

# GRNsight Expression Database *and* Expansion of Species for Gene Page Projects

Alexia Filler and Kevin Patterson

5/1/2020

---

# Some Helpful Terms!

---

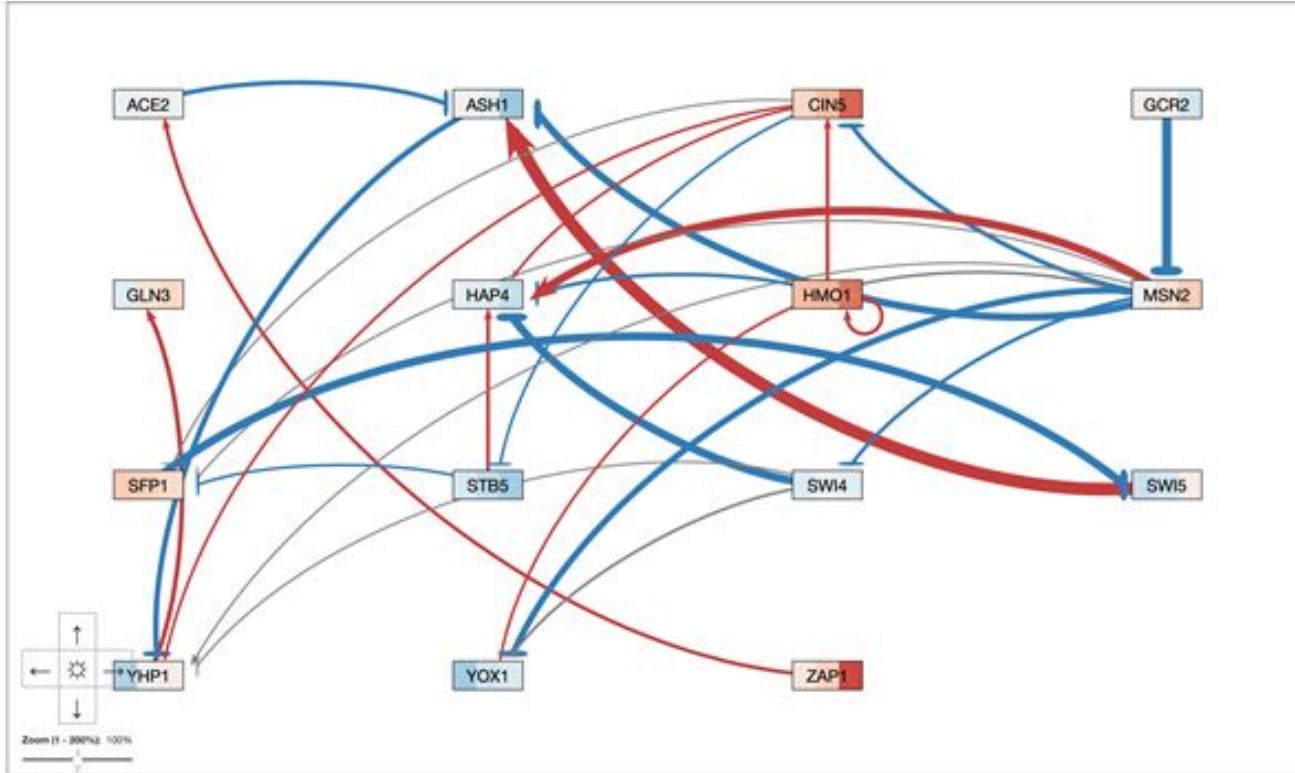
- A **gene regulatory network (GRN)** shows how different genes are connected, how they affect each other
- **Network data** is the baseline data required to visualize a GRN
  - Shows which genes are connected
- **Expression data** shows how much a gene is expressed or repressed over time

# GRNsight Overview

---

- GRNsight is an open source web application that allows users to better visualize the connections in gene regulatory networks
- Uses spreadsheets of data to show these connections

# GRNsight Overview (Cont.)



Node coloring is based on the change in expression over time for each gene

Arrow coloring is based on the regulatory relationship, activation or repression.

# Expression Database Project

- To visualize a colored network graph, users upload an Excel Workbook with multiple worksheets
- Network data shows the network connections
- Expression data shows the level of expression of genes, which we display with node coloring

## Network Data

	A	B	C	D	E	F	G	H
1	cols regulators	ACE2	ASH1	CIN5	GCR2	GLN3	HAP4	HMO1
2	ACE2	0	0	0	0	0	0	0
3	ASH1	1	0	0	0	0	0	0
4	CIN5	0	0	0	0	0	0	1
5	GCR2	0	0	0	0	0	0	0
6	GLN3	0	0	0	0	0	0	0
7	HAP4	0	0	1	0	0	0	1
8	HMO1	0	0	0	0	0	0	1

## Expression Data

	A	B	C	D	E	F
1	id	15	15	30	60	60
2	ACE2	0.6139	-1.0689	0.3377	0.817	0.5566
3	ASH1	0.97	0.3043	-0.7236	-1.3477	-1.0468
4	ZAP1	0.6594	0.6135	-0.394	2.9606	3.5569

# Expression Database Project

---

- The goal is to create a backend database populated with public expression data
- Users can use this data to color their GRNs
  - Allows node coloring even if user does not have their own expression data
- Produces uniform example for how expression data should be formatted

# Goals

---

- Create and utilize AWS RDS and EC2 instances
- Clean data and populate RDS with existing expression data
- Create UI that makes it easy for users to interact with expression database
- Write APIs that connect the front end functionality to the backend database

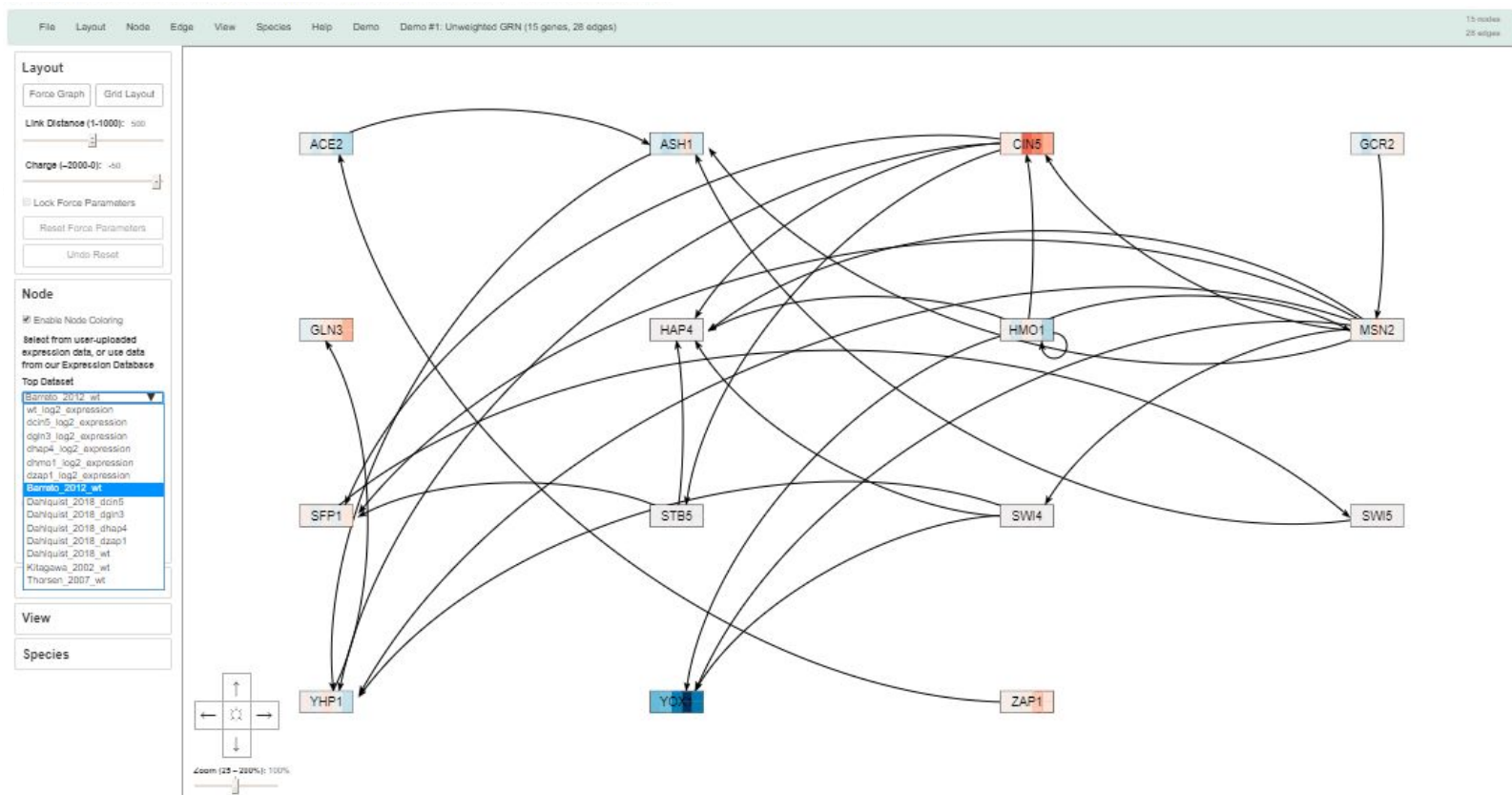
# Project Justification

---

- Worked on GRNsight project for a year, specifically with expression data
- Learning many new skills and concepts
  - Networks
  - AWS
  - REST API writing
- Working on a project that existed for a long time (similar to how it might be to work in industry)



# User Interface



# Expression Database Project Demo!

# Challenges

---

- General learning curve of working with a new set of tools (AWS, REST API writing)
- Challenging to work on a project that has existed for a long time

# Future Directions

---

- Add more data to the database, including network data
- Give users a meaningful way to access the other data already stored (production rates, degradation rates, publication data, etc.)

# Gene Page

- Right-clicking on a node opens a web page which displays information retrieved from various databases
- Currently hardcoded for yeast (*Saccharomyces cerevisiae*) only.

GLN3

*Saccharomyces cerevisiae* (strain ATCC 204508 / S288c)

General Information	Protein Information	Regulation	Gene Ontology	Sources
<u>General Information</u>				
SGD ID:	S000000842 <sup>[1]</sup>			
NCBI Gene ID:	856763 <sup>[4]</sup>			
Ensembl ID:	Not found <sup>[3]</sup>			
Uniprot ID:	GLN3_YEAST <sup>[2]</sup>			
JASPAR ID:	MA0307.1 <sup>[5]</sup>			
Description:	Transcriptional activator of genes regulated by nitrogen catabolite repression; localization and activity regulated by quality of nitrogen source and Ure2p <sup>[1]</sup>			
Species:	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) <sup>[2]</sup>			
Locus Tag:	YER040W <sup>[4]</sup>			
JASPAR Family:	GATA-type zinc fingers <sup>[5]</sup>			
JASPAR Class:	Other C4 zinc finger-type factors <sup>[5]</sup>			
Chromosome Sequence:	NC_001137.3 <sup>[4]</sup>			
<u>Protein Information</u>				
<u>Regulation</u>				
<u>Gene Ontology</u>				
<u>Sources</u>				

# Gene Page Species Expansion

---

- Add functionality for the gene page by expanding the species that GRNsight supports (additional 5)
- Allow users to upload a file with species data
- Update which species they are working with through dropdown menus in the application
- Dropdown menus display the currently selected species

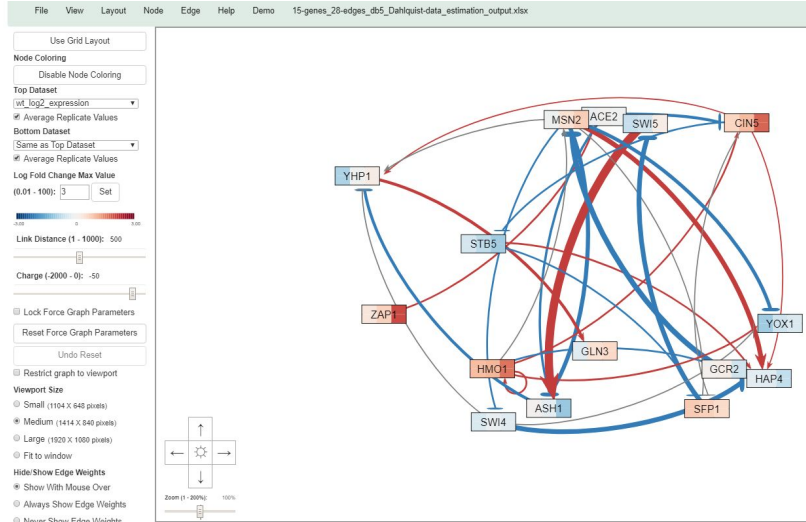
# Project Justification

---

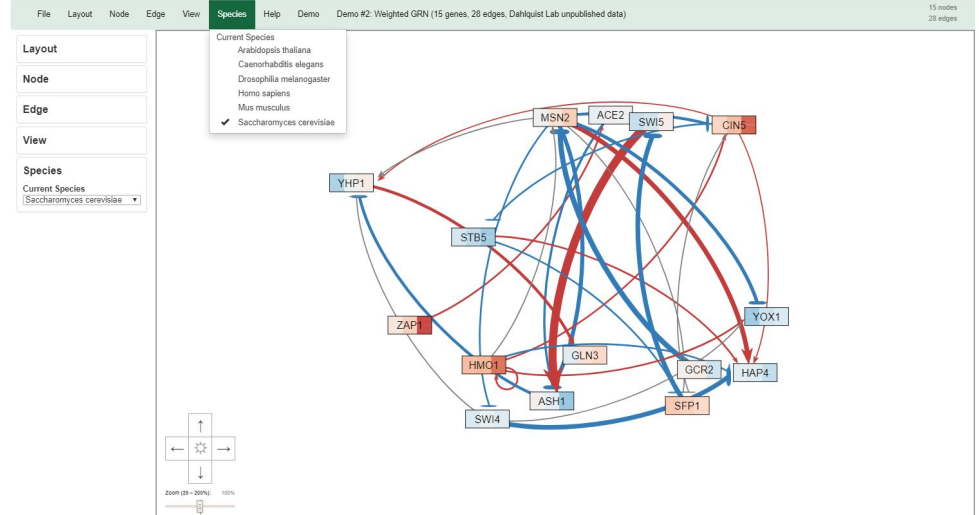
- Work in a environment similar to the real world as a member of a larger project team with weekly meetings and goals
- Full stack development working on both front and backend components
- Research project which Dr. Dahlquist uses each year in her biological databases class, meaning my work has an immediate impact

# Species Dropdown Menus

Live



Beta



Users can select from among six species in a panel or menu



# User Interface - Gene Pages

## GRNsight

Web app and service for visualizing models of gene regulatory networks.

A Joint Project of the  
LMU Bioinformatics and  
Biomathematics Groups

LMU|LA  
Loyola Marymount  
University

ZAP1

*Saccharomyces cerevisiae* (strain ATCC 204508 / S288c)

General Information

Protein Information

Regulation

Gene Ontology

Sources

### General Information

SGD ID:	S000003592 <sup>[1]</sup>
NCBI Gene ID:	853390 <sup>[4]</sup>
Ensembl ID:	YJL056C <sup>[3]</sup>
Uniprot ID:	ZAP1_YEAST <sup>[2]</sup>
JASPAR ID:	MA0440.1 <sup>[5]</sup>
Description:	Zinc-regulated transcription factor; binds to zinc-responsive promoters to induce transcription of certain genes in presence of zinc, represses other genes in low zinc; regulates its own transcription; contains seven zinc-finger domains <sup>[1]</sup>
Species:	<i>Saccharomyces cerevisiae</i> (strain ATCC 204508 / S288c) <sup>[2]</sup>
Locus Tag:	YJL056C <sup>[4]</sup>
JASPAR Family:	Factors with multiple dispersed zinc fingers <sup>[5]</sup>
JASPAR Class:	C2H2 zinc finger factors <sup>[5]</sup>
Chromosome Sequence:	NC_001142.9 <sup>[4]</sup>

### Protein Information

### Regulation

### Gene Ontology

### Sources

#### Sources

1. Saccharomyces Genome Database
2. UniProt
3. Ensembl
4. NCBI Gene Database
5. JASPAR Database

# Gene Page Species Expansion Project Demo!

# Challenges

---

- Learning and working on preexisting code that has been written by multiple different students over the years
- Rewriting code to fit a clients expectations, and all the bugs that come with a rewrite
- Multidisciplinary project, so I needed a biology refresh to understand the function of my code
- Working with a lot of different biological databases
- API errors and different databases for different species

# Future Directions

---

- Only calling species specific APIs when the species is selected
- API calls to different interMine databases for each species
- Adding extra species to GRNsight

# Acknowledgements

Thank you to Dr. Dahlquist of LMU Biology and Dr. Dionisio of LMU Computer Science for advising us on these projects.

Visit GRNsight at <http://dondi.github.io/GRNsight>

Visit our repository at <https://github.com/dondi/GRNsight>



**Thank you!**  
**Questions?**