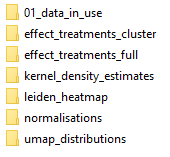
Manual for the codes

# General lay-out of the folders

All codes have been placed in a special folder called ‘R’. When using the different codes make sure that your working directory is the folder **above this folder** (*i.e.* the folder in which this manual is located). If this has been done correctly, you should be able to run all codes without errors and get the same tables and figures as I did.

All coding work has been subdivided into smaller bits, aimed at doing a small bit of the total analysis. These can be found in the different subfolders of the ‘R’ folder. The different folders are as follows:



Of these the folder ‘01\_data\_in\_use’ is special, in that it does not contain code, but that it is the storage place of the data frames all codes use. If you update one of the different analyses, leading to a different outcome, make sure to update the file in this folder! Otherwise your new analysis will not be carried over into the next.

All other folders contain parts of the analysis and each is subdivided into an ‘analysis’, ‘clean\_data’ and ‘figures’ folder:



The analysis folder is the one in which the codes themselves are stored. The code is usually split in an at least two parts:

* the libraries and variables;
* the code (in one or more files).

The clean\_data and the figures folders are filled with subfolders with the data on which an analysis is performed and these folders are filled with the output tables and figures, respectively.

# General use of the codes

As mentioned before, all code has been split into at least two parts: the libraries and variables part and the code itself. Make sure you clear the ‘Environment’ in R before running anything. Once the ‘Environment’ is cleared, run the entirety of the libraries and variables part. As you might notice, this is the part where you can change many of the variables for the following code. Also make sure you set the options ‘save\_figure’ and ‘save\_table’ to either “yes” or “no”, depending on what you whish to do with the output[[1]](#footnote-1).

Once the libraries and variables are all loaded, you can select the code part and run this in its totality. If all is well, the output should be stored in the clean\_data and figures folder if so desired.

# Analyses in the folders

## effect\_treatments\_cluster

In this folder you will find all the analyses on the cluster level. These are subdivided into analysis on:

* the antibody level, looking at differences between clusters;

This creates a heatmap for the different antibodies per cluster per cell, as well as feature plots and boxplots for individual antibodies

* the antibody level, looking at the differences between treatments;

This creates dotplots for the different fold changes per cluster when comparing vehicle with treatment, and another analysis leading to boxplots where the differences between clusters are compared

* the cell and signalling state level, looking at both the differences between clusters and treatments in one analysis.

This creates figures first detailing baseline signalling and cell state scores (based on the vehicle treated cells) and later fold changes of these scores upon treatment.

## effect\_treatments\_full

In this folder you will find all the analyses on the population level. These are subdivided into analysis on:

* finding the cell state markers, this is based on the Kruskal-Wallis test in order to compare all different treatments with each other;
* creating violin plots where one treatment is compared with the vehicle;
* creating volcano plots and finding the fold changes of each antibody upon a treatment on a population level.

## kernel\_density\_estimates

In this folder you will find the code to determine the kernel density estimates.

## leiden\_heatmap

In this folder you will find the code to determine the consensus clusters, as well as figures detailing the process. It is split up in two sperate parts. In the first part the custom clustering pipeline we came up with is run for 5 times. In the second part these different iterations are compared. It has outputs projected on umaps with assigned clusters, the distribution of said clusters over the different treatments, ambiguity of the cluster assignments and comparisons between the different iterations.

## normalisations

In this folder you will find the different normalisations as executed. These contain the following:

* a code to add the consensus clusters back to the TMM normalised data (prior to z-score normalisation);
* a code to normalise the values of all cells in a cluster on its cluster’s mean;
* a code the normalise the values of all cells in a cluster on its cluster’s median;
* a code to z-score normalise all values for the different antibodies.

## umap\_distributions

In this folder you will find a code that will lead to different umap plots, most importantly the different feature plots, but also distribution, cluster assignment and treatment plots.

1. When save\_figure == “no”, the figures are generally printed in the plots window of R-studio. [↑](#footnote-ref-1)