A graph of blue dots

Description automatically generated

Figure – MA plot visualising the differences in gene expression between conditions A and B. The X axis is representative of the log transformed mean expression, while the x axis shows the log 2 fold changes between conditions. Each point on the graph represents a discrete gene.

A graph with blue dots

Description automatically generated

Figure - Figure 3 – MA plot visualising the differences in gene expression between conditions A and C. The X axis is representative of the log transformed mean expression, while the x axis shows the log 2 fold changes between conditions. Each point on the graph represents a discrete gene.

Figures one and two give insights into the expression levels of genes across the two conditions, A vs B and A vs C. The log transformed mean expression shows the average expression of a given gene across two conditions. Both conditions displayed a large range of genes with log mean expression value of greater than one, with the A vs B condition having a slightly denser cluster. The log2 fold change, present on the Y axis, was able to quantify the up or downregulation of the genes across the two samples. Positive values are indicative of genes which have been upregulated, while negative values suggest downregulation. The A vs B condition showed to have slightly more upregulated and downregulated genes than the A vs C condition, with the highest and lowest log2 fold changes being 20 and -20 respectively. Significantly upregulated and downregulated genes present within both samples are strong candidates for further analysis.

A graph of a volcano plot

Description automatically generated

Figure – Volcano plot showing the relationship between expression changes and their respective statistical significance across conditions A vs B. The X axis shows the log2foldl change, which represents the magnitude and direction of any changes in gene expression. The Y axis shows the -log p value, which is indicative of the statistical significance of any observed changes in expression.

A graph with red dots

Description automatically generated

Figure - Volcano plot showing the relationship between expression changes and their respective statistical significance across conditions A vs C. The X axis shows the log2foldl change, which represents the magnitude and direction of any changes in gene expression. The Y axis shows the -log p value, which is indicative of the statistical significance of any observed changes in expression.

The volcano plots shown in figures 3 and 4 provides insights into genes within the two conditions A vs B and A vs C which have been upregulated and downregulated, as well as having statistical significance. P values were transformed using a negative log function to allow for visualisation of p values which were several orders of magnitude in difference. T he condition A vs B showed to have a greater number of significantly downregulated and upregulated genes than A vs C. This suggests that the A vs B conditions holds greater relevance for further analysis, as significantly upregulated or downregulated genes are of greater biological significance.

A screenshot of a graph

Description automatically generated

Figure – Showing a heatmap for the top log2fold changes in both conditions A vs B and A vs C. The X axis represents the log 2 fold changes for both the conditions, while the Y axis represents the top 6 gene names. The scale to the right of the graph shows which colour relates to what value.

A heatmap was constructed to identify the significantly upregulated genes for both conditions A vs B and A vs C. Both conditions showed that the gene YNCJ0019C had the greatest log2fold change and thus the greatest upregulation. Since this gene is heavily upregulated in both conditions it represents a string candidate for further analysis. Gene YNCD0031C also showed to be upregulated strongly in the A vs B condition, however this was not consistent with the A vs C condition which had a value 6 orders of magnitude lower. While this is a weaker correlation, further analysis would still be adequate due to YNCD0031C strong upregulation in one condition. The cluster of genes YL172C, YBR169C and YBR177C al showed to have similar log2fold changes across different conditions, which may represent a shared biological function. Overall the heatmap highlights the name of strongly upregulated genes which represent candidates for further analysis.

Interrogation of the YNCJ0019C gene revealed its function as a tRNA gene from the saccharomyces yest genome 1. Its molecular function is enable triplet codon- amino acid adaptor activity and involved in translational elongation 1. The up regulated YNCD0031C gene also showed to be a tRNA gene involved in translational elongation 1.

A graph of values for a person

Description automatically generated

Figure - Histogram of p values throughout the data for the condition A vs C. The X axis is representative of the respective p value of the data. The Y axis represents the frequency of genes which fall between a given p value range

A graph of a graph

Description automatically generated

Figure - Figure 7 - Histogram of p values throughout the data for the condition A vs C. The X axis is representative of the respective p value of the data. The Y axis represents the frequency of genes which fall between a given p value range.

Histograms were plotted to illustrate the frequency of different p values obtained throughout the gene expression data between conditions A vs B and A vs C. The p values of the gene expression data showed to be poison distributed, with the majority of genes a significant p value of less than 0.05. This suggests that a significant number of genes from both conditions are differentially expressed. Overall this suggests that the biological response from both condition is statistically significant.

**Bibliography**

(1)

*YNCJ0019C | SGD*. Yeastgenome.org. https://www.yeastgenome.org/locus/S000006773 (accessed 2025-01-28).