**Modeling of Six Related, Database-Derived GRNs (Slides 2-71)**

Six database-derived small gene regulatory networks (GRNs) with between 14-17 genes and 25-35 edges were modeled using GRNmap. Two separate networks (db2 and db3) were derived from the ∆cin5 strain data, as the removal of the transcription factor MCM1 from the larger 17-gene network (db3) further disconnected ACE2 and ZAP1 from the network. The naming scheme for the six networks is as follows:

* **db1: network derived from wild-type data**
  + 16 genes, 26 edges
* **db2: network derived from ∆cin5 data**
  + 14 genes, 25 edges
* **db3: network derived from ∆cin5 data**
  + 17 genes, 32 edges
* **db4: network derived from ∆gln3 data**
  + 14 genes, 35 edges
* **db5: network derived from ∆hap4 data\***
  + 15 genes, 28 edges
* **db6: network derived from ∆zap1 data**
  + 16 genes, 27 edges

Note: db5, marked with an asterisk (\*) above, was used as model network for comparison to randomly generated, small GRNs.

**I. Comparative Analysis of the Six Database-Derived GRNs**

**Table 1.** GRNmap network parameter outputs for db1-db6.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **db1** | **db2** | **db3** | **db4** | **db5** | **db6** |
| Penalty Term | 2.5923 | 1.8830 | 1.7570 | 1.7895 | 2.3443 | 1.8276 |
| LSE | 0.8194 | 0.6634 | 0.7864 | 0.6994 | 0.6919 | 0.8602 |
| minLSE | 0.5768 | 0.4885 | 0.5449 | 0.5379 | 0.4851 | 0.6156 |
| LSE:minLSE Ratio | 1.4206 | 1.3580 | 1.4100 | 1.3000 | 1.4263 | 1.3973 |
| Iteration Count | 109,718 | 53,862 | 118,921 | 78,124 | 62,139 | 76,769 |

* LSE:minLSE ratios for db1-db6 ranged from 1.3000 (db4) to 1.4263 (db5), indicating that GRNmap consistently modeled the dynamics of small GRNs well.
* The database-derived network with the lowest LSE:minLSE ratio was db4 (1.3000), indicating that this network best modeled the regulatory dynamics observed in the cold shock experiment from which the network was derived.
  + db4 contains the lowest number of genes (14) and the highest number of edges (35).
  + db4 was the only database-derived network to include the transcription factors CYC8 and TEC1 (Table 2), both of which demonstrated interesting connectivity to the network. CYC8 exhibited an in-degree of 4 and an out-degree of 0, while TEC1 showed an in-degree of 4 and out-degree of 5. The high betweenness centrality of TEC1 (24.3333) trailed only behind that of MSN2 in this network (26.8333), both of with exceeded the betweenness centrality measures of the nodes in the five other database-derived networks.
    - Removal of nodes with high betweenness centrality most disrupts the communications within a GRN, as these nodes lie along the largest number of shortest paths between other nodes in the network. MSN2 is present in all six database-derived networks and consistently exhibits a high betweenness centrality. However, TEC1 is only included in db4. If TEC1 is part of the “true” GRN controlling the response to cold shock in yeast, we would expect its exclusion from a network to substantially hinder the GRN’s ability to model cold shock experiment microarray data.
      * However, of TEC1 inclusion in this network is likely due to the fact that db4 was derived from a ∆gln3 deletion strain of *S. cerevisiae*. Deletion of gln3 is known to abolish pseudohyphal growth; and the transcription factor TEC1—which is known to regulate pseudohyphal growth—was substantially down-regulated during the ∆gln3 deletion strain cold shock experiment.
      * If TEC1 is not related to the response to cold shock, then its exclusion from other networks would not be relevant. Instead, its inclusion in db4 could merely have reduced the resulting LSE:minLSE ratio for db4 by increasing the number of parameters (35 edges) available within a 14-gene GRN (compared to 25 edges in db2, another 14-gene GRN).
* db5 exhibited the highest LSE:minLSE ratio (1.4263). Thus, it poorly modeled the cold shock microarray data from which it was derived compared to the other database-derived networks.
  + db5 contains an intermediate number of genes (15) and edges (28).

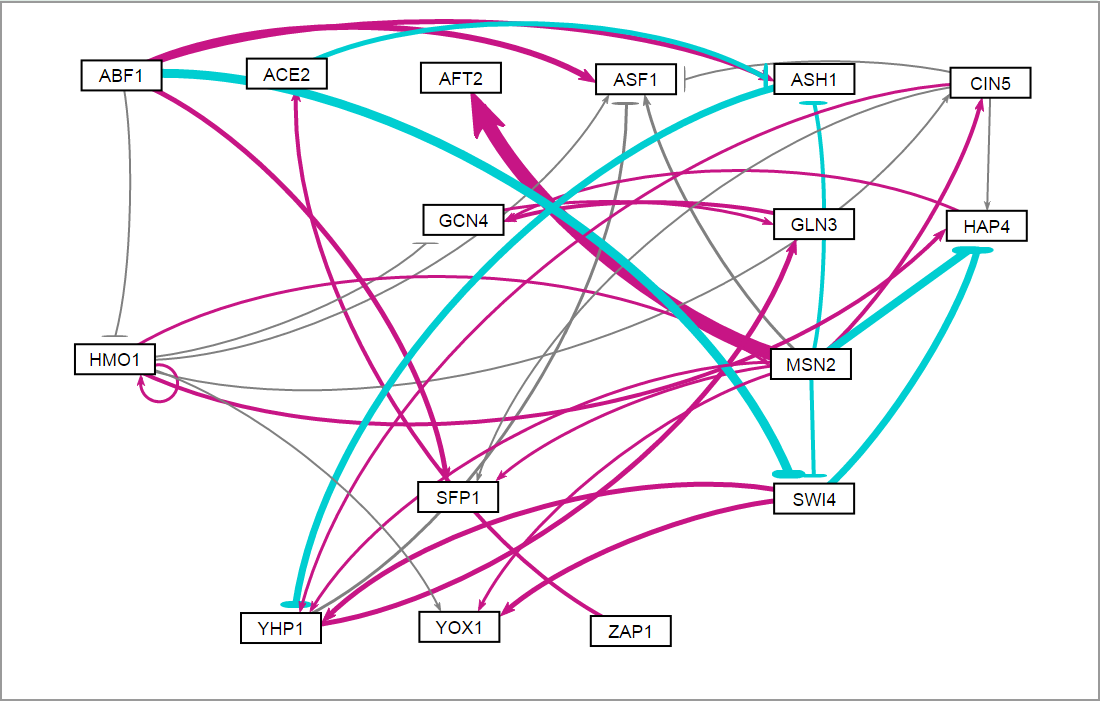
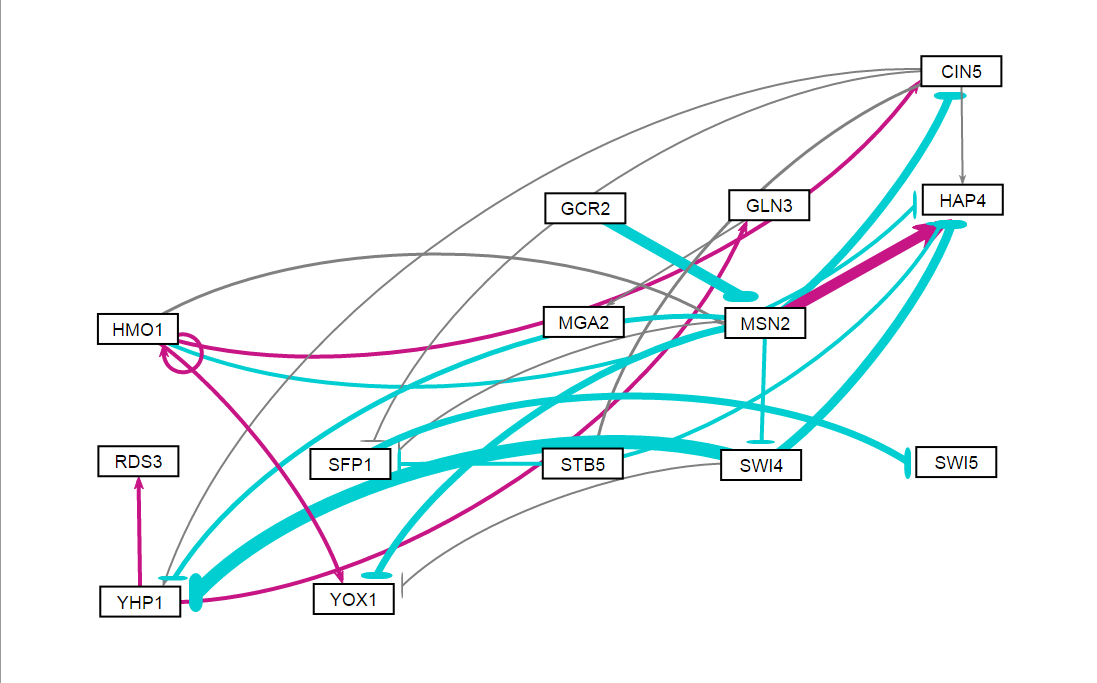
**Table 2.** Gene list for db1-db6. Factors that were deleted in tested strains of *Saccharomyces cerevisiae* are highlighted in yellow. Counts of the number of networks in which a gene is represented are displayed on the right.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **All Genes** | **db1** | **db2** | **db3** | **db4** | **db5** | **db6** | **Count** |
| ABF1 | ABF1 |  |  |  |  | ABF1 | 2 |
| ACE2 | ACE2 |  | ACE2 |  | ACE2 | ACE2 | 4 |
| AFT2 | AFT2 |  |  |  |  |  | 1 |
| ASF1 | ASF1 |  |  |  |  |  | 1 |
| ASH1 | ASH1 |  |  |  | ASH1 |  | 2 |
| CIN5 | CIN5 | CIN5 | CIN5 | CIN5 | CIN5 | CIN5 | 6 |
| CST6 |  |  |  |  |  | CST6 | 1 |
| CYC8 |  |  |  | CYC8 |  |  | 1 |
| GCN4 | GCN4 |  |  |  |  | GCN4 | 2 |
| GCR2 |  | GCR2 | GCR2 | GCR2 | GCR2 | GCR2 | 5 |
| GLN3 | GLN3 | GLN3 | GLN3 | GLN3 | GLN3 | GLN3 | 6 |
| HAP4 | HAP4 | HAP4 | HAP4 | HAP4 | HAP4 | HAP4 | 6 |
| HMO1 | HMO1 | HMO1 | HMO1 | HMO1 | HMO1 | HMO1 | 6 |
| HSF1 |  |  |  |  |  | HSF1 | 1 |
| MCM1 |  |  | MCM1 |  |  | MCM1 | 2 |
| MGA2 |  | MGA2 | MGA2 |  |  | MGA2 | 3 |
| MSN2 | MSN2 | MSN2 | MSN2 | MSN2 | MSN2 | MSN2 | 6 |
| MSN4 |  |  |  | MSN4 |  | MSN4 | 2 |
| RDS3 |  | RDS3 | RDS3 |  |  |  | 2 |
| SFP1 | SFP1 | SFP1 | SFP1 | SFP1 | SFP1 |  | 5 |
| STB5 |  | STB5 | STB5 |  | STB5 |  | 3 |
| SWI4 | SWI4 | SWI4 | SWI4 | SWI4 | SWI4 | SWI4 | 6 |
| SWI5 |  | SWI5 | SWI5 | SWI5 | SWI5 |  | 4 |
| TEC1 |  |  |  | TEC1 |  |  | 1 |
| YHP1 | YHP1 | YHP1 | YHP1 | YHP1 | YHP1 |  | 5 |
| YOX1 | YOX1 | YOX1 | YOX1 | YOX1 | YOX1 |  | 5 |
| ZAP1 | ZAP1 |  | ZAP1 |  | ZAP1 | ZAP1 | 4 |

* MSN2 and SWI4 were included in each of the six database-derived networks based on empirical evidence. CIN5, GLN3, HAP4, and HMO1 were manually included in all six GRNs, as deletion strain data was obtained from yeast in which these genes were knocked out.
* AFT2 and ASF1 only appear in db1.
  + Neither regulates other genes in the network (out-degrees of 0).
* CYC8 and TEC1 only appear in db4.
  + CYC8 does not regulate other genes in the network (out-degree of 0), but TEC1 regulates the expression of five other transcription factors in the network.
* CST6 and HSF1 only appear in db6.
  + Neither regulates other genes in the network (out-degrees of 0).
* SFP1, YHP1, and YOX1 are only *excluded* in db6.
  + db6 exhibited the second lowest LSE:min LSE ratio (1.4206).

**Table 3.** Characterization of weighted edges in db1-db6.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **db1** | **db2** | **db3** | **db4** | **db5** | **db6** |
| # Activation | 19 | 6 | 18 | 16 | 14 | 13 |
| # Repression | 7 | 11 | 10 | 12 | 10 | 11 |
| # Small Influence | 10 | 8 | 4 | 7 | 4 | 3 |
|  |  |  |  |  |  |  |
| **SUM:** | 8.1432 | -13.2390 | 1.5688 | -2.5750 | 1.2985 | -5.8876 |
| **SUM(activation):** | 23.4929 | 6.9575 | 18.8606 | 17.1864 | 17.7493 | 13.5913 |
| **SUM(repression):** | -15.3497 | -20.1966 | -17.2608 | -19.7614 | -16.4508 | -19.4790 |
| **SUM(abs):** | 38.8426 | 27.1541 | 36.1524 | 36.9479 | 34.2000 | 33.0703 |
| **MAX:** | 5.9424 | 3.0815 | 3.1810 | 2.4940 | 3.2167 | 3.7631 |
| **MIN:** | -3.3601 | -4.2947 | -3.6707 | -4.0281 | -4.0374 | -3.5141 |
| **MAX(abs):** | 5.9424 | 4.2947 | 3.6707 | 4.0281 | 4.0374 | 3.7631 |
| **SI Threshold:** | 0.2971 | 0.2147 | 0.1835 | 0.2014 | 0.2019 | 0.1882 |
| **MEAN:** | 0.2262 | -0.5296 | 0.0490 | -0.0736 | 0.0464 | -0.2181 |
| **STDEV:** | 1.6215 | 1.5085 | 1.5597 | 1.5026 | 1.6433 | 1.6812 |
| **MEDIAN:** | 0.3310 | -0.1214 | 0.2323 | 0.1139 | 0.2144 | 0.1276 |

* In **db1**, there was **substantially** **more activation than repression** of genes in the network.
  + Edge counts: 19 activation vs. 7 repression
  + Sum of activation edge weights = 23.4929
  + Sum of repression edge weights = -15.3497
* In **db2**, there was **substantially** **more repression that activation** of genes in the network.
  + Edge counts: 6 activation vs. 11 repression
  + Sum of activation edge weights = 6.9575
  + Sum of repression edge weights = -20.1966
* db1 and db2 exhibited a large number of small influence regulatory relationships (~30%), which we define as edges with regulatory weights < 5% of the maximum edge weight in the network. This contrasts to db3, db5, and db6, where < 15% of edges were classified as small influence.
  + In db1, the greatest activation relationship exhibited an edge weight of 5.9424, which far exceeded the weight of any other activation relationship in db2-db6 (next highest: 3.7631, db6).
    - The regulatory relationship in question is MSN2 🡪 AFT2.
    - Note: AFT2 only appears in db1, where MSN2 is its only input.
    - 
  + In db2, the greatest repression relationship exhibited an edge weight of -4.2947, which was the highest magnitude repression relationship of any network.
    - The regulatory relationship in question is SWI4 🡪 YHP1.
    - Note: This relationship appears in three other networks, although with lower magnitudes (and as a weak activation relationship in the case of db1).
    - 
  + These outliers contribute to the observed increase in “small influence” edges in db1 and db2. They might also influence the large imbalance of activation and repression in these two networks.
* In **db3-db5**, there were more activation than repression regulatory relationships, although the **overall magnitudes of activation and repression in the networks are similar**.
  + Activating edges: 18, 16, 14
  + Repressing edges: 10, 12, 10
  + Sum of activation edge weights = 18.8606, 17.1864, 17.7493
  + Sum of repression edge weights = -17.2608, -19.7614, -16.4508
* In **db6**, there were similar numbers of activation and repression regulatory relationships, although the **overall magnitude of repression exceeded that of activation in the network**.
  + Edge counts: 13 activation vs. 11 repression
  + Sum of activation edge weights = 13.5913
  + Sum of repression edge weights = -19.4790
* For db2-db5, the edge weights with the largest magnitudes cluster together in a relatively tight range: 3.6707-4.2947.

**Table 4.** Heat map comparing the edge weights of regulatory relationships in db1-db6.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **db1** | **db2** | **db3** | **db4** | **db5** | **db6** |  | **db1** | **db2** | **db3** | **db4** | **db5** | **db6** |
| ABF1->ASF1 |  |  |  |  |  |  | MCM1->SWI4 |  |  |  |  |  |  |
| ABF1->ASH1 |  |  |  |  |  |  | MCM1->SWI5 |  |  |  |  |  |  |
| ABF1->CST6 |  |  |  |  |  |  | MCM1->YHP1 |  |  |  |  |  |  |
| ABF1->HMO1 |  |  |  |  |  |  | MCM1->YOX1 |  |  |  |  |  |  |
| ABF1->MGA2 |  |  |  |  |  |  | MSN2->AFT2 |  |  |  |  |  |  |
| ABF1->MSN4 |  |  |  |  |  |  | MSN2->ASF1 |  |  |  |  |  |  |
| ABF1->SFP1 |  |  |  |  |  |  | MSN2->ASH1 |  |  |  |  |  |  |
| ABF1->SWI4 |  |  |  |  |  |  | MSN2->CIN5 |  |  |  |  |  |  |
| ACE2->ASH1 |  |  |  |  |  |  | MSN2->CYC8 |  |  |  |  |  |  |
| ASH1->YHP1 |  |  |  |  |  |  | MSN2->HAP4 |  |  |  |  |  |  |
| CIN5->ASF1 |  |  |  |  |  |  | MSN2->MSN4 |  |  |  |  |  |  |
| CIN5->CYC8 |  |  |  |  |  |  | MSN2->SFP1 |  |  |  |  |  |  |
| CIN5->HAP4 |  |  |  |  |  |  | MSN2->SWI4 |  |  |  |  |  |  |
| CIN5->SFP1 |  |  |  |  |  |  | MSN2->TEC1 |  |  |  |  |  |  |
| CIN5->STB5 |  |  |  |  |  |  | MSN2->YHP1 |  |  |  |  |  |  |
| CIN5->TEC1 |  |  |  |  |  |  | MSN2->YOX1 |  |  |  |  |  |  |
| CIN5->YHP1 |  |  |  |  |  |  | SFP1->SWI5 |  |  |  |  |  |  |
| GCN4->GLN3 |  |  |  |  |  |  | STB5->HAP4 |  |  |  |  |  |  |
| GLN3->GCN4 |  |  |  |  |  |  | STB5->SFP1 |  |  |  |  |  |  |
| GLN3->MGA2 |  |  |  |  |  |  | SWI4->CYC8 |  |  |  |  |  |  |
| GCR2->MSN2 |  |  |  |  |  |  | SWI4->HAP4 |  |  |  |  |  |  |
| HAP4->GCN4 |  |  |  |  |  |  | SWI4->YHP1 |  |  |  |  |  |  |
| HMO1->ASF1 |  |  |  |  |  |  | SWI4->YOX1 |  |  |  |  |  |  |
| HMO1->CIN5 |  |  |  |  |  |  | SWI5->ASH1 |  |  |  |  |  |  |
| HMO1->CYC8 |  |  |  |  |  |  | SWI5->TEC1 |  |  |  |  |  |  |
| HMO1->GCN4 |  |  |  |  |  |  | TEC1->CIN5 |  |  |  |  |  |  |
| HMO1->HAP4 |  |  |  |  |  |  | TEC1->HAP4 |  |  |  |  |  |  |
| HMO1->HMO1 |  |  |  |  |  |  | TEC1->MSN2 |  |  |  |  |  |  |
| HMO1->HSF1 |  |  |  |  |  |  | TEC1->SFP1 |  |  |  |  |  |  |
| HMO1->MCM1 |  |  |  |  |  |  | TEC1->YHP1 |  |  |  |  |  |  |
| HMO1->MSN2 |  |  |  |  |  |  | YHP1->ASF1 |  |  |  |  |  |  |
| HMO1->MSN4 |  |  |  |  |  |  | YHP1->GLN3 |  |  |  |  |  |  |
| HMO1->TEC1 |  |  |  |  |  |  | YHP1->RDS3 |  |  |  |  |  |  |
| HMO1->YOX1 |  |  |  |  |  |  | ZAP1->ACE2 |  |  |  |  |  |  |
| MCM1->ACE2 |  |  |  |  |  |  |  |  |  |  |  |  |  |



**Maximum: 1**

**Minimum: -1**

* The following regulatory relationships were conserved across the database-derived networks, always appearing as activation only or repression only when present:
  + Activation, CIN5🡪HAP4 (6 networks)
  + Activation, HMO1🡪CIN5 (6 networks)
  + Activation, HMO1🡪HMO1 (6 networks)
  + Activation, HMO1🡪MSN2 (6 networks)
  + Activation, HMO1🡪YOX1 (5 networks)
  + Activation, YHP1🡪GLN3 (5 networks)
  + Activation, ZAP1🡪ACE2 (4 networks)
  + Repression, SWI4🡪HAP4 (6 networks)
  + Repression, SFP1🡪SWI5 (4 networks)
  + The conserved activation relationships among CIN5, GLN3, HMO1, and YHP1 form the following regulatory chain: HMO1 🡪 CIN5 🡪 YHP1 🡪 GLN3.
* The following high magnitude edge weights appeared in the database-derived networks both as activation and repression relationships:
  + GCR2🡪MSN2 (2 Activation, 3 Repression)
    - Note: MSN2 is the only connection to GCR2 in these networks.
  + MSN2🡪CIN5 (2 Activation, 4 Repression)
  + MSN2🡪HAP4 (3 Activation, 3 Repression)
  + MSN2🡪SFP1 (3 Activation, 2 Repression)
  + MSN2🡪SWI4 (3 Activation, 3 Repression)
  + MSN2🡪YHP1 (3 Activation, 2 Repression)
  + MSN2🡪YOX1 (3 Activation, 2 Repression)
  + SWI4🡪YHP1 (2 Activation, 3 Repression)
  + SWI4🡪YOX1 (1 Activation, 2 Repression)
  + MCM1🡪SWI4 (1 Activation, 1 Repression)
  + All but one of these regulatory relationships involve either MSN2 or YHP1, which both exhibit high betweenness centralities.

[MSN2: high betweenness centrality, gatekeeper of downstream dynamics, dynamic modeling.

vs. HMO1… originator of dynamics, well-conserved?]

db1-db6.pdf

* Common edges between db5 & 30 random networks.
  + 3 Best
    - rand15 (LSE:minLSE = 1.4063)
      * 0 shared edges.
    - rand16 (LSE:minLSE = 1.4165))
      * 2 shared activation relationships.
      * 1 shared relationship w/ opposite weights.
    - **rand24 (LSE:minLSE = 1.3880)**
      * **1 shared activation relationship**
      * **1 shared relationship w/ opposite weights.**
  + 3 Most Similar (db5 Lse:minLSE = 1.4263)
    - rand2 (LSE:minLSE = 1.4278)
      * 4 shared relationships w/ opposite weights.
    - rand3 (LSE:minLSE = 1.4302)
      * 1 shared relationship w/ opposite weights.
    - rand9 (LSE:minLSE = 1.4247)
      * 1 shared activation relationship.
  + 3 Worst
    - **rand7 (LSE:minLSE = 1.5202)**
      * **1 shared relationship w/ opposite weights.**
    - rand12 (LSE:minLSE = 1.5080)
      * 1 shared activation relationship.
      * 1 shared relationship w/ opposite weights.
    - rand31 (LSE:minLSE = 1.5009)
      * 1 shared relationship w/ opposite weights.
  + Other Notables
    - rand23 (LSE:minLSE = 1.4332)
      * 3 shared activation relationships.
      * 3 shared repression relationships.
      * 1 shared relationship w/ opposite weights.
    - rand26 (LSE:minLSE = 1.4206)
      * 3 shared activation relationships.
      * 4 shared repression relationships.