Dynamical Systems Modeling of the Cold Shock Response in *Saccharomyces cerevisiae*

Katrina Sherbina[1], Nicholas A. Rohacz[2], Kam D. Dahlquist[3], Ben G. Fitzpatrick[1]

[1] Department of Mathematics, [2] Department of Chemistry and Biochemistry, [3] Department of Biology, Loyola Marymount University, 1 LMU Dr., Los Angeles, CA, 90045

Yeast respond to cold shock by changing gene expression

- Transcription factors are regulatory proteins that are encoded by genes.
  - Activators increase gene expression, while repressors decrease gene expression.
- Genes are turned on or off by the binding of transcription factors to regulatory DNA sequences located on the genes.
- Our goal is to determine the influence of a network of transcription factors on gene expression during cold shock.

DNA microarrays were used to determine expression for all 6000 genes in yeast

- Samples were subjected to cold shock for an hour, and then allowed to recover for an hour after being removed from the cold.
- Data was collected for 5 strains: wild type, Δ*cin5*, Δ*gln3*, Δ*hmo1*, and Δ*zap1*.
- Samples were harvested at 0,15, 30, 60, 90, and 120 minutes.
- 3-5 replicates were performed for each time point for a total of 103 microarrays.

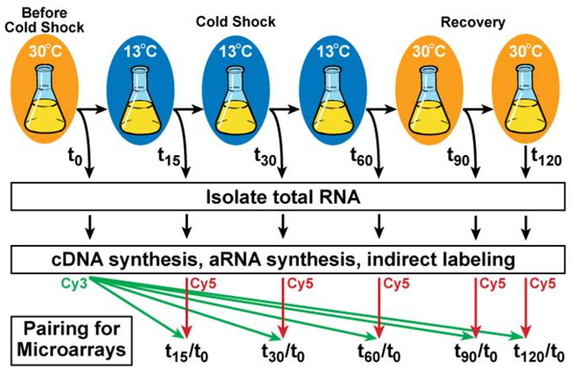

Figure 1. Preparation of microarrays from total RNA isolated from the each strain at six time points over the course of the cold shock experiment.

Dye intensity of individual spots on DNA microarrays are used to analyze gene expression

- The samples were labeled using Cy3 for the initial time point and Cy5 for the rest of the time points.
- These samples were then hybridized to the microarray slide.
- Each spot contains DNA from one gene in yeast.
- Images were analyzed with Genepix Pro software to get the ratio of fluorescent dye intensity for each spot.
- Ratios were converted to log base 2 to make the data more symmetric.
- Loess normalization and MAD scaling was used in R Statistical Software 2.7.2. to correct for dye intensity biases and to ensure each chip had the same weight in the analysis.

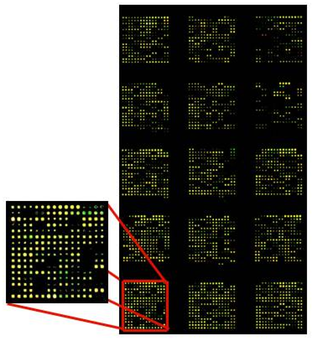

Figure 2. Labeled RNA from two time points hybridized onto a microarray.

21-gene network of transcription factors is involved in the cold shock response

- The network was constructed from transcription factors documented in the YEASTRACT database as regulating each other.
- Transcription factors were included in the network if their target genes were enriched in a list of genes that satisfied one of two criteria:
  - Had significant differential expression in the microarray data indicated by a *P* value < 0.05 after an ANOVA test and Benjamini & Hochberg correction.
  - Potentially involved in the cold shock response as suggested by other experimental evidence.

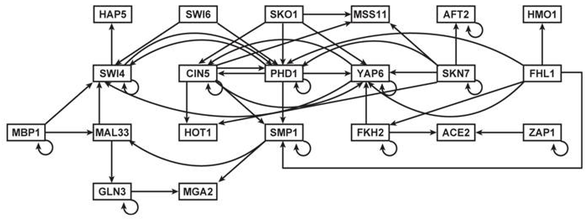

Figure 3. The network consists of 21 nodes representing the gene, mRNA, and transcription factor that it encodes, assuming that the gene is immediately translated after transcription. The nodes are connected by 50 edges which represent activation or repression depending on the sign of the weight of the regulatory effect.

Gene expression is a dynamic balance between production and degradation

- The rate of change of expression for each gene *i* in the network is modeled by a nonlinear differential equation consisting of a production term and a degradation term:

$$\frac{dx_i}{dt} = p(\vec{x}) - \lambda_i x_i$$

- The production of gene *i* depends on the network of regulating genes whose transcription factors activate and/or repress it.
- Degradation of a gene is represented by a linear function $\lambda_i x_i$ where λ_i is the degradation rate constant and x_i is the expression profile of the gene.

Production was initially modeled by a sigmoid function within the nonlinear differential equation

$$p(\vec{x}) = \frac{P_i}{1 + \exp\left(\sum_j -w_{ij}x_j + b_i\right)}$$

- The production of a target gene *i* is influenced by a rate constant (P_i) and the weighted (w_{ij}) concentrations of all transcription factors *j* that regulate it.
- The sign of *w* determines whether or not a target gene is activated or repressed.
  - +*w* indicates up-regulation or activation.
  - -*w* indicates down-regulation or repression.
- The position of the expression threshold of gene *i* is determined by the constant b_i.

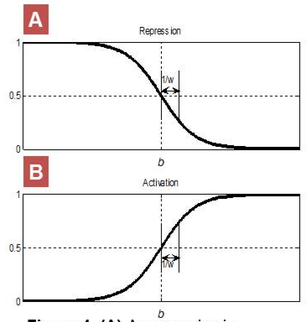

Figure 4. (A) A repression is represented by a standard sigmoid curve where the weight *w* is negative. **(B)** A standard sigmoid curve with a positive weight *w* represents an activation.

The sigmoid model does not accurately describe transcriptional regulation

- The sigmoid model inaccurately suggests that there is a high level of initial transcription of a target gene before the binding of a repressor.
- In the case of multiple regulators, the sigmoid model does not accurately represent either an "AND" or an "OR" transcriptional gate type.
  - An "AND" gate indicates that all transcription factors are required to regulate the target gene.
  - An "OR" gate indicates that either one of the transcription factors is enough to regulate the transcription of a target gene.

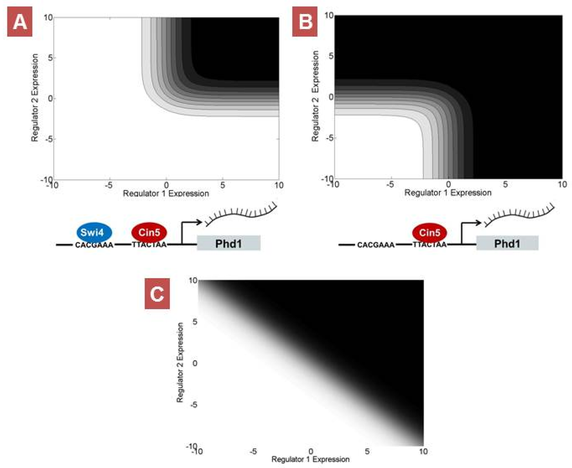

Figure 5. Representations of the regulation of a target gene controlled by two regulators in the case of **(A)** AND and **(B)** OR transcriptional gates in comparison to a **(C)** sigmoid model.

Michaelis-Menten kinetics more accurately describe transcriptional regulation

- The rate of gene expression is still determined by production and degradation.
- The production term now explicitly takes into account the "OR" nature of activation and repression:

$$p(\vec{x}) = P_i \cdot \left(\sum_j \left(\left[\frac{|w_{ij}x_j|}{\sum_k w_{ik}x_k} \right] \cdot \left[\frac{w_{ij}x_j}{1 + w_{ij}x_j} \right] I(w_{ij} > 0) \right) \right)$$

  - The first bracketed factor represents the relative weight of gene *j*.
  - The second bracketed factor represents the Michaelis-Menten reaction rate.
  - The third term models the effect of repression.
- The degradation rate of the target gene is the linear function $\lambda_i x_i$ as previously described.
- This production model treats repression and "OR" gate control more accurately.

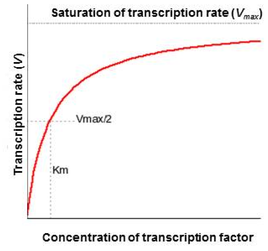

Figure 6. The transcription rate is a function of the transcription factor concentration. Km is the concentration of the transcription factor at which the transcription rate is at 50% of its maximum (V_{max}) and $V_{max}/2$ is the transcription rate at this half-saturation point.

Model parameters estimated based on microarray data

- The data and model were loaded into the MATLAB computational environment.
- MATLAB's function ODE45 solves the model differential equation.
- MATLAB's function FMINCON compares the model to the microarray data to estimate 50 *w*'s, 21 *b*'s (only for the sigmoid model), and 21 production rates.
- Degradation rates were taken from Belle *et al.* (2006).
- Fits were performed by comparing the model to strains individually and by comparing the model to all strains simultaneously.
- Deletion strains are modeled by removing the strain from the system.

PHD1 expression profile is different between the wild type and Δ*cin5* strains

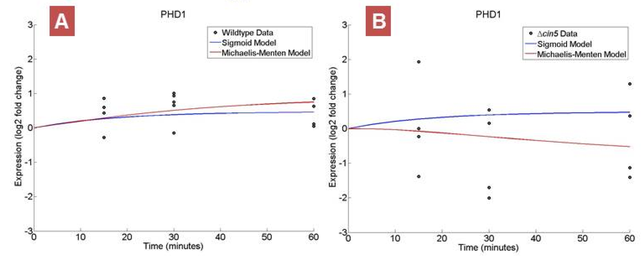

Figure 7. The **(A)** wild type and **(B)** Δ*cin5* models for *PHD1*. The models fitted by both the sigmoid function (blue line) and Michaelis-Menten function (red line) are shown.

Deletion may not be enough: weights need to be adjusted to provide a better fit

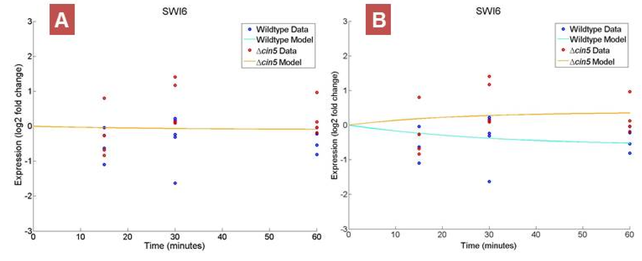

Figure 8. The wild type and Δ*cin5* models for *SWI6* based on Michaelis-Menten kinetics. The models were fitted using **(A)** data for all strains and for **(B)** individual strains in the simulation.

Conclusion

- Our new model, using "OR" gates and Michaelis-Menten kinetics, addresses repression and multiple regulators more accurately than the sigmoidal model.
- Differences are evident in the gene expression graphs between the model produced by a nonlinear differential equation defined by a sigmoid function and by one defined by Michaelis-Menten kinetics.
- Changes in expression were observed between the wild type and deletion strains when looking at graphs of both the sigmoidal model and Michaelis-Menten model for each gene.

Acknowledgements

- We would like to thank Dr. Fitzpatrick for all of his assistance with R and MATLAB code, understanding the deterministic model, and the statistical tests.
- We would like to thank Dr. Dahlquist for providing us with the data to normalize, the current regulatory network, and other background information regarding the experiment.
- We would like to thank lab members (too numerous to list here) who have contributed data.
- This work was made possible by the NSF-RUI grant 0921038 and the McLaughlin Chair of Biology.