

If you like Stator and use it, please consider citing the related article:

High order expression dependencies finely resolve cryptic states and subtypes in single cell data Abel Jansma, Yuelin Yao, et al., *BioRxiv*

**Stator** application code is available through GitHub

If you have any question, you can send an e-mail to

● About    Table    Heatmap-Cells    Heatmap-Genes    GO & KEGG    Using rrvg

Upset Plot    DE analysis    Find Markers    Automatic annotations    Markov Blanket

Dendrogram    2D Plot

## Explore Stator states by App

**Stator App** takes in scRNA-seq count matrix, estimated higher-order gene interactions from Stator nextflow pipeline and defines Stator states with user-specific settings. Here we show how explore Stator states.

### Data Visualization & Analysis

- Table - A summary statistics for deviating states.
- Heatmaps Cells - Over-representation test for MFTs and other external cell annotations.
- Heatmaps Genes - Over-representation test for cell-state genes and external gene annotations.
- GO & KEGG - For genes in each state.
- rrvgo - Simplifying the redundancy of GO sets.
- Upset Plot - Visualisation of how many cells sharing among states.
- DE analysis - Differential expression analysis for two mutually exclusive states.
- Find Markers - Identify marker genes for a give cell state.
- Automatic annotations - Identify marker genes for all Stator states and return DEGs in the provided gene list.
- 2D Plot - Visualisation of UMAP for given a cell state.
- Dendrogram - Visualisation of d-tuples dendrogram.

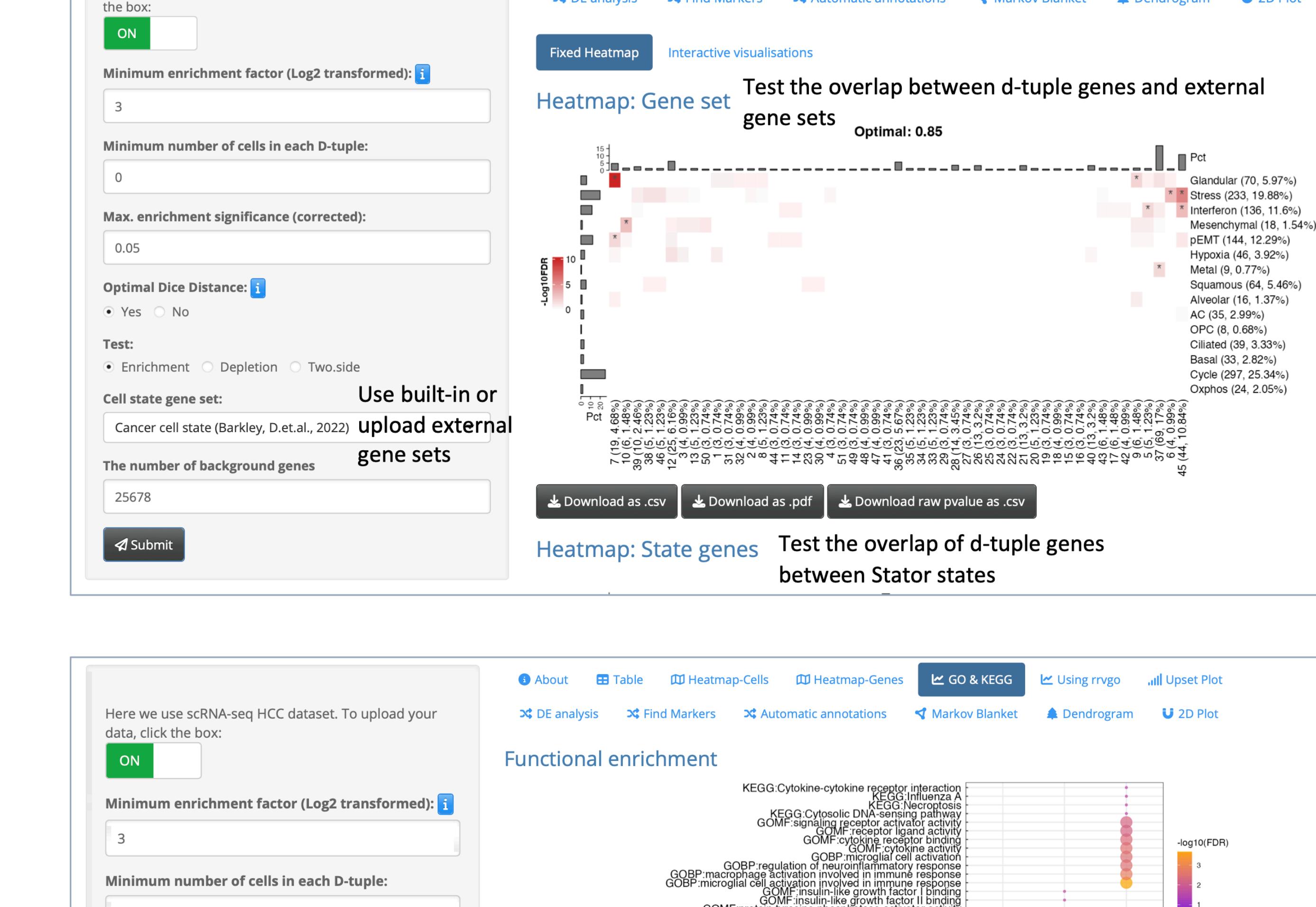
### Tutorial

#### Prepare and upload the dataset

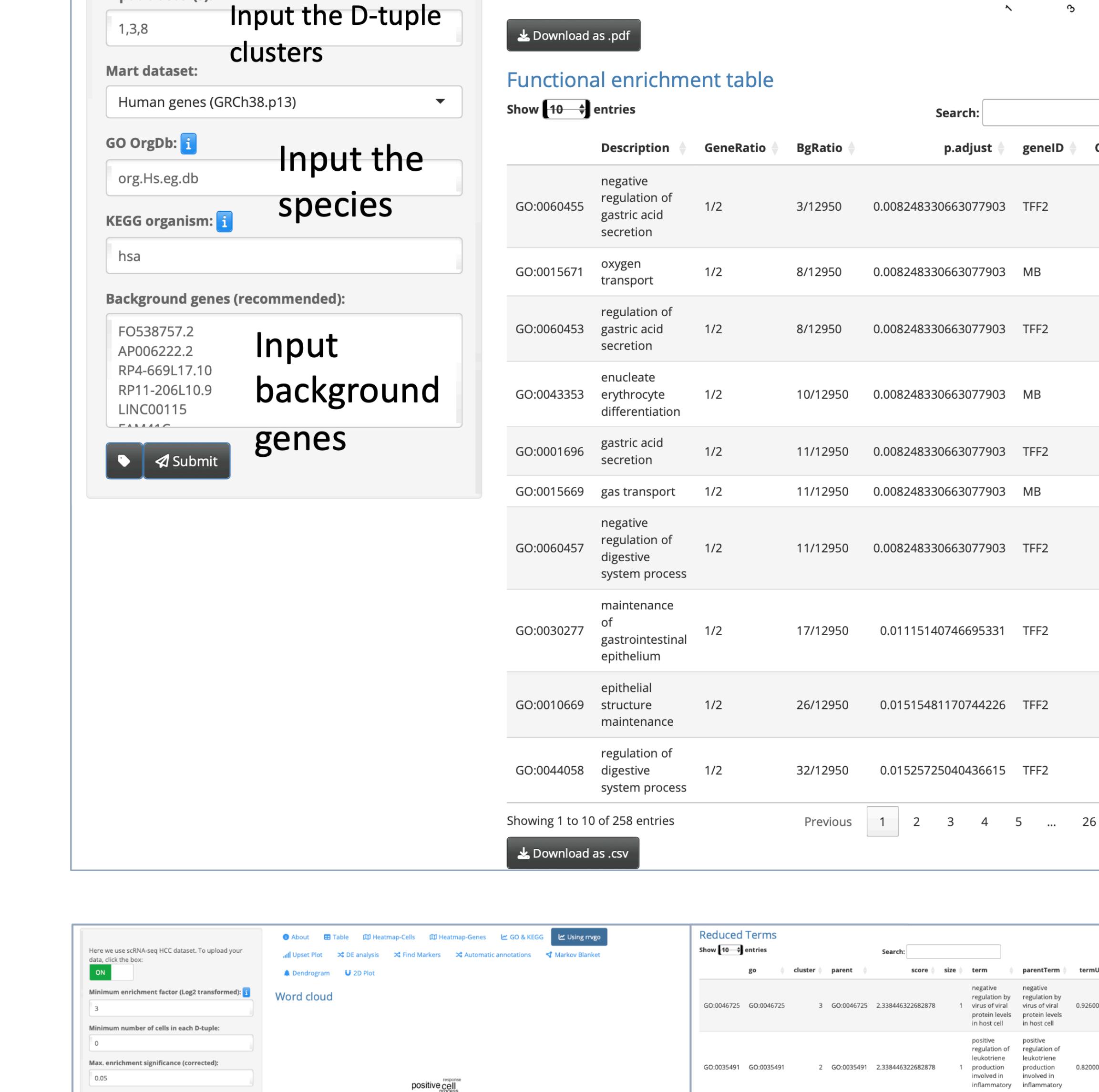
The input files include:

- Count\_matrix.csv: count matrix of scRNA-seq data, it can be the same file indicated in the [rawDataPath] in Stator nextflow JSON file. It is used for finding cells that satisfy the d-tuples within a group and differential expression analysis.
- Meta\_data.csv: cell's external annotations (e.g., from other tools clustering, NMFs, or experimental conditions. Two column csv file, example in ./data/ folder), is only used in the overrepresentation test heatmap (heatmap-cell tabs), if you don't want to plot this heatmap, you can skip this file.
- all\_Tuples.csv: located in the output from Stator nextflow pipeline (duplicates\_output folder)
- trainingData.csv: which is also in the output from Stator nextflow pipeline (output folder)
- GeneAnnotationSet.csv (optional): which is used in heatmap-genes, over-representation test between gene list. Example file format see ./data/CancerState.csv
- UMAP.csv (optional): which is used in 2D plot. Example file format see ./data/UMAP\_coords.csv
- The example of file can be found at: [here](#)
- MCMGraph.csv (optional): which is used in Markov Blanket Tab, located in the output folder from Stator nextflow pipeline. Example file format see ./data/MCMGraph\_14698Cells\_1000Genes.csv

#### Summary statistics: Table



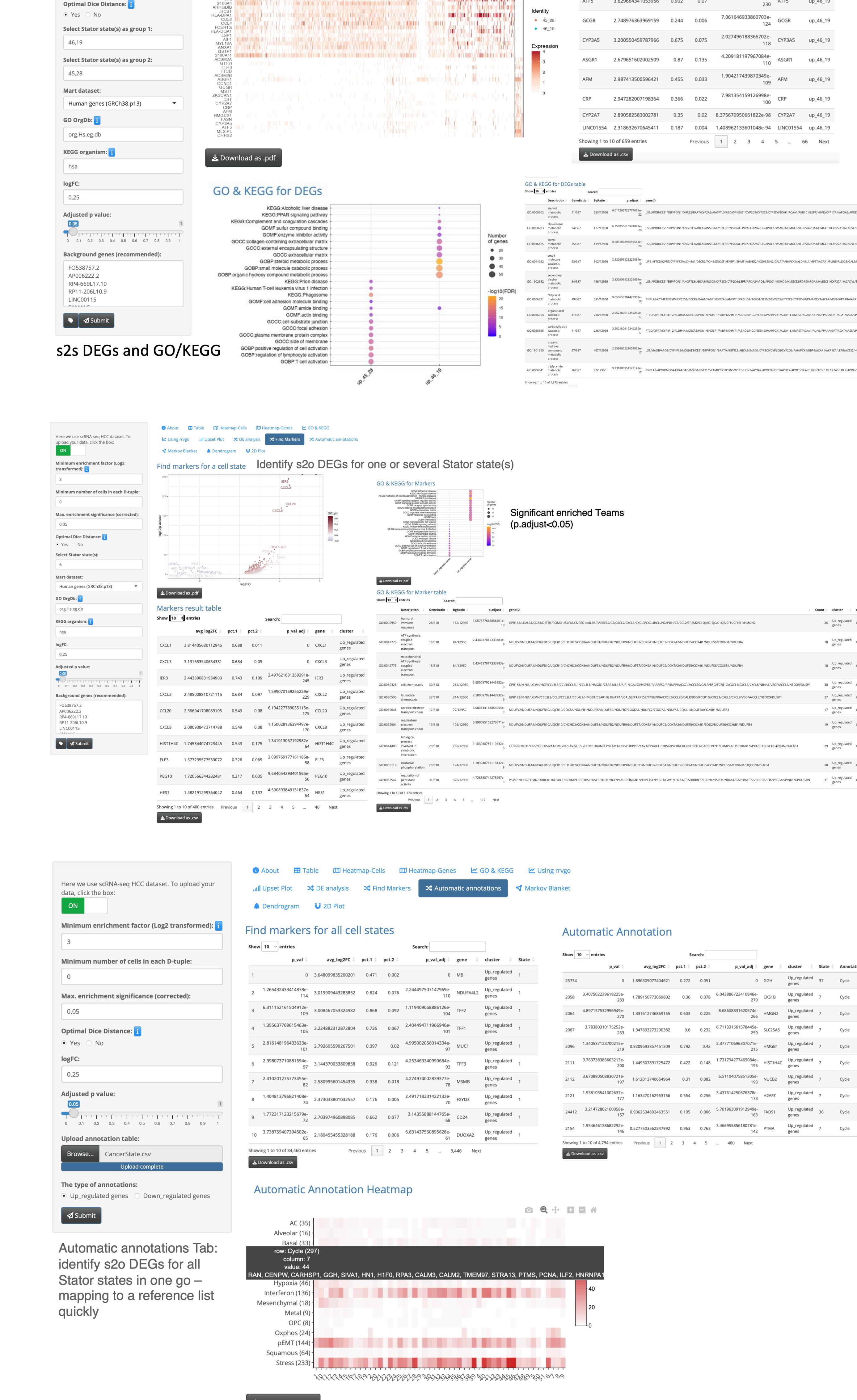
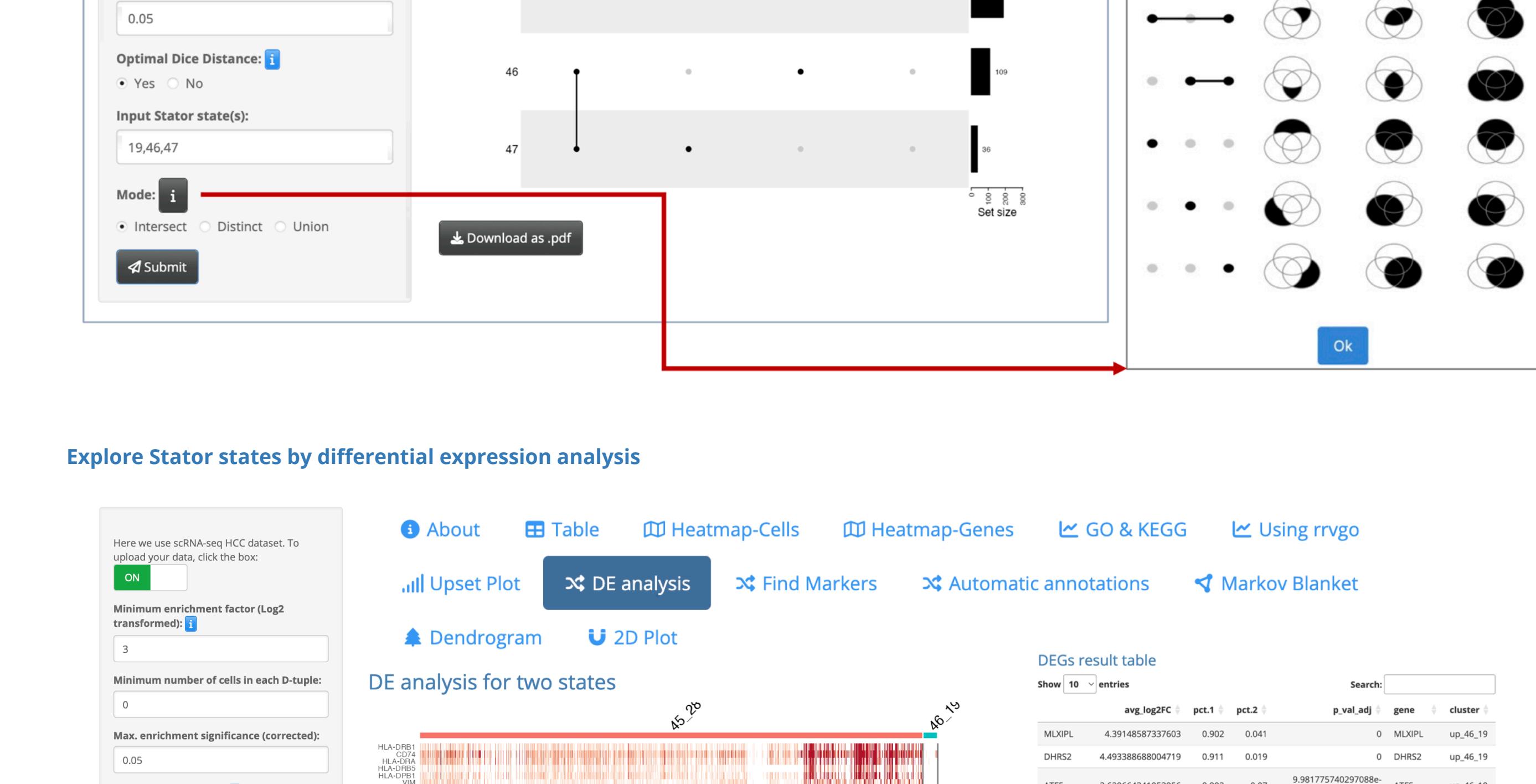
#### Extract phenotype-related Stator states: Heatmap-cell



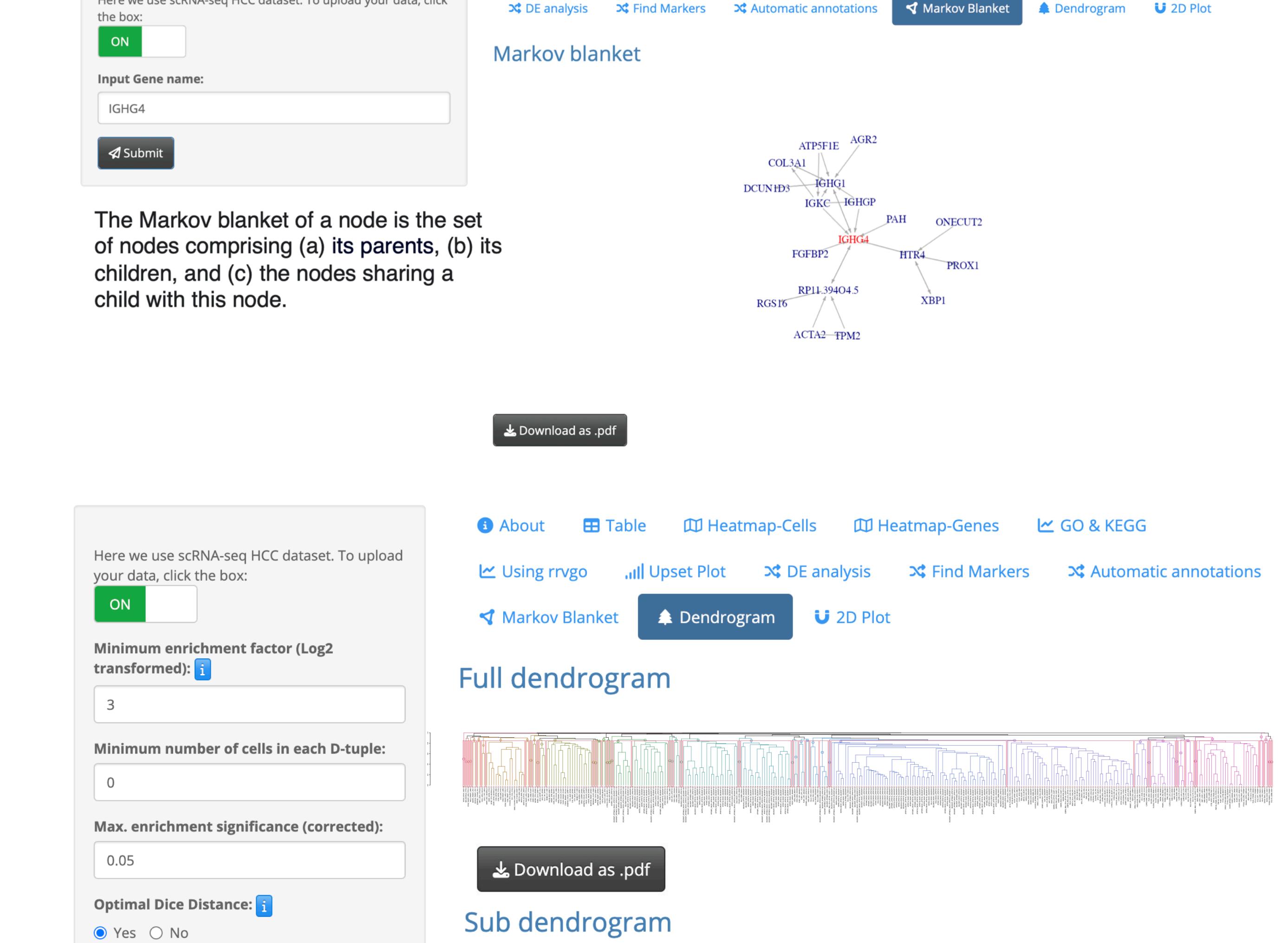
#### Test overlap (e.g., hypergeometric test or Fisher's test) between Stator states and external annotations:

- Experimental condition: e.g., health/disease, mutant/wild-type
- Developmental time point
- Tissues
- ...

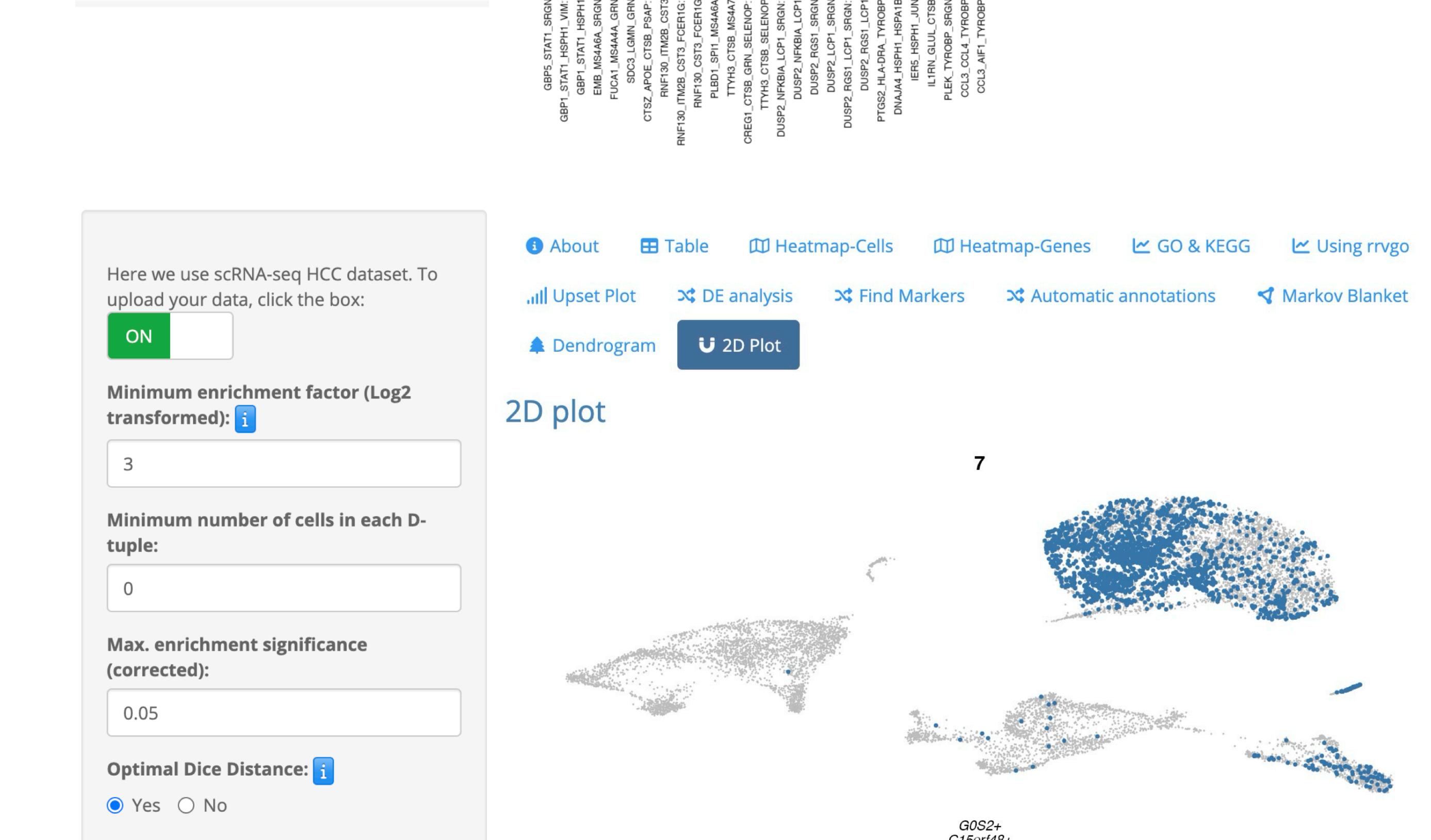
#### Explore d-tuple genes: Heatmap-Genes, GO&KEGG, Using rrvg



#### Cells cells sharing among states: Upset Plot



#### Explore Stator states by differential expression analysis



#### Other tools

