UTILIZING BIOLOGICAL NETWORK-BASED TOOLS TO BETTER UNDERSTAND DISEASE

NOLAN KEITH NEWMAN

IN-BIOS 5000 / 9000

NOVEMBER 6TH, 2024

OVERVIEW

How can we model complex diseases?

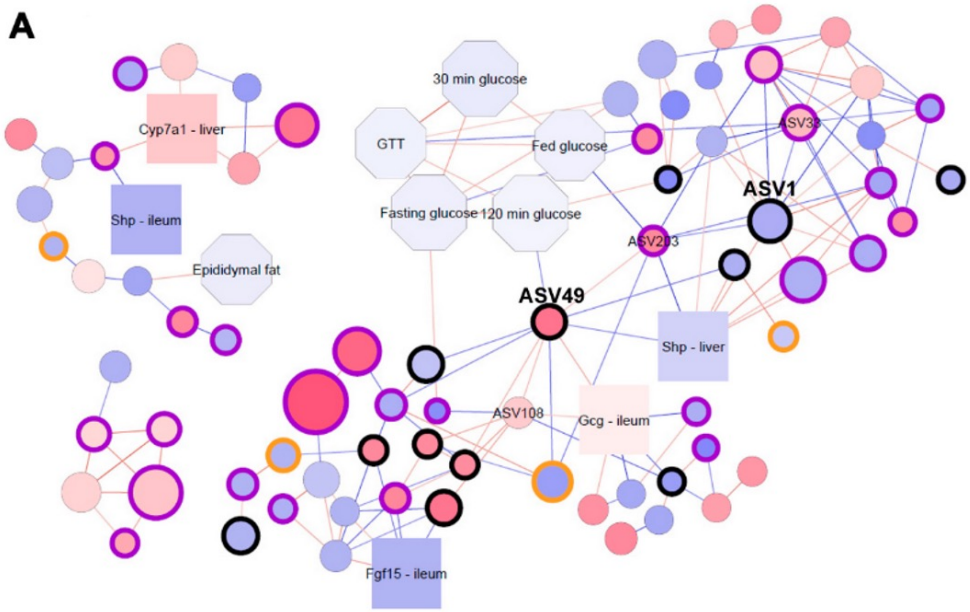
Using TkNA to identify potential drivers of disease

Gene regulation in disease

Identifying novel cancer subtypes using PANDA and LIONESS

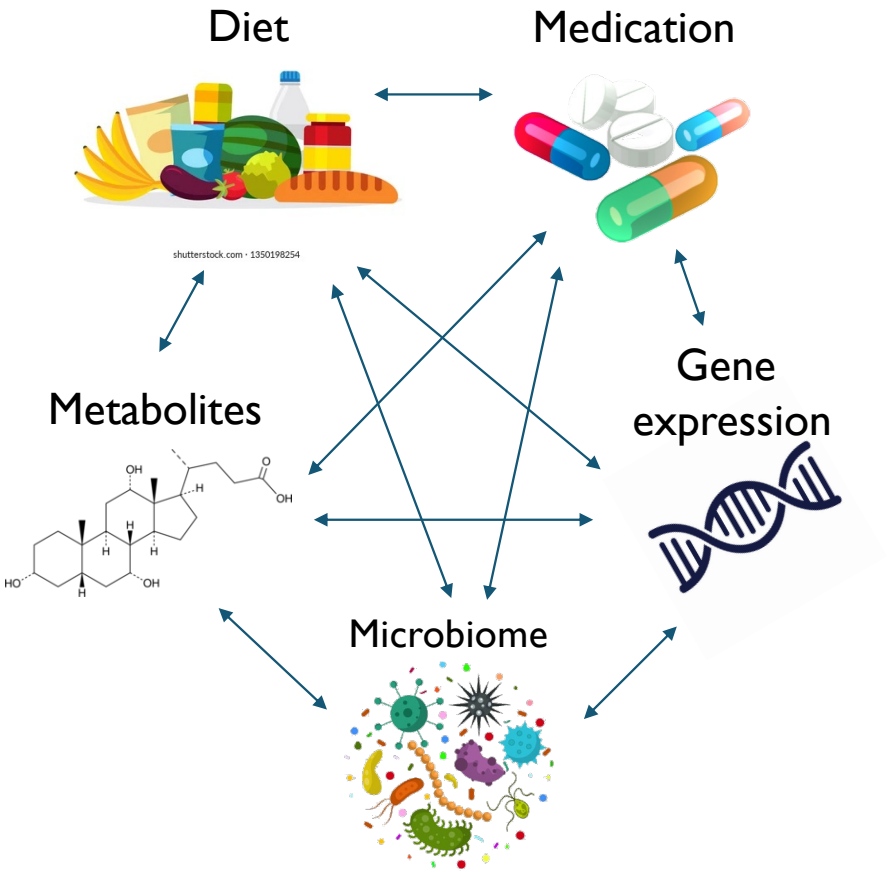
USING NETWORKS TO UNDERSTAND DISEASE

- There are many different types of networks
  - Social networks
  - Computer networks
  - Biological networks
  - etc,.
- Networks are composed of **nodes** and **edges**
  - These nodes and edges reflect different things in different networks



Newman et al., 2024

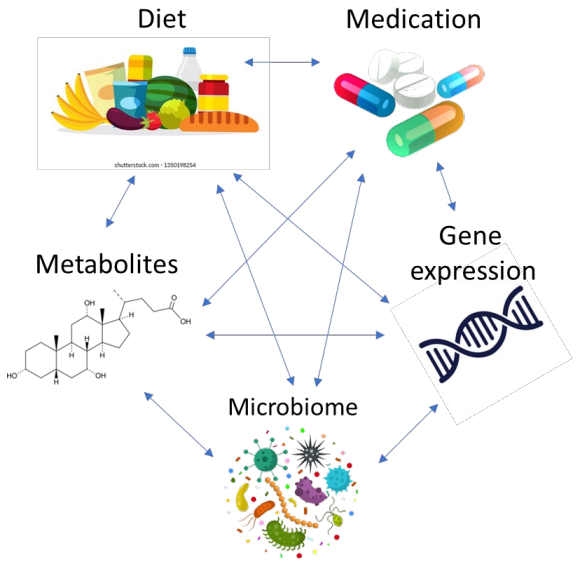
UNDERSTANDING DISEASE DEVELOPMENT IS NOT EASY



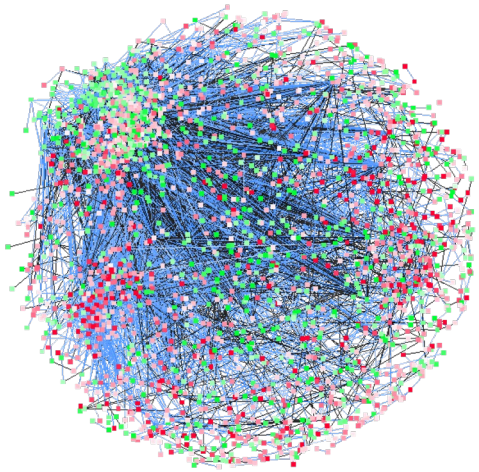
Disease

But how do we know which of these *cause* disease?

WHAT IS A NETWORK?



=

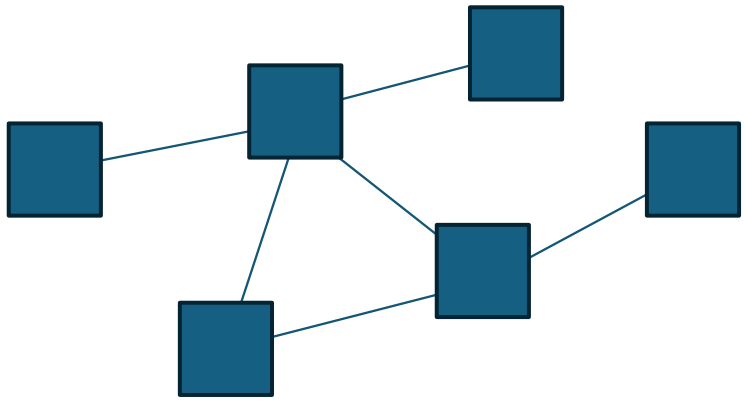


=

NETWORK STRUCTURES

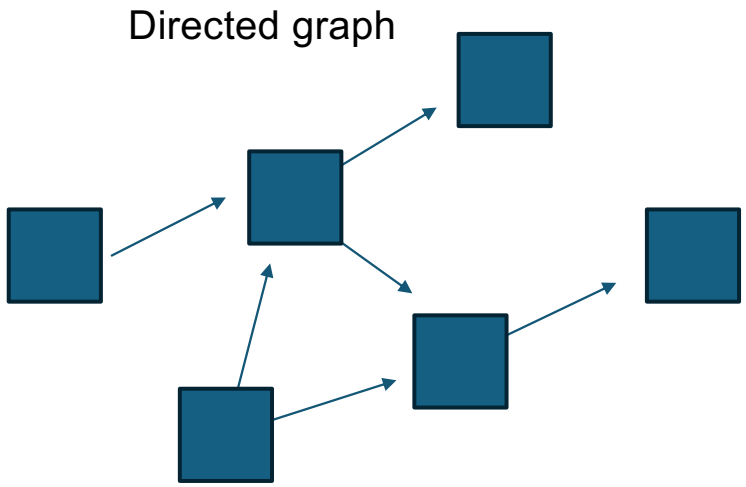
- Undirected graphs
- Directed graphs
- Directed acyclic graphs
- Bipartite graphs
- Complete graphs
- Disconnected graph

Undirected graph



NETWORK STRUCTURES

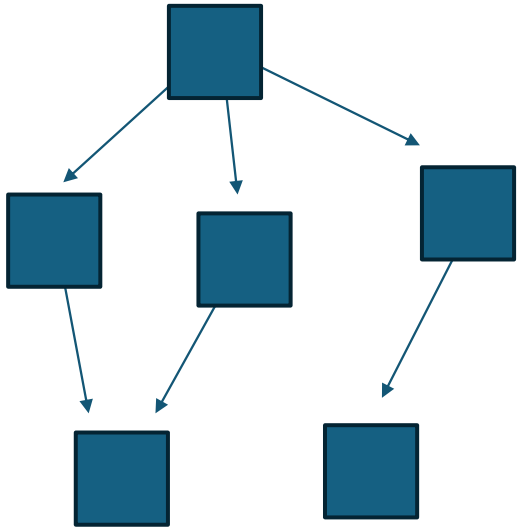
- Undirected graphs
- Directed graphs
- Directed acyclic graphs
- Bipartite graphs
- Complete graphs
- Disconnected graph



NETWORK STRUCTURES

- Undirected graphs
- Directed graphs
- Directed acyclic graphs
- Bipartite graphs
- Complete graphs
- Disconnected graph

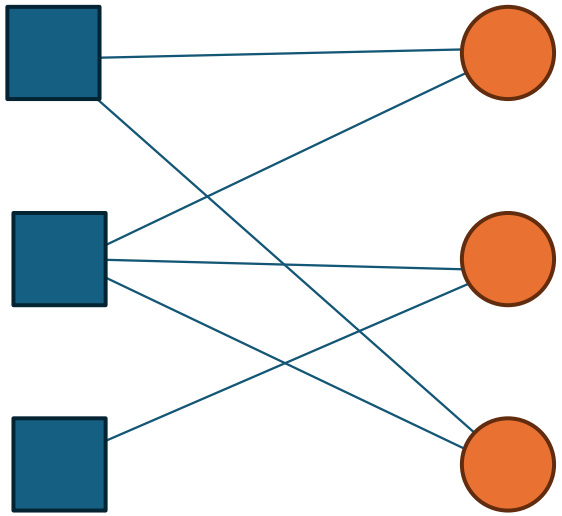
Directed acyclic graph



NETWORK STRUCTURES

- Undirected graphs
- Directed graphs
- Directed acyclic graphs
- Bipartite graphs
- Complete graphs
- Disconnected graph

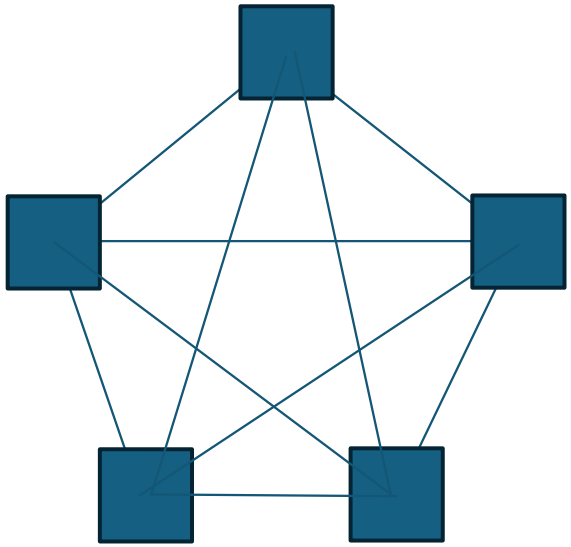
Bipartite graph



NETWORK STRUCTURES

- Undirected graphs
- Directed graphs
- Directed acyclic graphs
- Bipartite graphs
- Complete graphs
- Disconnected graph

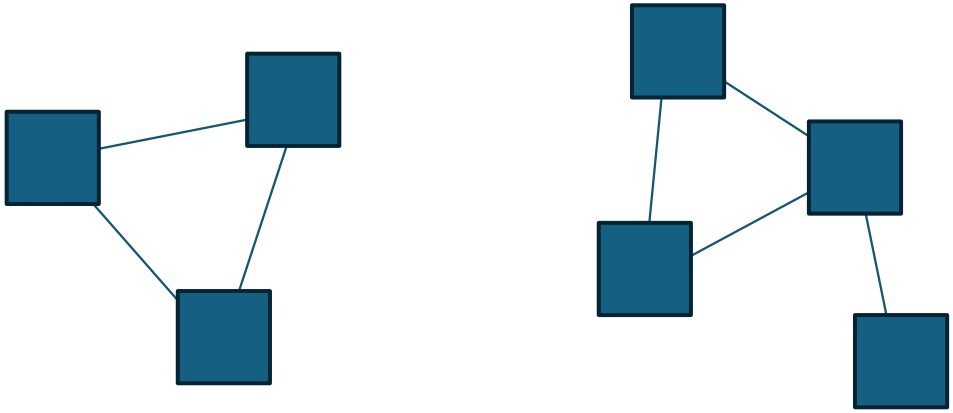
Complete graph



NETWORK STRUCTURES

- Undirected graphs
- Directed graphs
- Directed acyclic graphs
- Bipartite graphs
- Complete graphs
- Disconnected graph

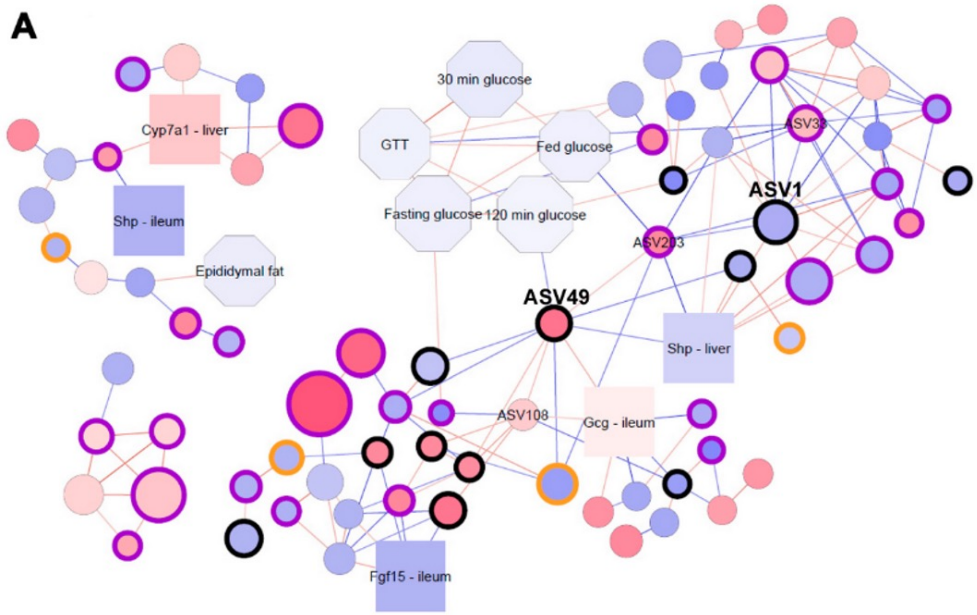
Disconnected graph



NETWORK STRUCTURES

- Undirected graphs
- Directed graphs
- Directed acyclic graphs
- Bipartite graphs
- Complete graphs
- Disconnected graph

Question: Based on these terms, how would you describe this network?



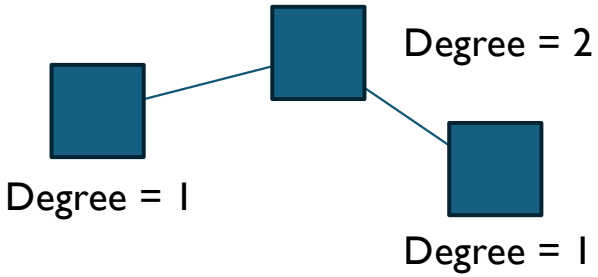
Newman et al., 2024

NETWORK INTERROGATION

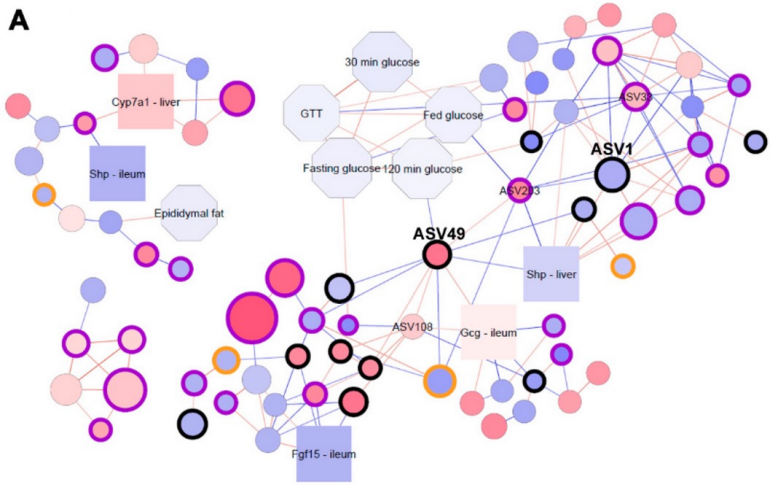
- **Network interrogation** refers to the analyses done on a network to understand the relationship between its parts
- We can calculate many different types of node/network properties
  - Degree
  - Degree assortativity
  - Density
  - Betweenness centrality
  - Bipartite betweenness centrality
  - Shortest paths
  - Modularity
  - Community detection
  - Clustering coefficient
  - Closeness centrality

DEGREE

- **Degree** refers to the number of connections a single node has to other nodes in the network



Question: In a network like the one below, what might a high degree indicate?



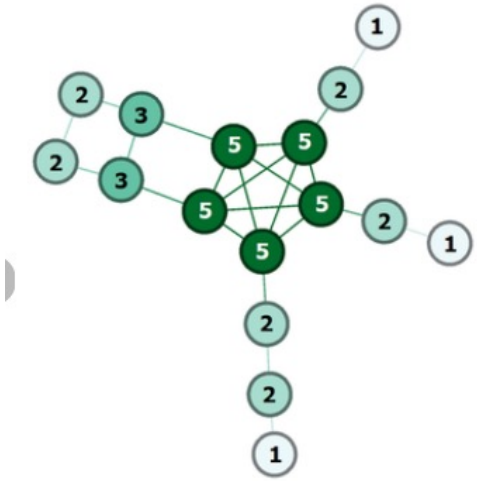
Newman et al., 2024

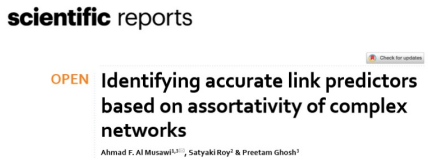
DEGREE ASSORTATIVITY

- Tells you whether nodes connect to other nodes with the same degree
  - Are high degree nodes attached to other high degree nodes?
- A complete network has a degree assortativity of 1.0

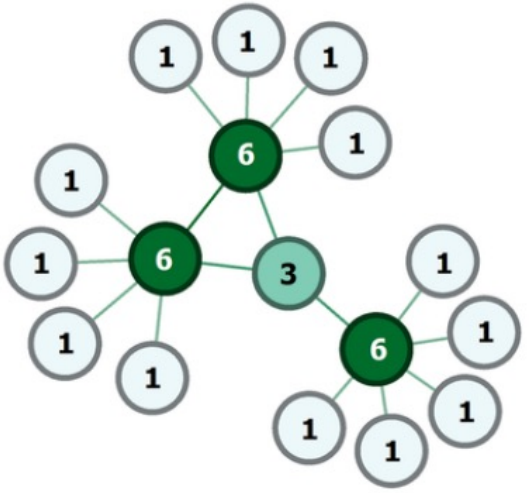
Question: Which of these two networks have a high assortativity?

Assortative





Disassortative



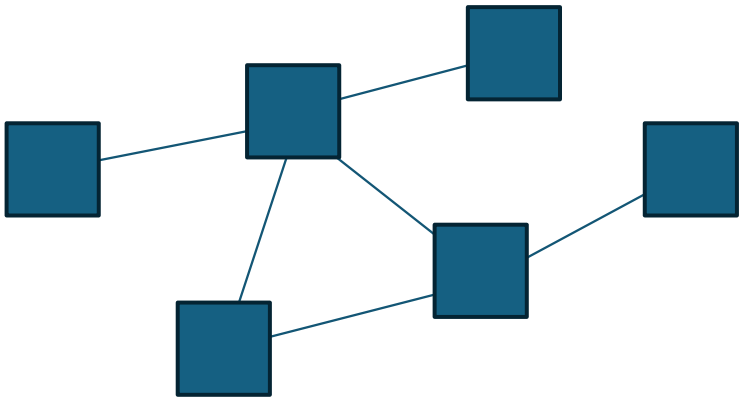
DENSITY

- Tells you how "full" the network is of edges by comparing the number of edges to the number of possible edges

$$D_{net} = \begin{cases} \frac{2E}{N*(N-1)}, \text{if undirected} \\ \frac{E}{N*(N-1)}, \text{directed} \end{cases}$$

Questions:

1. What would the density be of a complete network?
2. What is the density of the below network?



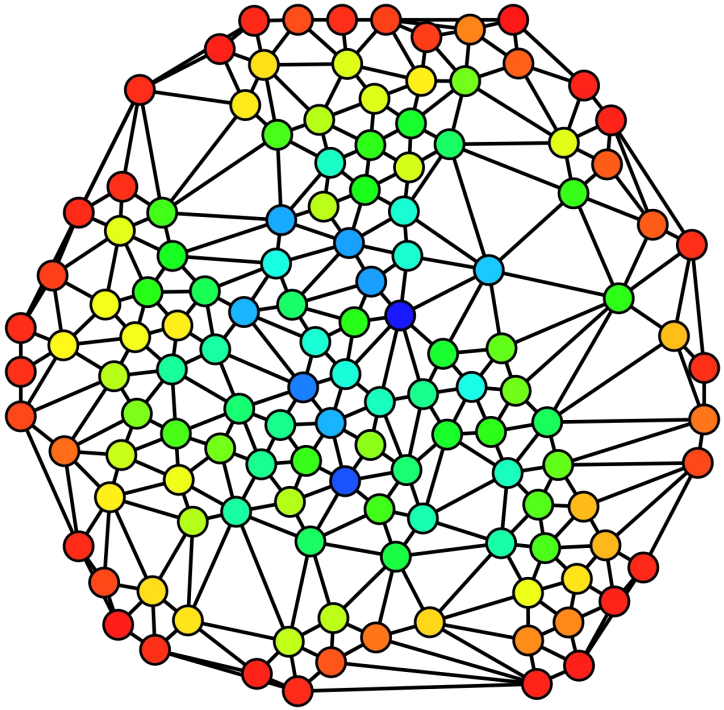
D = 2(6) / 6(5) = 12/30 = 0.4

BETWEENNESS CENTRALITY

- A measurement of how “central” a node is in a network
- Based on the shortest path calculation between **all** nodes in the network

Number of shortest paths between s and t that pass through node v

Number of shortest paths between s and t

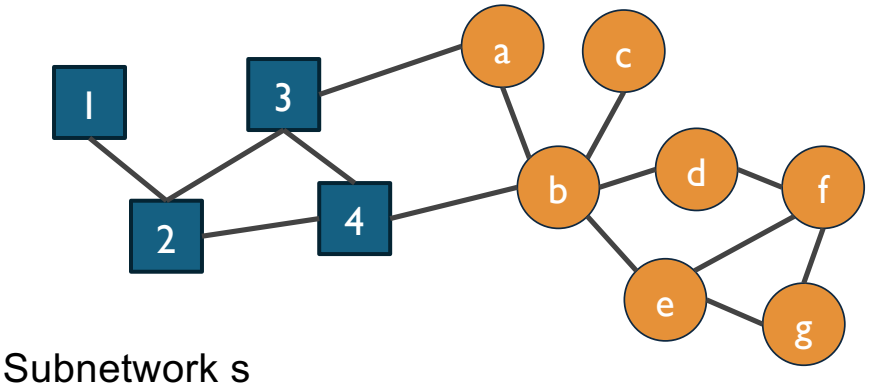


https://en.wikipedia.org/wiki/Betweenness_centrality

BIPARTITE BETWEENNESS CENTRALITY

- Not actually a metric of bipartite networks (oops)
- Similar to BC, but measures the bottleneck-ness of a node between two distinct regions of a network

Question: Which of these nodes do you think will have the largest BiBC? Which ones would have the lowest?

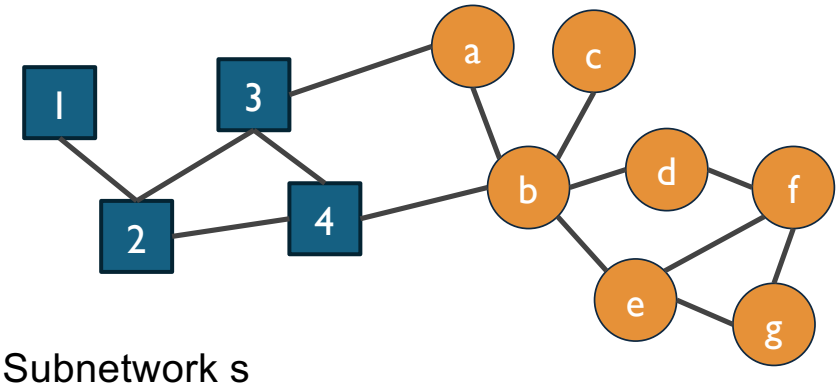


Subnetwork s

Subnetwork t

BIPARTITE BETWEENNESS CENTRALITY

EXAMPLE CALCULATION



Subnetwork s

Subnetwork t

Associated shortest paths that have node 3 within them:

1 - a (3 steps)

2 - a (2 steps)

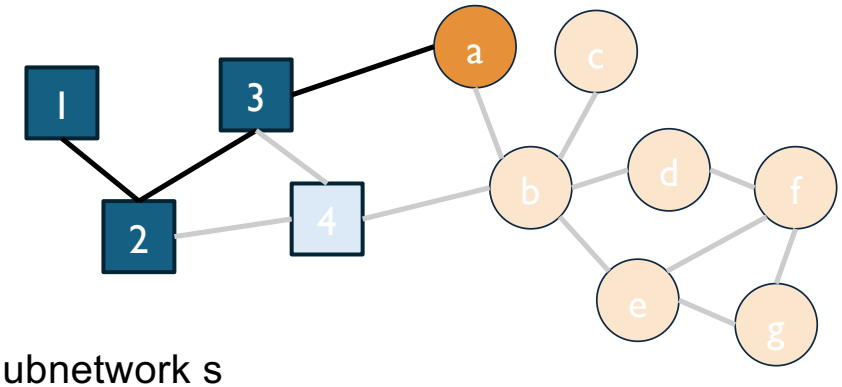
4 - a (2 steps)

(There are many other pairs of nodes that technically get calculated, but they do not have node 3 in them and thus will not contribute to the BiBC value)

Note: the shortest path cannot have the node of interest (v) on the starting or ending node in the path between node (s) or (t)

So it must be in this order (s)→(v)→(t)

BIPARTITE BETWEENNESS CENTRALITY



Subnetwork s

Subnetwork t

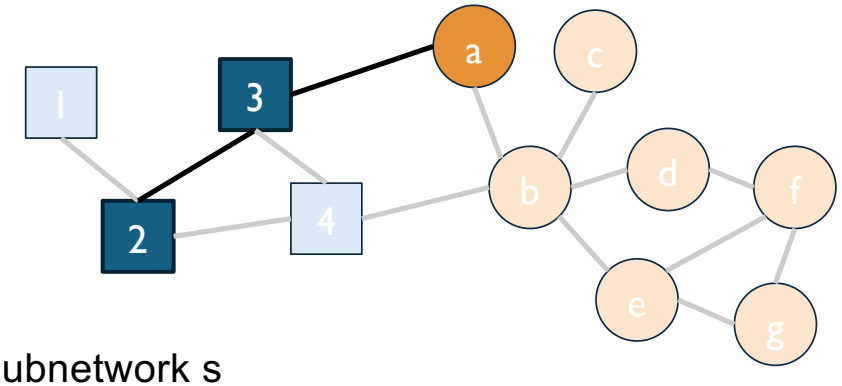
Associated shortest paths:

1 - a: 1/1 = 1

1 path passes through node 3

1 total shortest path

BIPARTITE BETWEENNESS CENTRALITY



Subnetwork s

Subnetwork t

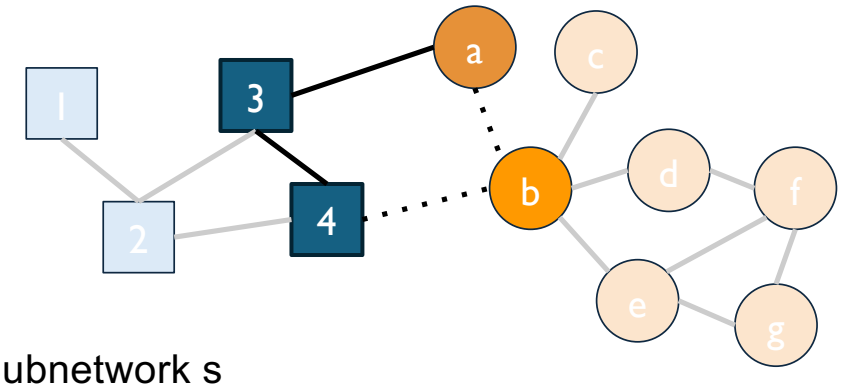
Associated shortest paths:

2 - a: 1/1 = 1

1 path passes through node 3

1 total shortest path

BIPARTITE BETWEENNESS CENTRALITY



Subnetwork s

Subnetwork t

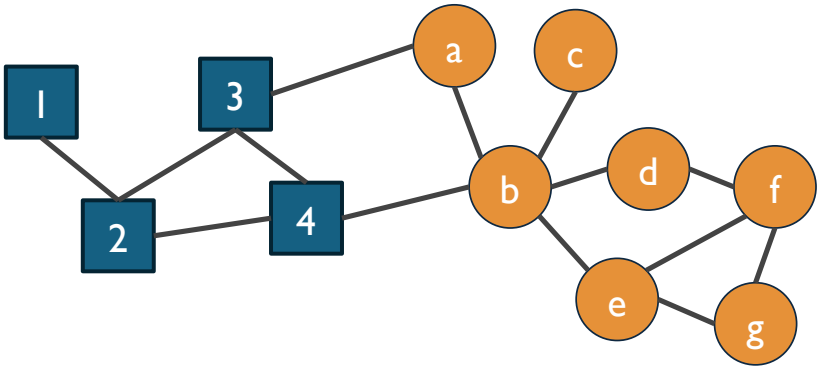
Associated shortest paths:

4 - a: 1/2 = .5

1 path passes through node 3

2 total shortest paths

BIPARTITE BETWEENNESS CENTRALITY



Subnetwork s

Subnetwork t

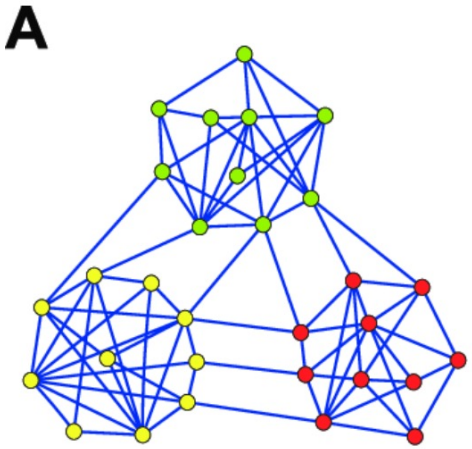
Sum:

1/1 + 1/1 + 1/2 +0/1 … +0/1 = **2.5**

Node 3 has a BiBC of 2.5

MODULARITY

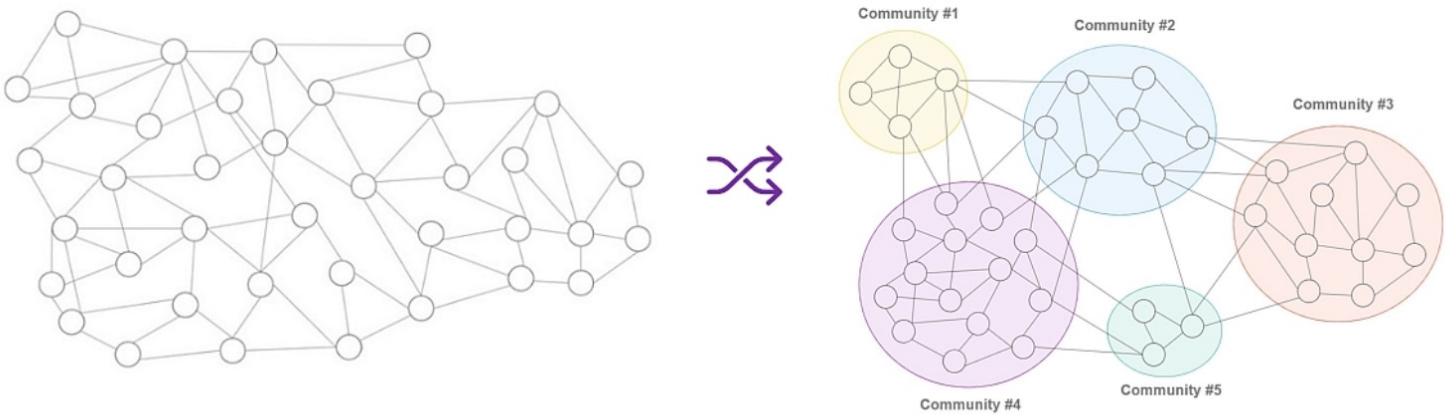
- How well the graph can be **partitioned** into smaller components
- A graph with high modularity means there are regions of the graph not densely connected to other regions
- The Louvain method is used to determine a network's modularity



https://www.researchgate.net/figure/An-Example-of-a-Small-Network-with-a-Modular-Structure-A-and-Its-Randomly-Rewired_fig1_25664666

COMMUNITY DETECTION ALGORITHMS

- Many types of community detection algorithms
  - Louvain/Leiden
  - Infomap
  - Girvan-Newman
  - And more!



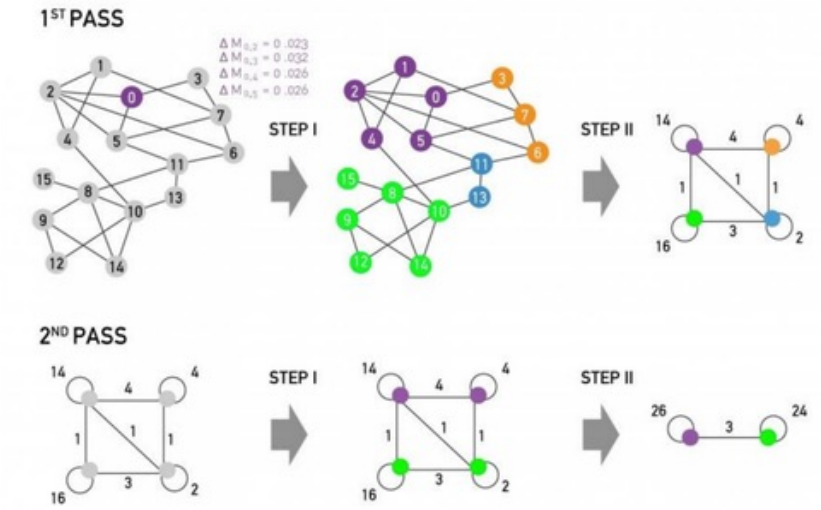
Network of Connected Entities

Entities clustered into communities

https://timbr.ai/community-detection-algorithm/

COMMUNITY DETECTION WITH LOUVAIN

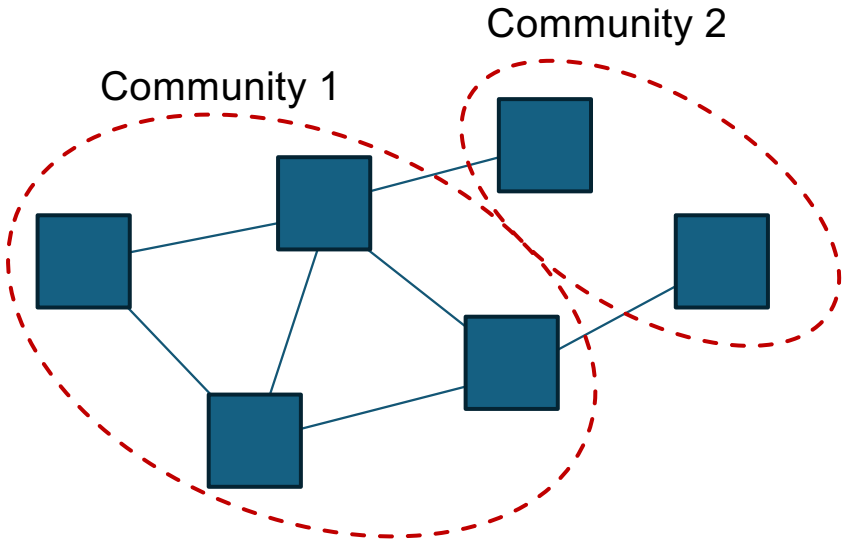
- Uses a process called modularity maximization
- Two-step, iterative process
  1. Assign every node its own community
  2. Loop through each node (i) and calculate the change in modularity (Q) if you add the node to a neighbor community
  3. Add node i to the community that most increased the modularity
  4. Aggregate the subnetworks into a single node and repeat until you can no longer achieve better modularity



https://wikidocs.net/93414

COMMUNITY DETECTION WITH LOUVAIN

- One major issue with Louvain is that it can create subnetworks of nodes that are not connected.
- The Leiden algorithm was later introduced to account for this, where it introduces an additional refinement step to ensure intra-community connectedness
  - Still utilizes the concept of modularity maximization, however



COMMUNITY DETECTION WITH INFOMAP

Eur. Phys. J. Special Topics **178**, 13–23 (2009)
© EDP Sciences, Springer-Verlag 2010
DOI: 10.1140/epjst/e2010-01179-1

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

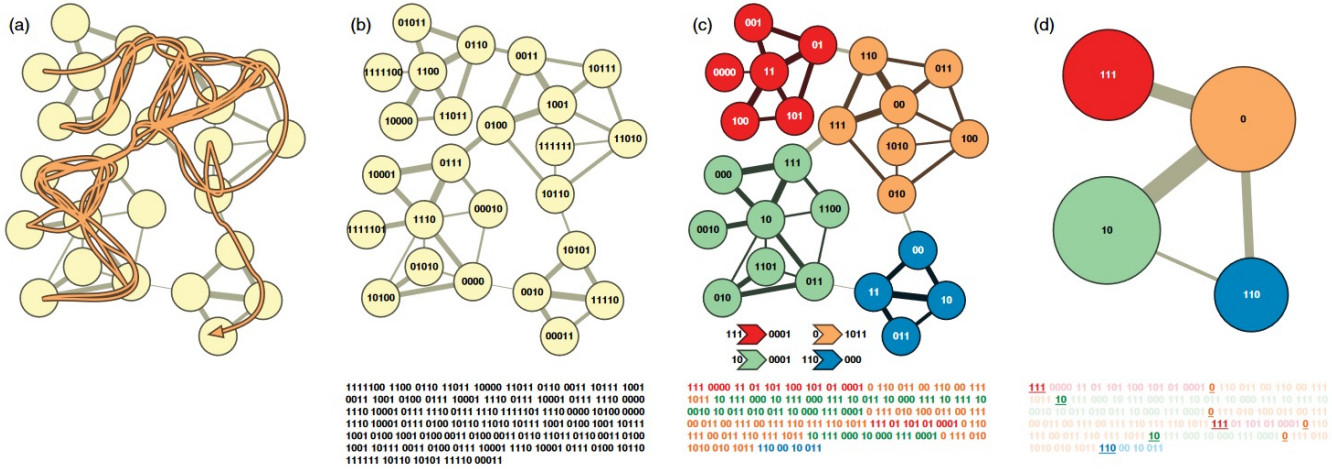
Regular Article

The map equation

M. Rosvall[1,a], D. Axelsson[1], and C.T. Bergstrom[2]

[1] Department of Physics, Umeå University, 901 87 Umeå, Sweden
[2] Department of Biology, University of Washington, Seattle, WA 98195-1800, USA

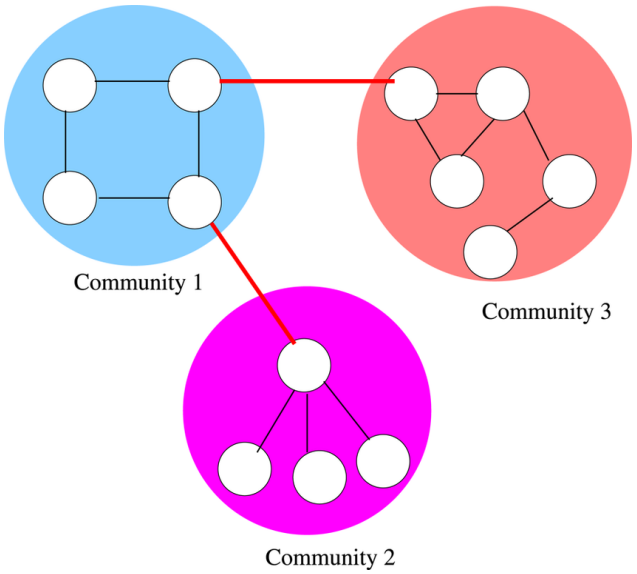
- Relies on random walks
- Codes nodes in what is called “Huffman code”, where short codewords correspond to common events (randomly walking on a node) and long codewords are uncommon, then
- Since no codes are prefixes of other codes, we can simply use a string of 0’s and 1’s to notate the exact path of the walker
- Then condense the map into regions where the walker spent the most time



COMMUNITY DETECTION WITH GIRVAN-NEWMAN

- Uses betweenness centrality to select edges that most separate regions of the network

1. Calculate betweenness of all edges
2. Remove edge with highest betweenness
3. Re-calculate the betweenness and remove edges again, then repeat until there are no remaining edges

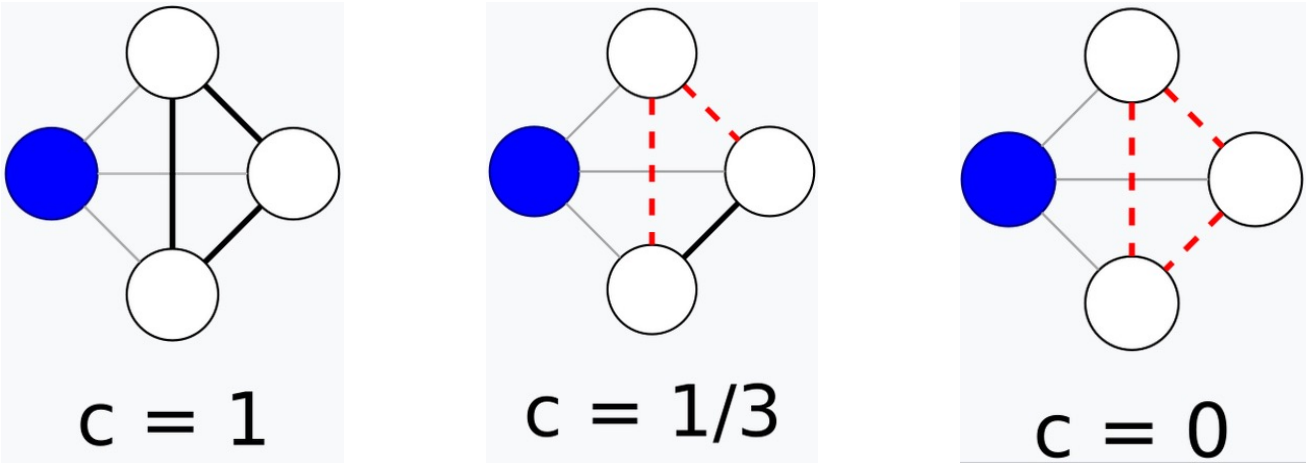


https://www.researchgate.net/figure/Illustrates-the-community-detection-process-of-the-Girvan-Newman-algorithm-The-two-red_fig3_371098970

Question: There is one major drawback to using the Girvan-Newman method. What do you think it is?

CLUSTERING COEFFICIENT

- Explains the degree to which the neighbors of a node are connected



https://en.wikipedia.org/wiki/Clustering_coefficient

CLOSENESS CENTRALITY

- How close a single node is to every other node in the network

Number of nodes in the graph

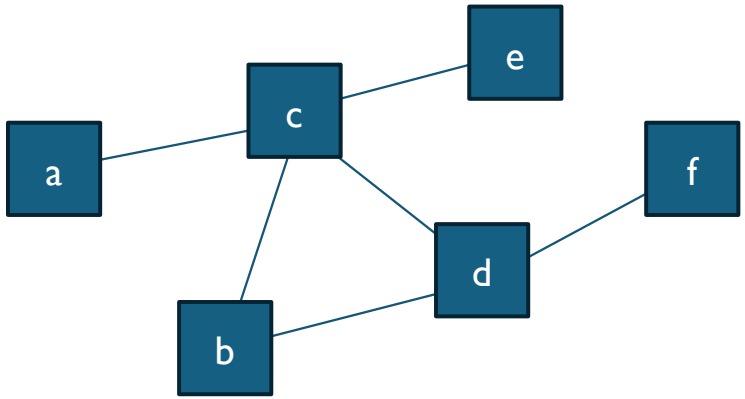
$$C(x) = \frac{N-1}{\sum_y d(y,x)}$$

"Starting" node

Distance between starting node and node y

Questions:

1. What is the closeness value of node b?
2. Would you expect node f to have a larger or smaller closeness?



C(b) = (6-1) / (2+1+1+2+2) = 5/8 = 0.625

CYTOSCAPE FOR NETWORK VISUALIZATION AND ANALYSIS

- Cytoscape is a Java-based application that allows you to visualize networks and perform basic analyses
- Requires a data file and a metadata file
  - Data file:
    - Two columns, one with source node and one with target
    - Also allows for additional columns with edge weights
  - Metadata file: maps your node names to classes as well as node properties

CYTOSCAPE LEARNING OBJECTIVES

1. Import data and metadata
2. Modify node and edge visualizations
3. Change the layout of your network
4. Create a subset of a network
5. Analyze a network using Cytoscape

OVERVIEW

How can we model complex diseases?

Using TkNA to identify potential drivers of disease

Gene regulation in disease

Identifying novel cancer subtypes using PANDA and LIONESS

THE TKNA SOFTWARE ALLOWS US TO MODEL AND UNDERSTAND COMPLEX DISEASES

- TkNA: **T**rans**k**ingdom **N**etwork **A**nalysis
  - Newman et al., 2024

Oregon State University

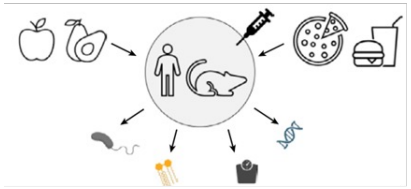
Protocol | Published: 12 March 2024

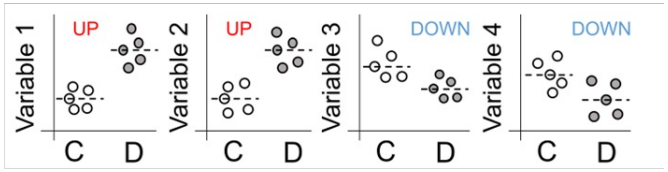
Transkingdom Network Analysis (TkNA): a systems framework for inferring causal factors underlying host–microbiota and other multi-omic interactions

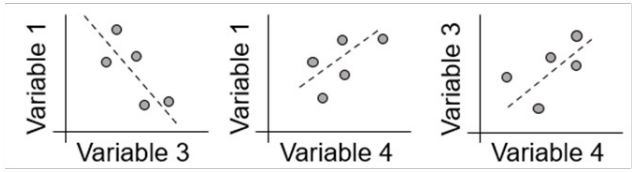
Nolan K. Newman, Matthew S. Macovsky, Richard R. Rodrigues, Amanda M. Bruce, Jacob W. Pederson, Jyothi Padiadpu, Jigui Shan, Joshua Williams, Sankalp S. Patil, Amiran K. Dzutsev, Natalia Shulzhenko, Giorgio Trinchieri ✉, Kevin Brown ✉ & Andrey Morgun ✉

Nature Protocols **19**, 1750–1778 (2024) | Cite this article

NETWORK RECONSTRUCTION OVERVIEW

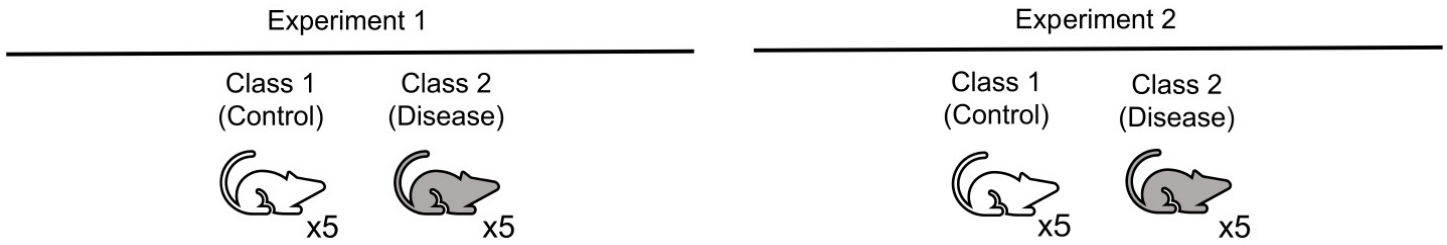




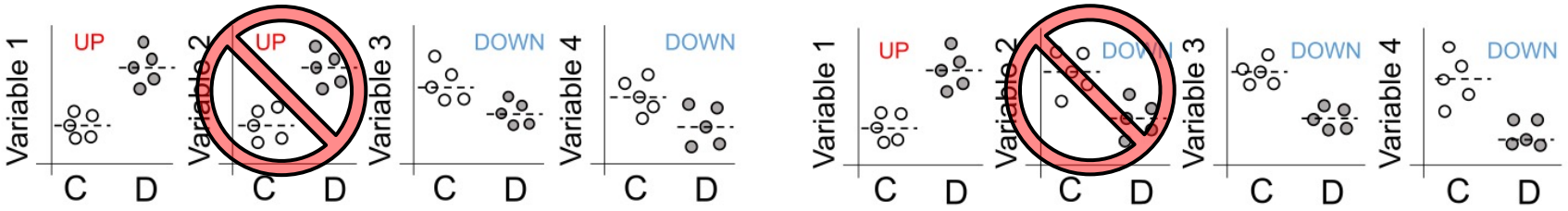


UNDERSTANDING THE STATISTICAL CONSIDERATIONS OF TKNA

Experimental design



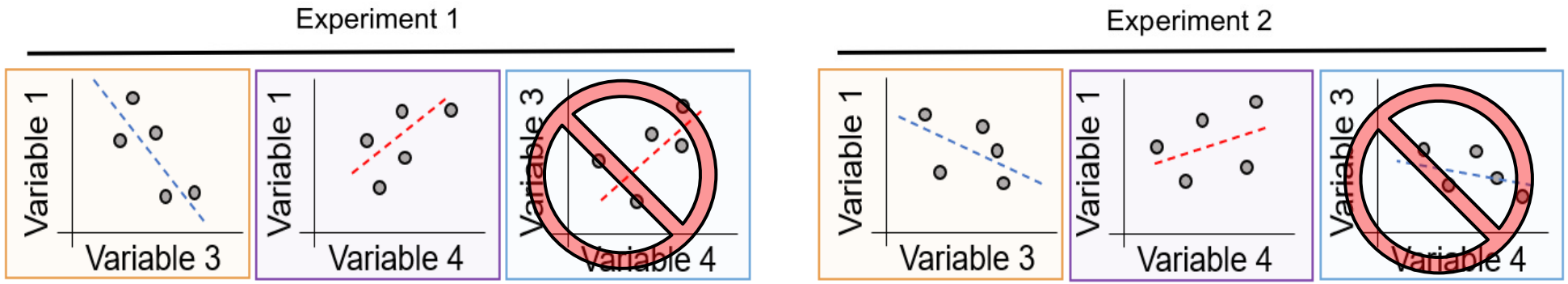
Identification of robust and reproducible variables across experiments



	p-value Expt1	p-value Expt2	combined p-value	FDR	log2 fold change Expt 1	log2 fold change Expt 2	Consistent fold change direction?
Variable 1	0.02	0.03	0.01	0.01	2.32	2.10	Yes ✓
Variable 2	0.09	0.04	0.02	0.02	2.21	-0.80	No ✗
Variable 3	0.07	0.01	0.01	0.01	-0.91	-1.38	Yes ✓
Variable 4	0.12	0.03	0.02	0.03	-0.54	-1.21	Yes ✓

UNDERSTANDING THE STATISTICAL CONSIDERATIONS OF TKNA

Identification of robust and reproducible correlations across experiments

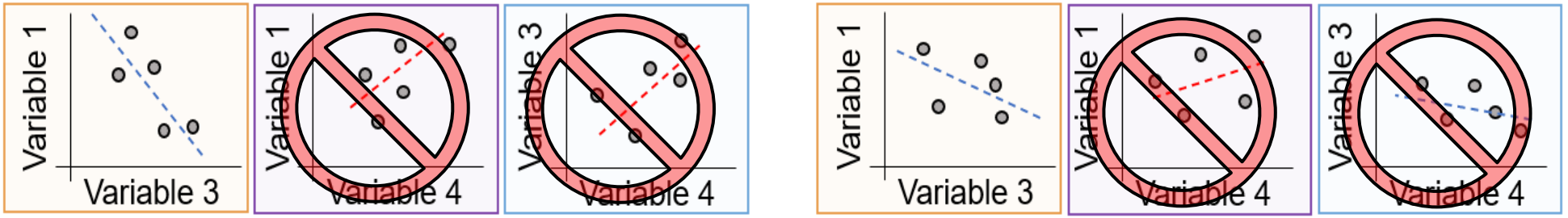


	p-value Expt1	p-value Expt2	rho coef. Expt 1	rho coef. Expt 2	combined p-value	FDR	Consistent correlation direction?
Variable 1 - Variable 3	0.01	0.04	-0.78	-0.45	0.0035	0.005	Yes ✓
Variable 1 - Variable 4	0.04	0.14	0.55	0.23	0.0346	0.005	Yes ✓
Variable 3 - Variable 4	0.12	0.18	0.73	-0.45	0.1044	0.08	No ✗
etc.	…	…	…	…	…	…	…

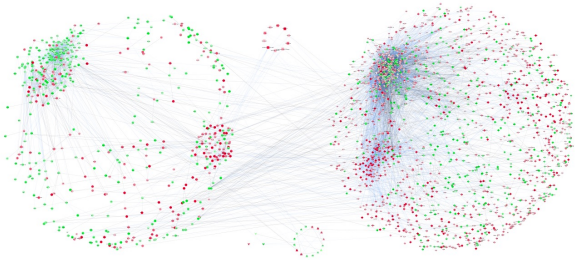
Accounts for Simpson's paradox as well

UNDERSTANDING THE STATISTICAL CONSIDERATIONS OF TKNA

Fold change: UP - DOWN UP - DOWN DOWN - DOWN UP - DOWN UP - DOWN DOWN - DOWN

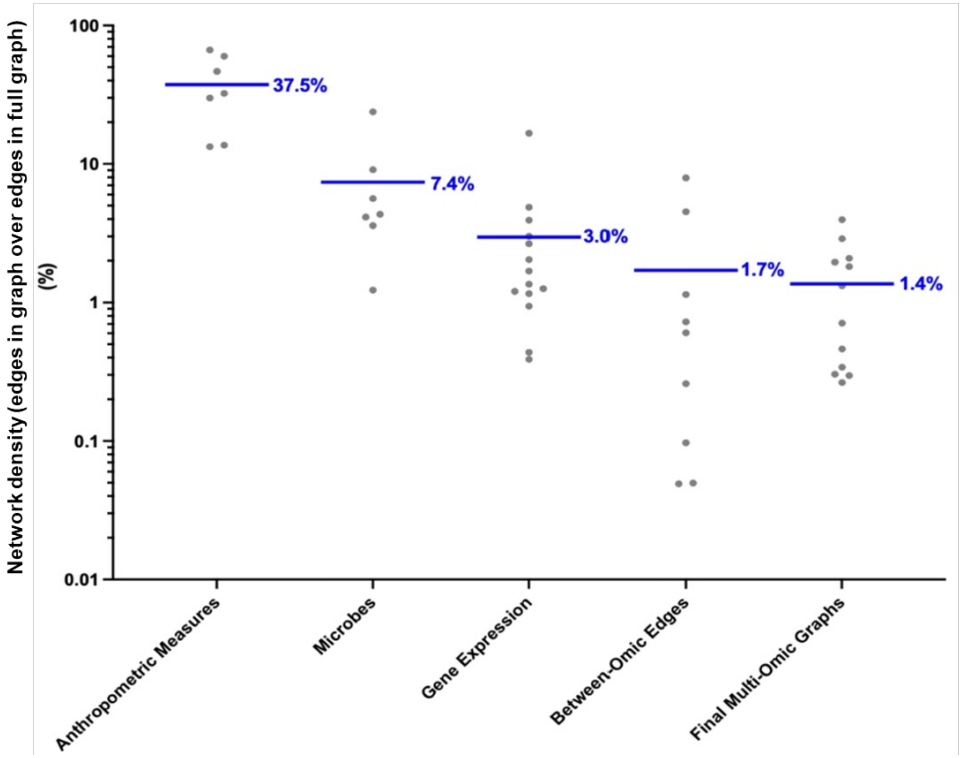


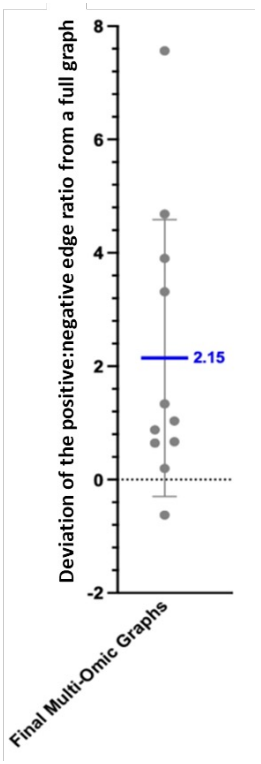
	p-value Expt1	p-value Expt2	rho coef. Expt 1	rho coef. Expt 2	combined p-value	FDR	Consistent correlation direction?	Expected correlation?
Variable 1 - Variable 3	0.01	0.04	-0.78	-0.45	0.0035	0.005	Yes ✓	Yes ✓
Variable 1 - Variable 4	0.04	0.14	0.55	0.23	0.0346	0.005	Yes ✓	No ✗
Variable 3 - Variable 4	0.12	0.18	0.73	-0.45	0.1044	0.08	No ✗	NA
etc.	…	…	…	…	…	…	…	…



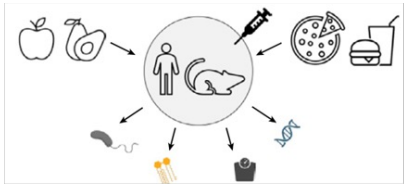
How do we know if this network is sufficient for downstream analysis?

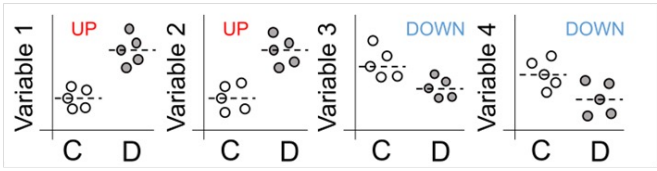
IS YOUR NETWORK SUFFICIENT FOR DOWNSTREAM ANALYSIS?

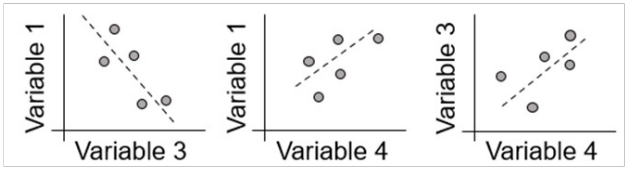


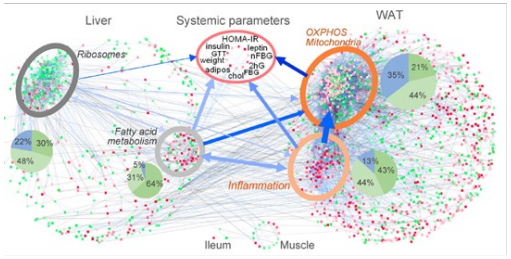


NETWORK RECONSTRUCTION OVERVIEW



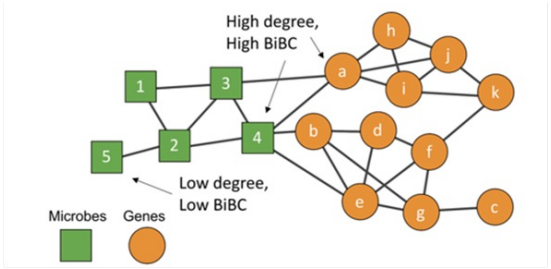


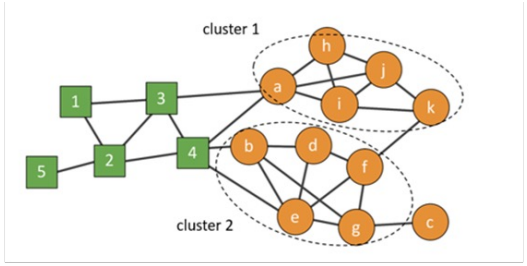


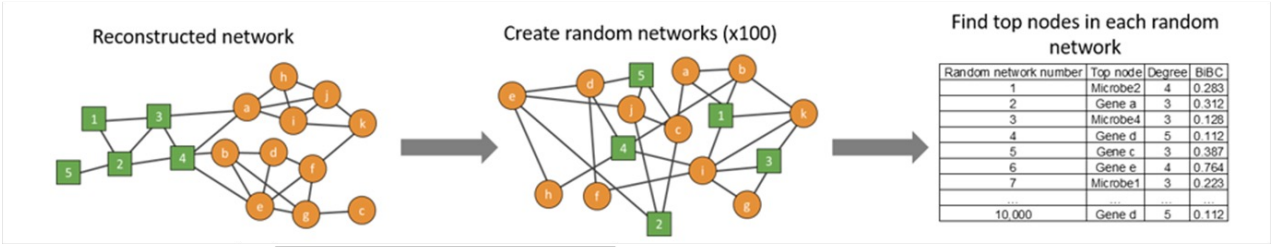


Adapted from Newman et al., 2024

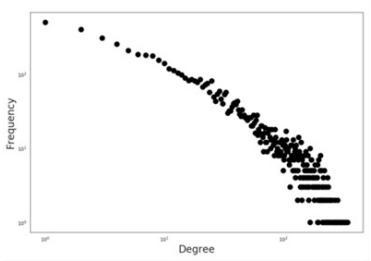
NETWORK RECONSTRUCTION OVERVIEW

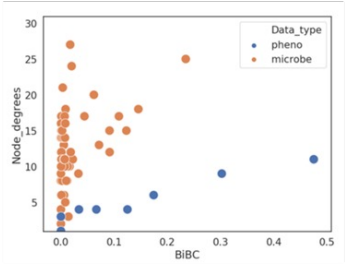


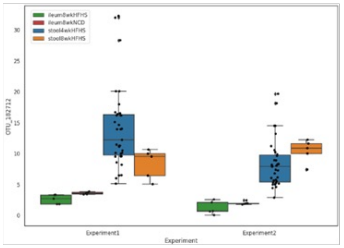


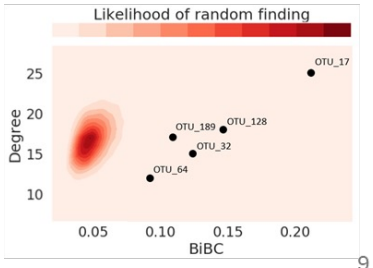


Random network number	Top node	Degree	BiBC
1	Microbe2	4	0.283
2	Gene a	3	0.312
3	Microbe4	3	0.128
4	Gene d	5	0.112
5	Gene c	3	0.387
6	Gene e	4	0.764
7	Microbe1	3	0.223
10,000	Gene d	5	0.112









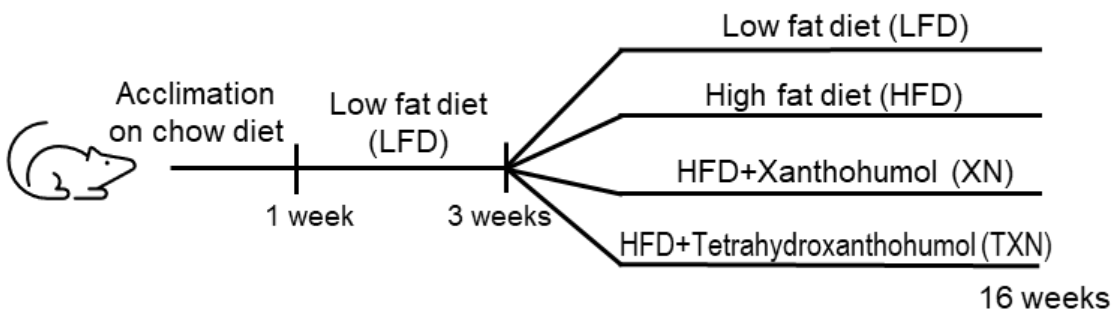
TKNA IN ACTION

Reversing gut microbiome-driven adipose tissue inflammation alleviates metabolic syndrome

N. K. Newman, Y. Zhang, J. Padiadpu, C. L. Miranda, A. A. Magana, C.P.Wong, K. A. Hioki, J.W. Pederson, Z. Li, M. Gurung, A. M. Bruce, K Brown, G. Bobe, T. J. Sharpton, N. Shulzhenko, C. S. Maier, J. F. Stevens, A. F. Gombart, A. Morgun

doi: https://doi.org/10.1101/2022.10.28.514267

Experimental question: Why do mice that are treated with Tetrahydroxanthohumol experience better phenotypic outcomes than those treated with Xanthohumol?



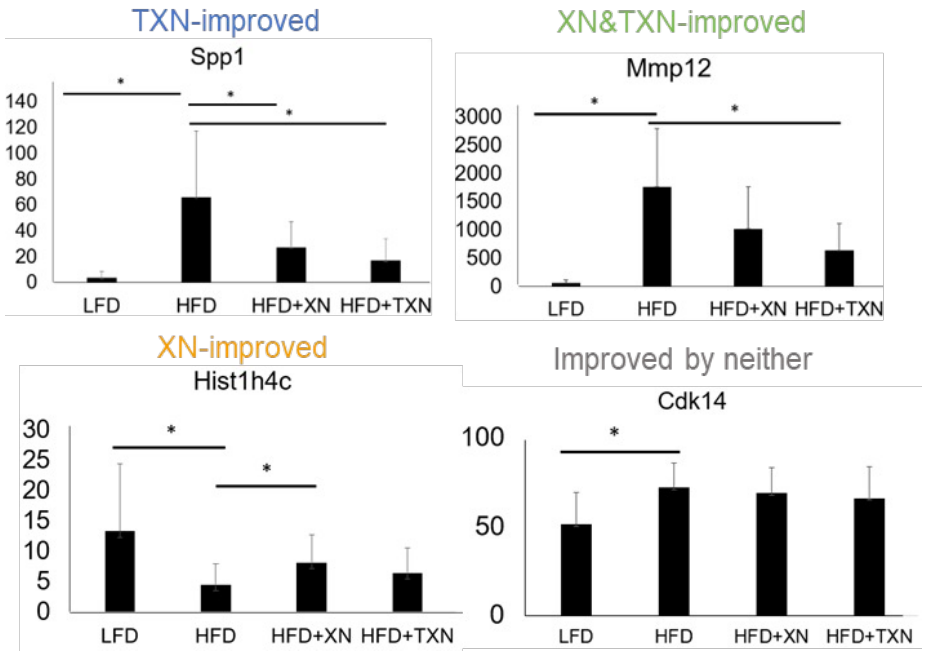
Measured parameters:

RNAseq:
- Liver
- Ileum
- White Adipose Tissue

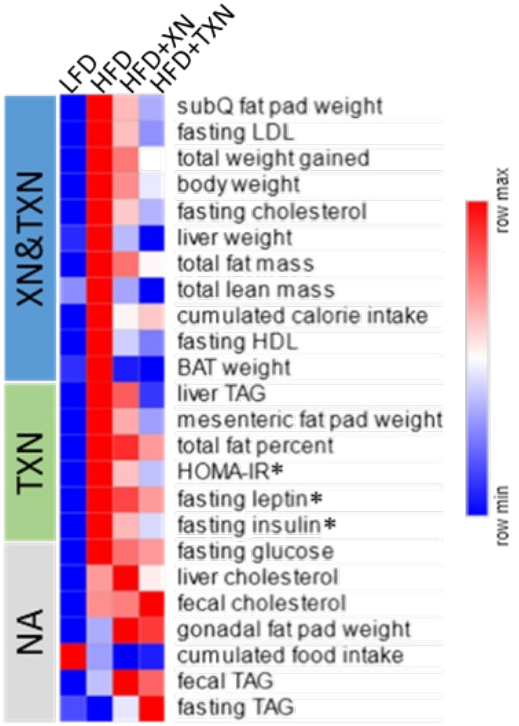
16S rRNA from stool
Bile acid concentration
Metabolic parameters

TKNA IN ACTION

Categorized features into four categories

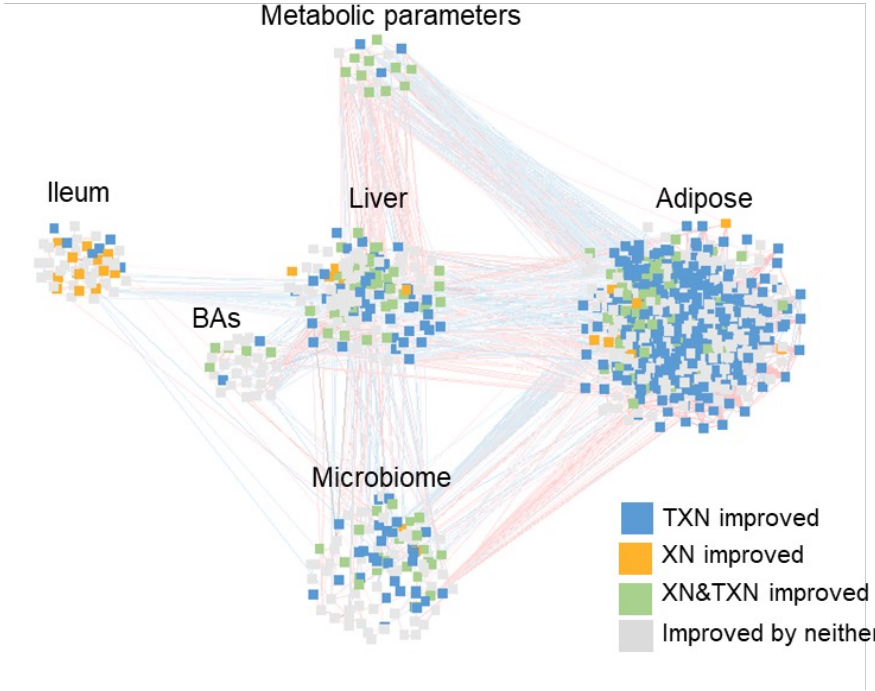


TXN reverses the levels of many phenotypic features related to metabolic disease

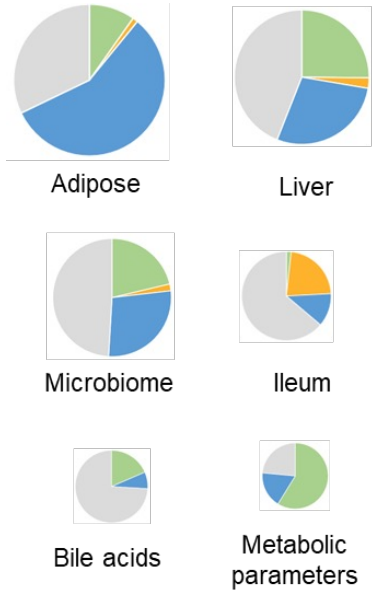


TKNA IN ACTION

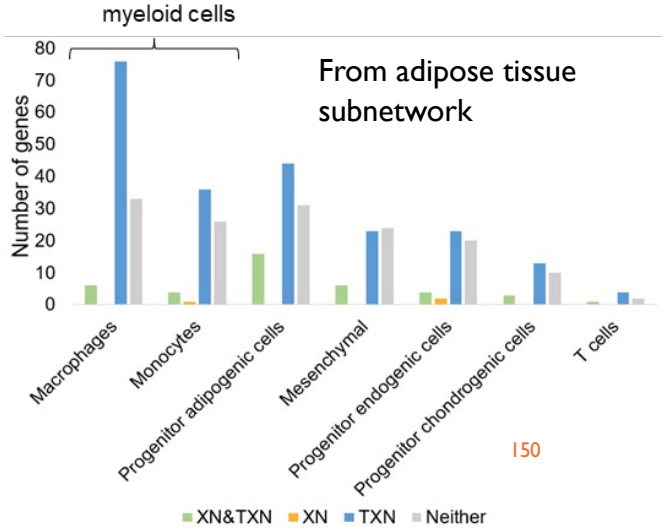
TkNA is used to reconstruct a co-variation network



Many parameters are improved by TXN treatment

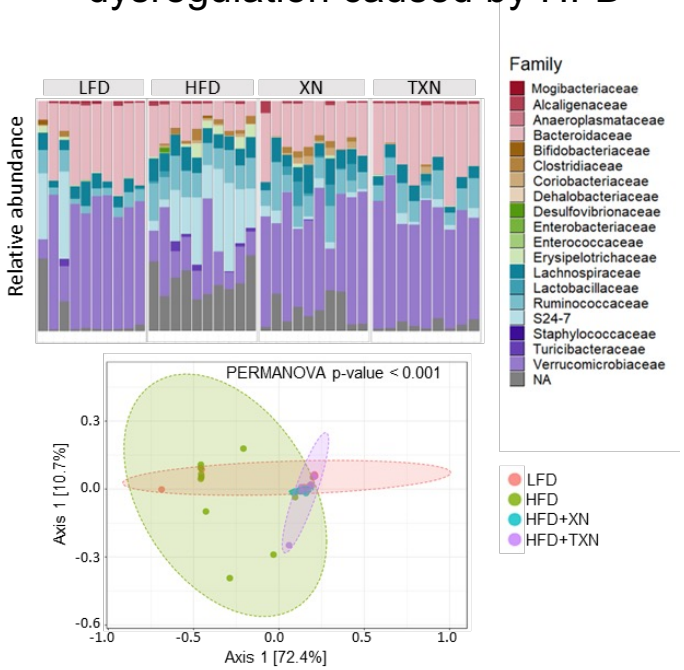


TXN primarily acts on macrophages and monocytes to ameliorate HFD-induced disease



TKNA IN ACTION

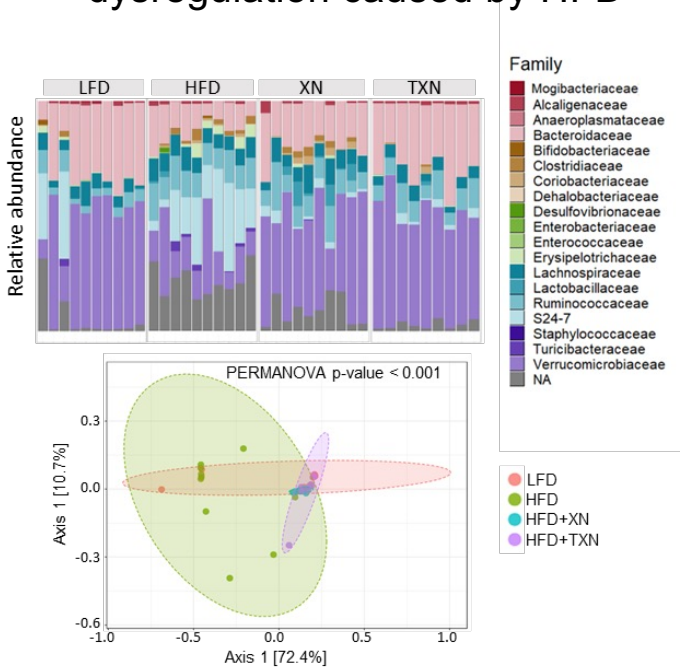
XN and TXN ameliorate dysregulation caused by HFD



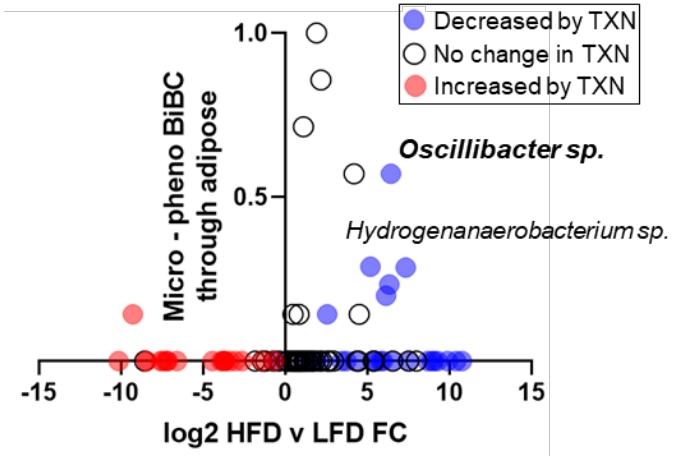
Question: At this point, we wanted to use network biology to find which specific microbes (ASVs) influenced the gene expression in the adipose tissue. Which network method(s) that we discussed earlier could we use at this step?

TKNA IN ACTION

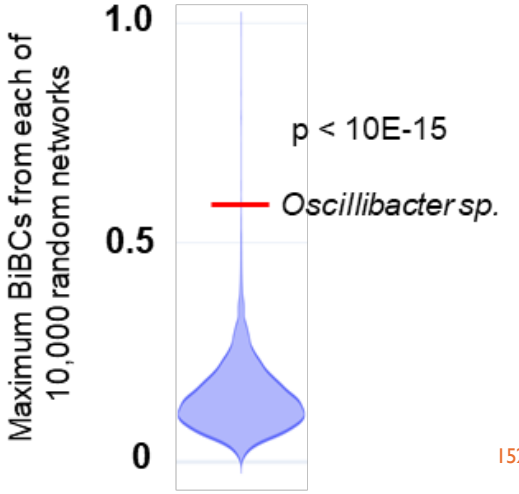
XN and TXN ameliorate dysregulation caused by HFD



According to TkNA, a microbe of Oscillibacter sp. regulates many of the dysregulated genes found in the adipose



Random network analysis via the TkNA pipeline confirms the microbe was likely not found due to random chance

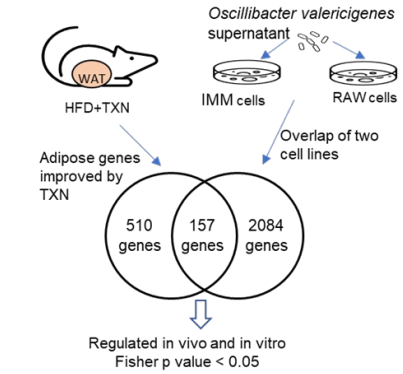


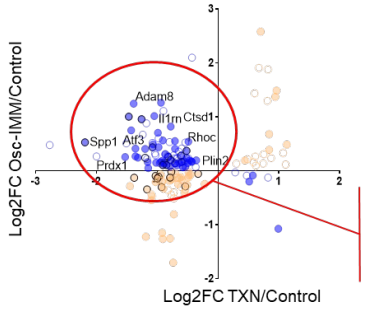
TKNA IN ACTION

Question: At this point, we had predicted the potentially regulatory microbe based on our network model. However, now we need to experimentally validate this result. Can you think of a method we could use to validate this prediction?

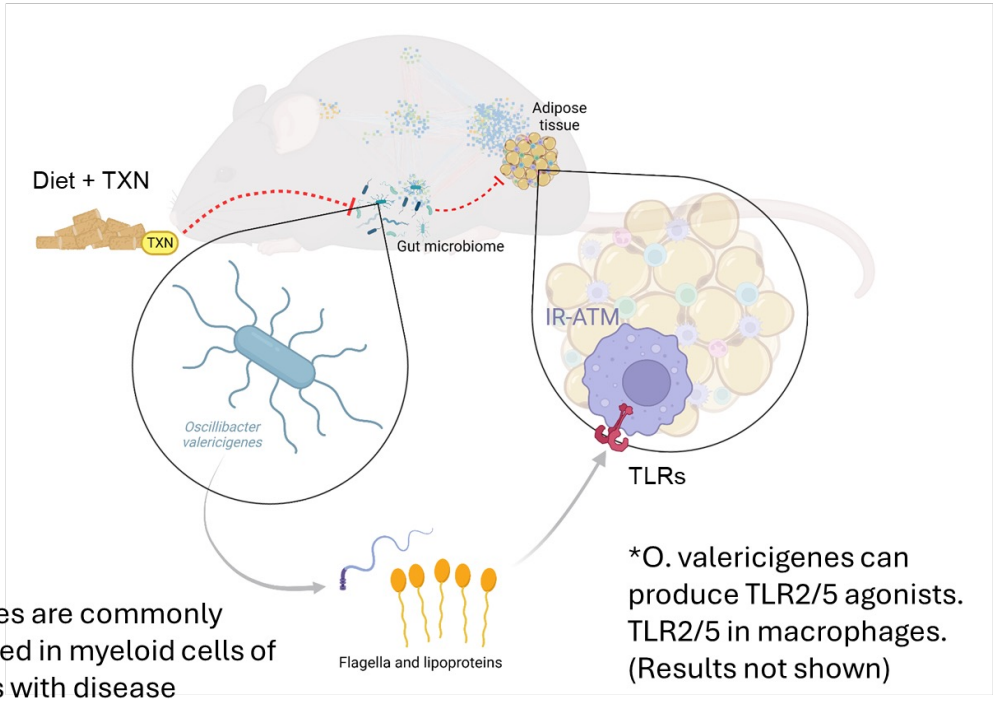
TKNA IN ACTION

In vivo validation finds *O. valericigenes* supernatant induces many inflammatory related genes, which TXN ameliorates





These genes are commonly dysregulated in myeloid cells of individuals with disease

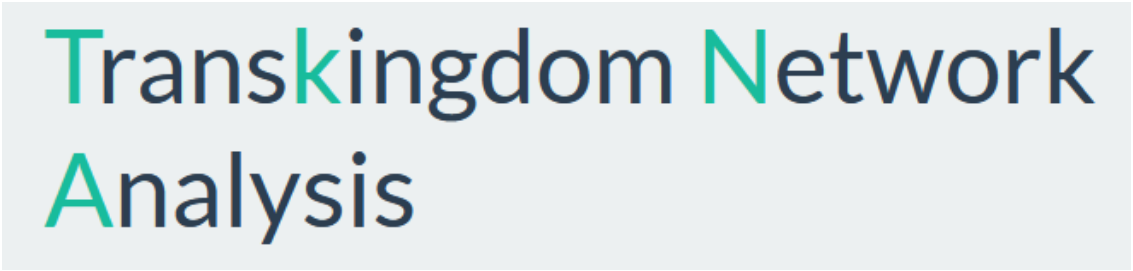


*O. valericigenes can produce TLR2/5 agonists. TLR2/5 in macrophages. (Results not shown)

THE TKNA WEB INTERFACE

- Now YOU get to use TkNA!
- The web interface greatly simplifies the process of using TkNA
- Available at https://bioinfo-abcc.ncifcrf.gov/TkNA/

***Full disclaimer: I am the co-author of TkNA, so I am a bit biased towards it ☺ There are other good network analysis tools out there as well!



TKNA LEARNING OBJECTIVES

Available at https://bioinfo-abcc.ncifcrf.gov/TkNA/

1. Familiarize yourself with the format of files TkNA uses
2. Successfully use the TkNA web interface
3. Compare your resulting network when you modify the statistical thresholds set
4. Be able to view and draw biological conclusions from your results

OVERVIEW

How can we model complex diseases?

Using TkNA to identify potential drivers of disease

Gene regulation in disease

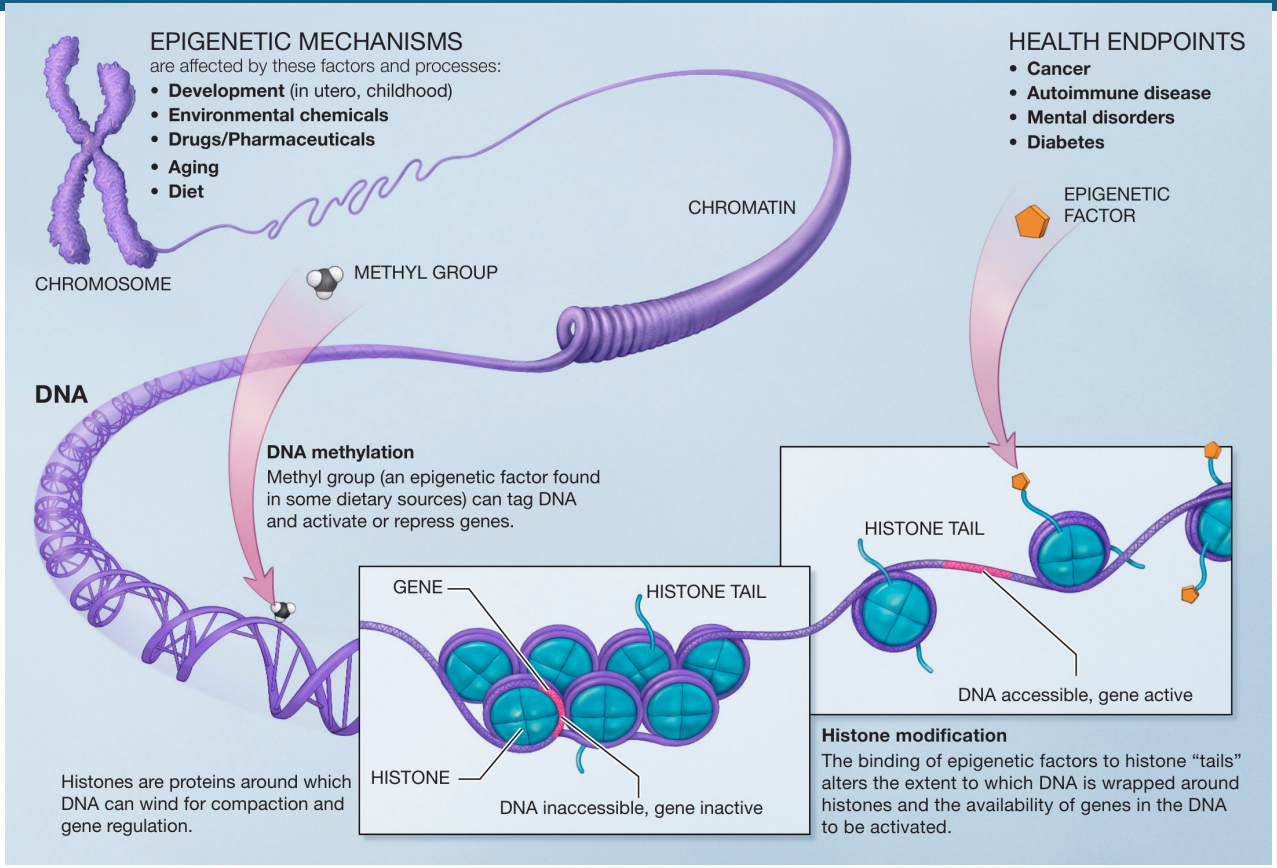
Identifying novel cancer subtypes using PANDA and LIONESS

NETWORKS MODEL OTHER BIOLOGICAL PROCESSES AS WELL

Regulatory networks model gene regulation, which occurs at multiple levels

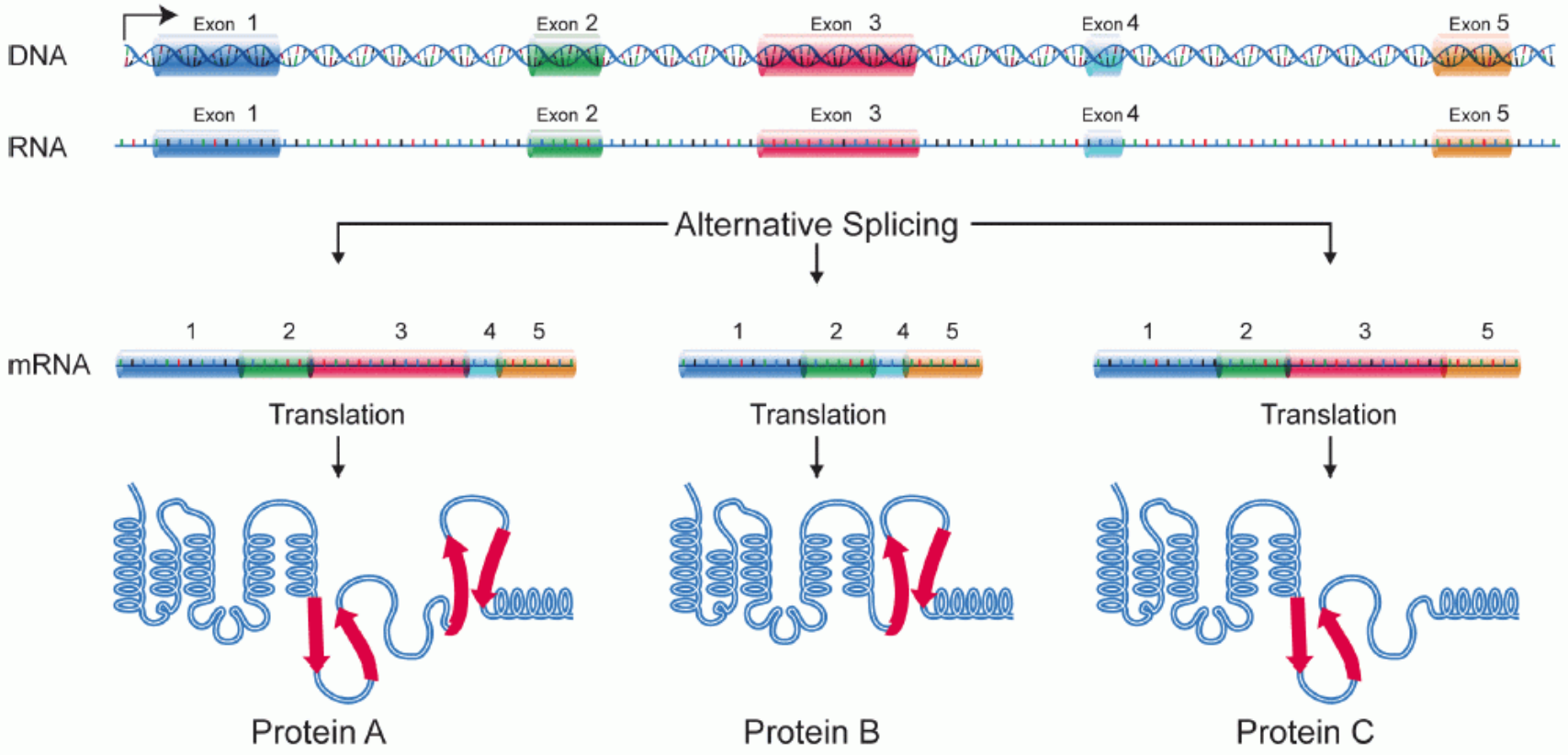
- Epigenetic level
  - Methylation
- Post-transcriptional modification
  - Alternative splicing
- Transcriptional level
  - Transcription factors

EPIGENETIC REGULATION



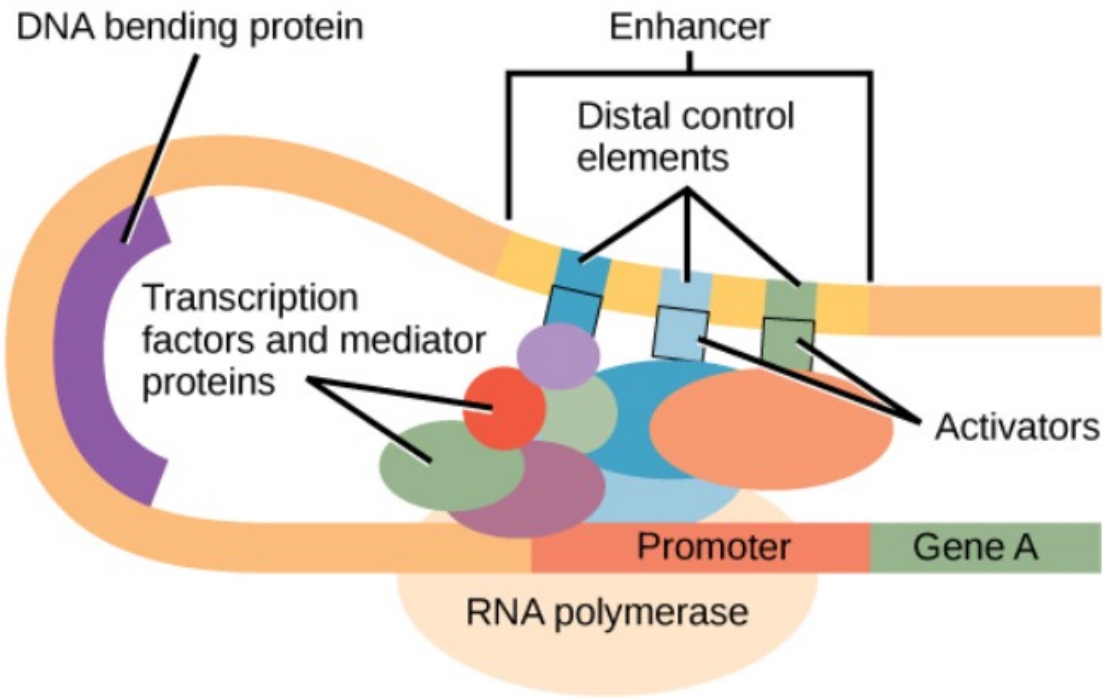
https://upload.wikimedia.org/wikipedia/commons/f/fc/Epigenetic_mechanisms.png

POST-TRANSCRIPTIONAL MODIFICATION



https://upload.wikimedia.org/wikipedia/commons/0/0a/DNA_alternative_splicing.gif

REGULATION VIA TRANSCRIPTION FACTORS



https://courses.lumenlearning.com/wm-biology1/chapter/reading-eukaryotic-transcription-gene-regulation/

OVERVIEW

How can we model complex diseases?

Using TkNA to identify potential drivers of disease

Gene regulation in disease

Identifying novel cancer subtypes using PANDA and LIONESS

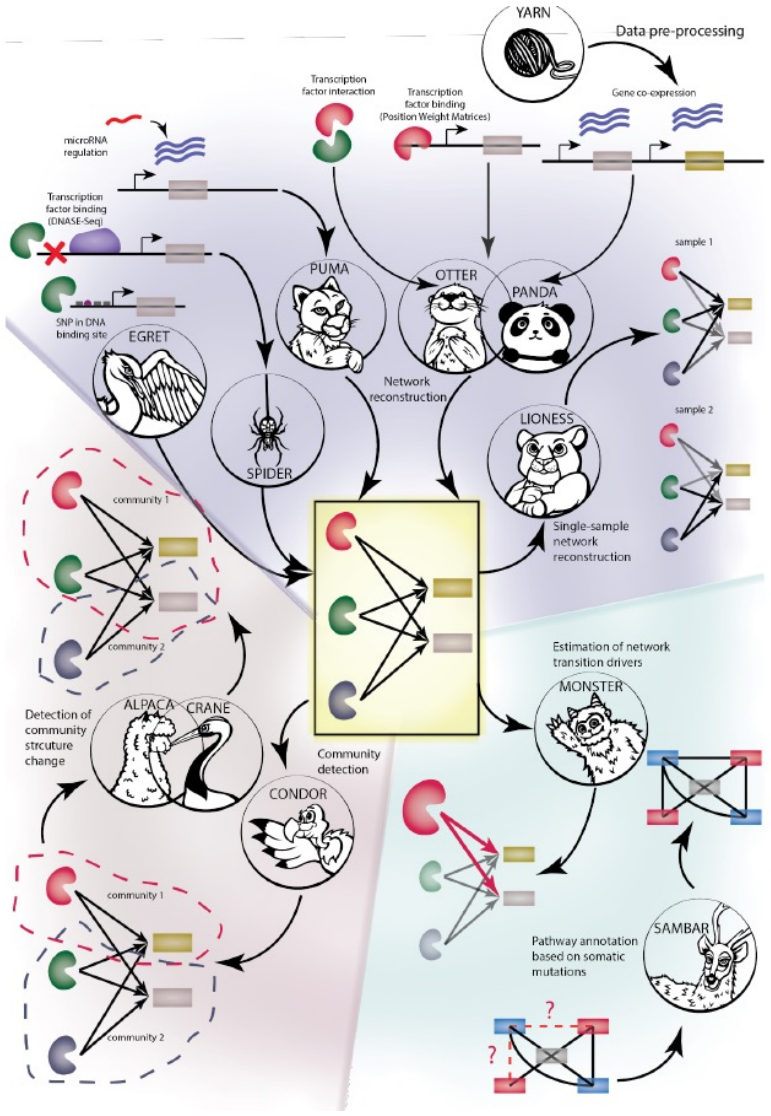
THE NETWORK ZOO

- An ecosystem of network reconstruction and analysis tools
- **PANDA** – models TF-gene regulatory interactions for a group of samples
- **LIONESS** – models TF-gene regulatory interactions for a **single sample**
- **PUMA** – models gene regulation via microRNA
- **PORCUPINE** – identifies pathways that display heterogeneous gene regulation across a patient population
- **SCORPION** – models regulatory networks on single-cell data
- **…and more!**

https://www.kuijjerlab.org/tools/

THE NETWORK ZOO

- An ecosystem of network reconstruction and analysis tools
- **PANDA** – models TF-gene regulatory interactions for a group of samples
- **LIONESS** – models TF-gene regulatory interactions for a **single sample**
- **PUMA** – models gene regulation via microRNA
- **PORCUPINE** – identifies pathways that display heterogeneous gene regulation across a patient population
- **SCORPION** – models regulatory networks on single-cell data
- **…and more!**



https://genomebiology.biomedcentral.com/articles/10.1186/s13059-023-02877-1

PANDA MODELS REGULATORY NETWORKS

PANDA: **P**assing **A**ttributes between **N**etworks for **D**ata **A**ssimilation

- Glass et al., 2013

OPEN ACCESS Freely available online

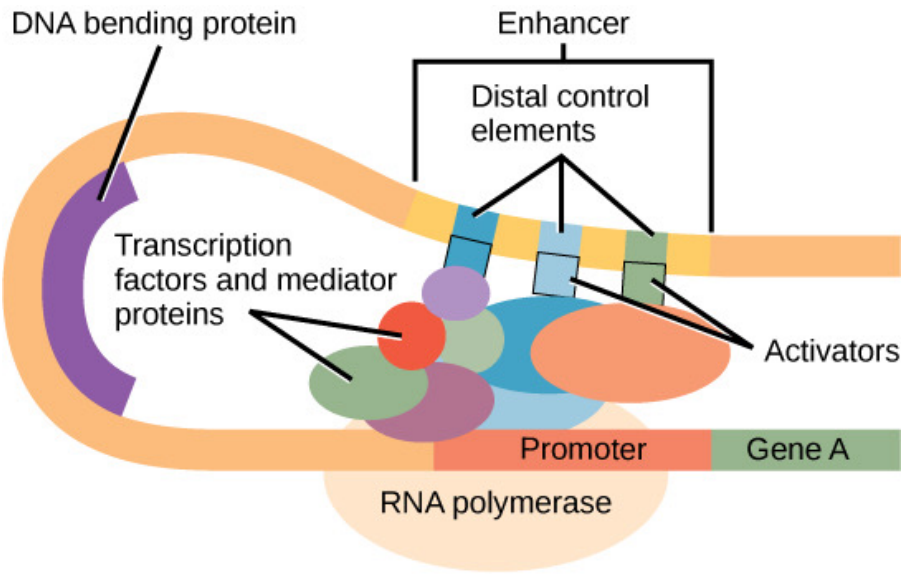
PLOS ONE

Passing Messages between Biological Networks to Refine Predicted Interactions

Kimberly Glass[1,2], Curtis Huttenhower[2], John Quackenbush[1,2], Guo-Cheng Yuan[1,2]*

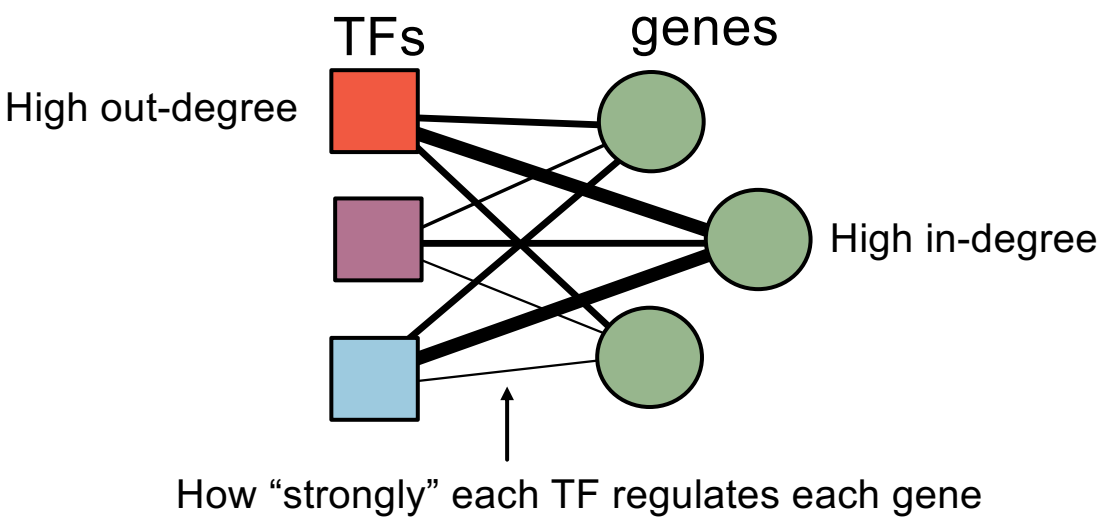
1 Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, United States of America, **2** Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, United States of America

PANDA MODELS REGULATORY NETWORKS

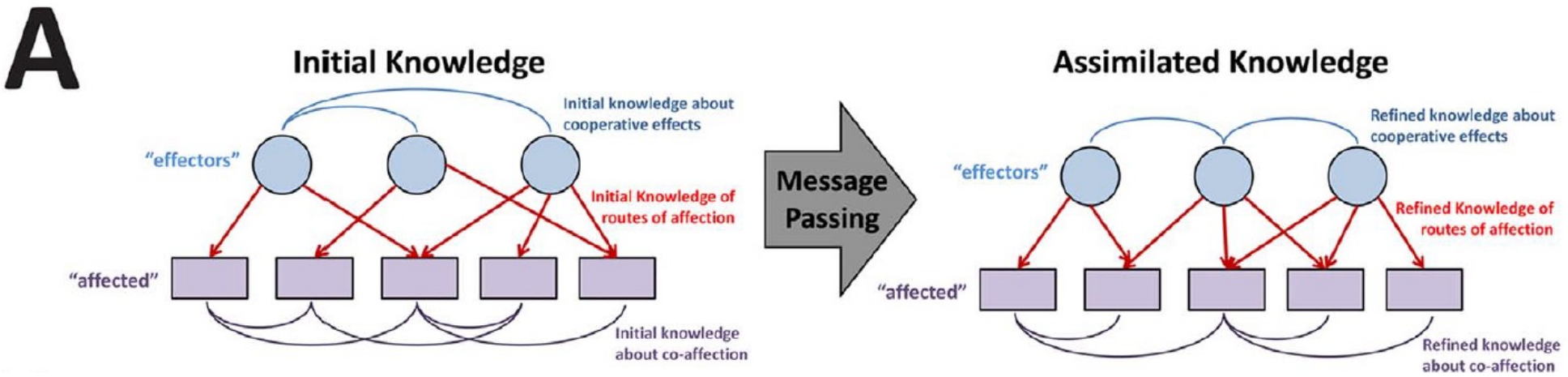


https://courses.lumenlearning.com/wm-biology1/chapter/reading-eukaryotic-transcription-gene-regulation/

Question: What type of graph does LIONESS produce?

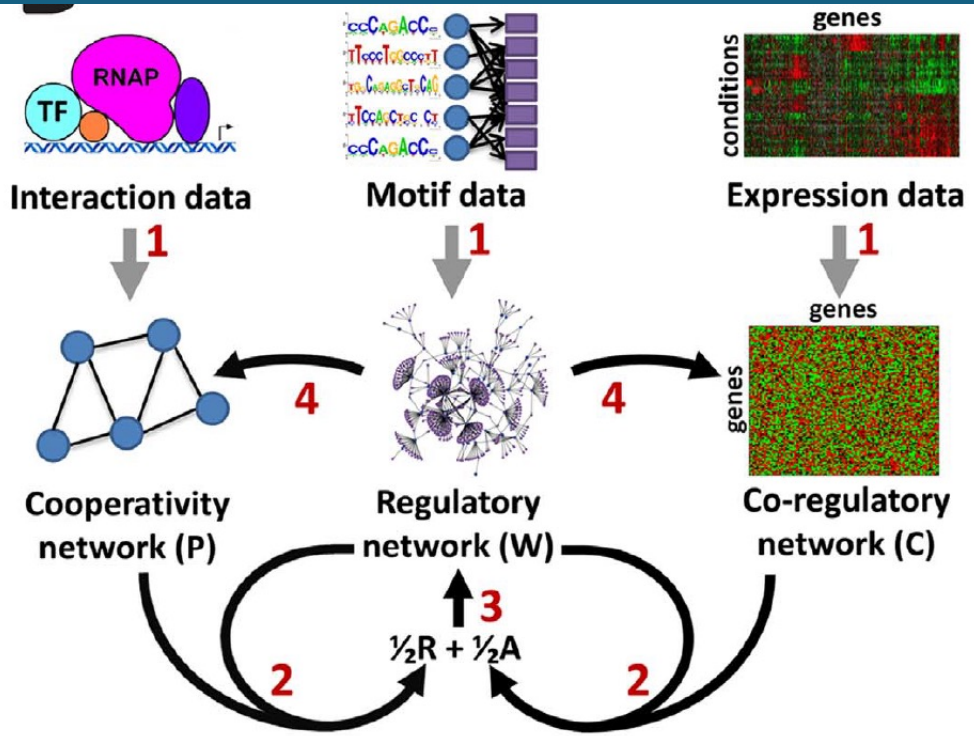


PANDA USES A MESSAGE PASSING APPROACH TO ACHIEVE THIS

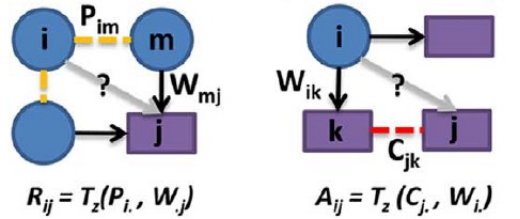


Glass et al., 2013

PANDA USES A MESSAGE PASSING APPROACH TO ACHIEVE THIS

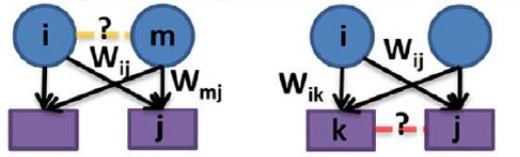


1. Initialize networks.
2. Estimate *responsibility* and *availability*.



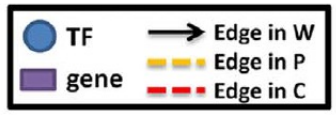
$R_{ij} = T_z(P_{i.}, W_{.j})$ $A_{ij} = T_z(C_{j.}, W_{i.})$

3. Average *R* and *A* and update *W*.
4. Estimate Protein-Cooperativity (*P*) and Co-regulatory (*C*) networks and update.



$P_{im} = T_z(W_{i.}, W_{m.})$ $C_{kj} = T_z(W_{.k}, W_{.j})$

5. Repeat 2-4.



Glass et al., 2013

LIONESS

- Similar to PANDA, but it can model transcriptional regulation for **individual samples**
- Lioness: **L**inear **I**nterpolation to **O**btain **N**etwork **E**stimates for **S**ingle **S**amples
  - Kuijjer et al., 2018

› iScience. 2019 Apr 26:14:226-240. doi: 10.1016/j.isci.2019.03.021. Epub 2019 Mar 28.

Estimating Sample-Specific Regulatory Networks

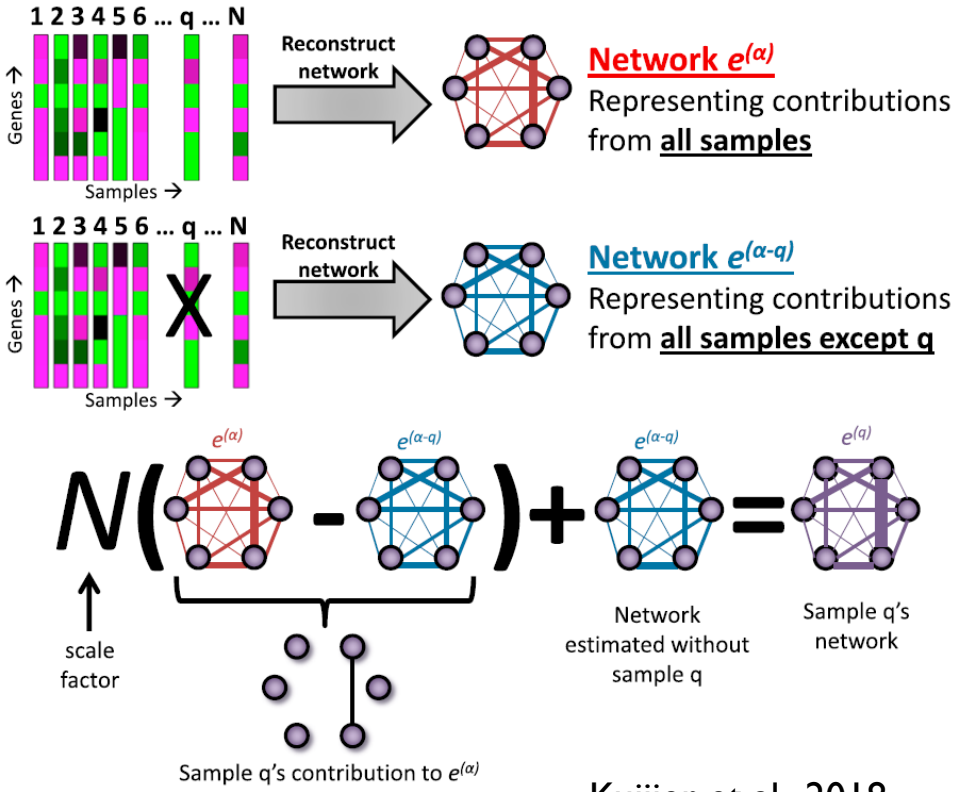
Marieke Lydia Kuijjer [1], Matthew George Tung [2], GuoCheng Yuan [3], John Quackenbush [4], Kimberly Glass [5]

Affiliations + expand

PMID: 30981959 PMCID: PMC6463816 DOI: 10.1016/j.isci.2019.03.021

LIONESS

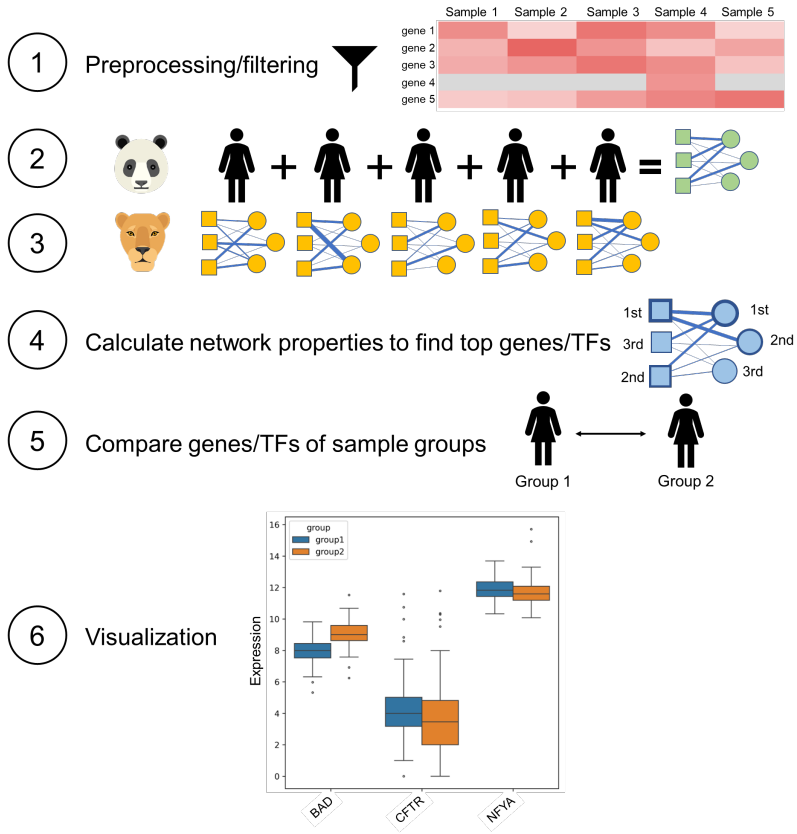
- Uses the PANDA algorithm in its network reconstruction
- Repeatedly uses the PANDA algorithm to make networks without a sample of interest
- Then compares the network without the sample to the network that included that sample and scales the difference



Kuijjer et al., 2018

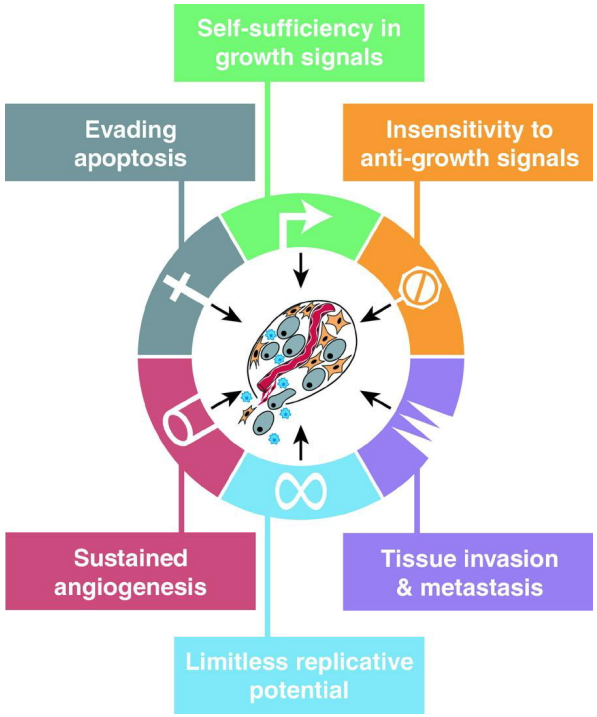
USING SISANA TO RUN PANDA/LIONESS AND ANALYZE NETWORKS

- Note: This and the following slides contain unpublished work. Please do not distribute.
- SiSaNA is a command line interface (CLI) that acts as a wrapper for PANDA and LIONESS
- Allows the user to reconstruct these networks, as well as easily perform downstream analysis and visualize their results



APPLYING THESE TOOLS TO UNDERSTAND THE REGULATION OF CANCER

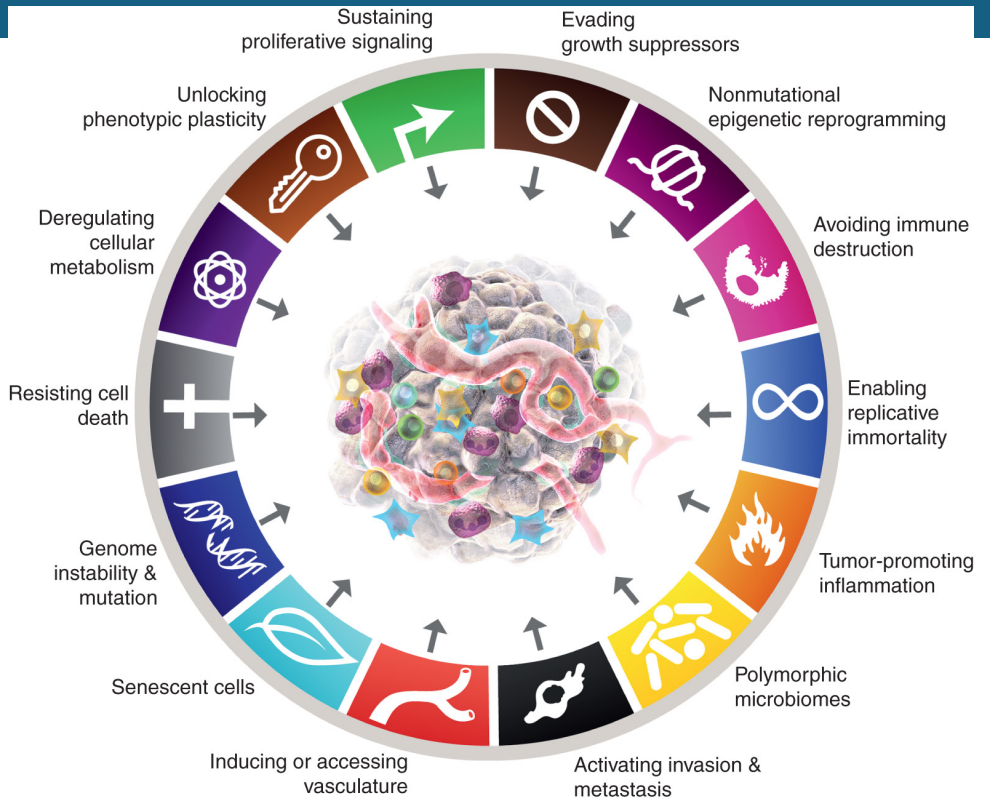
- Cancer is an incredibly difficult disease to study
- In 2000, Hanahan and Weinberg publish their famous paper “The Hallmarks of Cancer”
  - Lays out the required characteristics for a tumor to be considered cancerous



Hanahan & Weinberg, 2000
https://www.cell.com/cell/fulltext/S0092-8674(00)81683-9

APPLYING THESE TOOLS TO UNDERSTAND THE REGULATION OF CANCER

- They later expanded this model

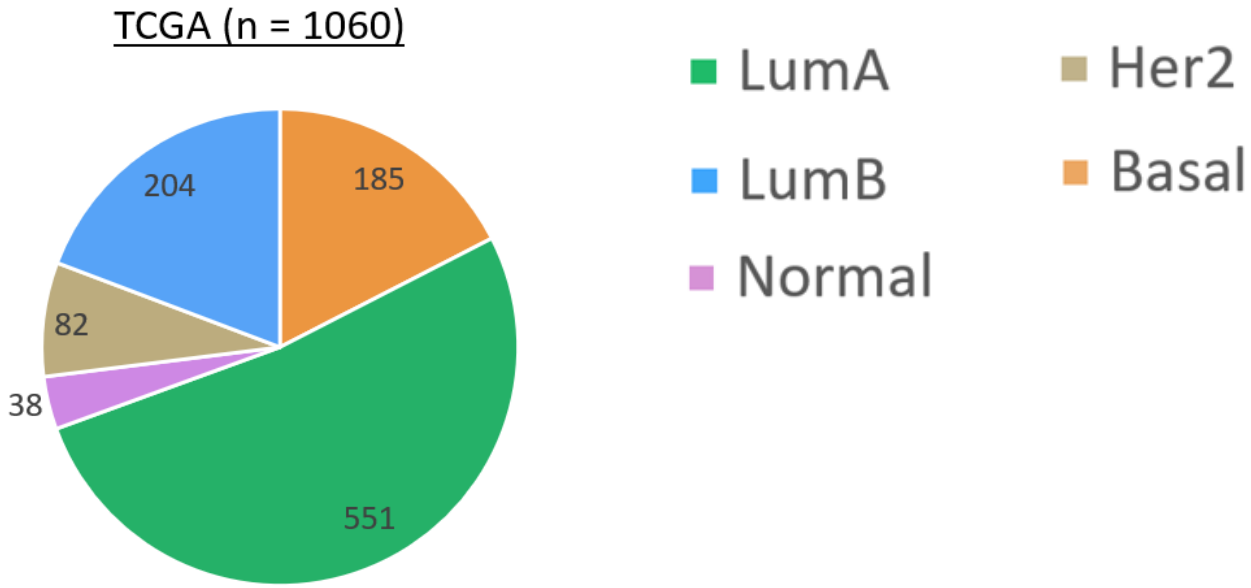


Hanahan & Weinberg, 2011
https://aacrjournals.org/cancerdiscovery/article/12/1/31/675608/Hallmarks-of-Cancer-New-DimensionsHallmarks-of

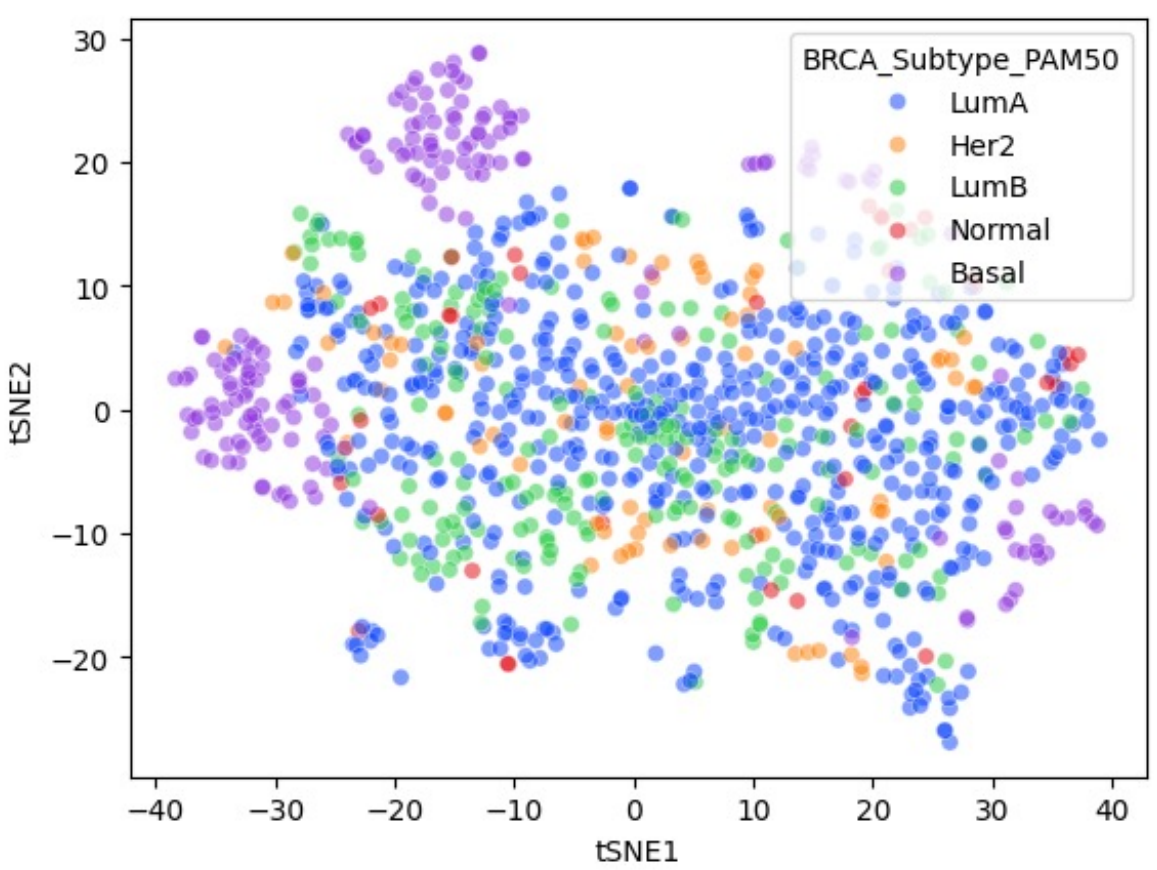
A key factor for ALL of these is transcriptional regulation!

USING LIONESS TO BETTER UNDERSTAND BREAST CANCER

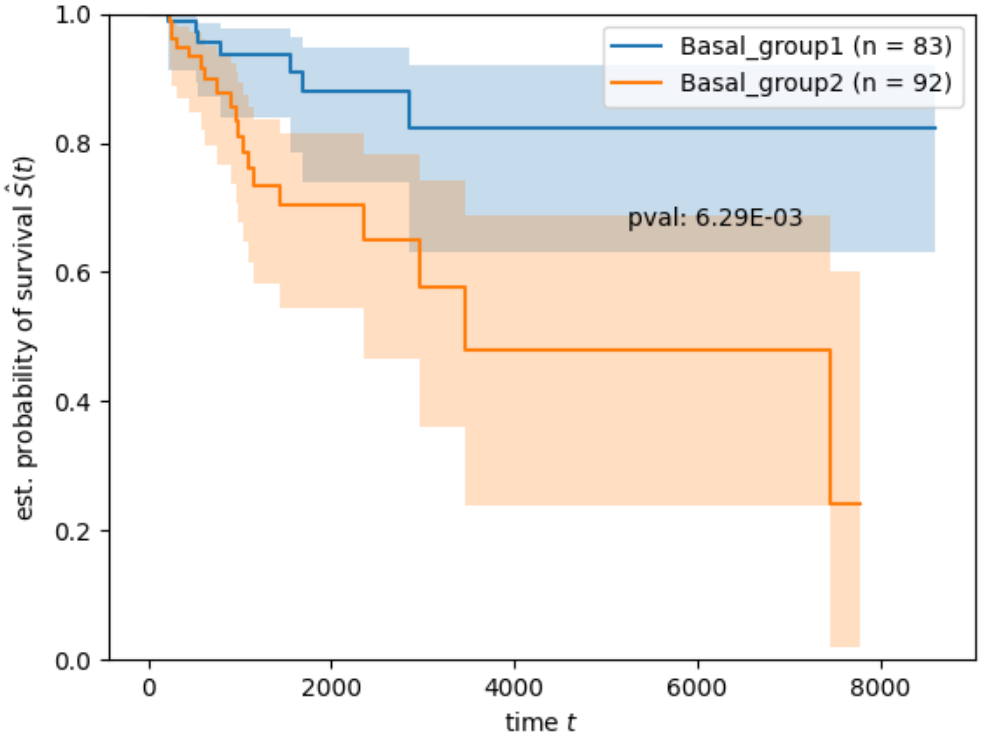
- PAM50: A set of 50 genes used to stratify patients into different subtypes of breast cancer
- Subtypes differ greatly in disease-free survival, response to treatment, and molecular classification of disease



CLUSTERING OF BREAST CANCER ON IN-DEGREES REVEALS TWO DISTINCT BASAL CLUSTERS

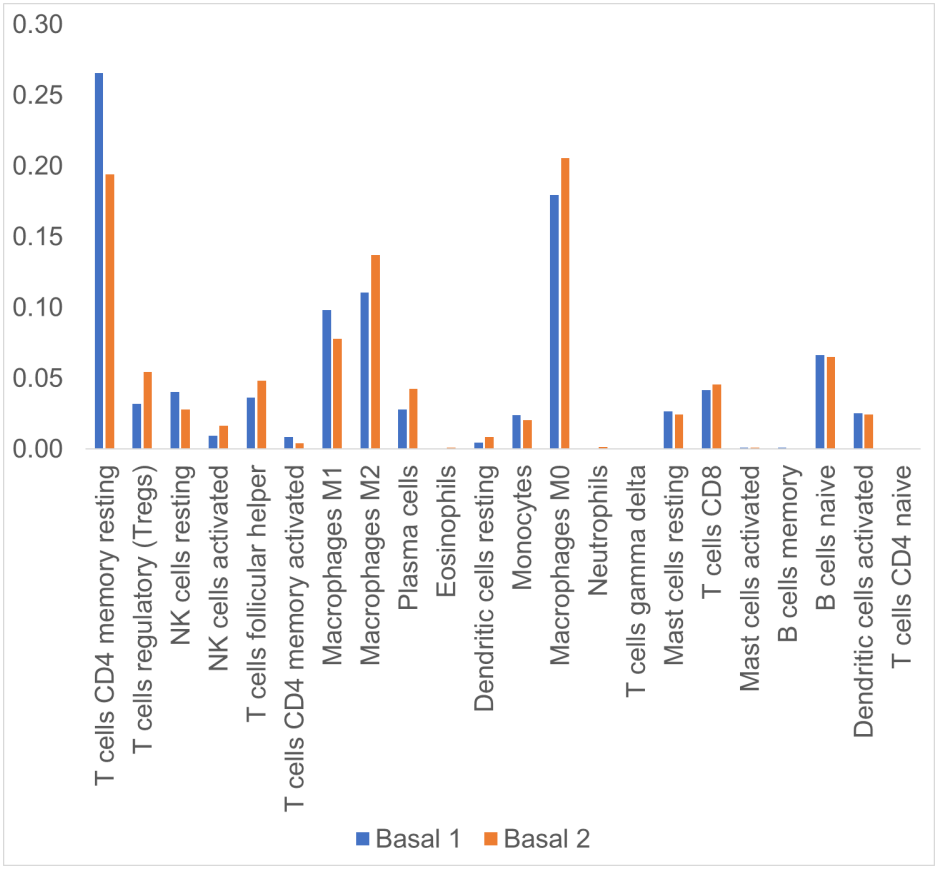


THESE TWO BASAL GROUPS GREATLY DIFFER IN SURVIVAL



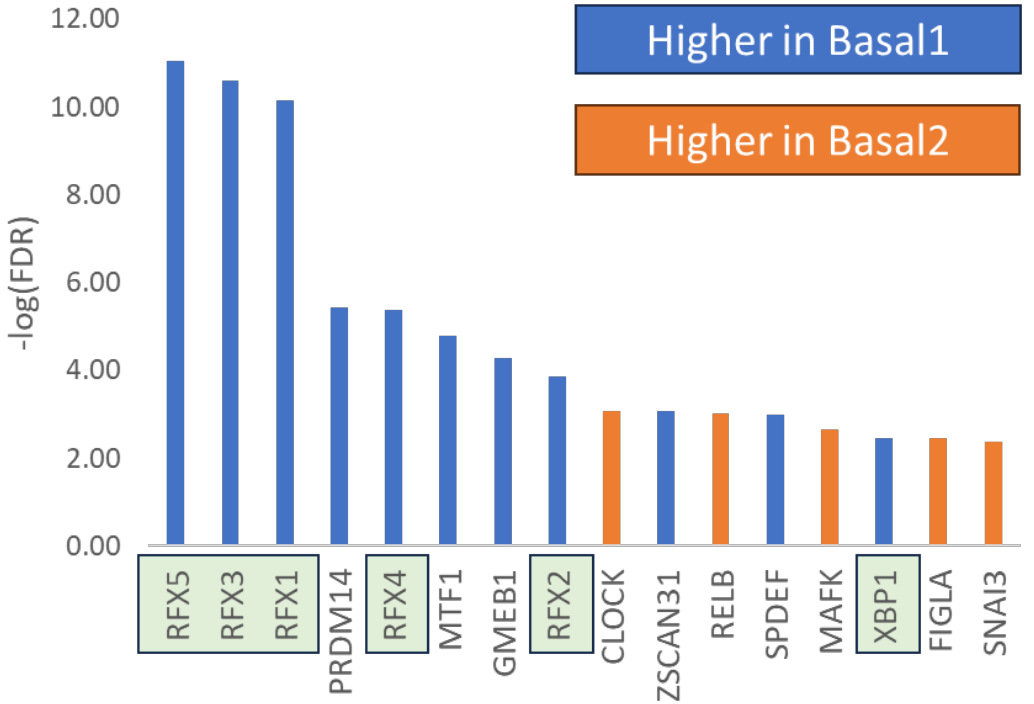
T CELL, NK CELLS, AND MACROPHAGES DISTINGUISH THE TWO BASAL GROUPS FROM ONE ANOTHER

CIBERSORTX



THE RFX FAMILY OF TRANSCRIPTION FACTORS SPECIFICALLY APPEAR TO PLAY A LARGE ROLE IN THE TRANSCRIPTIONAL LANDSCAPE THAT SEPARATES THE BASAL GROUPS

- RFX proteins
  - MHC class II regulatory factors (increases MHC class II gene expression)
  - Broadly recognized to be associated with various cancer types
  - Regulate genes through the x-box sequence motif
  - ESR1 promotes expression of RFX1-3,5,7 and ESR1 is also expressed higher in Basal1



SISANA LEARNING OBJECTIVES

1. Clone the NetZooPy and SiSaNA environments
2. Perform preprocessing of data and reconstruction of LIONESS networks
3. Understand how to interpret the LIONESS networks, including in-degree and out-degree
4. Visualize the results of your analysis and come to biological conclusions

SUMMARIZING WHAT WE LEARNED TODAY

16S microbiome analysis

- QIIME2 is a great piece of software that can allow us to process raw 16S reads
- There are many ways to study the microbiome
- MicriobiomeAnalyst (and other great tools) help us analyze the microbiome

Network reconstruction and analysis

- Biology is messy, but networks can help us understand the mess
- Many network analysis tools exist. We just scratched the surface!
- TkNA allows the user to predict regulatory features in their network that impact disease
- PANDA/LIONESS/SiSaNA allow for the modeling and analysis of regulatory networks so we can identify transcription factors with high activity or genes that are highly regulated in disease
- Cytoscape helps us visualize these networks, and also performs basic network analysis methods

THANK YOU!!!

nolankn@uio.no

github.com/newmanno