

## Use of the “SpatialDecon” DSP DA Script

### Intended use

This script was designed for data from GeoMx high-plex RNA assays, such as the CTA NGS readout application. It estimates the abundance of mixed cell types within each AOI in an experiment.

IMPORTANT: please use the appropriate cell profile matrix that represents the tissue type of interest as this will affect the cell abundance/proportion scores. The default matrix below is for Solid Tumor GeoMx data.

### Relevant Resources

A complete description of the SpatialDecon algorithm is at <https://www.biorxiv.org/content/10.1101/2020.08.04.235168v2>

The algorithm’s github page is <https://github.com/Nanostring-Biostats/SpatialDecon>

### Running the script

The SpatialDecon script needs two pieces of data to run:

1. A normalized expression matrix. You should use whatever normalization technique you have decided is best for your data.
2. A “cell profile matrix”. This is a .csv file giving the expected expression profiles of each cell type in your dataset. Example matrices can be found at <https://github.com/Nanostring-Biostats/CellProfileLibrary>. For tumor immune deconvolution, use the file “safeTME-for-tumor-immune.csv”, which is provided along with the script code. To use a custom matrix, make sure it matches the format of the matrices referenced above.

In addition, the script accepts five “arguments” that you can set by modifying the top of the script’s code. Instructions for how to use these arguments are in-line in the script’s R code.

Briefly, the arguments are:

1. *cell\_profile\_filename*: the .csv file containing the cell profile matrix. This is the name of whatever cell profile matrix .csv file you’ve uploaded to the DSP DA.
2. *pure\_tumor\_column\_name*: If you have tumor data with ROIs segmented into tumor and microenvironment, you can use this argument to specify which AOIs are almost pure tumor cells. The algorithm will use this information to fit a tumor cell profile and append it to the cell profile matrix. This optional step leads to slightly more accurate immune cell abundance estimates.

To use this argument, enter the name of a column in the segmentAnnotations. The code will look for entries in that column that say “tumor”.

3. *variables\_to\_plot*: Enter column names of any segmentAnnotations variables that you'd like to appear in plots. Use column names without special characters such as dashes or spaces; and begin column names with letters, not a number.
4. *custom\_annotation\_coloring*: Allows you to define custom coloring for the "*variables\_to\_plot*". To enable, set this argument to TRUE, and modify the example syntax provided in the code. Any typos here will cause the script to error out.
5. *hmcols*: Specify a color gradient for heatmaps.

## Interpretation of results

The deconvolution algorithm outputs a .xlsx of tables and a .pdf with multiple pages of plots.

### Data tables output:

*Abundance scores tab*: Gives the estimated abundance of each cell type in each AOI. These cell abundance scores are interpreted on the same scale as the normalized data. The scores give abundance of each cell type scaled by whatever quantity you used to normalize the data, e.g. cell abundance per unit of area, or cell abundance per unit of total expression if Q3 normalization was used.

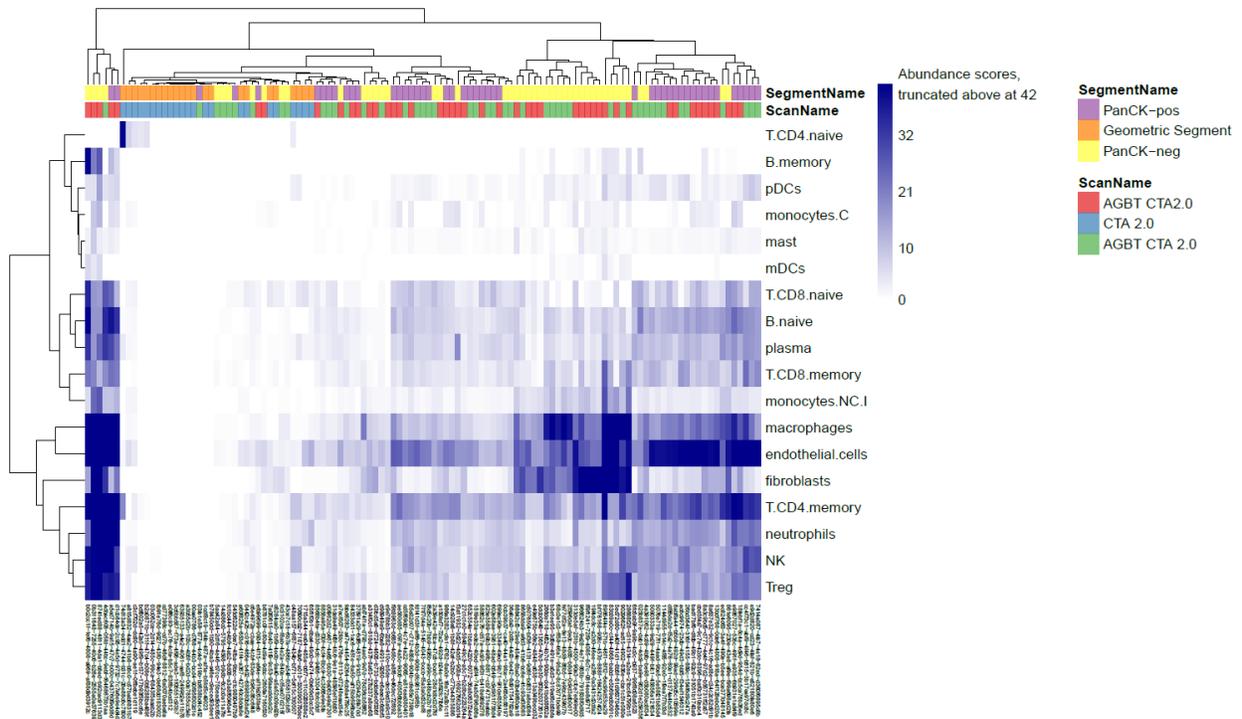
*Proportions tab*: Gives the proportion of each cell type in each AOI. Only "fitted" cells are included in this calculation. E.g. if tumor cells are present in the sample but the cell profile matrix only includes immune cells, then proportions will ignore the presence of tumor cells.

*Scaled abundance scores tab*: The same data as the "abundance scores" tab, but scaled so that each cell type has a maximum value of 1.

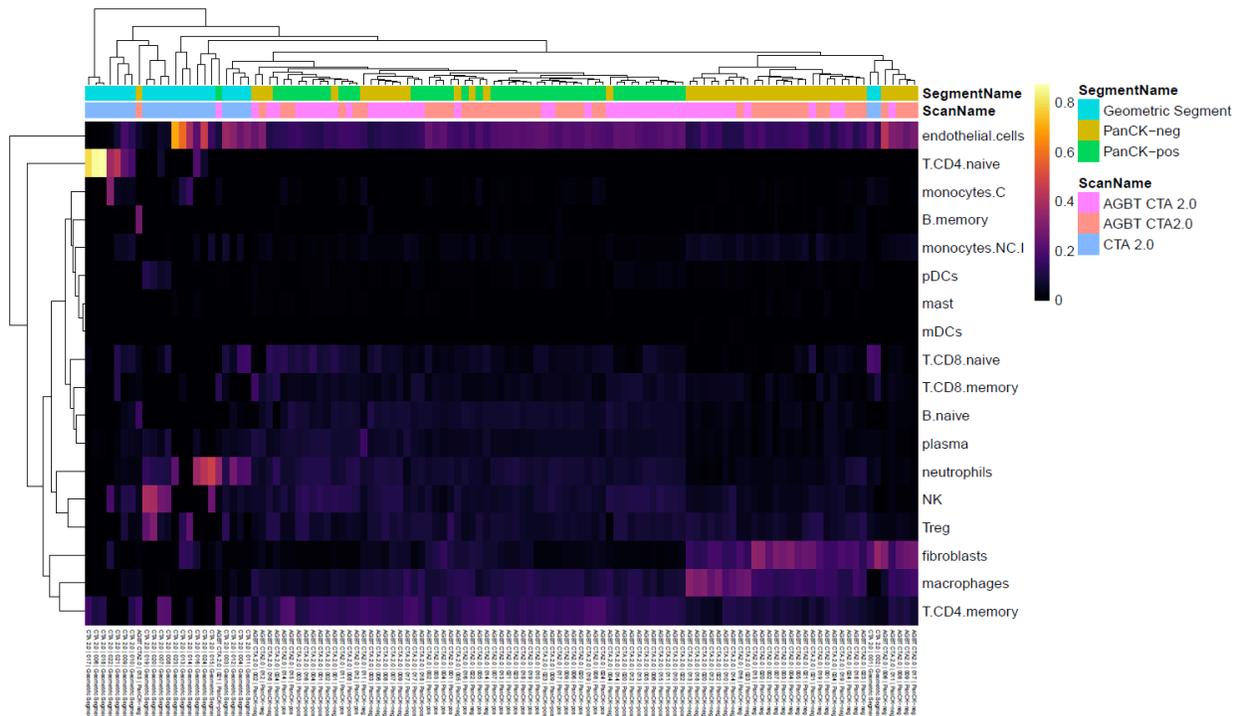
*Segment annotations tab*: gives the segment annotation data. Rows in this tab are aligned to columns in the other tabs.

### Plots:

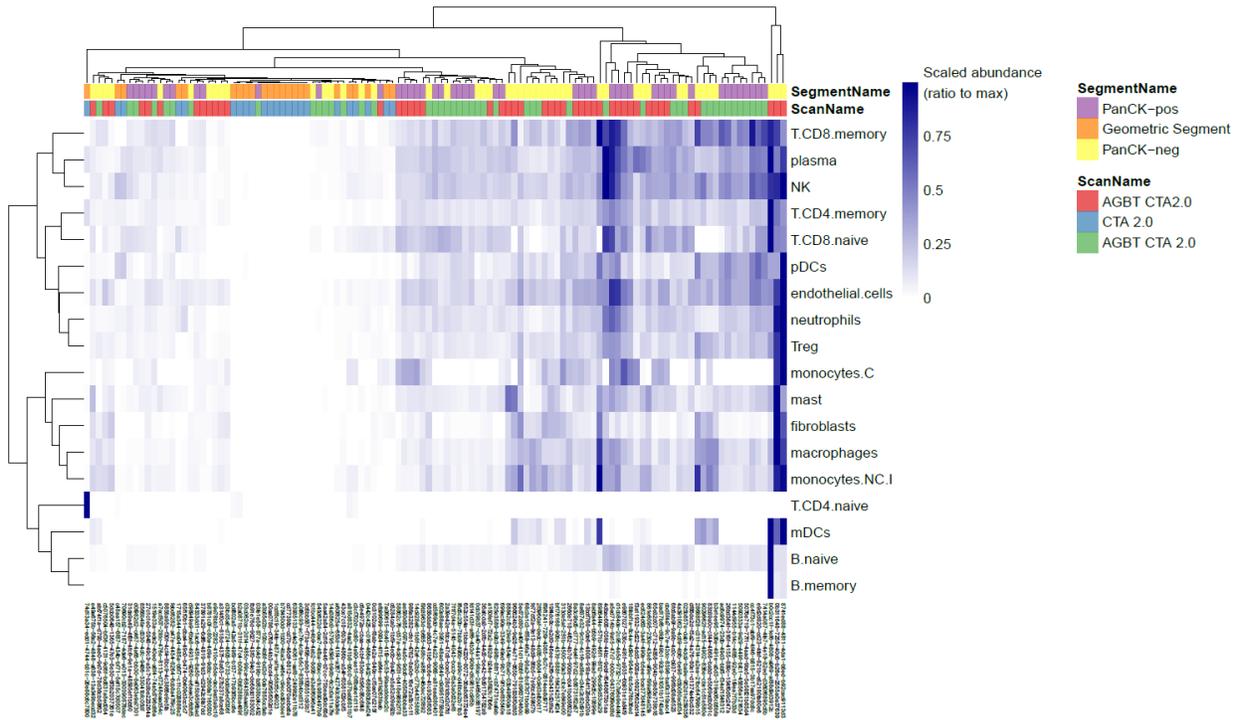
*Cell abundance scores heatmap*: Shows the results recorded in the "abundance scores" tab of the data tables output. Example below.



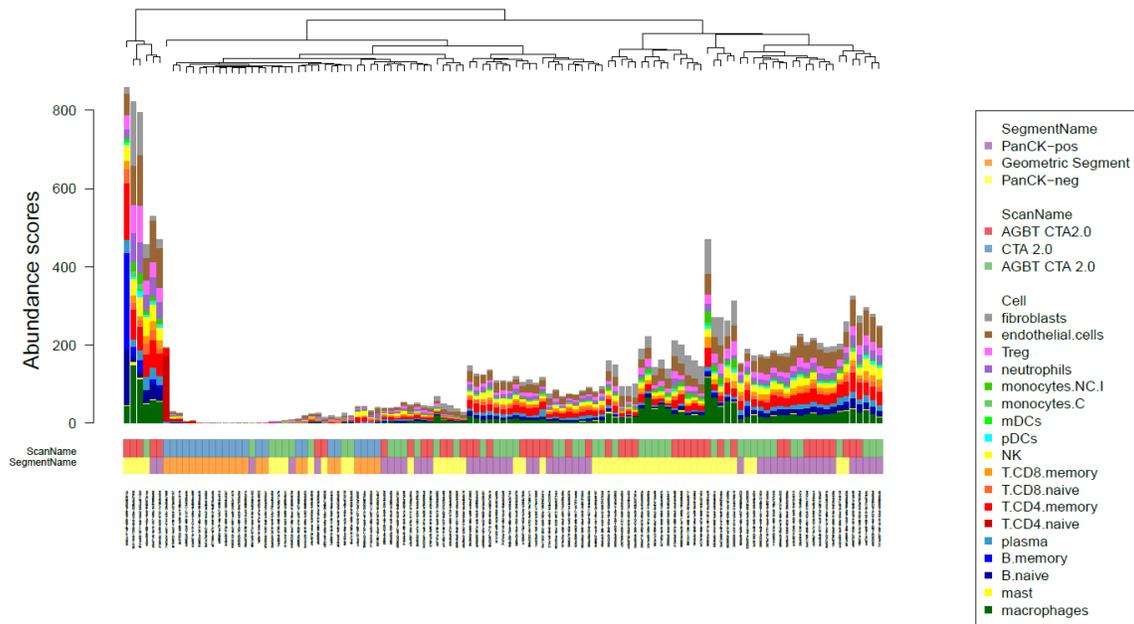
*Cell proportions:* Shows the results recorded in the “proportions” tab of the data tables output. Example below. This color scheme is activated by using the “viridis option B” option in the hmcpls argument.



**Scaled abundance scores heatmap:** Shows the results recorded in the “scaled abundance scores” tab of the data tables output. Example below.



**Cell abundance scores barplot:** See below for an example. Each column shows the cell abundances from a single AOI, with color denoting cell type.



*Cell proportions barplot:* See below for an example. Each column shows the cell abundances from a single AOI, with color denoting cell type.

