A pathway-centric approach to characterizing tumor heterogeneity and cell diversity across multiple cancer types

Ben Patterson[1], Cheng-Yi Chen[1], Nicolas Fernandez[1], Simran Kaushal[1], Jiang He[1], **Leiam Colbert[1]**

[1]Vizgen, Inc, 61 Moulton St, Cambridge, Massachusetts, United States

vizgen

Abstract: 5885

Introduction

The spatial information embedded in the heterogeneous tumor microenvironment plays a critical role in assessing patient prognosis. Vizgen's MERSCOPE® Platform is built on multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) technology and enables the direct profiling of the spatial organization of patient tumors with sub-cellular resolution, filling an important gap in previous iterations of transcriptomic evaluation. Here, we present a 500-gene pre-designed PanCancer Pathways Panel to comprehensively assess the genomic and cellular alterations found across diverse tumor types in human clinical samples using the MERSCOPE® Platform. **We demonstrate the ability of the PanCancer Pathways Panel to spatially profile gene expression across multiple tumor types, including breast, colon, prostate, ovarian, lung, and liver cancer**. Comparisons between experiments across multiple cancer types in fresh-frozen (FF) and formalin-fixed, paraffin-embedded (FFPE) preservation formats illustrate the robustness and reproducibility of this panel. To investigate how individual cell types in each tumor are dysregulated by cancer, we compared differentially expressed genes as detected by the MERFISH panel from normal and cancerous liver samples. We identified dysregulation in the *WNT*, *NOTCH*, and *TGFB* signaling pathways in the hepatocellular carcinoma samples, which are pathways known to play central roles in the development and progression of hepatocellular carcinoma. **These results demonstrate the power of the MERSCOPE® Platform and the PanCancer Pathways Panel to generate individualized, accurate cell atlases from patient-derived tumors and to enable further insights into the relationship between genomic profiles, dysregulated pathways, and disease phenotypes**. Such network-centric approaches are critical for assisting with identifying genotypic causes of diseases, classifying disease subtypes, and identifying drug targets.*

**RUO only, not approved for diagnostic or therapeutic purposes*

Materials and Methods

MERFISH enables the simultaneous measurement of the expression of hundreds of genes, giving insight into the quantity and distribution of RNA transcripts across a tissue slice. Target RNA species are labeled by tiling oligo probes to barcode each transcript in its native 3-dimensional cellular context. Each barcode is fluorescently detected in sequential rounds of imaging to resolve different RNA species (FIGURE 1A)[1].

We designed the MERSCOPE PanCancer Pathways gene panel to look at complex environments across multiple tumor types. The panel targets 500 pan-cancer genes including canonical signaling pathways of cancer, cancer type-specific genes, key immune genes, proto-oncogenes, and tumor-suppressor genes. The MERSCOPE FF and FFPE Sample Preparation Kits were used along with the MERSCOPE Cell Boundary Staining Kit to segment cells for further analysis of cell-cell interactions. FF and FFPE tissues were evaluated from the following human sample types: Normal Liver, Liver Cancer, Breast Cancer, Colon Cancer, Ovarian Cancer, Prostate Cancer, and Lung Cancer (FIGURE 1B). Bulk-RNA sequencing (RNA-seq) was performed on each tissue type for comparison.

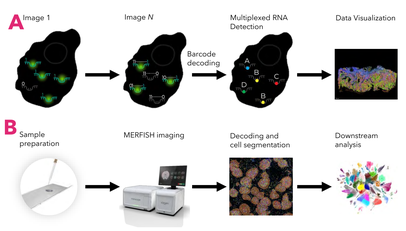

FIGURE 1. MERFISH process and study design. (A) During MERFISH, target RNA species are labeled by tiling oligo probes containing different barcodes; each barcode is fluorescently detected in sequential rounds of imaging to resolve the spatial location of different RNA species. (B) Prepared tissue samples were run on the MERSCOPE instrument for automated MERFISH imaging, decoding, and cell segmentation before downstream analysis.

Results

High correlation between replicates and across preservation methods

MERFISH counts across different gene targets in biological replicates have a Pearson's correlation coefficient above 0.9, with most samples from the same block sharing a correlation coefficient of ≥0.99 (FIGURE 2). The high levels of similarity across replicates demonstrates the MERSCOPE's capability for highly reproducible data capture, even in heterogenous tissues. Further, genes with low expression were reliably detected, revealing high detection efficiency. The transcriptional activity identified by MERFISH in each cancer type was consistent between samples preserved in FF or FFPE conditions. These results demonstrate the broad applicability of the PanCancer Pathways Panel across major tumor types.

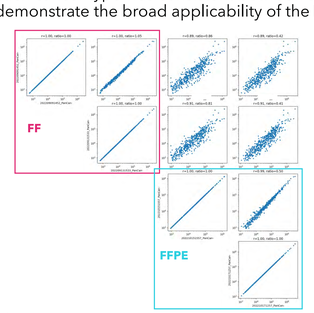

TABLE 1. Concordance values by sample type.

Sample Type	FF correlation	FFPE correlation	FF/FFPE correlation
Breast Cancer	0.99	0.99	0.76
Colon Cancer^	1.00	0.99	0.90
Lung Cancer^	0.99	0.99	0.72
Ovarian Cancer	0.99	0.99	0.79
Prostate Cancer	0.94	1.00	0.83
Liver Cancer	0.99	N/A	N/A
Normal Liver	0.99	N/A	N/A
^ matched			

FIGURE 2. High correlation between replicates and sample preservation methods in colon cancer tissue. MERFISH data from matched colon cancer tissue sections show high correlation between FF samples (pink box) and FFPE samples (blue box). FF and FFPE samples show relatively high correlation to each other (unboxed graphs, averages reported in Table 1). Concordance with RNA-seq data from each sample type was also calculated, data not shown. Reference data comparing MERFISH to RNA-seq and sc-RNA-seq in mouse brain and liver have been published[2].

© 2023 All rights in the trademarks, service marks, trade dress, logos and copyrights are owned by Vizgen, Inc. and fully reserved.

Identifying molecularly distinct cell types in tumor tissues

Our panel identified populations of endothelium, fibroblast, immune cell, and tumor cells in two fresh frozen colon cancer tissue samples (FIGURE 3). The UMAP and spatial projections are highly reproducible between these adjacent tissue samples (FIGURE 3A, 3C, 3F). While our panel is broad enough to be used across multiple cancer types, it is specific enough to identify cell types of interest to cancer researchers. For example, clear populations of fibroblasts surrounding immune cells can be found on the spatial projections of cell types (FIGURE 3F). To illustrate the broad applicability of this panel across multiple cancer types, we applied this same approach in the remaining tissue samples and preservation methods.

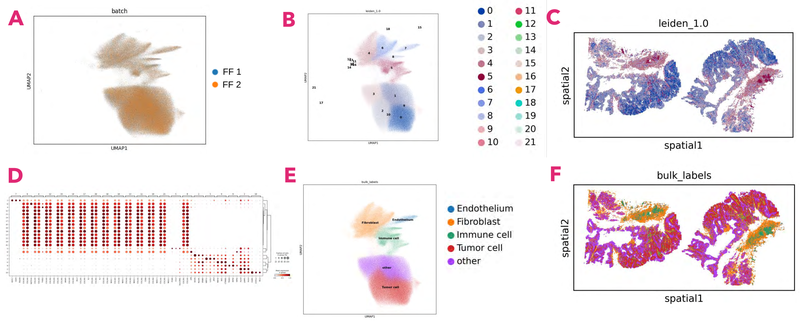

FIGURE 3. MERFISH identifies molecularly distinct cell types in fresh frozen colon cancer tissue. (A) UMAP of FF colon cancer tissues shows excellent overlap between samples. (B) UMAP with identified clusters (represented by different colors) revealed from unsupervised clustering analysis of MERFISH data. (C) Spatial mapping shows consistent cluster composition between samples. (D) Dot plot illustrating expression of top marker genes across different cell clusters. (E) Established marker genes were used to group cell clusters into major cell populations. (F) Cell type composition is similar in each tissue slice.

Applicability across tumor types and sample preservation methods

The transcriptional activity identified by MERFISH in each cancer type was relatively consistent between samples preserved in FF or FFPE conditions (Table 2). When designing this study, patient-matched samples of colon and lung cancer were included to guide general benchmarks for FF:FFPE correlation. When using the correlation scores, we observe that all matched and unmatched samples tested in both preservation types maintain a rather high FF:FFPE correlation of 0.7 and above. These results confirm the capability of our MERSCOPE PanCancer Pathways Panel to accurately detect gene expression in a variety of tissue contexts (FIGURE 4).

TABLE 2. Sample metrics.

Sample Type	Preservation Method	Average # Transcripts/Cell	Transcripts in Sample 1	Transcripts in Sample 2	Total Transcripts
Breast Cancer	FF	215	45,349,646	52,511,017	97,860,663
	FFPE	293	102,017,067	150,029,682	252,046,749
Colon Cancer^	FF	82	26,992,821	25,595,792	52,588,613
	FFPE	220	26,550,819	54,069,467	80,620,286
Lung Cancer^	FF	125	18,603,032	20,862,842	39,465,874
	FFPE	188	81,161,844	89,198,630	170,360,474
Ovarian Cancer	FF	159	39,723,603	53,246,183	92,969,786
	FFPE	84	4,747,710	25,627,136	30,374,846
Prostate Cancer	FF	87	17,581,245	39,005,065	56,586,310
	FFPE	168	104,538,240	67,744,301	172,282,541
Liver Cancer	FF	343	114,862,461	80,807,518	195,669,979
^ matched					

Identifying common cell types across multiple tumors in 3D

The genes included in the PanCancer Pathways Panel allow resolution of molecularly distinct cell populations across multiple tumour types. These molecularly distinct cell populations can then be mapped back to their 3-dimensional context in tumour samples, allowing insights into dominant cell types and cell-cell interactions.

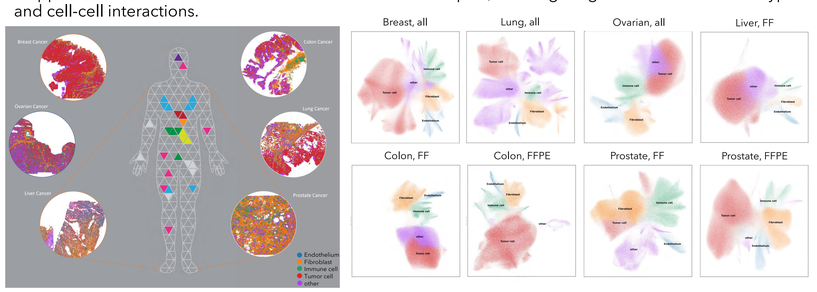

FIGURE 4. MERSCOPE data using the PanCancer Pathways Panel reveal molecularly distinct cell types across patient-derived tumor samples. UMAP showing the cell clusters revealed from unsupervised clustering analysis of MERSCOPE data, different colors represent different clusters. **Spatial projections of common cell types in tumors investigated in this study.** The MERSCOPE PanCancer Pathways Panel was used to profile the spatial transcriptomics of multiple tumor types with ease. We used gene expression data to cluster and label cell types, and project common cell types back onto the 3-dimensional spatial images of each tissue type, as illustrated.

Identifying patterns of dysregulation in liver cancer tissues

Using hepatocellular carcinoma liver cancer samples, the cell identities of the unsupervised clusters were assessed to provide a disease-specific view of the tissue biology (FIGURE 5A). The top expressed genes of each cluster were compared to databases such as UNCURL[7] and the Human Protein Atlas[8,9] to further refine the clusters (FIGURE 5B). Selective visualization of the tumor clusters was possible by returning to the spatial view of the unbiased cell clusters identified (FIGURE 6D). By visualizing the most populous tumor cell clusters in the samples, spatially defined gradients of gene expression in tissue were revealed (FIGURE 5D).

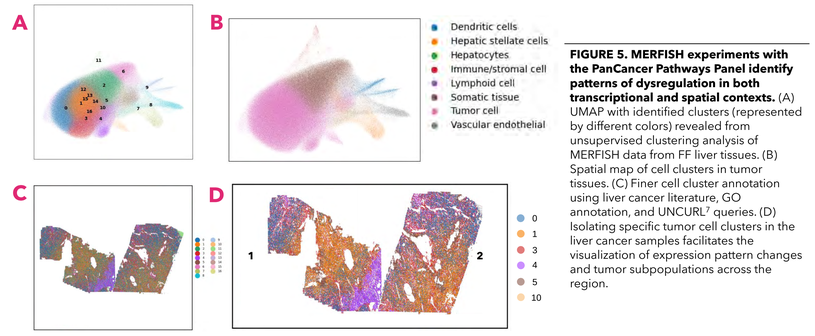

FIGURE 5. MERFISH experiments with the PanCancer Pathways Panel identify patterns of dysregulation in both transcriptional and spatial contexts. (A) UMAP with identified clusters (represented by different colors) revealed from unsupervised clustering analysis of MERFISH data from FF liver tissues. (B) Spatial map of cell clusters in tumor tissues. (C) Finer cell cluster annotation using liver cancer literature, GO annotation, and UNCURL[7] queries. (D) Isolating specific tumor cell clusters in the liver cancer samples facilitates the visualization of expression pattern changes and tumor subpopulations across the region.

Comparing gene expression in healthy liver and liver cancer tissues

We tested the PanCancer Pathways Panel on normal liver tissue and compared the gene expression of tumor clusters with the normal tissue clusters to identify altered gene activity (FIGURE 6A and B). We pooled the top 25 genes enriched in each tumor cell cluster and performed a network analysis with *STRING-DB*[3] and a gene ontology (GO) analysis (FIGURE 6C) to compare protein association networks across both sample types. We identified dysregulation in the *WNT*, *NOTCH*, and *TGFB* signaling pathways in the hepatocellular carcinoma samples, which are known to play central roles in the development and progression of hepatocellular carcinoma. The high differential expression between normal and tumor tissue highlights the use of the MERSCOPE PanCancer Pathways Panel to unravel complex tumor biology[4,5,6].

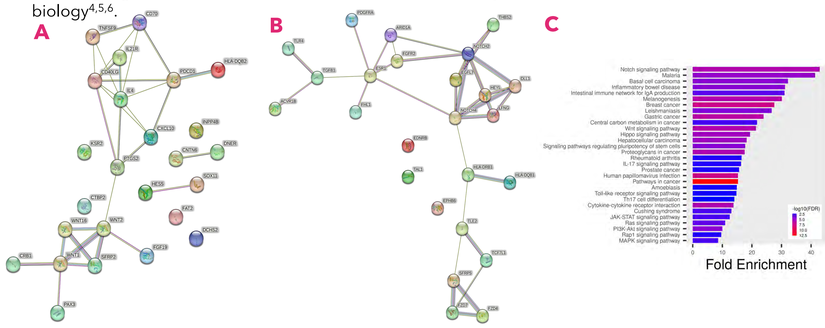

FIGURE 6. MERFISH data provide insights into gene expression of liver cancer samples. (A) Gene networks of the top 25 differentially expressed (DE) genes in normal sample cells and (B) cancer sample cells. (C) GO analysis of the pooled top 25 DE genes of all tumor cell clusters demonstrates enrichment of various canonical cancer pathways.

Conclusions

1. We designed a broadly applicable PanCancer Pathways Panel that was used to spatially profile several types of cancer using the MERSCOPE® Platform.
2. The MERSCOPE PanCancer Pathways Panel reproducibly detects transcripts in both normal tissues and cancer samples.
3. The MERSCOPE PanCancer Pathways Panel can reliably detect transcripts in both FF and FFPE-preserved tissue types. MERFISH measurements from FF and FFPE tissue blocks often correlate, regardless of whether they are matched samples. Comparing only matched samples may result in higher concordance between samples in future studies.
4. The same PanCancer Pathways Panel robustly detects target genes in 6 major tumor types.
5. The PanCancer Pathways Panel gives sufficient data to detect multiple tumor clusters in a spatial gradient in hepatocellular carcinoma patient samples.
6. We identified dysregulation of 3 major canonical drivers of hepatocellular carcinoma in our MERFISH data, highlighting the ability of the PanCancer Pathways Panel to capture critical cancer biology occurring in patient-derived samples.

References

1. Spatially resolved, highly multiplexed RNA profiling in single cells (2015). DOI: 10.1126/science.aaa6090
2. Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing. DOI: 10.26508/lsa.202201701.
3. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. DOI: 10.1093/nar/gkaa1074.
4. The Carcinogenic Role of the Notch Signaling Pathway in the Development of Hepatocellular Carcinoma. DOI: 10.7150/jca.26847.
5. Transforming growth factor-β as a therapeutic target in hepatocellular carcinoma. DOI: 10.1158/0008-5472.
6. Aberrant regulation of Wnt signaling in hepatocellular carcinoma. DOI: 10.3748/wjg.v22.i33.7486.
7. Zhang, Y., Mao, S., Mukherjee, S., Kannan, S. & Seelig, G. UNCURL-App: Interactive Database-Driven Analysis of Single Cell RNA Sequencing Data. bioRxiv, 2020.2004.2015.043737, doi:10.1101/2020.04.15.043737 (2020).
8. Karlsson, M. et al. A single-cell type transcriptomics map of human tissues. Sci Adv 7, doi:10.1126/sciadv.abh2169 (2021).
9. Uhlen, M. et al. Proteomics. Tissue-based map of the human proteome. Science 347, 1260419, doi:10.1126/science.1260419 (2015).

Scan me for more information