**MSB7102 Mini-project, Semester I, 2021**

**Background**

This is an assessed exercise that will contribute to coursework for this module. It is aimed at providing an experience with all that has been covered throughout the semester. The tasks are based on two Bioconductor packages; phyloseq and DEseq2. Endeavor to look at the documentation and the links indicated below may be useful. The data files are located in the shared google drive folder.

**Data sources and description**

The data used in this exercise is derived and was generated by Kolistic *et al* in their study “The dynamics of the human infant gut microbiome in development and progression toward Type 1 Diabetes”. See full publication at <https://doi.org/10.1016/j.chom.2015.01.001>. Briefly, this was a prospective analysis of developing gut microbiome in infants *en route* to type 1 diabetes. Infants from Finland and Estonia were recruited at birth based on predisposition to autoimmunity determined by human leukocyte antigen (HLA) genotyping. The cohort consists of 33 infants, 11 of which seroconverted to serum autoantibody positivity and of those, four developed T1D within the three-year time-frame of this study.

**Tasks**

1. Import the data described above into R, provide descriptive summaries of the subject data (using appropriate graphics and statistical summary measures) given in the *diabimmune\_16s\_t1d\_metadata.csv* file. In addition, use appropriate test(s) to check for association/independency between disease status and other variables (delivery mode, gender and age). Note that age is given in days.
2. Using phyloseq, create a phyloseq object. This will comprise the OTU abundance, taxonomy (provided in the *.txt* file) and sample data (provided in the *.csv* file).
3. Generate Alpha diversity plots and ordination plots. Examine any observed patterns by delivery mode, gender and disease status.
4. Perform differential abundance using DEseq2

**Useful links**

* Importing data: <https://joey711.github.io/phyloseq/import-data.html>
* Ordination: <https://joey711.github.io/phyloseq/plot_ordination-examples.html>
* Alpha diversity: <https://joey711.github.io/phyloseq/plot_richness-examples.html>
* Differential abundance: <http://joey711.github.io/phyloseq-extensions/DESeq2.html>

**Submission**

Use Rmarkdown to generate a report for your submission. Generate the report as a pdf or html file, upload it to your GitHub repository, send the link to [submissions.gronald@gmail.com](mailto:submissions.gronald@gmail.com) and copy [assekagiri@gmail.com](mailto:assekagiri@gmail.com) by **25th-06-2021** under the subject “**MSB7102 Mini-project, semester I, 2021**”.