[mariagiuli.bacalini2@unibo.it](mailto: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 | ALBERTI |
| 2 | AOF |
| 3 | BETTAZZI |
| 4 | BOJOVIĆ |
| 5 | BRICIC |
| 6 | CROITORU BRUCHENTAL |
| 7 | EL BOULAMI |
| 8 | ENRIQUEZ SANDOVAL |
| 9 | FORCONI |
| 10 | IARDUKHIN |
| 11 | LO CASCIO |
| 12 | MALATESTA |
| 13 | MASTROGIOVANNI |
| 14 | MIM |
| 15 | MOKHTARI |
| 16 | OSMANAGAOGLU |
| 17 | PICCOLO |
| 18 | POSTIGLIONE |
| 19 | RIVI |
| 20 | SERRAZANETTI |
| 21 | TACCALITI |
| 22 | VELIMIROVIC |
| 23 | VUONG |

**Step 3**

|  |  |
| --- | --- |
| Student | Address |
| 1 | 18756452 |
| 2 | 71773431 |
| 3 | 42639338 |
| 4 | 46801437 |
| 5 | 13673406 |
| 6 | 61760464 |
| 7 | 18744490 |
| 8 | 64638362 |
| 9 | 10737353 |
| 10 | 39802405 |
| 11 | 18756452 |
| 12 | 29643447 |
| 13 | 31763489 |
| 14 | 44666390 |
| 15 | 45652402 |
| 16 | 45652402 |
| 17 | 10715421 |
| 18 | 25710468 |
| 19 | 21695377 |
| 20 | 64689504 |
| 21 | 42639338 |
| 22 | 46801437 |
| 23 | 13673406 |

**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 | 0.01 |
| 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 |

**Step 7**

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

**Step 9**

|  |  |
| --- | --- |
| Student | Test |
| 1 | t-test |
| 2 | Mann-Whitney test |
| 3 | t-test |
| 4 | Mann-Whitney test |
| 5 | t-test |
| 6 | Mann-Whitney test |
| 7 | t-test |
| 8 | Mann-Whitney test |
| 9 | t-test |
| 10 | Mann-Whitney test |
| 11 | t-test |
| 12 | Mann-Whitney test |
| 13 | t-test |
| 14 | Mann-Whitney test |
| 15 | t-test |
| 16 | Mann-Whitney test |
| 17 | t-test |
| 18 | Mann-Whitney test |
| 19 | t-test |
| 20 | Mann-Whitney test |
| 21 | t-test |
| 22 | Mann-Whitney test |
| 23 | t-test |

**Pipeline steps**

1. Load raw data with minfi and create an object called RGset storing the RGChannelSet object
2. Create the dataframes Red and Green to store the red and green fluorescences respectively
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** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. Create the object MSet.raw
2. 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** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

1. Calculate raw beta and M values and plot the densities of mean methylation values, dividing the samples in WT and MUT (*suggestion: subset the beta and M values matrixes in order to retain WT or MUT subjects and apply the function mean to the 2 subsets*). Do you see any difference between the two groups?
2. 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 about the changes you observe. *Optional: do you think that the normalization approach that you used is appropriate considering this specific dataset? Try to color the boxplots according to the group (WT and MUT) and check whether the distribution of methylation values is different between the two groups, before and after normalization …*
3. Perform a PCA on the matrix of normalized beta values generated in step 7, after normalization*.* Comment the plot (do the samples divide according to the group? Do they divide according to the sex of the samples? Do they divide according to the batch, that is the column Sentrix\_ID?). *Only if we discussed PCA during Lesson 6!*
4. Using the matrix of normalized beta values generated in step 7, identify differentially methylated probes between group WT and group MUT using the function assigned to each student. *Note: it can take several minutes; if you encounter any problem you can run the differential methylation analysis only on a subset of probes (for example those on chromosome 1, 18 and 21)*
5. 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?
6. Produce a volcano plot and a Manhattan plot of the results of differential methylation analysis *Only if we discussed these plots during Lesson 6!*
7. Produce an heatmap of the top 100 differentially mehtylated probes *Only if we discussed these plots during Lesson 6!*