

W9 Assignment: Gene Set Analysis

We take the file with differentially expressed genes with FDR cutoff of 0.05 ordered from smallest p-value to highest p-value and select the row names(gene names) and save to a file.

Getting start and end coordinates for transcripts of genes:

We use the UCSC table browser to upload the list of differentially expressed genes and get the start and end coordinates of the transcripts.

We convert the above to a data frame and then to a GRanges object.

Promoters for the differentially expressed genes:

We use AnnotationHub promoters() function to get the promoters for the differentially expressed genes. We check to see the width of the promoters and we see all 2933 have a width of 2200.

Histone Modification Peak Data:

We use AnnotationHub to get H3K4me3 narrow peak data from the Epigenome Roadmap project

Fetal brain tissue H3k4me3 peak data

AH43854 | BI.Fetal_Brain.H3K4me3.UW_H-22510.narrowPeak.gz.

We run a quick summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
174	271	646	1162	1743	41740

Adult brain tissue H3k4me3 peak data

AH43565 | BI.Brain_Mid_Frontal_Lobe.H3K4me3.112.narrowPeak.gz

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
160.0	204.0	338.0	658.3	823.0	8001.0

Adult Liver tissue H3k4me3 peak data

AH43450 | BI.Adult_Liver.H3K4me3.3.narrowPeak.gz

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
144.0	151.0	206.0	366.6	355.0	9466.0

Odds Ratio: The odds ratio tells us how much more enriched the overlap between peaks and promoters is that many times more rich than we would expect

Fetal brain tissue- 46.24177

Adult brain tissue- 39.51774

Liver tissue - 21.50464

Yes there is a change in differential gene expression from fetal to adult brain.

We see that the differentially expressed genes are (46%) more enriched than we would expect in fetal brain tissue and slightly less, (40%) enriched in the adult brain tissues.

This would be consistent with what we were expecting to find as the fetal brain tissue has a higher number of differentially expressed genes and this is expected to reduce with age.

The promoters of genes that are differentially expressed between fetal and adult brain marked by H3K4me3 in the liver but to a much lesser extent in the liver when compared to the fetal and adult brain.

When looking at the liver tissue data we see that there is about 20% enrichment of the genes with H3K4me3 marks. This may be due to the fact that H3K4me3 modifications mark not only active promoters but also enhancers and other regulatory elements.