

Rodeo manual

1 Rodeo description

Rodeo (RObust DEcOnvolution) is a robust deconvolution tool that provides estimates of how strongly each cell type expresses each gene based on bulk expression data and cell type proportions. The current version 1.0 of the package includes one function intended for external usage: Rodeo. Please contact us (maria.jaakkola@utu.fi) in case of bugs, missing documentation, or problems with installation.

2 Installation and usage

Rodeo source package Rodeo_1.0.tar.gz can be downloaded from <https://elolab.utu.fi/software/> and also package MASS should be installed (comes with R by default) and activated with R command `library(MASS)` before using Rodeo. Rodeo is also available at github <https://github.com/elolab/Rodeo>. In case of installing from github, downloading the source package is not needed. There are several ways to install Rodeo:

Using command prompt:

Download the source package Rodeo_1.0.tar.gz and open command prompt. Navigate with command prompt to where the source package was saved. Type: `R CMD INSTALL Rodeo_1.0.tar.gz`

Using RStudio:

Download the source package Rodeo_1.0.tar.gz and open RStudio. Go to section 'Packages' in the bottom right window and select 'Install'. From the pop-up window set the option 'Install from' to option 'Package Archive File (.zip; .tar.gz)' using the drop down menu. Then set the path (will be asked automatically) to where you saved the source file and press 'Install'.

Using github and R:

Open R and type: `devtools::install_github("elolab/Rodeo")`

This requires package devtools to be installed.

Below is a code example of how to run Rodeo after it has been installed. The example is very simplistic and can not be directly copy pasted if e.g. the input data is stored differently.

```
# Load required packages
library(MASS)
library(Rodeo)

# Read Input data
setwd("path/to/my/input/files")
E = read.table("BulkData.txt", header=T, row.names=1)
C = read.table("CellTypeProportions.txt", header=T, row.names=1)

# Run Rodeo and save the results
S = Rodeo(E, C)
write.table(S, file="EstimatedS.txt", sep="\t", quote=F)
```

3 Input

- E** A bulk expression matrix with named rows (genes) and columns (samples)
- C** A cell type proportion matrix with named rows (cell types) and columns (samples) for the samples in **E**. Each sample (i.e. column) should sum to 1.

4 Output

The main function (and the only one meant for external use) Rodeo returns a matrix **S** where columns are cell types, rows are genes, and elements describe how strongly the cell type expresses the gene.