**Analysis of each node in db5 using graph statistics**

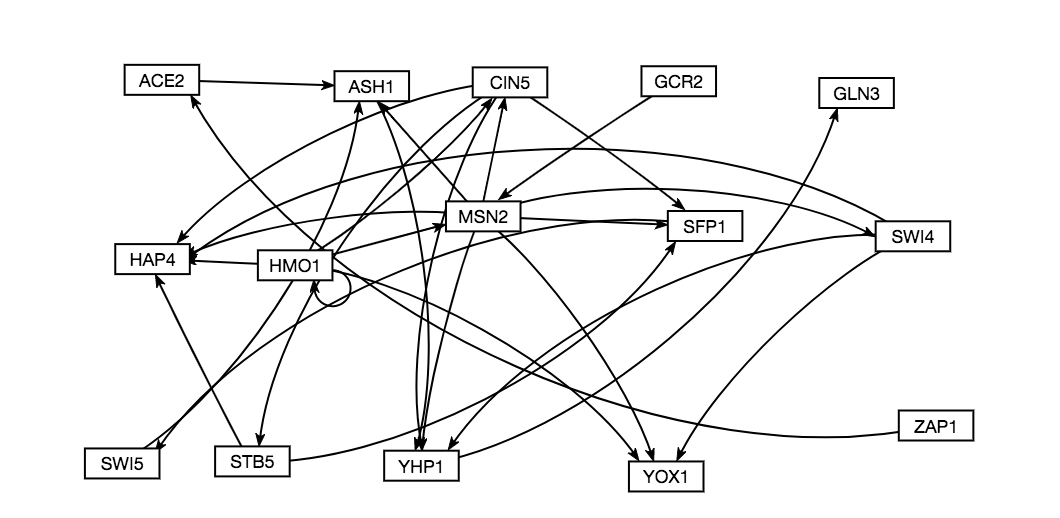


Table 1. Compilation of graph statistics as computed by Gephi for the network db5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gene | Closeness Centrality | Betweeness Centrality | Eigencentrality | Eccentricity |
| ACE2 | 0.5 | 3 | 0.008418 | 3 |
| ASH1 | 0.666667 | 10 | 0.575118 | 2 |
| CIN5 | 0.636364 | 5 | 0.249597 | 3 |
| GCR2 | 0.458333 | 0 | 0 | 3 |
| GLN3 | 0 | 0 | 0.8377 | 0 |
| HAP4 | 0 | 0 | 0.861994 | 0 |
| HMO1 | 0.55 | 0 | 0.11352 | 3 |
| MSN2 | 0.769231 | 14 | 0.121938 | 2 |
| SFP1 | 0.4 | 9 | 0.605438 | 4 |
| STB5 | 0.375 | 0 | 0.248138 | 5 |
| SWI4 | 0.8 | 0 | 0.136077 | 2 |
| SWI5 | 0.5 | 7 | 0.52969 | 3 |
| YHP1 | 1 | 11 | 1 | 1 |
| YOX1 | 0 | 0 | 0.392633 | 0 |
| ZAP1 | 0.4 | 0 | 0 | 4 |

**ACE2**

* Betweeness Centrality: 3
  + A Betweeness of 3 means that ACE2 is contained in 3 shortest paths on the network, so it is being used as a hub for a small number of nodes to reach other nodes in the graph.
* Closeness Centrality: 0.5
  + A Closeness of 0.5 is relatively high, which means many nodes have paths going to, or going through ACE2, which makes sense in that it acts as a hub for several nodes.
* Eccentricity: 3
  + An eccentricity of 3 means that compared to other nodes in the graph, ACE2 has a similar “reach” or influence. This moderate influence indicates that the hub nature of the node is of moderate importance to the network.
* Eigenvector Centrality: 0.008418
  + An eigenvector centrality of 0.008418 is incredibly low, and indicates that because the node is deemed central by other measures, and because the in:out degree ratio is 1:1, the eigenvector statistic is labeling the node as unimportant.

**ASH1**

* Betweeness Centrality: 10
  + With a Betweeness of 10, ASH1 is shown to be acting like a large hub for the network, with many shortest paths having to pass through ASH1
* Closeness Centrality: 0.666667
  + A closeness centrality of 0.666667 is relatively high, which makes sense in conjunction with the high Betweeness measure, as many paths in the network have to pass through or go to ASH1
* Eccentricity: 2
  + An eccentricity of 2 means that ASH1 has slightly less influence on the graph than the majority of the nodes, which would indicate that in acting as a hub, it is more of a way station than a command center in sending out activating or suppressing influences across the network.
* Eigenvector Centrality: 0.575118
  + With a relatively high eigenvector centrality of 0.575118 and a Betweeness value also so high, this means that ASH1 has more in degrees than out degrees, and that more nodes are trying to regulate it than it is regulating other nodes.

**CIN5**

* Betweeness Centrality: 5
  + With a Betweeness centrality of 5, CIN5 is operating as a moderately sized hub in the network, with several shortest paths passing through the node.
* Closeness Centrality: 0.636364
  + A closeness of 0.636364 makes sense in this node, as it means it is highly central to the network, which further confirms the hub nature of the node.
* Eccentricity: 3
  + An eccentricity of 3 indicates that CIN5 has a moderate level of influence over the network, as an eccentricity of 3 is about average for the network.
* Eigenvector Centrality: 0.249597
  + An eigenvector centrality of 0.249597 means that the node has more out degrees than in degrees, and is having more influence on other nodes in the network than nodes are having on it.

**GCR2**

* Betweeness Centrality: 0
  + With a Betweeness centrality of 0, and looking at the network, GCR2 is at the start of a regulatory chain, and not a hub in the network.
* Closeness Centrality: 0.458333
  + With a closeness centrality of 0.45833, this means that GCR2 is moderately connected to the rest of the network, and through it’s connection to MSN2, it has many shortest paths connecting it to other nodes.
* Eccentricity: 3
  + With an eccentricity of 3, it can be determined that while at the start of a regulatory pathway, GCR2 has an average level of influence over the rest of the network, which when compared to other nodes at the start of regulatory pathways, might help determine the importance of GCR2.
* Eigenvector Centrality: 0
  + GCR2 has an eigenvector centrality of 2, which means nothing is influencing the node (no in degrees), which makes sense seeing as how the node only has one out degree.

**GLN3:**

* Betweeness Centrality: 0
  + With a Betweeness centrality of 0 and looking at the graph, it can be determined that GLN3 is at the end of a regulatory pathway.
* Closeness Centrality: 0
  + A closeness centrality of 0 makes sense, as there are no out degrees for GLN3, and therefore no edges are emanating from it that form a shortest path.
* Eccentricity: 0
  + An eccentricity of 0 makes sense for GLN3, as there are no nodes for it to influence, since there is no out degree for this node.
* Eigenvector Centrality: 0.8377
  + As the in:out degree ratio for GLN3 is 1:0, it makes sense for this node to have a very high eigenvector centrality at 0.8377, as nodes are regulating it, and it is regulating no nodes.

**HAP4:**

* Betweeness Centrality: 0
  + With a Betweeness centrality of 0 and looking at the network, HAP4 is at the end of a regulatory pathway, which makes sense why no shortest paths are passing through the node.
* Closeness Centrality: 0
  + With a closeness centrality also at 0, this makes sense as there are no edges emanating out from the node.
* Eccentricity: 0
  + With an eccentricity of 0, it makes sense that HAP4 has no influence over other nodes in the network.
* Eigenvector Centrality: 0.861994
  + HAP4 has a high eigenvector centrality at 0.861994 because the ratio of in degree:out degree is 5:0, which shows many nodes are influencing HAP4.

**HMO1:**

* Betweeness Centrality: 0
  + With a Betweeness centrality of 0, it can be determined that the Gephi measure for Betweeness does not take self-regulation into account as an in degree. Therefore, HMO1 does not act as a hub, and there are no shortest paths that go through HMO1, making it the start of a regulatory pathway.
* Closeness Centrality: 0.55
  + With a closeness centrality of 0.55, HMO1 is moderately connected to the rest of the network, with many shortest paths emanating from HM01.
* Eccentricity: 3
  + With an eccentricity value at 3, HMO1 has the same level of influence as the majority of the genes in the network, indicating that while it is the start of a regulatory pathway, it is not necessarily of the most importance.
* Eigenvector Centrality: 0.11352
  + With a low eigenvector centrality, HMO1 is influencing more nodes than are influencing it, which makes sense with the other centrality measures calculated for this node/network.

**MSN2:**

* Betweeness Centrality: 14
  + With a Betweeness centrality of 14, MSN2 is shown to be the biggest hub in the network, with many nodes containing MSN2 on a shortest path. This means it is a highly important node in that it acts as a step stone that is incredibly central to the structure of the network.
* Closeness Centrality: 0.769231
  + With a very high closeness centrality, it is clear that not only do shortest paths go through MSN2, but it also has a number of shortest paths emanating from it, further showing the importance of this node as a hub for many edges in the network.
* Eccentricity: 2
  + With an eccentricity of 2, MSN2 might not have the farthest reach across the graph, but in being incredibly central to the graph, it is also possible that this number is lower than average because MSN2 does not have to reach as far as other nodes to get to the furthest node from it.
* Eigenvector Centrality: 0.121938
  + With many in and out degrees, it makes sense that MSN2 has a very low eigenvector centrality, as the number of nodes regulating it, as it is regulating is fairly similar.

**SFP1:**

* Betweeness Centrality: 9
  + With a Betweeness of 9, SFP1 is acting as a moderate sized hub for the network, with several shortest paths going through the node.
* Closeness Centrality: 0.4
  + With a closeness centrality of 0.4, SFP1 has a moderate closeness centrality, which means it is moderately connected to the rest of the network, which makes sense given the hub-like nature of the node being described by the high Betweeness centrality value.
* Eccentricity: 4
  + With an eccentricity of 4, SFP1 is able to reach and influence more nodes in the graph than is average of this network. This means it is a network of high influence, and based on the weight of the edges emanating from this node, it might be determined what the influence of this node is on the network.
* Eigenvector Centrality: 0.605438
  + With a relatively high eigenvector centrality of 0.605438, it can be determined that this hub has more in degrees than out degrees, and thus the few edges that are coming out of the node have a far reach over the graph, and are important to the structure of the network.

**STB5:**

* Betweeness Centrality: 0
  + With a Betweeness of 0, STB5 is not being used as a hub for the graph, which makes sense, as in the network STB5 is at the start of a regulatory pathway
* Closeness Centrality: 0.375
  + With a closeness centrality of 0.375, STB5 has a moderate closeness centrality, which makes sense as STB5 has two out degrees, which are both shortest paths.
* Eccentricity: 5
  + With an eccentricity of 5, STB5 has the highest eccentricity in the network. This means it has the highest level of influence over the network, and the farthest “reach” across the network. This makes sense, as the edge between STB5 and SFP1 connects STB5 to the rest of the network.
* Eigenvector Centrality: 0.248138
  + An eigenvector centrality of 0.248138 means that the node is not very important in the graph, which makes sense as it only has out degrees, with no in degrees. This means it is inherently influencing more nodes than are influencing it.

**SWI4:**

* Betweeness Centrality: 0
  + With a Betweeness centrality of 0, and both in and out degrees, this means that while pathways exist going through SWI4, those pathways are not the shortest pathways that exist between nodes, and so there exists more direct pathways between genes such as MSN2 and YOX1 (a pathway that goes through SWI4)
* Closeness Centrality: 0.8
  + SWI4 has the highest closeness centrality in the network, which would indicate that it has the largest amount of shortest paths emanating from the node. Following the edges coming from SWI4, this makes sense, as the edges lead to CIN5, and other nodes with a high Betweeness.
* Eccentricity: 2
  + With an eccentricity of 2, SWI4 does not have the furthest reach across the network. This might indicate that SWI4 is centrally located in the connection of edges in the network.
* Eigenvector Centrality: 0.136077
  + With an eigenvector centrality of 0.136077, SWI4 has a low eigenvector centrality, meaning that SWI4 is of low importance in the graph, and the ratio of in to out degrees is close to 1:1.

**SWI5:**

* Betweeness Centrality: 7
  + With a Betweeness of 7, SWI5 is used by the network as a fairly central hub between nodes, with many shortest paths going through SWI5.
* Closeness Centrality: 0.5
  + With a closeness of 0.5, SWI5 has a moderate number of paths emanating from the node, through a connection from SWI5 to ASH1.
* Eccentricity: 3
  + With an eccentricity of 3, SWI5 has a moderate level of influence over the rest of the network, with 3 being about average for eccentricity measures across the network.
* Eigenvector Centrality: 0.52969
  + With an eigenvector centrality of 0.52969, SWI5 is of moderate importance to the graph, which might indicate, given that its in:out degree ratio is 1:1, that SWI5 is closer to the end of the regulatory pathway it is on than to the beginning of the pathway.

**YHP1:**

* Betweeness Centrality: 11
  + With a Betweeness centrality of 11, YHP1 is being used frequently as a hub between nodes, which makes sense given that it has in degrees coming from MSN2, CIN5, and ASH1, which also are large hubs in the network.
* Closeness Centrality: 1
  + With a value of 1, YHP1 has the highest closeness centrality measure in the network. This might be due to the large number of hubs that it is connected to, which would lead to the largest number of edges emanating from YHP1 than any other node in the graph.
* Eccentricity: 1
  + With an eccentricity of 1, YHP1 does not have much reach across the network, which would make sense, as the other graph statistics seem to indicate that YHP1 is the most central hub in the graph, which would mean it doesn’t have to reach far to access the furthest node from it.
* Eigenvector Centrality: 1
  + With an eigenvector centrality of 1, YHP1 is evidently the most important node in the graph, which would make sense given the information provided by the other graph statistics.

**YOX1:**

* Betweeness Centrality:
  + With a Betweeness of 0, YOX1 is not being used as a hub for any nodes in the graph. This makes sense, as in looking at the graph it is at the end of a regulatory pathway.
* Closeness Centrality: 0
  + YOX1 has a closeness of 0, which makes sense as there are no out degrees emanating from the graph, and it is at the end of a regulatory pathway.
* Eccentricity: 0
  + With an eccentricity of 0, YOX1 has no reach across the network, which makes sense given that the node is at the end of a regulatory pathway, and thus has nothing that it can reach to.
* Eigenvector Centrality: 0.392633
  + With an eigenvector centrality of 0.392633, YOX1 has a much lower eigenvector centrality than other nodes that are also at the end of regulatory pathways. This might be because the nodes regulating YOX1 are not the most connected in the graph.

**ZAP1**:

* Betweeness Centrality: 0
  + With a Betweeness centrality of 0, ZAP1 is not being used as a hub for any nodes in the graph. Looking at the network, this is due to ZAP1 only having one out degree, regulating ACE2.
* Closeness Centrality: 0.4
  + With a closeness centrality of 0.4, the path emanating from ZAP1 is deemed to be moderately important to the network, which makes sense given that the edge eventually connects to YHP1, which statistically, seems to be the most important node in the network
* Eccentricity: 4
  + With an eccentricity of 4, ZAP1 has above average reach across the graph, which makes sense given that in regulating ACE2, it is indirectly influencing some of the important hubs in the network.
* Eigenvector Centrality: 0
  + With an eigenvector centrality of 0, ZAP1 is deemed to be very unimportant in the graph, which makes sense given that the in:out degree ratio for this node is 0:1.