The *E. coli* molecular phenotype under different growth conditions

Supplementary materials

Mehmet U. Caglar*, John R. Houser, Craig S. Barnhart, Daniel R. Boutz, Sean M. Carroll, Aurko Dasgupta, Walter F. Lenoir, Bartram L. Smith, Viswanadham Sridhara, Dariya K. Sydykova, Drew Vander Wood, Christopher J. Marx, Edward M. Marcotte*, Jeffrey E. Barrick*, Claus O. Wilke*

January 30, 2017

List of Figures

Supplementary Figures

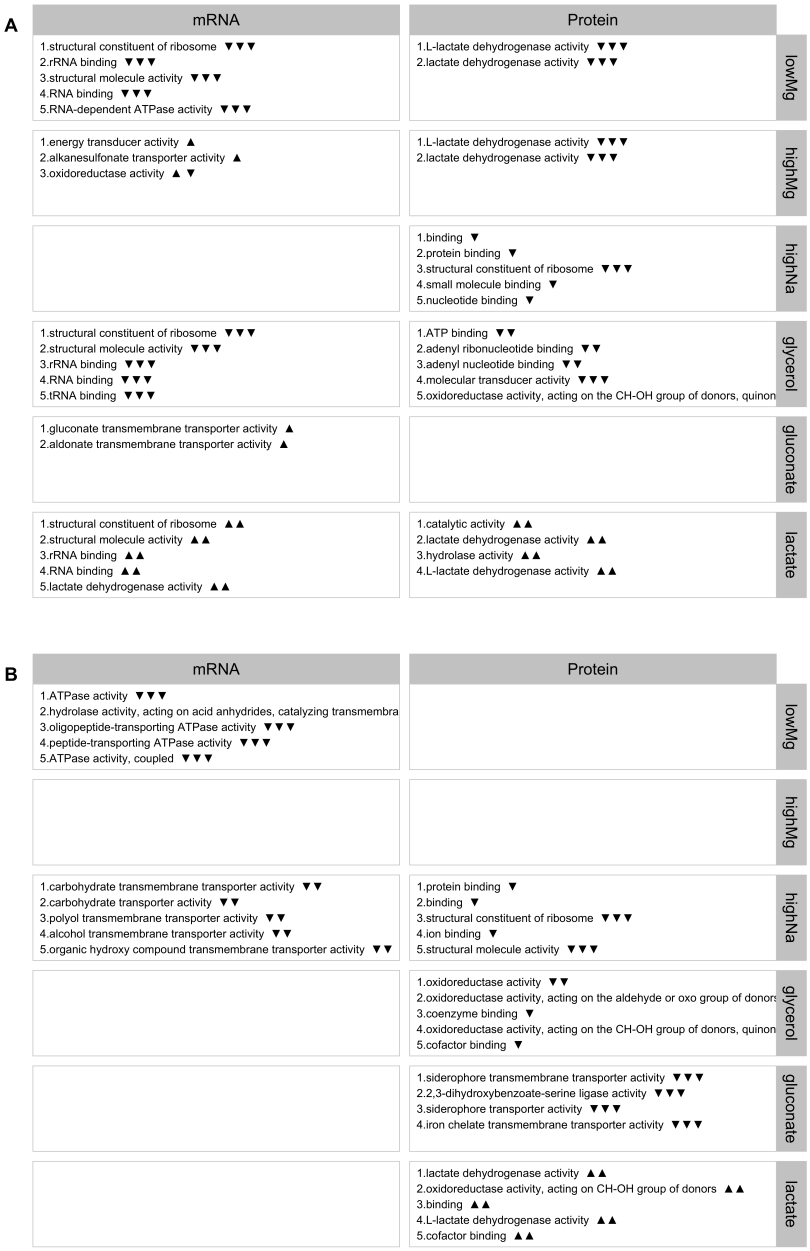

Figure S1: **Significantly differentially expressed molecular functions, as determined by GO annotations.** For each condition, we show the top-5 differentially expressed molecular functions according to either mRNA or protein abundances. Empty boxes indicate that no differentially expressed pathways were found. The arrows next to pathway names indicate the proportion of up- and down-regulated genes among the significantly differentially expressed genes in this pathway. One up arrow indicates that 60% or more of the genes are up-regulated, two arrows correspond to 80% or more genes, and three arrows correspond to 95% or more genes being up-regulated. Similarly, down arrows indicate the proportion of down-regulated genes. (A) Exponential phase. (B) Stationary phase.

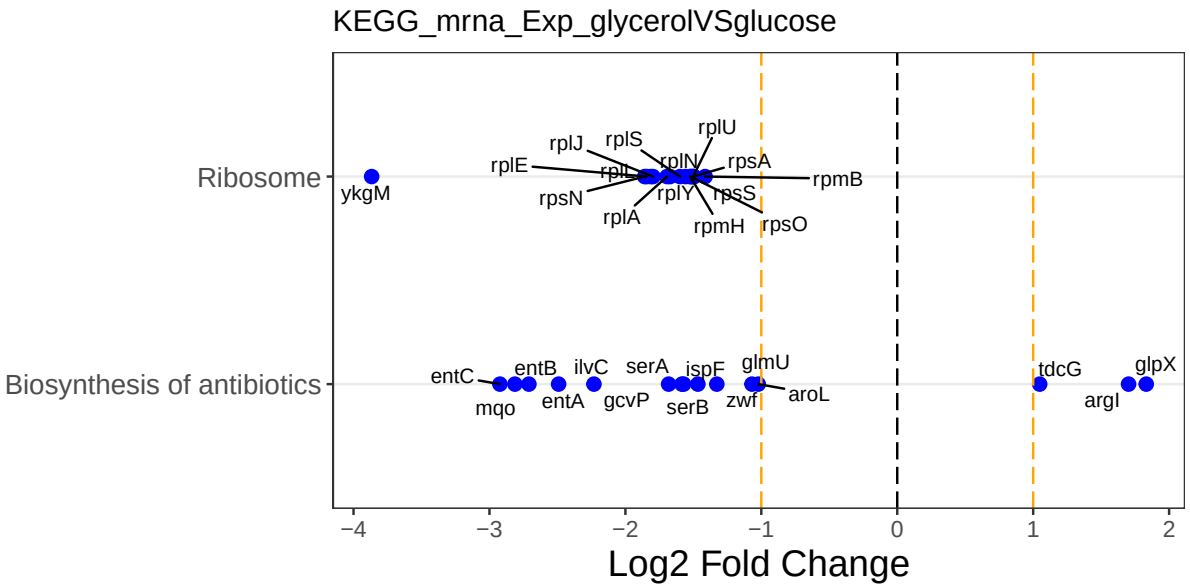

Figure S2: **Significantly differentially expressed KEGG pathways and associated genes with glycerol as carbon source, as determined by mRNA abundances in exponential phase.** The top 2 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway we show up to 15 of the most significantly changing genes.

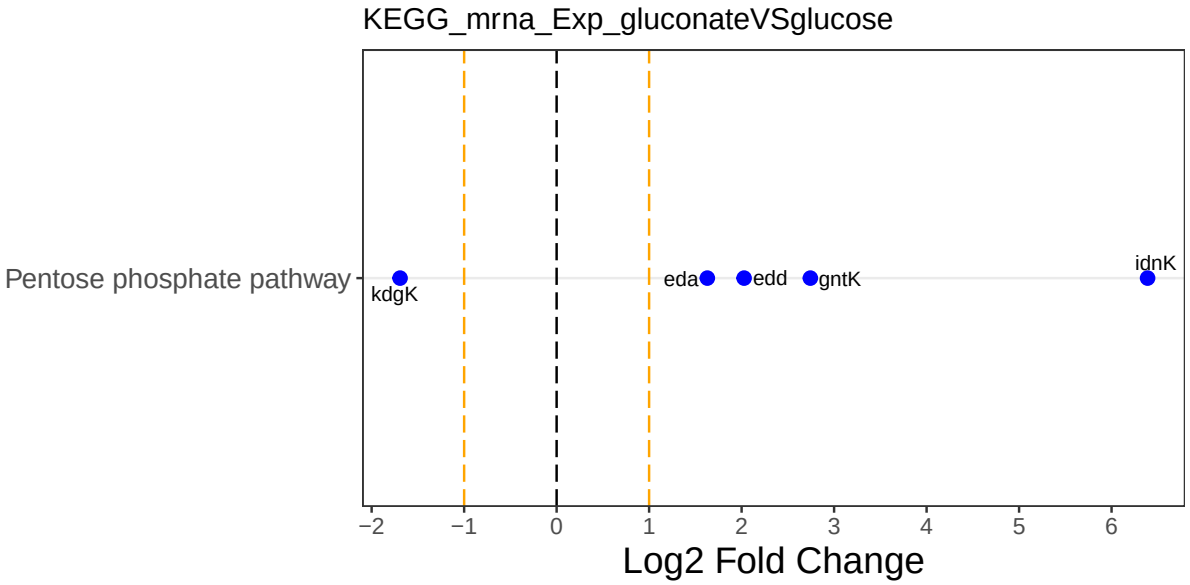

Figure S3: **Significantly differentially expressed KEGG pathway and associated genes with gluconate as carbon source, as determined by mRNA abundances in exponential phase.** The top differentially expressed KEGG pathway is shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway we show up to 15 of the most significantly changing genes.

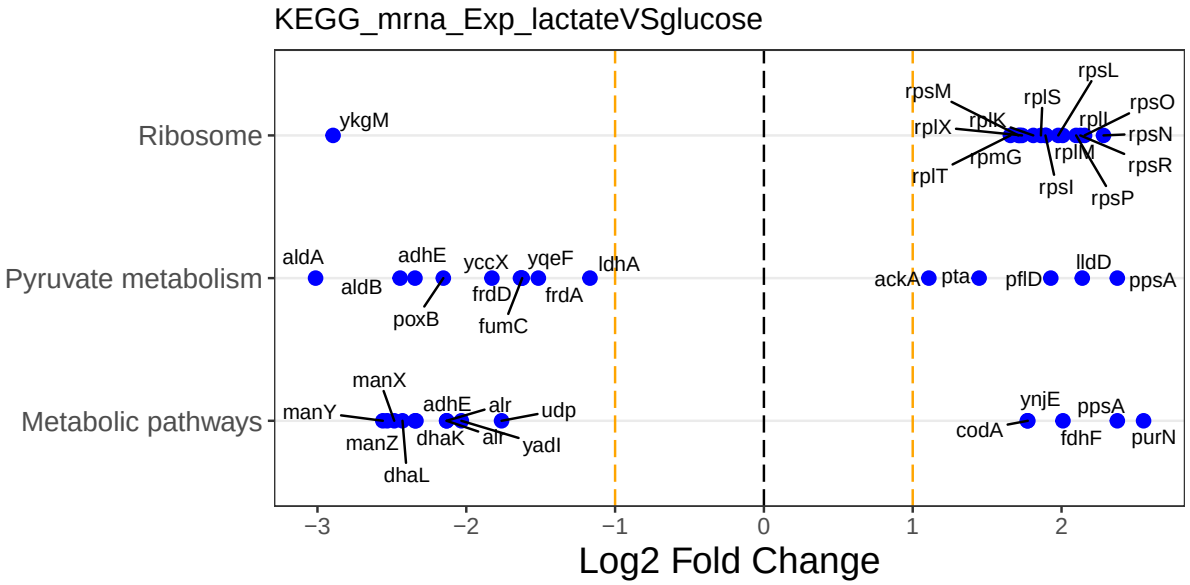

Figure S4: **Significantly differentially expressed KEGG pathways and associated genes with lactate as carbon source, as determined by mRNA abundances in exponential phase.** The top 3 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway we show up to 15 of the most significantly changing genes.

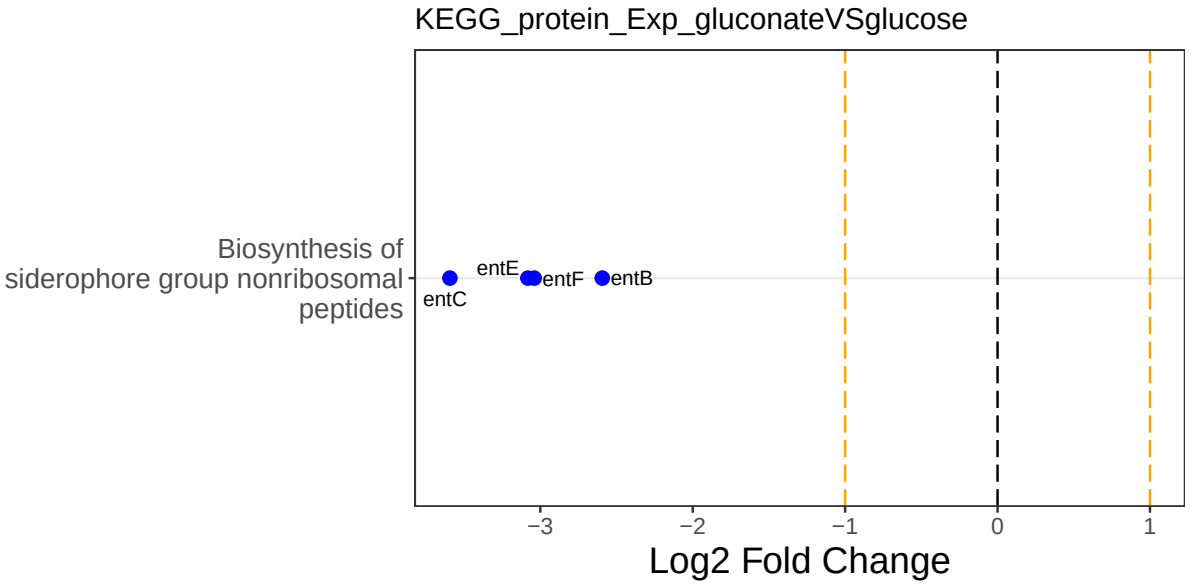

Figure S5: **Significantly differentially expressed KEGG pathway and associated genes with gluconate as carbon source, as determined by protein abundances in exponential phase.** The top differentially expressed KEGG pathway is shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway we show up to 15 of the most significantly changing genes.

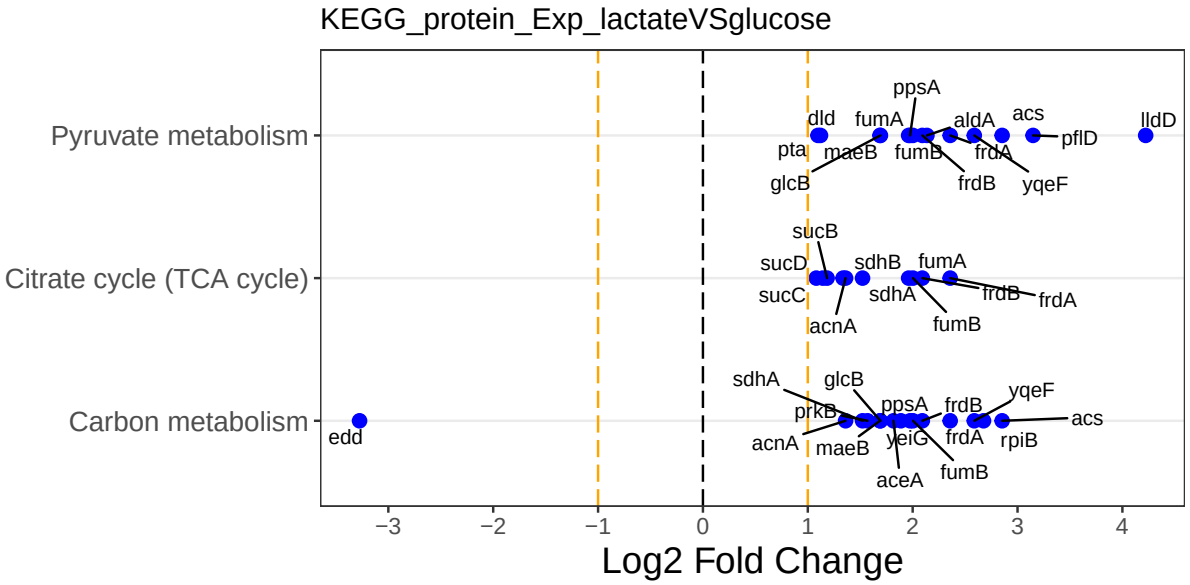

Figure S6: **Significantly differentially expressed KEGG pathways and associated genes with lactate as carbon source, as determined by protein abundances in exponential phase.** The top 3 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway we show up to 15 of the most significantly changing genes.

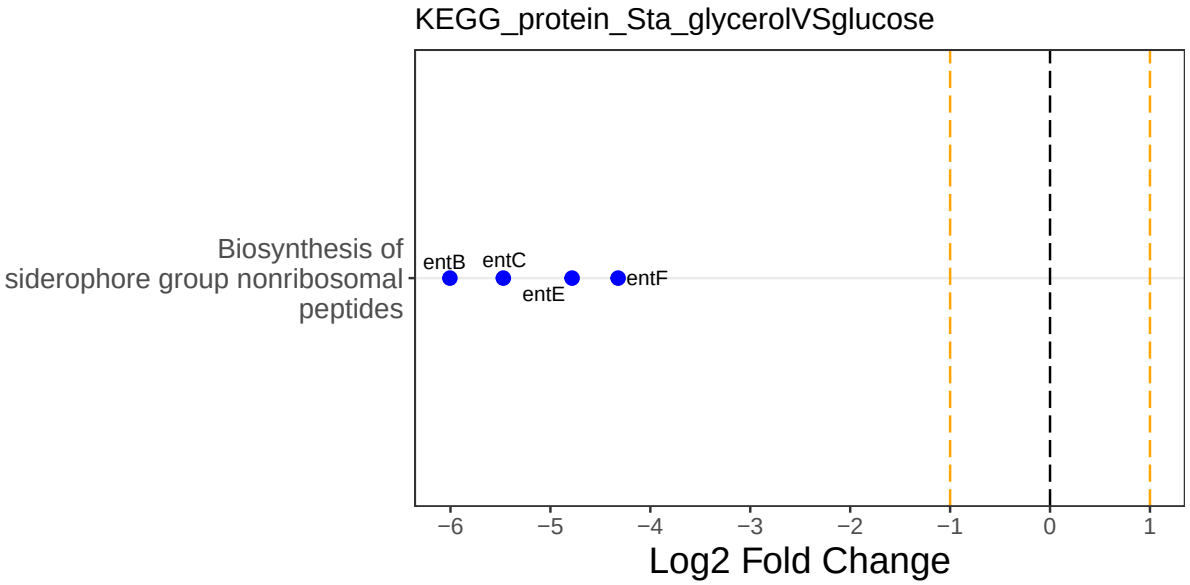

Figure S7: **Significantly differentially expressed KEGG pathway and associated genes with glycerol as carbon source, as determined by protein abundances in stationary phase.** The top differentially expressed KEGG pathway is shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

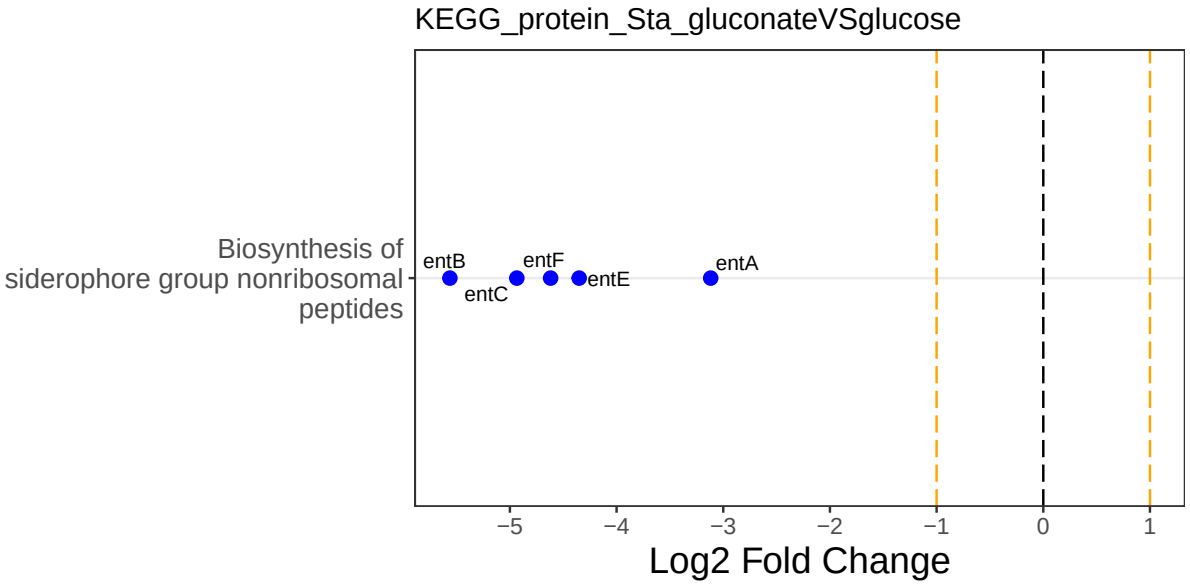

Figure S8: **Significantly differentially expressed KEGG pathway and associated genes with gluconate as carbon source, as determined by protein abundances in stationary phase.** The top differentially expressed KEGG pathway is shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

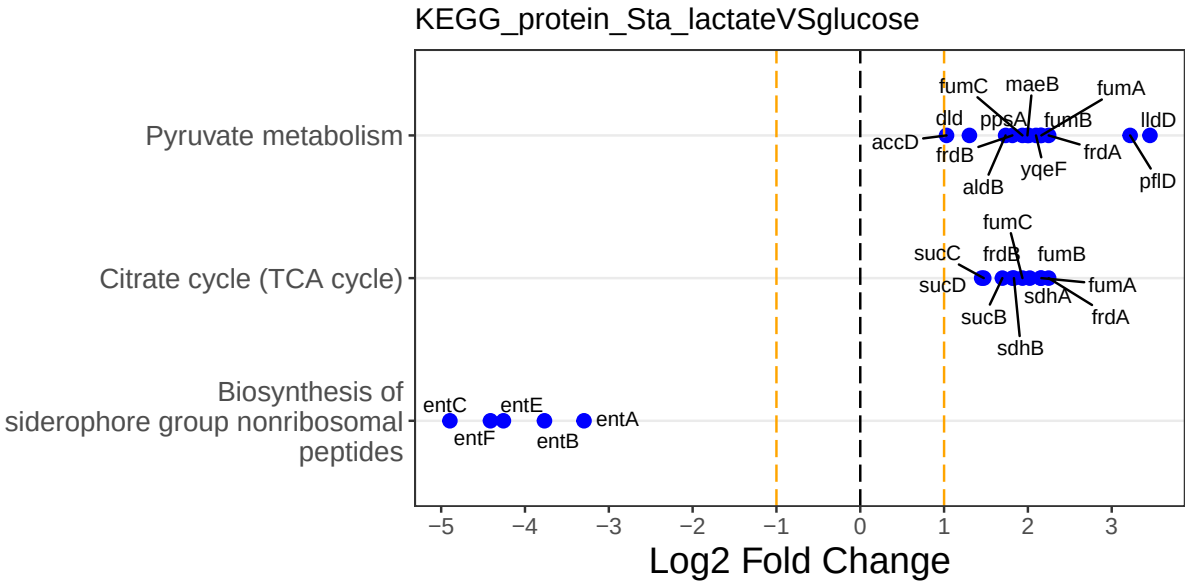

Figure S9: **Significantly differentially expressed KEGG pathways and associated genes with lactate as carbon source, as determined by protein abundances in stationary phase.** The top 3 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

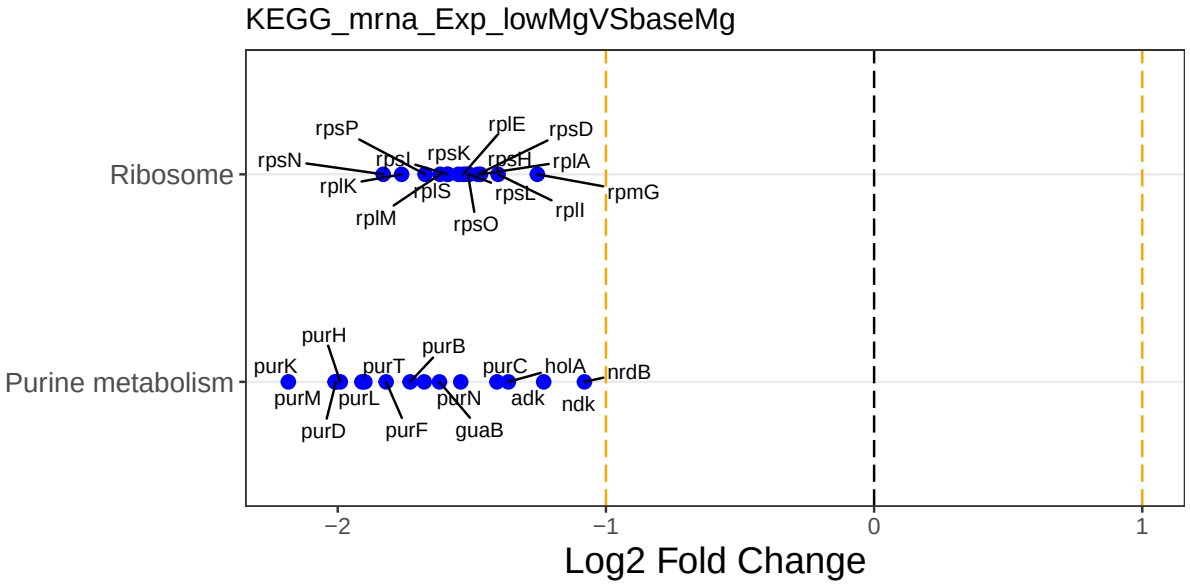

Figure S10: **Significantly differentially expressed KEGG pathways and associated genes with low Mg^{2+} levels, as determined by mRNA abundances in exponential phase.** The top 2 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

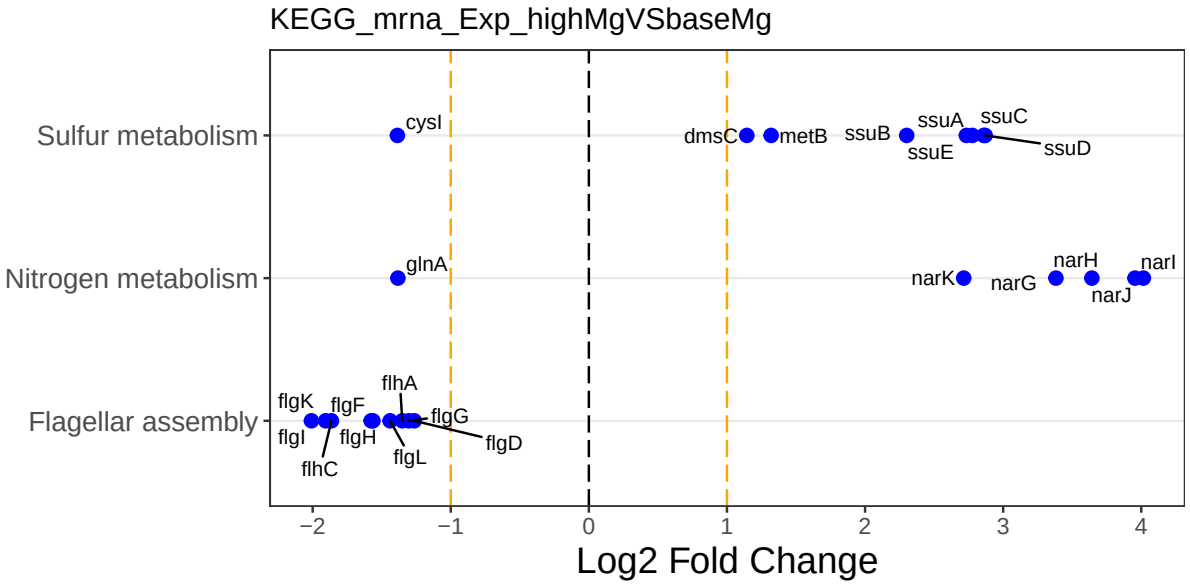

Figure S11: **Significantly differentially expressed KEGG pathways and associated genes with high Mg^{2+} levels, as determined by mRNA abundances in exponential phase.** The top 3 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

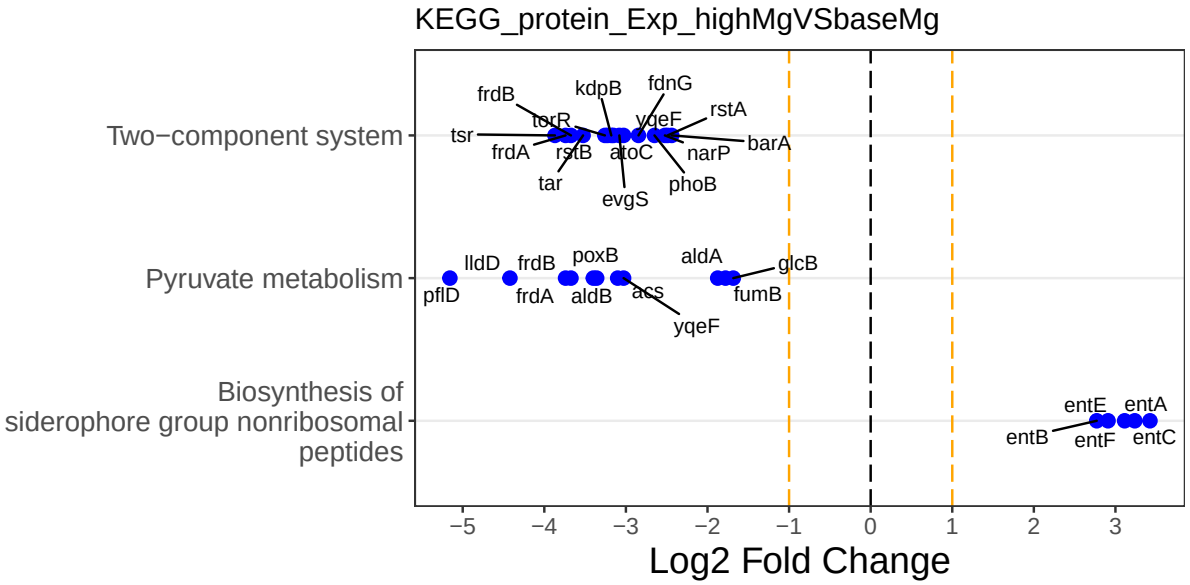

Figure S12: **Significantly differentially expressed KEGG pathways and associated genes with high Mg^{2+} levels, as determined by protein abundances in exponential phase.** The top 3 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

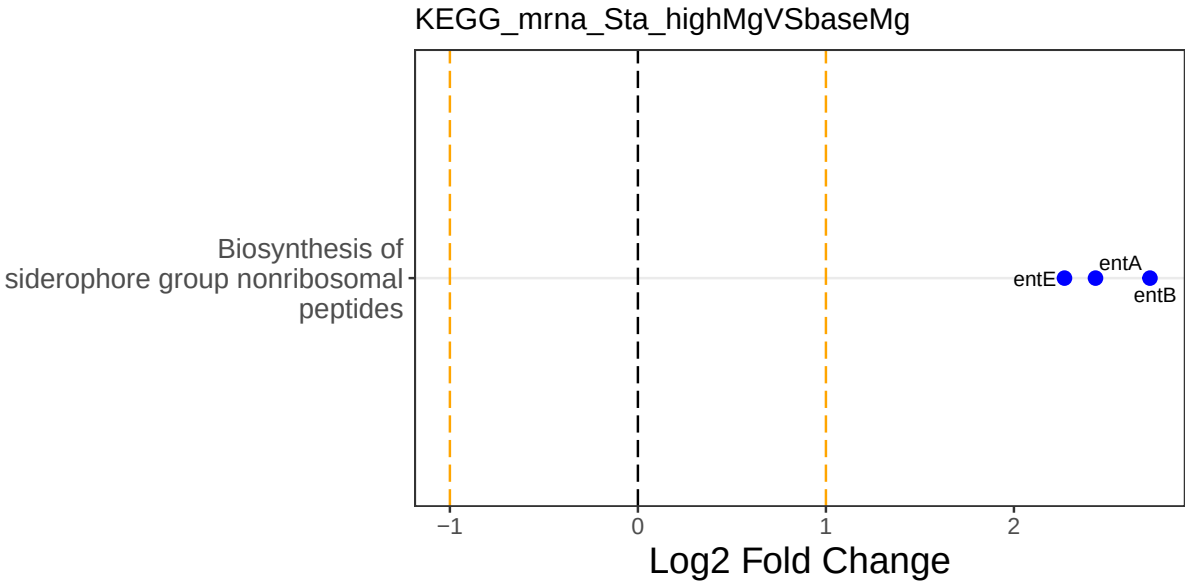

Figure S13: **Significantly differentially expressed KEGG pathway and associated genes with high Mg^{2+} levels, as determined by mRNA abundances in stationary phase.** The top differentially expressed KEGG pathway is shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

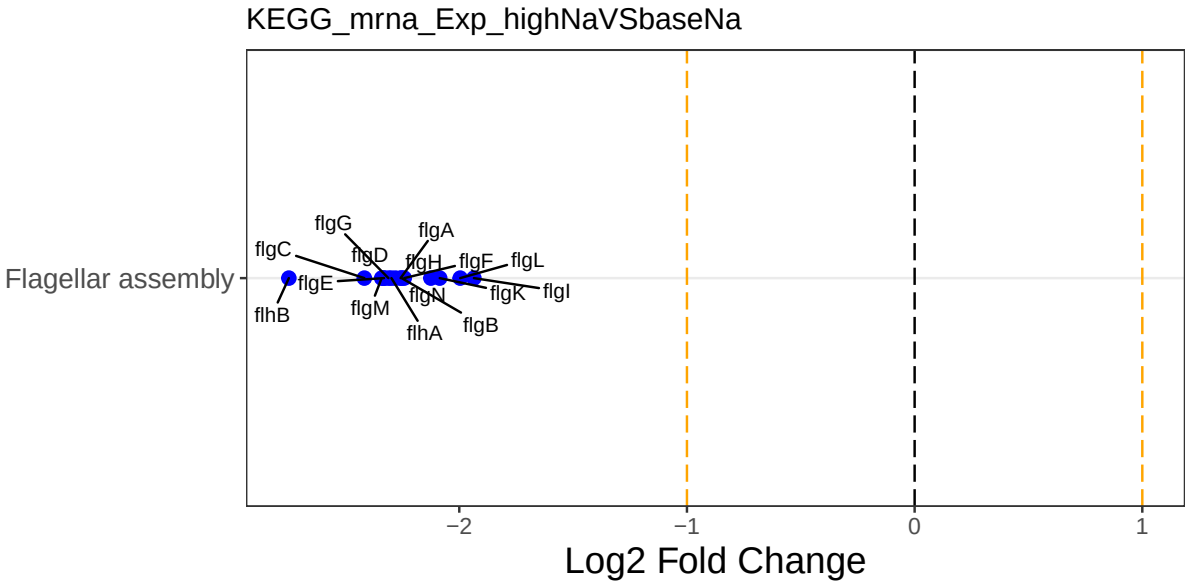

Figure S14: **Significantly differentially expressed KEGG pathway and associated genes with high Na^+ levels, as determined by mRNA abundances in exponential phase.** The top differentially expressed KEGG pathway is shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

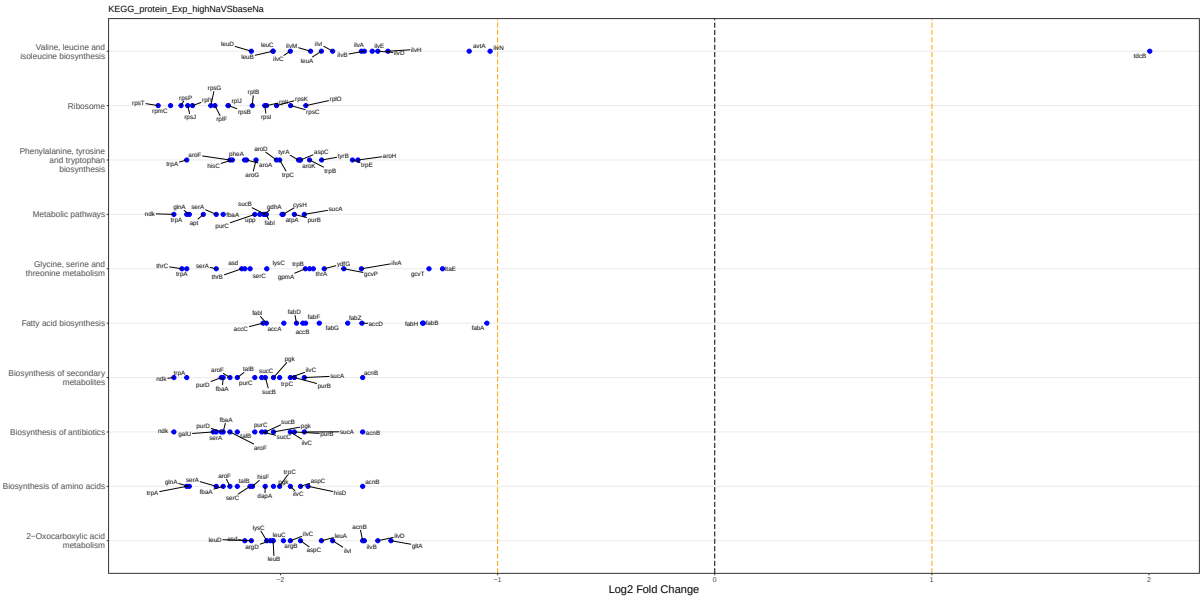

Figure S15: **Significantly differentially expressed KEGG pathways and associated genes with high Na^+ levels, as determined by protein abundances in exponential phase.** The top 10 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

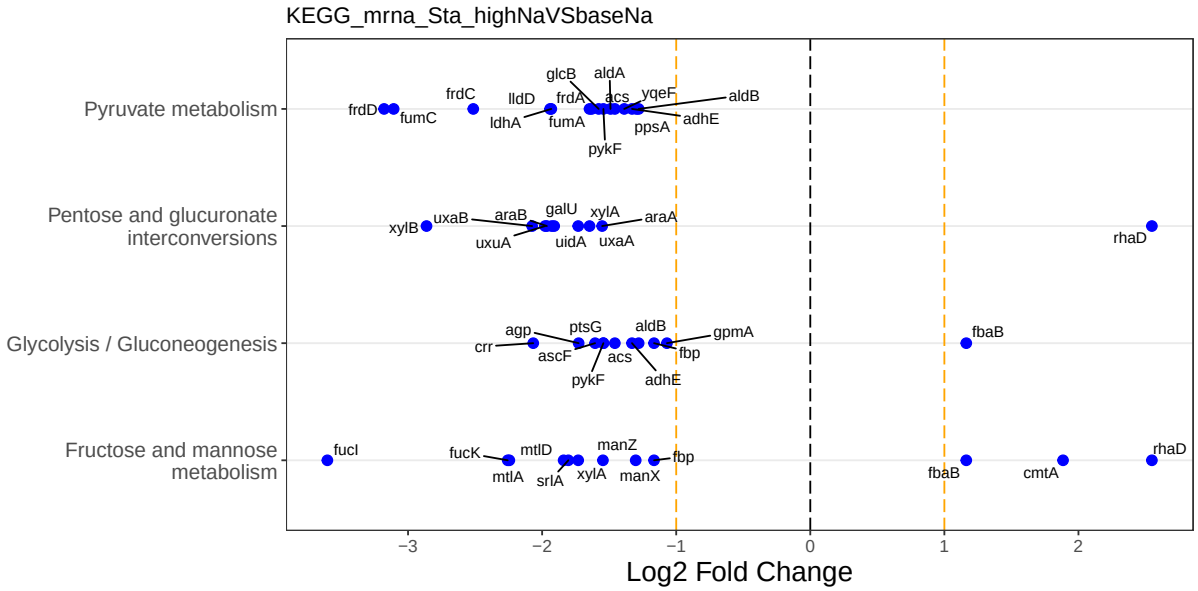

Figure S16: **Significantly differentially expressed KEGG pathways and associated genes with high Na^+ levels, as determined by mRNA abundances in stationary phase.** The top 4 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

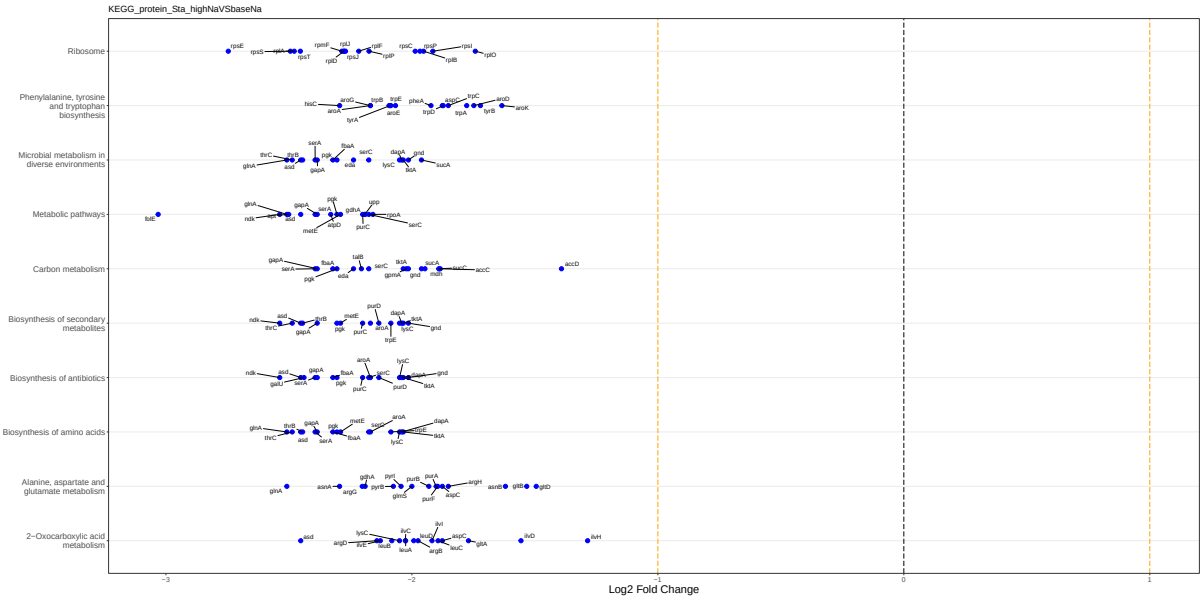

Figure S17: **Significantly differentially expressed KEGG pathways and associated genes with high Na^+ levels, as determined by protein abundances in stationary phase.** The top 10 differentially expressed KEGG pathways are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed pathways and for each pathway, we show up to 15 of the most significantly changing genes.

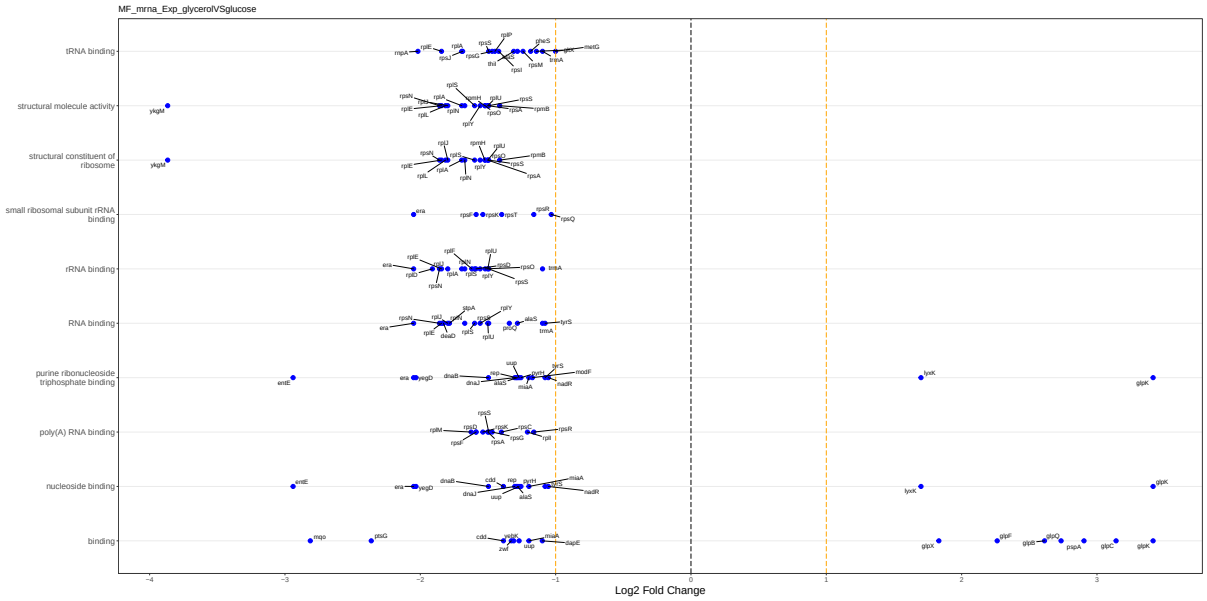

Figure S18: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with glycerol as carbon source, as determined by mRNA abundances in exponential phase.** The top 10 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

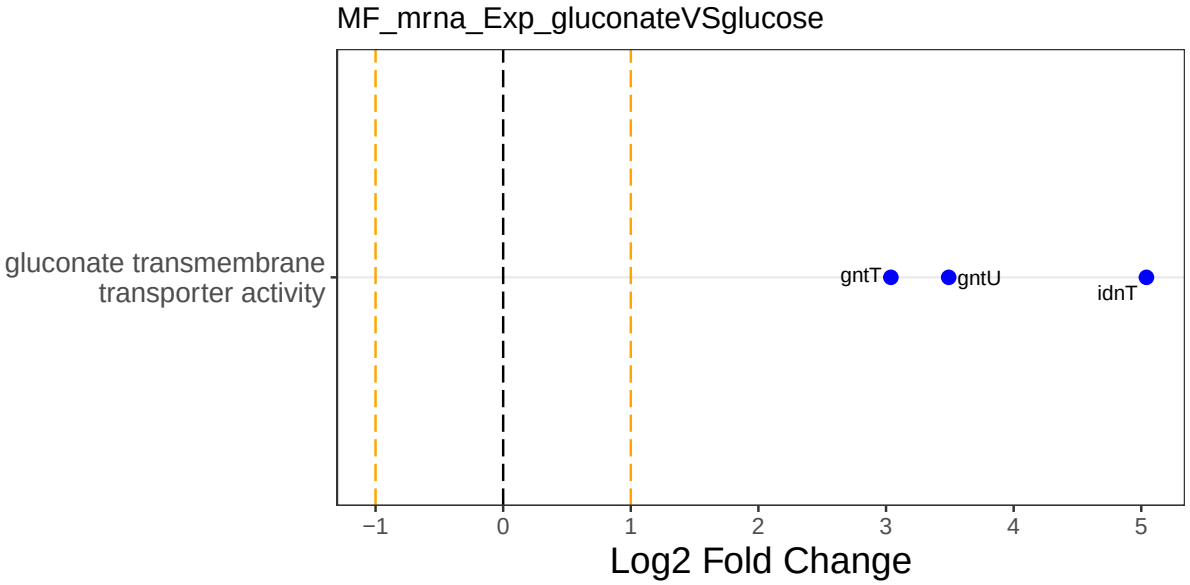

Figure S19: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with gluconate as carbon source, as determined by mRNA abundances in exponential phase.** The top 2 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

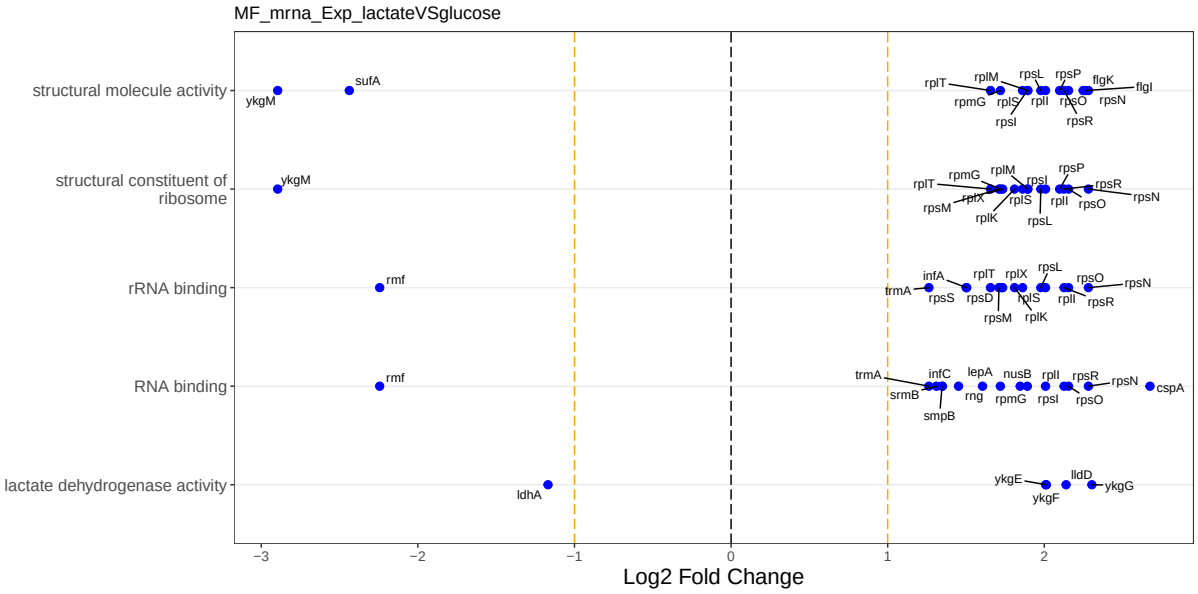

Figure S20: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with lactate as carbon source, as determined by mRNA abundances in exponential phase.** The top 5 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

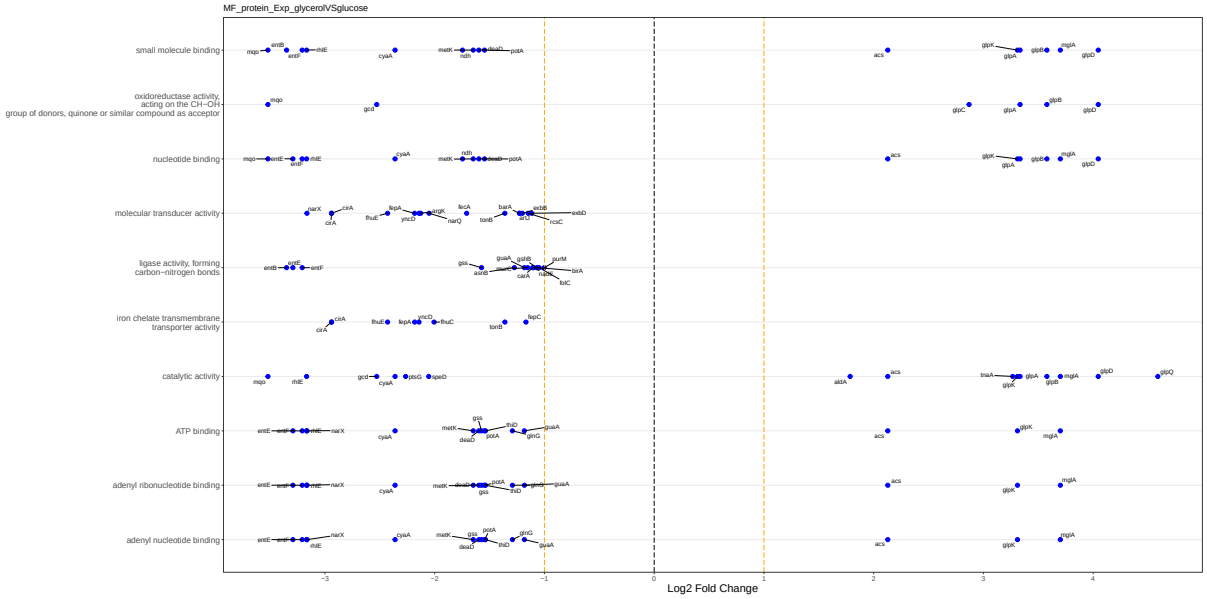

Figure S21: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with glycerol as carbon source, as determined by protein abundances in exponential phase.** The top 10 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

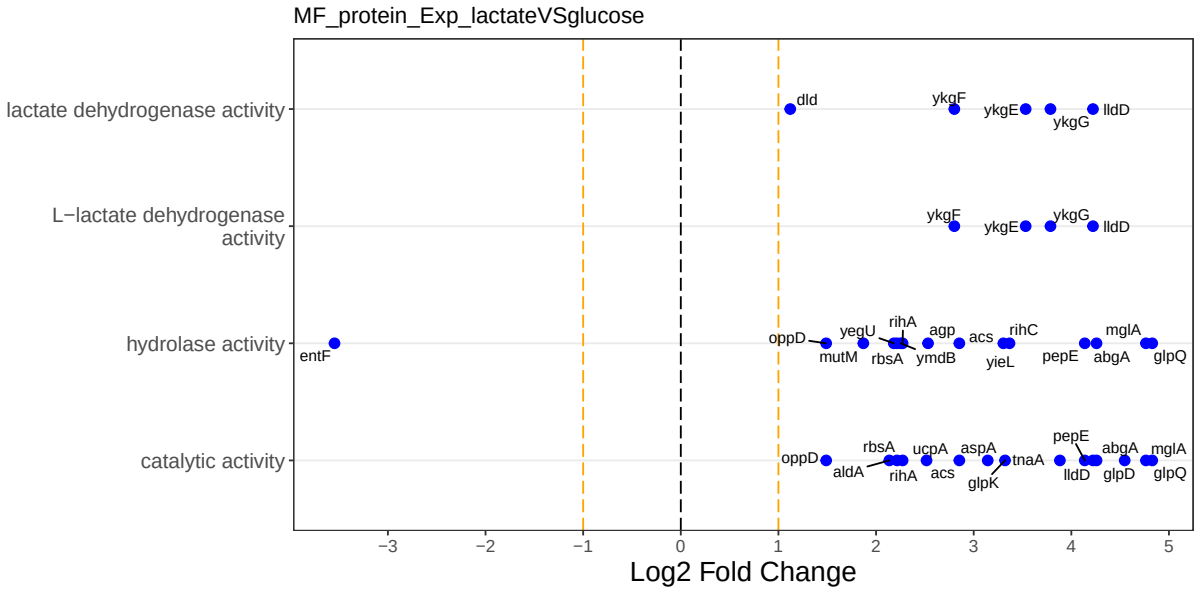

Figure S22: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with lactate as carbon source, as determined by protein abundances in exponential phase.** The top 4 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

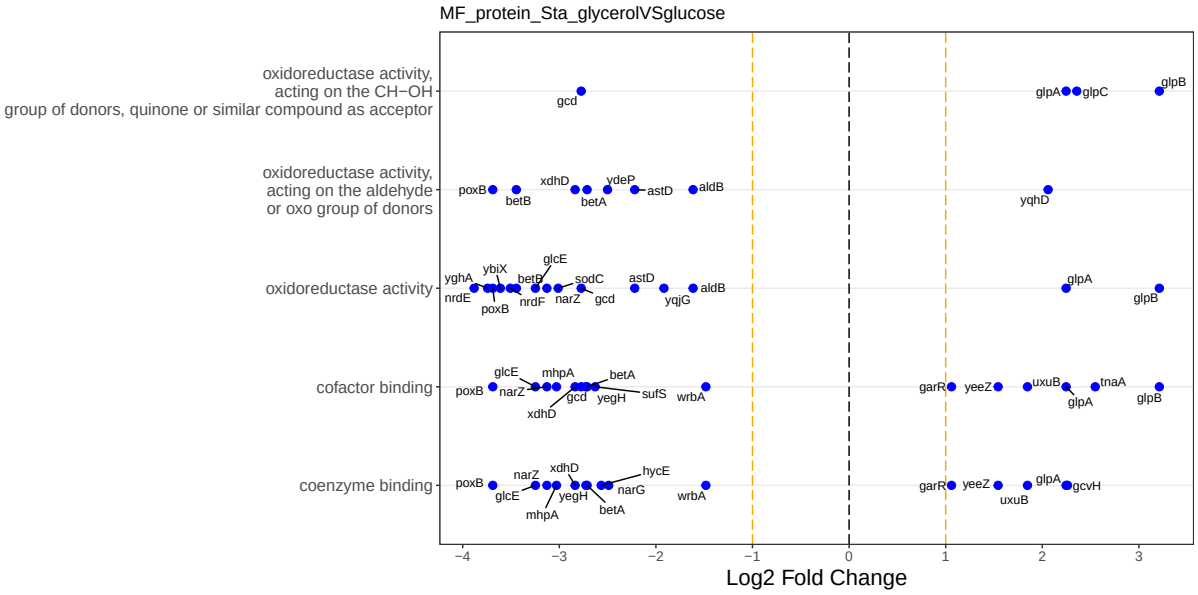

Figure S23: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with glycerol as carbon source, as determined by protein abundances in stationary phase.** The top 5 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

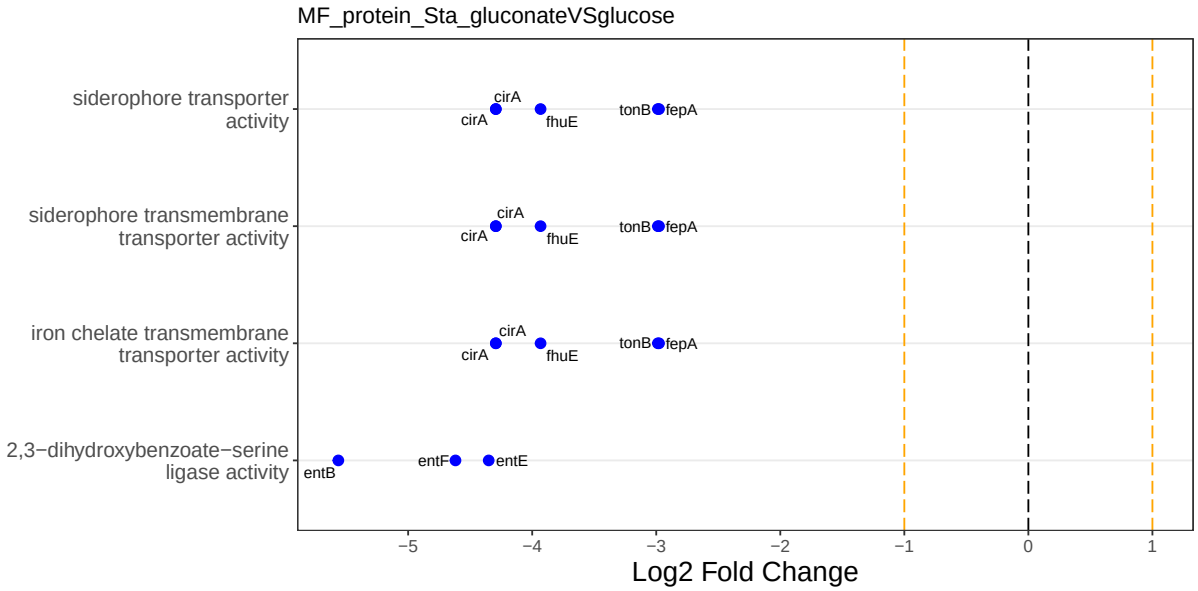

Figure S24: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with gluconate as carbon source, as determined by protein abundances in stationary phase.** The top 4 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

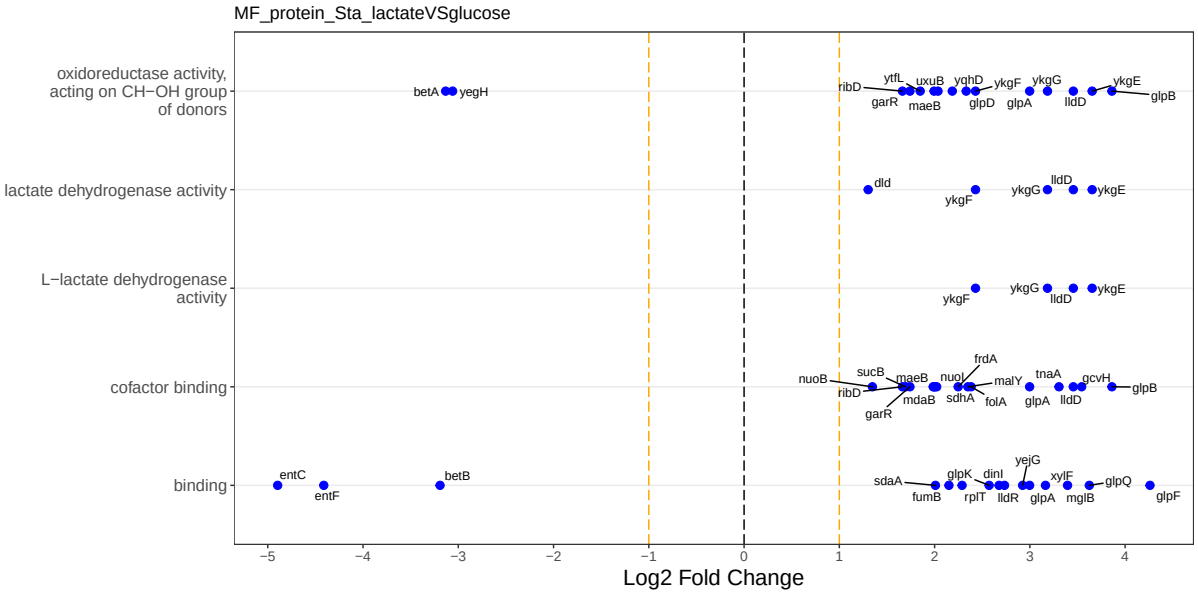

Figure S25: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with lactate as carbon source, as determined by protein abundances in stationary phase.** The top 5 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

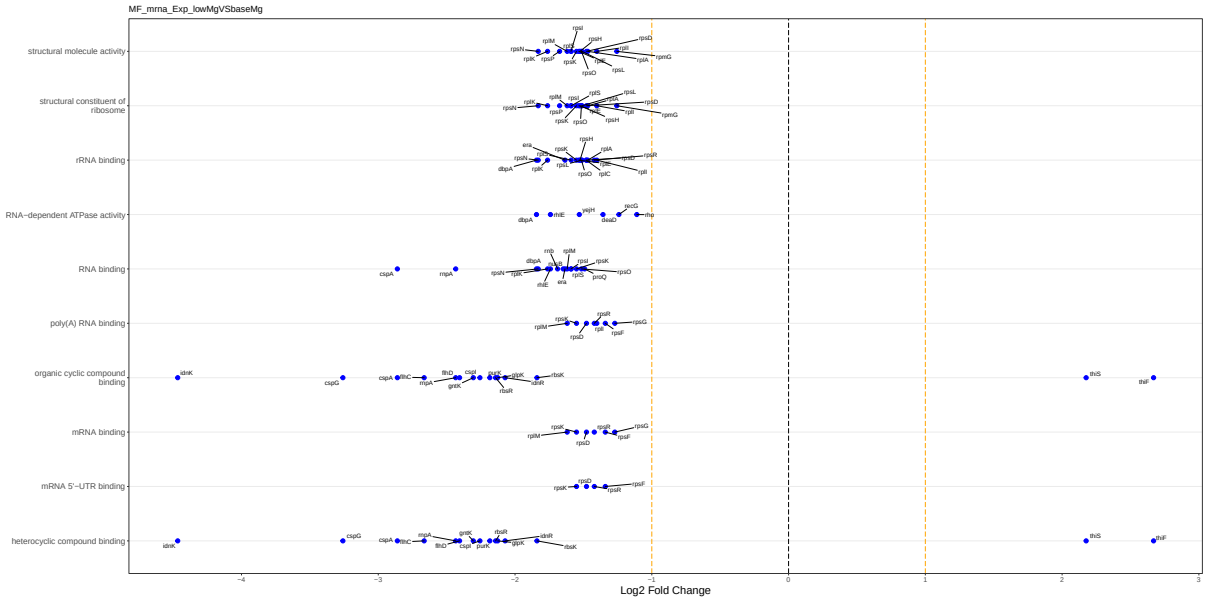

Figure S26: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with low Mg^{2+} levels, as determined by mRNA abundances in exponential phase.** The top 10 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

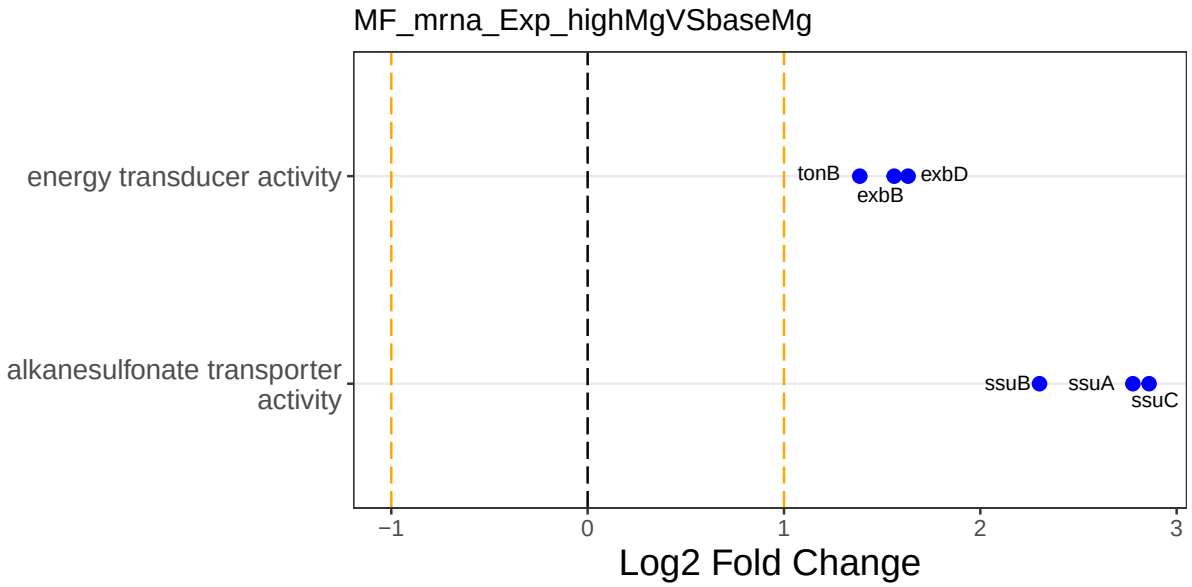

Figure S27: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with high Mg^{2+} levels, as determined by mRNA abundances in exponential phase.** The top 3 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

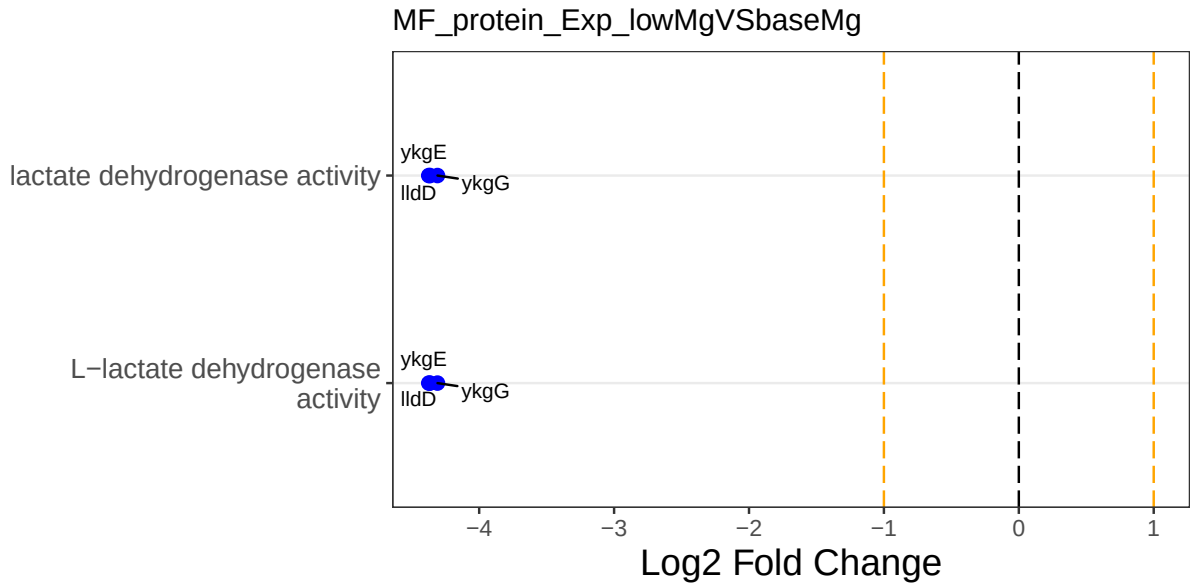

Figure S28: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with low Mg^{2+} levels, as determined by protein abundances in exponential phase.** The top 2 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

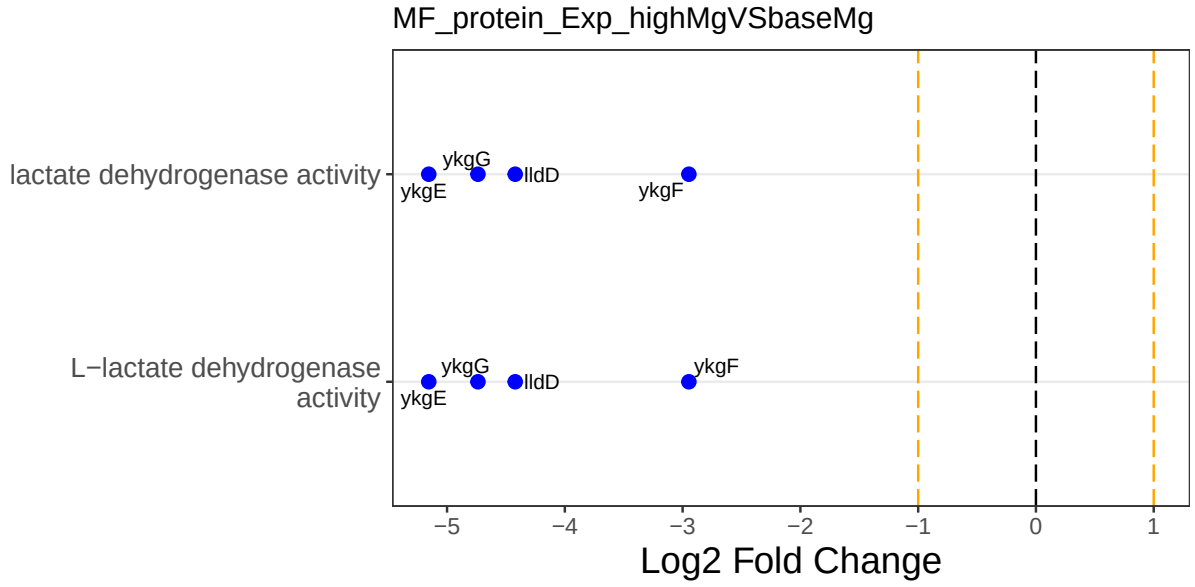

Figure S29: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with high Mg^{2+} levels, as determined by protein abundances in exponential phase.** The top 2 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

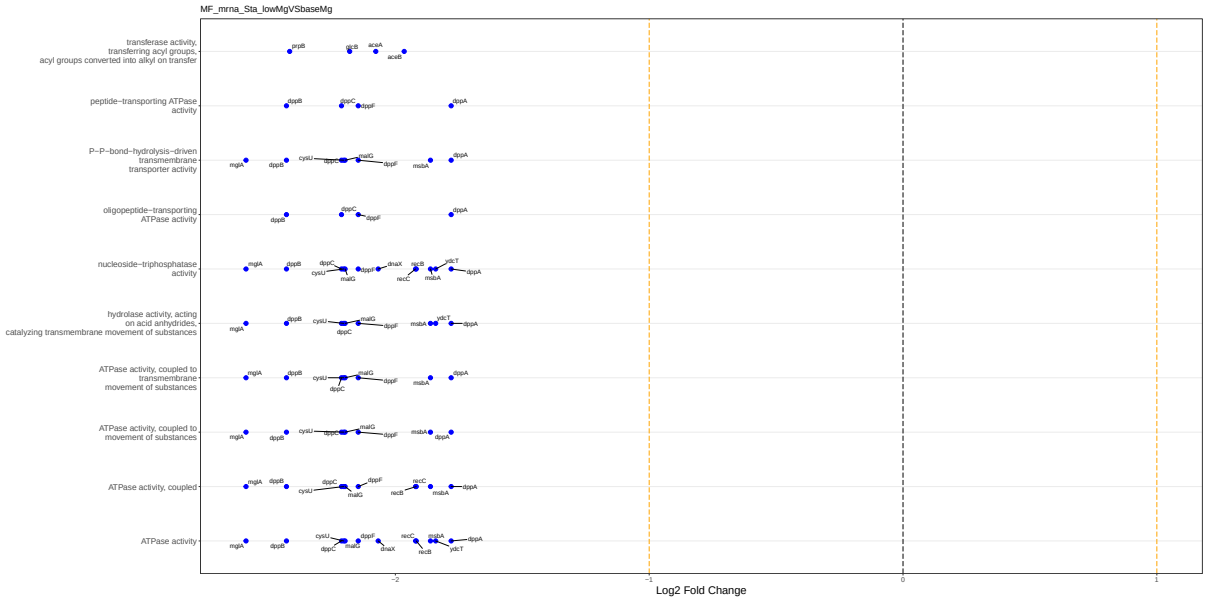

Figure S30: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with low Mg^{2+} levels, as determined by mRNA abundances in stationary phase.** The top 10 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

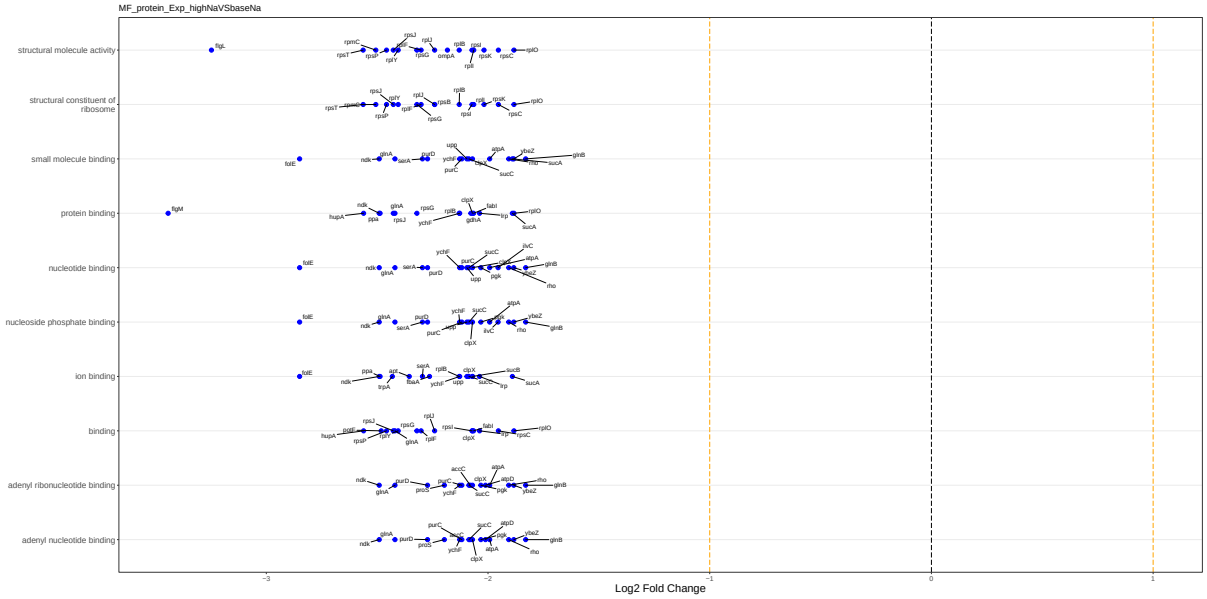

Figure S31: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with high Na^{+} levels, as determined by protein abundances in exponential phase.** The top differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

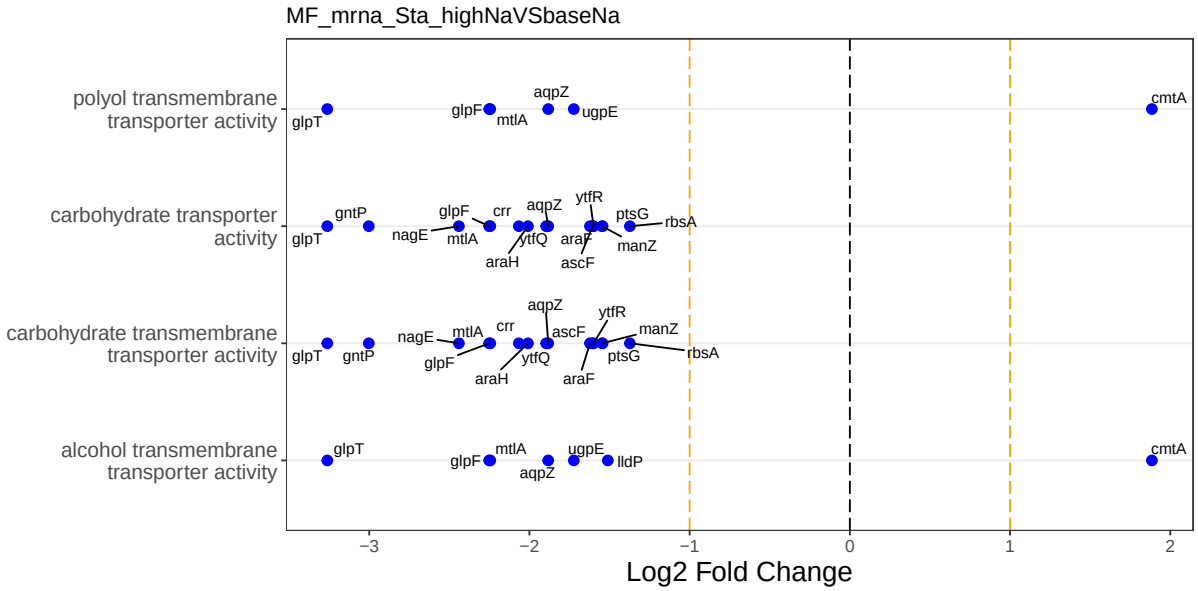

Figure S32: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with high Na^+ levels, as determined by mRNA abundances in stationary phase.** The top 5 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

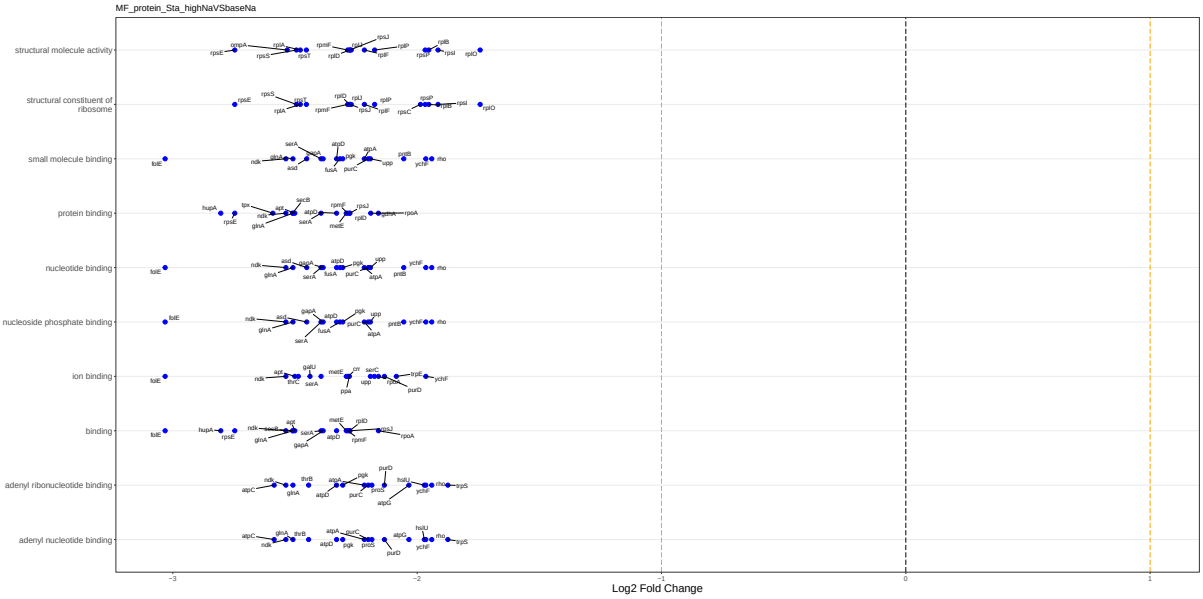

Figure S33: **Significantly differentially expressed GO annotations related with molecular functions and associated genes with high Na^+ levels, as determined by protein abundances in stationary phase.** The top 10 differentially expressed molecular functions are shown along the y axis, and the relative fold change of the corresponding genes is shown along the x axis. We show up to 10 of the most significantly changed molecular functions and for each molecular function, we show up to 15 of the most significantly changing genes.

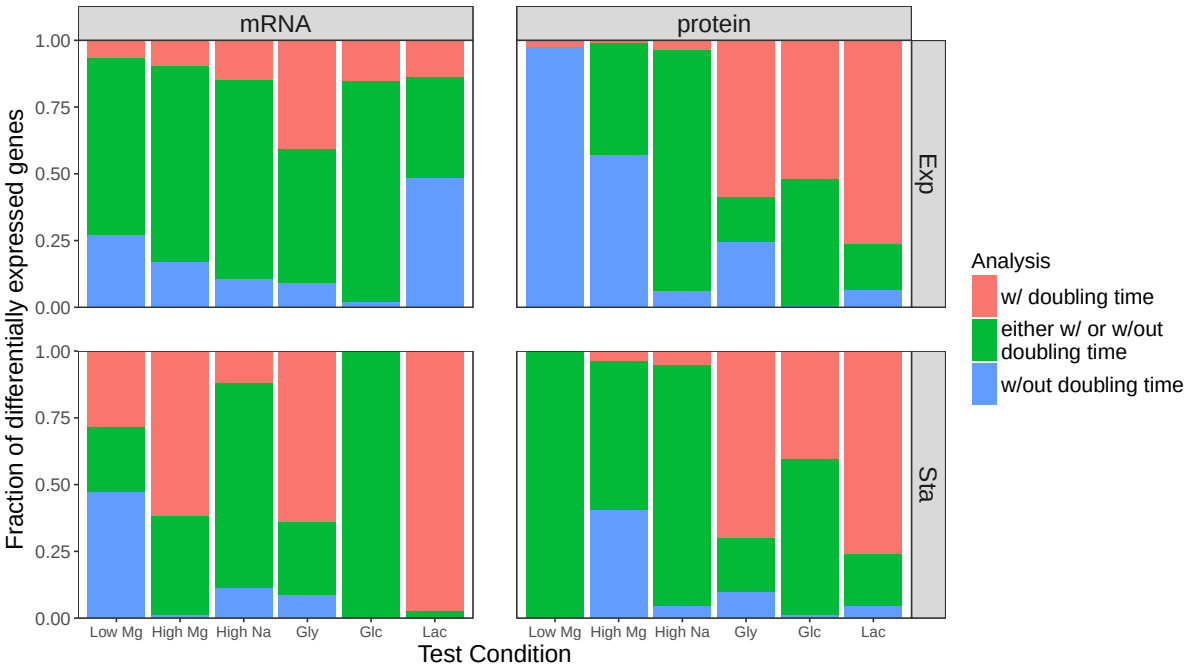

Figure S34: **Fraction of differentially expressed genes that are found in analyses with or without controlling for doubling time.** Shown are the fractions of genes identified as differentially expressed only when controlling for doubling time (red), only when not controlling for doubling time (blue), or in both cases (green). Combined with the absolute numbers of differentially expressed genes in the various conditions (Figure 5), we can see that the main differences in analyses with or without doubling time arise for protein abundances analyzed with respect to different carbon sources.

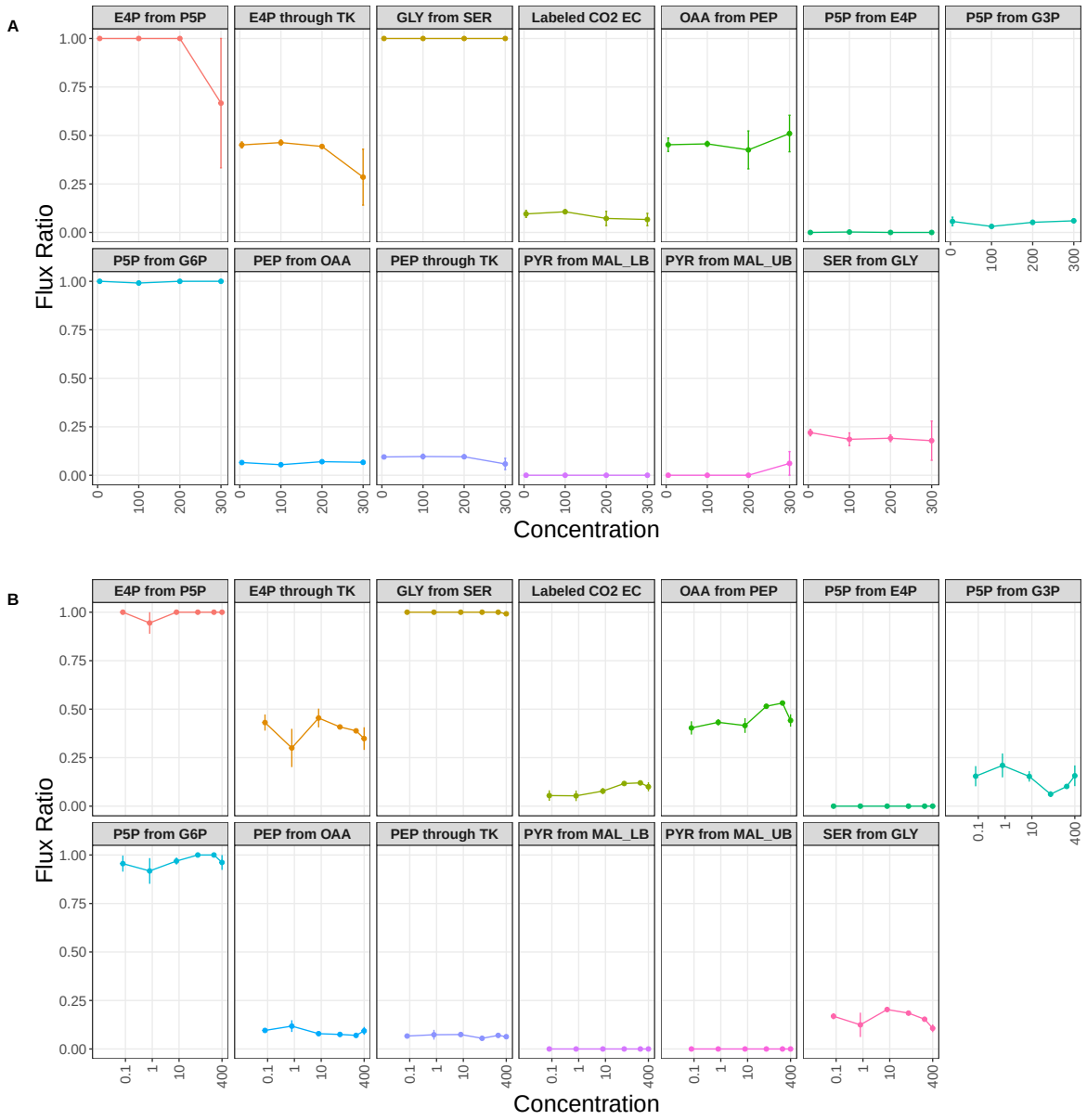

Figure S35: **Flux ratios versus ion concentrations.** 13 different flux ratios were measured with respect to four different Na^{+} and five different Mg^{2+} concentrations. (A) Concentrations with respect to changing Na+ concentrations. (B) Concentrations with respect to changing Mg^{2+} concentrations. There was no significant trend of increase or decrease in flux ratios with respect to either Na^{+} or Mg^{2+} concentrations (Supplementary Table 13).

Supplementary Tables

Supplementary Table S1

File name: tableS1_meta_data.csv

Meta data table that gives information about all samples. Includes information about MURI numbers, experiment name, the time that sample is collected, harvest date, number of RNA samples (tech. replicates), number of protein samples (tech. replicates), batch numbers, Mg 2+, Na + concentrations, growth phase, Mg2+ and Na+ levels, growth phases, doubling time related information (mean, +-95% confidence intervals, $r\hat{2}$)

Supplementary Table S2

File name: tableS2_mRNA_normalized_raw_data.csv

mRNA data after normalization with DeSeq2 for size factors. Includes information for 4197 distinct proteins, and for 105 samples

Supplementary Table S3

File name: tableS3_protein_normalized_raw_data.csv

Protein data after normalization with DeSeq2 for size factors. Includes information for 4197 distinct proteins, and for 152 samples

Supplementary Table S4

File name: tableS4_fluxData.csv

Raw flux data for 13 branches for exponential and stationary phase

Supplementary Table S5

File name: tableS5_doubling_times.csv

Includes doubling time information for different conditions and replicates, and ±95% confidence intervals.

Supplementary Table S6

File name: tableS6_clustering_mrna_cophenetic.csv

Influence of parameter changes on overall mRNA data. This is measured by clustering scores for individual conditions and sub-conditions including batches for mRNA data

Supplementary Table S7

File name: tableS7_clustering_protein_cophenetic.csv

Influence of parameter changes on overall protein data. This is measured by clustering scores for individual conditions and sub-conditions including batches for protein data

Supplementary Table S8

File name: tableS8_combinedOutputDF_DeSeq.csv

DeSeq2 enrichment analyse results for all genes and for all distinct tests. Table contains information about :

- Id (ECB number for mRNA and YP number for proteins), and corresponding gene name
- Direct outcomes of DeSeq2 such as base mean value, log2FoldChange, ifcSE, stat, pvalue, padj.
- Direction of the change wrt base level ("+1" for increase "-1" for decrease)
- Data type (mRNA or protein) growth phase

- What is tested; base value and contrast.
- Individual output file name
- Carbon Source, Mg^{2+} and Na^{+} levels and growth phase of test data
- Control parameters of the test (batch or batch + growth rate)

Supplementary Table S9

File name: tableS9_combinedDifferentiallyExpressedGenes_DeSeq.csv

Filtered version of Supplementary Table S8 with P<0.05 and log2FoldChange>2

Supplementary Table S10

File name: tableS10_combinedResultList_DAVID.csv

The DAVID web service outputs for KEGG and MF tests. Includes information for 24 tests which were controlled for batch effect

Supplementary Table S11

File name: tableS11_changed_protein_carbonSource_ExpSta.csv

Newly appeared genes when we control for "batch + growth rate" instead of "batch"; for proteins, and tested for carbon source. For both exponential and stationary phase

Supplementary Table S12

File name: tableS12_changed_DAVID_P05.csv

Enriched KEGG pathways and molecular functions based on the genes listed on Supplementary Table S11

Supplementary Table S13

File name: tableS13_flux_vs_conc_Pvalues.csv

P values obtained by likelihood maximization for flux vs concentration for individual flux branches with individual phase (either exponential or stationary) and with individual salt (Mg^{2+} or Na^{+})

Supplementary Table S14

File name: tableS14_flux_vs_doublingTime_Pvalues_tog.csv

P values obtained by likelihood maximization for flux vs doubling time for individual flux branches