mariagiuli.bacalini2@unibo.it
Via San Giacomo 12, secondo piano (Lab Franceschi)
051 2094748

Reference paper
https://www.ncbi.nlm.nih.gov/pubmed/25701644

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	ALKHANSA
2	BOLNER
3	BUZZAO
4	CICCONARDI
5	DE BERNARDINI
6	FANALISTA
7	FERRI
8	FOGLINI
9	FRATINI
10	IBISHOV
11	LUI
12	LUSSANA
13	MEI
14	MESTIZIA
15	MORLINO
16	MOSCA
17	OTTALEVI
18	PACE
19	PALMACCI
20	POLVERELLI MONTI
21	RAMBALDELLI
22	VIGNOLI
23	YAKHONTOV
24	ZUCCHI
25	KOVAL
26	KAHVECI

Step 3

Student	Address

1	42639338
2	10633381
3	10737353
4	61760464
5	71773431
6	39802405
7	64638362
8	18744490
9	72740437
10	31763489
11	41620492
12	29643447
13	10715421
14	13673406
15	18756452
16	59625465
17	45652402
18	59609499
19	25710468
20	52682510
21	20750401
22	64689504
23	46801437
23	21695377
25	32792369
26	44666390

Step 5

Student	detPvalue threshold
1	0.05
2	0.05
3	0.05
4	0.05
5	0.05
6	0.05
7	0.05
8	0.05
9	0.05
10	0.05
11	0.05
12	0.05

13	0.05
14	0.01
15	0.01
16	0.01
17	0.01
18	0.01
19	0.01
20	0.01
21	0.01
22	0.01
23	0.01
24	0.01
25	0.01
26	0.01

Step 7

Student	Normalization
1	preprocessNoob
2	preprocessNoob
3	preprocessNoob
4	preprocessNoob
5	preprocessNoob
6	preprocessSWAN
7	preprocessSWAN
8	preprocessSWAN
9	preprocessSWAN
10	preprocessSWAN
11	preprocessQuantile
12	preprocessQuantile
13	preprocessQuantile
14	preprocessQuantile
15	preprocessQuantile
16	preprocessFunnorm
17	preprocessFunnorm
18	preprocessFunnorm
19	preprocessFunnorm
20	preprocessFunnorm
21	preprocessFunnorm
22	preprocessFunnorm
23	preprocessNoob
24	preprocessSWAN
25	preprocessQuantile
26	preprocessFunnorm

Step 10

Student	Test
1	Anova
2	Anova correcting for Female
3	Mann-Whitney Test
4	Anova
5	Anova correcting for Female
6	Mann-Whitney Test
7	Anova
8	Anova correcting for Female
9	Mann-Whitney Test
10	Anova
11	Anova correcting for Female
12	Mann-Whitney Test
13	Anova
14	Anova correcting for Female
15	Mann-Whitney Test
16	Anova
17	Anova correcting for Female
18	Mann-Whitney Test
19	Anova
20	Anova correcting for Female
21	Mann-Whitney Test
22	Anova
23	Anova correcting for Female
24	Mann-Whitney Test
25	Anova
26	Anova correcting for Female]

Steps of the analysis pipeline

1. Load raw data with minfi and create an object called RGset storing the RGChannelSet object *(see Script Lesson3_bis_2019)*

2. Create the dataframes Red and Green to store the red and green fluorescence respectively

3. Fill the following table: what are the Red and Green fluorescence for the address assigned to you? (*suggestion: subset the Red and Green dataframes by setting rownames=="the Address you want")*. Optional: check from 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

4. Create the object MSet.raw

5. 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

   - create a vector of bad probes to be removed because they have a detection pValues higher that the assigned threshold (see above) in more than 1% of the samples.

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*).

7. Compare raw data with normalized data using the function assigned to each student (*see script Lesson_4_2019)*. Produce a plot with 6 panels (2 rows, 3 columns) 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.

8. Consider the normalized beta and M values and remove the probes previously defined as bad according to the detection pValue (*see script Lesson5_2019)*

9. Check beta and M matrixes generated in step 8 for homo/heteroschedasticity (*see script Lesson5_2019)*; comment the plot. Optional: plot the lowess line.

10. Perform a PCA on the beta matrix generated in step 8 to check for batch effects (*see script Lesson6_2019).* Comment the plot.

11. Using the beta matrix generated in step 8, identify differentially methylated probes between group A and group B using the functions assigned to each student (it will take several minutes) (*see script Lesson5_bis_2019)*. 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)

12. 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?

13. Produce a Manhattan plot and a volcano plot of your data (*see script Lesson6_2019)*

14. Produce an heatmap of the top 100 differentially mehtylated probes (*see script Lesson6_2019)*

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 (*suggestion: merge the beta values dataframe with the annotation dataframe,*

as we have done in lesson 6; then subset only 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?