



UNIVERSITÉ DE LILLE  
**FACULTÉ DE MÉDECINE HENRI WAREMBOURG**  
Année : 2022

THÈSE POUR LE DIPLÔME D'ÉTAT  
DE DOCTEUR EN MÉDECINE

**Transcriptomique spatiale de la voie urinaire saine. Identification et  
localisation de clusters et de sous types cellulaires.**

Présentée et soutenue publiquement le 10 Octobre 2022 à 18h  
au Pôle Formation  
par **Pierre-Emmanuel DESPREZ**

---

**JURY**

**Président :**

**Monsieur le Professeur Arnaud VILLERS**

**Assesseurs :**

**Madame le Docteur Guillemette MAROT**

**Monsieur le Docteur Jonathan OLIVIER**

**Directeur de thèse :**

**Monsieur le Docteur Gautier MARCQ**

---

## **AVERTISSEMENT**

**La Faculté n'entend donner aucune approbation aux opinions émises dans les thèses : celles-ci sont propres à leurs auteurs.**



## LISTE DES ABRÉVIATIONS

- Avg_log2FC	Average log 2 fold change
- BCL	Binary base call
- Bp	Base pair
- Bulk RNAseq	Bulk RNA sequencing
- cDNA	Complementary Desoxyribonucleic Acid
- H&E	Hematoxylin and Eosin staining
- IRB	Institutional Review Board
- mRNA	Messenger Ribonucleic Acid
- mtDNA	Mitochondrial Desoxyribonucleic Acid
- OCT	Optimal Cutting Temperature
- Pct.diff	Percentage difference
- qPCR	Quantitative Polymerase Chain Reaction
- RIN	RNA Integrity number
- scRNAseq	Single Cell RNA sequencing
- SSC	Sodium Citrate Solution
- UMI	Unique Molecular Identifier
- 10x	10x Genomics

## TABLE DES MATIÈRES

ABSTRACT.....	1
RÉSUMÉ (Traduction).....	3
INTRODUCTION.....	4
INTRODUCTION (Traduction) .....	6
MATERIALS AND METHODS.....	8
I.    Patients and tissue harvesting.....	8
II.   Visium experiment.....	9
1.  Optimization.....	9
2.  Spatial Gene Expression / Visium.....	9
III.  Sequencing.....	10
IV.  Data analysis.....	11
RESULTS.....	14
I.    Clustering results in the healthy ureter and bladder.....	14
II.   Cluster representation within the tissue and correlation with histology.....	17
III.  Identification of specific cell types based on marker genes.....	20
DISCUSSION.....	28
DISCUSSION (Traduction) .....	32
CONCLUSION.....	37
CONCLUSION (Traduction) .....	38
REFERENCES.....	39
SUPPLEMENTALS.....	41

## ABSTRACT

**Background:** RNA sequencing allows gene expression analysis, however, the spatial distribution is lost during RNA extraction. The main objective of our study is to report the spatial gene expression of the normal bladder and ureteral wall tissue layers. The secondary objective was to describe the localization of already known gene markers of urothelial mucosae, immune and stromal cell subtypes.

**Materials and methods:** Ex-vivo study of fresh human urothelial tissue samples collected from patients undergoing radical cystectomy or from organ donors. The collected tissue samples were frozen after inclusion. Each tissue sample (n=4 ureters, n=2 bladders) was studied with Spatial Gene Expression technology (Visium, 10xGenomics). Data from sequencing were analyzed by an unsupervised clustering procedure and compared to the corresponding histology using SpaceRanger, Seurat, and BayesSpace in R (v4.0.3).

**Results:** Our study identified 7 different clusters per ureter and 8 different clusters per bladder samples. The main gene markers overexpressed in both tissue types were KRT7, KRT19, SPINK1 for the urothelium, LUM, DCN, PECAM1 for the lamina propria, and ACTG2, ACTA2, MYH11 for the muscularis propria. Each cluster was coherent with the corresponding H&E tissue sample section. We identified areas where cell subtypes are overexpressed via co-expression of known gene markers.

**Conclusion:** We identified multiple gene clusters using spatial gene expression in the healthy urinary tract. These clusters were highly correlated with tissue samples layers. The Visium technology, allows the study of transcriptomics with spatial resolution and highlights a heterogeneous distribution of certain cell subtypes based on known gene markers. The study of pathological tissue is the next step.

## RÉSUMÉ

**Contexte** : Le séquençage de l'ARN permet d'analyser l'expression génique, cependant, leur distribution spatiale est perdue lors de l'extraction de l'ARN. L'objectif principal de notre étude était de rapporter l'expression spatiale des gènes des différentes couches de la paroi vésicale et urétérale saine. L'objectif secondaire était de décrire la localisation des marqueurs géniques déjà connus de sous-types de cellules urothéliales, immunitaires et stromales.

**Matériels et méthodes** : Etude ex-vivo de tissus urothéliaux humains frais prélevés chez des patients bénéficiant d'une cystectomie radicale ou chez des donneurs d'organes. Les tissus collectés ont été congelés après inclusion. Chaque tissu (n=4 uretères, n=2 vessies) a été étudié avec la technologie Spatial Gene Expression (Visium, 10xGenomics). Les données issues du séquençage ont été analysées par une procédure de clustering non supervisée et comparées à l'histologie correspondante à l'aide de SpaceRanger, Seurat et BayesSpace dans R (v4.0.3).

**Résultats** : Notre étude a identifié 7 clusters différents par uretère et 8 clusters différents par vessie. Les principaux marqueurs géniques surexprimés dans les deux types de tissus étaient KRT7, KRT19, SPINK1 pour l'urothélium, LUM, DCN, PECAM1 pour la lamina propria et ACTG2, ACTA2, MYH11 pour la muscularis propria. Chaque cluster était cohérent avec la section de tissu H&E correspondante. Nous avons identifié les zones où des sous types cellulaires sont surexprimées par la co-expression de marqueurs géniques connus.

**Conclusion** : Nous avons identifié de multiples clusters géniques en utilisant l'expression spatiale des gènes dans l'appareil urinaire sain. Ces clusters étaient fortement corrélés aux couches tissulaires. La technologie Visium, bien que limitée par

sa résolution, permet l'étude de la transcriptomique avec une résolution spatiale et met en évidence une distribution hétérogène de certains sous types cellulaires basée sur des marqueurs géniques connus. L'étude de tissus pathologiques est la prochaine étape.



## INTRODUCTION

In the past years, advances in genomics have unraveled a better understanding of human biology. Conventional RNA sequencing methods allowed the elaboration of transcriptomic atlases of many organs<sup>1 2 3</sup>, in particular the healthy<sup>3 4</sup> urinary tract. Some of these conventional RNA methods are known as bulk RNA sequencing (bulk RNAseq) or single cell level (scRNAseq). Bulk RNAseq<sup>5 6</sup> provides a global view of the transcriptomic profiles of the analyzed tissue, however, the spatial resolution is lost during RNA extraction. This problem has been partly solved thanks to scRNAseq<sup>3 7</sup>, but, even with this method, the cells are re-suspended into a solution losing again the spatial resolution.

Spatial transcriptomics techniques can solve this issue<sup>8 9</sup> by associating the transcriptomic data of a group of cells to its spatial location. These techniques may help to better understand the architectural organization of the different cell types and their interactions within the tissues at the gene level.

Other technologies such as microscopy, immunohistochemistry, and RNA in situ hybridization are also able to analyze tissue spatial resolution. However, the data provided by each technology are limited by the number of markers that can be studied at once. Recent spatial transcriptomics technologies allow overcoming this limitation by studying all expressed transcripts of the genome on a single slide. Messenger RNAs (mRNA) released from the tissue sample are captured and their localization integrated using oligonucleotides with a specific barcode, thus providing a transcriptional map of

the tissue. In this work, we will use the Spatial Gene Expression or Visum, developed by the company 10x Genomics (10x).

The principal objective of our study was to report the spatial gene expression of the normal bladder and ureteral tissue layers. The secondary objective was to describe the localization of already known gene markers of urothelial, immune and stromal cell subtypes.

## INTRODUCTION (Traduction)

Ces dernières années, les progrès de la génomique ont permis de mieux comprendre la biologie humaine. Les méthodes conventionnelles de séquençage de l'ARN ont permis l'élaboration d'atlas transcriptomiques de nombreux organes<sup>312</sup>, en particulier de l'appareil urinaire sain<sup>34</sup>. Certaines de ces méthodes conventionnelles utilisant l'ARN sont connues sous le nom de séquençage d'ARN de masse (bulk RNAseq) ou au niveau d'une cellule unique (scRNAseq). Le bulk RNAseq<sup>56</sup> fournit une vue globale des profils transcriptomiques du tissu analysé, mais la résolution spatiale est perdue pendant l'extraction de l'ARN. Ce problème a été partiellement résolu grâce au scRNAseq<sup>73</sup>, mais, même avec cette méthode, les cellules sont remises en suspension dans une solution, perdant à nouveau la résolution spatiale.

Les techniques de transcriptomique spatiale peuvent résoudre ce problème<sup>89</sup> en associant les données transcriptomiques d'un groupe de cellules à sa localisation spatiale. Ces techniques peuvent aider à mieux comprendre l'organisation architecturale des différents types cellulaires et leurs interactions au sein des tissus au niveau génique.

D'autres technologies telles que la microscopie, l'immunohistochimie et l'hybridation in situ de l'ARN sont également capables d'analyser la résolution spatiale des tissus. Cependant, les données fournies par chaque technologie sont limitées par le nombre de marqueurs qui peuvent être étudiés en même temps. Les récentes technologies de transcriptomique spatiale permettent de surmonter cette limitation en étudiant tous les transcrits exprimés sur une seule lame. Les ARN messagers (ARNm) libérés par le

tissu sont capturés et leur localisation intégrée à l'aide d'oligonucléotides possédant un code-barres spécifique, fournissant ainsi une cartographie transcriptionnelle du tissu. Dans ce travail, nous utiliserons le Spatial Gene Expression ou Visium, développé par la société 10x Genomics (10x).

L'objectif principal de notre étude était de rapporter l'expression spatiale des gènes des différentes couches du tissu vésical et urétéral sain. L'objectif secondaire était de décrire la localisation des marqueurs géniques déjà connus de sous-types de cellules urothéliales, immunitaires et stromales.

## MATERIALS AND METHODS

### I. Patients and tissue harvesting

Healthy ureteral tissues samples were collected from organ donors (n=1), or patients undergoing radical cystectomy for cancer but certified by a uropathologist as healthy (n=3). Normal bladder tissues samples were only issued from organ donors corresponding to the bladder dome (n=2), as it represents the major component of the bladder wall<sup>1</sup>. The tissue sample coming from brain-dead donors was harvested at the Cleveland Clinic, with a cold ischemia time between bladder samples harvesting and the beginning of lab manipulations shorter than 15 hours. All patients provided signed consent for the use of their tissue for research purposes. Approval was received by the Institutional Review Board (IRB) at the Cleveland Clinic, and subsequently signed informed consent from all patients was received.

All fresh samples were oriented to study their different layers (mucosa, lamina propria, muscle, and adventitia), then embedded in OCT (Optimal Temperature Cutting medium) and frozen.

The fully detailed protocol is available on the manufacturer's website at:

<https://www.10xgenomics.com/support/spatial-gene-expression-fresh-frozen>

## **II. Visium experiment**

### **1. Optimization**

As bladder and ureteral tissues samples have never been studied with the Spatial Gene Expression technology, an optimization step was needed prior to the actual Visium experiment. This issue optimization was necessary to determine the ideal permeabilization time. It corresponds to the time needed for the tissue sample to release the maximum amount of mRNA which will bind to the oligonucleotides printed on the optimization slide. A dead cell count after tissue dissociation was done for bladder tissue samples issued from organ donors to assess the cell viability. The fully detailed Optimization protocol is available on the manufacturer's website at <https://www.10xgenomics.com/support/spatial-gene-expression-fresh-frozen>. The optimization slide has 6 capture zones containing a 10 $\mu$ m tissue section, 1 positive control with mRNA, and 1 negative control with tissue. Permeabilization consisted in obtaining fluorescent complementary DNA (cDNA) from the mRNA captured in each tissue section by oligonucleotides printed on the slide, by adding a reverse transcriptase and fluorescent oligonucleotides. Different permeabilization times were defined for each capture area. The reference time was established visually using fluorescent imaging.

### **2. Spatial Gene Expression / Visium**

The Spatial Gene Expression slide has 4 capture zones containing this time printed oligonucleotides marked by a UMI (Unique Molecular Identifier) and a specific

sequence of its location (Spatial Barcode) within a 55 µm spot (Supplemental 1a). The released tissue RNA will bind to these printed oligonucleotides, which will capture their spatial localization, then cDNA and libraries was generated.

The 4 ureters and 2 bladders OCT embedded tissue samples have been cryosectioned at 10 µm of thickness and placed on the slides. After methanol fixation and H&E staining, the Visium slides were processed with all the standards required in the manufacturer protocol. The tissue sample was then permeabilized at the specific time previously defined during the optimization step and double-strand cDNA was then synthesized, amplified, and quantified for library generation. The libraries were stored at -20°C until they were sent for sequencing. All fully detailed protocols are available on the manufacturer's website at: <https://www.10xgenomics.com/support/spatial-gene-expression-fresh-frozen>

### **III. Sequencing**

The Illumina® sequencer-ready libraries were sent to the sequencing platform with an estimated depth of 50,000 reads per spot covered with tissue. We visually estimated the percentage of the capture area occupied by the tissue and we applied the following calculation:

Sequencing depth= (% coverage x 5000 spots) X 50,000 read pairs per spot.

There were 4992 spots of 55 µm diameter per capture area on the Visium slide, with within every million unique oligonucleotides sharing the same specific sequence (localization code) printed.

The sequencing was performed in an outside laboratory with the Illumina platform.

#### **IV. Data analysis:**

First of all, the mask of the capture area (which is strictly identical whatever the slide) was matched with the H&E tissue image taken before the permeabilization phase. This mapping was done as accurately as possible and felt essential to link the sequencing data to their location within the tissue. For this purpose, the Space Ranger v1.3 software developed by 10x was used with the “reorient-image” option. The algorithm allowed the automatic association of the spot map with the image of the tissue and was based on the fiducial circumscribing the capture area. Each corner of the fiducial contained a different pattern which was recognized by the software and allowed the alignment.

Thanks to this alignment, each spot was associated with the reads having the specific barcode of its spatial location. During this step, the software also identified the spots that were covered by tissue to restrict the analysis to these spots of interest. Using the grey levels of the image, the software estimated the placement of the tissue and assigned to each pixel either the tissue or background label.

Then the files from the Illumina platform sequencing that have been edited in binary base call (BCL) format were de-multiplexed to FastQ format for data analysis by Space Ranger.

Each sequenced read was aligned to a reference genome to identify it. This was also done by SpaceRanger, which uses the GRCh38-2020A reference genome and generates a summary of the integrated data for each specimen, allowing an overall assessment of their quality. The consortium used for transcriptomic annotation was GENECODE v32 (GRCh38.p13).



The data were analyzed via 2 main pipelines to define clusters, visualize them, and plot marker genes. First, the percentage of mitochondrial DNA (mtDNA) was measured per spot, then the data was normalized using SCTransform. This allowed the percentage of mtDNA in each spot to be negated using a negative binomial regression model or Poisson regression.

BayesSpace was used as the clustering method and for data visualization in this work. This method was based on the Bayes' theorem. This theorem allows computing the probability of an event by taking into account both already known information and data from new observations. Applied to Visium, this algorithm incorporates a spatial priority to encourage neighboring spots to cluster together. This improves the gene expression resolution at the spatial scale. The method can also enrich the resolution of the enhanced representation into "sub-spots", dividing each spot into 6 units and the resulting images are smoother, with reduced background noise and erratically expressed spots.

We aggregated all data per tissue type, to have the same clustering for each ureteral and vesical specimen. The "ideal" number of clusters was defined with the spatial clustering likelihood function. Once the clustering was performed, a label was put on each cluster to give them an identity in the form of a cell type. This identity was established according to the top 50-100 gene markers most expressed by these clusters. Based on these data, we then used the literature to match these top markers with a cellular identity. However, the 55 $\mu$ m spot size did not allow a single-cell resolution, especially in regions with high cell densities such as the urothelium.

The mapping of a marker gene spatial expression of interest was generated with the SpatialFeaturePlot option in Seurat v4.0.3 in R v4.0.3, which establishes a prediction score for its expression. Thus this tool allowed us to perform a mapping of the tissue sample based on gene expression, but also to search for genes of interest within the tissue.

In order to illustrate a potential application of this technology, we integrated data generated by scRNAseq performed in our lab and previously reported<sup>4</sup>: B cells (CD83, JCHAIN, IGKC, IGHA1, IGLC2), NK & T cells (GNLY, IL7R, CCL5, CXCR4, CCL4), monocytes (S100A9, LYZ, VCAN, THBS1, TIMP1), dendritic cells (HLA-DPB1, HLA-DRA, HLA-DPA1, CD74), macrophages (RNASE1, HMOX1, C1QA, C1QB, LYVE1), MAIT cells (, mast cells (TPSB2,TPSAB1, CPA3, IL1RL1, AREG), endothelial cells (TM4SF1, SPARCL1, ADAMTS1, CCL14, ACKR1), fibroblasts COL1A1-Hi (LUM, SPARCL1, A2M, MGP, C11orf96), fibroblasts APOE-Hi (APOD, MGP, MT2A, CCL2, PTX3), fibroblasts GAS1-Hi (DCN, GSN, IGFBP6, CFD, MAFP5), fibroblasts HAS1-Hi (DCN, GSN, IGFBP6, PLA2G2A, CFD), smooth muscle cells (TPM2, TAGLN, ACAT2, MYL9, C11orf96), basal cells (KRT17, S100A2, FABP4, AQP3, MIR205HG), intermediate cells (DHRS2, KRT19, CLDN4, S100P, CSTB), and umbrella cells (PSCA, KRT18,SNCG, KRT8, GDF15).

# RESULTS

Our spatial transcriptomic data reached all the minimum quality standards concerning sequencing depth, percentage of spots under tissue, or mapping to the reference genome for every specimen (Supplemental 1).

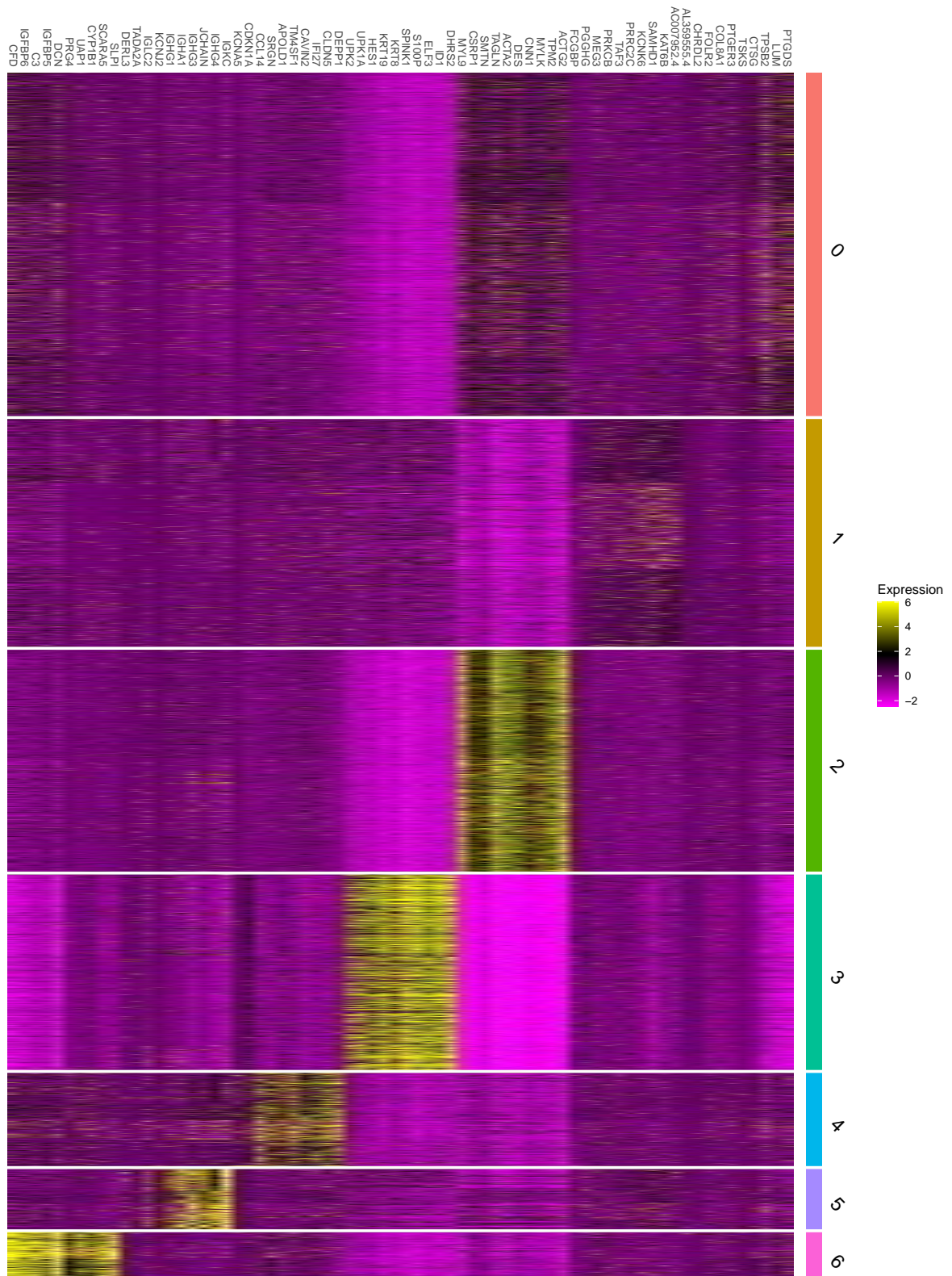
## I. **Clustering gene expression results in healthy ureter and bladder.**

A total of 7 different clusters have been identified in the ureter samples, and 8 different in the bladder specimens (Figure 1). Clusters identification and naming depend on the majority cell population constituting it or may be multiple if several strong markers of different cell types are expressed. The results presented below are based on the differential percentage of gene expression (pct. diff) of the cluster of interest compared to the rest of the clusters, but the results in log<sub>2</sub> fold change (avg\_log<sub>2</sub>FC) highlight the same cell types in total.

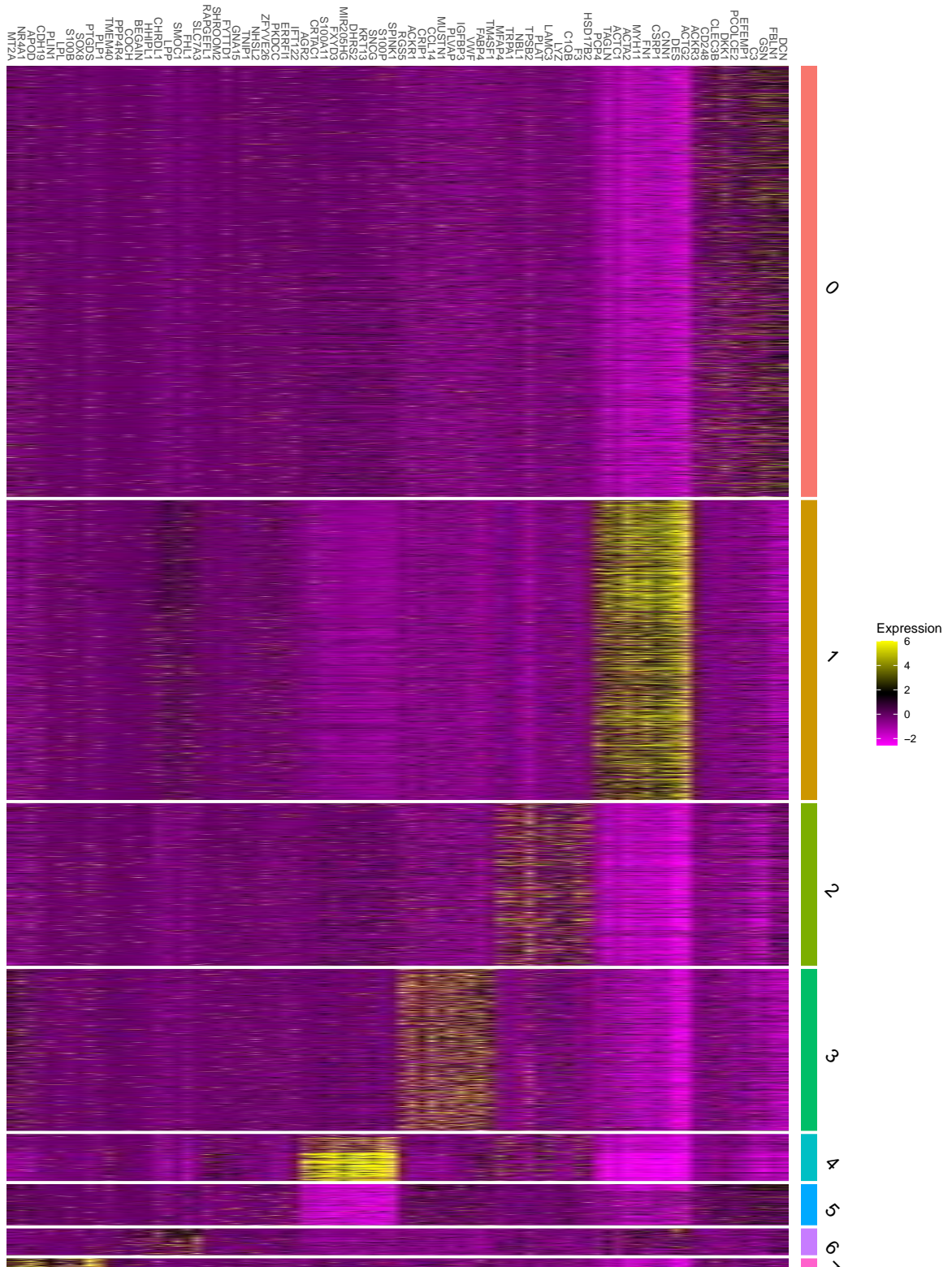
For the ureter samples, cluster 0 represents fibroblasts associated with smooth muscle cells (DCN, TAGLN, LUM, ACTA2), clusters 1, 5, and 6 represent fibroblasts and immune cells (SAMHD1, IGKC, JCHAIN), cluster 4 endothelial cells and fibroblasts (CLDN5, PECAM1), cluster 2 smooth muscle (ACTG2, MYH11, ACTA2), and cluster 3 urothelium (KRT7, KRT8, KRT19) (Figure 1a, Supplemental 2).

For the bladder samples, clusters 0, 6, and 7 are fibroblasts (DCN, COL12A1), cluster 1 smooth muscle (ACTG2, ACTA2, MYH11), cluster 2 and 5 immune cells, and fibroblasts (C1QB, LAMC3, CD163), cluster 3 endothelial and immune cells (ID1, PECAM1, CD93), and cluster 4 urothelial cells (KRT18, KRT19, KRT7) (Figure 1b, Supplemental 2).

**Figure 1: Top 10 average log<sub>2</sub>FC marker genes by cluster for ureter and bladder specimens.**



**Figure 1a: Heatmap of the top 10 average log<sub>2</sub>FC marker genes by cluster for ureter samples.**



**Figure 1b: Heatmap of the top 10 average log<sub>2</sub>FC marker genes by cluster for bladder samples.**

## **II. Cluster representation within the tissue and correlation with histology.**

For this first approach, the H&E sections were analyzed to identify each histological layer. The spatial projection of the clusters followed in a globally homogeneous way the 3 main histological layers (mucosa, lamina propria, and muscularis mucosa) for both ureter and bladder tissues (Figure 2). As described in the methods section, the analysis of the results with BayesSpace allows us to obtain an enhanced resolution at a "sub spot" level. The subspot resolution allowed, validated by the previous cluster identification, a gain in particular in the muscularis propria by separating smooth muscle and extracellular matrix. Moreover, circles erasure representing the spots allows a greater visualization finesse.

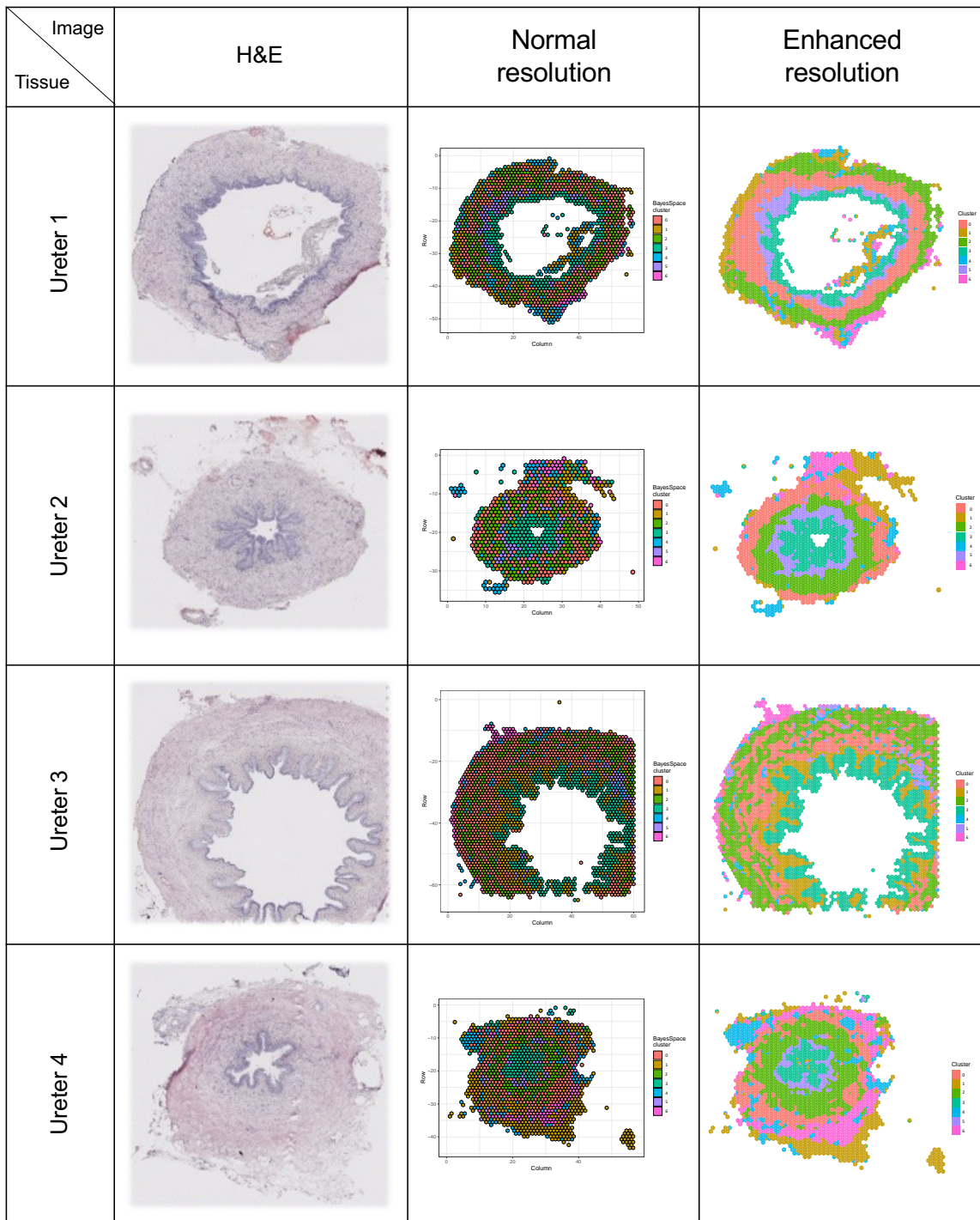
The identification of clusters allows us to establish the following organization for the ureter samples: urothelium corresponds to cluster 3, lamina propria cluster 5, muscularis propria clusters 0 and 2, stroma clusters 1 and 4, and the adventitia cluster 6 (Figure 2a).

For the bladder samples, the adventitia is not represented, but we find urothelium corresponding to cluster 4, lamina propria clusters 0, 2, and 7, muscularis propria cluster 1, and finally the stroma clusters 3, 5, and 6 (which also have a smooth muscle component, but in a minor proportion) (Figure 2b).

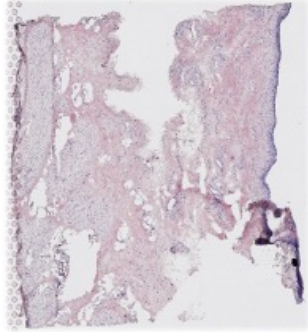
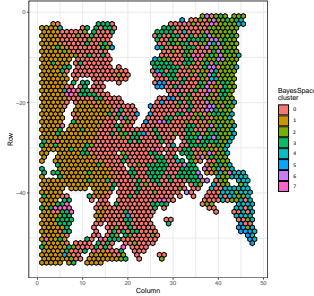
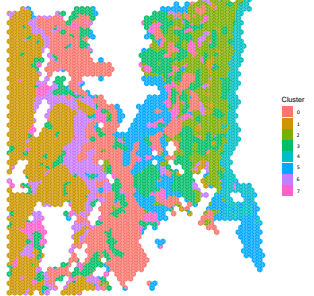

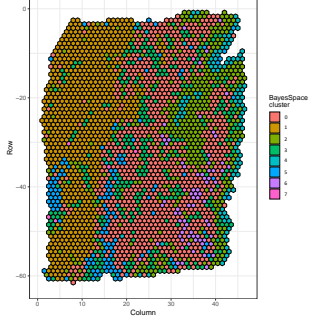
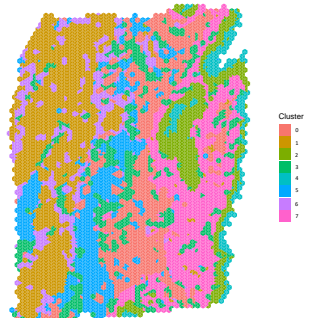
To emphasize this histological coherence on smaller clusters, there is a strong match between the endothelial/fibroblast cluster (4 for ureters and 3 for bladders) and the vessels in the H&E section.

Thus we find a spatial distribution of clusters consistent with the visual identification of the histological layers of the ureter and bladder wall.

**Figure 2: Histological coherence between H&E, BayesSpace normal and enhanced resolution of healthy ureter and bladder samples.**



**Figure 2a: H&E sections and BayesSpace normal and sub spot resolution for each ureter specimen.**

Image Tissue	H&E	Normal resolution	Enhanced resolution
Bladder 1			
Bladder 2			

**Figure 2b: H&E sections and BayesSpace normal and sub spot resolution for each bladder specimen.**



### **III. Identification of specific cell types based on marker genes.**

A spot contains the gene information of several distinct cell types, therefore reflecting the major cell populations within it. To identify specific cell types, with the known resolution limitation, we have implemented genes of interest specific to certain cell populations using the SpatialFeaturePlot function in Seurat. These genes of interest are derived from literature but also the scRNAseq database performed on the ureter within our laboratory<sup>4</sup>.

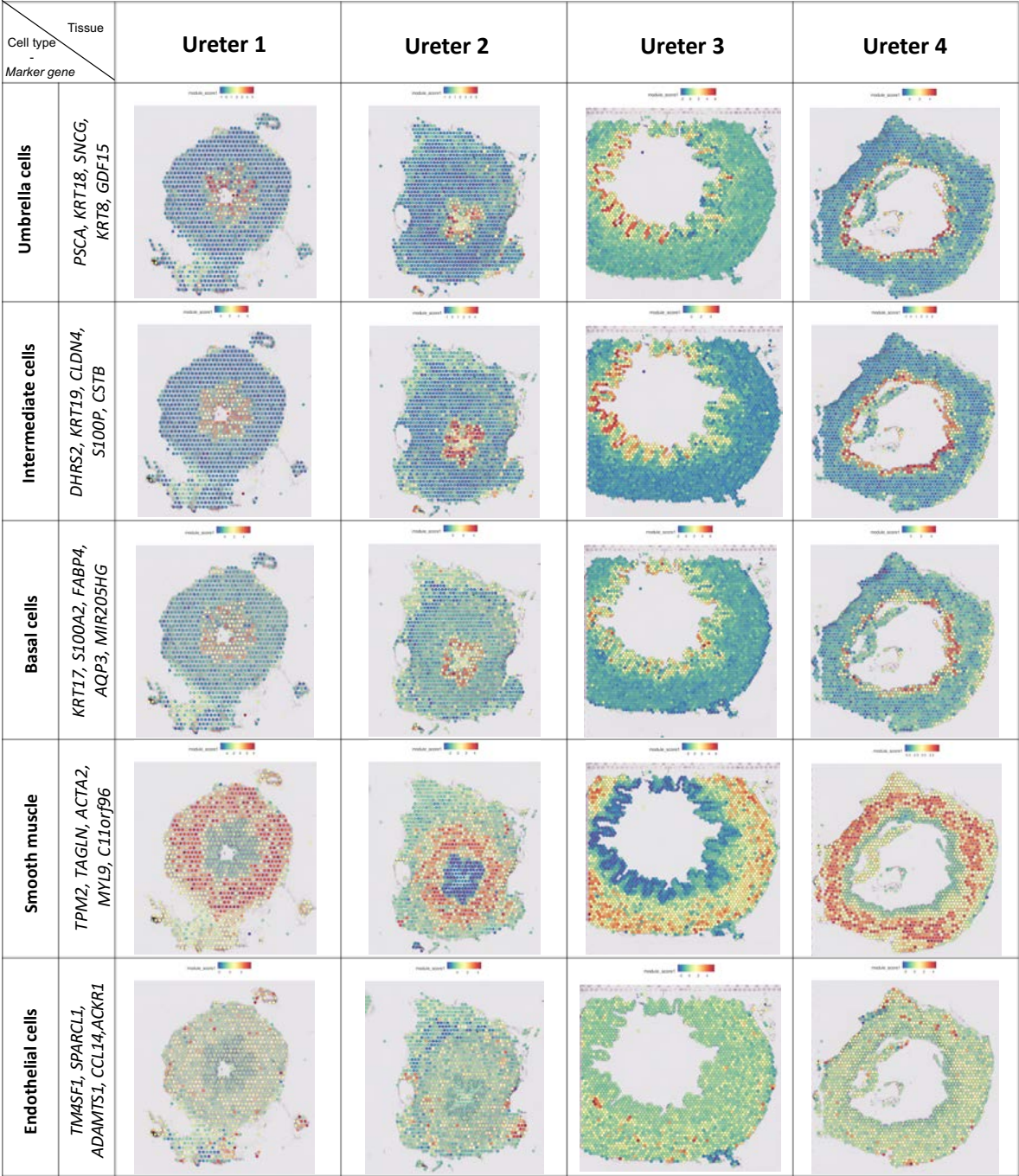
To strengthen the demonstration of coherence of transcriptomic data with histology, we first defined 5 cell types in ureteral and vesical samples according to 5 characteristic marker genes: Umbrella cells, Intermediate cells, Basal cells, smooth muscle, and endothelial cells (Figure 3).

The color scale indicates the level of gene expression within the spots and generates a result respecting the histology of the tissue. Smooth muscle cell types overlap very well with previously defined smooth muscle clusters and endothelial cell type match with vessels on H&E sections (Figure 2).

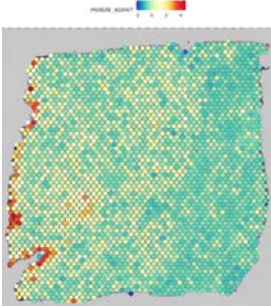

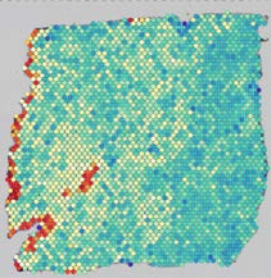
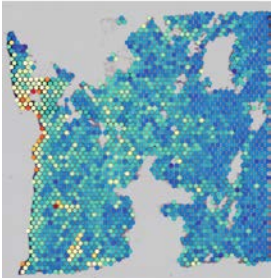
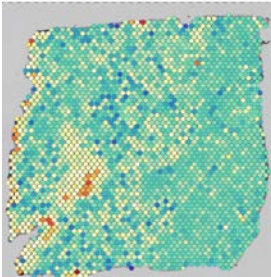
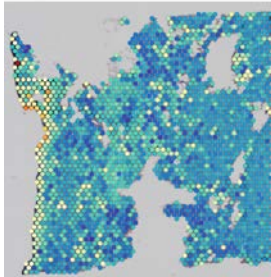
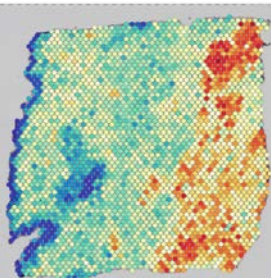
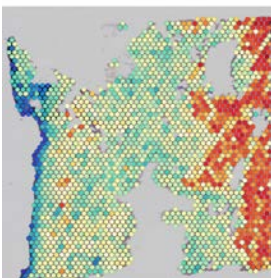
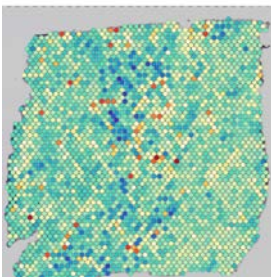

Within the urothelium itself, it is known that different cell types cohabit in a region of high cell density. We see then an overlap of urothelial cell populations (Umbrella, Intermediate, and basal cells) within the same spots. However, a more detailed view indicates a very high expression of umbrella cell marker genes on the luminal side of the urothelium.

Thus to look for gene expression levels in the tissue, this technology is interesting but does not have sufficient resolution to distinguish different cell types in regions of high cell density.

**Figure 3: Histological assessment based on the identification of 5 cell subtypes for healthy ureter and bladder samples.** The markers of interest are from the publication of Fink et al<sup>4</sup>.



**Figure 3a: Identification of cell subtypes based on specific marker genes on the ureter samples.**

Cell type - Marker gene		Tissue	
		Bladder 1	Bladder 2
<b>Umbrella cells</b>	<i>PSCA, KRT18, SNCG, KRT8, GDF15</i>		
<b>Intermediate cells</b>	<i>DHRS2, KRT19, CLDN4, S100P, CSTB</i>		
<b>Basal cells</b>	<i>KRT17, S100A2, FABP4, AQP3, MIR205HG</i>		
<b>Smooth muscle</b>	<i>TPM2, TAGLN, ACTA2, MYL9, C11orf96</i>		
<b>Endothelial cells</b>	<i>TM6SF1, SPARCL1, ADAMTS1, CCL14, ACKR1</i>		

**Figure 3b: Identification of cell subtypes based on specific marker genes on the bladder samples.**

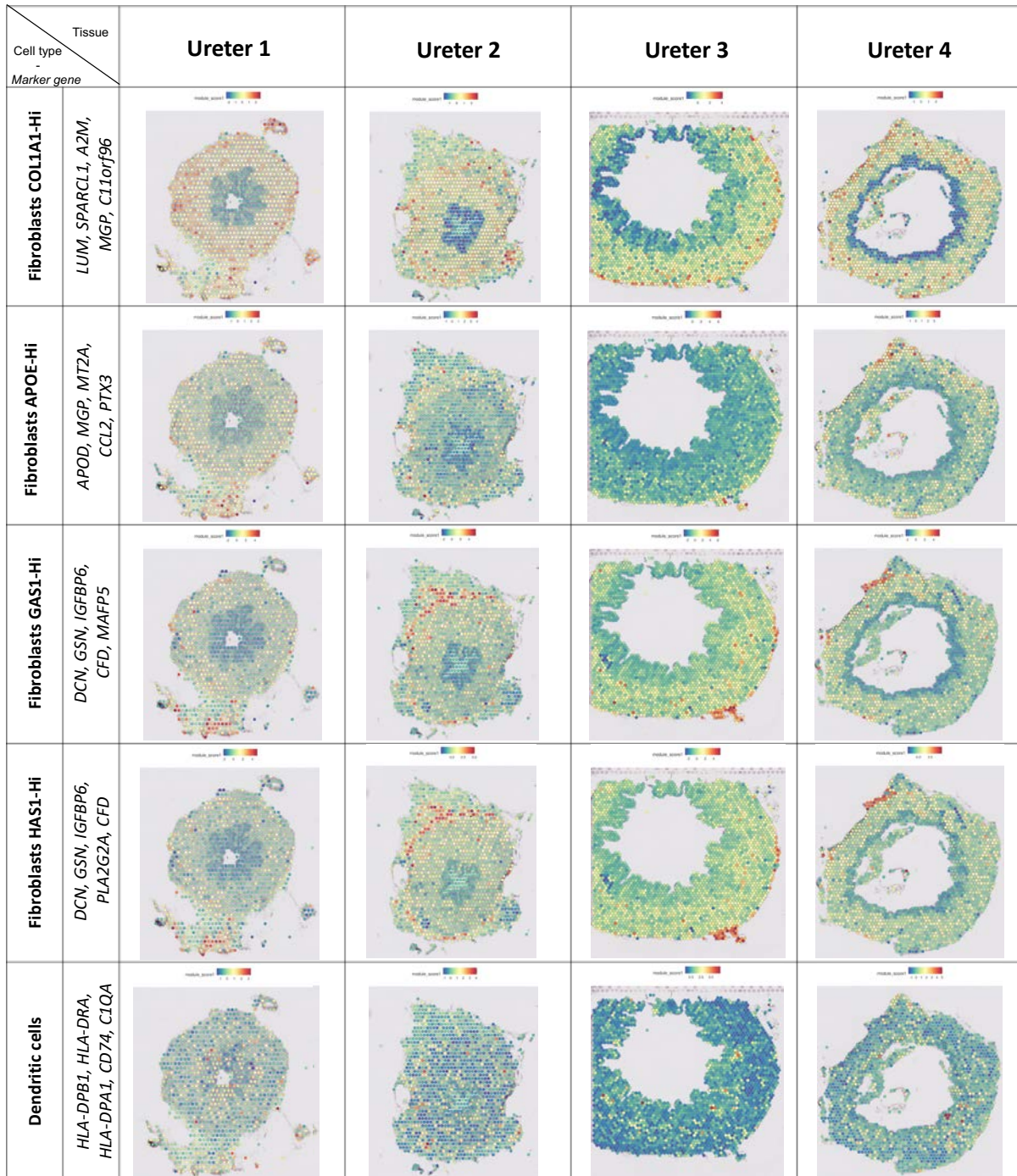
To localize even more specifically cell populations of interest, we defined 11 other cell types according to 5 characteristic gene markers (Figure 4).

Some fibroblast subpopulations also seem to have a specific organization within the tissues. HAS1-Hi and GAS1-Hi fibroblasts are located at the base of the lamina propria of the bladder tissue, and near the ureter adventice (Figure 4).

Lymphocytic immune cells (LB, LT/NK) are particularly localized in the peri urothelial and endothelial areas for the majority, whereas the other immune cell types do not seem to have a specific distribution in the tissue (Figure 4).

This spatial transcriptomics technology, therefore, allows the study of the spatial distribution of the different cell types within healthy ureteral and bladder tissue, with histological consistency. However, gene expression mapping is limited by the supra cellular resolution of the Visium.

**Figure 4: Identification of 11 cell subtypes for healthy ureter and bladder samples.** The markers of interest are from the publication of Fink et al<sup>4</sup>.



**Figure 4a: Identification of 5 cell types based on 5 marker genes on the ureter samples.**

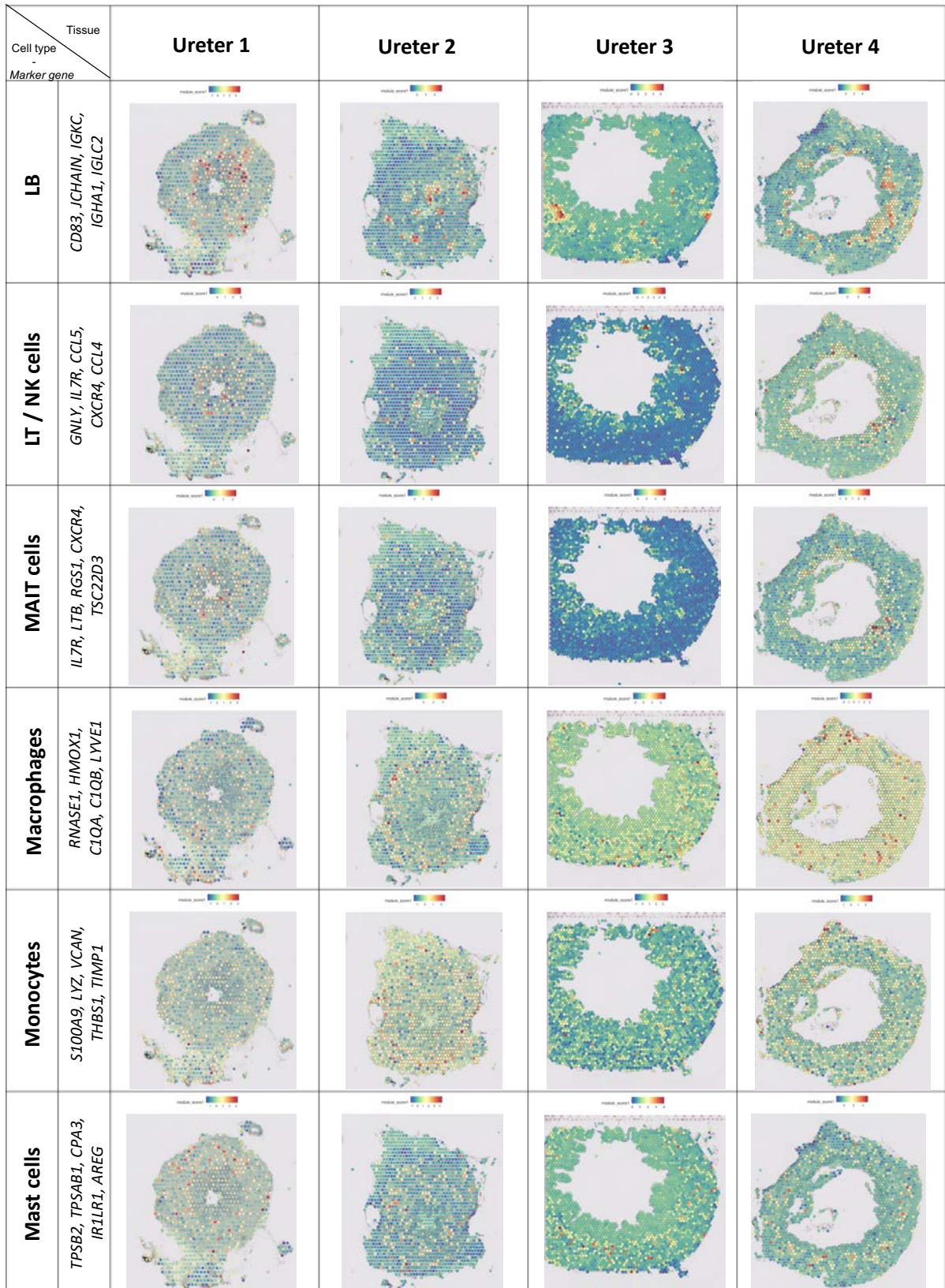
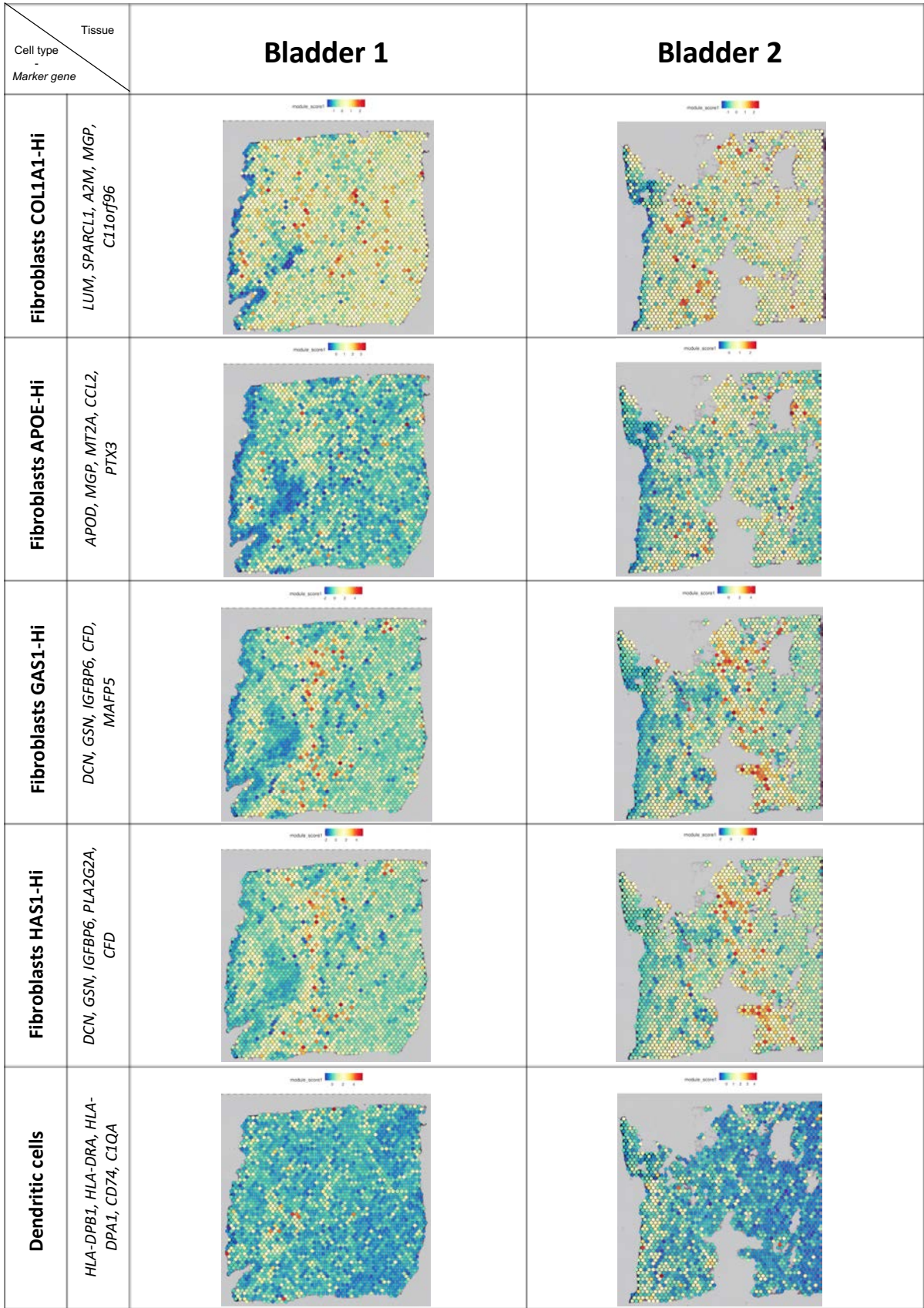


Figure 4b: Identification of 6 cell types based on 5 marker genes on the ureter samples.



**Figure 4c: Identification of 5 cell types based on 5 marker genes on the bladder samples.**

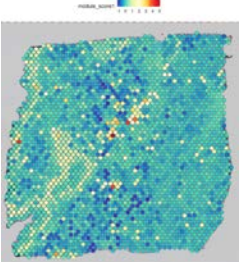
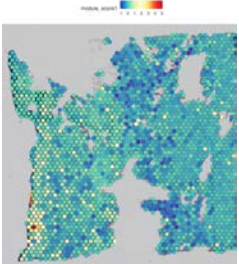
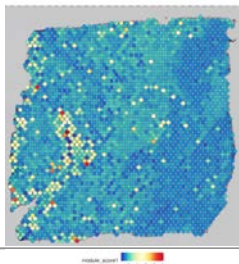
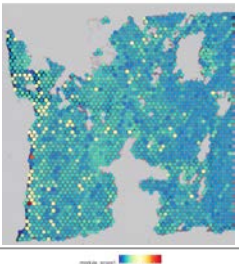
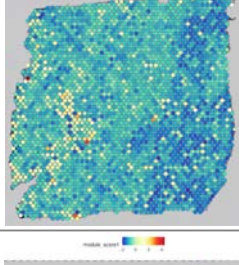
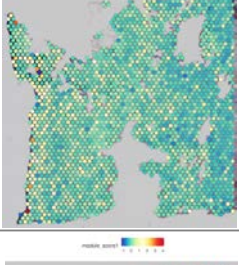
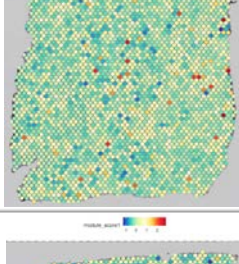

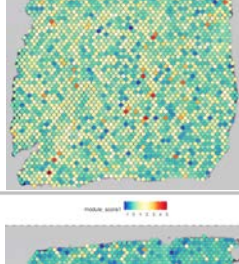
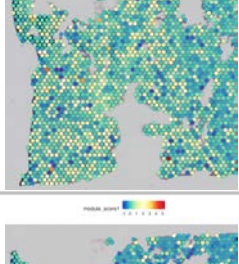
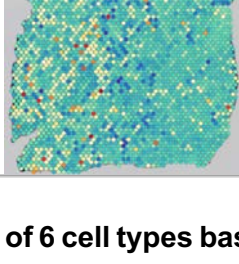
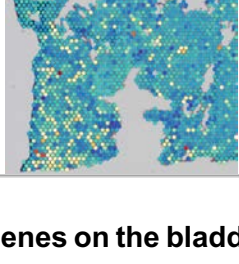
Cell type - Marker gene		Tissue	
		Bladder 1	Bladder 2
<b>LB</b>	<i>CD83, JCHAIN, IGKC, IGHA1, IGLC2</i>		
<b>LT / NK cells</b>	<i>GNLY, IL7R, CCL5, CXCR4, CCL4</i>		
<b>MAIT cells</b>	<i>IL7R, LTB, RGS1, CXCR4, TSC22D3</i>		
<b>Macrophages</b>	<i>RNASE1, HMOX1, C1QA, C1QB, LYVE1</i>		
<b>Monocytes</b>	<i>S100A9, LYZ, VCAN, THBS1, TIMP1</i>		
<b>Mast cells</b>	<i>TPSB2, TPSAB1, CPA3, IRLR1, AREG</i>		

Figure 4d: Identification of 6 cell types based on 5 marker genes on the bladder samples.



## DISCUSSION

In this study, we reported the spatial transcriptomic of the healthy ureter and bladder samples using Spatial Gene Expression. We identified 7 different gene clusters for the ureter and 8 different for the bladder samples. These clusters were highly correlated with the corresponding H&E. The technology can also identify the expression of various markers or cell types with their spatial resolution.

Elected method of the year 2020 by Nature Methods<sup>11</sup>, spatial transcriptomics technology generates a lot of interest in research teams worldwide. To our knowledge, we are the first to report spatial transcriptomics data focusing on the healthy ureter and bladder. Therefore external validation of our data cannot be performed.

Looking at our experiment data quality, we were able to make the most of this technology on these tissue samples. Summaries edited by Space Ranger for each specimen allow quality control of the sequencing data as well as the mapping (Supplemental 1). The average number of reads per spot under the tissue informs about the sequencing depth, with a quality criterion set by 10X Genomics largely achieved >40,000 reads per spot under the tissue (with an average of 110 278 reads in our study). The percentage of reads in spots under the tissue was > 80% for all specimens, which is a quality criterion according to the company. One of the reasons for this not being 100% is that during the permeabilization step, at the edge of the tissue, the released mRNAs move away and bind to oligonucleotides in an adjacent spot. The percentage of reads aligned to the reference genome was >84% for all

specimens, indicating the good quality of the library construction, sequencing process, and also the absence of major contamination.

In a parallel work<sup>4</sup>, the study of scRNA-seq data on ureteral tissue in our laboratory revealed 4 fibroblastic clusters: LAMC3hi, APOEhi, GAS1hi, and HAShi with different gene expression profiles and characteristic marker genes. A CellPhoneDB analysis<sup>12</sup> allowed studying the communication between cells through ligand-receptor interactions. Preliminary results show interactions between stromal and urothelial populations of the ureter, but more particularly between the fibroblastic cluster LAMC3hi (COL1A1hi in our work) and basal urothelial cells. This cluster appears in direct relation with the basement membrane (Figure 4), establishing anatomical coherence and reinforcing the hypothesis of the interaction of this particular cluster with the urothelial cells.

Spatial transcriptomic technologies have been used in other tissues. For example, associated with long-read sequencing, they helped to establish a spatially resolved splicing map in the mouse brain<sup>13</sup> or define the spatial topography of gene expression in the six-layered human dorsolateral prefrontal cortex (DLPFC)<sup>14</sup>. In the urological field, spatial transcriptomics allowed to identify gene expression gradients in prostate stroma adjacent to tumor regions leading to re-stratification of the tumor microenvironment<sup>15</sup>.

It is essential to picture the tissue as closely as possible to its real architecture. Despite the specimen collection realized as soon as possible after harvesting (<15h of cold ischemia), healthy bladder tissue from organ donors is known to be extremely fragile, especially the urothelium which is easily detached from its basement membrane<sup>16</sup>.

However, obtaining healthy human tissue samples is not easy and the essential information has been preserved. Moreover, the urothelium is a very fragile epithelium, and its manipulation during inclusion can induce damage and thus affect the representativeness of its organization. To counter-balance this donor tissue weakness, we did a cell viability control before the experiment which was established at <20% of dead cells.

The choice of these locations on the ureter and bladder (low ureter and bladder dome) has a consequence in the creation of an atlas. While the distribution of the different cell types is homogeneous along the ureter, it is not in the bladder<sup>10</sup>. For example, the trigone has a mesonephric embryological origin, with an over-representation of the smooth muscle layer, whereas the rest of the bladder has a mesodermal origin, even if the dome forms the major part of the bladder.

The spatial resolution of this technology is dependent on the spot size, which is 55µm in diameter for the Visium. Depending on the tissue and its cell density, this corresponds to approximately 1 to 30 cells<sup>17</sup>. Moreover, as the spots are round, part of the tissue is not covered by them at the junction between two spots. This is perfectly visible on the images linking the spot grid and H&E staining (Supplemental 1a), with a spot corresponding to about 10 cells in the urothelium, whereas it is 2 to 5 cells in the muscularis propria. Thus, the captured transcriptomic profiles may contain contributions from several cells, which induces a bias in the discovery of spatial patterns of localization and expression specific to each cell type.

While industries are working to improve this resolution by reducing the size of spots, many methods use data deconvolution algorithms specifically produced for spatial transcriptomics analyses<sup>18 19</sup>. These computational methods can exploit the genetic

profiles of known cell types from scRNAseq atlases and thus decompose the mixtures within each spot. While these methods require the input of scRNAseq data, the BayesSpace method<sup>20</sup> we used creates an improved resolution of clusters within the tissue with histological coherence and helps to draw the clusters more finely.

Based on the multiple cell types existing in a unique spot, this technology clustering gives a global view of their repartition as pools of cells within the tissue. It may be more useful with this spot resolution to tissues with less cellular density or to more homogeneous histological layers as shown by the previous work on the brain<sup>21</sup>. Indeed, it is difficult to differentiate distinct cell types in clusters if they are very dispersed or in small numbers because their marker genes won't be highlighted on the spot.

This technology also makes it possible to overcome the disadvantages of scRNAseq, which depends on the dissociation process of the tissue, and thus to highlight cell types that are absent<sup>22 23</sup>, present in small quantities, or are hard to dissociate.

More samples need to be processed to validate this cluster repartition to develop an atlas, also different locations in the bladder, even if the main limitation to this work is the access and high cost of the experiment.

A topic in the field of uro-oncology that could benefit from this technology and its improvements, is the understanding of the tumor microenvironment. It is the source of multiple research works because of its prognostic and therapeutic implication<sup>24 25 26 27</sup>

28.

## DISCUSSION (Traduction)

Dans cette étude, nous avons rapporté la transcriptomique spatiale de l'uretère et de la vessie sains en utilisant la technologie Spatial Gene Expression. Nous avons identifié 7 clusters géniques différents pour l'uretère et 8 différents pour la vessie. Ces groupes étaient fortement corrélés avec les sections H&E correspondantes. La technologie peut également identifier l'expression de divers marqueurs ou types cellulaires avec leur résolution spatiale.

Élue méthode de l'année 2020 par Nature Methods<sup>11</sup>, la technologie de transcriptomique spatiale suscite un grand intérêt auprès des équipes de recherche du monde entier. A notre connaissance, nous sommes les premiers à rapporter des données transcriptomiques spatiales centrées sur l'uretère et la vessie sains. Par conséquent, la validation externe de nos données ne peut être effectuée.

Au vu de la qualité des données de notre expérience, nous avons pu tirer le meilleur parti de cette technologie sur ces échantillons de tissus. Les résumés édités par Space Ranger pour chaque spécimen permettent un contrôle de qualité des données de séquençage ainsi que leur correspondance spatiale (Supplemental 1). Le nombre moyen de séquences par spot sous le tissu renseigne sur la profondeur de séquençage, avec un critère de qualité fixé par 10X Genomics largement atteint de >40 000 séquences par spot sous le tissu (avec une moyenne de 110 278 séquences dans notre étude). Le pourcentage de séquences dans les spots sous le tissu était > 80 % pour tous les spécimens, ce qui est un critère de qualité selon la société 10X. L'une des raisons pour lesquelles ce pourcentage n'est pas de 100 % est que pendant

l'étape de perméabilisation, au bord du tissu, les ARNm libérés s'éloignent et se lient aux oligonucléotides d'un spot adjacent. Le pourcentage de séquences alignées sur le génome de référence était >84% pour tous les spécimens, ce qui indique la bonne qualité de la construction de la bibliothèque, du processus de séquençage, et aussi l'absence de contamination majeure.

Dans un travail parallèle<sup>4</sup>, l'étude des données scRNA-seq sur le tissu urétéral dans notre laboratoire a révélé 4 clusters fibroblastiques : LAMC3hi, APOEhi, GAS1hi, et HASHi avec des profils d'expression génique différents et des marqueurs géniques caractéristiques. Une analyse CellPhoneDB<sup>12</sup> a permis d'étudier la communication inter cellulaire à travers les interactions ligand-récepteur. Les résultats préliminaires montrent des interactions entre les populations stromales et urothéliales de l'uretère, mais plus particulièrement entre le cluster fibroblastique LAMC3hi (COL1A1hi dans notre travail) et les cellules urothéliales basales. Ce cluster apparaît en relation directe avec la membrane basale (Figure 4), établissant une cohérence anatomique et renforçant l'hypothèse de l'interaction de ce cluster particulier avec les cellules urothéliales.

Les technologies transcriptomiques spatiales ont été utilisées dans d'autres tissus. Par exemple, associées au séquençage long, elles ont permis d'établir une carte d'épissage avec résolution spatiale dans le cerveau de la souris<sup>13</sup> ou de définir la topographie spatiale de l'expression génique dans le cortex préfrontal dorsolatéral humain à six couches (DLPFC)<sup>14</sup>. Dans le domaine urologique, la transcriptomique spatiale a permis d'identifier des gradients d'expression génique dans le stroma prostatique adjacent aux régions tumorales, ce qui a conduit à une re-stratification du microenvironnement tumoral<sup>15</sup>.

Il est essentiel de représenter le tissu aussi fidèlement que possible à son architecture réelle. Malgré le prélèvement réalisé le plus tôt possible (<15h d'ischémie froide), le tissu vésical sain provenant de donneurs d'organes est connu pour être extrêmement fragile, notamment l'urothélium qui se détache facilement de sa membrane basale<sup>16</sup>. Cependant, l'obtention de tissus humains sains n'est pas facile et les informations essentielles ont été préservées. De plus, l'urothélium est un épithélium très fragile, et sa manipulation lors de l'inclusion peut induire des dommages et ainsi affecter la représentativité de son organisation. Pour contrebalancer cette fragilité du tissu donneur, nous avons effectué un contrôle de viabilité cellulaire avant l'expérience qui a été établi à <20% de cellules mortes.

Le choix de ces localisations sur l'uretère et la vessie (bas d'uretère et dôme de vessie) a une conséquence dans la création d'un atlas. Alors que la distribution des différents types cellulaires est homogène le long de l'uretère, elle ne l'est pas dans la vessie<sup>10</sup>. Par exemple, le trigone a une origine embryologique mésonéphrique, avec une surreprésentation de la couche musculaire lisse, alors que le reste de la vessie a une origine mésodermique, même si le dôme constitue la majeure partie de la vessie.

La résolution spatiale de cette technologie dépend de la taille du spot, qui est de 55µm de diamètre pour le Visium. Selon le tissu et sa densité cellulaire, cela correspond à environ 1 à 30 cellules<sup>17</sup>. De plus, comme les spots sont ronds, une partie du tissu n'est pas couverte par ceux-ci à la jonction entre deux spots. Ceci est parfaitement visible sur les images reliant la grille de spots et la coloration H&E (Supplemental 1a), un spot correspondant à environ 10 cellules dans l'urothélium, alors qu'il s'agit de 2 à 5 cellules dans la muscularis propria. Ainsi, les profils transcriptomiques capturés peuvent contenir des contributions de plusieurs cellules, ce qui induit un biais dans la

découverte des modèles spatiaux de localisation et d'expression spécifiques à chaque type de cellule.

Alors que les industries s'efforcent d'améliorer cette résolution en réduisant la taille des spots, de nombreuses méthodes utilisent des algorithmes de déconvolution des données spécifiquement produits pour les analyses transcriptomiques spatiales<sup>1819</sup>. Ces méthodes informatiques peuvent exploiter les profils géniques de types cellulaires connus à partir d'atlas scRNAseq et ainsi décomposer les mélanges au sein de chaque spot. Alors que ces méthodes nécessitent l'entrée de données scRNAseq, la méthode BayesSpace<sup>20</sup> que nous avons utilisée crée une meilleure résolution des clusters au sein du tissu avec une cohérence histologique et aide à dessiner les clusters plus finement.

Alors que de multiples types cellulaires co-existent dans un même spot, cette technologie de clustering donne une vue globale de leur répartition en tant que pools de cellules dans le tissu. Elle peut être plus utile avec cette résolution pour des tissus avec une densité cellulaire moindre ou pour des couches histologiques plus homogènes comme l'ont montré les travaux précédents sur le cerveau<sup>21</sup>. En effet, il est difficile de différencier des types cellulaires distincts dans des amas, si elles sont très dispersées ou en petit nombre car leurs gènes marqueurs ne seront pas mis en évidence sur le spot.

Cette technologie permet également de pallier les inconvénients du scRNAseq, qui dépend du processus de dissociation du tissu, et ainsi de mettre en évidence des types cellulaires absents<sup>2223</sup>, présents en faible quantité, ou difficiles à dissocier.



Davantage d'échantillons doivent être traités pour valider cette répartition en clusters afin de développer un atlas, également à différents endroits de la vessie, même si la principale limitation de ce travail est l'accès et le coût élevé de l'expérience.

Un sujet dans le domaine de l'uro-oncologie qui pourrait bénéficier de cette technologie et de ses améliorations, est la compréhension du micro-environnement tumoral. Il est à l'origine de multiples travaux de recherche en raison de son implication pronostique et thérapeutique<sup>2425262728</sup>.

## CONCLUSION

Our spatial gene expression study identified 7 different clusters in the ureter and 8 different clusters in the bladder samples. The spatial distribution of these clusters reproduces well-known histological layers. This technology also enables the identification of cell types of interest based on selected gene markers but is limited by the resolution larger than a single cell. This emerging technology is promising, and future studies on pathological tissue are awaited.

## CONCLUSION (Traduction)

Notre étude de l'expression génique avec résolution spatiale a permis d'identifier 7 clusters différents dans l'uretère et 8 clusters différents dans la vessie. La distribution spatiale de ces clusters reproduit les couches histologiques déjà connues. Cette technologie permet également d'identifier des types cellulaires d'intérêt à partir de marqueurs géniques sélectionnés, mais elle est limitée par sa résolution supérieure à une cellule unique. Cette technologie émergente est prometteuse, et de futures études sur des tissus pathologiques sont attendues.

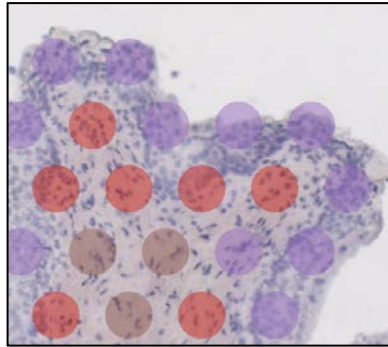
## REFERENCES

1. Wagner J, Rapsomaniki MA, Chevrier S, et al. A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell*. 2019;177(5):1330-1345.e18. doi:10.1016/j.cell.2019.03.005
2. Travaglini KJ, Nabhan AN, Penland L, et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature*. 2020;587(7835):619-625. doi:10.1038/s41586-020-2922-4
3. Yu Z, Liao J, Chen Y, et al. Single-Cell Transcriptomic Map of the Human and Mouse Bladders. *J Am Soc Nephrol*. 2019;30(11):2159-2176. doi:10.1681/ASN.2019040335
4. Fink EE, Sona S, Tran U, et al. Single-cell and spatial mapping Identify cell types and signaling Networks in the human ureter. *Developmental Cell*. 2022;57(15):1899-1916.e6. doi:10.1016/j.devcel.2022.07.004
5. Kamoun A, de Reyniès A, Allory Y, et al. A Consensus Molecular Classification of Muscle-invasive Bladder Cancer. *Eur Urol*. 2020;77(4):420-433. doi:10.1016/j.eururo.2019.09.006
6. Tan TZ, Rouanne M, Tan KT, Huang RYJ, Thierry JP. Molecular Subtypes of Urothelial Bladder Cancer: Results from a Meta-cohort Analysis of 2411 Tumors. *Eur Urol*. 2019;75(3):423-432. doi:10.1016/j.eururo.2018.08.027
7. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res*. 2014;42(14):8845-8860. doi:10.1093/nar/gku555
8. Teves JM, Won KJ. Mapping Cellular Coordinates through Advances in Spatial Transcriptomics Technology. *Mol Cells*. 2020;43(7):591-599. doi:10.14348/molcells.2020.0020
9. Moor AE, Itzkovitz S. Spatial transcriptomics: paving the way for tissue-level systems biology. *Curr Opin Biotechnol*. 2017;46:126-133. doi:10.1016/j.copbio.2017.02.004
10. Bolla SR, Odeluga N, Jetti R. Histology, Bladder. In: *StatPearls*. StatPearls Publishing; 2022. Accessed August 31, 2022. <http://www.ncbi.nlm.nih.gov/books/NBK540963/>
11. Marx V. Method of the Year: spatially resolved transcriptomics. *Nat Methods*. 2021;18(1):9-14. doi:10.1038/s41592-020-01033-y
12. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc*. 2020;15(4):1484-1506. doi:10.1038/s41596-020-0292-x
13. Joglekar A, Prjibelski A, Mahfouz A, et al. A spatially resolved brain region- and cell type-specific isoform atlas of the postnatal mouse brain. *Nat Commun*. 2021;12(1):463. doi:10.1038/s41467-020-20343-5
14. Maynard KR, Collado-Torres L, Weber LM, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat Neurosci*. 2021;24(3):425-436. doi:10.1038/s41593-020-00787-0
15. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity | Nature Communications. Accessed September 4, 2021. <https://www-nature-com.ezproxy.u-pec.fr/articles/s41467-018-04724-5>
16. Garthwaite M, Hinley J, Cross W, et al. Use of donor bladder tissues for in

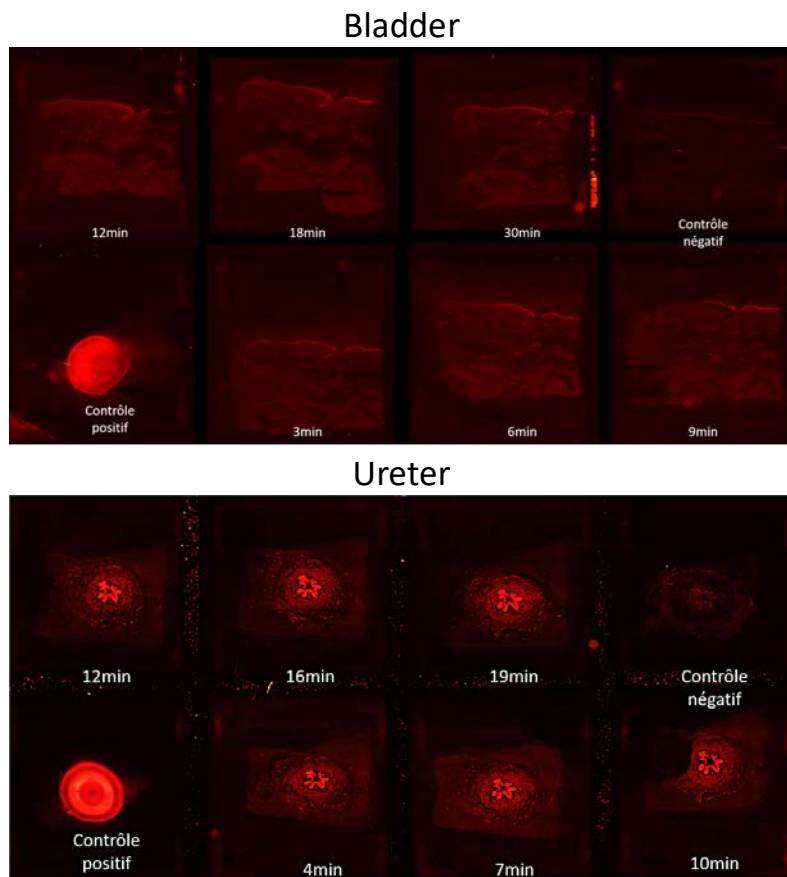
- vitro research. *BJU International*. 2014;113(1):160-166. doi:10.1111/bju.12285
17. Saiselet M, Rodrigues-Vitória J, Tourneur A, et al. Transcriptional output, cell-type densities, and normalization in spatial transcriptomics. *J Mol Cell Biol*. 2020;12(11):906-908. doi:10.1093/jmcb/mjaa028
  18. Cable DM, Murray E, Zou LS, et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat Biotechnol*. Published online February 18, 2021. doi:10.1038/s41587-021-00830-w
  19. Elosua-Bayes M, Nieto P, Mereu E, Gut I, Heyn H. SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res*. 2021;49(9):e50. doi:10.1093/nar/gkab043
  20. Zhao E, Stone MR, Ren X, et al. Spatial transcriptomics at subspot resolution with BayesSpace. *Nat Biotechnol*. Published online June 3, 2021. doi:10.1038/s41587-021-00935-2
  21. Ortiz C, Carlén M, Meletis K. Spatial Transcriptomics: Molecular Maps of the Mammalian Brain. *Annu Rev Neurosci*. 2021;44:547-562. doi:10.1146/annurev-neuro-100520-082639
  22. Akbar M, MacDonald L, Crowe LAN, et al. Single cell and spatial transcriptomics in human tendon disease indicate dysregulated immune homeostasis. *Ann Rheum Dis*. Published online May 17, 2021:annrheumdis-2021-220256. doi:10.1136/annrheumdis-2021-220256
  23. Single-cell genomics and spatial transcriptomics: Discovery of novel cell states and cellular interactions in liver physiology and disease biology - PubMed. Accessed September 6, 2021. <https://pubmed-ncbi-nlm-nih-gov.ezproxy.u-pec.fr/32534107/>
  24. Joseph M, Enting D. Immune Responses in Bladder Cancer-Role of Immune Cell Populations, Prognostic Factors and Therapeutic Implications. *Frontiers in Oncology*. 2019;9:1270. doi:10.3389/fonc.2019.01270
  25. Zhou X, Qiu S, Nie L, et al. Classification of Muscle-Invasive Bladder Cancer Based on Immunogenomic Profiling. *Front Oncol*. 2020;10:1429. doi:10.3389/fonc.2020.01429
  26. Liu Z, Zhou Q, Wang Z, et al. Intratumoral TIGIT+ CD8+ T-cell infiltration determines poor prognosis and immune evasion in patients with muscle-invasive bladder cancer. *J Immunother Cancer*. 2020;8(2):e000978. doi:10.1136/jitc-2020-000978
  27. Efstathiou JA, Mouw KW, Gibb EA, et al. Impact of Immune and Stromal Infiltration on Outcomes Following Bladder-Sparing Trimodality Therapy for Muscle-Invasive Bladder Cancer. *Eur Urol*. 2019;76(1):59-68. doi:10.1016/j.eururo.2019.01.011
  28. Necchi A, Raggi D, Gallina A, et al. Impact of Molecular Subtyping and Immune Infiltration on Pathological Response and Outcome Following Neoadjuvant Pembrolizumab in Muscle-invasive Bladder Cancer. *Eur Urol*. 2020;77(6):701-710. doi:10.1016/j.eururo.2020.02.028

# SUPPLEMENTALS

## Supplemental 1: Detailed experiment data.



## Supplemental 1a: Spatial Gene Expression spot resolution is over a single cell.



## Supplemental 1b: Ideal permeabilization time per tissue based on fluorescence imaging.

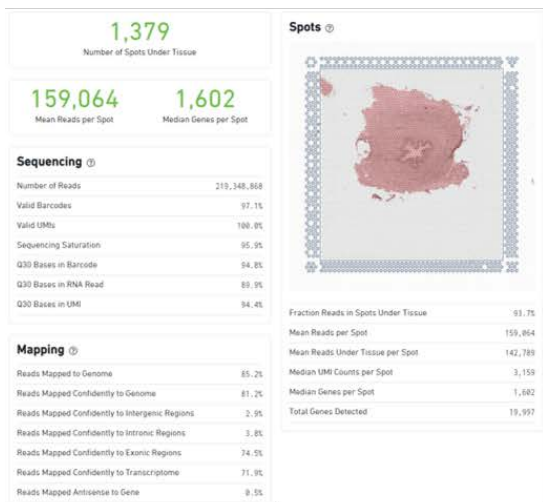
Specimen	cDNA qPCR Cq	cDNA amp cycle	CDNA QC (ng)	Library PCR cycle	Library QC (pg/uL)
Ureter 1	15,58	16	249,689	16	1748,97
Ureter 2	15,53	16	328,984	16	1247,9
Ureter 3	14,53	16	464,298	16	3082,14
Ureter 4	14,84	16	449,972	16	2116
Bladder 1	14,6	16	68,3	19	2999,69
Bladder 2	15,3	16	76,05	19	1731,67

**Supplemental 1c: Library construction data per specimen.**

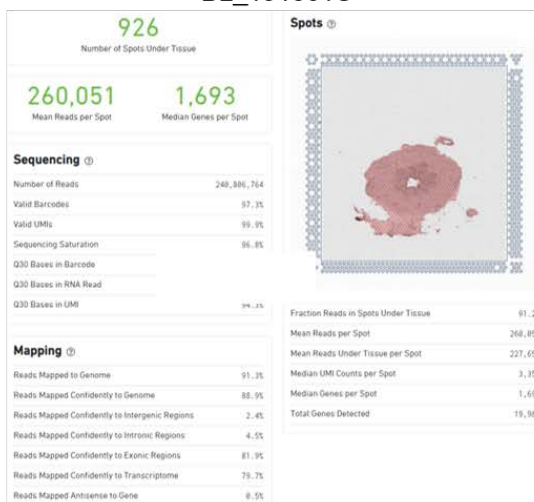
Specimen	Time (min)	Exposure Time (ms)	Gain	Ideal permeabilization time (min)
Bladder dome	3, 6, 9, 12, 18, 30	250	7	8
Ureter	4, 7, 10, 12, 16, 19	500	8	6

**Supplemental 1d: Summary quality data edited by SpaceRanger per specimen.**

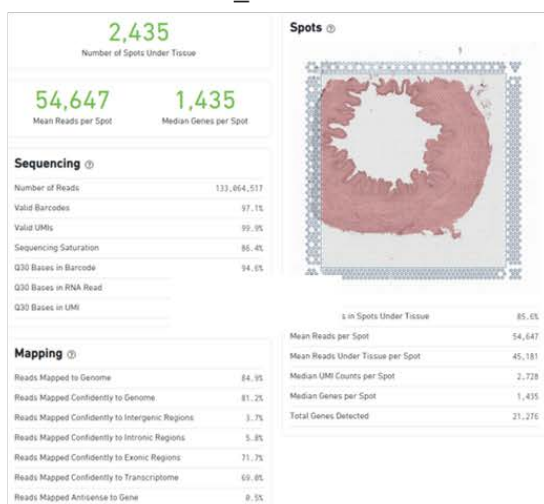
### Donor 4 Ureter



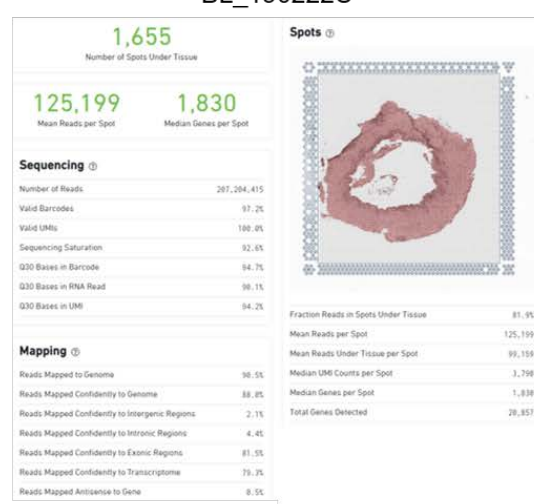
### BL\_191561U



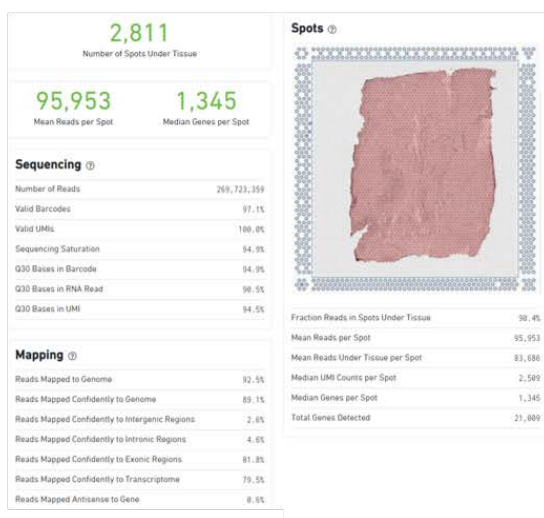
### BL\_1917129U



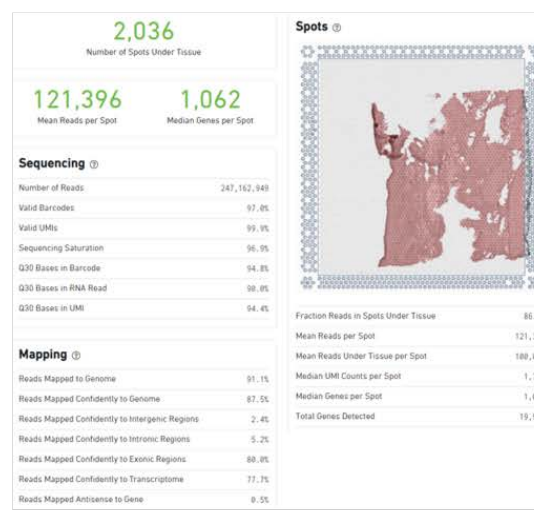
### BL\_190222U



### Donor 4 Dome



### Donor 3 Dome



Supplemental 1e: Summary quality data edited by SpaceRanger per specimen.



**Supplemental 2: Top 20 marker genes expressed in each cluster for ureter and bladder specimens.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
MGP		02.16057669	0.797	0.332	0.465
DCN	5.29E-248	-1.25303709	0.717	0.288	0.429
PTGDS	9.54E-283	3.261233716	0.728	0.301	0.427
FBLN1	5.35E-252	1.74624691	0.731	0.312	0.419
TAGLN	1.32E-143	-0.895401949	0.746	0.333	0.413
LUM	3.04E-259	3.277147068	0.693	0.299	0.394
CCDC80	3.51E-234	1.510340533	0.689	0.296	0.393
ACTA2	4.11E-138	-1.208332682	0.709	0.327	0.382
FTL	3.81E-226	2.324865098	0.717	0.339	0.378
MYL9	5.83E-123	-0.837114528	0.733	0.359	0.374
FLNA	4.15E-124	-0.64172561	0.731	0.361	0.37
MYH11	1.67E-145	-0.582128701	0.69	0.325	0.365
PI16	9.61E-201	2.49711559	0.644	0.289	0.355
C1R	5.07E-181	1.202616417	0.692	0.349	0.343
ACTG2	2.22E-121	-2.137338891	0.648	0.31	0.338
CFD	1.36E-170	-1.164914442	0.671	0.334	0.337
DPT	2.15E-149	2.274256672	0.632	0.295	0.337
TPM2	5.90E-124	-1.042126015	0.658	0.322	0.336
TPM1	6.24E-129	-0.820248594	0.672	0.337	0.335
C7	1.00E-184	2.626182848	0.687	0.358	0.329

**Supplemental 2a: Top 20 marker genes by percentage difference (pct.diff) for cluster 0 ureter.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
AC007952.4	1.54E-178	3.088562819	0.717	0.237	0.48
KAT6B	2.80E-188	3.723563513	0.671	0.212	0.459
RPL13	2.01E-159	1.574222711	0.766	0.369	0.397
SAMHD1	1.93E-176	3.987145612	0.727	0.338	0.389
TSPAN32	8.51E-253	0.771688271	0.576	0.192	0.384
MYRIP	1.96E-211	-0.865702255	0.491	0.124	0.367
KCNK6	1.53E-186	3.640172292	0.496	0.139	0.357
PRRC2C	5.92E-134	4.139664416	0.649	0.293	0.356
ARHGEF33	4.34E-165	-0.864487035	0.554	0.203	0.351
ABCA10	1.28E-199	1.342955641	0.504	0.157	0.347
RHOH	4.23E-207	0.668520163	0.562	0.218	0.344
RPL28	4.20E-113	0.686547813	0.68	0.338	0.342
TAF3	1.50E-189	3.069253178	0.509	0.174	0.335
UCP3	3.85E-105	-0.653480604	0.538	0.208	0.33
RYR2	7.20E-261	2.493711551	0.429	0.1	0.329
NUAK2	2.97E-243	-3.785396045	0.47	0.141	0.329
TMEM156	1.39E-265	-0.623404011	0.436	0.113	0.323
PRKCB	3.38E-158	3.423941336	0.482	0.16	0.322
TRAC	9.94E-176	2.491384434	0.485	0.171	0.314
KRT13	1.28E-120	-3.864308306	0.546	0.24	0.306

**Supplemental 2b: Top 20 marker genes by percentage difference (pct.diff) for cluster 1 ureter.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
ACTG2	0	5.419645408	0.996	0.273	0.723
TPM2	0	4.322681917	0.994	0.287	0.707
MYLK	0	4.687869922	0.982	0.282	0.7
CNN1	0	4.16629172	0.984	0.285	0.699
MYH11	0	3.877598281	0.997	0.301	0.696
DES	0	4.417296212	0.992	0.299	0.693
ACTA2	0	4.527280424	0.998	0.309	0.689
TPM1	0	3.843605184	0.982	0.308	0.674
TAGLN	0	4.23737351	0.997	0.327	0.67
SMTN	0	4.84516351	0.944	0.28	0.664
DSTN	0	3.194569316	0.964	0.3	0.664
CSRP1	0	4.092225302	0.984	0.326	0.658
MYL9	0	4.053254101	0.999	0.345	0.654
PCP4	0	3.830647973	0.884	0.233	0.651
FLNA	0	3.830796741	0.992	0.348	0.644
CARMN	0	3.48179502	0.895	0.257	0.638
MYL6	0	2.930892618	0.987	0.356	0.631
PLN	0	3.847906216	0.901	0.274	0.627
SYNPO2	0	3.535110304	0.913	0.297	0.616
ACTB	0	3.00363366	0.988	0.373	0.615

**Supplemental 2c: Top 20 marker genes by percentage difference (pct.diff) for cluster 2 ureter.**

gene	p_val	avg_log2FC	pct.cluster	Pct.other clusters	pct.diff
DHRS2	0	11.3728559	0.96	0.096	0.864
ID1	0	9.089439754	0.984	0.122	0.862
ELF3	0	10.04186607	0.971	0.115	0.856
S100P	0	10.89694274	0.991	0.136	0.855
SPINK1	0	8.889054432	0.98	0.152	0.828
KRT7	0	7.610454655	0.977	0.149	0.828
KRT8	0	9.045517334	0.94	0.116	0.824
KRT19	0	8.838126817	0.99	0.166	0.824
SDC1	0	7.939862345	0.951	0.139	0.812
FXYD3	0	6.693222225	0.964	0.159	0.805
PSCA	0	7.802660502	0.964	0.179	0.785
SNCG	0	7.56688486	0.951	0.168	0.783
HES1	0	8.984698886	0.935	0.153	0.782
TACSTD2	0	8.129698125	0.899	0.118	0.781
GPX2	0	6.271965341	0.931	0.153	0.778
S100A11	0	4.935460736	0.959	0.185	0.774
GDF15	0	8.532389602	0.865	0.096	0.769
S100A2	0	7.478628891	0.926	0.165	0.761
S100A6	0	4.640199856	0.978	0.218	0.76
CSTB	0	7.486616026	0.905	0.15	0.755

**Supplemental 2d: Top 20 marker genes by percentage difference (pct.diff) for cluster 3 ureter.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
DEPP1	4.92E-211	5.329961431	0.93	0.249	0.681
CLDN5	8.96E-164	5.379295236	0.854	0.204	0.65
IFI27	2.71E-156	5.029229231	0.848	0.215	0.633
CAVIN2	1.06E-160	5.21947379	0.892	0.274	0.618
TM4SF1	1.27E-138	5.734342459	0.79	0.176	0.614
S1PR1	1.52E-105	4.279297121	0.711	0.131	0.58
AL583785.1	2.54E-104	2.9762052	0.691	0.112	0.579
RAMP2	2.88E-111	3.458069627	0.768	0.192	0.576
APOLD1	1.59E-122	4.857637983	0.794	0.223	0.571
PTPRB	2.88E-104	3.377529311	0.695	0.135	0.56
SRGN	2.08E-154	4.952949106	0.858	0.299	0.559
SPRY1	3.55E-122	3.653916802	0.812	0.266	0.546
RCAN1	1.10E-106	4.075650668	0.75	0.214	0.536
HBA1	3.83E-84	1.700974387	0.705	0.173	0.532
PECAM1	9.18E-112	4.789005158	0.729	0.199	0.53
RERGL	1.78E-79	3.692130535	0.663	0.137	0.526
SOCS3	2.47E-98	3.377955993	0.745	0.223	0.522
GJA4	2.72E-80	3.536071733	0.671	0.151	0.52
NRN1	1.62E-76	3.552815297	0.669	0.152	0.517
KLF2	1.19E-146	4.7324073	0.882	0.368	0.514

**Supplemental 2e: Top 20 marker genes by percentage difference (pct.diff) for cluster 4 ureter.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
IGKC	1.27E-177	5.211842074	0.985	0.295	0.69
IGHG4	2.79E-118	5.057956833	0.861	0.256	0.605
JCHAIN	1.05E-98	3.934538631	0.772	0.192	0.58
IGHG3	6.78E-101	5.046041616	0.818	0.24	0.578
IGHA1	8.12E-100	4.48706349	0.759	0.228	0.531
PIEZO2	3.30E-79	0.741297298	0.627	0.152	0.475
AC097468.3	1.86E-85	0.817809278	0.577	0.113	0.464
TRHDE	7.63E-50	1.970812222	0.611	0.188	0.423
NDRG4	4.11E-69	0.58764399	0.54	0.14	0.4
OR51E2	1.69E-59	1.206749265	0.559	0.162	0.397
FAM27E3	5.18E-60	1.61396286	0.522	0.131	0.391
GPR146	3.38E-58	1.377460517	0.58	0.192	0.388
GDAP1	2.72E-42	2.399381803	0.556	0.17	0.386
IGHG1	3.81E-33	3.368441019	0.556	0.184	0.372
LCNL1	6.62E-30	1.096071824	0.574	0.202	0.372
C1QL1	1.47E-57	1.946173456	0.534	0.169	0.365
MYO15A	3.32E-54	1.158201194	0.494	0.129	0.365
TTN	2.67E-38	2.278490013	0.565	0.208	0.357
NR4A1	1.24E-35	1.693056009	0.769	0.412	0.357
UCN	2.34E-49	1.314658706	0.54	0.186	0.354

**Supplemental 2f: Top 20 marker genes by percentage difference (pct.diff) for cluster 5 ureter.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
SLPI	4.47E-150	8.898191926	1	0.196	0.804
SCARA5	5.21E-129	7.105702115	0.953	0.202	0.751
CYP1B1	2.88E-112	6.642676941	0.878	0.138	0.74
UAP1	4.34E-114	6.673006465	0.909	0.216	0.693
TIPARP	1.02E-113	5.86944889	0.929	0.244	0.685
PRG4	1.44E-86	5.947909331	0.772	0.093	0.679
HBA2	2.90E-106	3.656261273	0.917	0.239	0.678
SFRP4	1.66E-97	4.404075385	0.874	0.205	0.669
PODN	1.30E-97	5.53248927	0.854	0.187	0.667
C5AR2	1.40E-97	4.83029279	0.787	0.135	0.652
SHISA3	1.06E-91	4.485716129	0.776	0.13	0.646
ACKR3	1.04E-91	5.64511647	0.882	0.238	0.644
MT1A	8.02E-93	4.416791508	0.843	0.199	0.644
SFRP2	2.29E-110	3.952277304	0.965	0.323	0.642
HBA1	5.23E-87	4.185032083	0.831	0.19	0.641
EGR1	4.43E-124	3.317252971	0.992	0.354	0.638
CLEC3B	2.08E-120	5.484433009	0.98	0.352	0.628
RAMP2	2.21E-72	2.803745606	0.839	0.212	0.627
MYC	9.35E-96	4.118197952	0.87	0.244	0.626
WNT10B	2.57E-72	5.246372023	0.72	0.095	0.625

**Supplemental 2g: Top 20 marker genes by percentage difference (pt.diff) for cluster 6 ureter.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
CFD	0	3,437583	0,837	0,218	0,619
GPX3	0	3,731666	0,845	0,257	0,588
DCN	4,78E-308	4,376298	0,744	0,25	0,494
IGFBP6	8,80E-270	3,344395	0,753	0,281	0,472
FBLN1	6,79E-294	4,346055	0,706	0,252	0,454
ADH1B	3,93E-258	3,960602	0,693	0,259	0,434
AC044849.1	2,39E-226	-0,86348	0,666	0,233	0,433
MACC1	5,13E-237	-1,90901	0,61	0,196	0,414
GSN	1,47E-224	5,495897	0,691	0,283	0,408
ZNF66	0	-1,17603	0,499	0,095	0,404
RAB25	6,77E-257	-1,62368	0,553	0,153	0,4
C3	6,43E-229	4,672401	0,681	0,282	0,399
RNF39	0	-2,71973	0,495	0,096	0,399
SCARA5	1,91E-217	2,020832	0,635	0,24	0,395
ANXA8	5,54E-247	-1,23547	0,539	0,146	0,393
ITGB6	5,24E-280	2,594864	0,51	0,121	0,389
C1R	1,39E-207	3,843666	0,666	0,281	0,385
AC079140.6	0	-0,81128	0,488	0,103	0,385
TENM2	2,31E-299	-1,38338	0,487	0,107	0,38
FAM3B	5,67E-212	-2,52935	0,549	0,178	0,371

**Supplemental 2h: Top 20 marker genes by percentage difference (pct.diff) for cluster 0 bladder.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
ACTG2	0	10,40661	0,963	0,067	0,896
DES	0	7,045472	0,979	0,128	0,851
CNN1	0	9,179188	0,936	0,091	0,845
CSRP1	0	7,806259	0,94	0,111	0,829
FN1	0	9,948594	0,913	0,09	0,823
MYH11	0	6,801894	0,937	0,138	0,799
MYL9	0	5,988343	0,975	0,176	0,799
ACTA2	0	7,848402	0,932	0,138	0,794
FLNA	0	6,315028	0,96	0,168	0,792
ACTC1	0	8,455592	0,839	0,062	0,777
TPM2	0	6,73854	0,929	0,157	0,772
SMTN	0	5,492356	0,888	0,12	0,768
ACTB	0	4,6073	0,934	0,184	0,75
TAGLN	0	6,832041	0,931	0,185	0,746
TPM1	0	6,243122	0,886	0,161	0,725
PTGS1	0	5,454142	0,787	0,074	0,713
PCP4	0	8,546939	0,784	0,074	0,71
FLNC	0	5,853646	0,839	0,137	0,702
PGM5-AS1	0	5,986099	0,796	0,099	0,697
HSPB8	0	6,279384	0,788	0,113	0,675

**Supplemental 2i: Top 20 marker genes by percentage difference (pct.diff) for cluster 1 bladder.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
HSD17B2	1,58E-43	5,0246448	0,481	0,078	0,403
SELENOP	2,47E-104	3,9019016	0,726	0,328	0,398
F3	7,28E-51	5,1221254	0,513	0,123	0,39
LSP1	7,78E-43	4,712007	0,596	0,229	0,367
FTL	6,81E-96	2,9674251	0,765	0,4	0,365
C1QB	3,27E-65	5,0547006	0,653	0,3	0,353
C1QA	1,53E-75	4,4645636	0,639	0,303	0,336
LYZ	1,00E-24	4,9109056	0,481	0,148	0,333
TYROBP	9,78E-34	3,7907856	0,554	0,223	0,331
LAMC3	6,76E-17	5,0756257	0,449	0,123	0,326
PLAT	3,50E-44	4,7247137	0,497	0,171	0,326
C1QC	1,37E-39	1,8755081	0,567	0,248	0,319
FCER1G	8,35E-21	3,9494485	0,558	0,24	0,318
MRC1	1,44E-14	3,0128977	0,534	0,218	0,316
PDGFRA	4,22E-52	2,373582	0,623	0,308	0,315
TPSB2	4,61E-37	4,729844	0,557	0,244	0,313
NBL1	1,04E-72	5,6109538	0,63	0,323	0,307
TRPA1	1,39E-10	5,8138374	0,352	0,049	0,303
STAB1	2,94E-45	3,8002059	0,581	0,278	0,303
CD163	4,47E-40	4,1844297	0,561	0,26	0,301

**Supplemental 2j: Top 20 marker genes by percentage difference (pct.diff) for cluster 2 bladder.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
ID1	5,63E-231	3,670034	0,85	0,213	0,637
SOX18	2,32E-201	6,131371	0,827	0,197	0,63
NOTCH3	1,28E-161	4,843149	0,783	0,191	0,592
PECAM1	1,55E-203	6,034277	0,836	0,247	0,589
TM4SF1	1,43E-131	6,827952	0,71	0,132	0,578
IFI27	5,41E-166	4,896233	0,82	0,248	0,572
FABP4	1,18E-172	6,99195	0,759	0,207	0,552
CD93	9,41E-103	4,033688	0,704	0,155	0,549
EPAS1	5,81E-192	5,775433	0,82	0,272	0,548
TMSB10	2,23E-213	3,315315	0,886	0,341	0,545
ESAM	1,73E-88	3,802924	0,695	0,153	0,542
TIMP3	5,94E-190	5,387407	0,85	0,317	0,533
TINAGL1	1,90E-157	1,70815	0,771	0,242	0,529
HLA-B	1,58E-192	3,725222	0,842	0,316	0,526
EGFL7	3,33E-150	5,141874	0,792	0,267	0,525
CRIP2	6,75E-161	5,547635	0,803	0,28	0,523
PODXL	1,19E-61	5,110921	0,624	0,105	0,519
BCAM	2,38E-89	3,115822	0,666	0,158	0,508
SRGN	1,28E-97	4,969478	0,716	0,21	0,506
SLC9A3R2	1,79E-132	4,771136	0,729	0,225	0,504

**Supplemental 2k: Top 20 marker genes by percentage difference (pct.diff) for cluster 3 bladder.**

gene	p_val	avg_log2FC	pct.cluster	pct.other cluster	pct.diff
SPINK1	1,93E-80	8,7164869	0,839	0,242	0,597
FBP1	4,44E-68	8,4393385	0,793	0,203	0,59
C19orf33	6,48E-47	8,2951861	0,751	0,185	0,566
DMKN	2,08E-43	3,185083	0,725	0,167	0,558
PYGL	3,83E-52	4,387107	0,715	0,158	0,557
PERP	4,55E-56	7,6670851	0,751	0,198	0,553
GPX2	6,33E-53	8,4404881	0,767	0,223	0,544
TESC	5,22E-47	6,4177742	0,699	0,158	0,541
KRT18	3,94E-44	4,4984765	0,71	0,172	0,538
AKR1C2	2,08E-39	8,2017903	0,746	0,212	0,534
ID1	1,89E-54	3,4288354	0,813	0,279	0,534
SMIM22	2,25E-52	7,1784977	0,72	0,188	0,532
S100P	2,21E-57	9,0337004	0,782	0,251	0,531
KRT19	8,93E-56	8,5836249	0,772	0,241	0,531
PSCA	6,92E-63	8,3438466	0,788	0,264	0,524
FABP5	8,22E-69	5,1001708	0,767	0,244	0,523
KRT7	1,03E-54	7,7143206	0,746	0,228	0,518
SNCG	3,59E-72	9,558969	0,782	0,265	0,517
GATA3	7,44E-48	6,949985	0,715	0,199	0,516
BTG2	1,55E-46	3,3014655	0,715	0,202	0,513

**Supplemental 2l: Top 20 marker genes by percentage difference (pct.diff) for cluster 4 bladder.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
ACR	3,49E-92	-4,608237	0,953	0,065	0,888
ID2-AS1	3,52E-91	-4,772436	0,935	0,053	0,882
IL5RA	1,95E-91	-4,4065866	0,941	0,06	0,881
AL022067.1	8,93E-92	-4,3455152	0,947	0,069	0,878
AL139339.1	1,17E-88	-4,541893	0,947	0,069	0,878
TRABD2A	5,54E-90	-5,4047831	0,911	0,034	0,877
CBS	1,20E-86	-4,7204695	0,941	0,065	0,876
AC007220.1	1,20E-85	-4,8439809	0,941	0,066	0,875
TENM4	3,71E-82	-4,9920074	0,953	0,08	0,873
AL121933.2	1,55E-83	-4,7647595	0,953	0,081	0,872
FCER2	1,08E-84	-5,1615673	0,947	0,075	0,872
ISM1	2,71E-83	-4,5418906	0,953	0,084	0,869
ZNF497	1,21E-87	-4,6199828	0,935	0,067	0,868
ADCYAP1	1,10E-86	-4,6263429	0,953	0,088	0,865
AL390208.1	3,05E-90	-4,4870075	0,929	0,066	0,863
FNDC1	4,34E-82	-5,0185374	0,929	0,067	0,862
PGAM2	1,50E-82	-4,7204016	0,941	0,08	0,861
ADAMTS14	2,42E-83	-4,4223547	0,953	0,093	0,86
MIR155HG	1,06E-85	-4,3844845	0,929	0,072	0,857
CABP4	9,71E-84	-4,9633525	0,953	0,096	0,857

**Supplemental 2m: Top 20 marker genes by percentage difference (pct.diff) for cluster 5 bladder.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
CLDN11	5,20E-22	2,9892214	0,718	0,193	0,525
PDLIM3	1,28E-15	0,8570743	0,764	0,28	0,484
SLC7A3	5,55E-12	6,0579163	0,536	0,062	0,474
FHL1	1,31E-27	4,260086	0,827	0,358	0,469
PGM5	6,58E-17	-0,6245914	0,755	0,291	0,464
TAGLN	1,04E-17	-2,0564776	0,818	0,364	0,454
PRELP	1,87E-30	1,6984525	0,882	0,434	0,448
COL12A1	2,03E-14	3,171296	0,645	0,206	0,439
NEXN	8,46E-12	0,5098446	0,673	0,259	0,414
WNT2B	1,99E-18	2,1827918	0,618	0,205	0,413
IGFBP5	1,66E-18	1,0019355	0,8	0,387	0,413
CCN2	1,52E-16	-2,0007275	0,745	0,333	0,412
TNC	7,67E-14	3,3804578	0,609	0,198	0,411
AEBP1	1,34E-23	1,8166986	0,836	0,425	0,411
RGS2	1,89E-13	-1,008985	0,673	0,262	0,411
MYH11	3,71E-16	-1,126967	0,736	0,332	0,404
DES	7,94E-16	-1,405811	0,736	0,335	0,401
MYLK	2,73E-14	-0,5892022	0,736	0,344	0,392
SMOC1	1,14E-09	4,016003	0,445	0,06	0,385
ACTA2	3,16E-16	-2,9863505	0,709	0,331	0,378

**Supplemental 2n: Top 20 marker genes by percentage difference (pct.diff) for cluster 6 bladder.**

gene	p_val	avg_log2FC	pct.cluster	pct.other clusters	pct.diff
PLP1	4,29E-20	5,4078055	0,923	0,111	0,812
TMEM132B	8,58E-19	4,5423427	0,846	0,051	0,795
C5orf22	6,69E-17	-1,383282	0,872	0,097	0,775
PTGDS	3,95E-20	5,9539713	0,923	0,15	0,773
ZXDA	5,07E-16	-4,9761842	0,795	0,028	0,767
AFG1L	3,68E-17	1,4017048	0,821	0,056	0,765
FOXRED1	7,83E-18	0,3225967	0,846	0,093	0,753
KCNAB1	1,35E-14	-2,8865229	0,821	0,077	0,744
ENPEP	3,00E-16	3,2099793	0,846	0,103	0,743
CAMK2A	8,17E-17	3,5097825	0,795	0,058	0,737
SOX8	2,00E-14	5,5064673	0,795	0,062	0,733
HAGHL	6,49E-15	2,7662875	0,795	0,063	0,732
S100B	2,11E-18	5,3059762	0,821	0,091	0,73
EPHA5	1,03E-16	4,0195292	0,795	0,066	0,729
NRXN1	2,01E-14	2,7135831	0,795	0,071	0,724
SLC25A5-AS1	1,41E-14	3,6350097	0,769	0,046	0,723
TTC9C	5,81E-14	-1,0435652	0,821	0,099	0,722
TIGD1	1,00E-16	3,1683407	0,821	0,1	0,721
EVC2	2,59E-16	0,5705624	0,821	0,102	0,719
CYP7B1	6,26E-15	-1,6265698	0,821	0,103	0,718

**Supplemental 2o: Top 20 marker genes by percentage difference (pct.diff) for cluster 7 bladder.**



**AUTEUR : Nom :** DESPREZ

**Prénom :** Pierre-Emmanuel

**Date de soutenance :** 10 Octobre 2022

**Titre de la thèse :** Transcriptomique spatiale de la voie urinaire saine. Identification et localisation de clusters et de sous types cellulaires.

**Thèse - Médecine - Lille 2022**

**Cadre de classement :** Fondamental – Voies Urinaires

**DES + FST/option :** Urologie

**Mots-clés :** Transcriptomique spatiale, expression génique, voies urinaires saines, uretère, vessie.

**Résumé :**

**Contexte :** Le séquençage de l'ARN permet d'analyser l'expression génique, cependant, leur distribution spatiale est perdue lors de l'extraction de l'ARN. L'objectif principal de notre étude était de rapporter l'expression spatiale des gènes des différentes couches de la paroi vésicale et urétérale saine. L'objectif secondaire était de décrire la localisation des marqueurs géniques déjà connus de sous-types de cellules urothéliales, immunitaires et stromales.

**Matériels et méthodes :** Etude ex-vivo de tissus urothéliaux humains frais prélevés chez des patients bénéficiant d'une cystectomie radicale ou chez des donneurs d'organes. Les tissus collectés ont été congelés après inclusion. Chaque tissu (n=4 uretères, n=2 vessies) a été étudié avec la technologie Spatial Gene Expression (Visium, 10xGenomics). Les données issues du séquençage ont été analysées par une procédure de clustering non supervisée et comparées à l'histologie correspondante à l'aide de SpaceRanger, Seurat et BayesSpace dans R (v4.0.3).

**Résultats :** Notre étude a identifié 7 clusters différents par uretère et 8 clusters différents par vessie. Les principaux marqueurs géniques surexprimés dans les deux types de tissus étaient KRT7, KRT19, SPINK1 pour l'urothélium, LUM, DCN, PECAM1 pour la lamina propria et ACTG2, ACTA2, MYH11 pour la muscularis propria. Chaque cluster était cohérent avec la section de tissu H&E correspondante. Nous avons identifié les zones où des sous types cellulaires sont surexprimées par la co-expression de marqueurs géniques connus.

**Conclusion :** Nous avons identifié de multiples clusters géniques en utilisant l'expression spatiale des gènes dans l'appareil urinaire sain. Ces clusters étaient fortement corrélés aux couches tissulaires. La technologie Visium, bien que limitée par sa résolution, permet l'étude de la transcriptomique avec une résolution spatiale et met en évidence une distribution hétérogène de certains sous types cellulaires basée sur des marqueurs géniques connus. L'étude de tissus pathologiques est la prochaine étape.

**Composition du Jury :**

**Président :** Monsieur le Professeur Arnaud VILLERS

**Assesseurs :** Madame le Docteur Guillemette MAROT  
Monsieur le Docteur Jonathan OLIVIER

**Directeur de thèse :** Monsieur le Docteur Gautier MARCQ

