Integrative spatial and single-cell analysis elucidates tumor microenvironment heterogeneity in NSCLC

Qiming Zhou[1], Jiadong Zhou[2], Beixi Wang[3], Jingwen Shi[2], Han Yan[3], Xin Chen[4], Yun Zhang[5, #], Pei Zhang[3, #], Zhirong Shen[2, #]

[1]Bioinformatics, BeiGene (Shanghai) Co., Ltd., Shanghai, China; [2]Clin Biomarker Science and CDx Development, BeiGene (Beijing) Co., Ltd., Beijing, China;
[3]Bioinformatics, BeiGene (Beijing) Co., Ltd., Beijing, China; [4]Clin Biomarker Science and CDx Development, BeiGene USA, Inc., Albany, USA; [5]Biotech Unit CN-BJ, BeiGene (Beijing) Co., Ltd., Beijing, China; [#]Corresponding author.
#Correspondence: pei.zhang@beigene.com

Abstract No. 2090

Background

Tumor microenvironment (TME) heterogeneity plays a crucial role in cancer progression and therapeutic response. Advanced multi-omics spatial and single-cell technologies offer unprecedented resolution for dissecting TME complexity. However, the application of these technologies to formalin-fixed paraffin-embedded (FFPE) samples remains underexplored. This study aims to demonstrate the power and capability of these technologies in analyzing FFPE samples, benchmark and cross-validate different platforms/assays, and deepen our understanding of TME heterogeneity at the single-cell level and its spatial context.

Methods and Materials

We performed following assays including (1) High-plex multiple IHC platform: PhenoCycler Fusion (PCF) for protein-level spatial analysis (n=10), (2) FFPE Single-Cell RNA Sequencing (FFPE scRNA-seq) (n=4, including 3 paired-samples with PCF) and (3) Bulk-level RNAseq RNA, PD-L1 IHC and H&E for cross validation analysis with FFPE Non-small cell lung cancer (NSCLC) samples (**Figure 1**). PCF assay data were processed by HALO high-plex module for segmentation and marker co-localization and exported for Seurat single cell analysis. FFPE scRNAseq utilized 10x Genomics FLEX assay for tissue resection blocks. FFPE scRNAseq data were pre-processed by cellranger and downstream analysis were done as a Seurat object in R.

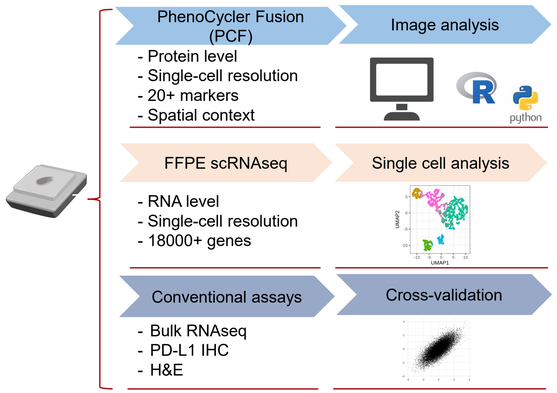

Figure 1. Study design to collectively elucidate tumor microenvironment at single cell level.

Conclusion

Overall, the study demonstrates the effectiveness and reliability of PCF along with ML-based H&E tumor area identification in evaluating PD-L1 expression. The high concordance across multi-omics approaches and the detailed transcriptome profiling provided by FFPE scRNAseq highlight the robustness of these latest methodology on FFPE samples. By integrating the detailed single-cell spatial context and composition provided by PCF with the single-cell transcriptional profiles from FFPE scRNA-seq, we enable a granular dissection of the tumor microenvironment (TME) and potential of unveiling molecular mechanisms underlying its heterogeneity. This multimodal approach advances our ability to decipher spatially resolved cellular interactions and functional states within complex tissue contexts.

Integrative imaging analysis help to identify tumor PD-L1 expression

Firstly, to evaluate the PCF marker expression pattern, cell density of PanCK+PD-L1+/PanCK+ cell ratio (%) was compared to PD-L1 IHC TC score, which showed a high correlation (*R*=0.73). Further, By image co-registration between H&E with ML-based tumor/stromal classification and PCF PanCK-PD-L1 staining, PD-L1+ cell density in tumor cell specifically can be precisely calculated, since PanCK+ cells can be non-tumor cells.

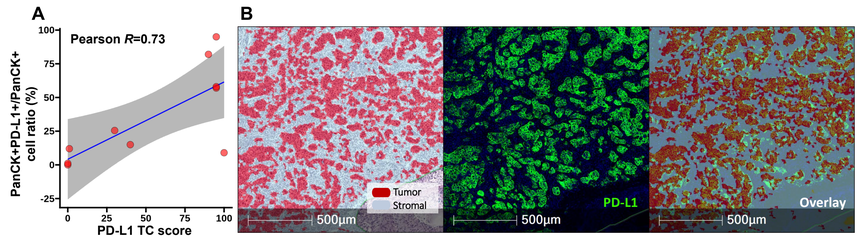

Figure 2. ML-based Tumor area classification to guide precise PD-L1 cell density calculation. (A) High Pearson's correlation between PD-L1 IHC TC score (x-axis) and PanCK+PD-L1+/PanCK+ cell ratio (y-axis) sample-wise; (B) ML-based Tumor area classification gave high overlap between PD-L1 expression and Tumor area. Left, ML-based tumor/stromal area classification on H&E image. Middle, PD-L1 staining in PCF assay. Right, Overlay image.

Abbreviation: ML: Machine Learning. TC: Tumor Cell. PanCK: Pan-Cytokeratin

High-resolution TME dissection enabled by whole slide cell annotation in PCF

Overall, 7 million single cells were segmented and profiled from 10 samples. A diverse TME was identified, highlighting the complexity and heterogeneity of TME components. This diversity was evident in the range of immune cell subsets and cancer-associated stromal cells detected with their markers specifically expressed, underscoring the importance of comprehensive TME analysis even in small cohorts. With Tumor/stromal area classified in H&E providing a spatial context of tumor cell distribution, PCF provided more spatial context with detailed cell type heterogeneity in different area of the tissue. The upper tissue were more showed tumor cell enriched with immune infiltration whereas the lower tissue area showed tumor/fibrotic mixed pattern. This TME partition difference has also been observed by KNN-neighborhood proximity analysis. There were PanCK-fibroblast enriched neighborhood partition, representing lower part of the tissue, and immune infiltrated partition with different immune cell harboring together, representing upper tissue part.

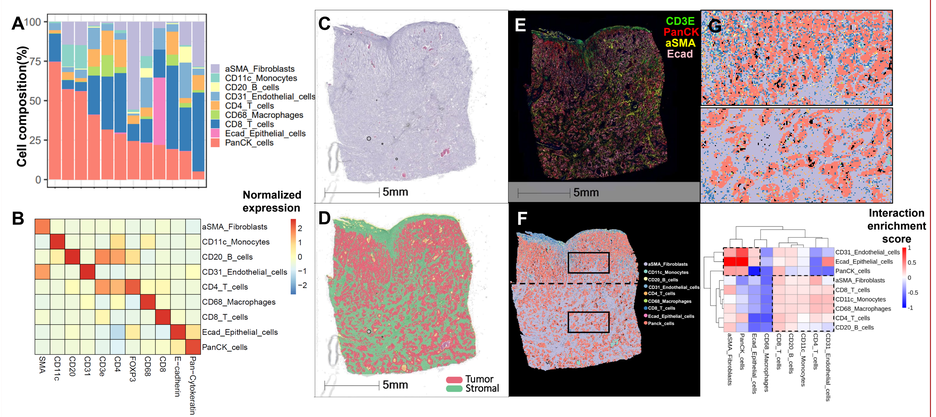

Figure 4. (A) Different TME composition identified in PCF; **(B)**. Cell type marker expression alignment with cell phenotype; **(C)** Whole slide H&E staining after PCF assay; **(D)** Tumor area classification by ML-model; **(E)** PCF cell type marker overlay; **(F)** Cell type spatial projection; **(G)** Zoom-in images of different TME within same tissue; **(H)** PCF spatial proximity enrichment by KNN.

High concordance across PCF, FFPE scRNAseq, and bulk RNAseq

To further evaluate the marker gene expression consistency, correlation was to conducted. High gene expression concordance was observed across PCF, FFPE scRNA-seq and bulk RNA-seq data for three paired samples. This overall high concordance (average Pearson's correlation *R*=0.84) validates the consistency and reliability of our multi-omics approach, ensuring that the data obtained from different sample types are comparable and robust.

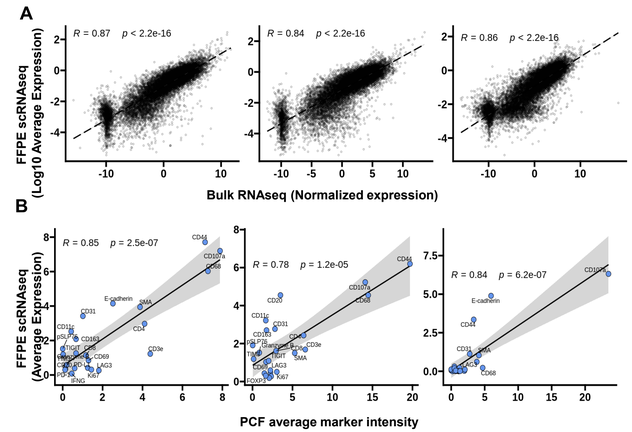

Figure 3. High concordance across PCF, FFPE scRNAseq and bulk RNAseq in paired samples. (A) Gene expression concordance: bulk RNAseq vs FFPE scRNAseq; **(B)** Marker expression concordance: PCF vs FFPE scRNAseq.

FFPE scRNAseq reveals diverse transcriptome profile

In addition to spatial context with detailed cell type level by PCF, FFPE scRNAseq enabled a deeper profiling at transcriptome level to facilitate the understanding the TME at the molecular level. Overall, about 36,000 cells were recovered from 4 different tissue block samples. In total, 12 cell populations were identified with a median gene recovery rate at 1130 genes per cell, which contribute to a more detailed cell lineage identification including small populations like mast cells and neutrophils. In tissue sample shown in Figure4, both spatially resolved cell-cell interaction in PCF and Ligand-receptor (LR) analysis by FFPE scRNAseq showed a trend of two distinct TME partition, PanCK-Epithelial-cell partition and Fibroblast-immune-cell partition with myeloid cells and PanCK-Epithelial cells being the most signific LR interaction communication.

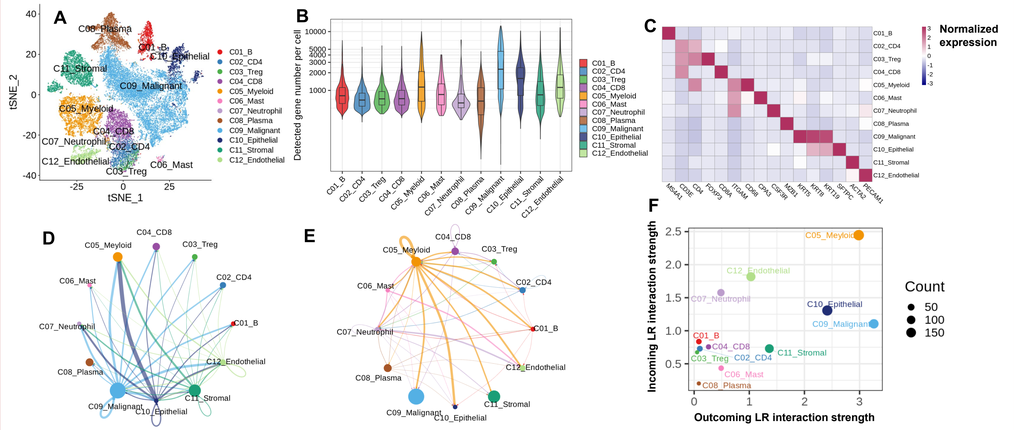

Figure 5. (A) Cell type tSNE embedding; **(B)** Gene detection number across cell types; **(C)** Average marker expression in annotated cell types; **(D)** Ligand-Receptor interaction outcoming by cells from C09 to C12; **(E)** Ligand-Receptor interaction outcoming by cells from C01 to C08; **(F)** Directions of LR interactions strength from different cell types.