Alice Finton

BIOL 588

January 31, 2020

**Outline**

**Title:** Modeling of Gene Regulatory Network Dynamics Reveals Important Regulatory Relationships that Control the Cold Shock Response in *Saccharomyces cerevisiae*

**Introduction**

* 1. Introduction to *Saccharomyces cerevisiae* and why it was used in this study
     1. The entire yeast genome has been sequenced and all 6189 genes are known. Yeast is an ideal model organism for use in gene expression experiments. The eukaryote is also ideal for use as a model for humans due to homologous genes (Botstein et al., 1997)
     2. Wide use of yeast deletion strains in experiments to determine the importance of each gene to the organism’s function.
        1. Gene expression data is easier to quantify now due to the development of DNA microarray technology
  2. Environmental stress response (salt, oxidative, … temperature) (Causton et al., 2001)
     1. Yeast must respond to environmental fluctuations in order to survive.
        1. The cells must respond both metabolically and physiologically in order to maintain homeostasis. (Madigan et al., 2006)
     2. The general environmental stress response is defined by changes to gene expression which help the organism adapt to the stressor. (Gasch et al., 2004)
  3. Cold-shock response in yeast (Aguilara et al., 2007)
     1. The optimal growth temperature for *S. cerevisiae* is 30°C, with cold-shock temperatures occurring at 10-18°C. (Salvado et al., 2011)
     2. Growth in cold-shock temperatures causes a decrease in the cell’s membrane fluidity, slowing of cellular processes and kinetics, and impaired protein synthesis.
        1. Yeast respond to cold-shock through global transcriptional changes (Murata et al., 2006)
        2. In a previous study, it was determined that low temperature conditions shunt yeast cell growth, causing a decrease in doubling time (Rodriguez-Vargas et al., 2002).
  4. DNA microarray
     1. DNA microarray technology measures the expression of all *S. cerevisiae* genes*,* which allows for the analysis of the global transcription response to various stimuli, including temperature.
     2. Each of the 6189 genes in the yeast genome are included on the DNA microarray slide, allowing for the transcriptional response in each of the genes to be determined. Using this technology, along with statistical analysis, the impact of cold shock on the expression of genes in the yeast genome can be determined.
  5. A gene regulatory network (GRN) depicts the regulatory relationship between transcription factors illustrating gene expression mechanism control. The dynamics of GRNs can offer information about the changes in gene expression over time in response to various environmental stimuli.
     1. Mathematical modeling of transcriptional response with GRNmap and the depiction of gene regulatory networks with GRNsight (Dahlquist et al., 2015)
        1. A previous study determined that a nonlinear approach is better at predicting the regulatory relationship between selected genes and determining the intensity of the activation or repression compared to a previously proposed linear differential equation approach. (Vu and Vohradsky, 2006).
        2. A different type of mathematical model which can be used to model regulatory relationships between transcription factors uses a differential equation based local dynamic bayesian network (DELDBN), which uses ordinary differential equations in combination with bayesian analysis to elucidate the dynamics of gene regulatory networks. This approach improved upon the ODE parameter estimation. (Li et al., 2011)
        3. Another approach to parameter estimation was proposed by Nariai et al. which uses additional data, such as protein-protein interactions in addition to functional categories and expression data to produce accurate gene regulatory networks. Results from the study indicate that the use of other data rather than solely the microarray data generates a more accurate GRN, with the majority of the gene regulatory relationships being found (Nariai et al., 2005).
  6. Lauren Kelly previous study
     1. In a previous study conducted by the Dalhquist lab, a systematic deletion of the edges in the db5 network was followed by a re-estimation of parameters. From the study, it was determined that some resulting models performed better, some performed worse, and some performed about the same as the intact network.
     2. To determine the importance of each edge in the intact network, the edges of the GRN were systematically deleted, one at a time, generating 28 new GRNs, whose parameters were then estimated by GRNmap. In these 28 edge-deletion networks, LSE:minLSE ratios indicated that five networks performed better than the intact network, while ten networks performed about the same, and thirteen performed worse. The edge-deletions involving the Hmo1, Msn2, and Cin5 transcription factors resulted in a poor performance of the model, indicating that those edges represent important regulatory relationships in the cold-shock response (Kelly, 2019).

|  |
| --- |
| **Fig 1:** Heat map of weight changes after systematic edge deletion. Red indicates an upregulation of the gene, while blue indicates downregulation. The heatmap was organized by increasing LSE:minLSE ratio and clustered by eye. The purple bar designates the intact network. |

* 1. The intact db5 network consists of 15 nodes, which represent transcription factors, and 28 edges, which indicate a regulatory relationship.

|  |
| --- |
| **Fig 2:** Intact (all-strain) db5 gene regulatory network generated using GRNsight. The GRN has 15 nodes (transcription factors) and 28 edges (regulatory relationships). |

* 1. A systematic deletion of mutant strain data resulting in 32 GRNs was conducted to determine the importance of the strain on the fit of the model. A re-estimation of model parameters, including production rates, expression thresholds, and regulatory weights, was performed using differential equations in GRNmap, a MATLAB software. In addition, to assess the sensitivity of the model, a series of runs were conducted in which production rate and/or expression threshold were either fixed values or estimated parameters. Further sensitivity tests involved the manipulation of optimized parameters (kk\_max, MaxIter, TolFun, MaxFunEval, TolX) by one degree of magnitude in order to assess the impact of these on the fit of the model.

**Materials and Methods**

* 1. Cold shock and DNA microarray
     1. Using DNA microarrays, the gene expression data of *Saccharomyces cerevisiae* wild type and deletion strains (dCIN5, dGLN3, dHAP4, dHMO1, and dZAP1) was collected from the Dahlquist lab. Thestrains were subjected to cold shock at 13oC for one hour and subsequent recovery at 30oC for another hour. Samples were taken after 15, 30, and 60 minutes of cold shock and 30 and 60 minutes of recovery. These t15, t30, t60, t90 and t120 samples were then used for analysis.
  2. STEM profile analysis
     1. ClueGO is a Cytoscape plugin that creates networks from the gene ontology terms and allows for a comparison between two lists of data. (Maere et al., 2005)
        1. Interprets biological information and visualizes the functional groups of terms via networks and charts
        2. Uses kappa statistics to create network pathways and link the terms
        3. Creates a network of genes considering the importance of genes initially
        4. The gene lists and gene ontology terms were input into ClueGO using the steps as written here: <https://openwetware.org/wiki/Alice_Finton_Online_Lab_Notebook#How_to_run_ClueGO>:.
        5. An example of a properly formatted gene list can be found here: <https://lmu.app.box.com/file/482634200651>
  3. Variable inclusion of strain data
     1. The list of models can be found here: <https://lmu.app.box.com/file/475332142365>.
     2. GRNmap
        1. GRNmap (Gene Regulatory Network modeling and parameter estimation) is a MATLAB application that uses differential equations to estimate GRN parameters including gene expression thresholds, production rates, and regulatory weights from DNA microarray data. GRNmap models gene expression change as the production of mRNA minus the degradation using a sigmoidal production function. The ratio of the least squares error (LSE) of the model to the minimum theoretical least squares error (minLSE) allows different models to be compared (Dahlquist et al., 2015).
           1. The degradation rates used as initial guesses in the study were derived from the Neymotin paper and production rates were set as two times the degradation rate (Neymotin et al., 2014)

|  |
| --- |
| **Fig 3:** The differential equation for expression over time used in GRNmap is modeled using production rate minus degradation rate. |

|  |
| --- |
| **Fig 4:** Least squares error (LSE) and minLSE offer information about the goodness of fit of the model, giving the difference between the experimental and model values. |

* + - 1. GRNsight is an open source web application that offers the visualization of gene regulatory networks (GRNs). GRNsight utilizes the GRNmap output spreadsheets to visualize GRNs and the intensity of the regulatory relationships (Dahlquist et al., 2016). In this study, GRNsight was used to visualize the GRNs for the variable inclusion of data, clustering of edge deletions, and Schade model runs. The default grid layout was used for all of the networks. The regulatory weights were normalized to the greatest magnitude (3.247) for the Schade model runs.
  1. K-means clustering
     1. MATLAB software was used to cluster using k-means clustering. Excel sheets containing the raw regulatory weights was loaded into MATLAB and clustered via k-means clustering. (MATLAB script can be found here: <https://lmu.app.box.com/file/607545631964>)

|  |
| --- |
| **Fig 5:** Heatmap of the systematic deletion study, clustered using k-means clustering in MATLAB using three clusters in the y-direction and four in the x-direction. |

* 1. Deletion of GCR2 and ZAP1 data
     1. The model run where GCR2 and ZAP1 data were deleted or floating did not work and MATLAB crashed. To fix this two more models were run with a reduced network where the edges that were not in the clusters were deleted from the network sheets. In addition, GCR2 and ZAP1 data were completely deleted from all of the sheets in the excel workbook. dZAP1 data was completely deleted. In another model run, the network was left intact, but GCR2 and ZAP1 were deleted from the matrix, making a 26-edge network.
     2. To analyze the fit of the models compared to the intact network, the LSE:minLSE ratios of each of the models were compared.
  2. Schade model runs
     1. Modification of the Schade input file from used in the Dahlquist et al. (2015) paper to fit the format of GRNmap. More runs were performed using modified models where the optimization parameters were changed to the following: the alpha to 0.02, MaxIter to 1.00E+08, TolFun to 1.00E-06, MaxFunEval to 1.00E+08, and TolX 1.00E-06 (This modification was done to all of the subsequent models). A model with the Schade network was run with replicate data added. Another model was run where the Belle et al. (2006) production rates, which were used in previous GRNmap models, was changed to the Neymotin et al. (2014) production rates.. The last model was run with the Schade network, with data changed to Dahlquist wild-type data. The Dahlquist lab data was found using the MS Access Database (Found at: <https://github.com/kdahlquist/DahlquistLab/blob/master/data/Spring2019/Expression-and-Degradation-rate-database_2019.accdb>).
  3. Fix/estimate parameters
     1. Using db1-db7 all-strain data input files, six new modified models per network were created: one where production rate (*P*), threshold (*b*), and regulatory weight (*w*) were all estimated; one where *P* and *b* were estimated and *w* was fixed; one where *w* and *b* were estimated and *P* was fixed; one where *b* was estimated and *P* and *w* were fixed; one where *b* and *w* were estimated and the fixed *P* values were Neymotin rates; and one were *b* was estimated and *w* and *P* (Neymotin rates) were fixed. Each of the models were run with GRNmap and then the LSE:minLSE ratio was used to analyse the difference in fit of the model to determine whether estimating/fixing the values impacts the model.
  4. Optimization parameters sensitivity tests
     + 1. Using db5 all-strain data, the optimization parameters in the input Excel workbook were changed one degree of magnitude each way for kk\_max, MaxIter, TolFun, MaxFunEval, and TolX. The resulting eight new model input files were run through GRNmap and the LSE:minLSE ratios and iteration counts of each model were then compared to determine if the change in value affected the fit of the model.

**Results/ Discussion**

1. Gene ontology analysis
   * 1. An analysis of the STEM profiles revealed that three profiles-- profile 45, 9, and 22-- were common to all of the deletion strains. Other profiles had overlap between the deletion strains; such as profiles 48, 2, 7, 40, 0, and 28; while others were unique to a single strain, like profiles 31 and 38. Profile 45 is characterized by an initial increase in gene expression followed by a decrease during the recovery period. Profile 9 shows consistent down regulation during cold-shock with a slight increase in expression towards the recovery period. Profile 22 shows no change in expression throughout the cold-shock, with upregulation of the genes during the recovery period.

|  |
| --- |
| **Fig 6:** The significant (p-value > 0.05) STEM results for the gene clustering profiles, with gene expression trend lines, indicating commonality between the genes in the cluster. Profiles included are listed in the top row. |

* + 1. ClueGO networks
       1. Using ClueGO software, a gene ontology analysis was conducted on each of the deletion strains within each profile. A comparison between wild-type and the rest of the deletion strains for each profile was conducted to determine if there was a difference. For profile 45, the overrepresented functional categories included those involved in ribosome biogenesis and RNA processing.
          1. In a previous study, ribosome biogenesis and transcription-related genes were found to be upregulated during cold shock, indicating their role in the cell’s adaptation to the low temperature (Al-Fageeh et al., 2006). In the present study, it was found that these were upregulated initially then downregulated during the recovery period. Protein synthesis and processing has been found to be downregulated during this period (Murata et al., 2006).

|  |
| --- |
| **Fig 7:** Overrepresented functional categories for each of the strains in profile 45. |

* + - * 1. The overrepresented functional categories that were common to the strains in profile 9 were involved in catabolic and metabolic processes, such as amino acid and carboxylic acid metabolism, and cell cycle regulation.

In a previous study, consistently upregulated genes during the time course of cold shock included those involved in energy and metabolism functions, which is inconsistent with these findings (Murata et al., 2006).

|  |
| --- |
| **Fig 8:** Overrepresented functional categories for each of the strains in profile 9. |

* + - * 1. Within profile 22, the common overrepresented functional categories included those involved in cell-aging and stress responses, such as desiccation and oxidative stress. Therefore, during the recovery period, the expression of these genes increased.

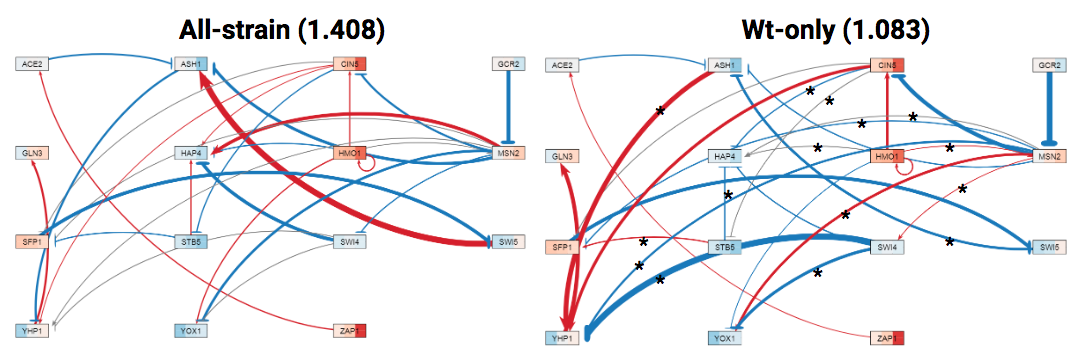
In a previous study, genes associated with cell rescue, defense, and virulence were consistently upregulated during the cold shock response (Murata et al., 2006). This supports the findings of this study.

|  |
| --- |
| **Fig 9:** Overrepresented functional categories for each of the strains in profile 22. |

1. The variable inclusion of strain data runs totaled to 32 new GRNs. The output was analyzed using the LSE:minLSE ratio, expression plots, and network parameters (production rate, regulatory weight, and threshold).
   * 1. In order to assess the fit of the model, the LSE:minLSE ratio was determined and compared between the 32 model runs. A ratio higher than that of the intact network indicates that the model is a worse fit, while a reduced ratio indicates that the model is a better fit than the intact network. The LSE:minLSE ratio increased for eight of the new models. Of these eight models, all of them included the GLN3 deletion strain. The LSE:minLSE ratio decreased for 23 new models, including the wild-type only model and all of the wild-type plus one deletion strain models. Overall, the model run with wild-type only data caused the best fit, while the inclusion of GLN3 deletion strain data caused the model to run about the same as the all-strain network or worse.
        1. The inclusion of dGLN3 data caused the model to perform worse than the all strain model run. Gln3 regulates glutamine metabolism and has been found to be associated with the nitrogen catabolite repression system (Tate et al., 2007, YeastGenome.com).

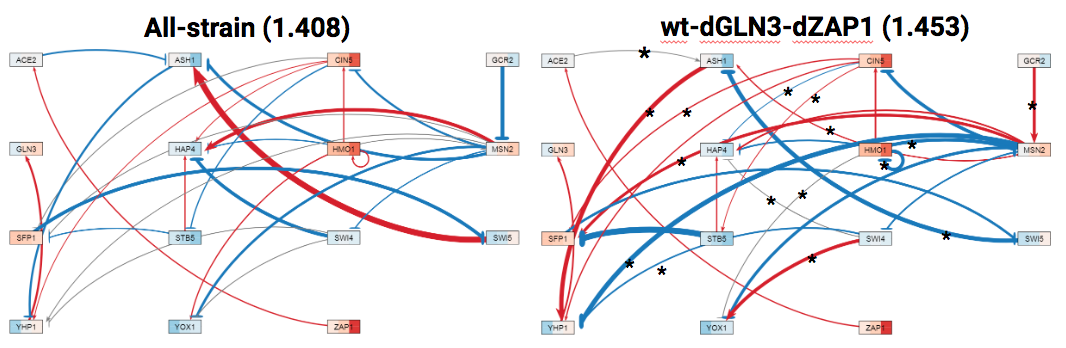
|  |
| --- |
| **Fig 10:** LSE:minLSE ratios for each of the model runs is given in the graph. The models are ordered by increasing LSE:minLSE ratio. The green bar represents the all-strain network and the red line indicates the ratio value for the all-strain network. |

* + 1. GRNs
       1. The intact network was compared to the model that ran the best, wt-only; the model that ran the worst, wt-dGLN3-dZAP1; and the network that did not include GLN3 deletion strain data, wt-dCIN5-dHAP4-dHMO1-dZAP1. Compared to the all-strain network, the wt-only GRN had 15 edges that changed.



**Fig 11:** Gene regulatory network (GRN) generated using GRNsight. The intact db5 network is compared to the wt-only network. The LSE:minLSE ratio is given in the parentheses. The asterisks on the edges of the wt-only GRN indicate edges that have changed from the intact network. 15 of the 28 edges changed.

* + - 1. The wt-dGLN3-dZAP1 model performed the worst, with the highest of the LSE:minLSE ratios. When compared to the all-strain network, 16 of the 28 edges changed.



**Fig 12:** Comparison of the all-strain and wt-dGLN3-dZAP1 GRNs reveals that 16 of the 28 edges changed when dCIN5, dHAP1, and dHMO1 strains were not included. The wt-dGLN3-dZAP1 network performed the worst, with the greatest LSE:minLSE ratio of the models. The ratios are given in the parentheses. The asterisks refer to the edges that changed.

* + - 1. When dGLN3 data was included in the input, the model consistently performed worse. Therefore, a comparison was made between the all-strain and model where only dGLN3 data was not included.

|  |
| --- |
| **(1.327)**  **(1.408)**  **Fig 13:** Comparison of all-strain and wt-dCIN5-dHAP4-dHMO1-dZAP1 GRNs. Four of the 28 edges changed when dGLN3 was not included in the model. The asterisks indicate edges that have changed in the new model. |

* + 1. Heat map clustered
       1. The initial heat map generated with models organized randomly showed no major clusters. Therefore, it was necessary to cluster the heat map using k-means clustering in MATLAB. In clustering the heat map in both dimensions, a cluster containing the all-strain, wt-dCIN5-dHAP4-dHMO1-dZAP1, and wt-dCIN5- dHAP4-dHMO1 model runs was generated. The two models that were most similar to the all-strain’s regulatory weights were the models that either excluded dGLN3 strain data or dGLN3 and dZAP1 data.

|  |
| --- |
| **Fig 14:** Clustered heatmap of regulatory weights with corresponding LSE:minLSE ratio. The heatmap was clustered in the x and y axes, using ten clusters. The purple bar indicates the intact network. Clusters are indicated by the black line. |

* + 1. Expression plots
       1. The inclusion of dCIN5, dHMO1, and dZAP1 data did not cause the simulated model data to diverge, indicating a better fit.
       2. When dCIN5, dHMO1, and dZAP1 data were not included, the simulated model data did diverge, indicating a worse fit.

|  |
| --- |
| **Fig 15:** Expression plots for candidate model runs. The inclusion of dCIN5, dHMO1, or dZAP1 data caused the simulated model to diverge. Divergence is indicated by multiple expression lines, while no divergence results in a single expression line. |

1. Clustering
   * 1. Clustering of Lauren Kelly’s heat map using k-means clustering
        1. K-means clustering was performed on the edge weight values from the intact network and edge-deletion networks. An examination of the clusters showed that weight values deviated from those of the intact network for 6 of 9 edge-deletions involving Msn2, 4 of 5 edge-deletions involving Hmo1, and 3 of 6 edge-deletions involving Cin5. In addition, the edge-deletions involving Gcr2 and Zap1 also caused major deviation in weight values from the intact network. This edge-deletion analysis indicates the importance of regulatory relationships in the cold shock response.
           1. In a previous study, it was found that Msn2 and Msn4 are key transcription factors in the environmental stress response in yeast (Gasch et al., 2000). These TFs recognize the stress response element (STRE) promoter sequence which causes the induction of the stress response (Gasch et al., 2000). Therefore, when the edge deletions included Msn2, they deviated from the all-strain due to the importance of this TF in controlling the stress response.
           2. Gcr2 is a transcription factor which activates glycolytic genes, thereby increasing carbohydrate metabolism and energy production. In a previous study, Gcr2 was found to be the only consistently upregulated stress-response regulator in all of the environmental stress conditions (Chen et al., 2009). Therefore, it is an important regulator of the environmental stress response.
           3. In a previous study, the deletion of the Hmo1 gene caused a decrease in the transcription of genes indirectly associated with ribosome biogenesis (Berger et al., 2007). Therefore, Hmo1 is an important transcription factor in the cold shock response, as ribosome biogenesis has been shown to be an overrepresented upregulated functional category.
           4. In a previous study, Cin5 has been found to be activated when there is a change in the environment, whether that be temperature, carbohydrate availability, etc.. It has also been associated with increasing ribosome biogenesis and protein synthesis (Uniprot, 2019).

|  |
| --- |
| **Fig 16:** k-means clustering resulted in three distinct clusters. The intact network is included in the second cluster. Edge deletions including Gcr2, Msn2, and Zap1 were consistently found in clusters one and three, indicating deviation from the intact network. |

1. Exclusion of Gcr2 and Zap1 from the data and re-run
   * 1. From the clustering, edge deletions that included Gcr2 and Zap1 caused major deviations from the intact network. Therefore, a new set of model runs were conducted, with the exclusion of Gcr2 and Zap1 from the network. When all edges that were not included in the main cluster were deleted from the network, the LSE:minLSE ratio was about the same as the intact network, indicating that there was not a major difference in the fit of the model. When only two edges involving Gcr2 and Zap1 were deleted, the LSE:minLSE ratio decreased, indicating a better fit of the model. When dZAP1 data was deleted from the model, the LSE:minLSE ratio decreased slightly, indicating a slight better fit.

|  |
| --- |
| **Fig 17:** LSE:minLSE ratios for the model runs: the 16-edge reduced network (16-edge), intact network without GCR2 and ZAP1 (26-edge), no dZAP1 data included (28-edge no Zap1), and the intact db5 network. |

|  |
| --- |
| **Fig 18:** Modified gene regulatory networks where: intact network (1), weights in clusters 1 and 3 were put close to zero (2), and edges not in cluster 2 deleted (4). |

1. Schade model runs
   * 1. A new set of models was run using the Schade network and input sheets with the following changes: optimization parameters changed, replicate data added, Neymotin production and degradation rates used with no replicate data, Neymotin production and degradation rates used with replicate data, and Dahlquist lab wt-only data with Neymotin degradation and production rates. When comparing the resulting GRNs with the original Schade network, there were changes in edge weights for each of the new model runs. Certain edges changed through the intensity of the weight, while others flipped activation/repression completely.

|  |
| --- |
| **Fig 19:** Schade model runs where: demo network (1), optimization parameters changed (2), replicate data added (3), Neymotin production and degradation rates with no replicate data (4), Neymotin production and degradation rates with replicate data (5), and Dahlquist wt-only DNA microarray data with Neymotin production and degradation rates (6). |

* 1. Production rate comparison with the Neymotin values
     + 1. Production rates
          1. The initial guess for production rate was consistently lower than the Neymotin and all-strain estimated production rate for ten of the fifteen included genes. (All production rate graphs are found in the appendix)

|  |
| --- |
| **Fig 20:** Production rates of the variable inclusion of data model runs along with the initial guess (orange), all-strain production rate (green), and the Neymotin production rate (red). |

* + - 1. Using a scatter plot, the similarities between the initial guess or estimated production rates and the Neymotin rates could be visualized. A slight linear relationship could be seen, indicating there is some commonality between the production rates, but this relationship is small with R2 values of 0.1594 and 0.1404 for the estimated vs. and initial guess vs. Neymotin production rates.

|  |
| --- |
| **Fig 21:** Neymotin production rates are compared with the estimated (top) and initial guess (bottom) production rates. |

1. Fix/Estimate Parameters runs

|  |
| --- |
| **Fig 22:** LSE:minLSE ratio for the estimate/fix production rate (P), weight (w), and threshold (b) model runs. Production rates used were the initial guess rate or the Neymotin (2014) production rate (designated by “Neymotin”). For db1-db7, fixing production rate caused a greater ratio, the greatest of which were fixing the Neymotin production rates. |

* + 1. It was determined that the use of the Neymotin values made the LSE:minLSE ratio higher, indicating a worse fit. For all db1-7, the runs with Neymotin values showed an increase in ratio value compared to the initial guess runs. The ratio was the least for the runs where all parameters were estimated.
    2. Overall, it was determined that fixing any parameter would make the model run worse. Estimating all of the parameters had the lowest LSE:minLSE ratio. The "Fix P" runs performed the worst. Estimating or fixing the threshold did not change the LSE:minLSE value greatly, which indicates that threshold does not have a major impact on the fit of the model.
    3. When the Neymotin production rate was used as the initial guess for production rate in the "Fix P" runs, the resulting LSE:minLSE was greater than the initial fix/estimate runs.

1. Optimized parameters test runs (change kk\_max, MaxIter, TolFun, MaxFunEval, and TolX by order of magnitude)
   * 1. The LSE:minLSE ratio for the run under "normal" conditions (mid-point value for each parameter) is 1.4081816. When the parameters were changed by one magnitude in either direction, the ratio changed only for kk\_Max at 10 and TolX at 1.00E-07. The other runs caused no difference in the ratio. The change that did occur for kkMax and TolX was very small though.
     2. The iteration count for the run under "normal" conditions (mid-point value for each parameter) is 111,242. Again, only changing the kk\_Max to 10 and the TolX to 1.00E-07 caused a change in the iteration count. However, the change in iteration count for these parameters was greater than the change that occurred for the LSE:minLSE ratio, especially for kk\_Max.
     3. These results indicate that changing the optimization parameters does not have a great impact on the fit of the model, except for when kk\_max and TolX were changed.

|  |
| --- |
| **Fig 23:** The LSE:min LSE ratio remained the same for all of the optimization parameter tests except for when kk\_Max was set to 10 and TolX was set to 1.00e-07 in which case, the ratio decreased. Likewise, the iteration count remained the same except for these same optimization parameter changes. The iteration count increased in both of these cases. |

**References**

Aguilera, J., Randez-Gil, F., & Prieto, J. A. (2007). Cold response in *Saccharomyces cerevisiae*:

new functions for old mechanisms. *FEMS microbiology reviews*, *31*(3), 327-341. Doi: 10.111/j.1574-6976.2007.00066.x

Al-Fageeh MB, Smales CM. Control and regulation of the cellular responses to cold shock: the

responses in yeast and mammalian systems. *Biochem J*. 2006;397(2):247–259. doi:10.1042/BJ20060166

Belle, A., Tanay, A., Bitincka, L., Shamir, R., & O’Shea, E. K. (2006). Quantification of protein

half-lives in the budding yeast proteome. *Proceedings of the National Academy of Sciences*, *103*(35), 13004-13009.

Berger, A. B., Decourty, L., Badis, G., Nehrbass, U., Jacquier, A., & Gadal, O. (2007). Hmo1 is

required for TOR-dependent regulation of ribosomal protein gene transcription. *Molecular and cellular biology*, *27*(22), 8015-8026.

Botstein, D., Chervitz, S. A., & Cherry, M. (1997). Yeast as a model

organism. *Science*, *277*(5330), 1259-1260.

Causton, H. C., Ren, B., Koh, S. S., Harbison, C. T., Kanin, E., Jennings, E. G., … Young, R. A.

(2001). Remodeling of yeast genome expression in response to environmental changes. *Molecular biology of the cell*, *12*(2), 323–337. doi:10.1091/mbc.12.2.323

Chen, T., Li, F., & Chen, B. S. (2009). Cross-talks of sensory transcription networks in response

to various environmental stresses. *Interdisciplinary sciences: computational life sciences*, *1*(1), 46-54.

Dahlquist, K. D., Fitzpatrick, B. G., Camacho, E. T., Entzminger, S. D., & Wanner, N. C. (2015).

Parameter estimation for gene regulatory networks from microarray data: cold shock response in *Saccharomyces cerevisiae*. *Bulletin of mathematical biology*, *77*(8), 1457-1492.

Dahlquist, K. D., Dionisio, J. D. N., Fitzpatrick, B. G., Anguiano, N. A., Varshneya, A.,

Southwick, B. J., & Samdarshi, M. (2016). GRNsight: a web application and service for visualizing models of small-to medium-scale gene regulatory networks. *PeerJ Computer Science*, *2*, e85.

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., ... &

Brown, P. O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Molecular biology of the cell*, *11*(12), 4241-4257.

GLN3. (n.d.). Retrieved from https://www.yeastgenome.org/locus/S000000842

Kelly, L. M. (2019). *Modeling of Gene Regulatory Network Dynamics Predicts which*

*Regulatory Relationships are Important for Controlling the Cold Shock Response in*

*Saccharomyces cerevisiae* (Unpublished doctoral dissertation). Loyola Marymount University, Los Angeles, CA.

Li, Z., Li, P., Krishnan, A., & Liu, J. (2011). Large-scale dynamic gene regulatory network

inference combining differential equation models with local dynamic Bayesian network analysis. *Bioinformatics*, *27*(19), 2686-2691. doi:10.1093/bioinformatics/btr454

Madigan MT, Martinko JM. 2006. Brock biology of microorganisms. Prentice Hall, Upper

Saddle River, NJ

Maere S, et al. (2005) ClueGO: a Cytoscape plugin to assess overrepresentation of Gene

Ontology categories in biological networks. *Bioinformatics*. 21:3448–3449.

Murata, Y., Homma, T., Kitagawa, E. et al. Extremophiles: Genome-wide expression analysis of

yeast response during exposure to 4C (2006) 10: 117. doi:10.1007/s00792-005-0480-1

Nariai, N., Tamada, Y., Imoto, S., & Miyano, S. (2005). Estimating gene regulatory networks

and protein–protein interactions of *Saccharomyces cerevisiae* from multiple genome-wide data. Bioinformatics, 21(suppl\_2), ii206-ii212. doi: 10.1093/bioinformatics/bti1133

Neymotin, B., Athanasiadou, R., & Gresham, D. (2014). Determination of in vivo RNA kinetics

using RATE-seq. *Rna*, *20*(10), 1645-1652. doi: 10.1261

Rodriguez-Vargas, S., Estruch, F., & Randez-Gil, F. (2002). Gene expression analysis of cold

and freeze stress in Baker's yeast. *Appl. Environ. Microbiol.*, *68*(6), 3024-3030.DOI: 10.1128/AEM.68.6.3024-3030.2002

Salvado Z, Arroyo-Lopez FN, Guillamon JM, Salazar G, Querol A, Barrio E. 2011. Temperature

adaptation markedly determines evolution within the genus Saccharomyces. Appl. Environ. Microbiol. 77:2292–2302. http://dx.doi.org/10.1128/AEM.01861-10

Schade, B., Jansen, G., Whiteway, M., Entian, K. D., & Thomas, D. Y. (2004). Cold adaptation

in budding yeast. *Molecular biology of the cell*, *15*(12), 5492–5502. doi:10.1091/mbc.e04-03-0167

Tate, J. J., & Cooper, T. G. (2007). Stress-responsive Gln3 localization in Saccharomyces

cerevisiae is separable from and can overwhelm nitrogen source regulation. Journal of Biological Chemistry, 282(25), 18467-18480.

UniProt ConsortiumEuropean Bioinformatics InstituteProtein Information ResourceSIB Swiss

Institute of Bioinformatics. (2019, December 11). AP-1-like transcription factor YAP4. Retrieved from https://www.uniprot.org/uniprot/P40917

Vu, T. T., & Vohradsky, J. (2006). Nonlinear differential equation model for quantification of

transcriptional regulation applied to microarray data of *Saccharomyces cerevisiae*. *Nucleic acids research*, *35*(1), 279-287. doi: 10.1093/nar/gkl1001

**Appendix:**

<https://docs.google.com/document/d/1wLl84bKPZzHhDU8if-JbpDY-ZwwNRsNavYkZaoVv8xw/edit?usp=sharing>