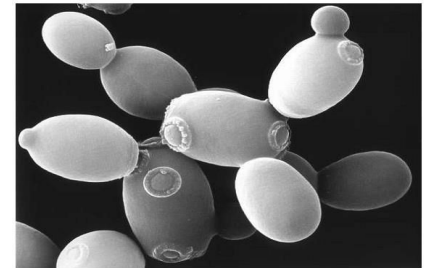


Transcriptional Response of Transcription Factor Deletion Strain dHAP4 of *Saccharomyces cerevisiae* to the Environmental Stress of Cold Shock and Subsequent Recovery

Fatimah Alghanem, Desiree Gonzalez, Ava Lekander, &
Brianna Samuels

Biology 388
March 21st, 2019



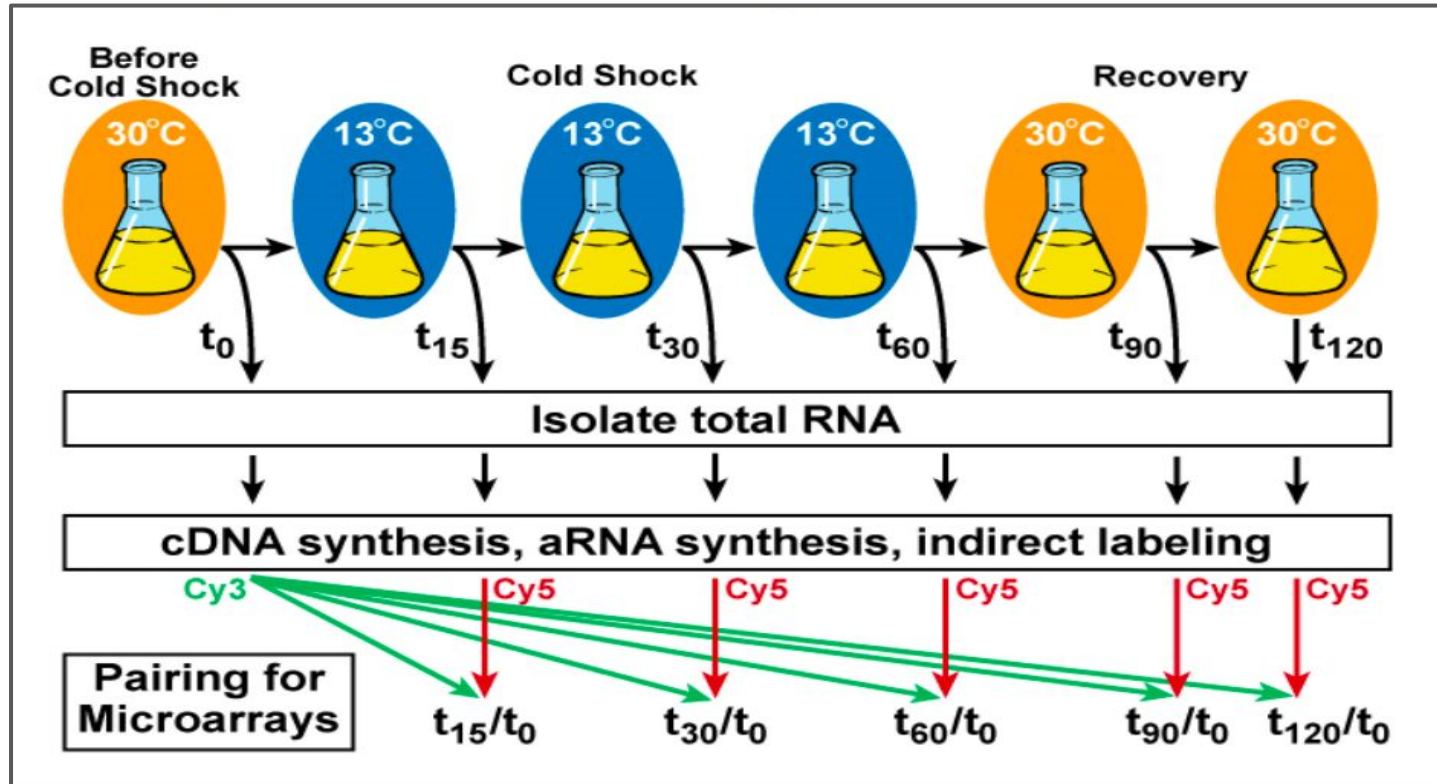
Outline

- Importance of gene regulation in response to cold shock
 - How microarray data helps us understand regulation and expression and how to analyze it
- Flow of methods throughout microarray modeling
- ANOVA table
- Clusters generated
- Unweighted/weighted GRNsight networks
- Modeling results
 - Optimized w , b , and p rates
- Individual expression plots
- Discussion of genes that most likely regulate the cold shock response

Gene Regulation in Response to Cold Shock

- Eukaryotes like yeast use **combinatorial control** to control gene regulation.
 - Yeast typically have 4-5 reg. transcription factors.
 - Humans typically have 12+ reg. transcription factors
- **DNA microarray experiment:**
 - DNA microarrays measure the expression of thousands of genes simultaneously.
 - **Green** = decreased expression relative to control
 - **Red** = increased expression
 - **Yellow** = no change in expression
- Little is known about which transcription factors regulate the early response to cold shock.
- The regulatory model will be used to estimate the relative contribution of each transcription factor to the regulation of gene expression.

Methods for Expression Profiling by Array



Steps Used to Analyze DNA Microarray Data

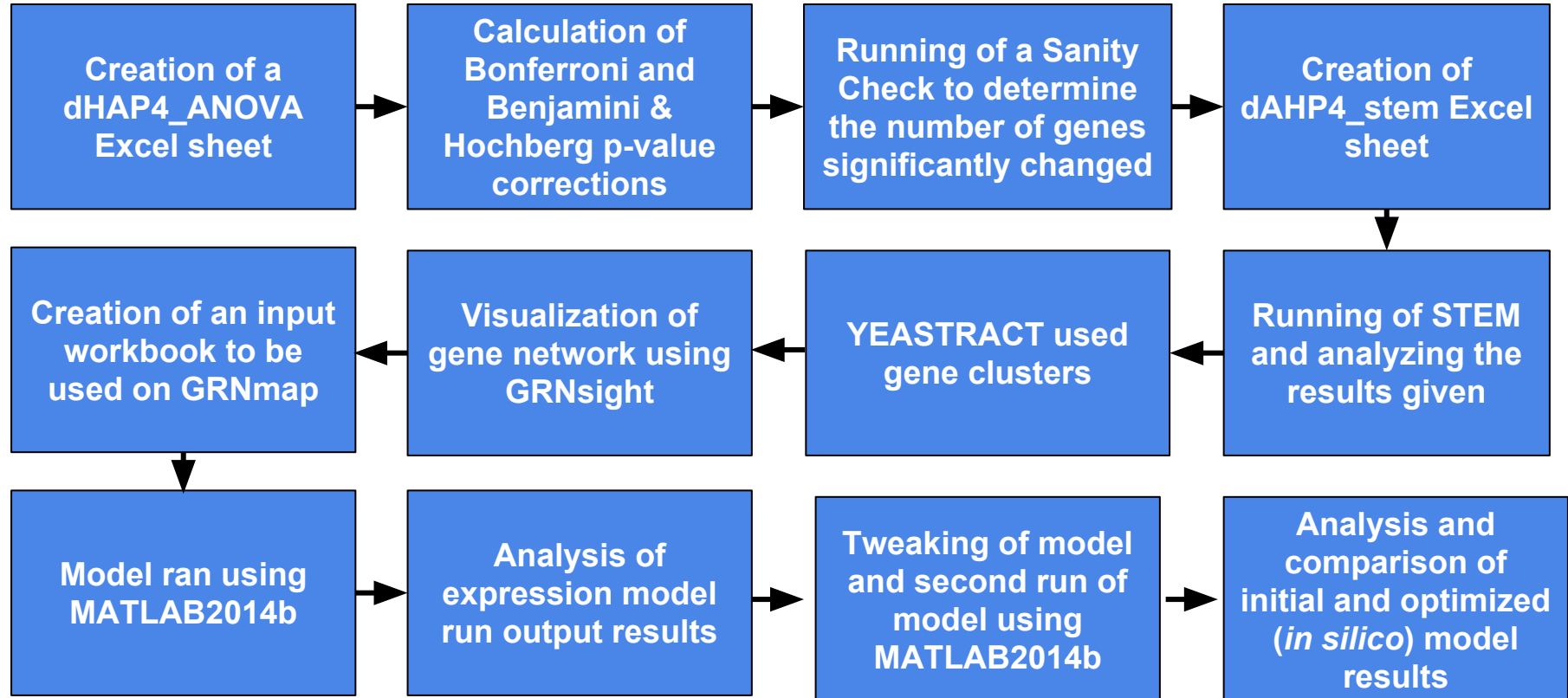
1. Quantitate the fluorescence signal in each spot
2. Calculate the ratio of red/green fluorescence
3. Log2 transform the ratios
4. Normalize the ratios on each microarray slide
5. Normalize the ratios for a set of slides in an experiment
6. Perform statistical analysis on the ratios
7. Compare individual genes with known data
8. Pattern finding algorithms (clustering)
9. Mapping onto biological pathways
10. Identifying regulatory transcription factors responsible for observed changes in gene expression
11. Dynamical systems modeling

Differential Equation

- The differential equation and least squares equation were used to create the input workbook in order to perform initial estimation for parameters from P , w , and b :

$$P_i / 1 + e^{-(\sum w_{ij}x_j - d_{xi})} = dx_i/dt$$

Flow of Methods Throughout Microarray Modeling



ANOVA Results

Table 1: Numerical counts of genes for the strain dHAP4 at the differing significance levels

ANOVA	dHAP4
p<0.05	2479 (40%)
p<0.01	1583 (26%)
p<0.001	739 (12%)
p<0.0001	280 (4.5%)
B & H p<0.05	1735 (28%)
Bonferroni p<0.05	75 (1.2%)

Comparison of Clusters

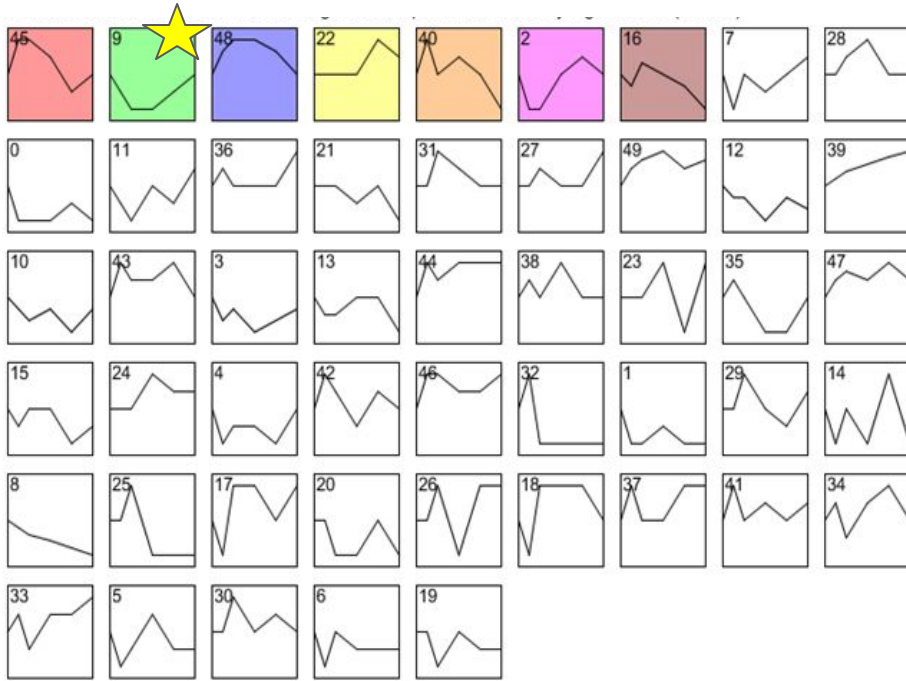


Fig.1: Cluster from the STEM Data Analysis

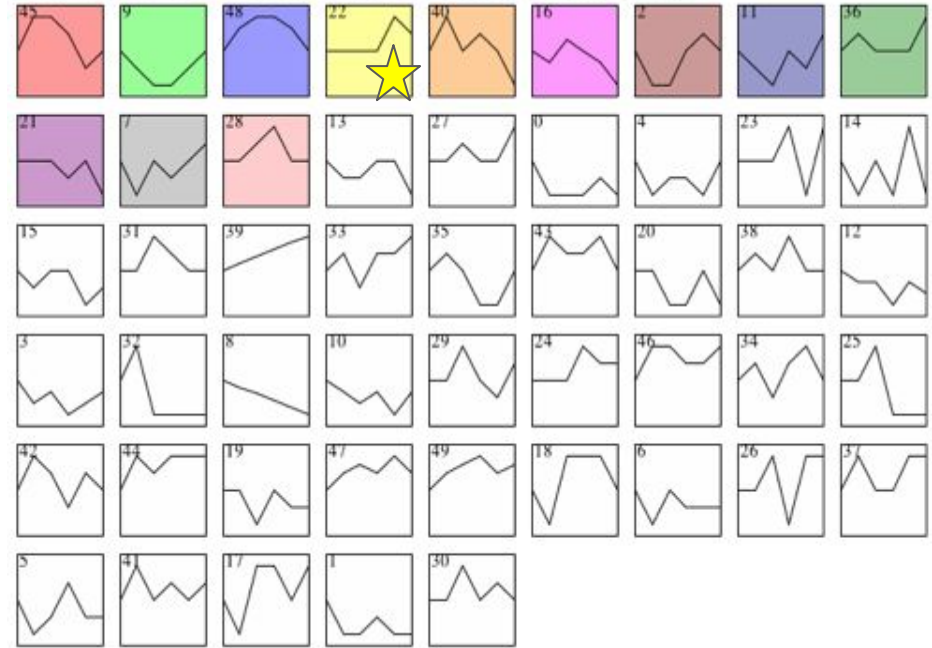


Fig. 2: Fatimah's Cluster from the STEM Data Analysis

Interpretation of GO Results

- 15 transcription factors deemed "significant"
- Expected Gene Number in HAP4: 56.1
- Actual Gene Number in HAP4: 289.0
- GO terms are associated with this profile at $p < 0.05$:
 - 25/88 were found to have a $p < 0.05$.

Determining Fit By Looking at Output Graphs

- Closest Fit Genes:
 - AFT2, ERT1, FHL1, GAL3, GCN4, GLN3, HAP4, IFH1, SUT2, TOD6, YGR067C
- Worst Fit Genes:
 - MPB1, PHO2, SUM1
- There was a total of 15 nodes
- Relationship to Microarray Data
 - Graphical Representation of Log2 Fold Changes over time
 - Slopes relate to activation or repression of genes

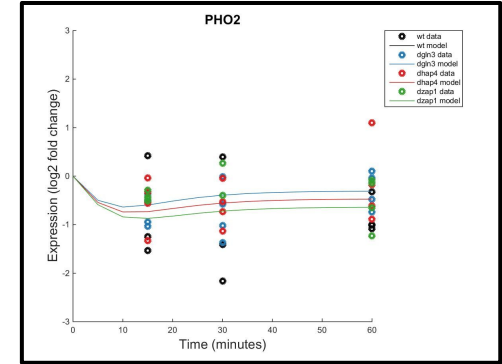


Fig 3: PHO2 expression graph generated through MATLAB.

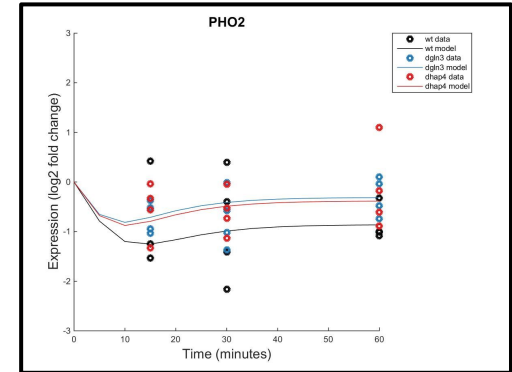


Fig 4: PHO2 optimized expression graph generated with MATLAB.

Initial Gene Network

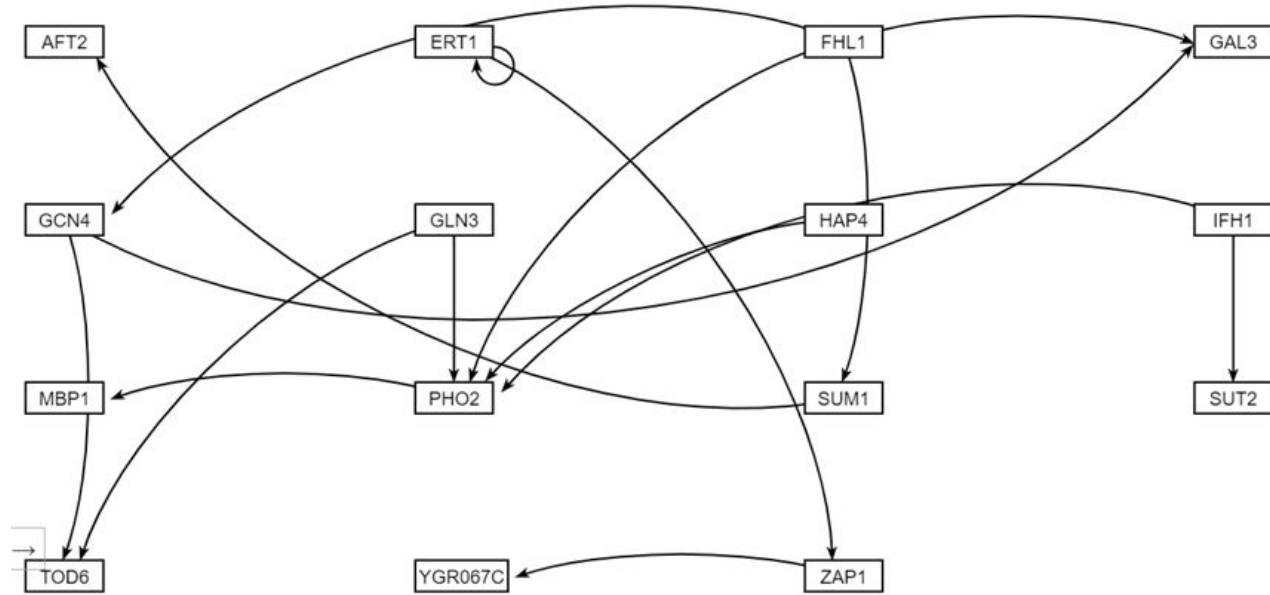


Fig. 5: Desiree's initial gene network obtained using GRNsight (before running the model).

GRNsight Network After running the model

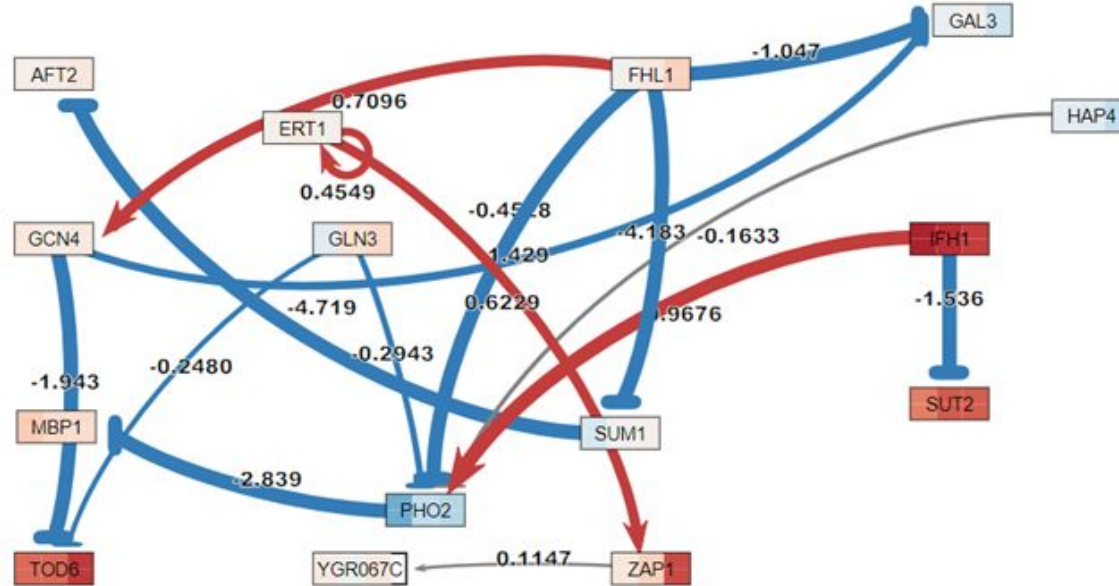


Fig 6: Desiree's GRNsight network after running the model including the weights from one gene to another and the activation/repression relationship between genes.

Optimized Thresholds versus Gene IDs for Profile 9

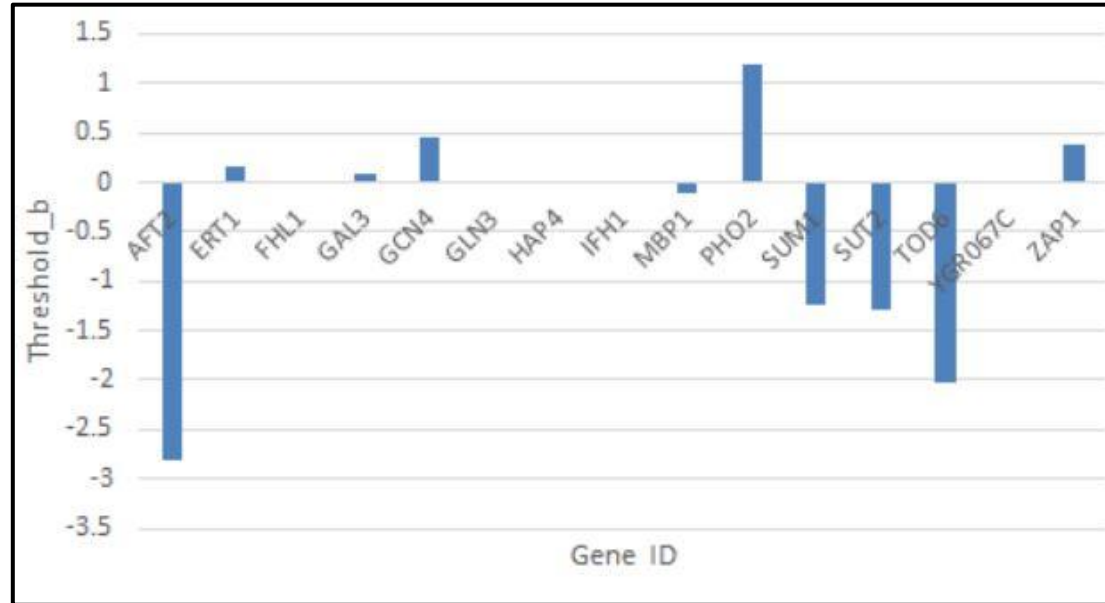


Fig.7 : Bar graph of the optimized Thresholds vs. Gene IDs for dHAP4 Cluster's Profile 9.

Production Rates vs. Gene IDs

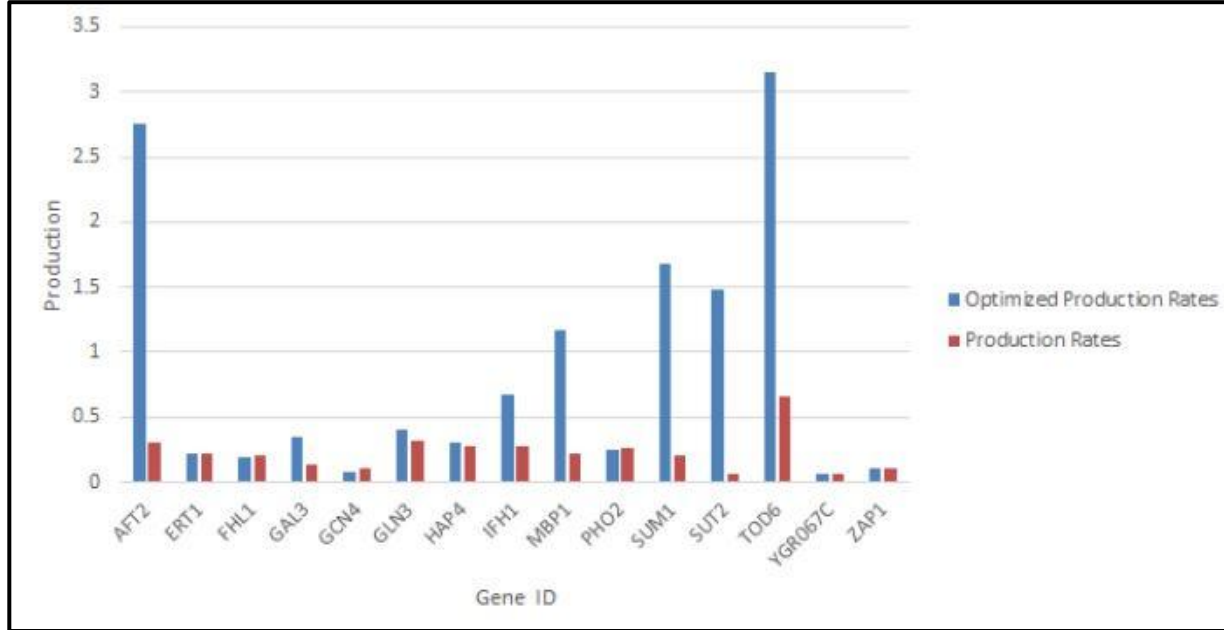


Fig. 8: bar graph of Production Rates vs. Gene IDs in which it shows major differences between the two runs for AFT2, MBP1, SUM1, SUT2, and TOD6.

Optimized Weights vs Gene IDs for Profile 9

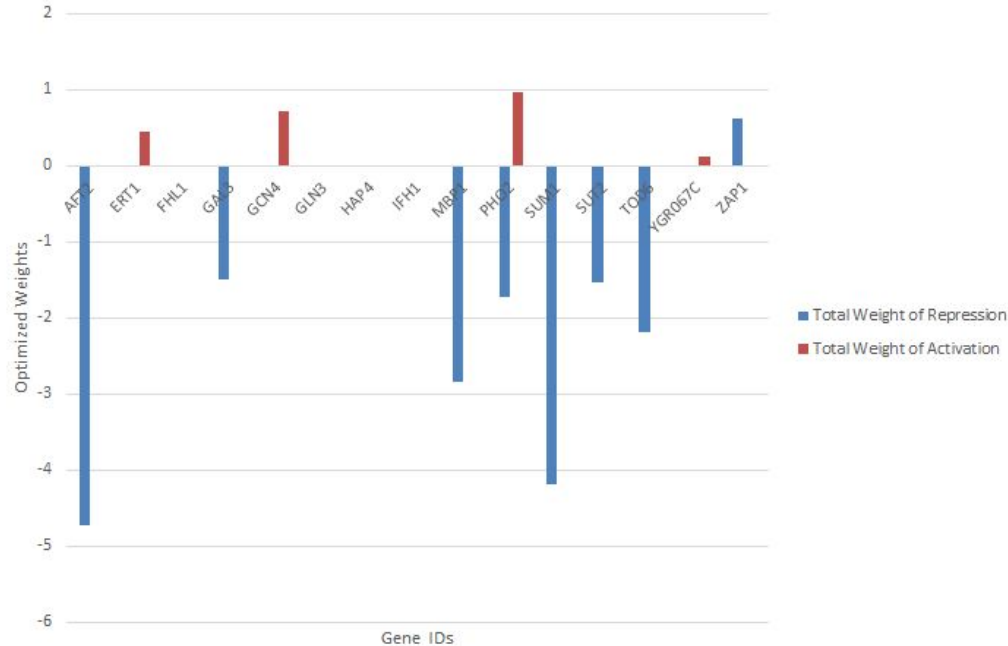


Fig. 9: Optimized weights (with up and down regulations are shown) versus genes IDs for dHAP4 Cluster's Profile 9.

Comparison Between HAP4 & Wildtype

- Genes that showed differences between wild type and dHAP4:
 - FLH1: wt activated; HAP4-repressed
 - SUM1: wt repressed then activated; HAP4 activated then repressed
 - ZAP1: wt activated; HAP4 repressed
 - GAL3: wt repressed, HAP4 activated
- Looking at matlab it does adequately capture the differences
- FLH1, SUM1, and GAL3 all have connections with one

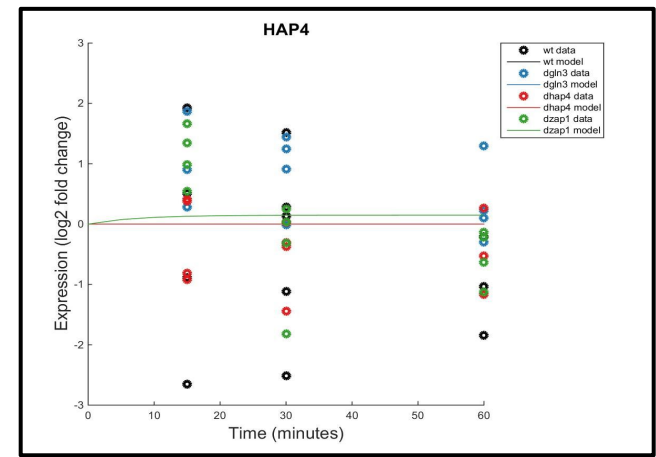


Fig 10: Optimized HAP4 graph generated through MATLAB

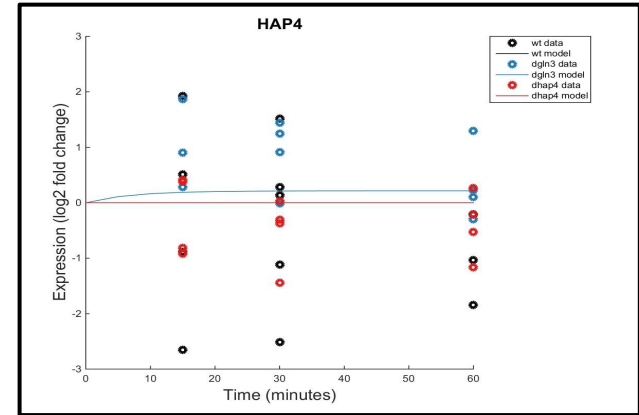


Fig 11: HAP4 graph generated through MATLAB

Genes that Showed the Largest Dynamics Over Time

- MGA2, MSN4, MSN2, GLN3, HAP4, GCN4, GIS1, AFT2, SUT1, RGM1

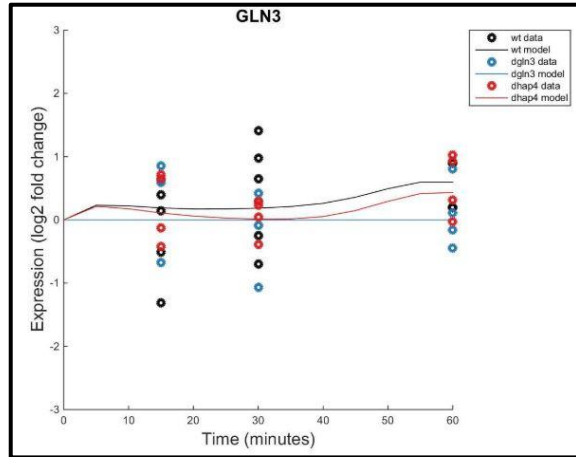


Figure 12: GLN3 expression vs. Time

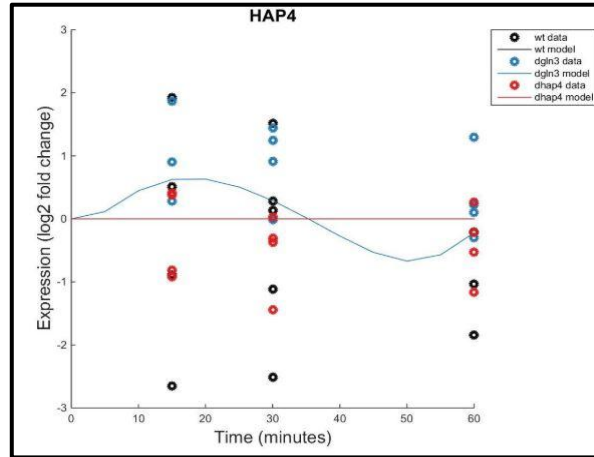


Figure 13: HAP4 expression vs. Time

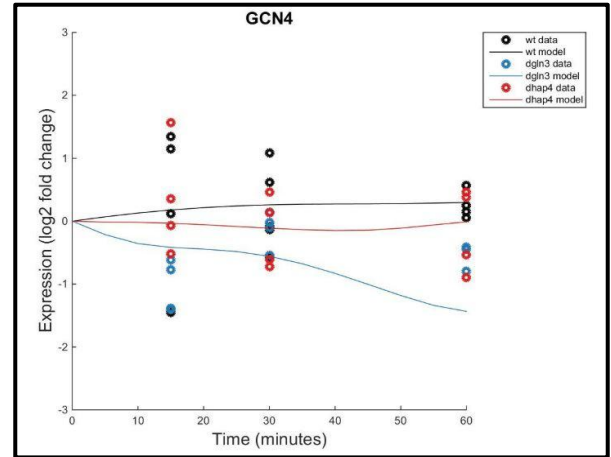


Figure 14: GCN4 expression vs. Time

In Silico Experiments: Deleting ZAP1 and CUP2

- Collectively decided to delete specific transcription factors
 - Brianna, Desiree, Ava- ZAP1
 - Fatimah- CUP2
- Hypothesis
 - No significant effect on the rest of the data if deleted
 - Deletion will result in more precise values for b and P

Least Square Errors

Table 2: The LSE values are represented and shows how well assimilated the actual data is.

	LSE	minLSE	LSE/minLSE ratio
Desiree	0.981743	0.696258	1.41002729
Fatimah	1.0236	0.667	1.53

Before and After Deleting ZAP1

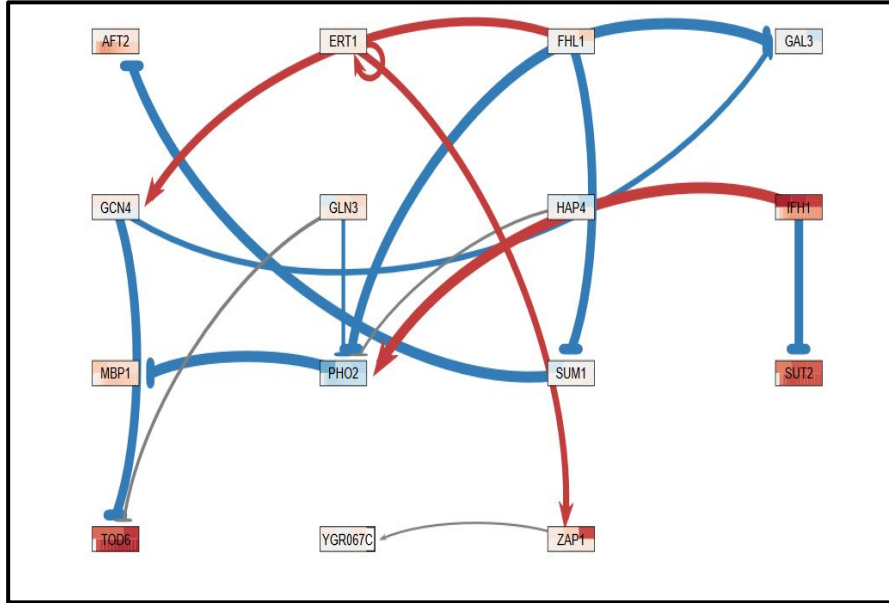


Fig. 15: GRNsight network before deleting ZAP1

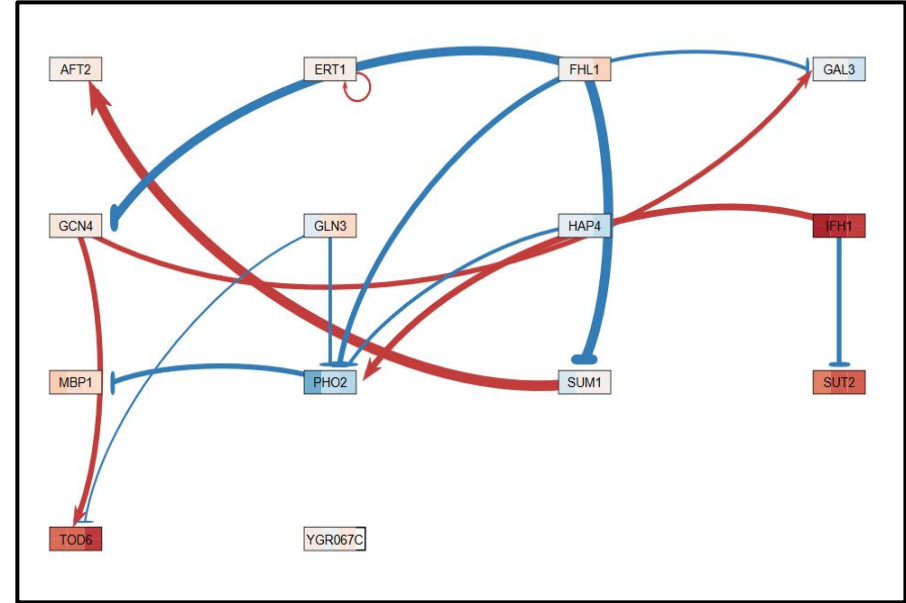


Fig. 16: GRNsight network after deleting ZAP1

Before and After Deleting CUP2

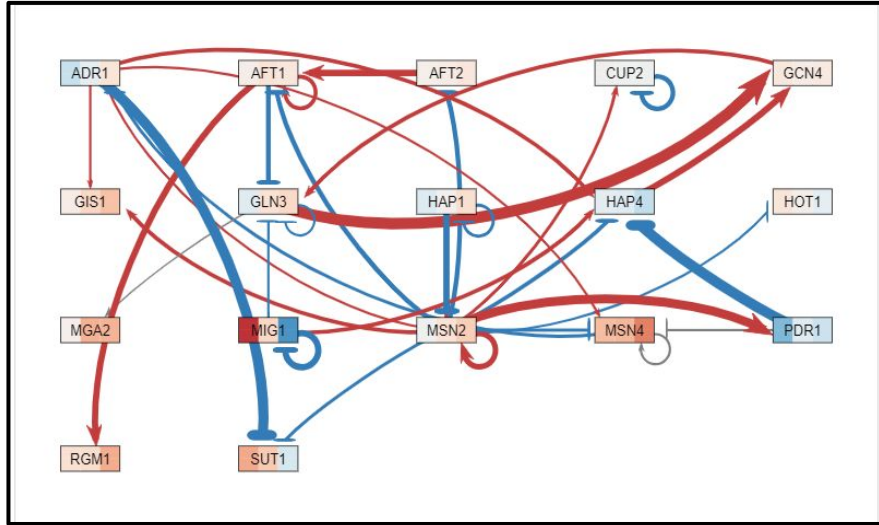


Fig. 17: GRNsight network before deleting CUP2.

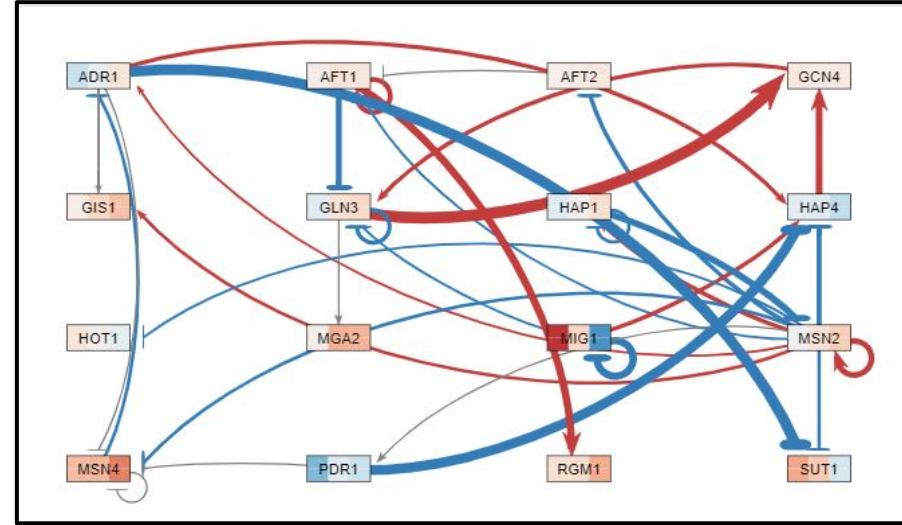


Fig. 18: GRNsight network after deleting CUP2.

Conclusions

- GCN4 and GLN3 most likely regulate the cold shock response when data of all is compared amongst the four of us
 - GCN4: present in $\frac{3}{4}$ gene networks.
 - GLN3: present in all 4 networks and made the most connections with other genes.
- From the *in silico* experiment: the hypothesis was confirmed as the deletion of ZAP1 and CUP1 had no significant effects on the rest of the data.

Future Directions to Continue Experiment

- Get better optimized parameters:
 - Get rid of seemingly unnecessary edges.
 - Use fixed instead estimated parameters.
 - Include info from t90 and t120.
- Run a similar experiment in a different strain of yeast.

Acknowledgments

Department of Biology

Biology 388

Loyola Marymount University

Kam D. Dahlquist, Ph.D.

Ben G. Fitzpatrick, Ph.D.

Dahlquist Lab



Frank R. Seaver College
of Science and Engineering

References

- Dahlquist, K., & Fitzpatrick, B. (n.d.). BIOL388/S19: Week 6. Retrieved from https://openwetware.org/wiki/BIOL388/S19:Week_6
- Dahlquist, K., & Fitzpatrick, B. (n.d.). BIOL388/S19: Week 7. Retrieved from https://openwetware.org/wiki/BIOL388/S19:Week_7
- 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