



EFFICIENT DRUG OPTIMIZATION:

Cell deconvolution to estimate cell type abundance from spatial transcriptomic data within heterogeneous tissues

SUMMARY

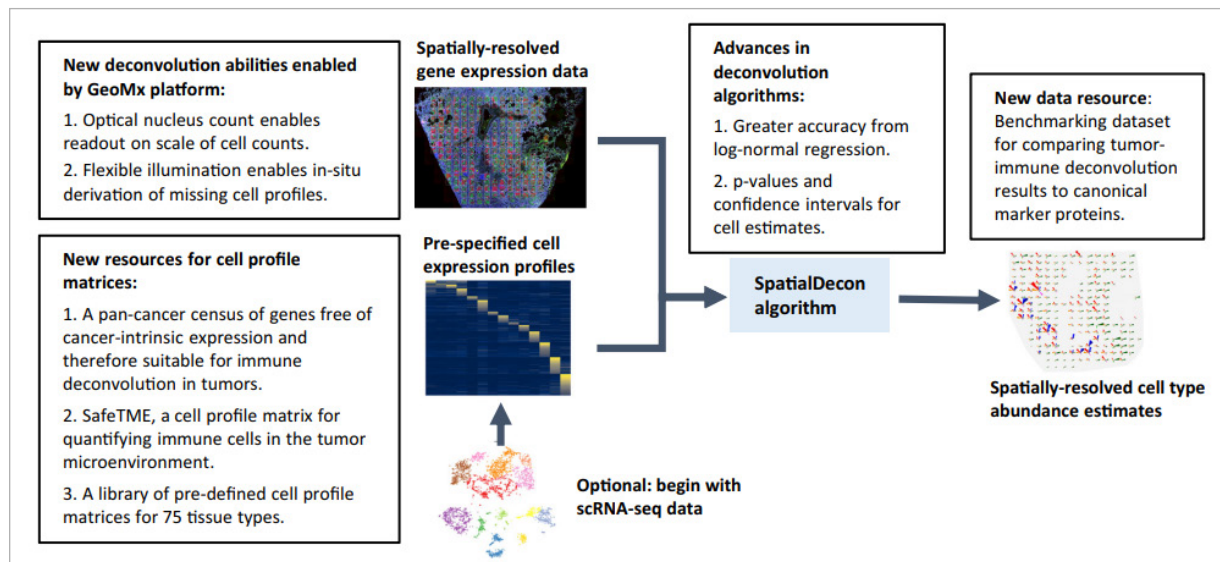
Mapping cell types across a heterogeneous tissue is a central issue of spatial biology as it helps understand cells relationship in the disease context and therefore their involvement in treatment response. The application of deconvolution to quantify cell populations defined by single cell sequencing within the regions of spatial gene expression studies has thus become an indispensable tool for in depth tissue biomarker questions. In this perspective, we highlight the use of the SpatialDecon R library by NanoString to quantify baseline cell type abundance defined by spatial transcriptomic data in non-small-cell lung cancer (NSCLC) samples that differentially responded to checkpoint inhibitor therapy.

APPROACH

Here we give an overview of the cell deconvolution workflow (Fig1) as described in Danaher & Kim (2022) paper titled “Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data.”

The workflow that reports spatially resolved cell type abundance estimates from spatially resolved gene expression data starts with estimation of the background from the negative control probes, then the selection of the reference matrix also called cell profile matrix and finally the application of the SpatialDecon algorithm to run the deconvolution and thus obtain the cell type abundance information in the heterogeneous tissue.

FIGURE 1. Deconvolution workflow summary





CASE ILLUSTRATION STUDY

In the present study, we report abundance estimates of various immune cell types on two resected NSCLC patients that differentially responded to anti-PD1 treatment (CR and PD) using the cell deconvolution algorithm.

Spatial gene expression data from RNA Immune Pathway Panel (86 transcripts) were obtained from six (6) regions of interest (ROIs) (Fig2) further segmented into tumor and stroma compartments on each tissue. A total of 20 segments normalized against data matrix that passed QC were then exported from the GeoMx[®] DSP analysis suite to initiate the cell deconvolution.

Background estimation

Each data points' expected background from the negative control probes were estimated from its corresponding observation. As a result, a matrix of expected background for all data points in the normalized data matrix was obtained.

Choice of the cell profile matrix

A cell profile matrix is a pre-defined matrix that specifies the expected expression profiles of each cell type in the experiment. In this perspective, the Census Adult Immune x10 (Tab2) was selected to perform cell deconvolution considering the higher number of shared genes and the numerous immune cell types (Fig3) in common with the immune spatial profiling panel used for the NSCLC tissues.

FIGURE 2. GEOMx image of the tumor with ROIs superimposed.

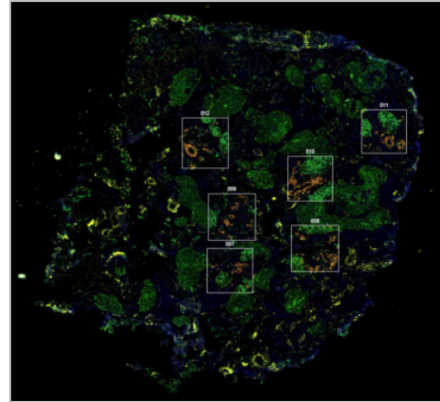


TABLE 2. Cell profile matrix information

	CENSUS ADULT IMMUNE X10
Database	Human Cell Atlas (HCA)
Type	Immune + Stroma
Number of genes	42 831
Number of cell types	22
Number of shared genes	75

FIGURE 3. Census Adult Immune x10 cell types (n=22)

	T.helper.cell
	pro.B.cell
	precursor.B.cell
	plasmacytoid.dendritic.cell
	plasma.cell
	natural.killer.cell
	naive.CD8..T.cell
	naive.B.cell
	mesenchymal.stem.cell
	memory.B.cell
	megakaryocyte
	hematopoietic.stem.cell
	erythroid.cell.type.2
	erythroid.cell.type.1
	cytotoxic.T.cell.type.2
	cytotoxic.T.cell.type.1
	conventional.dendritic.cell
	CD4..naive.T.cell
	CD16..monocyte
	CD14..monocyte.type.2
	CD14..monocyte.type.1
	B.cell.T.cell.doublet



Cell deconvolution

Cell deconvolution using the spatialdecon algorithm was then performed to estimate mixed cell type abundance in the spatially-resolved gene expression defined segments.

The minimal required input is:

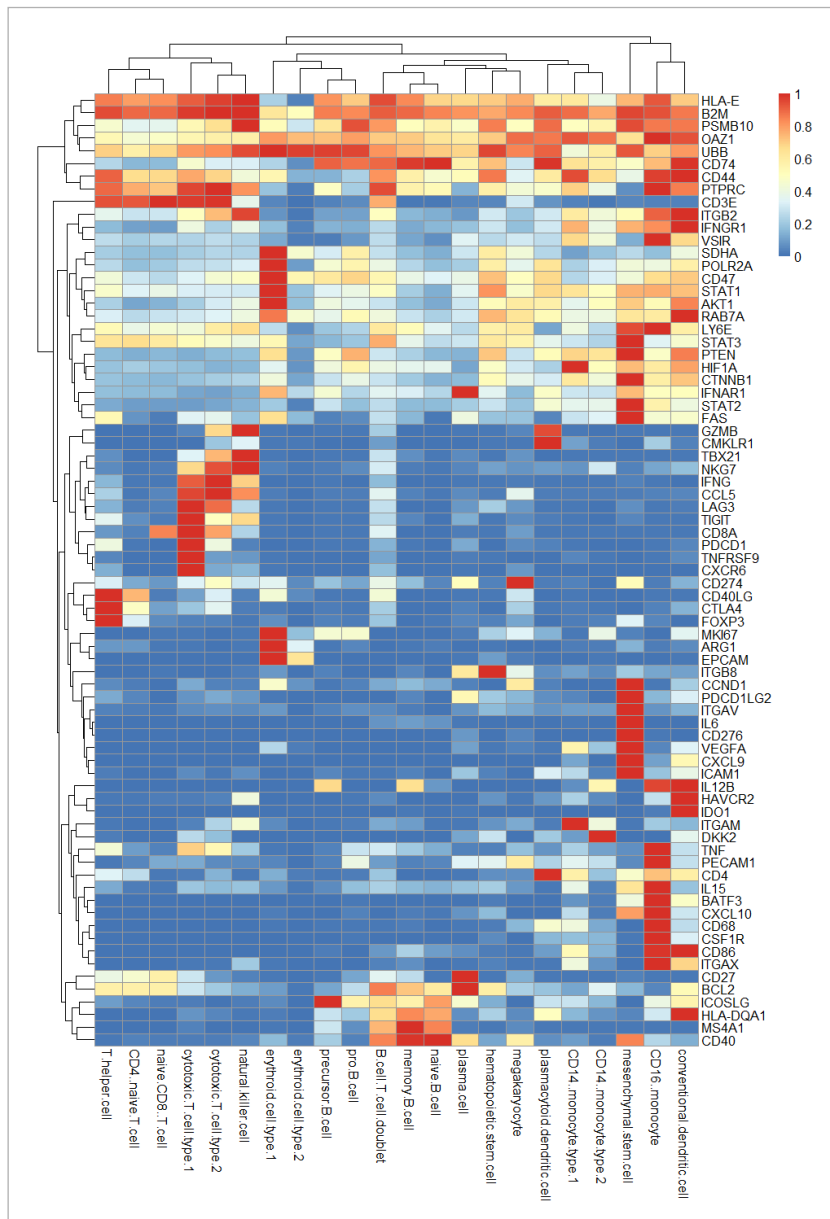
- The normalized data matrix
- The background estimation matrix previously calculated
- The Census Adult Immune x10 cell profile matrix

In addition, two advanced settings of the algorithm were used to go beyond the basic deconvolution:

- The nuclei counts for each region were provided to estimate the number of cells of each cell type
- The raw data matrix

The first output (Fig4) is the matrix of expected expression profiles of each cell type. PDCD1 (PD1) and CD8a gene are seen over-expressed in cytotoxic T cell type 1 & 2 as expected.

FIGURE 4. Heatmap of expected expression profiles of each cell type in the experiment

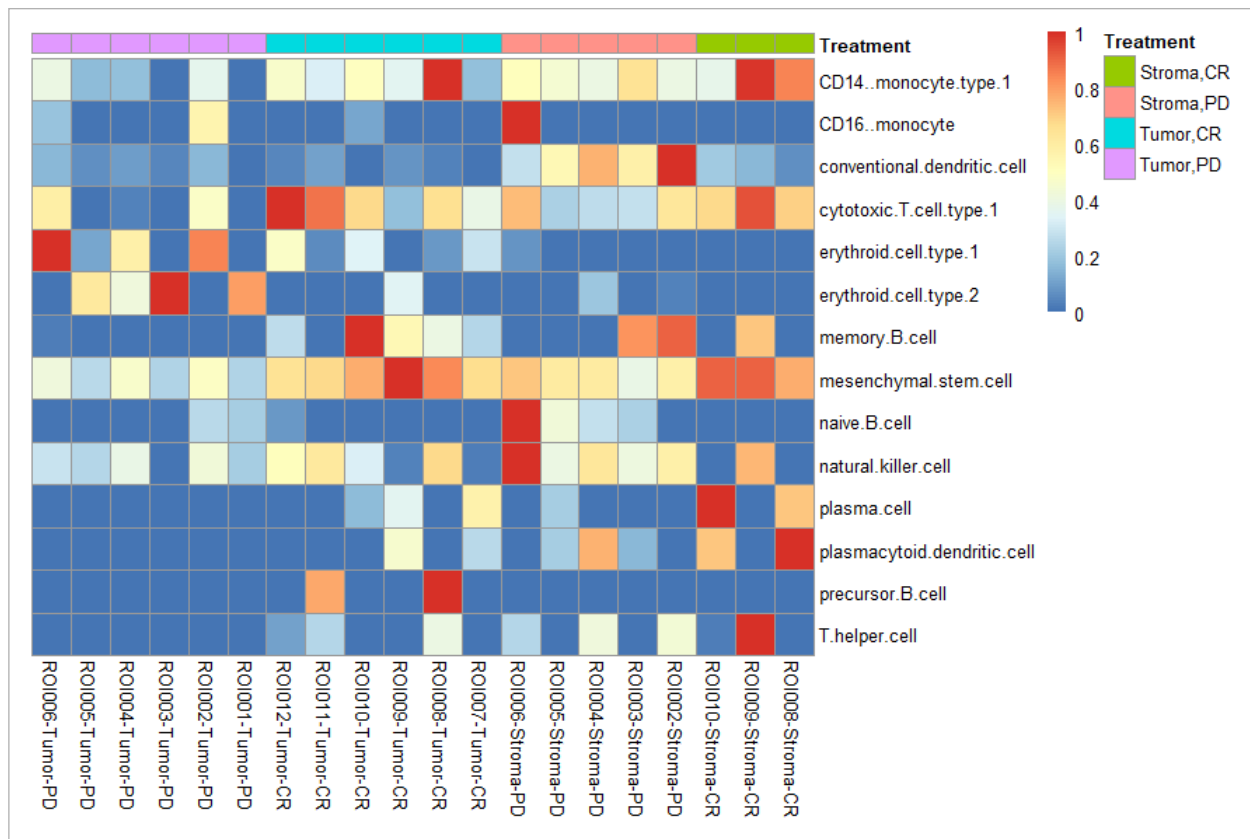




The second output (Fig5) is the cell abundance scores of the cell types for each segment. The cell types that had a score of 0 for all ROIs/segments were removed from the heatmap.

As expected, stroma segments in both CR and PD tissues contain a great number of immune cells. On the contrary, only the tumor segments from the CR patient had immune cells. Estimation of immune cells in the tumor segments from the PD patient was very low. Erythroid cells were abundant in this patient indicating their ability at repressing immune response.

FIGURE 5. Heatmap of estimated cell abundance in each region



Barplot of cell type proportion

A plotting function of the SpatialDecon package consists of drawing barplot of cell type abundance or proportion.

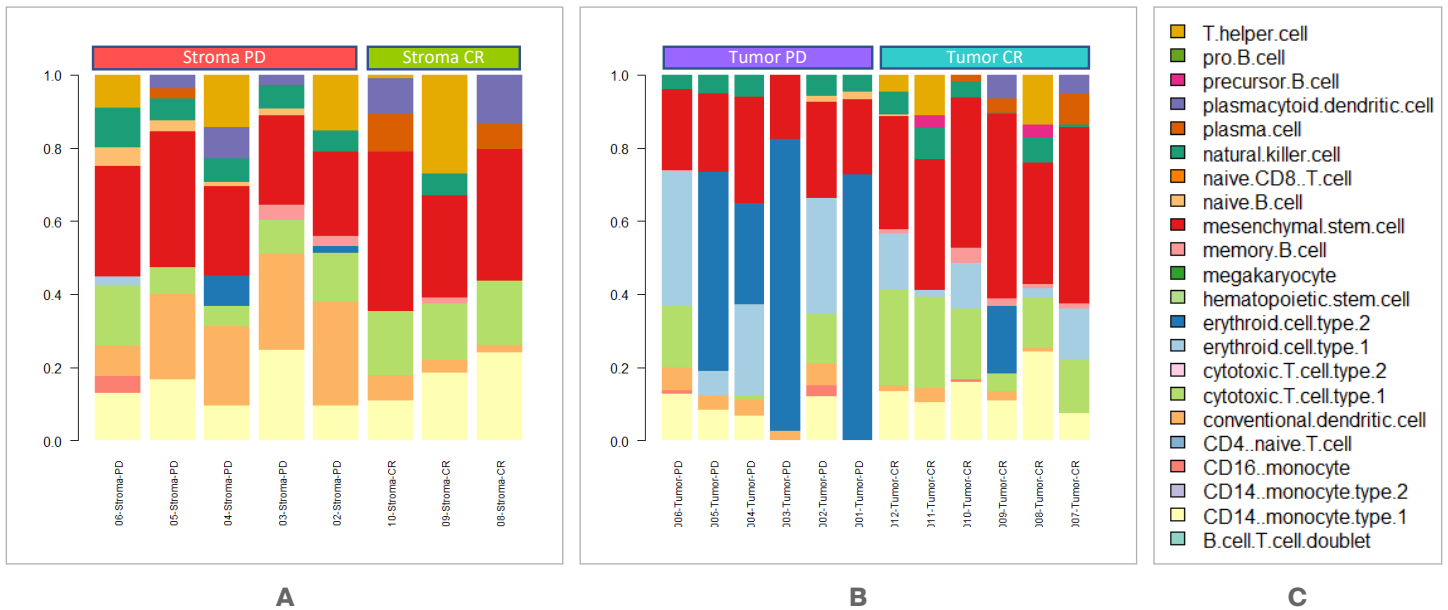
In the Stroma segments (Fig6A), the major cell types are mesenchymal stem cell, conventional dendritic cells, cytotoxic cells type 1 and CD14+ monocyte type 1. The differentiation between CR and PD is subtle : 2/3 of CR have plasma cells versus 1/6 of PD, all PD have NK cells versus only one CR.

The presence of erythroid cells in half of PD segments was also noticed.

In the Tumor segments (Fig6B), mesenchymal stem cells are present as well in both patients; however, a clear difference is seen at the level of erythroid cells abundance and cytotoxic T cells type 1 abundance indicating a role for these cell types in the response to checkpoint inhibitor treatment.



FIGURE 6. Barplot of cell type proportion. A- In Stroma Segments. B- In Tumor Segments. C – Cell types



Florets

The other plotting function found in the SpatialDecon package consists of the representation as “floret”: for each coordinate of a 2D plan, a circular barplot is drawn indicating the localized abundance of cell types. Fig7A&B show the cell abundances only in the stroma segments of each analyzed tissue. Another way to visualize the floret is to perform PCA on the normalized data and plot the cell abundances on top of the first two principal components obtained.

The PCA (Fig8) displays a strong separation between the four groups : Stroma-PD on top, Stroma-CR on the right, Tumor CR on the bottom and Tumor-PD on the left. Moreover, the first component opposes clearly the Tumor-PD to the other groups.

FIGURE 7. Stroma segments' florets in CR (A) and PD (B)

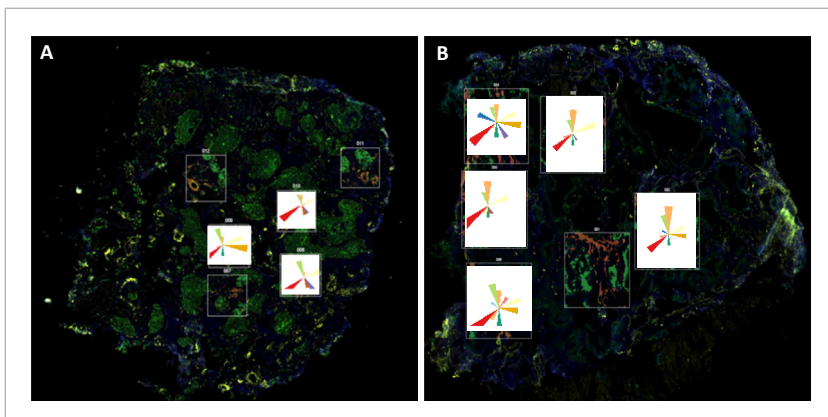
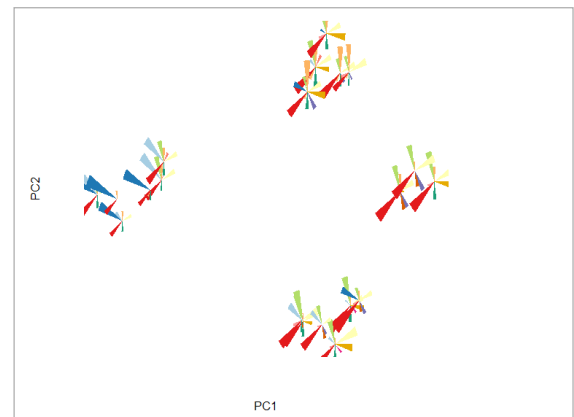


FIGURE 9. Floret PCA on all segments





APPLICATIONS

Based on cell abundances, we identified different microenvironment cell subtypes within NSCLC tissues. This heterogeneity raises the prospect that tumors could be classified not just by their overall cell abundance, but by the localized microenvironment subtypes they contain.

What role do cellular abundance play in promoting treatment response? These microenvironment phenotypes based on cell types abundance are providing new insights into disease mechanism and context and can reveal:

Novel drug targets

Novel efficacy biomarkers/
predictive biomarkers

Novel impact of drug on tissue

For more information visit www.aliribio.com.