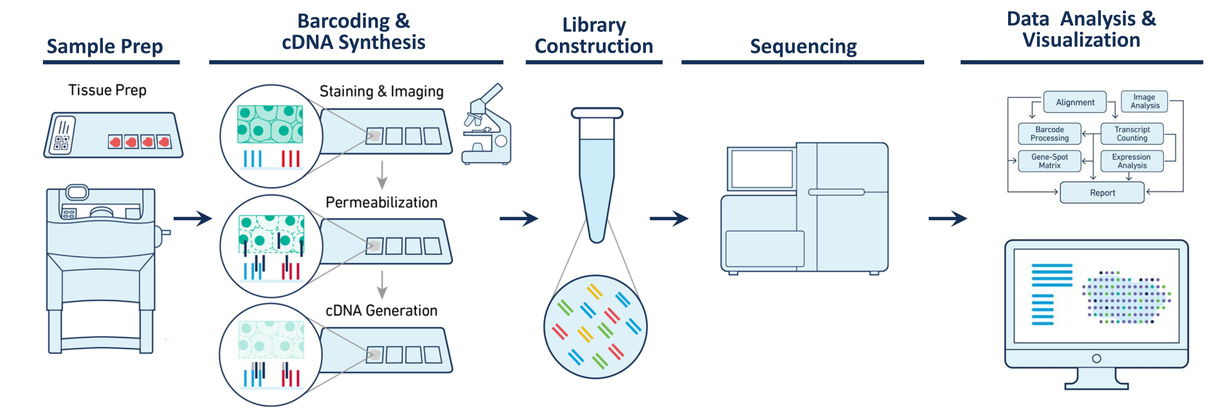
Anatomical and Transcriptional Characterization of Breast Tumor Heterogeneity Using Spatial RNA Sequencing

Solongo Ziraldo, Cedric Uytingco[1], Stephen Williams[2], Jennifer Chew[1], Meghan Frey[1], YiFeng Yin[1], Francesca Meschi[1], Andrej Hartnett[1], Eswar Iyer[1], Stefania Giacomello[1], Aleksandra Jurek[1], James Chell[1], Erik Borgstrom[2], Nigel Delany[2], Neil Weisenfeld[2], Zachary Bent[1], and Tarjei S. Mikkelsen[1]
[1]Molecular Biology, [2]Computational Biology, 10X Genomics Inc., Pleasanton, CA.

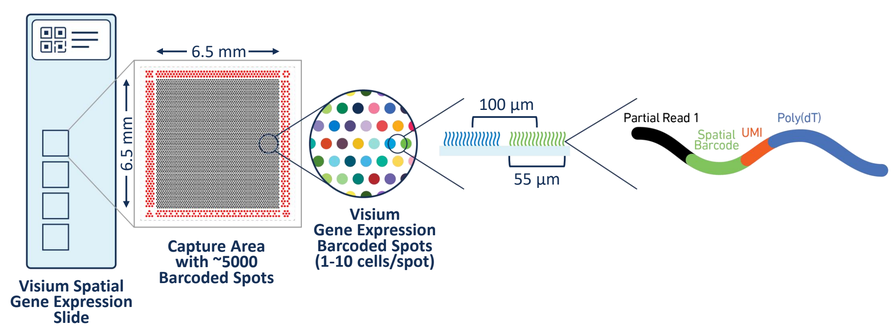
Introduction

The tumor microenvironment is consisted of highly heterogeneous cellular components that dynamically interact and communicate with each other. Significant advancement in single-cell RNA sequencing allowed capturing thousands of cells and has revealed many subpopulations of cells in tumor tissues. However, dissociation of the tissue into single cells results in the loss of its important architectural information. The recently introduced spatial transcriptomics technology resolves spatial localization of cells within a tissue section. Here we present an improved version of this spatial gene expression technology with increased tissue coverage, higher spatial resolution, and significantly improved sensitivity. We applied our improved technology to tumor tissue sections from human breast tumors and analyzed tissue-wide transcriptomics profiles to locate cancer related genes within spatial context and reveal intra-tumor heterogeneity within a tissue section. Elucidation of the spatial heterogeneity of tumor cells can shed light into understanding the disease states and progression, aiding treatment decisions.

Overview of unbiased spatial gene expression methodology

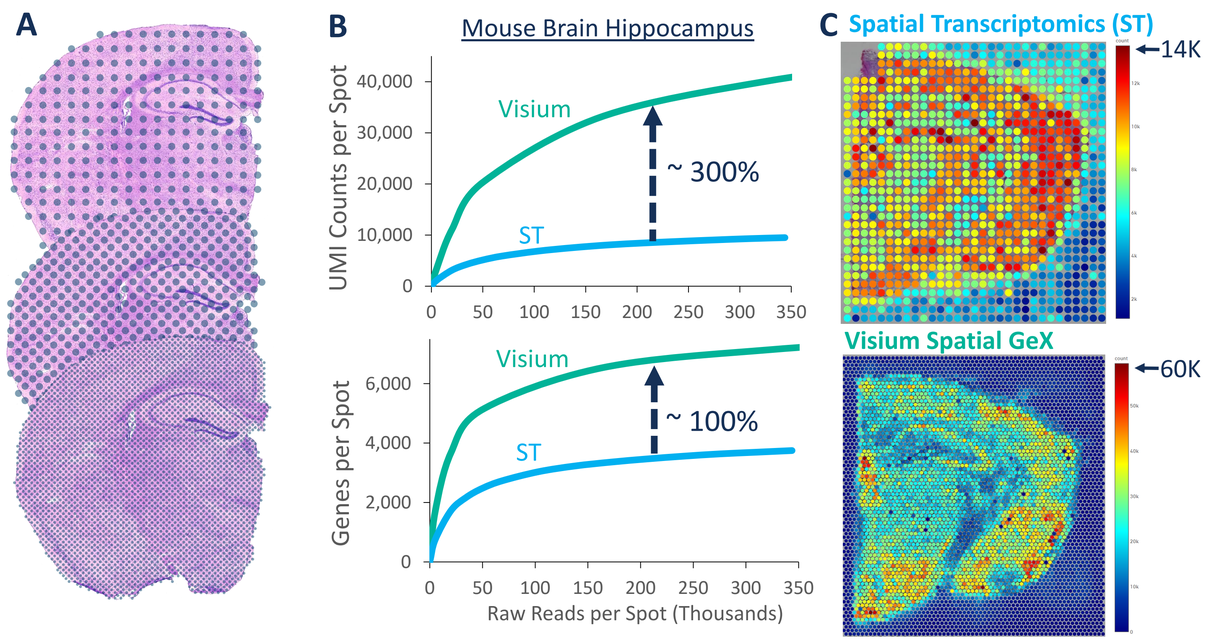


Fresh-frozen tissue is sectioned, placed onto a Visium Gene Expression slide, then fixed, stained, and permeabilized, releasing mRNA which binds to spatially barcoded capture probes, allowing for the capture of gene expression information. cDNA is then synthesized from captured mRNA and sequencing libraries prepared. The libraries are then sequenced, data analyzed and visualized.



Each Capture Area in Visium Spatial Gene Expression Slide contains ~ 5000 spots. Each spot is functionalized with sequencing primers, spatially barcoded oligonucleotides that allow tracing RNA-seq data back to the location in the tissue section, unique molecular identifier (UMI), and poly dT to prime polyadenylated RNA transcripts. The diameter of each spot is 55 µm and the spot center to center distance is 100 µm. The frame, also known as fiducials, allow for easy alignment of tissues.

Spatial resolution is significantly improved by increasing spots and by improving the sensitivity.

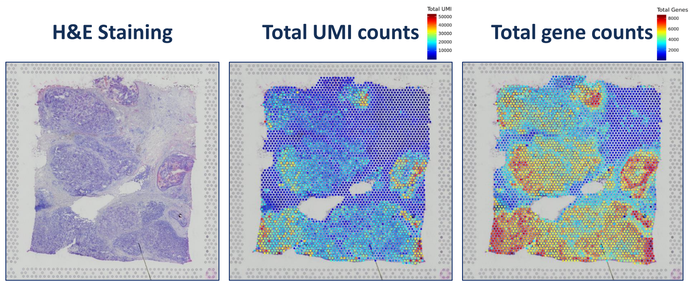


Denser spots improve the tissue coverage area and smaller spots improve the spatial resolution (A). The tissue section from the same mouse brain was processed with either Spatial Transcriptomics (ST)[1] or Visium workflow. Using ST method, 615 spots were covered under tissue and 3,748 genes per spot and 9,489 transcripts per spot observed at ~344,000 read pairs per spot. With Visium Spatial Gene Expression (Gex) Solution, 2,990 genes per spot and 40,132 transcripts per spot observed at ~327,000 reads per spot. With mouse hippocampus, ~300% and ~100% increase in transcript counts and in gene counts, respectively, observed with Visium Spatial Gene Expression compared to Spatial Transcriptomics (B). The heatmap clearly illustrates that the Visium workflow results in a significantly higher structural resolution of the mouse tissue along with higher UMI counts (C).

Characterization of Human Breast Invasive Ductal Carcinoma

We profiled the spatial gene expression of invasive ductal carcinoma tissue from a female patient (ER+, PR-, HER2+) [2]. As a control, the healthy tissue sections adjacent to the tumor were obtained. 4 replicates were used for each tissue type.

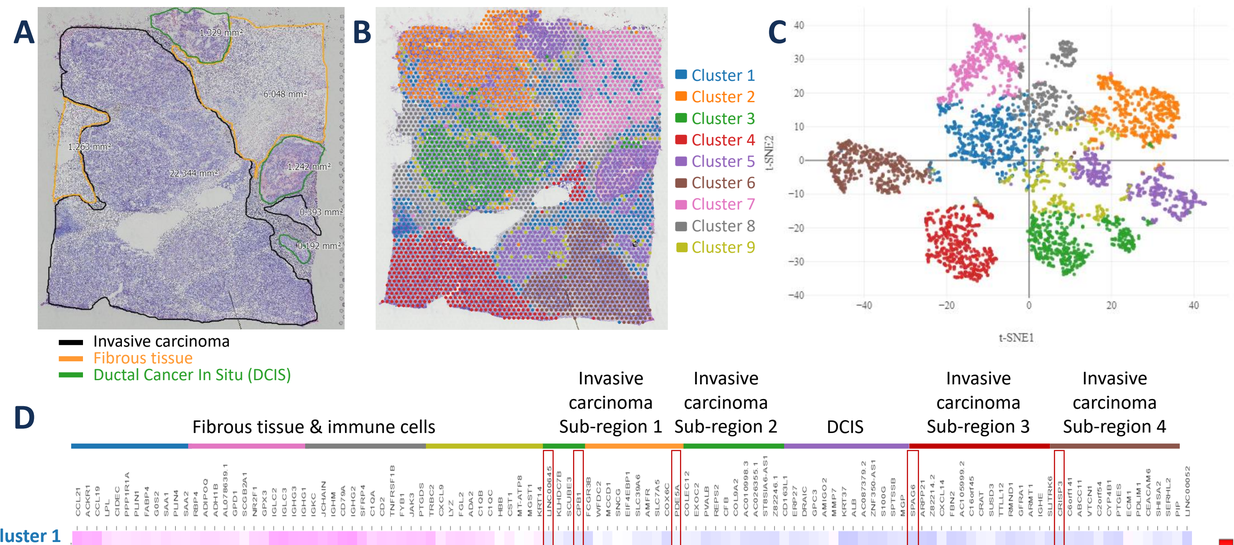
Tissue Type	Spots covered by tissue	Mean reads per spot (sequencing saturation)	Median UMI counts per spot	Median genes per spot
Invasive ductal carcinoma	~ 4,000	~ 70,000 (75%)	~ 10,500	~ 3,900
Normal	~ 1,800	~ 170,000 (90%)	~ 4,900	~ 2,100



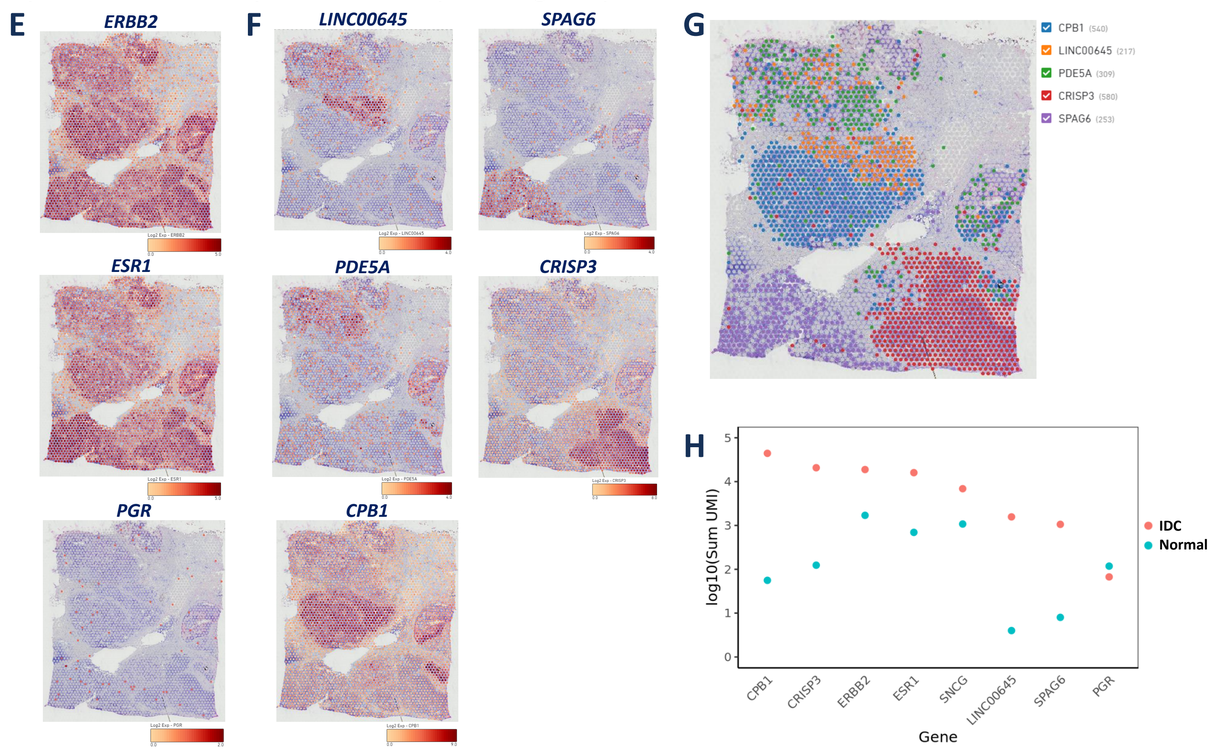
Please contact info@10xgenomics.com for more information.
https://www.10xgenomics.com/spatial-transcriptomics/

Spatially-resolved gene expression and clustering in invasive ductal carcinoma reveal intra-tumor heterogeneity.

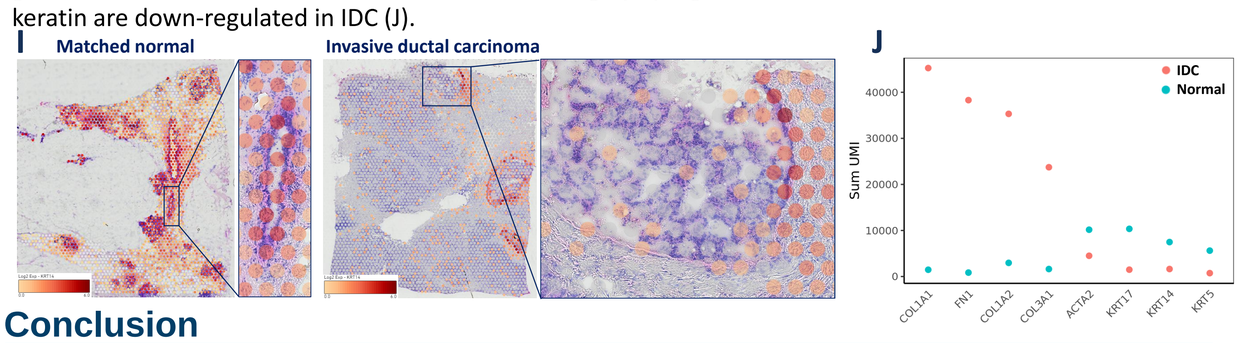
(A) Histological section of an invasive ductal carcinoma annotated by a pathologist. The section contains a large proportion of invasive carcinoma (black), three separate ductal cancer in situ regions (green), and fibrous tissue. (B) Tissue plot with spots colored by unsupervised clustering. (C) tSNE plot of spots colored by unsupervised clustering. (D) Gene expression heat map of the most variable genes between 9 clusters. The region defined as fibrous tissue mostly corresponds to cluster 1, 7 and 8. Interestingly, a large region annotated as invasive carcinoma contained spatial spots that were assigned to DCIS (cluster 5). In addition, four subtypes of invasive carcinoma with distinct molecular properties (cluster 2, 3, 4, and 6) were identified, revealing intra-tumor heterogeneity.



The expression levels of genes corresponding to human epidermal growth factor receptor 2 (Her2), estrogen receptor (ER), and progesterone receptor (PR) in the tissue section are shown in E. It is clearly visible that Her2 and ER are highly expressed in the invasive carcinoma and DCIS regions while the expression of PR is absent, consistent with the patient's diagnosis. One of the top differentially expressed genes from each cluster in the invasive carcinoma region is selected (red rectangular box in D), and its expression levels are located in the tissue (F) and overlapped in one plot (G). With the exception of *PGR*, all of these genes are highly up-regulated in the carcinoma tissue compared to the adjacent normal tissue (H). Analysis revealed that all of these up-regulated genes have implication in cancer progression. Interestingly, in the subset of cluster 3 a long non-coding RNA, of which abnormal expression has recently been implicated in tumor development [2], is one of the top differentially expressed genes. In glioblastoma, *LINC00645* promotes epithelial-to-mesenchymal transition by inducing TGF-β[3].



During breast cancer progression, the myoepithelial cells, which continue to surround preinvasive *in situ* carcinoma, gradually disappear[5]. This phenomena is clearly visualized below where *KRT14* (a gene signature of myoepithelial cells) is highly expressed around the lining of the duct in the normal tissue while it is disappear-ing in the DCIS region in IDC tissue (I). The extracellular matrix genes such as COL1A1 and FN1, key genes associated with invasion and metastasis, are highly upregulated while smooth muscles and basal keratin are down-regulated in IDC (J).



Conclusion

- Spatial resolution and the assay sensitivity are significantly improved in Visium Spatial Gene Expression Solution.
- Tissue-wide transcriptomics of human breast invasive ductal carcinoma are profiled to locate cancer related genes within spatial context and revealed intra-tumor heterogeneity within a tissue section.
- May enhance understanding of the molecular mechanism underlying carcinoma development and may shed light into a potential novel strategy for the treatment of human breast cancer.

References

1. Stahl, P. *et al*. Visualization and analysis of gene expression in tissue sections by Spatial Transcriptomics. *Science*. **2016**, *353*, 78.
2. BioIVT: Asterand (https://www.bioivt.com) - Case ID 66320; Specimen ID 116899F.
3. Zhang T, *et al*. Long Non-Coding RNA and Breast Cancer. *Technol Cancer Res Treat*. **2019**, *18*,1533033819843889
4. Li, C. *et al*. Long non-coding RNA linc00645 promotes TGF-β-induced epithelial–mesenchymal transition by regulating miR-205-3p-ZEB1 axis in glioma. *Cell Death & Dis.* **2019**, *10*, 272.
5. Gudjonsson, T. *et al.* Myoepithelial Cells: Their Origin and Function in Breast Morphogenesis and Neoplasia. *J. Mammary Gland Biol. Neoplasia.* **2009,** *10*, 261.