Computational reconstruction of the Transcriptional Regulatory Modules (TRMs) in yeast

Wei-Sheng Wu
Yeast Bioinformatician

Associate Professor
Department of Electrical Engineering (EE)
National Cheng Kung University
Taiwan

Visiting Scholar
Department of Ecology and Evolution (EE)
University of Chicago
USA
Where is Taiwan?

http://www.taroko.gov.tw/English/?mm=2&sm=2&page=1

https://travel.state.gov/content/passports/en/country/taiwan.html
Outlines

• Background of Transcriptional Regulatory Modules (TRMs)

• My algorithm for reconstructing TRMs (published in BMC Genomics)
  - using yeast heat shock response as an example

• My other expertise: databases and web tools development
Outlines

• Background of *Transcriptional Regulatory Modules* (TRMs)

• My algorithm for reconstructing TRMs
  - using yeast heat shock response as an example

• My other expertise: databases and web tools development
What is a transcriptional regulatory module (TRM)?

- **TRM**: a set of genes that is co-regulated by the same set of TFs.

\[ \text{TF1, TF2, TF3} \rightarrow \text{Gene1, Gene2, Gene3, Gene4} \]
Why study TRM? (using yeast heat shock response as an example)

• Single-cell organisms such as yeasts constantly face changing or even harsh environments such as high temperature that threaten their survival.

• By organizing the genome into transcriptional regulatory modules (TRMs), a yeast cell can coordinate the activities of many genes and carry out complex functions in response to high temperature.

• Therefore, identifying TRMs of heat response is instrumental for understanding cellular responses to heat shock.
Current knowledge of yeast heat shock response (from numerous experimental studies for many years)

- Disruption of a large number of cellular assemblies and processes, an increased protein unfolding and aggregation, and membrane structure alterations are paramount in cells exposed to high temperature.
- Heat shock response serves to counteract these deleterious effects.
- Through it cells increase their thermostolerance or ability to withstand heat stress.
Many events occur in yeast cells during heat shock response (from numerous experimental studies for many years)

- Cell cycle transiently arrests during a heat shock stress.
- Heat shock proteins (HSPs) are rapidly synthesized.
- Many HSPs function as protein chaperones, so named because of their ability to bind to partially unfolded proteins to protect them from degradation or aggregation.
Many events occur in yeast cells during heat shock response (from numerous experimental studies for many years)

- Heat shock cells induce a variety of genes related to carbohydrate metabolism, fatty acid metabolism, respiration and others.
Many events occur in yeast cells during heat shock response (from numerous experimental studies for many years)

- Heat shock causes the extremely rapid accumulation of a large cytoplasmic pool of trehalose.

- Trehalose is one of the most effective substances known for preservation of membranous structures and enzyme activities during heating.
Many events occur in yeast cells during heat shock response (from numerous experimental studies for many years)

- Heat shock causes an increased protein unfolding and aggregation in a cell.
- Thus, many genes that are involved in protein (re)folding are induced to bind to partially unfolded proteins to protect them from degradation or aggregation.
TFs and genes known to be involved in heat shock response (from numerous experimental studies for many years)
TFs and genes known to be involved in heat shock response (from numerous experimental studies for many years)
Computational reconstruction of TRMs

• Identifying direct regulations one by one using traditional experimental approaches is very time-consuming and labor-intensive to reconstruct TRMs.

- Binding evidence (DNA footprinting, Giardina et al. 1995)
- Regulation evidence (Northern blotting WT vs. hsf1Δ; Eastmond and Nelson 2006)

• Since various kinds of high-throughput experimental technologies (e.g. DNA microarray/RNA-seq, ChIP-chip/ChIP-seq, ...) are available now, it is possible to reconstruct TRMs using computational approaches.
Outlines

• Background of Transcriptional Regulatory Modules (TRMs)

• My algorithm for reconstructing TRMs
  - using yeast heat shock response as an example

• My other expertise: databases and web tools development
My algorithm for reconstructing TRMs of the yeast heat shock response

- ChIP-chip/ChIP-seq of yeast cells with heat shock
- DNA microarray/RNA-seq of yeast cells with heat shock
- DNA microarray/RNA-seq of yeast cells without heat shock
- TF binding information
- Gene expression information
- TRMs of the yeast heat shock response
My algorithm for reconstructing TRMs of the yeast heat shock response

**Step 1:** For each gene in the yeast genome (~6500 genes), identify its promoter-binding TFs under heat shock.

**Step 2:** For each gene in the yeast genome, extract its regulatory TFs from its promoter-binding TFs under heat shock.

**Step 3:** Identify heat-responsive genes from the yeast genome.

  e.g. Gene1, Gene2, Gene3, Gene4, Gene5, Gene6,...

**Step 4:** Identify heat-responsive TF sets.

  e.g. [TF1,TF2], [TF3,TF4,TF5],...

**Step 5:** Reconstruct heat-responsive TRMs.

  e.g. \{[TF1,TF2] ⇒ [Gene1, Gene2, Gene3]\}
**Step 1:** For each gene in the yeast genome (~6500 genes), identify its promoter-binding TFs under heat shock

- Using **ChIP-chip/ChIP-seq** technology, researchers can know the **genome-wide binding target genes of a specific TF**.

- For example, Harbison et al. (Nature 2004) used the ChIP-chip technology to determine the **genome-wide binding target genes of 200 yeast TFs in rich media conditions** and 7 TFs (Adr1, Gat1, Hsf1, Msn2, Skn7, Xbp1, Yap1) under heat shock.

- So, for each yeast gene, we know its promoter-binding TFs under heat shock from Harbison’s ChIP-chip data.
Step 2: For each gene in the yeast genome, extract its regulatory TFs from its promoter-binding TFs under heat shock.

- Using **DNA microarray/RNA-seq** technology, researchers can have the **mRNA time profile of each gene** in the genome.

- For example, Causton et al. (MBC 2001) used DNA microarray technology to measure the **mRNA time profile of each gene in the yeast genome at 0 (before heat shock; 25°C), 15, 30, 45, 60, 120 min (after heat shock; 37°C).**
Step 2: For each gene in the yeast genome, extract its regulatory TFs from its promoter-binding TFs under heat shock.

\[ y[t + 1] = \left(b_1 \cdot x_1[t] + b_2 \cdot x_2[t] + b_3 \cdot x_3[t] + k\right) - a \cdot y[t] + \epsilon[t] \]

If maximum likelihood (ML) parameter estimation results show that \(|b_1| \& |b_2|\) are >> 0, but not \(|b_3|\)

\{TF1, TF2\} are regulatory TFs of gene y.
After running Step 2, we will know

- **The regulatory TFs of each gene** in the yeast genome under heat shock

  ![Diagram of regulatory TFs and target genes](image)

<table>
<thead>
<tr>
<th>TF1</th>
<th>TF2</th>
<th>...</th>
<th>TFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene1</td>
<td>v</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>Gene2</td>
<td>v</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GeneK</td>
<td>v</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **The regulatory target genes of each TF** under heat shock

  ![Diagram of regulatory target genes](image)

<table>
<thead>
<tr>
<th>TF1</th>
<th>TF2</th>
<th>...</th>
<th>TFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene1</td>
<td>v</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>Gene2</td>
<td>v</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GeneK</td>
<td>v</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Step 3: Identify heat-responsive genes from the yeast genome

**DNA microarray/RNA-seq** data of yeast cells without heat shock

**DNA microarray/RNA-seq** data of yeast cells with heat shock

Differential expressed gene identification tools

Heat-responsive genes
Step 4: Identify heat-responsive TF sets

- A TF set (e.g. [TF1,TF2]) is said to be heat-responsive only if a significant portion of the regulatory target genes of the TF set is heat-responsive.
- The number of TFs in a TF set could be one, two or more.

The regulatory targets of a TF set [TF1,TF2] identified in Step 3

<table>
<thead>
<tr>
<th></th>
<th>TF1</th>
<th>TF2</th>
<th>...</th>
<th>TFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene1</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene2</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GeneK</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The hypergeometric distribution is used to test the statistical significance.
Step 5: Reconstruct heat-responsive TRMs

For each heat-responsive TF set (e.g. [TF1, TF2]), we collect all their regulatory targets that are heat-inducible \& highly co-expressed to form a TRM.

The regulatory targets of a TF set [TF1, TF2]

Heat-responsive genes identified in Step 3

[Gene1, Gene2, Gene3, Gene4, Gene5, Gene6]

[Gene1, Gene2, Gene3, Gene4]

TRM:

\{[TF1, TF2] \rightarrow [Gene1, Gene2, Gene3, Gene4] \}
My algorithm for reconstructing TRMs of the yeast heat shock response

Harbison 2004 ChIP-chips of yeast cells with heat shock

Causton 2001 DNA microarrays of yeast cells with heat shock

Caustion 2001 DNA microarrays of yeast cells without heat shock

TF binding information

Gene expression information

algorithm

TRMs of the yeast heat shock response
• **Msn2** colored blue: known to be involved in heat shock response

• The periphery of a rectangle colored **purple**: the module has at least one enriched MIPS functional category

• An oval colored **green**: the TF's function is consistent with at least one of the module's enriched MIPS functional categories

• Physical interaction

• Transcriptional regulation

<table>
<thead>
<tr>
<th>(3,8) Module 9 regs:</th>
<th>Hsf1</th>
<th>Msn2</th>
<th>Msn4</th>
</tr>
</thead>
<tbody>
<tr>
<td>YBR117C</td>
<td>YBR117C TKL2</td>
<td>Transketolase, similar to TRM</td>
<td></td>
</tr>
<tr>
<td>YDR343C</td>
<td>YDR343C HXT6</td>
<td>High-affinity glucose transport</td>
<td></td>
</tr>
<tr>
<td>YDR533C</td>
<td>YDR533C HSP31</td>
<td>Possible chaperone activity</td>
<td></td>
</tr>
<tr>
<td>YEL039C</td>
<td>YEL039C CYC7</td>
<td>Cytochrome c isoform</td>
<td></td>
</tr>
<tr>
<td>YER103W</td>
<td>YER103W SSA4</td>
<td>Heat shock protein thioredoxin</td>
<td></td>
</tr>
<tr>
<td>YER150W</td>
<td>YER150W SPI1</td>
<td>GPI-anchored cell wall</td>
<td></td>
</tr>
<tr>
<td>YKR076W</td>
<td>YKR076W ECM4</td>
<td>Omega class glutathione</td>
<td></td>
</tr>
<tr>
<td>YLL026W</td>
<td>YLL026W HSP104</td>
<td>Heat shock protein thioredoxin</td>
<td></td>
</tr>
<tr>
<td><strong>P-VALUE</strong></td>
<td><strong>ENRICHED FUNCTIONAL CATEGORY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.58E-03</td>
<td>32.01 stress response</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.35E-03</td>
<td>32.01.07 unfolded protein response (e.g. ER stress)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.44E-03</td>
<td>32 CELL RESCUE, DEFENSE AND VIRUS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.01E-03</td>
<td>14.01 protein folding and stabilization</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TRMs of yeast heat shock response

• We identify 29 GRMs, containing 182 heat-responsive genes regulated by 12 heat-responsive TFs.

• 108/182 genes and 7/12 TFs are known to be involved in heat shock response.
Identification of 12 heat-responsive TFs

- 7 TFs (Msn2, Msn4, Hsf1, Yap1, Hac1, Rlm1 and Cad1) known to be involved in heat shock response.
- Five novel heat-responsive TFs (Rox1, Cst6, Ume6, Ste12 and Dig1) have also been identified.
- These five novel heat-responsive TFs and the other seven known heat-responsive TFs form a highly connected network of interactions, suggest

1. these five novel heat-responsive TFs may play a role in heat shock response.
2. different combinations of a fairly small number of heat-responsive TFs may be sufficient to regulate a large number of genes involved in heat shock response.
Identification of known heat-responsive genes

- Genes code for heat shock proteins (HSPs)
  - [HSP31, HSP104, SSA1, SSA4, SSE2]

- Genes involved in the protein folding or refolding
  - [HSP12, HSP78]
  - [HSP10, HSP60]

- Genes code for trehalose synthase subunits
  - [EUG1, LHS1, SCJ1, ERO1]
  - [TSL1]
  - [TSP1, NTH1]
Identification of known heat-responsive genes

- Genes known to be involved in protein degradation
  - Hac1
  - [DER1, PBI2]
  - Msn2
  - [LAP4, ASI1]
  - Rlm1
  - [ATG8, YSP3]
  - Hsf1
  - [APG12, JEM1, UBC8, UBI4]

- Genes known to be involved in the cell wall biogenesis and maintenance
  - Rlm1
  - [CWP1, CHS3, FLC2, GFA1, HKR1, KTR2, SLT2]

- Genes involved in glucose metabolism
  - Msn2
  - [GLK1, GPH1, HXK1]

- Genes involved in fatty acid metabolism
  - Msn4
  - [COX20, CYC7]

- Genes involved in respiration
  - Hsf1
  - [FAA1, FAA4]
Annotating 68 uncharacterized genes

- Among the 182 identified heat-responsive genes, 68 genes have unknown function according to the Saccharomyces Genome Database.
- We suggest that these genes are involved in heat shock response.

- Our predictions are supported by the fact that
  1. all these 68 genes are induced by more than three folds at least at two time points of their expression profiles under heat shock
  2. all these 68 genes are regulated by known heat-responsive TFs: 14 of these genes are regulated by Hac1, 25 by Hsf1, 24 by Msn2, 20 by Msn4, 17 by Yap1, and so on

- However, further experimental validations are needed to confidently annotate these uncharacterized genes as heat-responsive genes.
Refining the clusters of the genes involved in the protein (re)folding

- Heat shock causes an increased protein unfolding and aggregation in a cell.
- Thus, many genes that are involved in protein (re)folding are induced to bind to partially unfolded proteins to protect them from degradation or aggregation.
- Among the 182 identified heat-responsive genes, 20 genes are known to be involved in protein (re)folding.
- Although these genes are functionally similar, they may be under different transcriptional controls. Indeed, my algorithm assigns these 20 genes into five modules.
- My algorithm can refine the cluster of genes involved in protein (re)folding and degradation and can provide a better understanding of how the cell regulates the complex expression program of these genes.
Summary

• I developed an algorithm which combine ChIP-chip/ChIP-seq and DNA microarray/RNA-seq data to reconstruct TRMs of the yeast heat shock response.

• My algorithm identified 29 GRMs, which in total contain 182 heat-responsive genes regulated by 12 heat-responsive TFs.

• The literature indicates that 108 of the 182 genes and 7 of the 12 TFs are known to be involved in heat shock response.

• My algorithm suggested that 68 uncharacterized genes may be involved in heat shock response and it also identified their plausible heat-responsive regulators.

• My algorithm refined the cluster of genes that are involved in the protein (re)folding and provided a better understanding of how the complex expression program of heat shock is regulated.
Outlines

• Background of Transcriptional Regulatory Modules (TRMs)

• My algorithm for reconstructing TRMs
  - using yeast heat shock response as an example

• My other expertise: databases and web tools development
## Yeast databases

<table>
<thead>
<tr>
<th><strong>YPA (Yeast Promoter Atlas)</strong></th>
<th><strong>NAR</strong> featured article</th>
<th><strong>NAR 2011</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>YTRP (Yeast Transcriptional Regulatory Pathway)</strong></td>
<td></td>
<td><strong>Database 2014</strong></td>
</tr>
<tr>
<td><strong>YNA (Yeast Nucleosome Atlas)</strong></td>
<td></td>
<td><strong>BMC Genomics 2014</strong></td>
</tr>
<tr>
<td><strong>CoopTFD (Cooperative TFs Database)</strong></td>
<td></td>
<td><strong>Database 2016</strong></td>
</tr>
<tr>
<td><strong>YCRD (Yeast Combinatorial Regulation Database)</strong></td>
<td></td>
<td><strong>Plos One 2016</strong></td>
</tr>
<tr>
<td><strong>YLBP (Yeast Lipid-Binding Proteins)</strong></td>
<td></td>
<td><strong>Under construction</strong></td>
</tr>
<tr>
<td><strong>YARG (Yeast Arsenic-Related Genes)</strong></td>
<td></td>
<td><strong>Under construction</strong></td>
</tr>
<tr>
<td><strong>YPRG (Yeast Prion-Related Genes)</strong></td>
<td></td>
<td><strong>Planning to do</strong></td>
</tr>
</tbody>
</table>
## Yeast web tools

<table>
<thead>
<tr>
<th>Tool Name</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>YGA (Yeast Genes Analyzer)</td>
<td>Gene 2012</td>
</tr>
<tr>
<td>YNA (Yeast Nucleosome Atlas)</td>
<td>BMC Genomics 2014</td>
</tr>
<tr>
<td>YAGM (Yeast Associated Gene Miner)</td>
<td>BMC Systems Biology 2015</td>
</tr>
<tr>
<td>PCTFPeval (Predicted Cooperative TF Pairs evaluator)</td>
<td>BMC Bioinformatics 2015</td>
</tr>
<tr>
<td>YGSE (Yeast Gene Set Enrichment)</td>
<td>Under construction</td>
</tr>
</tbody>
</table>
### Other databases and tools

- **Drosophila database**
  - cis-MEP (cis-regulatory Module Epigenetic Profile Database for Drosophila Melanogaster)
    - BMC Systems Biology 2014

- **Human database**
  - CSmirTar (Condition-Specific miRNA Targets)
    - Ready for submission
  - IRES (Internal Ribosome Entry Zone)
    - Ready for submission
  - RPDV (Ribosome Profiling Data Viewer)
    - Under construction
  - p53BLV (p53 Binding Location Viewer)
    - Under construction

- **Phosphoproteomic tool**
  - iPhos (a toolkit to streamline the alkaline phosphatase-assisted comprehensive LC-MS phosphoproteome investigation)
    - BMC Bioinformatics 2014

- **Microarray missing value imputation**
  - MissVIA (Missing Value Imputation Atlas)
    - BMC Systems Biology 2013
  - MVIAeval (Missing Value Imputation Algorithm evaluator)
    - Submitted
A tool for yeast gene sets enrichment & comparison

• Nowadays, yeast biologists can have a set of genes (e.g. differentially expressed genes under heat shock) or two sets of genes (induced genes under heat shock vs. repressed genes under heat shock) easily by using **RNA-seq technology**.

• How to make biological sense of these genes is challenging.

• My lab is developing a web tool for yeast biologists.
Web interface of a single list of genes

<table>
<thead>
<tr>
<th>Input</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>YLR178C</td>
<td>Essential</td>
</tr>
<tr>
<td>YJR096W</td>
<td>Essential</td>
</tr>
<tr>
<td>YGR248W</td>
<td>Essential</td>
</tr>
<tr>
<td>YPL230W</td>
<td>Essential</td>
</tr>
<tr>
<td>YLR149C</td>
<td>Essential</td>
</tr>
<tr>
<td>YGR088W</td>
<td>Essential</td>
</tr>
<tr>
<td>YDL204W</td>
<td>Essential</td>
</tr>
<tr>
<td>YBR053C</td>
<td>Essential</td>
</tr>
<tr>
<td>YBR056W</td>
<td>Essential</td>
</tr>
<tr>
<td>YBR126C</td>
<td>Essential</td>
</tr>
<tr>
<td>YCL042W</td>
<td>Essential</td>
</tr>
<tr>
<td>YCL040W</td>
<td>Essential</td>
</tr>
<tr>
<td>YGL037C</td>
<td>Essential</td>
</tr>
<tr>
<td>YML100W</td>
<td>Essential</td>
</tr>
<tr>
<td>YBR169C</td>
<td>Essential</td>
</tr>
<tr>
<td>YLR258W</td>
<td>Essential</td>
</tr>
<tr>
<td>YMR105C</td>
<td>Essential</td>
</tr>
<tr>
<td>YFR053C</td>
<td>Essential</td>
</tr>
<tr>
<td>YMR250W</td>
<td>Essential</td>
</tr>
<tr>
<td>YHL021C</td>
<td>Essential</td>
</tr>
<tr>
<td>YDR171W</td>
<td>Essential</td>
</tr>
<tr>
<td>YBL064C</td>
<td>Essential</td>
</tr>
<tr>
<td>YKR076C</td>
<td>Essential</td>
</tr>
<tr>
<td>YMR090W</td>
<td>Essential</td>
</tr>
<tr>
<td>YIR038C</td>
<td>Essential</td>
</tr>
<tr>
<td>YIR039C</td>
<td>Essential</td>
</tr>
<tr>
<td>YER079W</td>
<td>Essential</td>
</tr>
</tbody>
</table>

Compare to yeast genome

- Plasticity
- 5'UTR length
- 3'UTR length
- Transcription level
- Expression level
- Half life
- Transcriptional freq
- Translational efficiency
- Ing1
- Ger1
- Mcm1
- Sub1
- mTIF
- MW
- PI
- Protein Length
- TFs defined by B
- TFs defined by K
- TFs defined by B and K
- TFs defined by B or K
- Physical PPI
- Genetic PPI
- P and G PPI
- P or G PPI

Enriched TF P-value cutoff 10^-2

- TF_B
- TF_K
- TF_BK

Domains P-value cutoff 10^-2

- Literature P-value cutoff 10^-2

- GO P-value cutoff 10^-2

- phenotype P-value cutoff 10^-2

- pathway P-value cutoff 10^-2

- Interaction P-value cutoff 10^-2

- Physical
- Genetic
**Essential genes**

<table>
<thead>
<tr>
<th>Essential genes / # of genes in the yeast genome</th>
<th>Intersects / # of input genes</th>
<th>Hypergeometric P-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>1117/6572 (17.00%)</td>
<td>5/281 (1.78%)</td>
<td>1.000e+00</td>
</tr>
</tbody>
</table>

**Enriched TF P-value cut off** 10^-2
- TF_B
- TF_K
- TF_BK
- Domains P-value cut off 10^-2
- Literature P-value cut off 10^-2
- GO P-value cut off 10^-2
- phenotype P-value cut off 10^-2
- pathway P-value cut off 10^-2
- Interaction P-value cut off 10^-2

- Physical
- Genetic

- Ger1
- Mcm1
- Sub1
- mTIF
- MW
- PI
- Protein Length
- TFs defined by B
- TFs defined by K
- TFs defined by B and K
- TFs defined by B or K
- Physical PPI
- genetic PPI
- P and G PPI
- P or G PPI
<table>
<thead>
<tr>
<th>TF_BK</th>
<th>Name</th>
<th>TF_BKs / # of genes in the yeast genome</th>
<th>Intersects / # of input genes</th>
<th>Hypergeometric P-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSN2</td>
<td>692.0/6572.0 (10.53%)</td>
<td>102.0/281.0 (36.30%)</td>
<td>4.061e-30</td>
<td></td>
</tr>
<tr>
<td>PDR1</td>
<td>109.0/6572.0 (1.66%)</td>
<td>28.0/281.0 (9.96%)</td>
<td>7.181e-13</td>
<td></td>
</tr>
<tr>
<td>SOK2</td>
<td>374.0/6572.0 (5.69%)</td>
<td>52.0/281.0 (18.51%)</td>
<td>1.455e-12</td>
<td></td>
</tr>
<tr>
<td>YAP1</td>
<td>342.0/6572.0 (5.20%)</td>
<td>47.0/281.0 (16.73%)</td>
<td>5.389e-11</td>
<td></td>
</tr>
<tr>
<td>MSN4</td>
<td>120.0/6572.0 (1.83%)</td>
<td>26.0/281.0 (9.25%)</td>
<td>5.299e-10</td>
<td></td>
</tr>
<tr>
<td>HSF1</td>
<td>116.0/6572.0 (1.77%)</td>
<td>23.0/281.0 (8.19%)</td>
<td>6.638e-08</td>
<td></td>
</tr>
<tr>
<td>CIN5</td>
<td>200.0/6572.0 (3.04%)</td>
<td>27.0/281.0 (9.61%)</td>
<td>1.269e-05</td>
<td></td>
</tr>
</tbody>
</table>
The table shows the analysis of iESR genes in the Yeast Genome.

### Plasticity

**Input1**

<table>
<thead>
<tr>
<th># of input genes</th>
<th>Median</th>
<th>Mean</th>
<th>Yeast Genome</th>
<th># of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>279/281 (99.288 %)</td>
<td>0.72</td>
<td>0.938</td>
<td>6105/6572 (92.894 %)</td>
<td>0.252</td>
</tr>
</tbody>
</table>

**K-test P-value**

<table>
<thead>
<tr>
<th>A ≠ B</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>A &gt; B</td>
<td>7.012837e-77</td>
</tr>
<tr>
<td>A &lt; B</td>
<td>1</td>
</tr>
</tbody>
</table>

**U-test P-value**

<table>
<thead>
<tr>
<th>A ≠ B</th>
<th>2.528223e-89</th>
</tr>
</thead>
<tbody>
<tr>
<td>A &gt; B</td>
<td>1.264112e-89</td>
</tr>
<tr>
<td>A &lt; B</td>
<td>1</td>
</tr>
</tbody>
</table>

**T-test P-value**

<table>
<thead>
<tr>
<th>A ≠ B</th>
<th>6.159119e-19</th>
</tr>
</thead>
<tbody>
<tr>
<td>A &gt; B</td>
<td>3.07956e-19</td>
</tr>
<tr>
<td>A &lt; B</td>
<td>1</td>
</tr>
</tbody>
</table>
Web interface of comparing two lists of genes

<table>
<thead>
<tr>
<th>Input</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>YDR255C, YIL097W, YHR171W, YKL124W, YMR041C, YDR254W, YGL156W, YFL124W, YER1, YDR0, YDL1, YPR1, YPR1, YPL1, YKL172W, YNL308C, YKL082C, YGR272C, YGL292W, YPL266W, YQL072C</td>
<td>Essential, iESR, rESR, Singleton, Duplicate, KEGG ribosome</td>
</tr>
</tbody>
</table>

- rESR genes
- iESR genes
Analyses results of rESR vs. iESR

• See webpage
Be my collaborators

• I am familiar with many kinds of yeast genome-wide data (sequence, gene expression, ChIP-chip, TF knockout, nucleosome occupancy, histone modification, PPI, genetic interaction, ribosome profiling, protein phosphorylation, mutant phenotype, GO, pathway, ...)

• My lab has eight master students who are good at coding.

• We are happy to develop databases or tools for you.

• Yeast biologists, tell me your needs!
Any Questions or Comments?