

732A51 Bioinformatics

# LAB 4 Bioinformatics

Hugo Morvan  
William Wiik

STIMA  
Department of Computer and Information Science  
Linköpings universitet  
2024-12-05

## Contents

1	Question 1	1
2	Question 2	1
3	Question 3	1
4	Question 4	1

## 1 Question 1

Run all the R code and reproduce the graphics. Go carefully through the R code and explain in your words what each step does. HINT Recall what a design/model matrix is from linear regression.

## 2 Question 2

In the presented analysis, there are no plots of raw paired data. In the section where the contrasts are defined find the three contrasts. Present the variables versus each other original, log-scaled and MA-plot for each considered pair both before and after normalization. A cluster analysis is performed on the page but not reported. Present plots and also draw heatmaps.

## 3 Question 3

The volcano plot is only for huvec versus choroid. Provide volcano plots for the other pairs. Indicate significantly differentially expressed genes. Explain how they are found.

## 4 Question 4

Try to find more information on the genes that are reported to be significantly differentially expressed. The place to start off is <https://www.ncbi.nlm.nih.gov/gene/>, remember that the data is from the species human. Try to look also for other databases where (some) information on the genes may be found. Try to follow on some of the provided links. Report in your own words on what you find. Report all the Gene Ontology (GO) terms associated with each gene. Are any of the GO terms common between genes? If so do the common GO terms seem to be related to anything particular? Try to present GO analysis in an informative manner, if possible visualize.