

## Exercise 5

### Osimertinib experiment

- PC9 (EGFR mutated, sensitive to Osimertinib)
- MOCK (DMSO)
- Acute 500 nM osimertinib for 24 hrs
- drug tolerant persister (DTP) 500 nM osimertinib for 21 days
- We have the count table for the RNAseq counts.
- Exercise 1:
  - Install in Docker edgeR, which is a Bioconductor package and use the cpm function implemented in edgeR

home.

- 5) Give the permission to run the file with `chmod +x edgeR.R`
- 6) Make it executable with `Rscript edgeR.R` → it requires some time to install edgeR

- 1) Open RStudio and write a file, named “edgeR.R”, with this string, and save it in a folder, such as “lesson5\_data\_analysis”, on the desktop.
- 2) Copy the path of this file on the home of the docker.
- 3) Do a commit
- 4) Run the docker and verify that the file is present in the

```
edgeR.R* x L5_script.R x
1 if (!require("BiocManager", quietly = TRUE))
2   install.packages("BiocManager")
3   BiocManager::install("edgeR")
4
```

```
Windows PowerShell
Windows PowerShell
Copyright (C) Microsoft Corporation. Tutti i diritti riservati.

Prova la nuova PowerShell multiplatforma https://aka.ms/pscore6

PS C:\Users\cauci> docker run -i -t r4:v.0.01
root@fedf380f2cc0:/# cd /home
root@fedf380f2cc0:/home# ls
command_umap.R  docker
root@fedf380f2cc0:/home# cd ..
root@fedf380f2cc0:/# exit
exit
PS C:\Users\cauci> cd C:\Users\cauci\Desktop
PS C:\Users\cauci\Desktop> ls
```

```
PS C:\Users\cauci\Desktop> docker cp C:/Users/cauci/Desktop/lesson5_data_analysis/edgeR.R fedf380f2cc0:/home/
PS C:\Users\cauci\Desktop> docker commit fedf380f2cc0 r4:v.0.01
sha256:e086ccd76852337eab3f9e9f8bbaad4423e100da627a25f484b471b7b107fc06
PS C:\Users\cauci\Desktop> docker run -i -t r4:v.0.01
root@c51f3307a4a8:/# cd /home
root@c51f3307a4a8:/home# ls
command_umap.R  docker  edgeR.R
root@c51f3307a4a8:/home# chmod +x edgeR.R
root@c51f3307a4a8:/home# Rscript edgeR.R
Installing package into '/usr/local/lib/R/site-library'
(as 'lib' is unspecified)
trying URL 'https://cloud.r-project.org/src/contrib/BiocManager_1.30.17.tar.gz'
Content type 'application/x-gzip' length 287948 bytes (281 KB)
```

- 7) Now, move to RStudio and run the edgeR.R command, in order to easily see the eventual errors. Write a script L5\_script.R with all the commands that are necessary to convert the dataset into a matrix containing the CPM values. Remember to set as working directory the folder where the file is located. In the exercise, the L5dataset.txt has been converted into L5cpm.txt.

```
edgeR.R* x L5_script.R x
1 getwd()
2 x <- read.table("L5dataset.txt", sep = "\t", header = T, row.names = 1)
3 class(x)
4 # Now, I convert the data frame into a numeric matrix
5 y <- data.matrix(x, rownames.force = NA)
6 library(edgeR)
7 z <- edgeR::cpm(y)
8 class(z)
9 head(y)
10 head(z)
11 write.table(z, file = "L5cpm.txt", sep = "\t", col.names = NA)
12
```

Once this has been done, I can make the same thing on the docker.

- 8) Mount the folder lesson5\_data\_analysis on the folder data\_analysis, that I create in the home of the docker.
- 9) Do a commit.

```

root@83399fac1430:/home# cd ..
root@83399fac1430:/# exit
exit
PS C:\Users\cauci\Desktop> docker run -i -t -v C:/Users/cauci/Desktop/lesson5_data_analysis:/home/data_analysis r4:v.0.0
root@5b2a0d73f045:/# cd /home
root@5b2a0d73f045:/home# ls
command_umap.R  docker  edgeR.R
root@5b2a0d73f045:/home# cd data_analysis
root@5b2a0d73f045:/home/data_analysis# ls
edgeR.R  L5cpm.txt  L5dataset.txt  L5_script.R
root@5b2a0d73f045:/home/data_analysis# cd ..
root@5b2a0d73f045:/home# cd ..
root@5b2a0d73f045:/# exit
exit
PS C:\Users\cauci\Desktop> docker commit 5b2a0d73f045 r4:v.0.01
sha256:9f16200a12c8c95cc1bb8cc8e090004ca66a9a69bc93ba785ee7c4474cdb7da5
PS C:\Users\cauci\Desktop>

```

- 10) In order to install devtools, I have to create a file.sh (that I call "command\_devtools.sh") with nano: here, I write the following code without RUN. Save it with Ctrl+s → exit with Ctrl+x. Make it executable with `chmod +x command_devtools.sh` → run it with `bash command_devtools.sh`. In the same folder, write a .R file, called devtools.R, with nano → the content is `install.packages("dev.tools")` → save it → run it with `Rscript devtools.R`. Both passages require some time!

```

RUN apt-get update
RUN apt-get update
RUN apt-get install -y \
  apt-transport-https \
  ca-certificates \
  curl \
  gnupg \
  lsb-release
RUN apt-get update && apt install -y libudunits2-dev libgdal-dev
RUN apt-get update
RUN apt-get -y install gfortran
RUN apt-get -y install build-essential
RUN apt-get -y install fort77
RUN apt-get -y install xorg-dev
RUN apt-get -y install libzmq-dev libblas-dev gfortran
RUN apt-get -y install gobjc++
RUN apt-get -y install aptitude
RUN apt-get -y install libbz2-dev
RUN apt-get -y install libpcre3-dev
RUN aptitude -y install libreadline-dev
RUN apt-get -y install libcurl4-openssl-dev

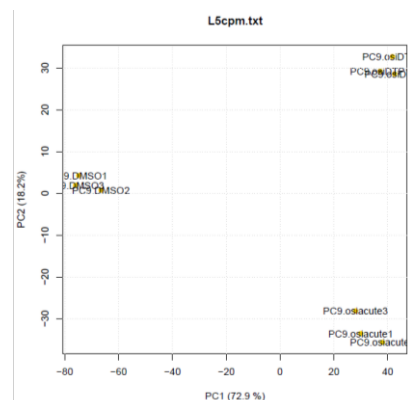
```

- 11) Finally, I install docker4seq in the same way I installed the previous packages, copying and pasting the string from <https://github.com/kendomaniac/docker4seq> → ctrl+s → ctrl+x → `Rscript docker4seq.R`
- 12) Run the pca function on the L5cpm.txt file. To do so, create a pca.R file in the "lesson5\_data\_analysis" folder and set these parameters (remember to set the folder as my working directory):

```

1 #pca function
2 library(docker4seq)
3 pca(experiment.table = "L5cpm.txt", type = "counts", covariatesInNames = F,
4     samplesName = T, principal.components = c(1, 2), legend.position = "topright",
5     pdf = T, output.folder = getwd())
6

```



As output, a PDF file will be created; in my case, it is like that: