mariagiuli.bacalini2@unibo.it

Instructions:

- For each step of the analytical pipeline provide the R code you used
- In some steps, student specific instructions are given; refer to the tables below for assignments

Student	**Surname**
1	Davide Abbondandolo
2	Tommaso Becchi
3	Federica Brando
4	Alessia Campo
5	Ludovica Cataneo
6	Alessandro Caula
7	Gaia Cervi
8	Simone Del Motto
9	Immanuela Antigone Engländer
10	Giorgia Gandolfi
11	Ana Cristina Gonzalez Sanchez
12	Fidan Gurbanova
13	Sahar Heidaribakavoli
14	Biagio Iacolare
15	Thomas James Isaac
16	Naiara Landeta González
17	Davide Lisi
18	Daniele Lucarelli
19	Gennaro Luciano
20	Vanessa Mucci
21	Matteo Orlandi
22	Fabiana Patalano
23	Lorenzo Pedroni
24	Saul Pierotti
25	Ilaria Pirona
26	Stefano Roncelli
27	Aigerim Rymbekova
28	Andrea Sambugaro
29	Óscar San José Rodríguez
30	Rosaria Tornisiello
31	

Pipeline steps

Step 1
Load raw data with minfi and create an object called RGset storing the RGChannelSet object

Step 2
Create the dataframes Red and Green to store the red and green fluorescences respectively

Step 3
Fill the following table: what are the Red and Green fluorescences for the address assigned to you? **Optional**: check in the manifest file if the address corresponds to a Type I or a Type II probe and, in case of Type I probe, report its color.

Sample	**Red fluor**	**Green fluor**	**Type**	**Color**

Student	Address
3	42639338

Step 4
Create the object MSet.raw

Step 5
1. Perform the following quality checks and provide a brief comment to each step:
   - QCplot
   - check the intensity of negative controls using minfi
   - calculate detection pValues; for each sample, how many probes have a detection p-value higher than the threshold assigned to each student?

Sample	**Failed positions**

Student	detPvalue threshold
3	0.05

Step 6
Calculate raw beta and M values and plot the densities of mean methylation values, dividing the samples in DS and WT (*suggestion: subset the beta and M values matrixes in order to retain DS or WT subjects and apply the function mean to the 2 subsets*).

Step 7
Normalize the data using the function assigned to each student and compare raw data and normalized data. Produce a plot with 6 panels in which, for both raw and normalized data, you show the density plots of beta mean values according to the chemistry of the probes, the density plot of beta standard deviation values according to the chemistry of the probes and the boxplot of beta values. Provide a short comment regarding the changes you observe.

Student	Normalization
3	preprocessQuantile

Step 8
Perform a PCA on the beta matrix generated in step 7. Comment the plot.

Step 9
Using the matrix of normalized beta values generated in step 7, identify differentially methylated probes between group DS and group WT using the functions assigned to each student. Note; it can take several minutes; if you encounter any problem you can run the differential methylated analysis only on a subset of probes (for example those on chromosome 1, 18 and 21)

Student	Normalization
3	t-test

Step 10
Apply multiple test correction and set a significant threshold of 0.05. How many probes do you identify as differentially methylated considering nominal pValues? How many after Bonferroni correction? How many after BH correction?

Step 11
Produce an heatmap of the top 100 differentially mehtylated probes

Step 12
Produce a volcano plot and a Manhattan plot of the results of differential methylation analysis

Optional

As DS is caused by the trisomy of chromosome 21, try also to plot the density of the methylation values of the probes mapping on chromosome 21. Do you see a very clear difference between the samples? How many differentially methylated probes do you find on chromosome 21?