## FIGURES and TABLES



Figure 1. Experimental setup. The illustration shows the structure of the experimental setup for RNA-Seq & Protein data after cleaning out inconsistent data points in a general perspective. There are seven different sets of experiments; glucose time course experiment, glycerol time course experiment, gluconate experiment, lactate experiment, high sodium experiment, low magnesium experiment and high magnesium experiment. We have RNA-Seq data for all seven experiments and protein data for four of them. The four different carbon sources used for the experiments and varying MG and Na levels are also shown in figure. Number of test tubes represents the number of different batches for each experimental setup so basically there are 3 replicas for each experiment and there are 21 different batches in total. The colors of test tubes refer to the carbon sources used in experiments, glucose, glycerol, gluconate and lactate are represented with light purple to dark purple. The green tones in the background represent the phase of the experiment. Light green is exponential phase, darker green is stationary phase and dark green is late-stationary phase. The x-axis shows time and for glucose, glycerol, gluconate and lactate experiments we show the exact hours that the samples are collected. On the other hand for Na and Mg experiments the x-axis is divided into pieces representing different concentrations of Na and Mg. The asterisk “[\*]” near the concentrations of Mg and Na levels represent the base value. The circular points with an x on them, above the axis represents the time or phase that the RNA sample is collected and diamond shaped points, below the axis represents the time or phase that the protein sample is collected. Colors of points represent how many batches we have at that point. Red represents a single batch, yellow represent two batches and green represent three batches.

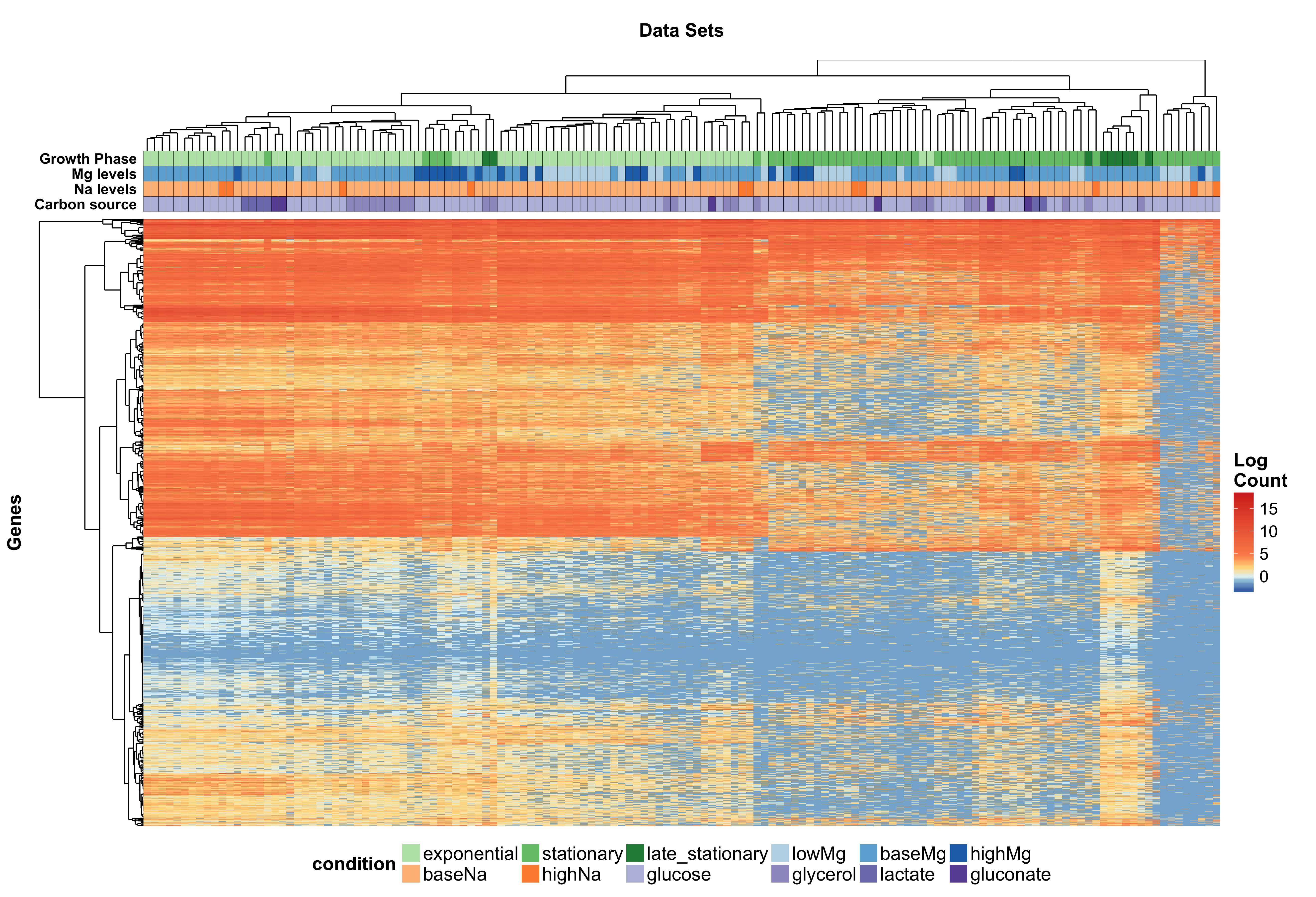


Figure 2: Heatmap of all RNA data. The figure represents the heatmap of all 143 samples with 3698 different normalized RNA levels. The x-axis shows different samples and the dendogram of x-axis calculates the clustering of the data using Euclidian distance. A similar approach is used to cluster different genes also and generated dendogram is shown on y-axis. The reddish area represents the highly abundant genes and bluish area represents less abundant genes. The clustered data sets are color coded with 12 different conditions with in 4 different categories. The four different variables are represented in a color table at top. The variables are ordered with respect to quality of clustering. The order of categories is growth time, Mg level, Na level and carbon source from best clustered to least clustered category.

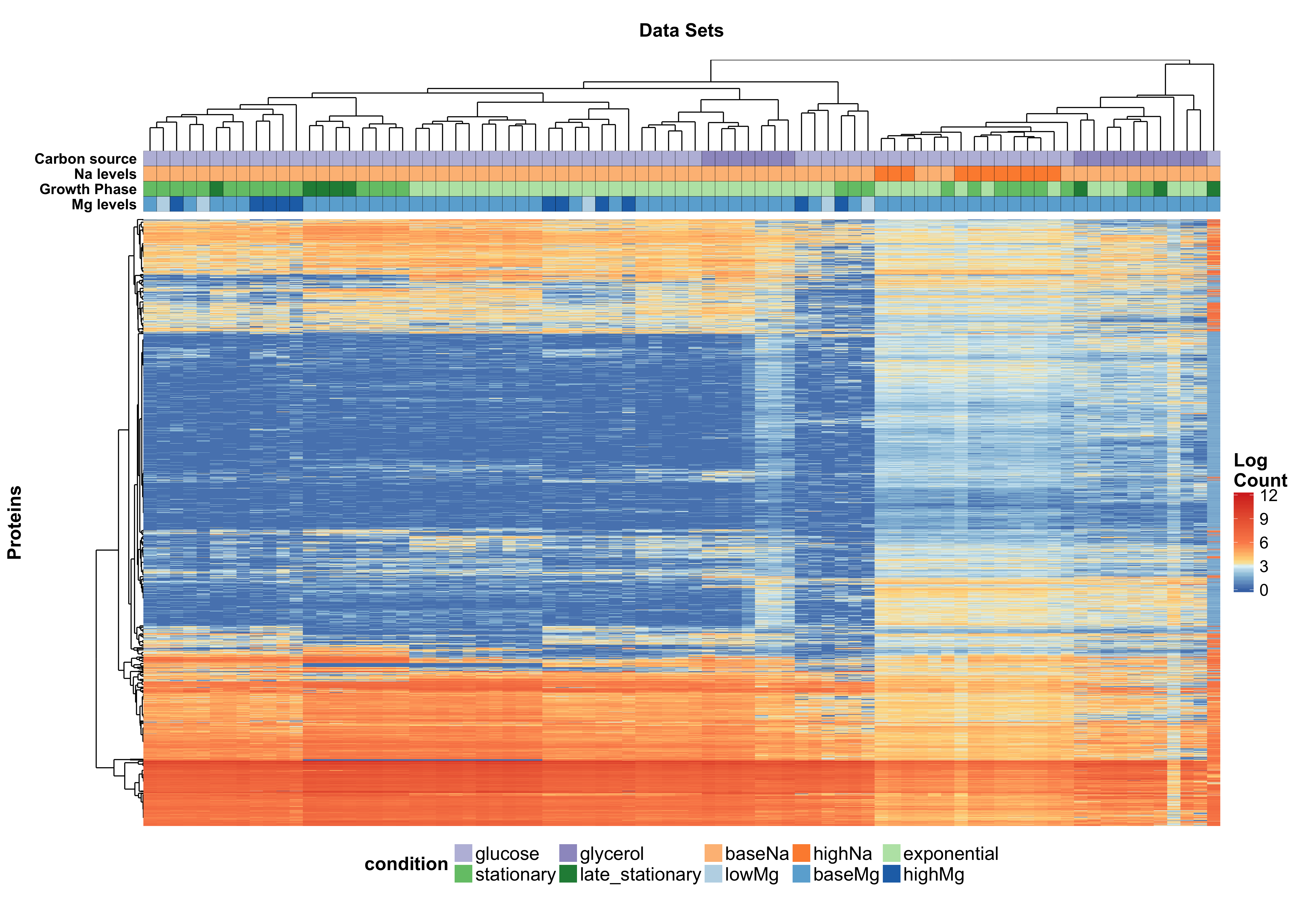


Figure 3. Heatmap of all protein data. The figure represents the heatmap of all 81 samples with 3117 different normalized protein levels. The x-axis shows different samples and the dendogram of x-axis calculates the clustering of the data using Euclidian distance. A similar approach is used to cluster different genes also and generated dendogram is shown on y-axis. The reddish area represents the highly abundant genes and bluish area represents less abundant genes. The clustered data sets are color coded with 10 different conditions with in 4 different categories. The four different variables are represented in a color table at top. The variables are ordered with respect to quality of clustering. The order of categories is carbon source, Na-level, growth phase and Mg-level from best clustered, to least clustered category.

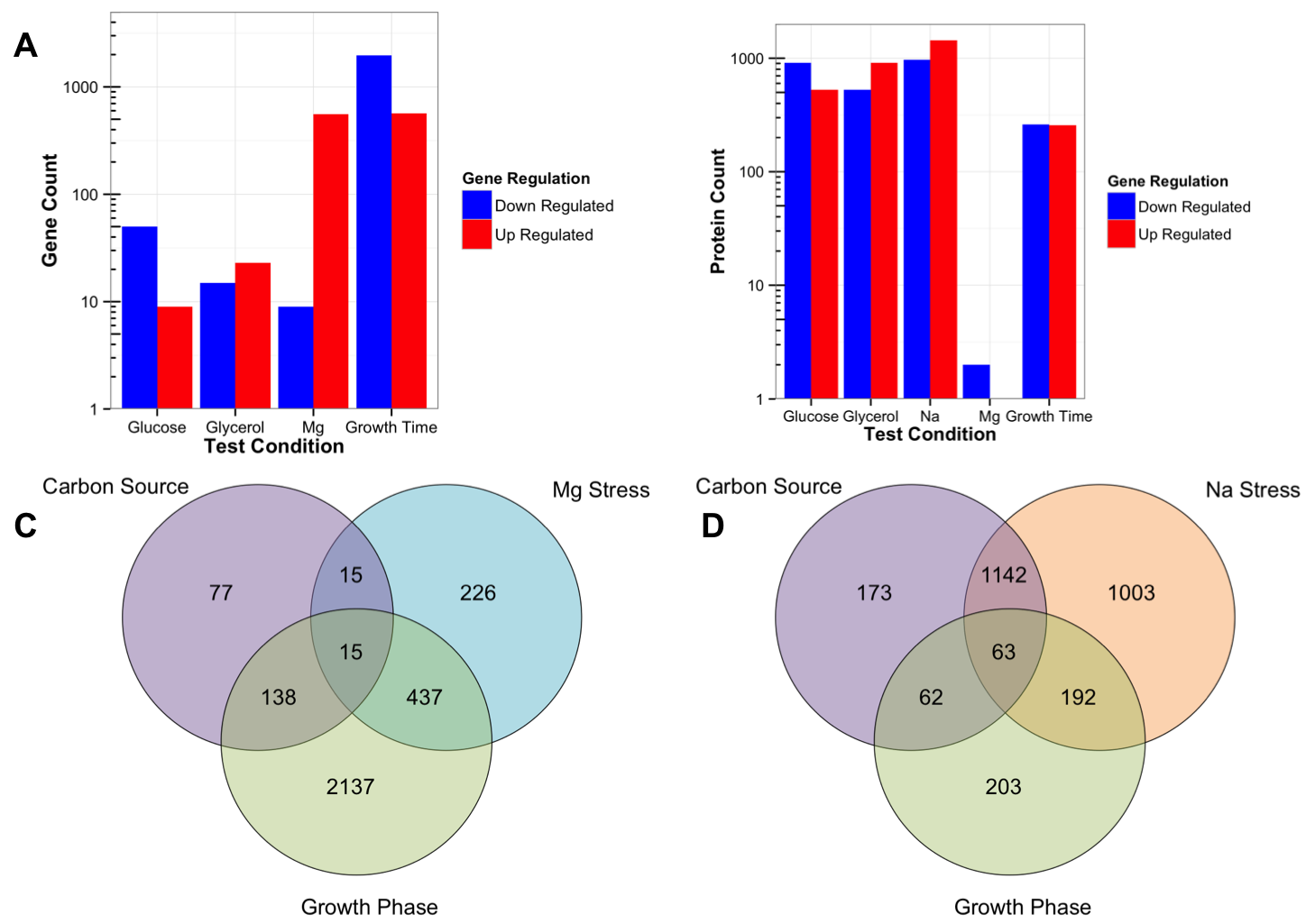


Figure 4. Cell response comparison between different categories of variables. The comparison of response to different categories of variables can be analyzed from different perspectives. For the comparisons we calculate the Spearman's rank correlation for each gene or protein with respect to each category of variables and determine the significance of response in terms of P’. We do the same analysis for both proteins and RNA's. (A) Amount of significantly responding genes. Figure represents the number of significantly up and significantly down regulated genes (P’<0.05) with respect to two different carbon sources, Mg levels and growth time. There are no significantly responding genes to switching to gluconate and lactate as carbon source and, with respect to changing Na levels. (B) Amount of significantly responding proteins. Figure represents the number of significantly up and significantly down regulated proteins (P’<0.05) with respect to two different carbon sources, Na, Mg levels and growth time. There are no significantly responding genes to switching to gluconate and lactate. It can be seen that cells do not give a significant response to varying Mg levels. (C) Co-altered genes with respect to different categories of variables. The figure shows similarities and differences in the significantly responding genes (P’<0.05) between different categories. Venn diagram shows the intersections of three different categories of variables (Carbon Source, Mg Stress, Growth Phase) and the numbers represent the number of genes that are significantly altered. The carbon source set includes responses to two different carbon sources glucose and glycerol. There is no significantly responding genes to carbon sources gluconate and lactate. (D) Co-altered proteins with respect to different categories of variables. The figure shows similarities and differences in the significantly responding proteins (P’<0.05) between different categories. Venn diagram shows the intersections of three different categories of variables (Carbon Source, Na Stress, Growth Phase) and the numbers represent the number of genes that are significantly altered. The carbon source set includes responses to two different carbon sources glucose and glycerol. The 2 genes altered because of varying Mg stress are omitted from the figure.

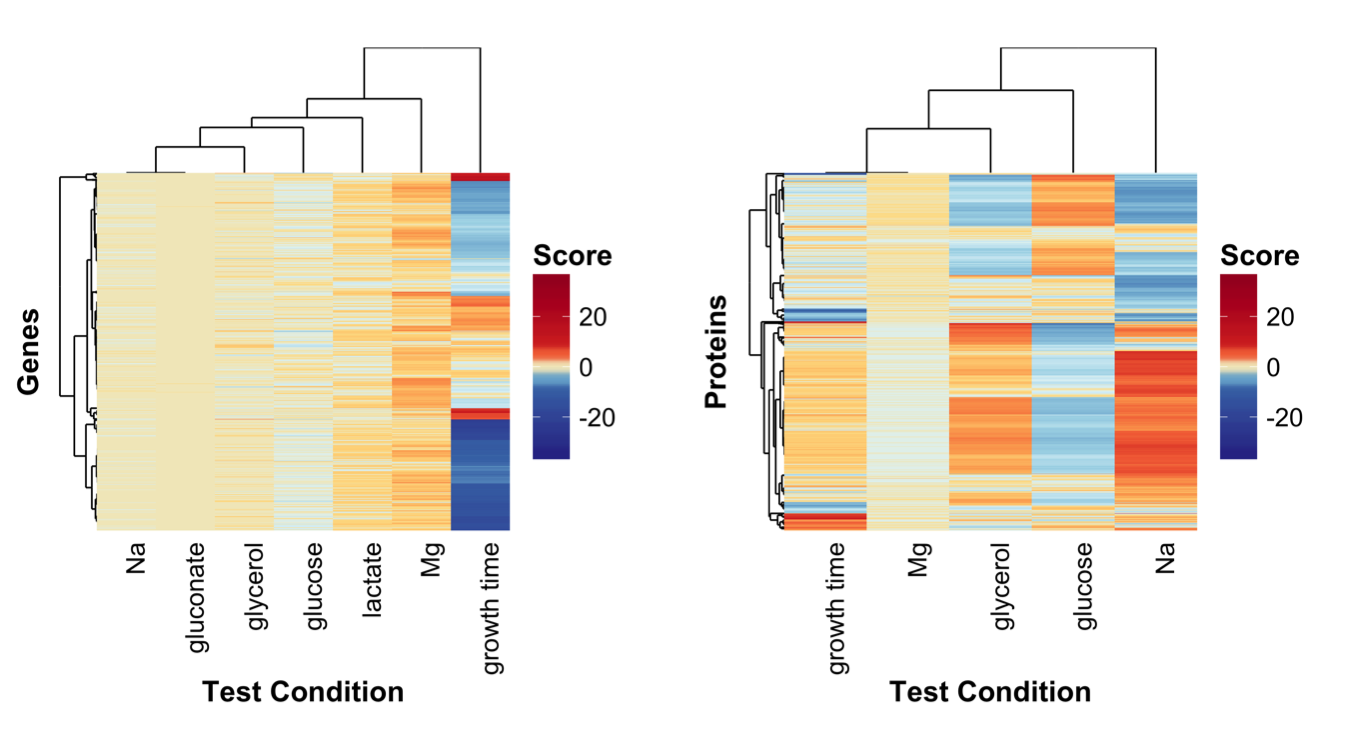


Figure 5. Histograms of cell response to external conditions. Figures show how the responses of genes with respect to different conditions clustered with dendograms based on Euclidian distance between responses. Reddish color represents up regulation and bluish color represents down regulation with respect to the related condition. Score is a variable derived from *P’* by the function *“sign(corr.coeff \* -log10(P’)*”. It is a big positive or negative number if the considered gene or protein is up regulated or down regulated significantly with respect to change in the related category. (A) Clustering between responses to conditions in terms of RNA amounts. The clustering shows carbon sources are clustered together and salt stresses are separated from each other, in addition to them growth time is the most unusual of all conditions investigated. That means the response to four different carbon sources affects similar sets of genes, but the response to two different salt stresses effects different sets of genes. Moreover the responses of genes are most significant with respect to growth phase and almost all of the genes are altered (either up or down regulated) with respect to growth time. (B) Clustering between responses to conditions in terms of protein amounts. The clustering shows two carbon sources are clustered together and salt stresses are separated from each other. That means the response to two different carbon sources affects similar sets of genes, but the response to two different salt stresses effects different sets of genes. Moreover the responses of genes are most significant with respect to Na amount and almost half of the genes are altered (either up or down regulated) with respect to Na amount. In addition to these Na and carbon sources have similar responding protein clusters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Overall z-score** | **Condition** | **Z-score** | **# elements** |
|  |  | exponential | -14.15 | 77 |
| Growth Phase | -2.82 | stationary | 3.14 | 57 |
|  |  | Late-stationary | -2.34 | 9 |
|  |  | lowMg | 1.52 | 35 |
| Mg Levels | -0.70 | baseMg | -2.33 | 85 |
|  |  | highMg | -0.80 | 23 |
| Batch number | -0.14 |  |  |  |
| Na Levels | -0.05 | baseNa | -0.76 | 132 |
|  |  | highNa | 0.83 | 11 |
|  |  | glucose | 1.68 | 107 |
| Carbon Source | 0.19 | glycerol | -1.39 | 24 |
|  |  | lactate | -2.49 | 6 |
|  |  | gluconate | -0.28 | 6 |

Table 1: Clustering quality of different categories. Table shows different categories of parameters and their z-scores in terms of clustering quality. To calculate the clustering quality with respect to a category, we calculate cophenetic distance between all pairs and calculate the mean value for all categories and conditions. We repeat the same calculation thousand times with mixed data and calculate the z-score of original mean value with respect to mean values of randomly mixed runs.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **RNA** | |  | **Proteins** | |
| **Variable** | **Z-score** |  | **Variable** | **Z-score** |
| Growth Phase | -3.34 |  | Carbon Source | -1.55 |
| Mg Levels | -0.79 |  | Na Levels | -1.22 |
| Na Levels | -0.18 |  | Growth Phase | -0.67 |
| Carbon Source | 0.13 |  | Mg Levels | -0.07 |
|  |  |  |  |  |
| Batch Number | -0.16 |  | Batch Number | -4.62 |