1. Check whether you have the correct number of samples and proteins.

What status variables are available for the samples?

1. Yes loaded data in r and it matches

2. Perform the significance test for Spearman’s correlation coefficient

between Age and Dose (cor.test() function with Spearman method). Create

a scatter plot of the two variables in log scale. What can you say about the

correlation?

2.There seems to be correlation between the two variables and is positive in nature as the rho=0.7258823,also the p-value is 8.329e-06 for the null hypothesis that correlation=0

3. Create two linear models for each of the 1281 proteins for

status$Dose.mGy and status$Age. Calculate the BIC for each protein for

the two models. Create a vector modVec of length 1281, where you enter

“dose” if BIC for the dose model is smaller than for the age model, and

“age” otherwise.

1. In the r code

4. How many proteins are better explained by age than by dose? You can

remove them from further analysis:

4. Age Dose

422 859

422 are such where age explains better than dose out of 1281

5. Perform the Kruskal-Wallis test for each of the proteins and record the

p-values in a vector.

1. In the R code

6. How many adjusted p-values are significant below the level of 0.01?

Remove the insignificant proteins from the data using the subset()

function and keep only the proteins which are significant in at least one

dose group.

6.

Non-Significant 682

Significant 177

7. How many proteins are significantly deregulated with regard to the

control samples below the level of 0.01 in at least one of the dose groups?

Remove the insignificant proteins from the data using the subset()

function.

8. How do the samples in the dose groups cluster? Comment on the

similarities and differences between samples in relation to the exact dose

in mGy.

8. The high does groups are clustered together and same can be said about medium and low, also the dose mGY suggests a similar story

9. The three most significant pathways all related to one biological process.

What process is it? Click on one of the pathways in the table and you will

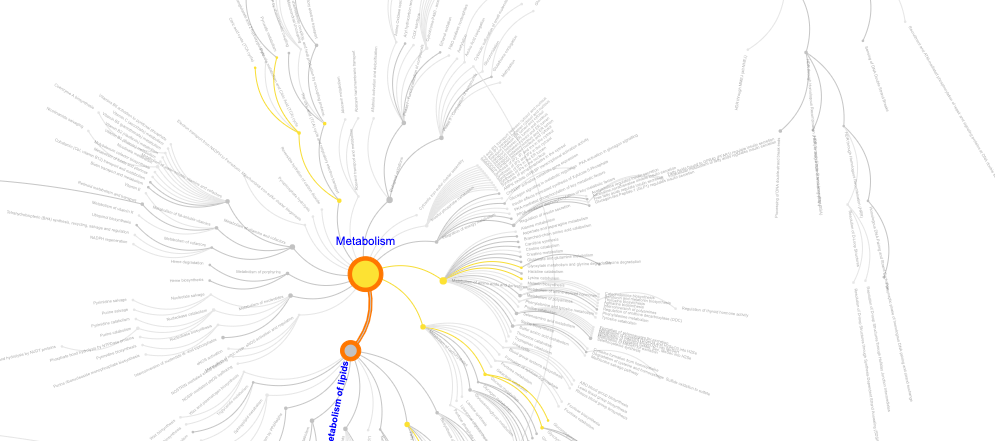
zoom in on it in the graph. To which major pathway group does this

pathway belong to?

9.Metabolism

10. Click on the ‘Downloads tab’ at the top of the table. In the ‘Pathway

Diagram section’ you can download a diagram in graphics formats of the

pathway you marked. Include this diagram in the report.

11. What kind of information can you find in the report?

11. Pathways details , pathway names and other details

12. Look at the Reactome Pathways table within the Analysis tab. Is the

pathway from question 9. there? Does the KEGG Pathways database

confirm overrepresentation of this pathway?

12.Drug Metabolism is the top path

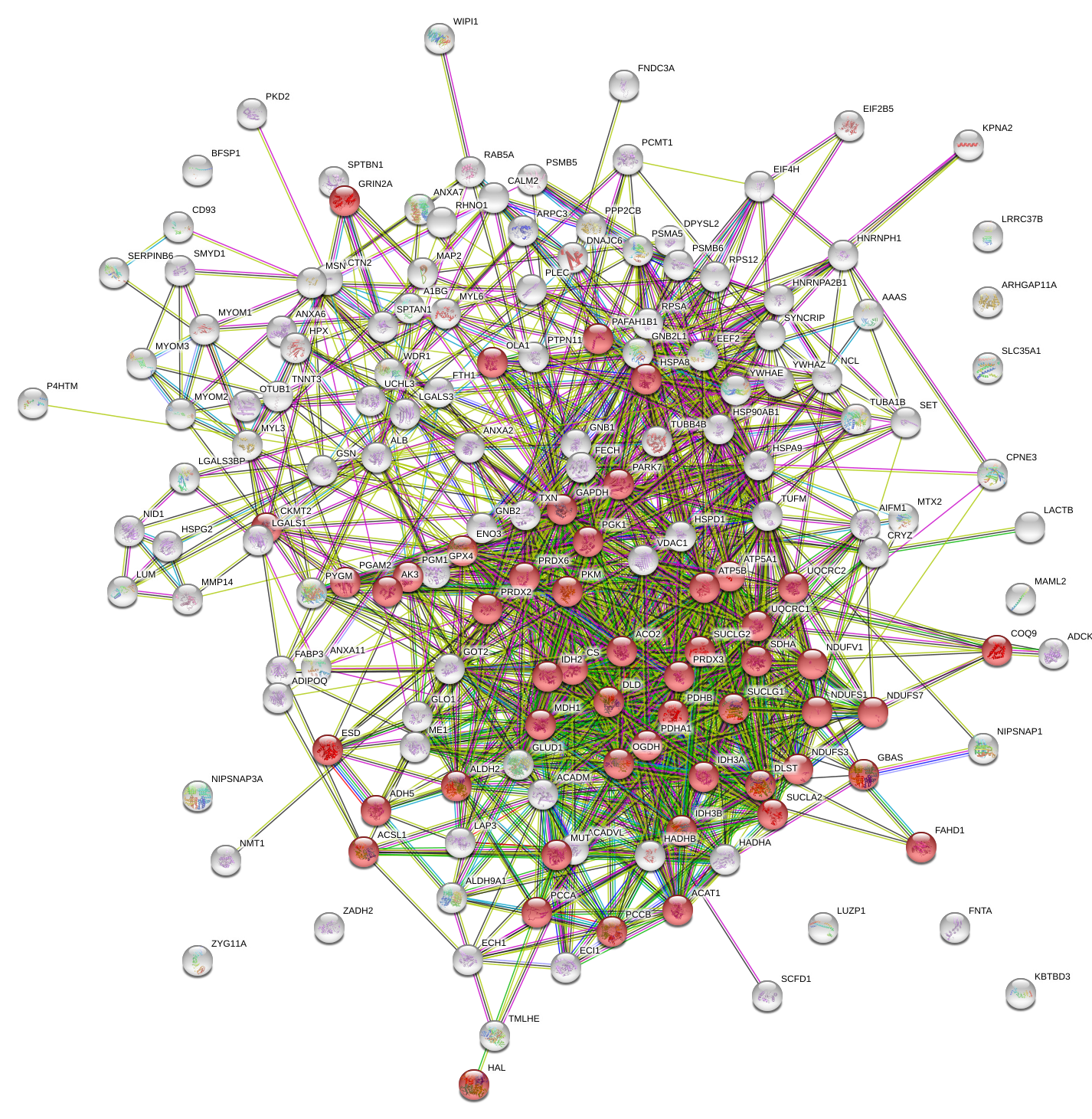
13. When you click on a pathway in the analysis tables it receives a color

assignment which allows you to follow the protein involved on the graph.

Mark the Reactome and KEGG pathway from question 12. and insert a

picture of the highlighted part of the graph into your report.

13.



14. At the bottom of the page, download the Reference publications table file.

Are there any publications related to ischemia/ischemic heart disease? If

you can find some, insert their citations into your report.

14.