells with a high-quality Cell-ID, being on average 86.3-fold more abundant than the next most common index (Fig. 1b–e and Extended Data Fig. 1g–l). Among these, 99.7% of human cells and 93.8% of mouse cells had the correct species-specific Cell-ID label (Fig. 1f), reflecting a high labeling specificity of cells.

Finally, we confirmed that CCI-seq preserves the integrity of the original cell clumps. Only about 0.3% (2 of 682) clumps contained both human and mouse cells (Fig. 1g), reflecting low cross-contamination among cell clumps. The vast majority of cells (~95% or 3,831/4,042) were assigned to a clump of two or more cells, providing information for cell–cell interactions (Fig. 1h). Together, these findings demonstrate that CCI-seq achieves specific cell labeling and accurate identification of cell–cell interactions.

CCI-seq resolves the mouse kidney cell interactome with deep transcriptome coverage

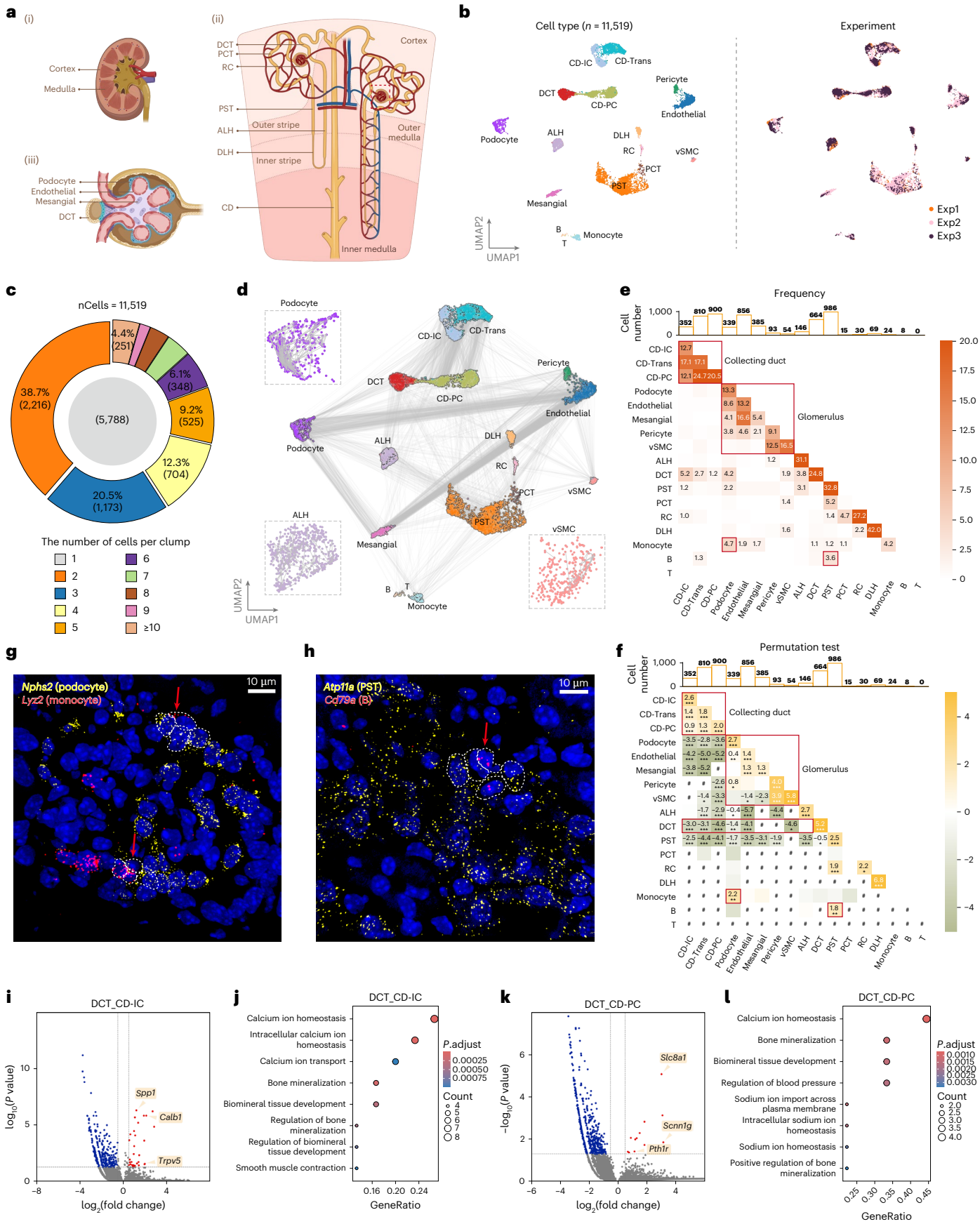
We applied CCI-seq to the mouse kidney, a complex organ with a well-characterized anatomy[36] (Fig. 2a and Methods). After stringent quality control and filtering, we retained 11,519 high-quality single-cell profiles from three biological replicates, detecting a median of 2,822 genes per cell (Fig. 2b and Extended Data Fig. 2a–i). This represents a substantial increase in transcriptomic depth over established spatial technologies like Slide-seq-V2 (ref. 37) (436 genes, 6.5-fold) and seqFISH[38] (115 genes, 24.5-fold) (Extended Data Fig. 2j). From this rich dataset, we performed unsupervised clustering and identified 17 distinct cell types encompassing all major renal lineages (Fig. 2b and Extended Data Fig. 2k).

Nearly half of the cells were captured in small clumps of 2–5 cells, indicating direct physical contacts and providing a robust foundation for interaction analysis (Fig. 2c,d and Supplementary Fig. 1). We then calculated a normalized 'interaction frequency' between cell types and assessed its statistical significance to generate a comprehensive interaction map (Fig. 2e,f and Methods). This map proved highly robust and accurate, as evidenced by three lines of observations: first, it was highly reproducible across biological replicates (cosine similarity = 0.68–0.84; Extended Data Fig. 3a); second, it recapitulated known anatomical structures, such as intra-glomerular (podocytes, endothelial, and mesangial cells)[39] and collecting duct (CD-ICs, CD-PCs and CD-Trans)[40,41] associations (Fig. 2e,f and Extended Data Fig. 3b,c); and third, it showed a high degree of concordance with a high-resolution seqFISH dataset[38] (cosine similarity = 0.86; Extended Data Fig. 3d,e), demonstrating that CCI-seq captures cell–cell interactions without apparent bias.

Beyond confirming known structures, the map revealed previously unrecognized interactions in the kidney. Notably, we identified enriched interactions between immune and nonimmune cells,

Fig. 2 | CCI-seq reveals cell–cell interaction networks for the mouse kidney. **a**, Illustration of the main cell types and their spatial arrangement in the mouse kidney[36]. (i), Gross regional organization of the adult kidney. (ii), Anatomy of the nephron. PCT, proximal convoluted tubule; RC, renal corpuscle; ALH, ascending limb of loop of Henle; DLH, descending limb of loop of Henle; CD, collecting duct. (iii), High-resolution overview of the renal corpuscle and associated structures (region boxed in (ii)). **b**, Uniform Manifold Approximation and Projection (UMAP) visualization of the CCI-seq's scRNA data of the mouse kidney. Cells (dots) are colored by cell type (left) and experiment (right). CD-PC, collecting duct principal cell; CD-IC, collecting duct intercalated cell; CD-Trans, collecting duct transitional cell; vSMC, vascular smooth muscle cell. Three collecting duct cell types: CD-ICs, CD-Trans and CD-PCs; four glomerular cell types: podocytes, endothelial cells, pericytes and mesangial cells; five renal tubule cell types: PSTs, PCTs, ALHs, DLHs and DCTs; three immune cell types: T cells, B cells and monocytes. **c**, Pie plot showing the relative abundance of cells assigned to cell clumps of the indicated sizes (n = 11,519 cells with high-quality Cell-IDs). **d**, The cell–cell interaction network of the mouse kidney. Cells (nodes) are colored by cell type, and gray edges represent detected cell–cell interactions. Three dashed-line boxes highlight the interactions within three specific cell types. **e**, Heatmap plot showing the interaction frequencies between cell types, with a barplot (top) showing the number of cells for each cell type. Red boxes highlight interactions discussed in the manuscript. Interaction frequencies below 1.0 are not shown. **f**, Heatmap plot showing the enrichment and depletion of interaction frequencies between cell types, with a barplot (top) showing the number of cells for each cell type. Statistical analysis was conducted using a one-sided permutation test (n = 1,000). The values have been $\log_2$-transformed. *P < 0.05, **P < 0.01, ***P < 0.001. Red boxes highlight interactions discussed in the manuscript. #Indicates that interacted cell type pair was detected in fewer than three cell clumps. **g**, AsmFISH image of the glomerulus with 4′,6-diamidino-2-phenylindole (DAPI) counterstaining revealing cell–cell interactions between podocytes and monocytes. *Nphs2* (yellow), podocytes; *Lyz2* (red), monocytes; dashed lines indicate cell borders, and red arrows indicate interactions between podocytes and monocytes. **h**, AsmFISH image of the kidney with DAPI counterstaining revealing cell–cell interactions between PSTs and B cells. *Atp11a* (yellow), PSTs; *Cd79a* (red), B cells. The dashed lines indicate cell borders, and red arrows indicate interactions between PSTs and B cells. **i**, Volcano plot showing the DEGs between CD-IC-interacting and CD-IC-non-interacting DCT cells. Red and blue dots represent genes up- and downregulated in CD-IC-interacting DCT cells, respectively. The P values were determined using a two-sided Student's t-test with Benjamini–Hochberg correction. The dashed lines indicate thresholds: adjusted P value < 0.05, $|\log_2(\text{fold change})|>0.5$. **j**, Dotplot showing the top eight significantly enriched GO terms for DEGs of CD-IC-interacting DCT cells. The P values were determined using a hypergeometric test with Benjamini–Hochberg correction. Dot size represents gene counts, and dot color represents adjusted P value. **k**, Volcano plot showing the DEGs between CD-PC-interacting and CD-PC-non-interacting DCT cells. Red and blue dots represent genes up- and downregulated in CD-PC-interacting DCT cells. The P values were determined using a two-sided Student's t-test with Benjamini–Hochberg correction. The dashed lines indicate thresholds: adjusted P value < 0.05, $|\log_2(\text{fold change})|>0.5$. **l**, Dotplot showing the top eight significantly enriched GO terms for DEGs of CD-PC-interacting DCT cells. The P values were determined using a hypergeometric test with Benjamini–Hochberg correction. Dot size represents gene counts, and dot color represents adjusted P value. Panel **a** created in BioRender; Tang, L. https://biorender.com/oswzpcz (2026).

https://doi.org/10.1038/s41592-026-03237-0



including monocytes with podocytes and between B cells with proximal straight tubule (PST) cells (Fig. 2e,f). We validated the direct physical proximity of these new pairs in situ using immunofluorescence and asmFISH[42] (Fig. 2g,h, Extended Data Fig. 3f and Supplementary Table 1). As monocytes and B cells have been implicated in kidney function and immune diseases[43–45], these interactions may contribute to local immune surveillance and epithelial regulation.

The unique power of CCI-seq lies in linking these physical interactions to molecular function. By performing differential gene expression analysis, we uncovered notable interaction-dependent cellular functional specialization (Methods). This was exemplified by distal convoluted tubule (DCT) cells, which adopted divergent transcriptional programs depending on their collecting duct partner (Fig. 2i–l). Specifically, DCT cells interacting with CD-IC cells upregulated calcium transport genes such as *Trpv5* and *Calb1*, indicating that they adapt their calcium machinery in response to the local CD-IC microenvironment[46,47]. In contrast, DCT cells interacting with CD-PC cells upregulated blood pressure regulation genes such as *Scnn1g* and *Slc8a1*, highlighting how this interaction influences the DCT's role in sodium and water balance[48,49]. This principle of interaction-driven programming was widespread. For instance, podocytes interacting with DCT cells showed markedly altered transcriptional profiles enriched in oxidative phosphorylation (Extended Data Fig. 4a–c).

Beyond transcriptional programming, CCI-seq uniquely grounds predicted signaling events in verified physical proximity. By performing LIANA-based ligand–receptor analysis[50] on cells within the same clumps (Methods), we uncovered highly specific local communication networks. For example, podocytes interacting with endothelial cells uniquely engaged the Timp3–Kdr and Vegfa–Kdr axes (Extended Data Fig. 4d), yielding pivotal mechanistic insights into glomerular microvascular homeostasis[51]. Furthermore, CD-Trans cells exhibited elevated Spp1–Itgav/Itgb1 ligand–receptor pairs when contacting CD-ICs, while showed stronger L1cam–Cd9 interactions when engaging CD-PCs (Extended Data Fig. 4e,f and Methods), highlighting context-dependent modulation of epithelial adhesion. Similarly, monocytes interacting with podocytes exhibited enhanced Thbs1–Sdc4 and H2-K1–Aplp2 signaling (Extended Data Fig. 4g and Methods), indicating direct involvement in immune regulation and cell adhesion[52–54]. Overall, CCI-seq successfully captured known and new cell–cell interactions within the mouse kidney, revealing their gene-level impacts on cellular function and crosstalk.

CCI-seq reconstructs the crypt–villus architecture of the mouse small intestine

We next assessed the mouse small intestine, which features a well-characterized crypt–villus structure composed of intestinal epithelial cells (IECs)[55] (Fig. 3a and Methods). Briefly, the crypts house *Lgr5*$^+$ intestinal stem cells (ISCs), the transit amplifying (TA) progeny and supportive Paneth cells. The villi are predominantly lined by absorptive enterocytes, whereas secretory cell types (including goblet, enteroendocrine (EEC) and tuft cells) are distributed along the crypt–villus axis[55].

After data quality control and filtering, we obtained 6,412 high-quality single-cell profiles from three biological replicates, with an average of 4,127 genes detected per cell (Fig. 3b and Supplementary Fig. 2a–h). This substantially surpassed 10x Visium HD (438 genes, 9.4-fold) in gene detection for the same tissue (Supplementary Fig. 2i). We performed unsupervised clustering and identified eight distinct cell types, including *Lgr5*$^+$ ISCs, TA cells, enterocytes, B cells and various secretory cells, with proportions highly consistent with published scRNA-seq data[56] (cosine similarity = 0.98; Fig. 3b and Supplementary Fig. 3a,b). To resolve spatial positions along the villus, we calculated a 'landmark score' for enterocytes and goblet cells based on established spatially-varying 'landmark genes'[28,57], classifying them into spatially distinct subtypes (villus-bottom, villus-middle, and villus-tip enterocytes, and crypt and villus goblets) (Supplementary Fig. 3c–h, Supplementary Table 2 and Methods).

Among the 6,412 cells, 56.6% (3,631 cells) were captured in clumps, providing direct physical interaction data (Fig. 3c and Supplementary Fig. 4). The resulting interaction map was highly reproducible across biological replicates (cosine similarity scores = 0.68–0.78) and faithfully recapitulated the known tissue architecture (Fig. 3d,e and Supplementary Fig. 5). We observed preferential interactions between spatially adjacent enterocyte subtypes, correctly ordered along the villus-bottom-to-tip axis. Likewise, villus goblet cells preferentially interacted with villus enterocytes, whereas crypt goblet cells preferentially interacted with crypt-resident TA cells. And we also detected canonical niche interactions between *Lgr5*$^+$ ISCs and Paneth cells[58].

We further leveraged cell clump data to explore higher-order tissue architecture. Clustering clumps by their cellular compositions revealed biologically meaningful groupings (Methods). A large cluster was enriched for *Lgr5*$^+$ ISCs, Paneth cells, crypt goblet cells and TA cells, supporting that these clumps were derived from the intestinal crypt (Extended Data Fig. 5a,b). Building on this, we stratified clumps along a hierarchical assembly trajectory using clump size as a proxy for structural complexity, which effectively reconstructed the crypt–villus axis (Extended Data Fig. 5c). Together, these analyses demonstrate that clump composition profiles capture biologically meaningful, multicellular structural motifs reflective of higher-order tissue organization.

CCI-seq uncovers a spatial and differentiation continuum in TA cells

Originating from *Lgr5*$^+$ ISCs, TA cells are crucial for ensuring the uniquely continuous high turnover and regeneration of the intestinal epithelium[58–60]. It is well-established that TA cells are located throughout the crypt, with extensions reaching into the villus[55] (Fig. 3a); however, the landmark scores from previous studies[28,57] do not contain TA cells. Whether TA cells have specific spatial subtypes with different functions

Fig. 3 | CCI-seq reveals cell–cell interaction networks and TA spatial subtypes for the mouse intestine. a, Illustration of the main cell types and their spatial arrangement in the mouse intestine[55]. **b**, The cell–cell interaction network of the mouse intestine. Cells (nodes) are colored by cell type (left) and experiment (right), and gray edges represent detected cell–cell interactions. The dashed-line box highlights the interactions within *Lgr5*$^+$ ISCs. **c**, Pie plot showing the relative abundance of cells assigned to cell clumps of the indicated sizes (n = 6,412 cells with high-quality Cell-IDs). **d**, Heatmap plot showing the interaction frequencies between cell types, with a barplot (top) showing the number of cells for each cell type. Red boxes highlight interactions discussed in the manuscript. Interaction frequencies below 1.0 are not shown. **e**, Heatmap plot showing the enrichment and depletion of interaction frequencies between cell types, with a barplot (top) showing the number of cells for each cell type. Statistical analysis was conducted using a one-sided permutation test (n = 1,000). The values have been $\log_2$-transformed. *P < 0.05, **P < 0.01, ***P < 0.001. Red boxes highlight interactions discussed in the manuscript. Hash symbol indicates that interacted cell type pair was detected in fewer than three cell clumps. **f,g**, UMAP visualization of the CCI-seq's scRNA data of TA cells interacting with other cell types. Cells are colored by interaction partner (**f**) and spatial subtype (**g**). **h**, Left: asmFISH image of the intestinal crypt with DAPI counterstaining. *Lgr5* (green), *Lgr5*$^+$ ISCs; *Mki67* (yellow), proliferation marker; *Sorbs2* (magenta), crypt-bottom TA; *Rbp7* (cyan), crypt-top TA. The dashed lines indicate the crypt and cell borders. Right: line plot showing the fluorescence signal strength from the crypt bottom to the crypt top region. **i,j**, UMAP visualization of the CCI-seq's scRNA data of TA cells interacting with other cell types. Cells are colored by the landmark score (**i**) and pseudotime (**j**). In **i**, black arrow indicates the inferred spatial distribution of TA cells from crypt bottom to crypt top. **k**, Correlation analysis of pseudotime versus the landmark score of TA cells (n = 477). Statistical analysis was conducted using a two-sided Wald Test with t-distribution. The dashed line represents the fitted linear regression curve. Panel **a** created in BioRender; Tang, L. https://biorender.com/ukxkwa6 (2026).

https://doi.org/10.1038/s41592-026-03237-0

is therefore not known. Our CCI-seq interaction map revealed that TA cells frequently interacted with both villus-bottom enterocytes and crypt-resident cells (for example, *Lgr5*$^+$ ISCs, Paneth cells and crypt goblet cells) (Fig. 3d and Supplementary Fig. 5b,c). We therefore assigned TA cells to two spatial subtypes: crypt-top TA and crypt-bottom TA, based on distinct interaction partners (Fig. 3f,g and Methods).

These two subtypes displayed distinct transcriptional profiles that mirrored their location. Crypt-bottom TA expressed high levels of *Lgr5*$^+$ ISCs marker genes, including *Sorbs2* and *Olfm4*, while crypt-top TA showed elevated expression of genes linked to mature enterocytes, such as *Rbp7* and *Smim24* (Extended Data Fig. 6a). We validated this spatial segregation in situ using asmFISH, confirming that TA cells expressing *Sorbs2* were confined to the crypt bottom, whereas TA cells expressing *Rbp7* were restricted to the crypt top (Fig. 3h).

To quantify this spatial axis, we defined landmark gene sets from the top differentially expressed genes (DEGs) of each subtype and calculated a spatial score for each TA cell, following the approach used to discern enterocyte spatial subtypes[57] (Supplementary Table 3 and

Methods). This spatial score correlated well with the differentiation pseudotime trajectory inferred from single-cell analysis (Pearson's correlation coefficient = 0.74; Fig. 3i–k and Methods), confirming that a cell's interaction-defined location reflects its developmental state. Finally, Gene Ontology (GO) analysis further illuminated their distinct functions: crypt-bottom TA cells were enriched in Wnt-signaling-related terms, consistent with their proximity to Wnt ligands niche at the crypt base[59], whereas crypt-top TA cells exhibited enrichment for metabolic and energy production pathways, reflecting their similarity to mature enterocytes[61] (Extended Data Fig. 6b,c). Collectively, these data show that CCI-seq resolves TA cells into spatial subtypes that reflect a differentiation continuum along the crypt–villus axis, governed by distinct microenvironmental signals.

CCI-seq reveals pathological rewiring of the interactome in an *Apc*-KO adenoma model

To investigate how cell interactions are altered during tumorigenesis, we applied CCI-seq to analyze small intestine samples from a mouse model with conditional *Apc* knockout (*Apc*-KO), which drives adenoma formation[62–64] (Methods). We obtained 18,986 high-quality cells after quality control, with 52.7% captured in interacting clumps (Fig. 4a,b and Supplementary Figs. 6–8). We calculated the interaction frequencies among distinct cell types in the *Apc*-KO samples (Fig. 4c and Supplementary Fig. 9a).

We immediately noted dramatic changes in cell–cell interactions for the *Apc*-KO versus wild-type (WT) data (Figs. 3d,e and 4c,d, Supplementary Fig. 9b–d and Methods). Notably, in stark contrast to WT intestines, crypt-resident cells (*Lgr5*$^+$ ISCs, Paneth cells and TA cells) were no longer confined to the crypt and showed aberrant interaction with villus-tip and -middle enterocytes (Fig. 4c,d versus Fig. 3d,e). We confirmed this ectopic localization of Paneth cells using asmFISH, observing that they scattered throughout the villus and directly contact villus-tip enterocytes in the *Apc*-KO samples (Fig. 4e and Supplementary Fig. 9e). This pathological rewiring of the cellular landscape highlights a profound loss of spatial organization at the earliest stages of adenoma development[65].

Within this disorganized environment, we identified new, disease-associated cell states. A population of cells expressing canonical TA markers (*Stmn1* and *Tubb5*) seemed transcriptionally distinct from WT TA cells, which we termed 'TA-like' cells (Extended Data Fig. 7a and Supplementary Fig. 7b). Further analysis of these cells uncovered two unique subtypes characterized by strong expression of either the Hippo–YAP/TAZ gene *Ly6a*[66,67] or the Wnt antagonist gene *Notum*[68,69] (Fig. 4f and Extended Data Fig. 7b–d). We designated these as TA-like_*Ly6a*$^+$ and TA-like_*Notum*$^+$ cells. Targeted genotyping confirmed that the TA-like_*Ly6a*$^+$ cells harbor the *Apc* exon 15 deletion, establishing them as bona fide adenoma cells (Extended Data Fig. 7e,f). Functionally, TA-like_*Ly6a*$^+$ cells specifically highly expressed repair-related genes, whereas TA-like_*Notum*$^+$ cells specifically highly expressed Wnt-signaling-related genes, suggesting they have distinct roles in adenoma progression[67,68,70] (Extended Data Fig. 7g,h and Supplementary Table 4).

Of note, CCI-seq revealed that these new precancerous cell states establish an aberrant interaction network. Compared to WT TA cells, both TA-like_*Ly6a*$^+$ and TA-like_*Notum*$^+$ subtypes lost their interaction with *Lgr5*$^+$ ISCs and instead gained frequent interactions with villus-tip enterocytes and, notably, with each other (Fig. 4g and Extended Data Fig. 8a–c). This suggests that they form a new, ectopic signaling hub in the villus. We validated this finding in situ, confirming that TA-like_*Ly6a*$^+$ and TA-like_*Notum*$^+$ tended to colocalize within the villus region of *Apc*-KO samples (Fig. 4h,i and Extended Data Fig. 8d). Concurrently, the tumor microenvironment showed increased immune infiltration, with CCI-seq capturing a broad spectrum of immune-epithelial and immune-immune interactions, such as direct B–T cell contacts, which were more extensive than those detected by other methods[16] (Extended Data Fig. 8e–h). Villus-tip enterocytes interacting with T cells displayed enhanced antigen-presentation signatures, featuring stronger MHC-I–related ligand–receptor interactions (for example, *H2-K1–Cd3g*; Extended Data Fig. 8i), consistent with active immune surveillance[71–73]. Together, these results demonstrate that CCI-seq can deconstruct the complex remodeling of a precancerous niche, identifying new pathological cell states and the aberrant interaction networks they form to drive disease.

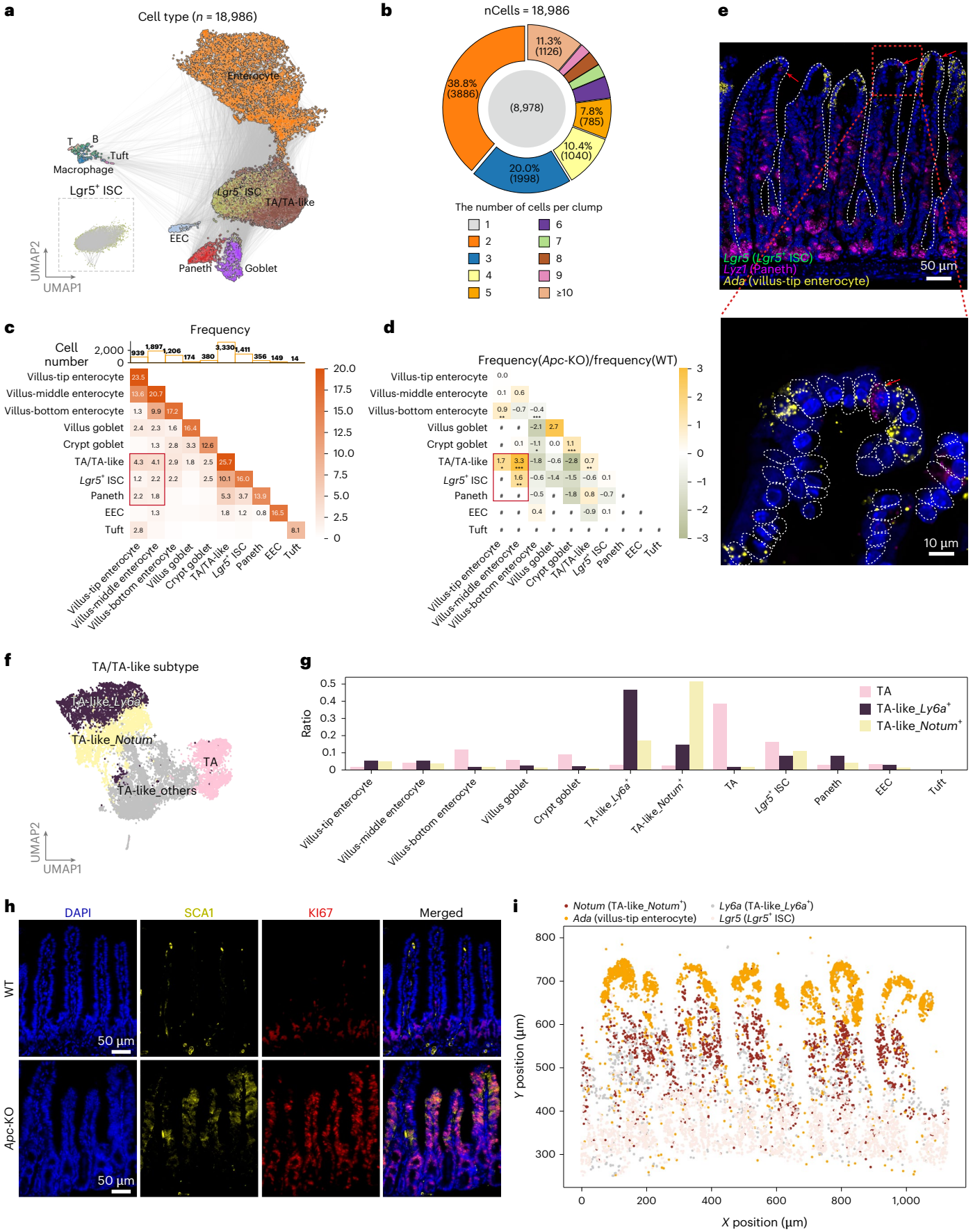
CCI-seq dissects subtype-specific cell–cell interactions in human CRC samples

Human CRC is one of the most prevalent cancers worldwide[74]. CRC tumors can be characterized by integrating epithelial intrinsic-consensus molecular subtypes (iCMS2 and iCMS3)[75]. This classification provides information about tumor biology, improves prognostic accuracy, and can be used to tailor therapeutic strategies. To map the cellular interactomes that define these subtypes, we applied CCI-seq to five CRC tumor samples (T1–T5) from five patients, with three adjacent normal samples (N2, N4 and N5) from three of the patients (nos. 2, 4 and 5), and obtained 37,897 high-quality cells (Fig. 5a, Extended Data Fig. 9a–d and Methods). Among these, 71.9% (27,248 of 37,897) of the cells were captured within clumps, providing cell interactions information about the tumor microenvironment (TME) (Fig. 5b). We transferred cell type labels from an annotated reference CRC dataset[76] to our scRNA-seq data obtained from CCI-seq using SCALEX (Methods), and identified 16 distinct cell types (Fig. 5a and Extended Data Fig. 9e,f).

We focused primarily on the 34,193 epithelial cells that underpin CRC molecular classification. Consistent with previous work[75], we identified two subtypes of tumor cells, namely iCMS2 and iCMS3 with distinct gene expression profiles and copy number profiles (Fig. 5c, Extended Data Fig. 9g–i and Methods). Specifically, iCMS2 cells, predominantly from T2, T4 and T5, were characterized by high expression of *MYC* pathway activation and frequent gains of chromosomal arms including 7pq, 8q, 13q and 20pq, whereas iCMS3 cells, mainly

Fig. 4 | CCI-seq reveals precancerous TA-like cell subtypes in the villus for the *Apc*-KO mouse intestine. a, The cell–cell interaction network of the *Apc*-KO mouse intestine. Cells (nodes) are colored by cell type, and gray edges represent detected cell–cell interactions. The dashed-line box highlights the interactions within *Lgr5*$^+$ ISCs. **b**, Pie plot showing the relative abundance of cells assigned to cell clumps of the indicated sizes (n = 18,986 cells with high-quality Cell-IDs). **c**, Heatmap plot showing the interaction frequencies between cell types, with a barplot (top) showing the number of cells for each cell type. Red box highlights interactions discussed in the manuscript. Interaction frequencies below 1.0 are not shown. **d**, Heatmap plot showing the enrichment and depletion of interaction frequencies for each indicated pair of cell types between the WT and *Apc*-KO samples. Statistical analysis was conducted using a one-sided permutation test (n = 1,000). The values have been $\log_2$-transformed. *P < 0.05, **P < 0.01, ***P < 0.001. Red box highlights interactions discussed in the manuscript. #Indicates that interacted cell type pair was detected in fewer than three cell clumps in WT samples. **e**, AsmFISH image of the intestinal crypt–villus axis with DAPI counterstaining. *Lgr5* (green), *Lgr5*$^+$ ISCs; *Lyz1* (magenta), Paneth cells; *Ada* (yellow), villus-tip enterocytes. The dashed lines indicate the crypt–villus axis borders (top) and cell borders (bottom), and red arrows indicate Paneth cells located at the villus-tip region. **f**, UMAP visualization of the CCI-seq's scRNA data of the total TA/TA-like cells. Cells are colored by TA/TA-like subtype. **g**, Barplot showing the relative interaction frequencies between each TA/TA-like subtype and other cell types. **h**, Representative examples of the mouse intestinal crypt–villus axis stained by immunofluorescence showing the spatial distribution of the TA-like_*Ly6a*$^+$. DAPI (blue), SCA1 (yellow), KI67 (red). Cells double-positive for SCA1 (encoded by *Ly6a*) and KI67 represent the TA-like_*Ly6a*$^+$ population. **i**, Dotplot showing the spatial locations of four selected genes, representing four cell types, along the mouse intestinal crypt–villus axis. Corresponding asmFISH image presented in Extended Data Fig. 8d.

https://doi.org/10.1038/s41592-026-03237-0



https://doi.org/10.1038/s41592-026-03237-0

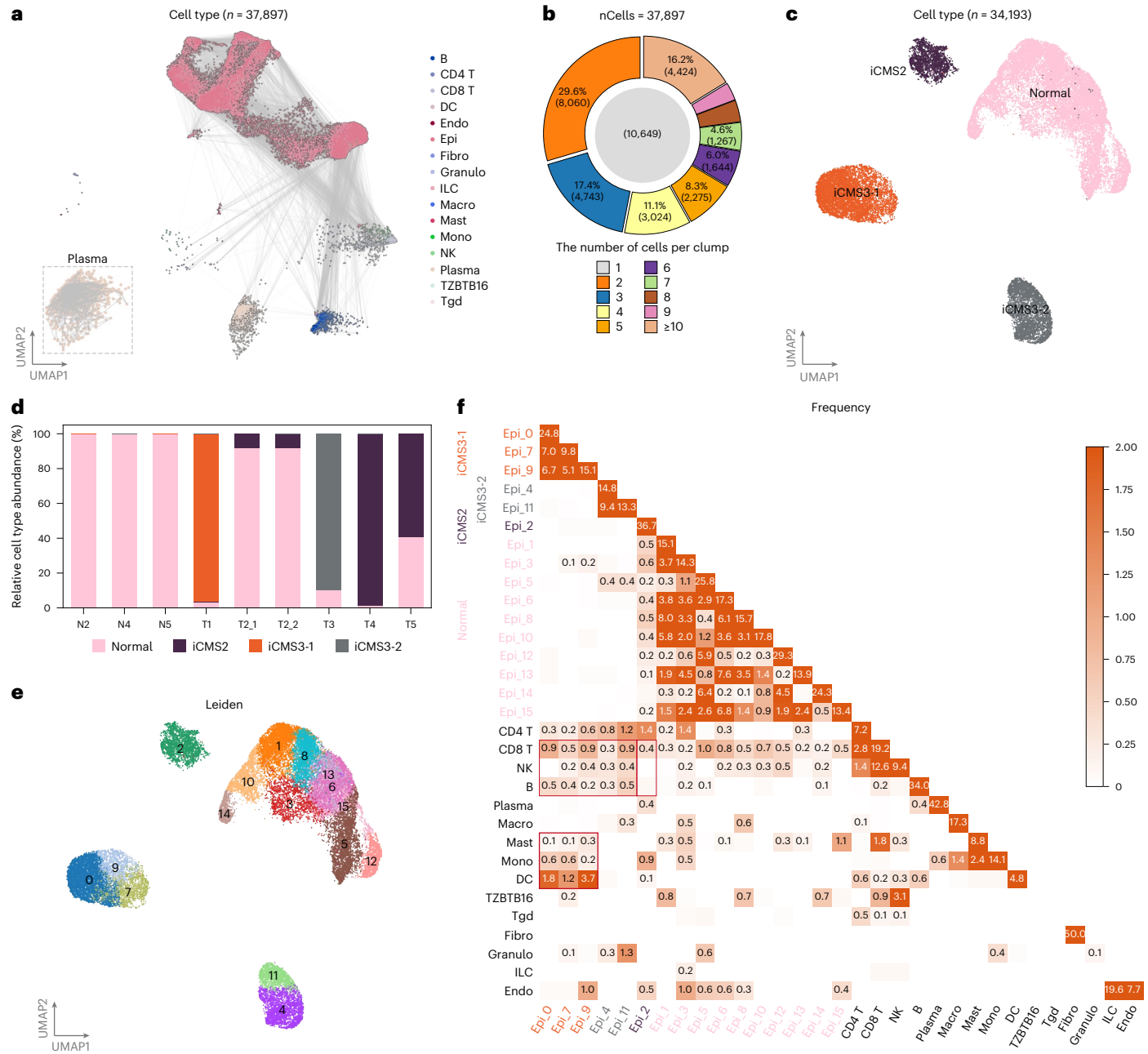


Fig. 5 | CCI-seq detects consensus molecular subtypes and subtype-specific interactions in human CRC clinical samples. a, The cell–cell interaction network. Cells (nodes) are colored by cell type, and gray edges represent detected cell–cell interactions. The dashed-line box highlights the interactions within plasma. **b**, Pie plot showing the relative abundance of cells assigned to cell clumps of the indicated sizes (n = 37,897 cells with high-quality Cell-IDs). **c**, UMAP visualization of the CCI-seq's scRNA data of epithelial cells in human CRC. Cells are colored by epithelial subtype. **d**, Stacked barplot showing the relative abundance of epithelial subtypes across different donors. **e**, UMAP visualization of the CCI-seq's scRNA data of epithelial cells in human CRC. Colors indicate the Leiden clusters, which were derived from unsupervised clustering of single-cell gene expression data. **f**, Heatmap plot showing the interaction frequencies between cell types. Red boxes highlight interactions discussed in the manuscript. DC, dendritic cells; Endo, endothelial cells; Epi, epithelial cells; Fibro, fibroblasts; Granulo, granulocytes; ILC, innate lymphoid cells; Macro, macrophages; Mono, monocytes; NK, natural killer cells; Tgd, γδ T cells.

from T1 and T3 (designated as iCMS3-1 and iCMS3-2), showed high expression of the iCMS3 signatures including *FUT8* and were either diploid or showed infrequent and inconsistent copy number alterations (Fig. 5d, Extended Data Fig. 9g–i and Supplementary Table 5).

With these subtypes defined, CCI-seq revealed a stark dichotomy in their immune interactions. The iCMS2 tumors were largely 'immune deserts', showing minimal interaction with immune cells. In stark contrast, iCMS3 tumors exhibited extensive immune infiltration, with tumor cells frequently engaging $CD8^+$ T cells, B cells and natural killer cells (Fig. 5e,f and Extended Data Fig. 10a). This finding directly visualizes the cellular basis for the 'immune-hot' signature previously reported for the iCMS3 subtype[75]. Further, CCI-seq uncovered heterogeneity within the iCMS3 subtype, detecting interactions between myeloid cells (including mast cells, monocytes and dendritic cells) and iCMS3-1 tumor cells that were absent in iCMS3-2, underscoring distinct immune milieus even within the same major subtype.

To understand the functional consequences of these interactions, we examined how direct contact with monocytes reshapes iCMS3-1 tumor cells. Monocyte-interacting iCMS3-1 tumor cells specifically upregulated NF-κB target genes, including the proinflammatory

cytokines *CXCL8* and *IL1B* (Extended Data Fig. 10b–d). This suggests that monocyte-derived signals activate NF-κB signaling in tumor cells, hijacking inflammatory machinery to promote invasion and immune evasion[77]. This pro-tumorigenic reprogramming was further linked to the induction of stem-like traits, as ligand–receptor analysis revealed stronger CD44-related signaling in these monocyte-interacting tumor cells—a pathway critical for cancer stem cell maintenance[78,79] (Extended Data Fig. 10e).

Overall, these findings demonstrate that CCI-seq effectively captures cell interactomes in human CRC, linking clinically relevant tumor subtypes to their distinct immune environments, and uncovers how tumor–immune interactions can reprogram cancer cells to foster a pro-tumorigenic, stem-like state, offering new insights into the molecular logic governing tumor–immune ecosystems.

Discussion

We have developed CCI-seq, a method to unbiasedly map large-scale cell–cell interaction networks for complex tissues by combining partial tissue dissociation, combinatorial cell clump indexing and single-cell sequencing. As a single-cell-based method, CCI-seq naturally overcomes inherent limitations of many ST technologies, including restricted gene detection and difficulties in defining cell boundaries, while remaining comparable in time and costs (Supplementary Table 6). CCI-seq also addresses several critical limitations of other single-cell-based interaction detection methods. First, unlike proximity labeling techniques, CCI-seq preserves all interacting partners to generate unbiased interaction networks. Second, in contrast to spatial nuclei indexing technologies, CCI-seq has a high cell recovery rate and provides high-quality whole-cell transcriptome data. Third, CCI-seq circumvents the need for probabilistic computational deconvolution and previous surface markers-based cell sorting required by methods like paired-cell sequencing[25], establishing it as a broad discovery platform for the mapping of large-scale cellular association landscapes. Finally, CCI-seq surpasses the low-resolution analysis by methods such as fragment-seq[80], which is restricted to analyzing large tissue fragments (Ø 200–450 µm) and lacks the spatial resolution necessary for interaction studies.

Beyond constructing interaction maps, CCI-seq enables each cell's molecular state to be examined in relation to the identities of its interaction partners. This partner-aware analysis provides opportunities to investigate how cellular neighborhoods are associated with functional specialization and tissue organization, and how these relationships are remodeled during disease. As demonstrated here, the deep transcriptomic coverage of CCI-seq supports differential gene expression, pathway enrichment and ligand–receptor analyses that associate local cellular interactions with inflammatory, metabolic and stem-like programs, generating mechanistic hypotheses for further investigation in cancer and other heterogeneous tissues.

CCI-seq showed robust performance across biological replicates and tissue types, supported by a standardized protocol incorporating gentle dissociation and multistage quality control (Methods). Concordance with published scRNA-seq and single-nuclear RNA sequencing datasets further confirmed the biological representativeness of the captured cell populations. Moreover, the enrichment of highly adhesive cell types in larger kidney clumps—contrasting with the stable cell-type composition across clump sizes in the structurally uniform intestine—suggests that clump composition reflects tissue-specific adhesion and architecture rather than a uniform technical bias.

As a platform, CCI-seq captures cellular associations across a proximity spectrum (from direct physical contact to short-range spatial colocalization) which we view as a strength that gives the method flexibility for diverse biological questions; however, the implementation in the present study was deliberately designed to enrich for interactions of physical contact. Specifically, we applied stringent clump-size filtering to retain only small clumps, after which over 50% of captured clumps consist of just two cells. Notably, the interaction patterns focusing exclusively on two-cell clumps show strong concordance with our full dataset (cosine similarities: 0.95 for mouse kidney, 0.97 for WT mouse intestine, and 0.99 for *Apc*-KO mouse intestine). In addition, the high concordance between CCI-seq interaction profiles and these closest-neighbor seqFISH pairs (cosine similarity 0.86) provides orthogonal evidence that our filtered pipeline recovers physical interactions rather than incidental colocalization.

In summary, CCI-seq offers a versatile and scalable platform to directly measure cell–cell interactions, while retaining complete single-cell transcriptomic data. While this study establishes a robust framework, the current cell throughput per experiment may limit the resolution of rare or highly complex interaction states. This constraint could be alleviated through advanced scaling strategies, such as combinatorial indexing-based[81] or droplet overloading strategies[82,83], to enable deeper coverage of cellular interactions at the tissue scale. The present implementation does not yet incorporate additional molecular layers, but the standard single-cell workflow of CCI-seq makes it inherently compatible with multimodal extensions. Future iterations could seamlessly integrate genomic (wellDR-seq[84]), epigenomic (Parallel-seq[83] and scEpi^2-seq[85]), protein level measurements (CITE-seq[86]) and immune repertoire sequencing (TCR/BCR-seq[87,88]) to more comprehensively characterize the regulatory mechanisms underlying cellular crosstalk. Its potential can be further expanded by incorporating cell enrichment strategies such as fluorescence-activated cell sorting or magnetic-activated cell sorting to target rare populations. By simultaneously mapping cell–cell interaction networks and resolving their functional impacts, CCI-seq enables deeper investigations into how cellular crosstalk shapes tissue homeostasis and disease, advancing efforts to uncover new therapeutic targets in complex diseases.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-026-03237-0.

References

1. Tabula Sapiens, C. et al. The Tabula Sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).
2. Armingol, E., Baghdassarian, H. M. & Lewis, N. E. The diversification of methods for studying cell–cell interactions and communication. *Nat. Rev. Genet.* **25**, 381–400 (2024).
3. Longo, S. K., Guo, M. G., Ji, A. L. & Khavari, P. A. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* **22**, 627–644 (2021).
4. Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nat. Methods* **19**, 534–546 (2022).
5. Cho, C. S. et al. Microscopic examination of spatial transcriptome using Seq-Scope. *Cell* **184**, 3559–3572.e22 (2021).
6. Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792.e21 (2022).
7. Fu, X. et al. Polony gels enable amplifiable DNA stamping and spatial transcriptomics of chronic pain. *Cell* **185**, 4621–4633.e17 (2022).
8. Zhang, M. et al. Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598**, 137–143 (2021).
9. Janesick, A. et al. High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat. Commun.* **14**, 8353 (2023).
10. He, S. et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat. Biotechnol.* **40**, 1794–1806 (2022).

11. Cao, J. et al. Decoder-seq enhances mRNA capture efficiency in spatial RNA sequencing. *Nat. Biotechnol.* **42**, 1735–1746 (2024).
12. Kleshchevnikov, V. et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).
13. Biancalani, T. et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat. Methods* **18**, 1352–1362 (2021).
14. Si, Y. et al. FICTURE: scalable segmentation-free analysis of submicron-resolution spatial transcriptomics. *Nat. Methods* https://doi.org/10.1038/s41592-024-02415-2 (2024).
15. Zhu, J. et al. Custom microfluidic chip design enables cost-effective three-dimensional spatiotemporal transcriptomics with a wide field of view. *Nat. Genet.* **56**, 2259–2270 (2024).
16. Nakandakari-Higa, S. et al. Universal recording of immune cell interactions in vivo. *Nature* **627**, 399–406 (2024).
17. Pasqual, G. et al. Monitoring T cell-dendritic cell interactions in vivo by intercellular enzymatic labelling. *Nature* **553**, 496–500 (2018).
18. Ge, Y. et al. Enzyme-mediated intercellular proximity labeling for detecting cell–cell interactions. *J. Am. Chem. Soc.* **141**, 1833–1837 (2019).
19. Liu, Q. et al. A proximity-tagging system to identify membrane protein–protein interactions. *Nat. Methods* **15**, 715–722 (2018).
20. Liu, Z. et al. Detecting tumor antigen-specific T cells via interaction-dependent fucosyl-biotinylation. *Cell* **183**, 1117–1133. e19 (2020).
21. Oslund, R. C. et al. Detection of cell–cell interactions via photocatalytic cell tagging. *Nat. Chem. Biol.* **18**, 850–858 (2022).
22. Liu, H. et al. Antigen-specific T cell detection via photocatalytic proximity cell labeling (PhoXCELL). *J. Am. Chem. Soc.* **144**, 5517–5526 (2022).
23. Zhang, X. et al. Cellular-scale proximity labeling for recording cell spatial organization in mouse tissues. *Sci. Adv.* **9**, eadg6388 (2023).
24. Boisset, J. C. et al. Mapping the physical network of cellular interactions. *Nat. Methods* **15**, 547–553 (2018).
25. Halpern, K. B. et al. Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. *Nat. Biotechnol.* **36**, 962–970 (2018).
26. Giladi, A. et al. Dissecting cellular crosstalk by sequencing physically interacting cells. *Nat. Biotechnol.* **38**, 629–637 (2020).
27. Andrews, N. et al. An unsupervised method for physical cell interaction profiling of complex tissues. *Nat. Methods* **18**, 912–920 (2021).
28. Manco, R. et al. Clump sequencing exposes the spatial expression programs of intestinal secretory cells. *Nat. Commun.* **12**, 3074 (2021).
29. Srivatsan, S. R. et al. Embryo-scale, single-cell spatial transcriptomics. *Science* **373**, 111–117 (2021).
30. Lee, Y. et al. XYZeq: spatially resolved single-cell RNA sequencing reveals expression heterogeneity in the tumor microenvironment. *Sci. Adv.* https://doi.org/10.1126/sciadv.abg4755 (2021).
31. Russell, A. J. C. et al. Slide-tags enables single-nucleus barcoding for multimodal spatial genomics. *Nature* **625**, 101–109 (2024).
32. Bakken, T. E. et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS ONE* **13**, e0209648 (2018).
33. McGinnis, C. S. et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).
34. Quinodoz, S. A. et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757. e24 (2018).
35. Mereu, E. et al. Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat. Biotechnol.* **38**, 747–755 (2020).
36. Wallace, M. A. Anatomy and physiology of the kidney. *AORN J.* **68**, 799–800, 803–804, 806, 808, 810–811, 813–816, 819–820 (1998).
37. Marshall, J. L. et al. High-resolution Slide-seqV2 spatial transcriptomics enables discovery of disease-specific cell neighborhoods and pathways. *iScience* **25**, 104097 (2022).
38. Polonsky, M. et al. Spatial transcriptomics defines injury specific microenvironments and cellular interactions in kidney regeneration and disease. *Nat. Commun.* **15**, 7010 (2024).
39. Schlondorff, D. & Banas, B. The mesangial cell revisited: no cell is an island. *J. Am. Soc. Nephrol.* **20**, 1179–1187 (2009).
40. Park, J. et al. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763 (2018).
41. Gueutin, V. et al. Renal β-intercalated cells maintain body fluid and electrolyte balance. *J. Clin. Invest.* **123**, 4219–4231 (2013).
42. Yang, M. et al. Spatiotemporal insight into early pregnancy governed by immune-featured stromal cells. *Cell* **186**, 4271–4288. e24 (2023).
43. Marshall, R. J. & MacIver, A. G. The monocyte/macrophage population of the normal human kidney. *J. Pathol.* **143**, 275–280 (1984).
44. Devi, S. et al. Multiphoton imaging reveals a new leukocyte recruitment paradigm in the glomerulus. *Nat. Med.* **19**, 107–112 (2013).
45. Sampangi, S. et al. Human proximal tubule epithelial cells modulate autologous B-cell function. *Nephrol. Dial. Transplant.* **30**, 1674–1683 (2015).
46. de Groot, T., Bindels, R. J. & Hoenderop, J. G. TRPV5: an ingeniously controlled calcium channel. *Kidney Int.* **74**, 1241–1246 (2008).
47. Alexander, R. T., Cordat, E., Chambrey, R., Dimke, H. & Eladari, D. Acidosis and urinary calcium excretion: insights from genetic disorders. *J. Am. Soc. Nephrol.* **27**, 3511–3520 (2016).
48. Chen, L. et al. Renal-tubule epithelial cell nomenclature for single-cell RNA-sequencing studies. *J. Am. Soc. Nephrol.* **30**, 1358–1364 (2019).
49. Castaneda-Bueno, M., Ellison, D. H. & Gamba, G. Molecular mechanisms for the modulation of blood pressure and potassium homeostasis by the distal convoluted tubule. *EMBO Mol. Med.* **14**, e14273 (2022).
50. Dimitrov, D. et al. LIANA+ provides an all-in-one framework for cell–cell communication inference. *Nat. Cell Biol.* **26**, 1613–1622 (2024).
51. Qi, J. H. et al. A novel function for tissue inhibitor of metalloproteinases-3 (TIMP3): inhibition of angiogenesis by blockage of VEGF binding to VEGF receptor-2. *Nat. Med.* **9**, 407–415 (2003).
52. Liu, S. et al. Single-cell and spatial transcriptomics unveil M-MDSC–tumor crosstalk driving EMT in HCC via THBS1-SDC4 axis. Preprint at *Research Square* https://doi.org/10.21203/rs.3.rs-6912640/v1 (2025).
53. Tuli, A. et al. Amyloid precursor-like protein 2 increases the endocytosis, instability, and turnover of the H2-K(d) MHC class I molecule. *J. Immunol.* **181**, 1978–1987 (2008).
54. Sanin, D. E. et al. A common framework of monocyte-derived macrophage activation. *Sci. Immunol.* **7**, eabl7482 (2022).
55. Beumer, J. & Clevers, H. Cell fate specification and differentiation in the adult mammalian intestine. *Nat. Rev. Mol. Cell Biol.* **22**, 39–53 (2021).
56. Haber, A. L. et al. A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
57. Moor, A. E. et al. Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *Cell* **175**, 1156–1167.e15 (2018).

58. Barker, N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat. Rev. Mol. Cell Biol.* **15**, 19–33 (2014).
59. Gehart, H. & Clevers, H. Tales from the crypt: new insights into intestinal stem cells. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 19–34 (2019).
60. Sanman, L. E. et al. Transit-amplifying cells coordinate changes in intestinal epithelial cell-type composition. *Dev. Cell* **56**, 356–365. e9 (2021).
61. Guerbette, T., Boudry, G. & Lan, A. Mitochondrial function in intestinal epithelium homeostasis and modulation in diet-induced obesity. *Mol. Metab.* **63**, 101546 (2022).
62. Barker, N. et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009).
63. Sansom, O. J. et al. Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev.* **18**, 1385–1390 (2004).
64. Andreu, P. et al. Crypt-restricted proliferation and commitment to the Paneth cell lineage following Apc loss in the mouse intestine. *Development* **132**, 1443–1451 (2005).
65. Chen, Q. et al. Paneth cell-derived growth factors support tumorigenesis in the small intestine. *Life Sci. Alliance* https://doi.org/10.26508/lsa.202000934 (2021).
66. Yui, S. et al. YAP/TAZ-dependent reprogramming of colonic epithelium links ECM remodeling to tissue regeneration. *Cell Stem Cell* **22**, 35–49.e7 (2018).
67. Mzoughi, S. et al. Oncofetal reprogramming drives phenotypic plasticity in WNT-dependent colorectal cancer. *Nat. Genet.* **57**, 402–412 (2025).
68. Flanagan, D. J. et al. NOTUM from Apc-mutant cells biases clonal competition to initiate cancer. *Nature* **594**, 430–435 (2021).
69. van Neerven, S. M. et al. Apc-mutant cells act as supercompetitors in intestinal tumour initiation. *Nature* **594**, 436–441 (2021).
70. Bala, P. et al. Aberrant cell state plasticity mediated by developmental reprogramming precedes colorectal cancer initiation. *Sci. Adv.* **9**, eadf0927 (2023).
71. Wu, X. et al. Targeting MHC-I molecules for cancer: function, mechanism, and therapeutic prospects. *Mol. Cancer* **22**, 194 (2023).
72. Anderson, P., Aptsiauri, N., Ruiz-Cabello, F. & Garrido, F. HLA class I loss in colorectal cancer: implications for immune escape and immunotherapy. *Cell. Mol. Immunol.* **18**, 556–565 (2021).
73. Glawe, J. D. et al. Genetic deficiency of Itgb2 or ItgaL prevents autoimmune diabetes through distinctly different mechanisms in NOD/LtJ mice. *Diabetes* **58**, 1292–1301 (2009).
74. Sung, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249 (2021).
75. Joanito, I. et al. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. *Nat. Genet.* **54**, 963–975 (2022).
76. Pelka, K. et al. Spatially organized multicellular immune hubs in human colorectal cancer. *Cell* **184**, 4734–4752.e20 (2021).
77. Chen, D. P. et al. Peritumoral monocytes induce cancer cell autophagy to facilitate the progression of human hepatocellular carcinoma. *Autophagy* **14**, 1335–1346 (2018).
78. Ziranu, P. et al. CD44: a new prognostic marker in colorectal cancer? *Cancers* https://doi.org/10.3390/cancers16081569 (2024).
79. Sonnentag, S. J., Ibrahim, N. S. M. & Orian-Rousseau, V. CD44: a stemness driver, regulator, and marker-all in one? *Stem Cells* **42**, 1031–1039 (2024).
80. Handler, K. et al. Fragment-sequencing unveils local tissue microenvironments at single-cell resolution. *Nat. Commun.* **14**, 7775 (2023).
81. Lu, Z. et al. Organism-wide cellular dynamics and epigenomic remodeling in mammalian aging. *Science* **391**, eadw6273 (2026).
82. Datlinger, P. et al. Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing. *Nat. Methods* **18**, 635–642 (2021).
83. Tang, L. et al. Joint analysis of chromatin accessibility and gene expression in the same single cells reveals cancer-specific regulatory programs. *Cell Syst.* **16**, 101266 (2025).
84. Wang, K. et al. Coalescing single-cell genomes and transcriptomes to decode breast cancer progression. *Cell* **188**, 6355–6369.e16 (2025).
85. Geisenberger, C. et al. Single-cell multi-omic detection of DNA methylation and histone modifications reconstructs the dynamics of epigenomic maintenance. *Nat. Methods* **22**, 2042–2051 (2025).
86. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
87. Tu, A. A. et al. TCR sequencing paired with massively parallel 3′ RNA-seq reveals clonotypic T cell signatures. *Nat. Immunol.* **20**, 1692–1699 (2019).
88. Goldstein, L. D. et al. Massively parallel single-cell B-cell receptor sequencing enables rapid discovery of diverse antigen-reactive antibodies. *Commun. Biol.* **2**, 304 (2019).

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026

Methods

Ethics statement

All research complies with all relevant ethical regulations. Animal experiments were performed with the approval of the Institutional Animal Care and Use Committee of Tsinghua University (17-ZQF1.G23-1). Human CRC samples were collected and used with the approval of the Medical Research Ethics Committee of Beijing Chao-Yang Hospital, Capital Medical University, and written informed consent was obtained from all participants.

Cell culture

The human HEK293T (SCSP-502) and mouse NIH/3T3 (SCSP-515) cell lines were newly purchased from the Cell Bank of the Chinese Academy of Sciences. The identities of these cell lines were authenticated by the vendor before purchase, and they were routinely tested negative for *Mycoplasma* contamination. Neither line is listed in the International Cell Line Authentication Committee database of misidentified cell lines. HEK293T and NIH/3T3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (C11995500BT, Thermo Fisher) supplemented with 10% fetal bovine serum (FBS) (P30-3302, PAN BIOTECH) at 37 °C with 5% CO_2. Cells were rinsed with PBS (C10010500BT, Thermo Fisher) and detached by incubating for 3–5 min at 37 °C with 1 ml of 0.25% Trypsin-EDTA (25200114, Thermo Fisher). Detached HEK293T and NIH/3T3 cells suspensions were collected via centrifugation, washed with PBS and counted using a C-Chip Disposable Hemocytometer (DHC-N01N, As One).

Three-dimensional spheroids culture

HEK293T and NIH/3T3 cells were cultured as mentioned above. AggreWellTM400 six-well microwell plates (34421, StemCell) were prepared following the manufacturer's instructions. A total of 42,000 cells were seeded per well (approximately six cells per microwell) and cultured in the humidified incubator at 37 °C with 5% CO_2 for 24 h. HEK293T and NIH/3T3 spheroids were carefully collected, and single cells were separated by passing through a 10-µm strainer. Approximately 20,000 spheroids per type were prepared and labeled with human- or mouse-specific indexed CMOs, followed by three rounds of split-pooling barcoding.

Mouse studies

Six WT C57BL/6J male mice (6–8 weeks) were purchased from the Laboratory Animal Research Center at Tsinghua University. Mice were housed in a specific-pathogen-free facility under a standard 12-h light–dark cycle. The ambient temperature was maintained at 20–24 °C, and the humidity was kept at 40–60%. Three *Apc*$^{flox/flox}$; *Villin*-CreERT2 male mice[64] (about 12 weeks old), kindly provided by Y.-G. Chen (Tsinghua University), were intraperitoneally injected with 200 µl of tamoxifen dissolved in sunflower oil at a concentration of 10 mg ml^{-1}, and the small intestine was isolated 7 days later for CCI-seq or tissue fixation.

Clump isolation from mouse kidney

Kidneys from WT C57BL/6J mice were collected and transferred to a 1.5-ml EP tube. We added 1 ml of dissociation solution (4.8 mg ml^{-1} Dispase II, 3 mg ml^{-1} Collagenase IV, 0.1 mg ml^{-1} DNase I and 10 mM $CaCl_2$ in PBS) and used scissors to repeatedly mince the tissue on ice for about 2 min until most of the tissue fragments were approximately 1 mm in size. Then, we added 4 ml dissociation solution and incubated at 37 °C for 20 min while shaking at 100 rpm. We transferred the minced tissue to a 70-µm cell strainer and rinsed with PBS repeatedly to collect the filtrate. We centrifuged the filtered cells at ~188*g* (1,000 rpm) for 3 min at 4 °C, discarded the supernatant, and added 1 ml of red blood cell lysis buffer (A1049201, Thermo Fisher). We let it sit at room temperature for 5 min, then centrifuged and discarded the supernatant. We resuspended the cell clumps in PBS, and they were then ready for subsequent experiments.

Clump isolation from mouse intestine

We removed the small intestines of WT C57BL/6J and *Apc*$^{flox/flox}$; *Villin*-CreERT2 mice, and kept them on ice in PBS. We washed the lumen three times with PBS and removed any adipose tissue. We opened the small intestine longitudinally and gently rubbed off any remaining mucus. We washed the tissue once in PBS before cutting it into 4-mm long fragments. We immersed the fragments in 10 mM EDTA-PBS and incubated for 15 min on ice. We allowed the fragments to settle at the bottom and discarded the supernatant. We then added cold PBS and shook the sample vigorously. After allowing the fragments to settle at the bottom, the supernatant was collected as fraction 1 in a new 15-ml tube. We repeated this procedure once or twice to collect fraction 2 and fraction 3. The three fractions were strained through a 70-µm filter, respectively. Using light microscopy, we checked which fraction enriched crypts, and centrifuged the fractions at 300*g* for 5 min.

CRC samples collection and ethics statement

Tumor samples were obtained from the resected specimens of five patients (three males, two females; age range 40–70 years) who were pathologically diagnosed with CRC and underwent surgery at the Department of General Surgery, Beijing Chao-Yang Hospital, Capital Medical University. Paired normal tissue samples were collected from areas at least 5 cm away from the tumor margin during the same surgical procedure. Participants were not compensated for their participation in this study.

Clump isolation from human CRC samples

We collected CRC tumor samples and normal samples from more than 5 cm away from the tumor. We quickly transferred them into tissue storage solution (130-100-008, Miltenyi Biotec), ensuring they were fully submerged, and placed them on ice. We transferred them promptly to the laboratory for processing. Tissues were removed and then rinsed three times with PBS to eliminate the preservation solution. We transferred the tissues to a 1.5-ml EP tube, added 200 µl of PBS, and minced them with scissors for approximately 2 min until the fragments were approximately 1 mm^3 in size. We transferred the tissue fragments to a 70-µm cell strainer, rinsed with PBS repeatedly, and collected the filtrate. We centrifuged the collected cell clumps at 500*g* for 3 min at 4 °C, discarded the supernatant, and resuspended the cell pellet in 1 ml of red blood cell lysis buffer using a wide-bore pipette tip. We let the sample sit at room temperature for 5 min, then centrifuged and discarded the supernatant. We resuspended the cell clumps with a wide-bore pipette tip for subsequent experiments.

A standardized protocol with measures and controls of CCI-seq

CCI-seq employs multiple measures and controls to ensure the reproducibility and robustness of the dissociation process. Specifically, we applied gentle dissociation to preserve cell clumps in their native conditions, using low trypsin activity, or mechanical partial dissociation. We also carefully handled clumps with wide-bore pipette tips to minimize unwanted mechanical dissociation. After that, we used a 70-µm cell strainer for standardized-size clump selection, removing large cell clumps (>20 cells, this also to avoid potential artifacts from indirect interactions in large clumps). Of note, we performed rigorous microscopic inspection as key controls to ensure all samples met predefined quality criteria (including cell viability, CMO labeling and clump sizes). Using the mouse kidney as an example, we established the following QC thresholds (1) >80% of cell viability post-dissociation; (2) effective CMO labeling of >95% cells within each clump; and (3) a clump size of <20 cells (Extended Data Fig. 2a).

CCI-seq protocol

Preparation of ligation plates. Oligonucleotides were ordered from Sangon Biotech (Shanghai) Co. Odd and even ligation adaptor

plates were prepared as per the SPRITE protocol and are listed in Supplementary Table 7. We annealed the top and bottom strands of the odd and even plates separately in 1× Annealing Buffer (100 mM Tris-HCl, pH 7.5 and 1.5 mM EDTA) with 10 µM final concentration for each strand. The annealing conditions were 95 °C for 2 min, followed by a slow cooling to room temperature at a rate of 0.1 °C s^{-1}. The terminal ligation adaptor plate was dissolved in nuclease-free water to create a 100 µM stock solution, which was then diluted to 10 µM. The odd, even and terminal plates were distributed into 96-well plates, with each well containing 10 µl of the adaptor.

Anchor and co-anchor labeling. CCI-seq labeled DNA barcodes to plasma membranes by hybridization to an 'anchor' CMO. The anchor and co-anchor CMOs designs were adapted from Multi-seq[33] and ordered from Sangon Biotech (Shanghai) Co. We conjugated a cholesterol via a triethylene glycol (TEG) linker to the 3′ end of co-anchor CMO or 5′ end of anchor CMO, respectively. The up/down oligonucleotide duplex was prepared by annealing 10 µM up and 10 µM down DNA oligonucleotides (Supplementary Table 7). The oligonucleotides were heated to 95 °C for 2 min, then gradually cooled to 20 °C at 0.1 °C s^{-1}, yielding a 10 µM duplex solution. For the 10× Anchor solution, 2.2 µl of the annealed 10 µM up/down duplex was mixed with 2.2 µl of 10 µM anchor CMO and 17.6 µl of PBS. This solution (20 µl) was then diluted with 180 µl of PBS and used to resuspend cell clumps, followed by incubation on ice for 5 min. Next, 20 µl of 2 µM co-anchor was added, mixed thoroughly and incubated on ice for another 5 min. Finally, 1 ml of ice-cold 1% BSA in PBS was added, and the sample was centrifuged at ~188*g* (1,000 rpm) for 3 min using a vertical rotor. The supernatant was carefully removed.

Combinatorial index ligation. The ligation reaction was carried out using a modified SPLiT-seq protocol. We prepared 2 ml of 1× NEBuffer 3.1 and 2 ml of ligation solution, which contained 500 µl of 10× T4 DNA ligation buffer, 100 µl of T4 DNA ligase, 50 µl of 20% BSA and 1,350 µl of UltraPure DNase/RNase-free distilled water. The pooled cell clumps were resuspended in the 1× NEBuffer 3.1 and thoroughly mixed with the ligation solution using wide-bore pipette tips. Next, 40 µl of the cell-ligation mixture was added to each well of the prepared odd-numbered ligation plate and incubated at room temperature with rotation at 15 rpm for 5 min. Following ligation, 2 µl of 500 µM EDTA was added to each well, and the cell clumps were mixed. Then, 50 µl of 20% BSA was added, and the mixture was centrifuged at 500 g for 5 min at 4 °C. The supernatant was removed, and the cell clumps were washed with 950 µl of PBS containing 1% BSA. The cell clumps were then resuspended in 2 ml of 1× NEBuffer 3.1 and mixed with 2 ml of ligation solution using wide-bore pipette tips. Subsequently, 40 µl of the cell-ligation mixture was distributed into the even ligation plate, and the ligation and wash steps were repeated. This process was repeated with the terminal plates. Finally, the cell clumps were pooled and washed with 950 µl of PBS containing 1% BSA.

Clump dissociation. About 10,000 clumps were counted and dissociated with TrypLE Express Enzyme (1×), no phenol red (12604039) at the 37 °C for 5 min with shaking at 800 rpm. During enzymatic dissociation, light microscopy was used to monitor the process and obtain an appropriate number of single cells and multiplets. Cells were filtered through 35-µm strainer.

Sequencing library generation. All the filtered single cells were loaded into 10x Genomics 3′ V3.1 kit. mRNA library and Cell-ID library were prepared as 10x Genomics protocol described.

Libraries sequencing. Libraries of mRNA and Cell-ID were sequenced to 400 million and 200 million reads, respectively, with NovaSeq 6000 System S4 kit and PE150 sequencing mode.

scRNA-seq data analysis

After sequencing, CellRanger software (v.7.1.0, 10x Genomics) was used to prepare fastq files, align reads to the mm10 genome for mouse samples and hg38 genome for human samples, and ultimately generate gene-by-cell matrices. The Python package scanpy (v.1.9.1) was used to filter cells and genes. We excluded cells with fewer than 200 detected genes or had more than 20% mitochondrially mapped reads, and excluded genes with fewer than three detected cells. The Python package scalex (v.1.0.2) was used to batch-correct and integrate data from different experiments. The Leiden algorithm[89] was performed for clustering, and UMAP was used for two-dimensional visualization. Cell types were manually annotated based on the expression of canonical marker genes from previous literature.

Combinatorial index data analysis

After sequencing, combinatorial index reads were first deduplicated using seqkit[90] (v.2.5.1). Next, umi_tools[91] (v.1.1.4) was used to extract cell barcodes based on the pattern of the index library. Then cutadapt[92] (v.4.4) tool was used to identify and remove adaptor sequences, outputting standardized index reads for each cell.

Theoretically, the most abundant index (Cell-ID) is considered as the cell identifier and is used to infer whether cells are from the same clump. To reduce noise, we performed quality control and filtering for the Cell-IDs. Specifically, we defined droplets with qualified gene expression data as 'Valid droplets' and those without as 'Empty droplets'. Cells were considered to have a qualified Cell-ID if they met all three of the following criteria:

1. $N_{\text{Valid droplet}} > \{0.75 \text{ quantile of } N_{\text{Empty droplet}}\}$, where N represents the number of combinatorial indices.
2. $R_1_{\text{Valid droplet}} > \{0.75 \text{ quantile of } R_1_{\text{Empty droplet}}\}$, where R_1 represents the ratio of the rank 1 index.
3. $N_1_{\text{Valid droplet}}/N_2_{\text{Valid droplet}} > 2$, where N_1 and N_2 represent the number of the rank 1 index and the rank 2 index.

Finally, we retained the cells assigned to clumps with no more than 20 cells.

Clustering analysis of clump compositions

We collected all cell clumps (≥2 cells) and generated a fixed-length vector representing the cell composition for each clump, followed by Leiden clustering and UMAP visualization.

Quantification of interaction strength

We introduce two metrics to quantify the interaction strength for any pairs of cell types.

Frequency. For any two cell types, the interaction frequency is calculated as the ratio of multiplied cell numbers (for different cell types) or pairwise combinations (for the same cell type) to the total pairwise combinations within each clump, then summarizing these ratios across all clumps and normalizing by their total cell numbers. This metric captures the relative abundance of co-occurring cells across clumps and serves as a reference for identifying functional cell–cell interactions.

$$\text{Frequency}_{AB} = \begin{cases} 100 \times (N_A \times N_B)^{-0.5} \times \sum_{i=0}^{K} \frac{n_{iA} \times n_{iB}}{\binom{2}{n_i}}, \; if A \neq B \\ 100 \times (N_A \times N_B)^{-0.5} \times \sum_{i=0}^{K} \frac{\binom{2}{n_A}}{\binom{2}{n_i}}, \; if A = B \end{cases} \quad (1)$$

Where N_A and N_B represent the total number of cells of cell type A and B across all clumps, while n_{iA} and n_{iB} represent the total number of cells of cell type A and B in clump i. n_i represents the total number of cells in clump i, and K represents the total number of clumps.

Clump. It is calculated as the co-occurrence of two cell types across all cell clumps.

$$\text{Clump}_{AB} = \sum_{i=0}^{K} I(\text{clumps with both } A \text{ and } B) \quad (2)$$

where K represents the number of clumps.

Significance analysis of cell–cell interactions

To test the statistical significance for interactions across each cell type pairs, we conducted a permutation-based test. In brief, we scrambled cell type labels while preserving the distribution of clump sizes and re-computed the interaction frequency across cell type pairs in each permutation (default $n = 1{,}000$). The enrichment score was calculated as the true interaction frequency divided by the mean of the permutated interaction frequency, and the P value was calculated based on the proportion of the permutated interaction frequency higher than the true interaction frequency across 1,000 permutations.

Orthogonal validation using seqFISH spatial transcriptomics data

Orthogonal validation of CCI-seq-identified networks was performed using a high-resolution mouse kidney seqFISH dataset. To minimize the inclusion of nondirect-contacting, colocalized cells, we restricted the spatial interaction matrix to include only the single nearest neighbor for each cell. Utilizing the squidpy package, we defined spatial neighbors via gr.spatial_neighbors with parameter n_neighs = 1 and computed the interaction matrix using gr.interaction_matrix. Interaction frequencies were calculated by dividing raw interaction counts by the total number of cells per cell type, enabling a comparison between the seqFISH-derived frequencies and our CCI-seq results.

AsmFISH and immunofluorescence staining

AsmFISH was performed following the manual[42] provided by Suzhou Dynamic Biosystems Co. In brief, tissues were fixed in 10 ml ISH Fixative Solution (animal) (G1113-500ML, Servicebio) for 24 h at room temperature and then embedded into paraffin. Tissues were sectioned at 5 µm and permeabilized for probe hybridization. The fluorescent signals were amplified by probe ligation, circularization and rolling circle amplification. The slides were mounted with an antifade mounting medium containing DAPI.

Multiplex immunofluorescence utilized tyramide signal amplification (TSA; Servicebio)[93]. Formalin-fixed paraffin-embedded sections underwent dewaxing, dehydration and antigen retrieval. Following endogenous peroxidase inactivation (3% H_2O_2, 25 min, dark) and blocking (3% BSA, 30 min), sections were incubated with the first primary antibody (anti-MS4A1, Servicebio, GB155721, 1:4,000 dilution; or anti-SCA1, Abcam, ab109211, 1:1,000 dilution; overnight, 4 °C), an HRP-conjugated secondary antibody (50 min, room temperature) and a TSA fluorophore (10 min, dark). After microwave-mediated antibody stripping, this cycle was repeated for the second target using appropriate primary/secondary antibodies (anti-SLC13A3, Servicebio, GB115110, 1:5,000 dilution; or anti-Ki67, Servicebio, GB111141, 1:3,000 dilution) and a distinct TSA fluorophore. Finally, slides were DAPI-counterstained (10 min), quenched for autofluorescence (5 min) and mounted with an antifade medium.

Whole-slide images were acquired using a VS200 slide scanner (Olympus), and high-resolution fluorescence images were captured using a STELLARIS 8 Falcon confocal microscope (Leica Microsystems). Following image acquisition, all fluorescence images were visualized and processed using Imaris software (v.10.1, Oxford Instruments).

Single-cell genotyping

Individual SCA1$^+$EPCAM$^+$ cells were sorted from the intestines of tamoxifen-treated or untreated *Apc*$^{flox/flox}$; *Villin*-CreERT2 mice. To genotype the *Apc* locus, single cells were subjected to a nested three-primer PCR assay (Supplementary Table 7). In the second round of amplification, primers 1 and 2 were used to flank the first *loxP* site, generating a ~380-bp product for the WT allele. Simultaneously, primers 1 and 3 were designed to span the deletion junction, generating a ~240-bp product for the recombined allele (*Apc*-KO). PCR products were resolved by gel electrophoresis and validated by Sanger sequencing.

Differential gene and Gene Ontology term enrichment analysis

The ranking for the highly differential genes in each cluster was computed by rank_genes_groups function with method = 't-test' in the scanpy package. Genes with $\log_2$(fold change) > 0.5 and Benjamini–Hochberg adjusted P value < 0.05 were defined as DEGs. We performed gene set enrichment analysis for the top 200 DEGs sorted by gene scores (implemented in scanpy) using the enrichGO function with pAdjustMethod = BH in the clusterProfiler package[94] (v.4.10.1).

Landmark score

Inspired by Moor et al.[57], the landmark (LM) score for each cell i is calculated as:

$$\text{LM_score}_i = \frac{\sum_{g \in \text{tLM}} E_{g,i}}{\sum_{g \in \text{bLM}} E_{g,i} + \sum_{g \in \text{tLM}} E_{g,i} + 1e^{-8}} \quad (3)$$

Where $E_{g,i}$ represents the expression of gene g in cell i, tLM is a set of top landmark genes, and bLM is a set of bottom landmark genes.

Definition of spatial subtypes of enterocytes and goblets

The annotations of villus-tip, villus-middle and villus-bottom enterocytes relied exclusively on spatially relevant landmark genes (a set of genes specifically expressed at different villus region) identified in a published study[57], and independently of any clump information from our data. As a proof of concept, these three enterocyte spatial subtypes, with their well-established spatial organization, serve as a ground truth to validate the accuracy of the 'interacting cells' inferred by CCI-seq based on cell clump information. Spatial subtypes of goblets were delineated using the same approach.

Definition of TA spatial subtypes

We extracted double-cell clumps composed of one TA cell and one cell of a different type. Crypt-top TA cells were defined as TA cells interacting with villus-located cells, including villus-bottom enterocytes, villus-middle enterocytes, villus-tip enterocytes and villus goblet cells, whereas crypt-bottom TA cells were defined as TA cells interacting with crypt-located cells, including crypt goblet cells, Paneth cells and *Lgr5*$^+$ ISCs.

Pseudotime analysis

We performed pseudotime inference using the partition-based graph abstraction algorithm implemented in the scanpy package, which was also used for all major preprocessing steps. To root the trajectory, we selected a villus-bottom enterocyte based on known spatial organization and differentiation patterns in the small intestine epithelium. This cell was chosen randomly from the villus-bottom region to serve as the starting point for the pseudotime computation.

Comparative analysis of CCI-seq and uLIPSTIC

A key advantage of CCI-seq is its ability to identify the full cell–cell interaction networks within a tissue in an unbiased manner, which sets it apart from proximity labeling methods. To demonstrate this, we compared CCI-seq and uLIPSTIC—an advanced proximity labeling-based method for detecting cell–cell interactions using mouse small intestine

samples. For quantitative comparison, we normalized the interaction signals detected by uLIPSTIC by constructing a virtual cell clump consisting of one IEC and one immune cell, and then calculated the interaction frequency metric.

CRC annotation by label transfer

Python package scalex (v.1.0.2) was used to annotate cell types for human CRC data. Specifically, we first integrated a publicly available CRC scRNA-seq dataset[76] with our CRC data and projected all cells into a common cell embedding space. Next, we used the KNeighborsClassifier function from the scikit-learn package to train a prediction model by the cell embedding representations of the public dataset with known cell type labels. The trained model was then applied to predict cell types for the unlabeled cells in our CRC dataset based on their embeddings.

Infer CNVs from CRC dataset

We used the Python package infercnvpy (v.0.4.2) to infer copy-number variations (CNVs) by averaging gene expression over genomic regions for our CRC dataset. All cell types except for epithelial cells were annotated as normal cells. The window_size parameter in the cnv.tl.infercnv function was set to 250. All other parameters were default. The per-gene copy number scores calculated for each cell of each cohort were visualized using the cnv.pl.chromosome_heatmap function.

Ligand–receptor analysis by LIANA+

We used the Python package liana (v.1.5.0) to infer ligand–receptor interactions from scRNA-seq data. Following the recommended pipeline, we first conducted an anndata object containing processed single-cell transcriptomics data with pre-annotated cell types. We then computed the consensus ligand–receptor scores using the rank_aggregate function, aggregating predictions from multiple methods. The parameter expr_prop was set to 0.1, while all other settings were default.

Statistics and reproducibility

No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment. For all quantitative experiments (for example, sequencing and data analysis), the findings were consistent across at least three independent biological replicates. For all representative micrographs shown in the Figures, Extended Data Figures and Supplementary Figures, the experiments were repeated independently twice (biological replicates) with similar results obtained.

Reporting summary

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

Data availability

Raw sequence reads and count matrices generated in this study are available in the Genome Sequence Archive at the National Genomics Data Center under accession number PRJCA034558. All processed data supporting the key findings of this study are available in the Zenodo repository at https://doi.org/10.5281/zenodo.19544434 (ref. 95) and the OMIX database at the NGDC under accession number OMIX016046. Additionally, the processed data specifically for the human CRC samples are also available in the Gene Expression Omnibus under accession number GSE327054. All processed data are publicly available as described above; however, to comply with the Regulations of the People's Republic of China on the Administration of Human Genetic Resources and to protect patient privacy, access to the raw sequence reads of human patients is available upon request for scientific research. The publicly available mouse small intestine Visium HD dataset can be accessed from the 10x Genomics (https://www.10xgenomics.com/datasets/visium-hd-cytassist-gene-expression-libraries-of-mouse-intestine). Mouse kidney Slide-seq-V2 and seqFISH data are available at GSE190094 and https://doi.org/10.5061/dryad.bnzs7h4hj (ref. 96), respectively. Source data are provided with this paper.

Code availability

The source code for the analysis is available on GitHub (https://github.com/zhangqf-lab/CCI-seq) and at the Zenodo repository (https://doi.org/10.5281/zenodo.19544434)[95] with an MIT License. For reproducibility, the Jupyter notebooks for data analyses are also available in the above repository.

References

89. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
90. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, e0163962 (2016).
91. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
92. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
93. Yu, M. et al. NSUN6-mediated 5-methylcytosine modification of NDRG1 mRNA promotes radioresistance in cervical cancer. *Mol. Cancer* **23**, 139 (2024).
94. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).
95. Tian, K. & Tang, L. Code for: resolving cell–cell interaction networks and their molecular logic in complex tissues. *Zenodo* https://doi.org/10.5281/zenodo.19544434 (2026).
96. Polonsky, M. et al. Data from: spatial transcriptomics defines injury specific microenvironments and cellular interactions in kidney regeneration and disease. *Dryad* https://doi.org/10.5061/dryad.bnzs7h4hj (2024).

Acknowledgements

We thank S. Fu (Guangzhou Laboratory) for conceptual input; P. Wang, Q. Xue and H. Li for their assistance with the mouse experiments; and the Laboratory Animal Research Center at Tsinghua University for animal husbandry and management. We thank the following core facilities at Tsinghua University for their technical support: Y. Dang (THU-IDG/McGovern Open Laboratory of Shared Instruments for Brain Science) for assistance with 10x Genomics platforms, Y. Sun (Center of Biomedical Analysis) for confocal microscopy, B. Liu (Imaging Core Facility, Technology Center for Protein Sciences) for assistance with microscopy image visualization, and P. Jiao (Center of Biomedical Analysis) for cell sorting. Finally, we thank all members of the Q.C.Z. laboratories for their valuable discussions and contributions.

Author contributions

Q.C.Z. conceived the project, and supervised the project with C.F. and W.W. L.T. designed the experimental pipeline with the help of K.T. L.T., X.F. and Y.X. implemented the CCI-seq experiments. Y.X., L.T., X.F. and J.W. implemented the mouse experiments. C.F. and C.Y. collected patient tissues. K.T. constructed the computational pipeline and implemented data analysis with the help of L.T., X.F., Y.X. and J.W. J.Z., X.-W.W. and Q.W. participated in project discussions and provided critical intellectual input. Q.C.Z., L.T. and K.T. wrote the paper, with inputs from all the authors.

Funding

Q.C.Z. discloses support for the research of this work from the National Natural Science Foundation of China (grant nos. 32125007 and 32230018), the Science & Technology Fundamental Resources

Investigation Program (grant no. 2025FY100300), the Tsinghua University Dushi Program (grant no. 20251080102), the TsienTang Life Science Development Fund at Tsinghua University, and the New Cornerstone Science Foundation through the XPLORER PRIZE. All other authors declare no relevant funding.

Competing interests

Q.C.Z., L.T., K.T., Y.X. and X.F. are inventors on one or more patents or patent applications filed by Tsinghua University covering the methods, reagents and data described in this paper. The other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41592-026-03237-0.

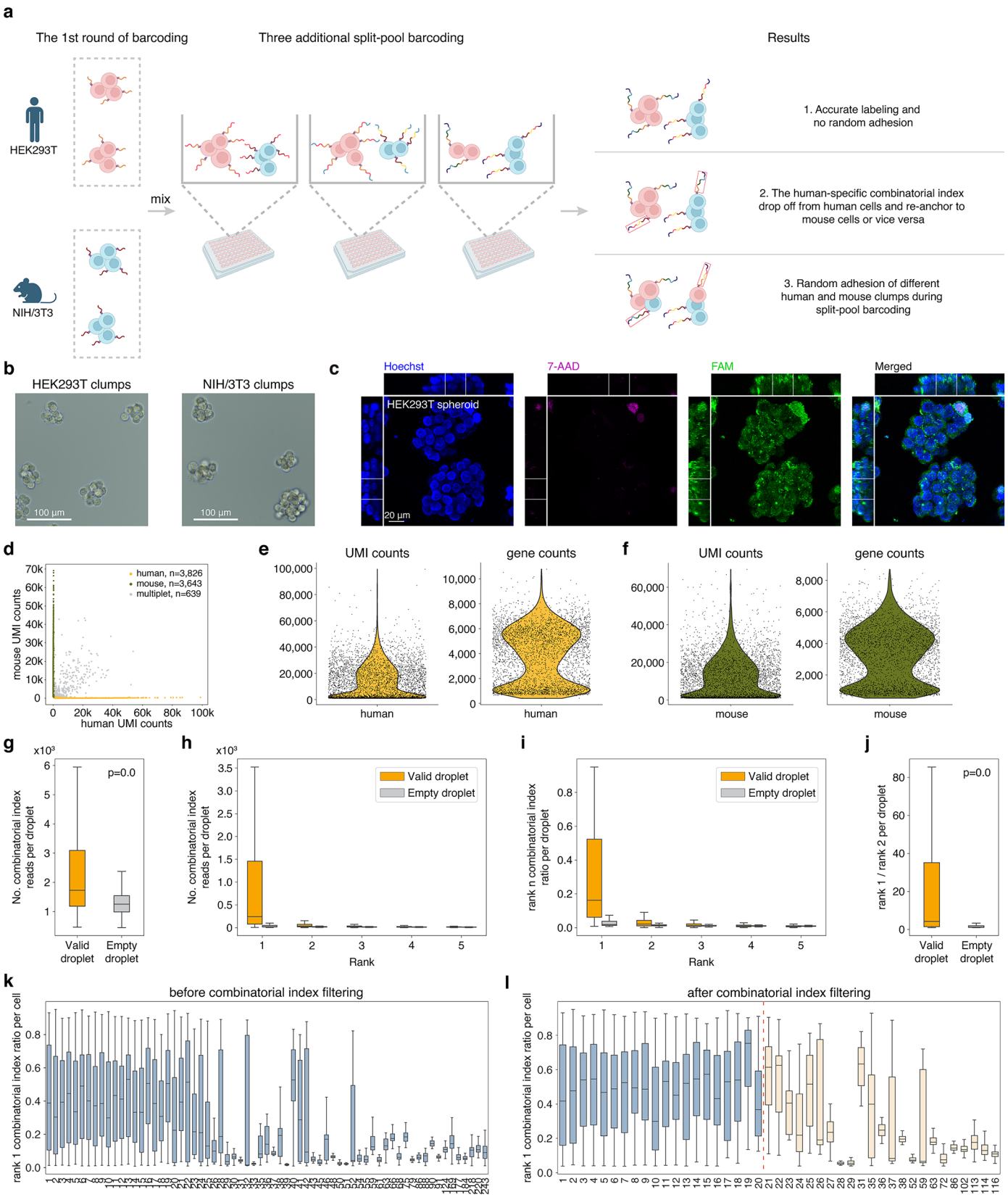
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-026-03237-0.

Correspondence and requests for materials should be addressed to Kang Tian, Wei Wu, Changjiang Feng or Qiangfeng Cliff Zhang.

Peer review information *Nature Methods* thanks Louis Vermeulen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available. Primary Handling Editors: Lei Tang and Madhura Mukhopadhyay, in collaboration with the *Nature Methods* team.

Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41592-026-03237-0

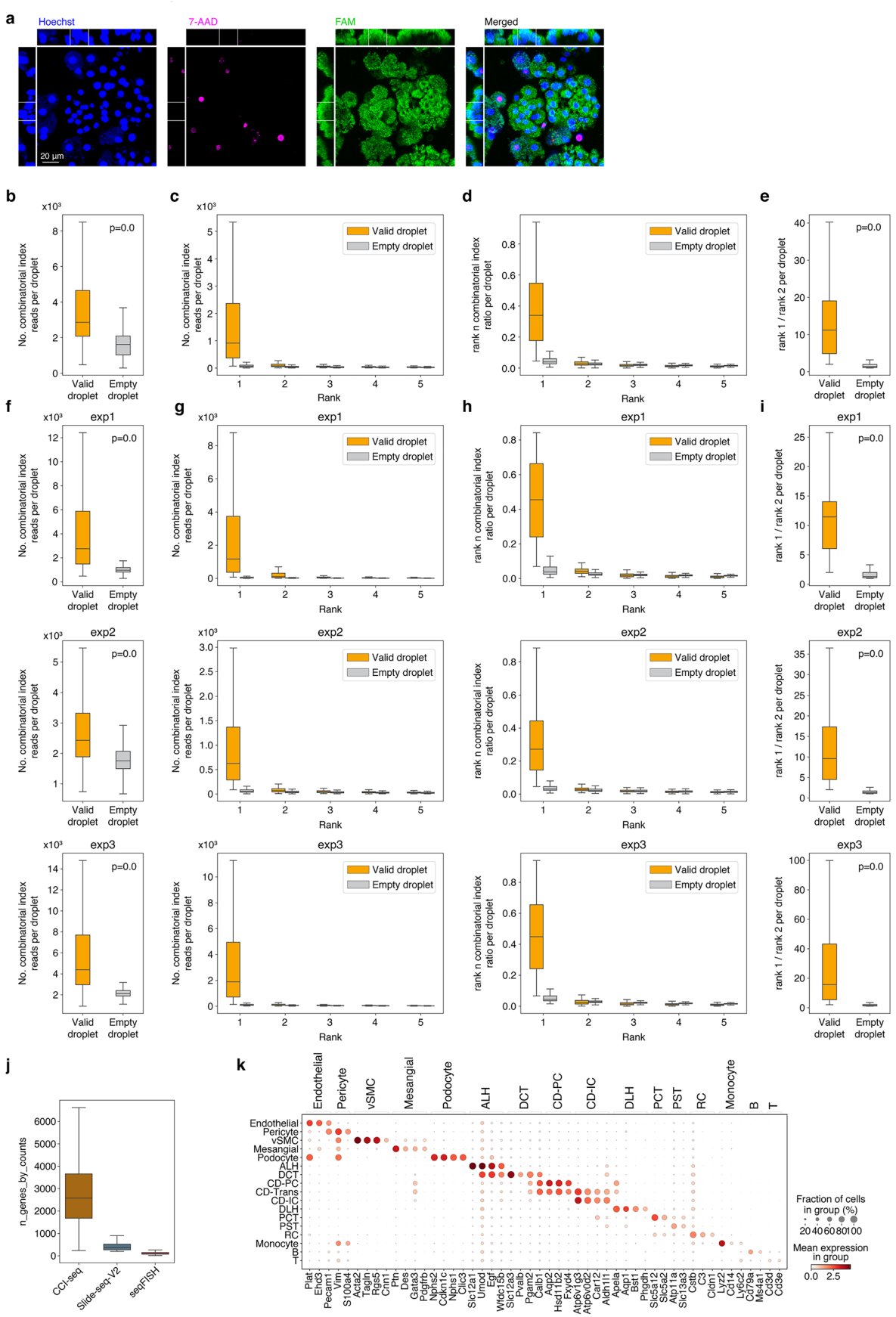


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41592-026-03237-0

Extended Data Fig. 1 | The human-mouse mixing experiment design and CCI-seq data quality control. a, Overview of the human-mouse mixing experiment design. The human HEK293T and mouse NIH/3T3 three-dimensional spheroids were cultured to approximately 20 cells in size. The first round of barcoding labeled them with human- and mouse-specific CMO barcodes. Then, these cell clumps were mixed together and uniquely labeled through three rounds of split-pool barcoding, followed by complete dissociation and single-cell sequencing. Three possible results: 1. Accurate labeling and no random adhesion. 2. The human-specific combinatorial index drop off from human cells and re-anchor to mouse cells or vice versa. 3. Random adhesion of different human and mouse cell clumps during split-pool barcoding. **b**, Microscopic images of human HEK293T (left) and mouse NIH/3T3 (right) spheroids before labeling. **c**, Confocal image of the FAM-conjugated barcodes (green) labeled on cells within HEK293T spheroids. FAM, Fluorescein amidites; Blue, Hoechst-stained nuclei; Magenta, 7-AAD stained dead cells. **d**, The UMI counts mapped to human and mouse genomes. Define the 10th percentile of barcodes where (mouse > human UMI counts) as the threshold for calling mouse cells, and the 10th percentile of barcodes where (human > mouse UMI counts) as the threshold for calling human cells. Barcodes exceeding both thresholds were classified as multiplets. Cells with UMI counts less than 1,000 were designated as low quality and filtered out. **e, f**, Violin plots showing the UMI counts and gene counts for human HEK293T cells (**e**) and mouse NIH/3T3 cells (**f**). **g-j**, Boxplot showing the following metrics within each droplet before filtering: the total number of combinatorial indices (**g**), the number of top 5 indices (**h**), the ratio of top 5 indices (**i**), and the ratio of rank 1 to rank 2 indices (**j**). Valid droplets, droplets with qualified RNAs, $n = 9{,}080$. Empty droplets, droplets without qualified RNAs, $n = 85{,}449$. Boxes, interquartile range. Midline, median. Whiskers, 1.5× interquartile range. Significance estimated using a one-sided Student's t-test. **k,l**, Boxplot showing the ratio of rank 1 combinatorial indices for cells in clumps with varying sizes before (**k**) and after (**l**) filtering. Detailed information for n is provided in the Source Data. Boxes, interquartile range. Midline, median. Whiskers, 1.5× interquartile range. Dashed line indicates the size threshold for cell clumps filtering. Panel **a** created in BioRender; Tang, L. https://biorender.com/2tqsuuv (2026).

https://doi.org/10.1038/s41592-026-03237-0

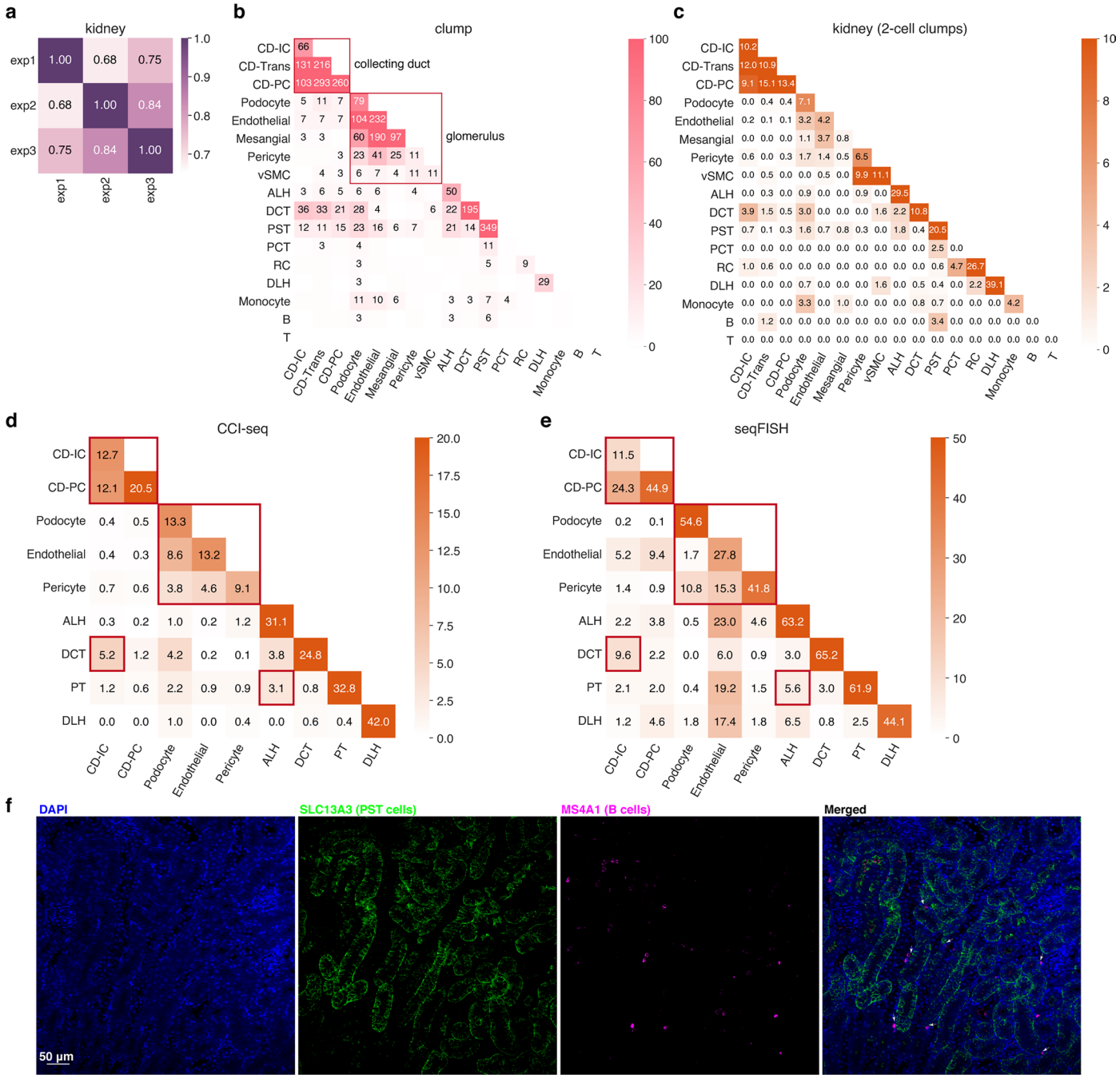


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41592-026-03237-0

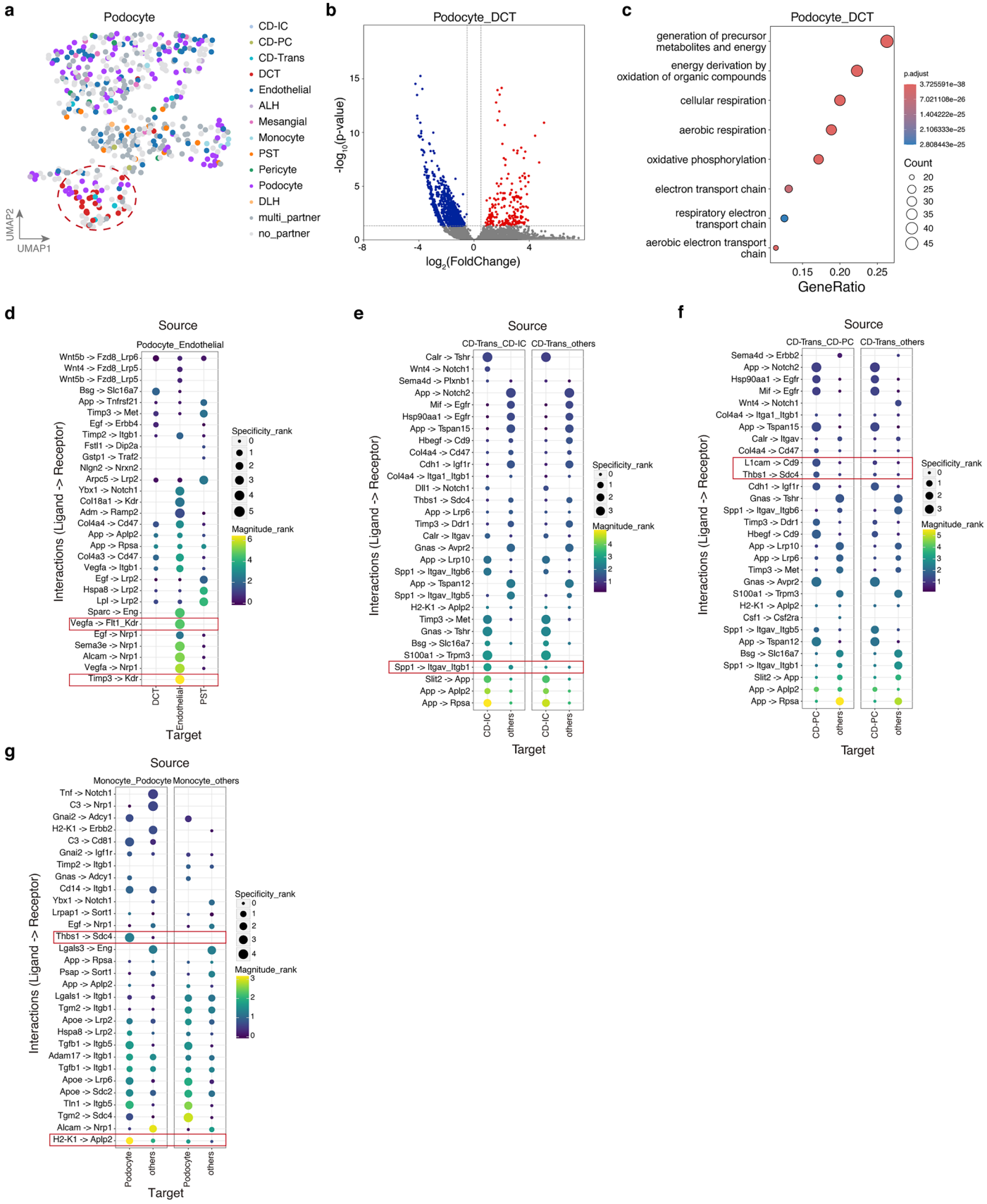
Extended Data Fig. 2 | CCI-seq data quality control for the mouse kidney. **a**, Confocal images of the FAM-conjugated barcodes (green) labeled on cells within clumps of the mouse kidney after CMO labeling. FAM, Fluorescein amidites; Blue, Hoechst-stained nuclei; Magenta, 7-AAD stained dead cells. **b**–**e**, Boxplot showing the following metrics within each droplet after filtering: the total number of combinatorial indices (**b**), the number of top 5 indices (**c**), the ratio of top 5 indices (**d**), and the ratio of rank 1 to rank 2 indices (**e**). Valid droplets, droplets with qualified RNAs, n = 11,519. Empty droplets, droplets without qualified RNAs, n = 213,286. Boxes, interquartile range. Midline, median. Whiskers, 1.5× interquartile range. Significance estimated using a one-sided Student's t-test. **f**-**i**, The same plots as in **b**–**e** when applied to three independent experiments. exp1, Valid droplets, n = 1,496, Empty droplets, n = 89,596; exp2, Valid droplets, n = 6,452, Empty droplets, n = 52,594; exp3, Valid droplets, n = 3,571, Empty droplets, n = 71,096. **j**, Boxplot showing the number of genes detected by CCI-seq (n = 11,519), Slide-seq-V2 (n = 372,732), and seqFISH (n = 55,856) in mouse kidney samples. Boxes, interquartile range. Midline, median. Whiskers, 1.5× interquartile range. Significance estimated using a one-sided Student's t-test. CCI-seq vs. Slide-seq-V2, p-value < 0.001; CCI-seq vs. seqFISH, p-value < 0.001. **k**, Dotplot showing the expression of marker genes for different cell types. Each dot represents the mean expression within each cell type (visualized by color) and the fraction of cells expressing the marker gene in each cell type (visualized by the size of the dot).

https://doi.org/10.1038/s41592-026-03237-0



Extended Data Fig. 3 | Cell-cell interaction analysis for the mouse kidney. **a**, Heatmaps showing the cosine similarity of interaction frequencies between different biological replicates for the mouse kidney. **b**, Heatmap plot showing the cell–cell interactions quantified by *clump* (Methods) between cell types. Values below 3 are not shown. Red boxes highlight interactions discussed in the manuscript. **c**, Heatmaps showing the interaction frequencies between cell types within 2-cell clumps for the mouse kidney. **d,e**, Heatmaps showing the interaction frequencies between cell types in the mouse kidney data identified by CCI-seq (**d**) and seqFISH (**e**). Red boxes highlight consistent interactions within the collecting duct and glomerular regions identified by both techniques, as well as consistent interactions between cell types, such as CD-IC with DCT and ALH with PT. For any pair of cell types in the seqFISH dataset, we used the *gr.interaction_matrix* function from *squidpy* package to compute interaction matrix, and normalized the values by the total number of cells in each cell type to calculate the interaction frequencies. **f**, Representative examples of the mouse kidney stained by immunofluorescence showing the spatial distribution of the PST and B cells. DAPI (blue); SLC13A3 (green), PST; MS4A1 (magenta), B cells.

https://doi.org/10.1038/s41592-026-03237-0

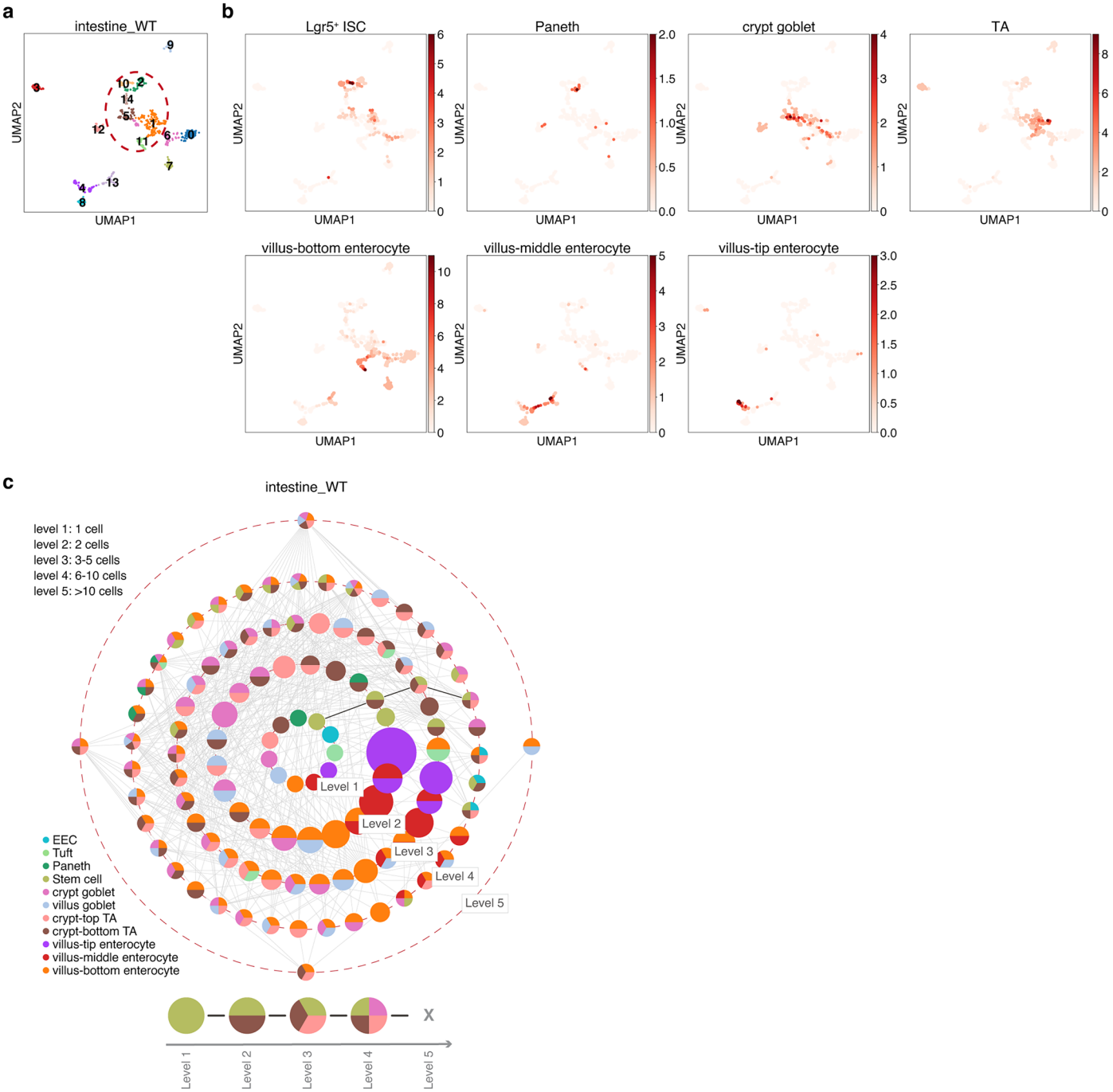


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41592-026-03237-0

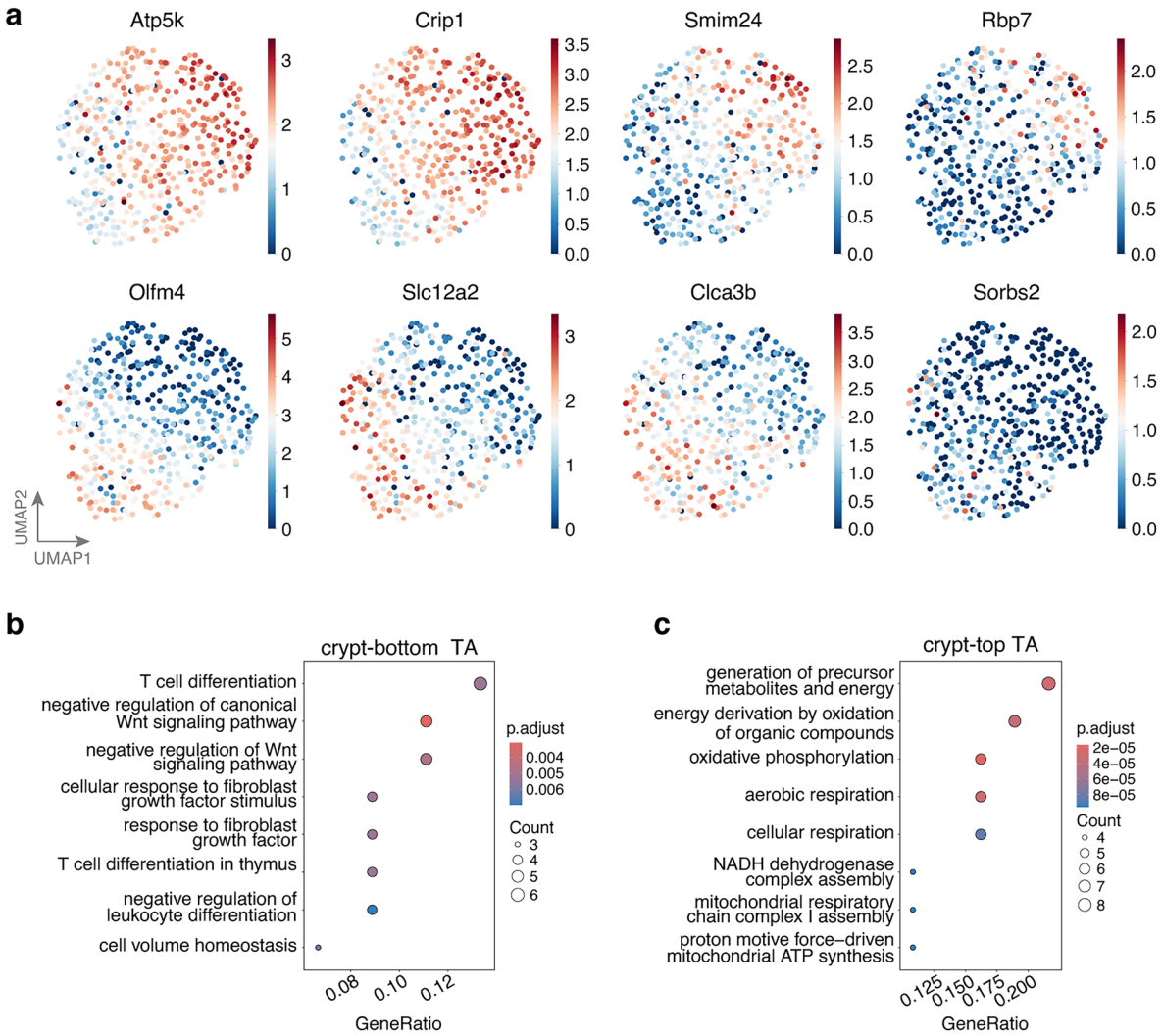
Extended Data Fig. 4 | Characterization of gene level impact of cell–cell interactions in the mouse kidney. a, UMAP visualization of podocytes from the mouse kidney dataset. Cells (dots) of podocytes are colored by their interaction partners (single-type, multi-type, and no interactions). **b**, Volcano plot showing the DEGs between DCT-interacting and DCT-non-interacting podocytes. Red and blue dots represent genes up- and down-regulated in DCT-interacting podocytes, respectively. The p-values were determined using a two-sided Student's *t*-test with Benjamini-Hochberg correction. Dashed lines indicate thresholds: adjusted p-value < 0.05, $|\log_2(\text{fold change})|>0.5$. **c**, Dotplot showing the top 8 significantly enriched GO terms for DEGs of DCT-interacting podocytes. The p-values were determined using a hypergeometric test with Benjamini-Hochberg correction. Dot size represents gene counts, and dot color represents adjusted p-value. **d**, Dotplot showing the aggregate rank of ligand–receptor scores between podocytes and endothelial cells, derived using multiple computational methods implemented by the LIANA tool. Each row represents a ligand–receptor pair. Podocyte_Endothelial represents endothelial cell-interacting podocytes. Red boxes highlight interactions discussed in the manuscript. **e**, Dotplot comparing the aggregate rank of ligand–receptor scores in CD-Trans cells interacting with CD-ICs versus those interacting with other cell types (others). The analysis was derived using the LIANA tool. CD-Trans_CD-IC represents CD-IC-interacting CD-Trans. Red box highlights interactions discussed in the manuscript. **f**, The same plot as in **e** when applied to CD-Trans and CD-PC. **g**, The same plot as in **e** when applied to monocytes and podocytes.

https://doi.org/10.1038/s41592-026-03237-0



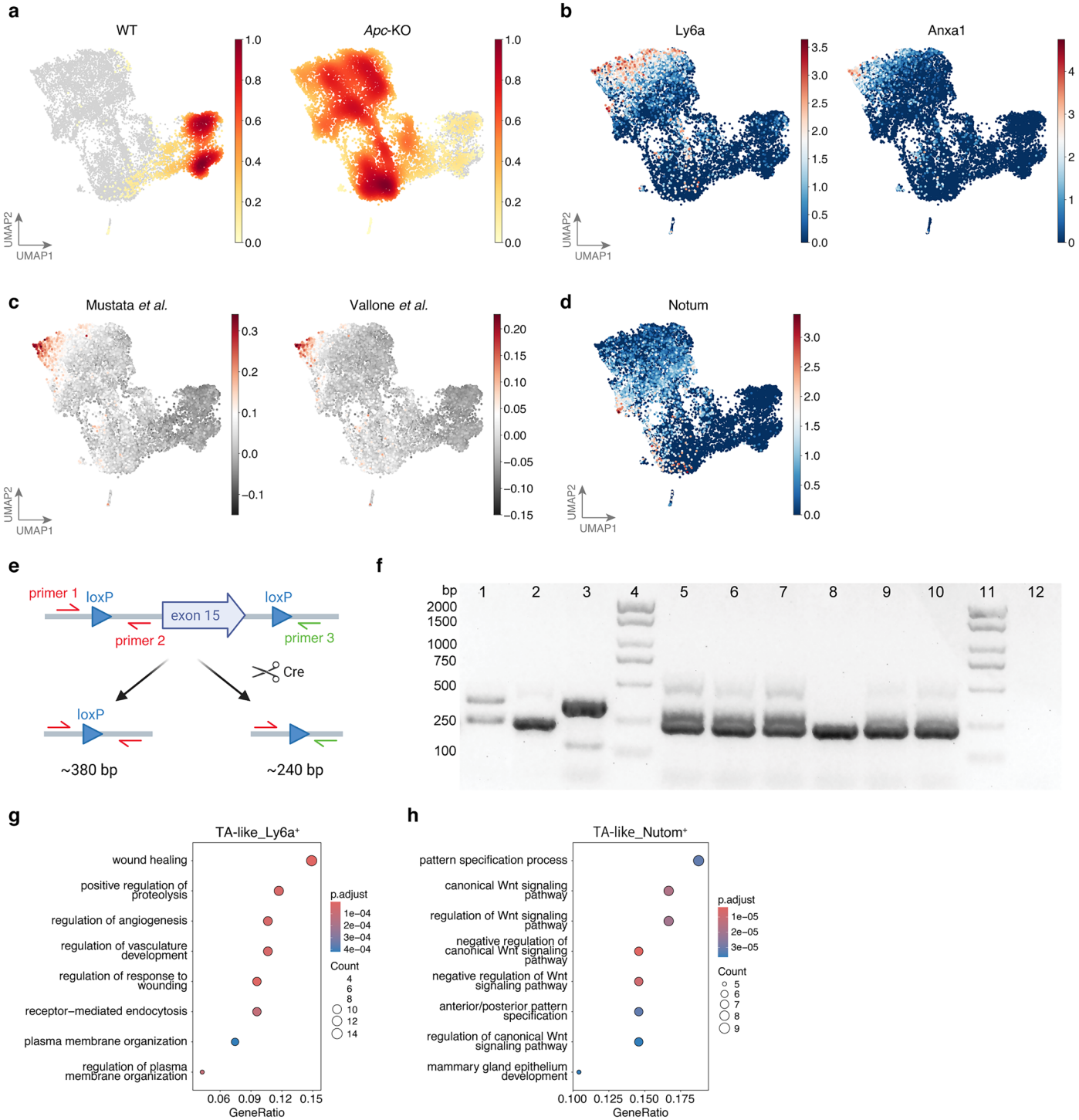
Extended Data Fig. 5 | Clump analysis for the mouse intestine. a,b, UMAP visualization of clump compositions from the mouse intestine dataset. Clumps (dots) are colored by Leiden clusters (**a**), and by the number of cells from selected cell types they contain (**b**). The red dashed-circle indicates the clumps discussed in the manuscript. **c**, The pie charts and connecting lines illustrating the clumps in the mouse intestine dataset and their hierarchical relationships. Each pie represents a group of clumps with the same cell compositions, with the pie's area indicating the number of clumps, and the size and color of each sector denoting the proportion and identity of cell types, respectively. Clumps are classified into five levels based on the number of constituent cells. Lines connecting pies across adjacent levels indicate inclusion relationships between clumps, defined by cell type compositions (*for example*, a clump composed of a–b–c includes one composed of a–b). The schematic at the bottom highlights one example of the hierarchical organization of clumps corresponding the crypt structure of the intestine.

https://doi.org/10.1038/s41592-026-03237-0



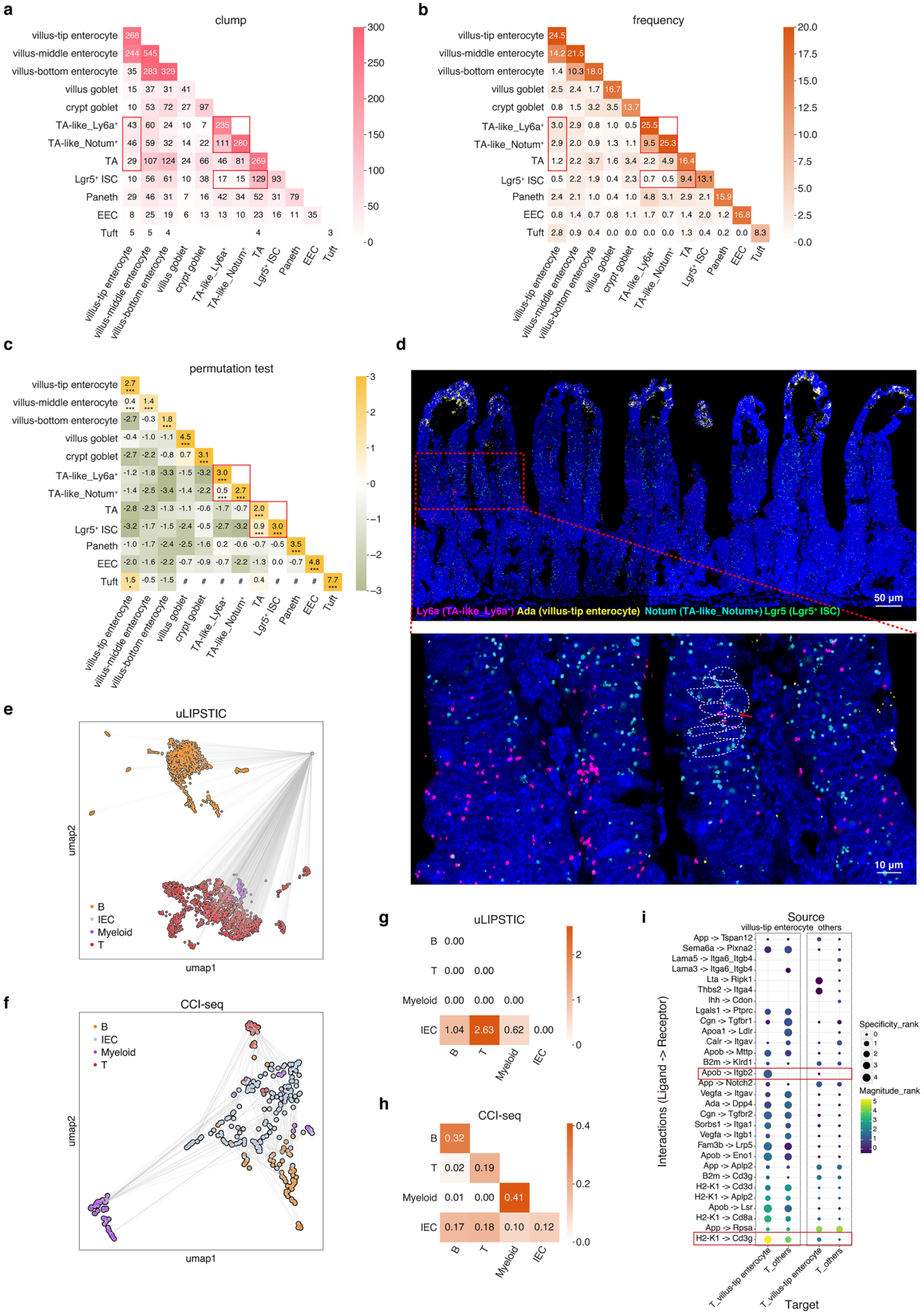
Extended Data Fig. 6 | Gene expression characteristics of TA spatial subtypes. **a**, UMAP visualization of the CCI-seq's scRNA data of TA cells interacting with other cell types. Cells are colored by the expression of selected differentially expressed genes (DEGs) of crypt-top TA (top) and crypt-bottom TA (bottom). **b,c**, Dotplot showing the top 8 significantly enriched Gene Ontology (GO) terms for DEGs of crypt-bottom TA (**b**) and crypt-top TA (**c**). The p-values were determined using a hypergeometric test with Benjamini-Hochberg correction. Dot size represents gene counts, and dot color represents adjusted p-value.

https://doi.org/10.1038/s41592-026-03237-0



Extended Data Fig. 7 | Gene expression characteristics of *Apc*-KO mouse intestinal TA/TA-like cells. a-d, UMAP visualization of the CCI-seq's scRNA data of TA/TA-like cells. Cells are colored by cell density of the WT (left) and *Apc*-KO (right) samples (**a**), the expression of *Ly6a* (left) and *Anxa1* (right) (**b**), the gene scores of two sets of fetal-like intestinal gene signatures (**c**), and the expression of *Notum* (**d**). **e**, Overview of the experiments design for single-cell genotyping. A nested three-primer PCR assay was used to genotype the *Apc* locus. In both rounds, primers 1 and 2 flank the first loxP site and amplify the wild-type allele, while primers 1 and 3 span the deletion junction and amplify the recombined allele. Nested three-primer PCR of the *Apc* locus yields ~380 bp (WT) and ~240 bp (recombined) products in the second round. **f**, Nested PCR showing the size of PCR products. Lane 4 and 11, 2,000 bp marker; lane 1, a mixture of $Apc^{KO/KO}$ ($Apc^{flox/flox}$; Villin-CreERT2 mouse with tamoxifen treatment) cells and $Apc^{flox/flox}$ ($Apc^{flox/flox}$; Villin-CreERT2 mouse with no tamoxifen treatment) cells; lane 2, $Apc^{KO/KO}$ cells; lane 3, $Apc^{flox/flox}$ cells; lane 5–10, six SCA1+EPCAM+ single cells sorted from the intestine of tamoxifen-treated $Apc^{flox/flox}$; Villin-CreERT2 mouse; lane 12, no template. The expected sizes of PCR products are ~240 bp and ~380 bp for $Apc^{KO/KO}$ and $Apc^{flox/flox}$ cell, respectively. All fragments were validated by sequencing. **g,h**, Dotplot showing the top 8 significant enriched GO terms for DEGs of TA-like_*Ly6a*$^+$ (**g**) and TA-like_*Notum*$^+$ (**h**). The p-values were determined using a hypergeometric test with Benjamini-Hochberg correction. Dot size represents gene counts, and dot color represents adjusted p-value.

https://doi.org/10.1038/s41592-026-03237-0

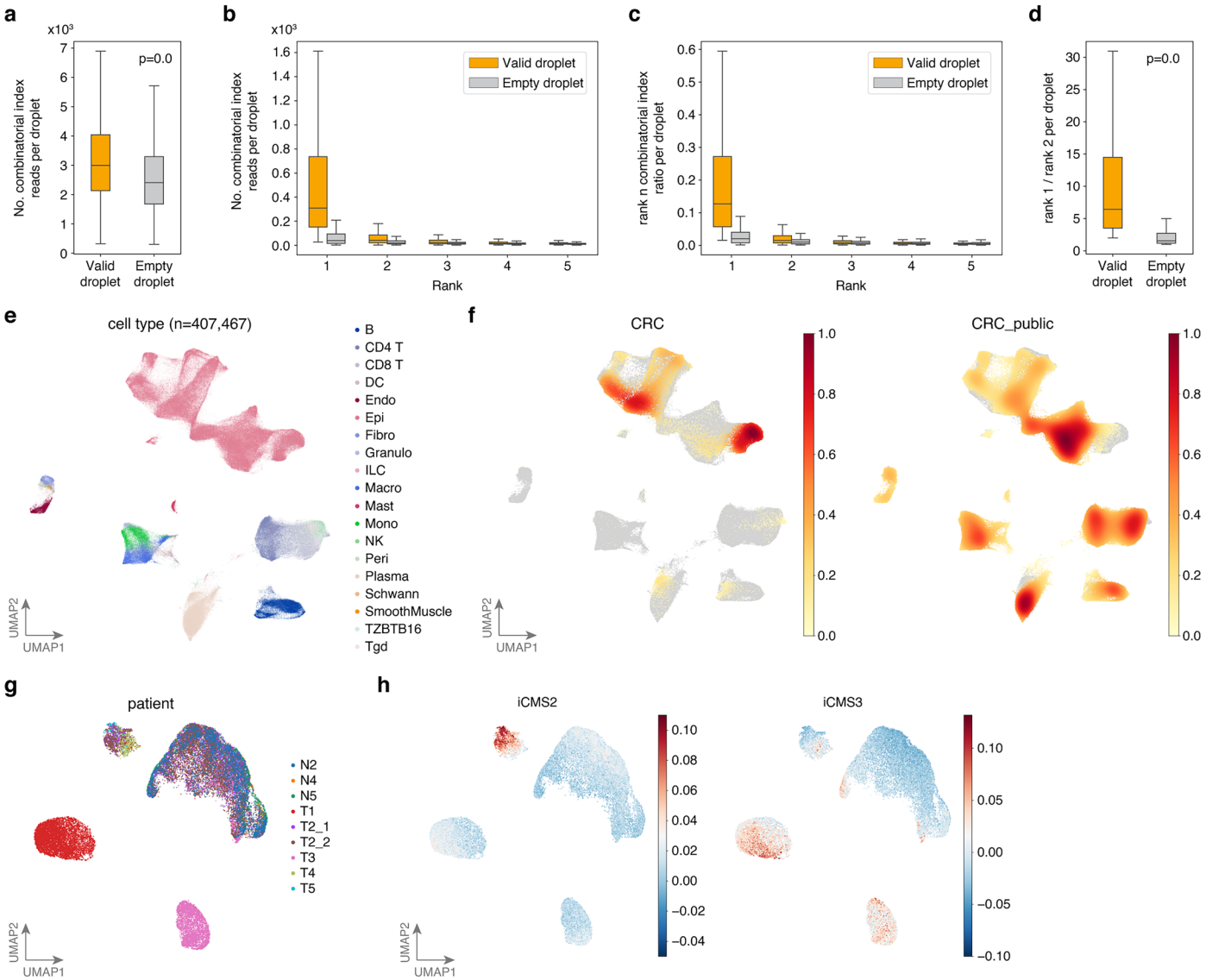


Extended Data Fig. 8 | See next page for caption.

https://doi.org/10.1038/s41592-026-03237-0

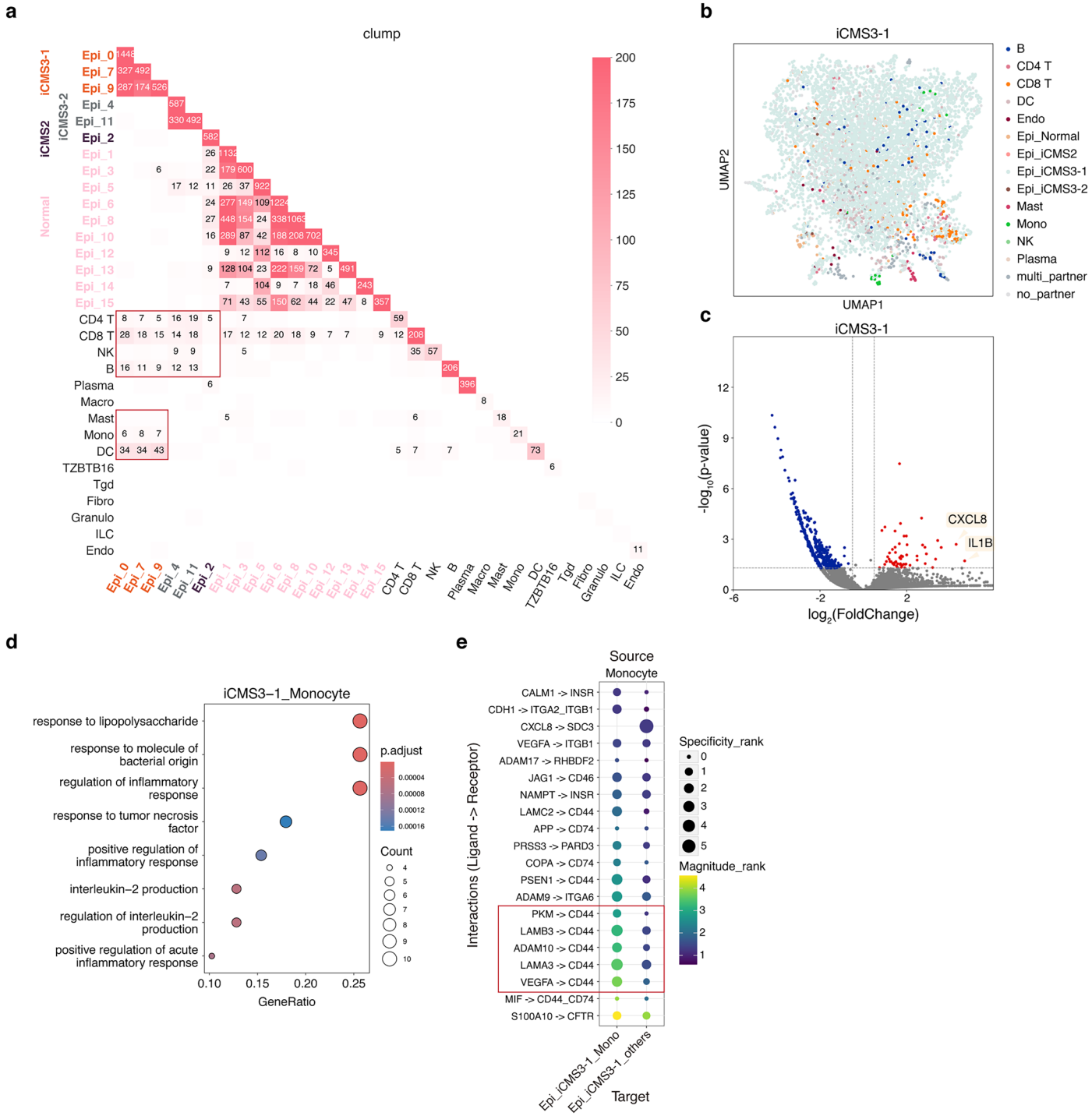
Extended Data Fig. 8 | Identification of cell–cell interactions between TA/TA-like subtypes and other cell types in the *Apc*-KO mouse intestine. a, Heatmap plot showing the cell–cell interactions quantified by *clump* (Methods) between cell types. Values below 3 are not shown. Red boxes highlight interactions discussed in the manuscript. **b**, Heatmap plot showing the interaction frequencies between cell types. Red boxes highlight interactions discussed in the manuscript. **c**, Heatmap plot showing the enrichment and depletion of interaction frequencies between cell types. Statistical analysis was conducted using a one-sided permutation test (n = 1,000). The values have been log2-transformed. *p < 0.05, **p < 0.01, ***p < 0.001. Red boxes highlight interactions discussed in the manuscript. # indicates that interacted cell type pair was detected in fewer than 3 cell clumps. **d**, AsmFISH image of the intestinal crypt–villus axis with DAPI counterstaining. *Ly6a* (magenta), TA-like_*Ly6a*$^+$; *Ada* (yellow), villus-tip enterocyte; *Notum* (cyan), TA-like_*Notum*$^+$; *Lgr5* (green), *Lgr5*$^+$ ISC. Dashed lines indicate cell borders, and red arrow indicates interactions between TA-like_*Ly6a*$^+$ and TA-like_*Notum*$^+$. **e**, The cell–cell interaction network of immune cells from uLIPSTIC mouse intestine samples. An intestinal epithelial cell (IEC) was artificially introduced as a bait cell. Cells (nodes) are colored by cell type, and gray edges represent detected cell–cell interactions. **f**, The cell–cell interaction network of immune cells from CCI-seq *Apc*-KO mouse intestine samples. Cells (nodes) are colored by cell type, and gray edges represent detected cell–cell interactions. **g**,**h**, Heatmap plots showing the interaction frequencies between cell types in mouse intestine samples detected by uLIPSTIC (**g**) and CCI-seq (**h**). **i**, Dotplot showing the aggregate rank of ligand–receptor scores between T cells and villus-tip enterocytes, derived using multiple computational methods implemented by the LIANA tool. Each row represents a ligand–receptor pair. T_villus-tip enterocytes represent villus-tip enterocyte–interacting T cells. Red box highlights interactions discussed in the manuscript.

https://doi.org/10.1038/s41592-026-03237-0



Extended Data Fig. 9 | Gene expression characteristics and CCI-seq data quality control for human CRC. a-d, Boxplot showing the following metrics within each droplet after filtering: the total number of combinatorial indices (**a**), the number of top 5 indices (**b**), the ratio of top 5 indices (**c**), and the ratio of rank 1 to rank 2 indices (**d**). Valid droplets, droplets with qualified RNAs, n = 37,897. Empty droplets, droplets without qualified RNAs, n = 664,197. Boxes, interquartile range. Midline, median. Whiskers, 1.5× interquartile range. Significance estimated using a one-sided Student's *t*-test. **e,f**, UMAP visualization of the CCI-seq's scRNA data of our human CRC dataset integrated with a public CRC dataset. Cells are colored by cell type (**e**), and cell density of our human CRC dataset (left) and the public CRC dataset (right) (**f**). Pericyte, Schwann, and Smooth Muscle were not detected in our CRC data and are therefore not shown in Fig. 5a. **g,h**, UMAP visualization of the CCI-seq's scRNA data of epithelial cells in human CRC. Cells are colored by patient (**g**), and the gene scores of iCMS2 (left) and iCMS3 (right) gene signatures (**h**). **i**, Heatmap plot showing the CNV scores of epithelial cells inferred by *inferCNV*. Columns are genes sorted based on their chromosomal position, while rows are cells clustered by their subtypes.

https://doi.org/10.1038/s41592-026-03237-0



Extended Data Fig. 10 | Cell-cell interaction analysis for human CRC. a, Heatmap plot showing the cell–cell interactions quantified by *clump* (Methods) between cell types. Values below 3 are not shown. Red boxes highlight interactions discussed in the manuscript. **b**, UMAP visualization of iCMS3-1 tumor cells from the human CRC dataset. Cells (dots) of iCMS3-1 are colored by their interaction partners (single-type, multi-type, and no interactions). **c**, Volcano plot showing the DEGs between monocyte-interacting and monocyte-non-interacting iCMS3-1 tumor cells. Red and blue dots represent genes up- and down-regulated in monocyte-interacting iCMS3-1 tumor cells, respectively. The p-values were determined using a two-sided Student's *t*-test with Benjamini-Hochberg correction. Dashed lines indicate thresholds: adjusted p-value < 0.05, |log2(fold change)|>0.5. **d**, Dotplot showing the top 8 significant enriched GO terms enriched for DEGs of monocyte-interacting iCMS3-1 tumor cells. The p-values were determined using a hypergeometric test with Benjamini-Hochberg correction. Dot size indicates gene count; color indicates adjusted p-value. **e**, Dotplot showing the aggregate rank of ligand–receptor scores between monocytes and iCMS3-1 tumor cells, derived using multiple computational methods implemented by the LIANA tool. Each row represents a ligand–receptor pair. Epi_iCMS3-1_Mono represents monocyte-interacting iCMS3-1 tumor cells. Red box highlights interactions discussed in the manuscript.

nature portfolio

Corresponding author(s): Kang Tian, Wei Wu, Changjiang Feng, Qiangfeng Cliff Zhang.

Last updated by author(s): Apr 16, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No specialized software was used for data collection
Data analysis	A detailed description of the data analysis is provided in the "Methods" section, specifically in the subsection entitled: "scRNA-seq data analysis", "Differential gene expression analysis and Gene Ontology term enrichment analysis", "Cell-ID data analysis", "Quantification of interaction strength", "Significance analysis of physical interactions", "Landmark gene score", "Definition of TA subtypes". The computational pipeline is available at https://github.com/zhangqf-lab/CCI-seq and at the Zenodo repository (https://doi.org/10.5281/zenodo.19544434).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw sequence reads and count matrices are available at GSA (PRJCA034558). Processed data are fully publicly available at Zenodo (10.5281/zenodo.19544434),

GEO (), and OMIX (OMIX016046). To comply with the Regulations of the People's Republic of China on the Administration of Human Genetic Resources, only the raw sequence reads of human patients are restricted and available upon request for scientific research to protect privacy. External datasets used are Visium HD (10x Genomics), Slide-seq-V2 (GSE190094), and seqFISH (Dryad: 10.5061/dryad.bnzs7h4hj).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Since this study focuses on technology development, neither sex nor gender was considered in the study design.
Reporting on race, ethnicity, or other socially relevant groupings	Since this study primarily focuses on technology development, race, ethnicity, and other socially relevant groupings were not considered in the study design.
Population characteristics	The study used human colon samples collected from participants aged between 40-70. We didn't seek participant with specific gender or ethnicity, no PHI information were obtained from the participants.
Recruitment	Samples were used without any biased selection.
Ethics oversight	This study was approved by the Medical Research Ethics Committee of Beijing Chao-Yang Hospital (2018-2-26-2), Capital Medical University, and written informed consent was obtained from all participants.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The study focus on methods development. No biological sample size calculations were performed.
Data exclusions	For scRNA-seq data, we excluded cells with fewer than 200 detected genes or had more than 20% mitochondrially mapped reads, and excluded genes with fewer than 3 detected cells. We also excluded cells with low Cell-ID data, details were described at "Combinatorial index data analysis" in the "Methods" section.
Replication	For the mouse experiments, we conducted the study with a minimum of three biological replicates to obtain the data.
Randomization	The experiments were not randomized.
Blinding	Clustering for each dataset were performed in an unbiased, unsupervised manner.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Anti-Ki67 Rabbit pAb (RRID: AB_3096315, Servicebio#GB111141, 1:3000); Anti-Ly-6A/E (Sca-1) antibody (RRID: AB_10862573, abcam#ab109211, 1:1000); Anti-NaDC-3 Rabbit pAb (RRID: AB_3741560, Servicebio#GB115110, 1:5000); Recombinant Anti-CD20 antibody (Rabbit mAb) (RRID: AB_3741559, Servicebio#GB155721, 1:4000); HRP conjugated Goat Anti-Rabbit IgG (H+L) (Servicebio#GB23303, 1:300); Brilliant Violet 421™ anti-mouse Ly-6A/E (Sca-1) Antibody (RRID: AB_10898327, Biolegend#108127); PE anti-mouse CD326 (Ep-CAM) Antibody (RRID: AB_1134172, Biolegend#118206).
Validation	The used antibodies are commercially available, with detailed validation provided on the vendor's website.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	All cell lines were newly purchased from Cell Bank, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. (cat #; HEK293T: SCSP-502; NIH/3T3: SCSP-515)
Authentication	Cells were authenticated at Cell Bank, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences.
Mycoplasma contamination	Cell lines were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	Species: Mus musculus (House Mouse) Strains: Wild type and Apc^flox/flox; Villin-CreERT2 Age: 6-8 weeks for Wild type and 12 weeks for Apc^flox/flox; Villin-CreERT2 Sex: Male Source: Wild-type mice were obtained from the Laboratory Animal Research Center at Tsinghua University. Apc^flox/flox; Villin-CreERT2 mice were kindly provided by Prof. Ye-Guang Chen (Tsinghua University). All animal experiments were performed with the approval of the Institutional Animal Care and Use Committee of Tsinghua University.
Wild animals	Not applicable, as the study involved only laboratory animals.
Reporting on sex	All experiments were conducted using male mice. The study's focus was on the effects of the experimental manipulations, and sex was not a variable of interest in this context. No data regarding female mice were collected, and sex-based analyses were not performed as they were not relevant to the study's objectives.
Field-collected samples	Not applicable, as the study did not involve samples collected from the field.
Ethics oversight	The study was approved by the Institutional Animal Care and Use Committee of Tsinghua University (17-ZQF1.G23-1).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*