**Post Genomics – Fall 2025**

**Homework 5**

**Due: Thursday, October 2, 2025**

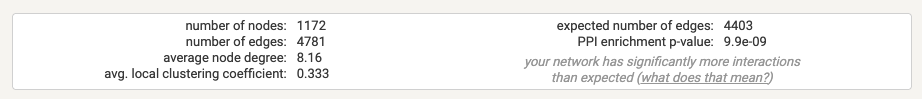
**Part 1. Protein-Protein Interaction Networks. (Non-Programming)** *(60 points)*

1.1 Use the CSV (AML\_gene.csv), to determine, create a list, of the unique genes within this dataset. Submit these genes into String-DB for the Protein-Protein Interaction Network:

<https://string-db.org/cgi/input?sessionId=byv6kGAGpNla&input_page_active_form=multiple_identifiers>

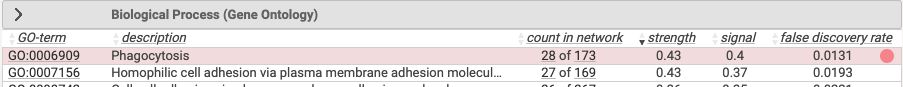
1.2 Under the Analysis Tab,

1.2.1 Take a screen shot of the Network Stats



1.2.2 For each Gene Ontology terms (Biological Process, Molecular Function, and Cellular Component) what is the top strength term? What is the count of network for each of these terms?

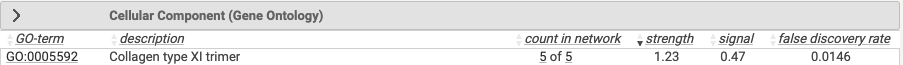
For Biological process:



For molecular function:



For Cellular Component:



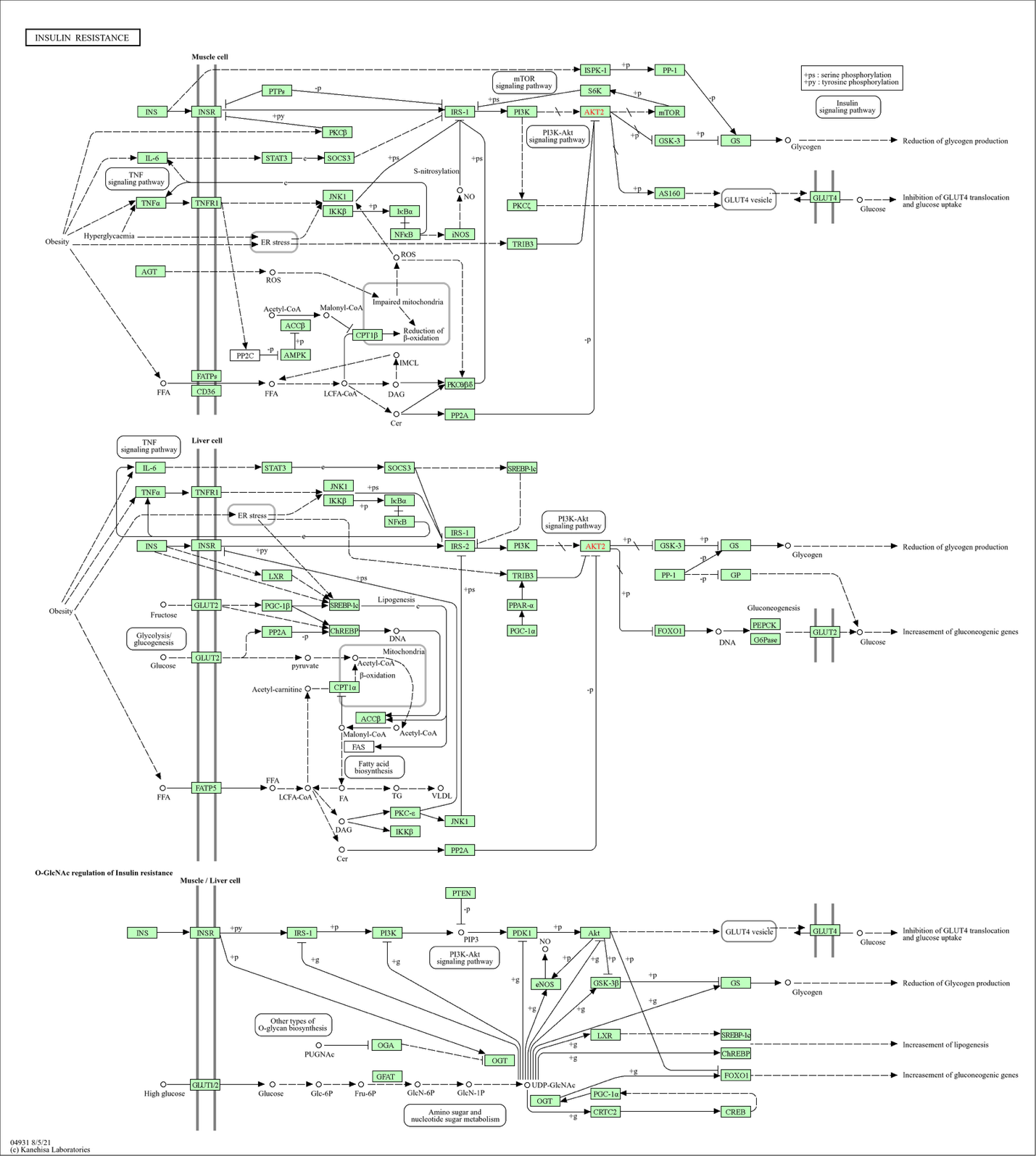
1.2.3 The format is #A of #B, what does #A represent and what does #B represent?

The first number (#A) indicates how many proteins in the network are annotated with a particular GO-term. The second number (#B) indicates how many proteins in total (in the network and in the background) have this term assigned.

1.2.4 What is the top strength KEGG pathway? What is its count of network? Click the link for the pathway and submit a screen shot of the pathway.

A screenshot of a computer

AI-generated content may be incorrect.



1.3 Under the Exports Tab,

1.3.1 Download the “protein node degrees” file. How many proteins are there that have a node degree above the found average from the Network Stats?

There are 410 proteins with the node degrees > 8.16

1.3.2 Download the “as short tabular text output” file. Within this file, what information contributes to the Combined Score? What is the highest combine scored interaction?

The combined score depends on all previous rows: (“neighborhood\_on\_chromosome, gene\_fusio, phylogenetic\_cooccurrence, homology, coexpression, experimentally\_determined\_interaction, database\_annotated, automated\_textmining”)

The highest combine score is: 0.999

1.4 Under the Settings Tab,

1.4.1 Submit a screen shot of the network as is.

A large cluster of colorful dots

AI-generated content may be incorrect.

1.4.2 Select the “hide disconnected nodes in the network” and update the network with this feature. Submit a screen shot of the new network.

A close-up of a network

AI-generated content may be incorrect.

1.4.3 In your own words, what changes did you observe, if any.

The network is clearer and shows less points that leads to nothing.

**Part 2. Node Network Practice. (Non-Programming)** *(40 points)*

Using the figure below, answer the following questions (assume that each edge is of length 1).

Start

A picture containing indoor, wire, accessory, lamp

Description automatically generated

End

2.1 What is the longest path to get from the Start to End node? You can either: Draw by hand and submit image, edit image virtually on iPad/Computer, or give list of directionals (down, right, left, etc. )

Start

A diagram of a molecule

AI-generated content may be incorrect.

End

2.2 How many edges are on the longest path?

There are 8 edges in the longest path.

2.3 Repeat 2.1 and 2.2 for the shortest path.

A green and grey line with dots and lines

AI-generated content may be incorrect.

End

Start

There are 5 edges in the shortest path.

**What to Submit:**

1. **A single PDF with responses and screen shots from Part 1 and 2. (Your Last Name) \_HW5.pdf *\*\* Submit on Blackboard \*\****