Integration and disentanglement of single-cell and spatial transcriptomics in health and disease

Novella Rausell, C.

ABSTRACT

Single-cell and spatial transcriptomics have significantly enhanced our understanding of cellular diversity and function in complex biological systems. These technologies enable researchers to examine gene expression patterns in individual cells, revealing tissue heterogeneity that is key to understanding health and disease. This work focuses on improving single-cell and spatial transcriptomics data analysis methods and applying them to gain deeper insights into kidney biology and disease, with a particular emphasis on Autosomal Dominant Polycystic Kidney Disease (ADPKD). ADPKD is a genetic disorder characterised by the progressive formation of fluid-filled cysts in the kidneys, often leading to kidney failure.

We began by creating a comprehensive Mouse Kidney Atlas, integrating multiple single-cell RNA sequencing datasets to establish a reference map of kidney cell types. This atlas provides a hierarchical model of kidney cell populations, enhancing cell type classification and identifying robust markers for both well-known and previously overlooked cell populations. We then applied spatial transcriptomics techniques to healthy and diseased mouse kidneys, using our Mouse Kidney Atlas to improve the resolution and accuracy of cell type identification. This approach revealed an increase of specific cell types in ADPKD, including fibroblasts, injury-repair associated cells, and immune cells. Notably, we identified failed-repair proximal tubule cells exclusively in diseased kidneys, concentrated near cyst formations. By analysing cellular communication patterns in the cyst microenvironment, we uncovered key ligand-receptor interactions that may contribute to disease progression.

Recognizing the challenges in integrating diverse single-cell datasets, we designed a novel computational framework called spVIPES. This method separates shared and private factors in single-cell data, allowing for more robust integration of datasets from different conditions, species, or experimental batches. We demonstrated spVIPES' improved performance on both simulated and real-world datasets, applying it to cross-species comparisons, kidney injury models, and immune cell stimulation studies. Finally, we conducted pilot studies to optimise single-nuclei RNA sequencing for early-stage ADPKD, analysing kidney samples at various disease stages. This work provided preliminary insights into the early cellular changes and molecular pathways involved in ADPKD progression, setting the stage for more comprehensive future studies.

Collectively, these improvements in single-cell analysis methods and their application to kidney biology offer new perspectives on the cellular and molecular landscape of ADPKD.