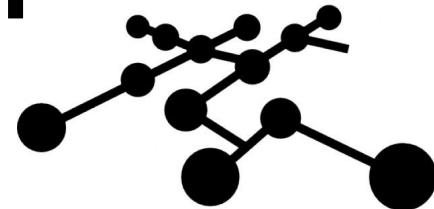


# Computational Group

-

# Network Analyses



July 2023 TBRU Meeting

Presentation by

Ben Polacco Krogan Lab UCSF

Max Bennett Krogan Lab UCSF

Ellie Reinhart Cox Lab UC Berkeley

Featuring work by

Mehdi Bouhaddou Krogan Lab UCSF

Assistant Professor of Microbiology, Immunology, and Molecular Genetics at UCLA

# Looking for Drivers of Innate Immunity using a Network Model

Network Propagation  
Shortest Pathway Analysis

All this data, how does it connect?

## CLINICAL

SNP  
ATACseq  
Methylation  
RNAseq  
Proteomics

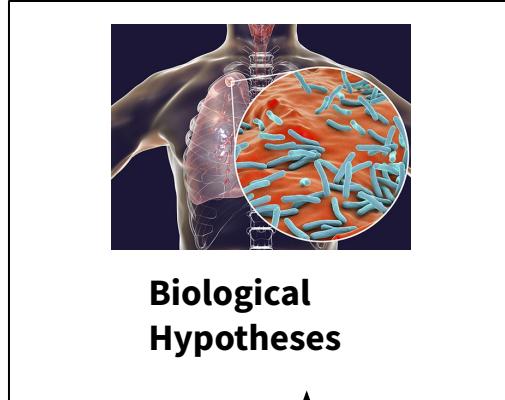
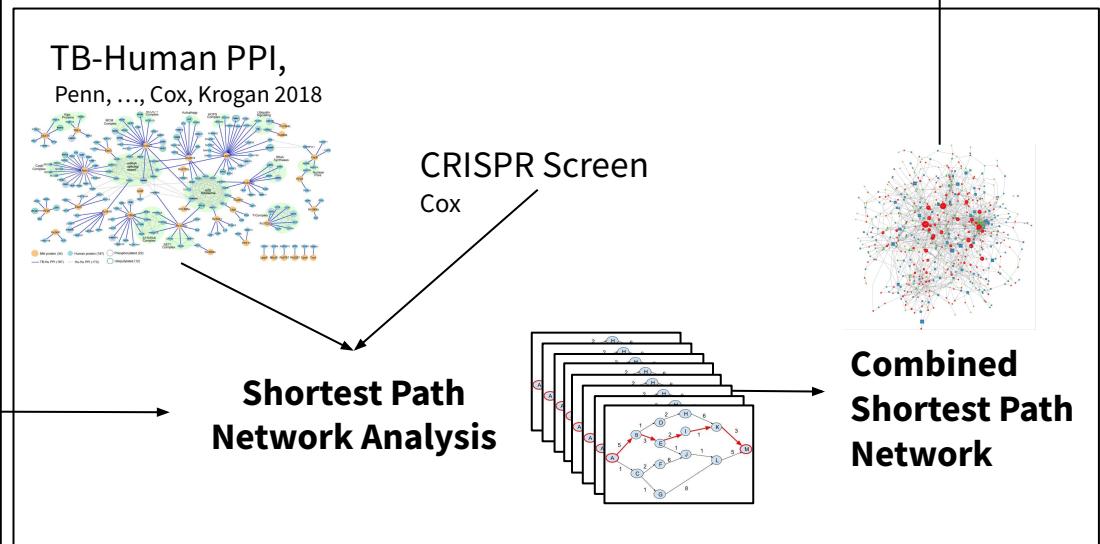
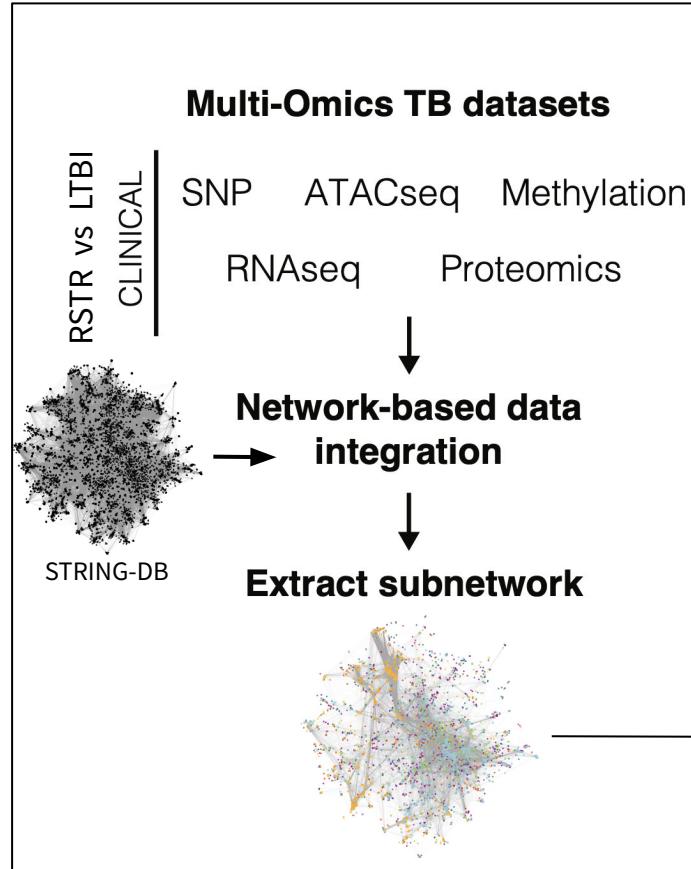
## EXPERIMENTAL

PPI  
CRISPR screen  
PTM proteomics  
Proteomics

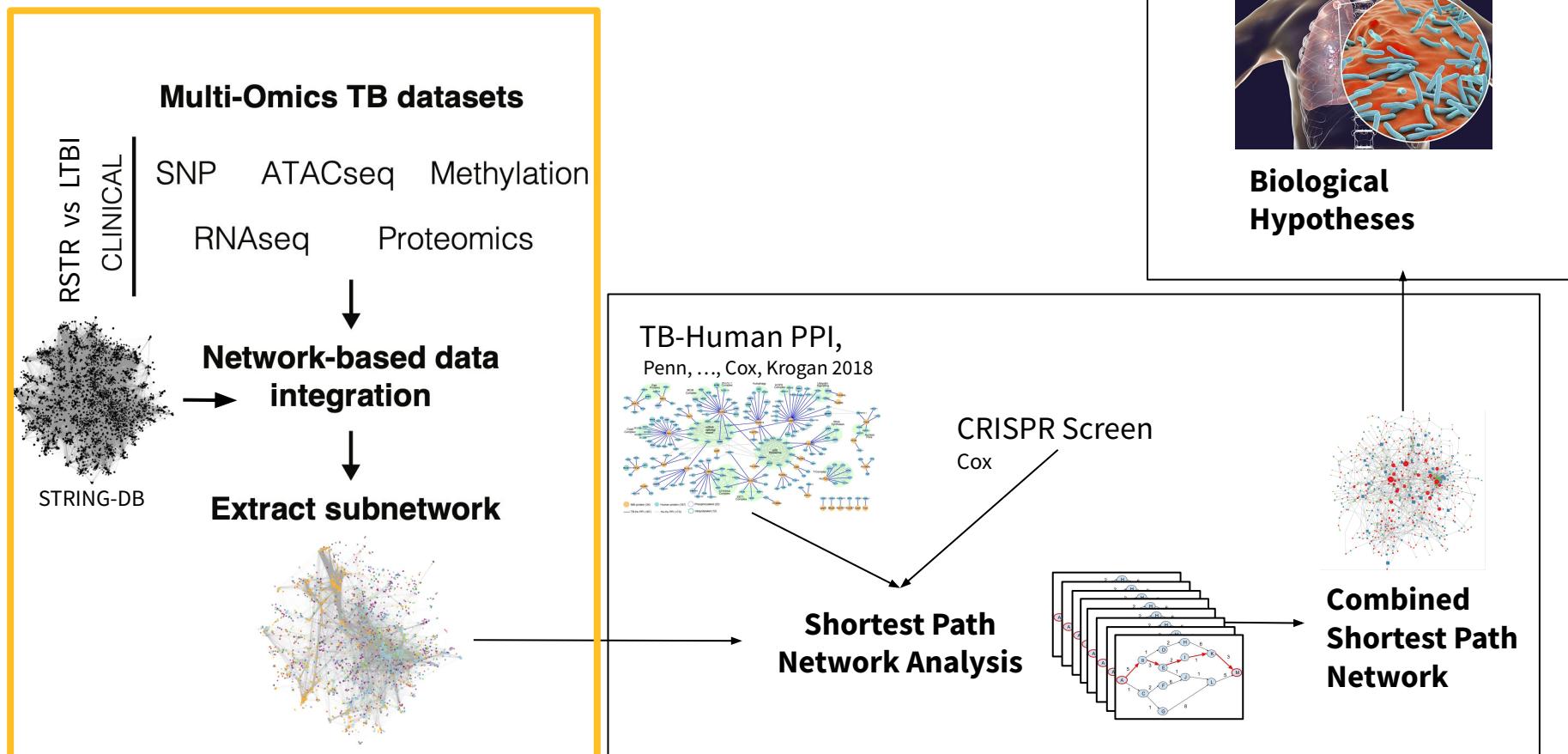
Strategy: Leverage public PPI  
databases to find strongest  
connections

---

# Overview



# Overview

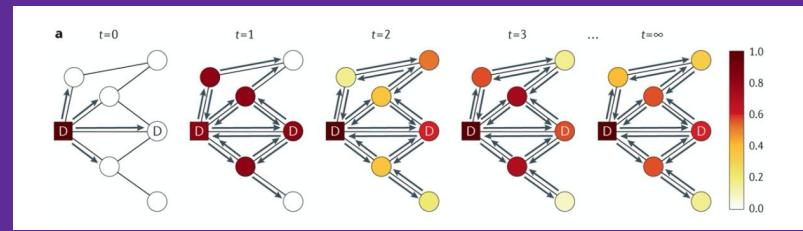


# Part 1, Network Propagation

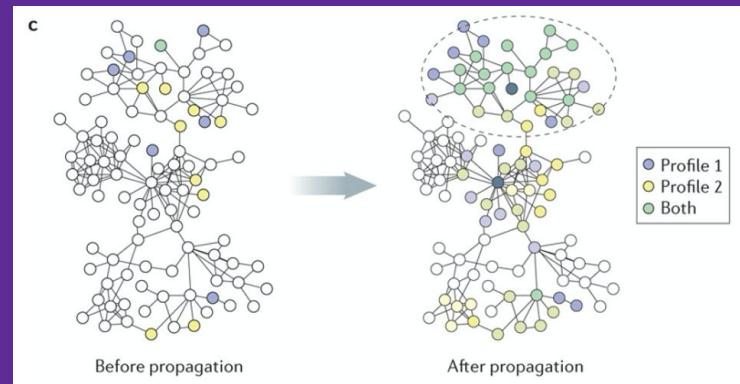
Majority by Mehdi Bouhaddou\*,  
Follow up and Presentation by  
Ben Polacco  
Krogan Lab, UCSF

\* now at UCLA

Diffusion on interaction network (STRING):  
1) Extends observed data to important network proteins,



2) Corroborates network neighbors within and between orthogonal datasets (PPI, RNAseq, etc)



Adapted from Cowen et al, 2017

# Background Network : “Bigger is Better”\*

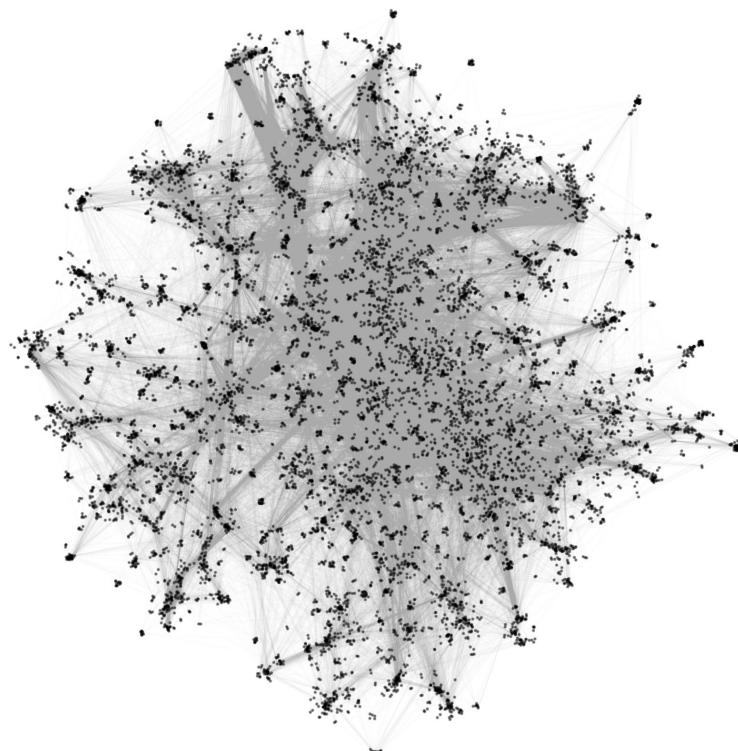
STRING-DB

BG Network:

**STRING-DB (combined score > 400)**

**19,076 Human Proteins**

**888,774 Edges (Undirected, Unweighted)**



\*Approximate quote; Huang, ..., Ideker, 2018

# Protein Abundance Input Heat

STRING-DB + TBR Protein

BG Network:

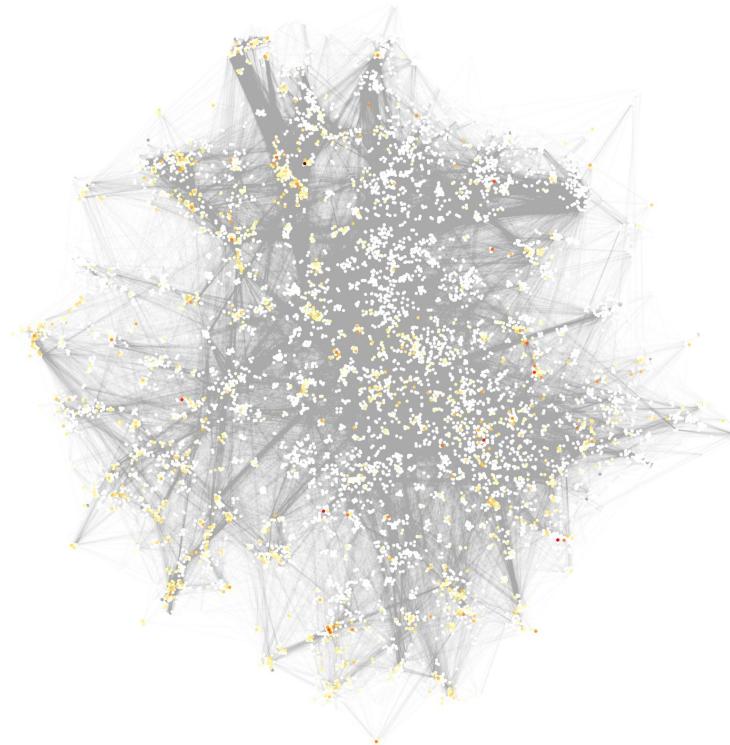
STRING-DB, combined score > 400

19,076 Nodes (Proteins)

888,774 Edges (Undirected, Unweighted)

**Input Heat: a measure of statistical significance  
and magnitude for any protein change**

$$\text{sqrt}(-\log_{10}(p) \times \text{abs}(\log_{2}\text{FC}))$$



# Propagated Protein Effects

STRING-DB + TBR Protein, Network Propagation

BG Network:

STRING-DB, combined score > 400

19,076 Nodes (Proteins)

888,774 Edges (Undirected, Unweighted)

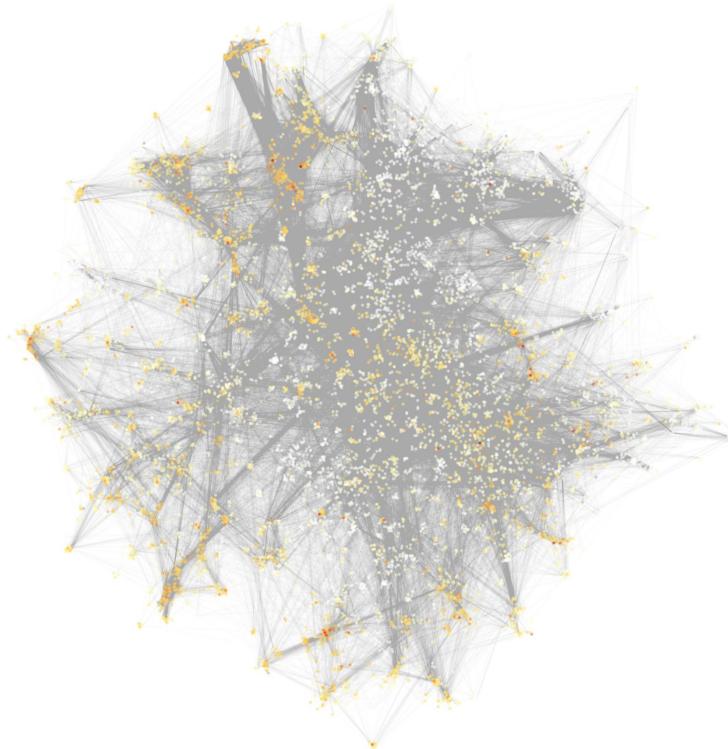
Input Heat: a measure of statistical significance and magnitude for any protein change

$$\text{sqrt}(-\log_{10}(p) \times \text{abs(log2FC)})$$

**Network Propagation:**

**Each gene gets a score (not sparse)**

**Neighbors corroborate (less noisy)**



# Protein significance by Network Propagation

STRING-DB + TBR Protein, Network Propagation,  $p < 0.05$

BG Network:

STRING-DB, combined score  $> 400$

19,076 Nodes (Proteins)

888,774 Edges (Undirected, Unweighted)

Input Heat: a measure of statistical significance and magnitude for any protein change

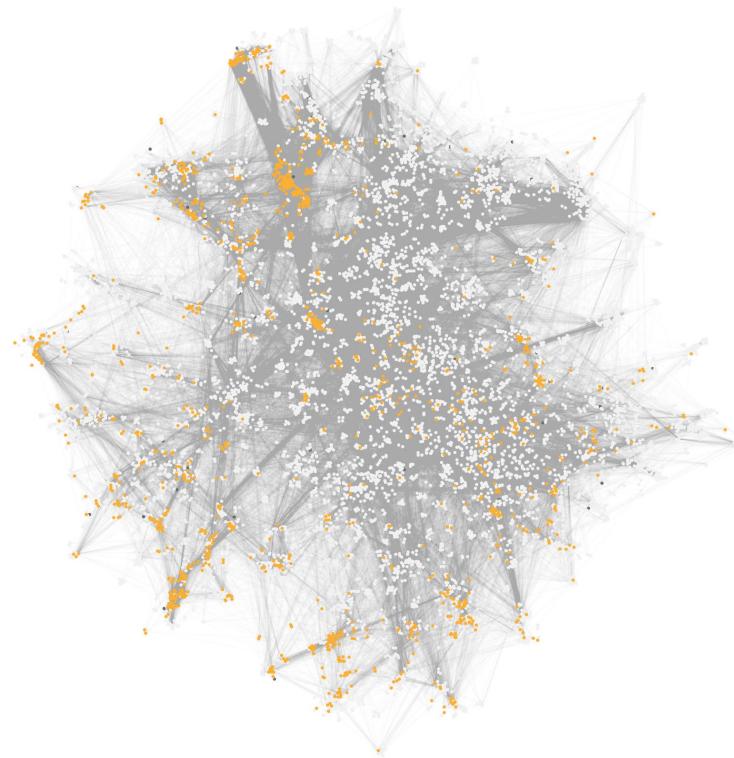
$$\text{sqrt}(-\log_{10}(p) \times \text{abs}(\log_{10}FC))$$

Network Propagation:

Each gene gets a score (not sparse)

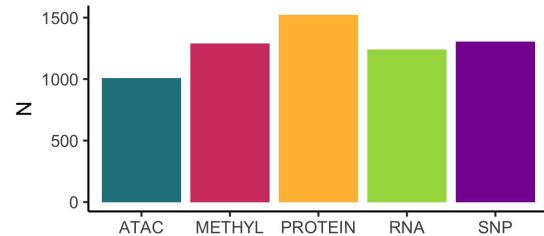
Neighbors corroborate (less noisy)

**Threshold  $P < 0.05$  (permutation analysis)**

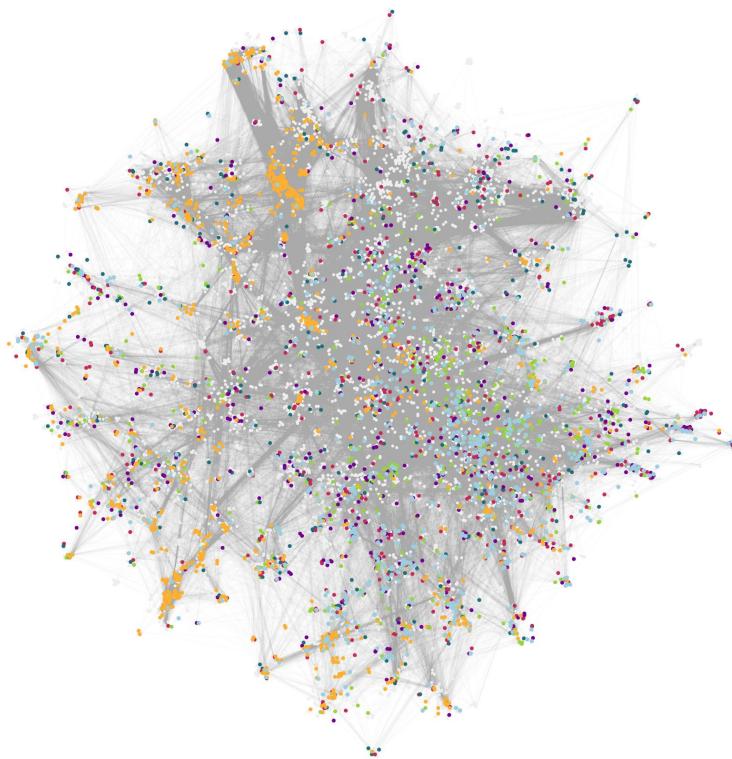
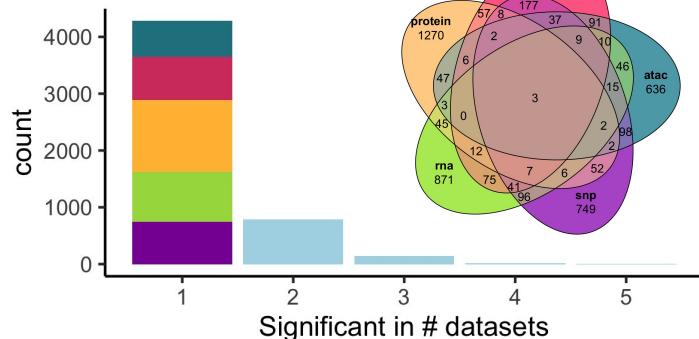


# Integrate output of NP on all Omics

All NP  $P < 0.05$

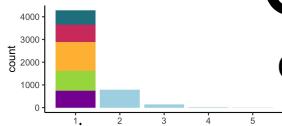


Usually look at overlap between datasets,  
but low numbers of overlap here.



# Genes and Gene Sets in Sub-Network

## Gene Ontology Enriched Sets



-log<sub>10</sub>(adj.p)

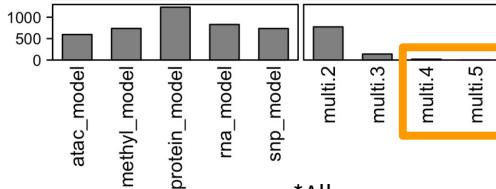
# genes matching term and group

53	white: p < 0.05
27	gray: p > 0.05

11	13	30	83	10
8	15	32	89	13
17	10	83	10	8
10	9	117	7	13
15	28	31	53	39
14	41	21	27	25
14	29	19	25	41
1	1	8	2	
20	24	40	31	25
2	20	2	1	9
4	6	2	23	
4	13	10	22	34

69	14	4	2	
54	7	2	1	
12	3	1	1	
18				
58	26	6	2	
35	19	5	1	
46	17	7	1	
7	11			
44	5	3	3	
13	1			
12	7	1		
18	4	1	1	

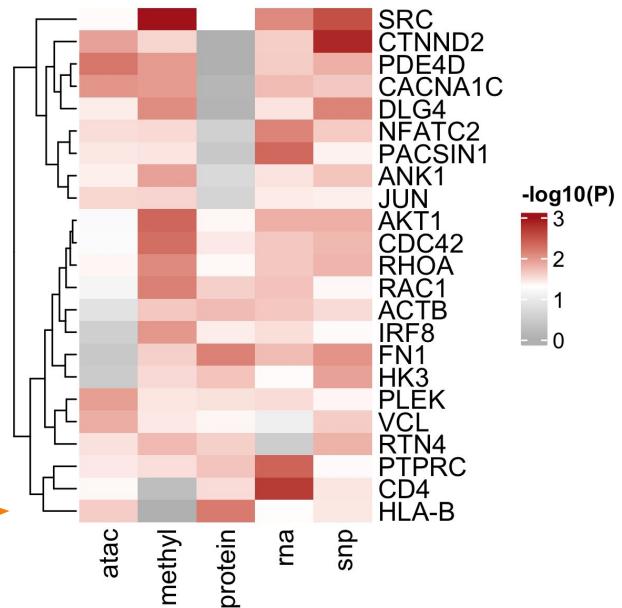
- regulation of lymphocyte activation
- adaptive immune response
- regulation of translation
- mitochondrial matrix
- positive regulation of cell migration
- cell-substrate adhesion
- cell junction assembly
- phosphotyrosine residue binding
- negative regulation of intracellular signal transduction
- collagen trimer
- endocrine system development
- vascular process in circulatory system



\*All groups are mutually exclusive, no gene is double-counted

Key genes and pathways may be among those observed in multiple datasets.

## Genes significant in 4 or 5 NP



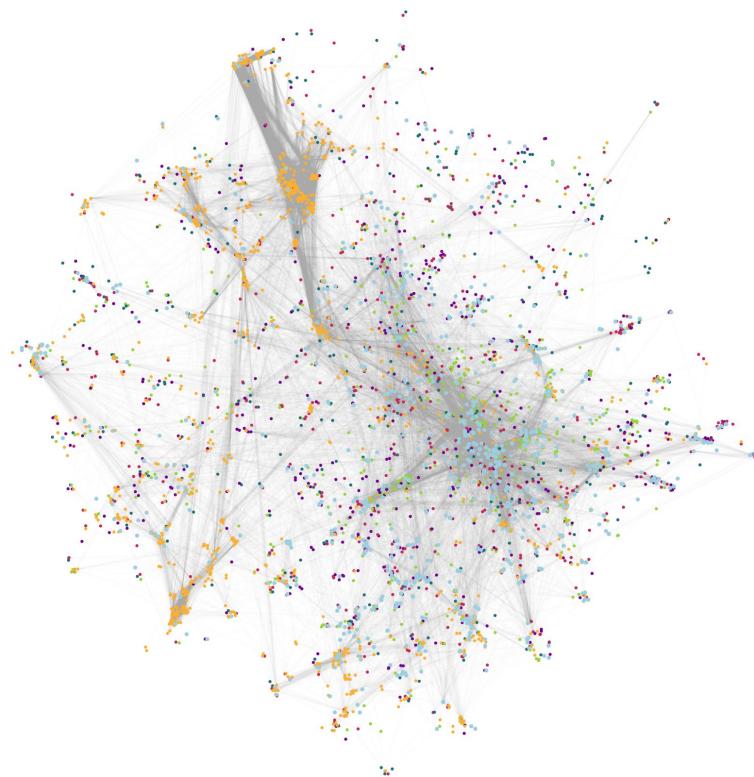
# Network Subset

All NP P < 0.05

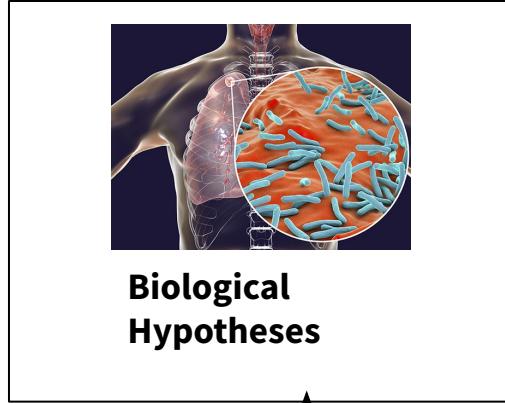
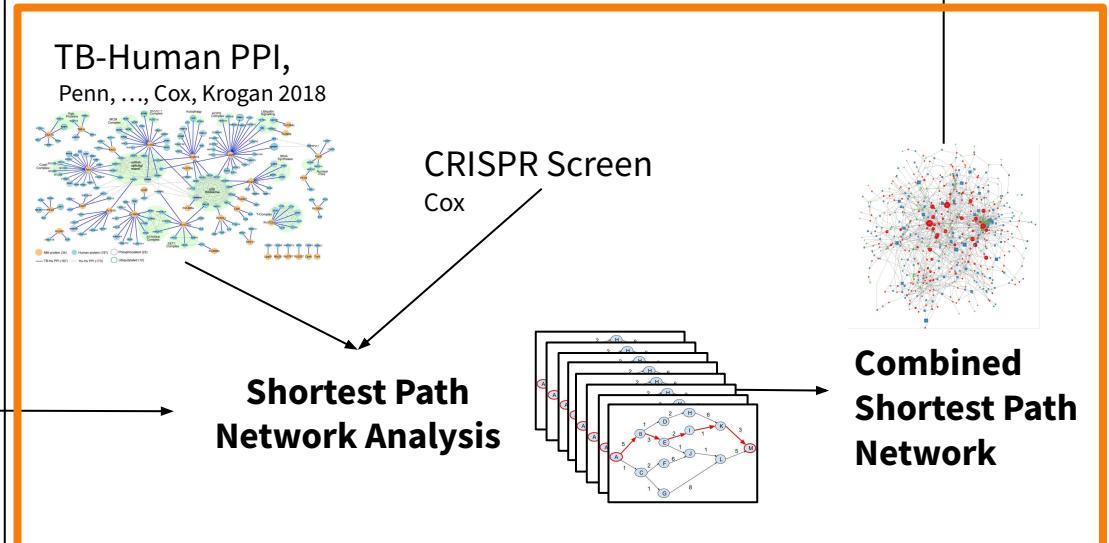
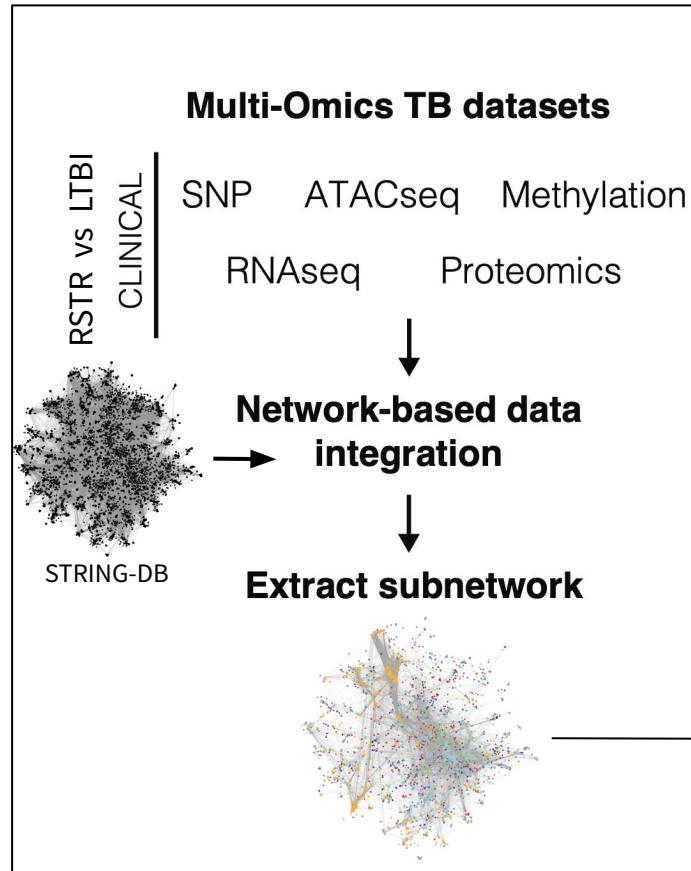
**Union of all: 5205 genes**

**Represents all genes and pathways hypothesized important in TB infection (RSTR vs LTBI phenotypes)**

**This will return in next section.**



# Overview



# The Data: AP-MS and CRISPR Screen

## Mtb-secreted factor host interactors

- Identified secreted proteins through mass spec of culture filtrates
  - 105 proteins
- Expressed proteins and purified
- Added purified proteins in lysates from U937 macrophages
- Putative interacting partners identified by mass spec

### An *Mtb*-Human Protein-Protein Interaction Map Identifies a Switch Between Host Anti-Viral and Anti-Bacterial Responses

Bennett H. Penn<sup>1,2,†</sup>, Zoe Netter<sup>2</sup>, Jeffrey R. Johnson<sup>3</sup>, John Von Dollen<sup>3</sup>, Gwendolyn M. Jang<sup>3</sup>, Tasha Johnson<sup>3</sup>, Yamin M. Ohol<sup>4</sup>, Cyrus Maher<sup>5</sup>, Samantha L. Bell<sup>1,4</sup>, Kristina Geiger<sup>2</sup>, Guillaume Golovkine<sup>2</sup>, Xiaotang Du<sup>4</sup>, Alex Choi<sup>4</sup>, Trevor Parry<sup>4</sup>, Bhopal C. Mohapatra<sup>6</sup>, Matthew D. Storck<sup>6</sup>, Hamid Band<sup>6</sup>, Chen Chen<sup>2</sup>, Stefanie Jäger<sup>3</sup>, Michael Shales<sup>3</sup>, Dan A. Portnoy<sup>2,7</sup>, Ryan Hernandez<sup>5</sup>, Laurent Coscoy<sup>2</sup>, Jeffery S. Cox<sup>#2,8,\*</sup>, and Nevan J. Krogan<sup>#3,8,9,\*</sup>

## TNFA/iNOS and IFNb sorted CRISPR KD/KOs

Two genome wide crispr screens based in macrophages

Both FACs based

IFNb readout

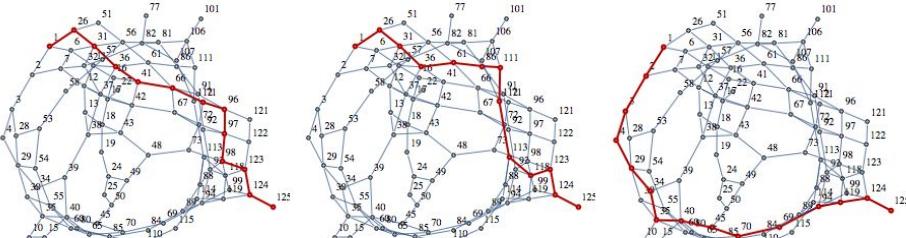
- Thp1s
- CRISPRi

TNF $\alpha$  readout -

- CIMs
- CRISPRko
- Also sorted by iNOS+

# Part 2, Pathway Analysis

Genes critical to phenotype across experiments should be related closely in the PPI Network.



AP-MS PPI hits represent the beginning of TB disease pathology in the host cell

CRISPR hits are likely drivers of macrophage intracellular immune phenotypes (+/- TNF, iNOS, and IFNb)

To connect **PPI** & **CRISPR KD/KO** datasets  
we calculated shortest paths  
*from each CRISPR hit to any PPI hit*

# Creating the Subnetwork(s)

- Experimental PPI evidence from STRING is used as the base network.
- Genes that aren't in the network propagation output are excluded.
- Shortest paths between the PPI and CRISPR hits are found and combined.
- Information about the subnetwork's structure, its overlap with clinical and experimental datasets, and the individual shortest paths that make up the subnetwork are all collated and further analyzed.

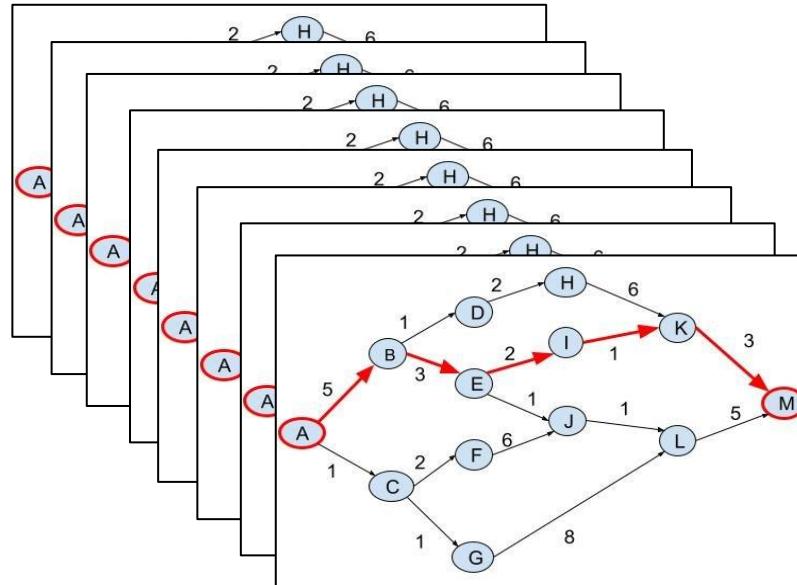
From around 19,000 genes and 200,000 experimentally derived interactions in STRING



To around 5000 genes and 60,000 edges in the network propagation results

# Creating the Subnetwork(s)

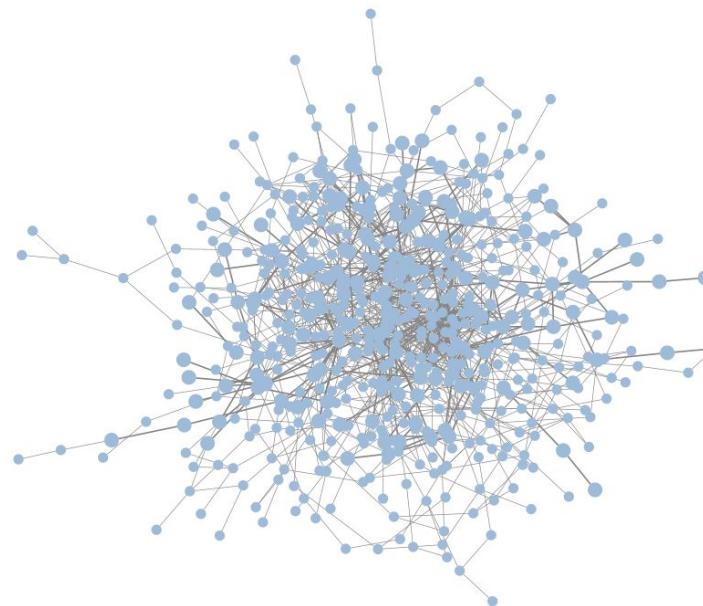
- Experimental PPI evidence from STRING is used as the base network.
- Genes that aren't in the network propagation output are excluded.
- Shortest paths between the PPI and CRISPR hits are found and combined.
- Information about the subnetwork's structure, its overlap with clinical and experimental datasets, and the individual shortest paths that make up the subnetwork are normalized using randomized source and target nodes and further analyzed.



# Creating the Subnetwork

- Experimental PPI evidence from STRING is used as the base network.
- Genes that aren't in the network propagation output are excluded.
- Shortest paths between the PPI and CRISPR hits are found and combined.
- Information about the subnetwork's structure, its overlap with clinical and experimental datasets, and the individual shortest paths that make up the subnetwork are all collated and further analyzed.

561 genes and 1254 edges in the simplified union of shortest paths.



Network is generated with random source and targets 1000 times

```
classes 'data.table' and 'data.frame': 561 obs. of  20 variables:  
$ string  
$ betweenness  
$ degree  
$ eigen  
$ reach  
$ npaths  
$ normBetwn  
$ normDegree  
$ normEigen  
$ normReach  
$ normPaths
```

: chr "9606.ENSP00000300589" "9606.ENSP00000358335" "9606.ENSP0000222005" "9606.ENSP00000392147" ...  
: num 0 4 0 0 2 0 4 2 0 0 ...  
: num 1 4 12 4 2 3 3 2 14 1 ...  
: num 0.000549 0.006875 0.051019 0.034032 0.000142 ...  
: num 0.00714 0.03036 0.13393 0.075 0.01071 ...  
: int 2 8 28 10 6 6 4 21 2 ...  
: num 0 0.671 0 0 0.439 ...  
: num 0.0982 0.9066 2.5125 1.4785 0.5894 ...  
: num 0.00113 0.17939 1.77535 16.93649 0.03135 ...  
: num 0.0832 0.7042 3.0376 6.7285 0.5948 ...  
: num 0.13 0.887 4.063 2.623 1.085 ...



MS Protein Abundance



RNA Sequencing



ATAC Sequencing

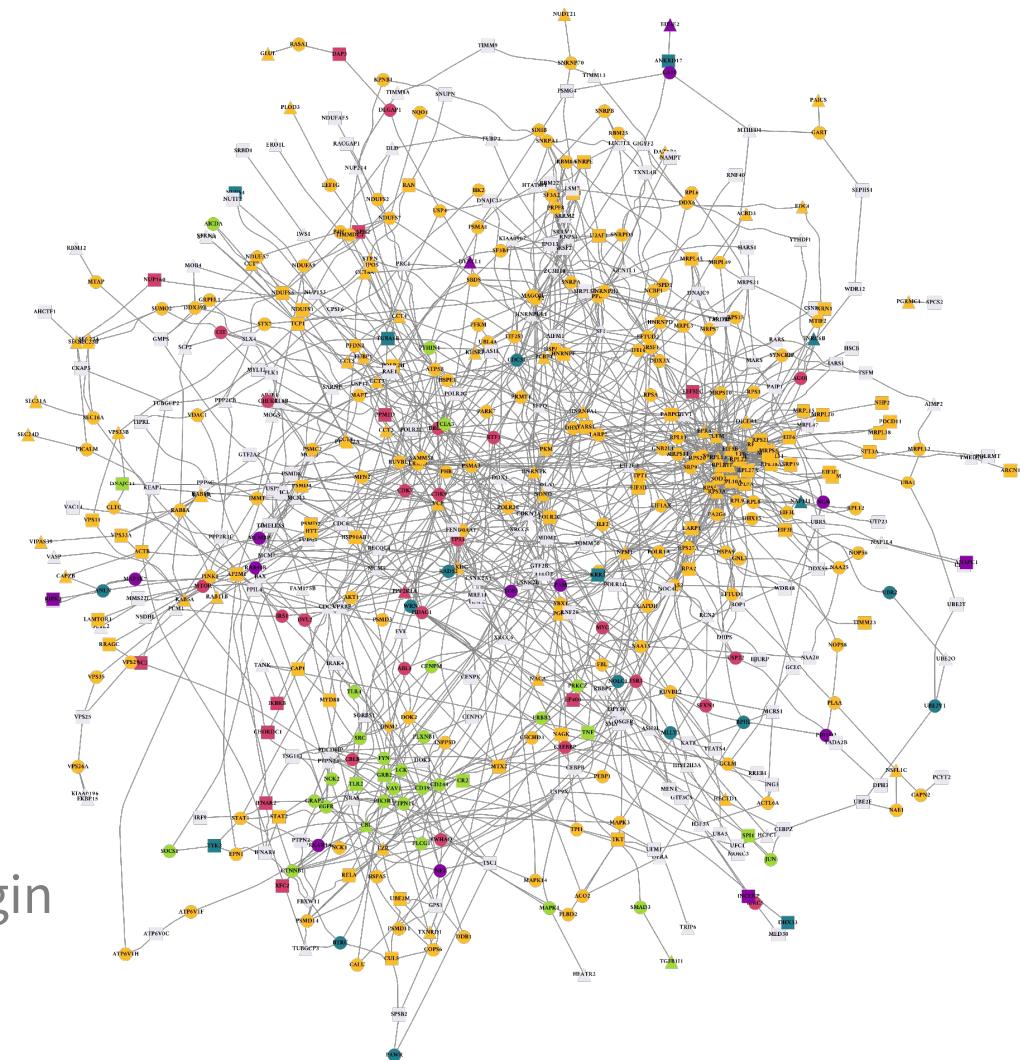


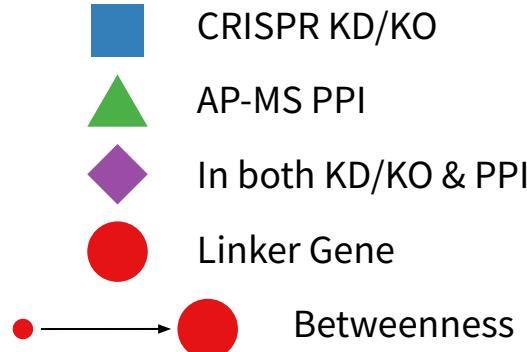
SNP Analysis



Methylation Analysis

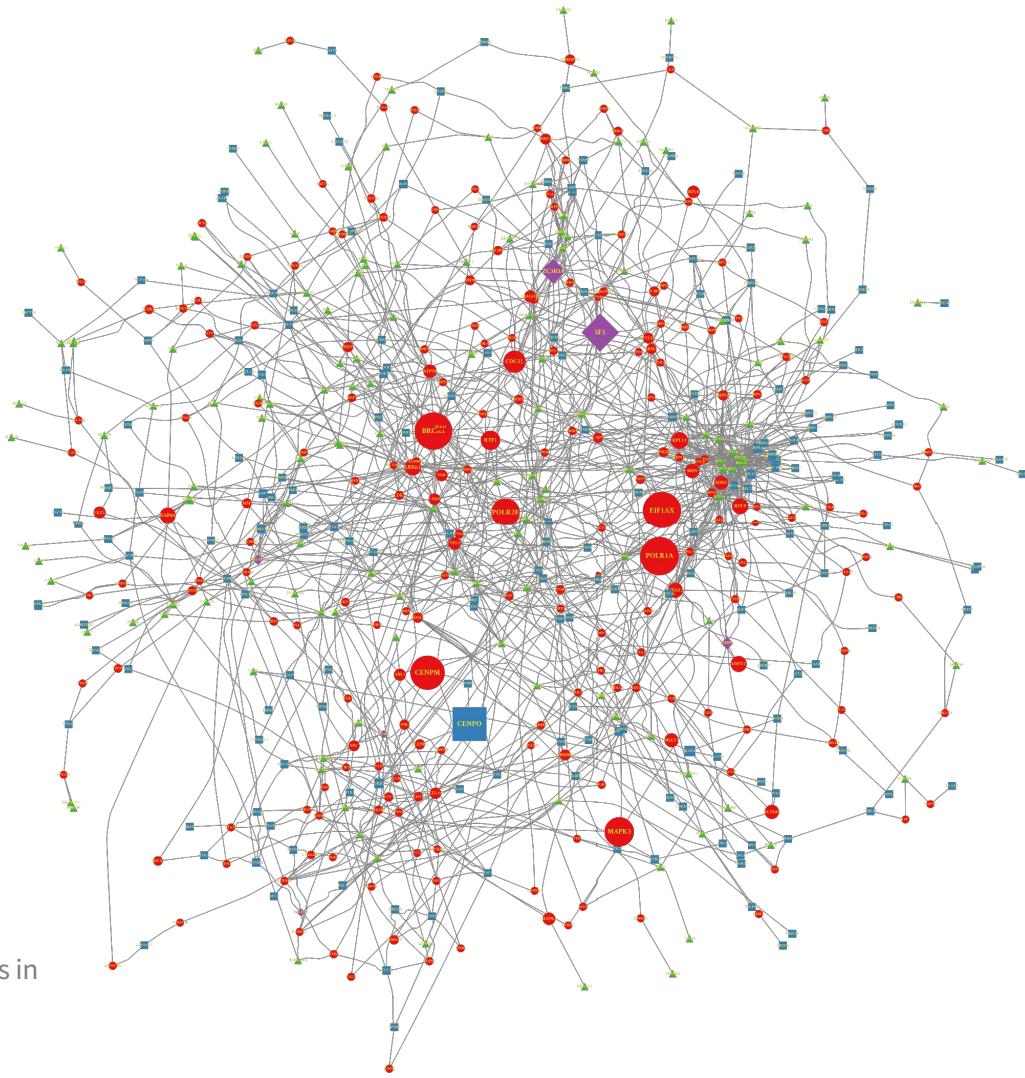
Network colored by Dataset of Origin

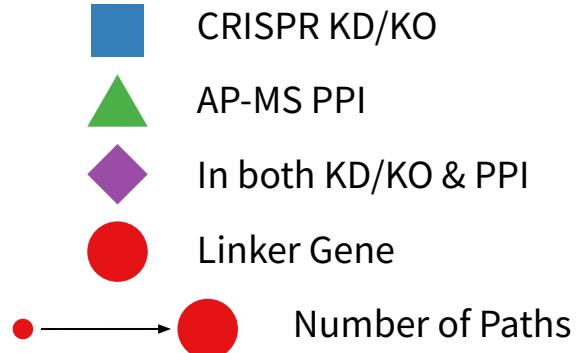




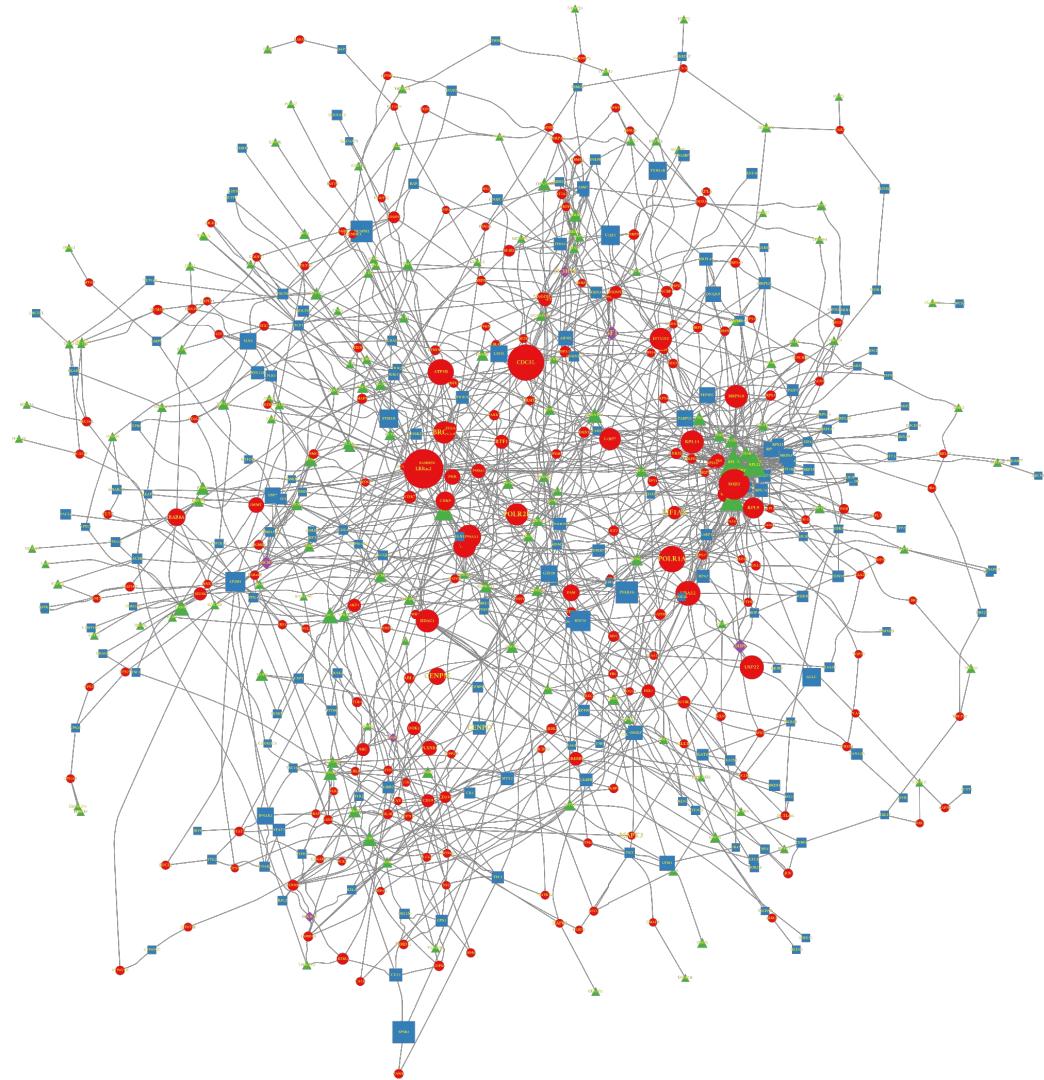
## Network scaled by Betweenness

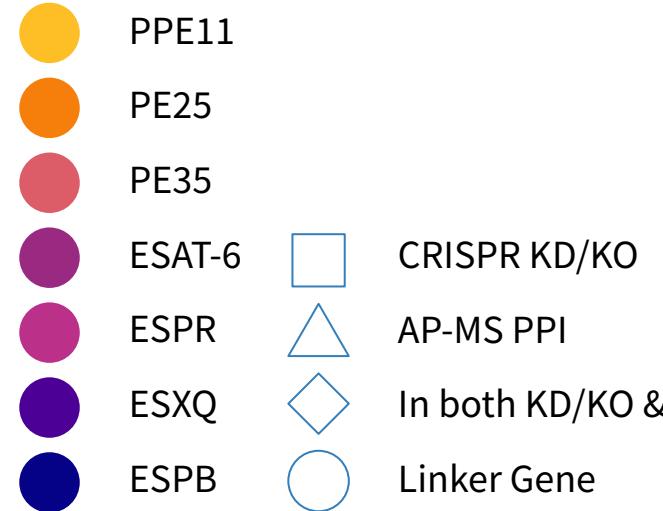
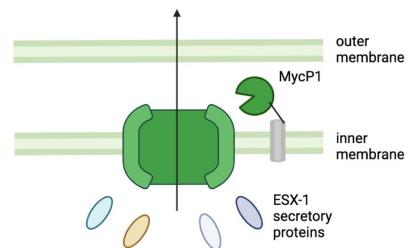
Score normalized for each node by the node's mean betweenness in 1000 simulated networks with random source and target nodes



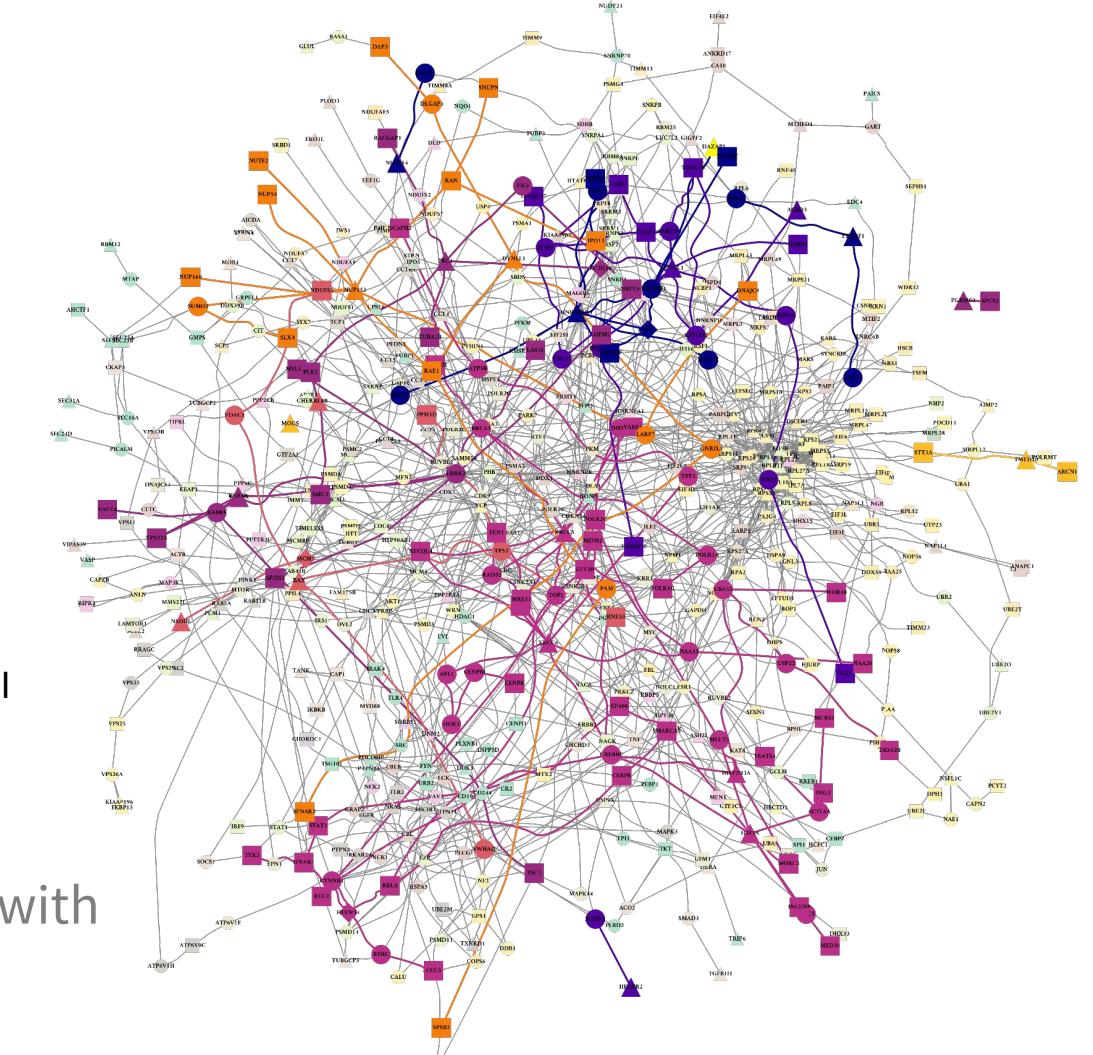


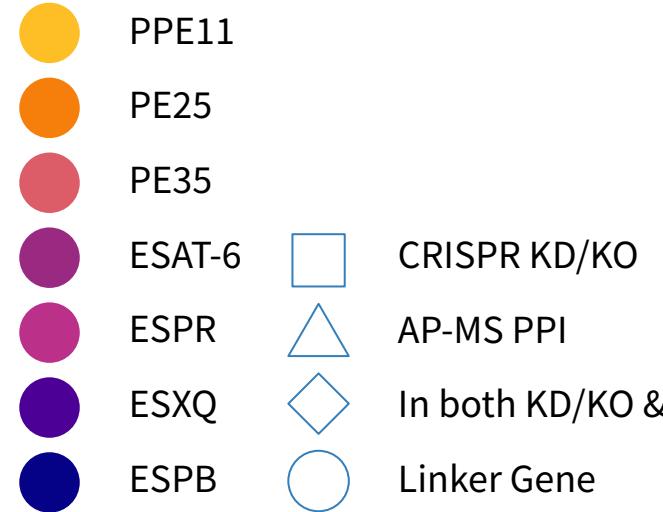
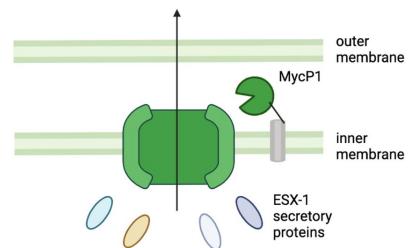
Network scaled by Number  
of Shortest paths



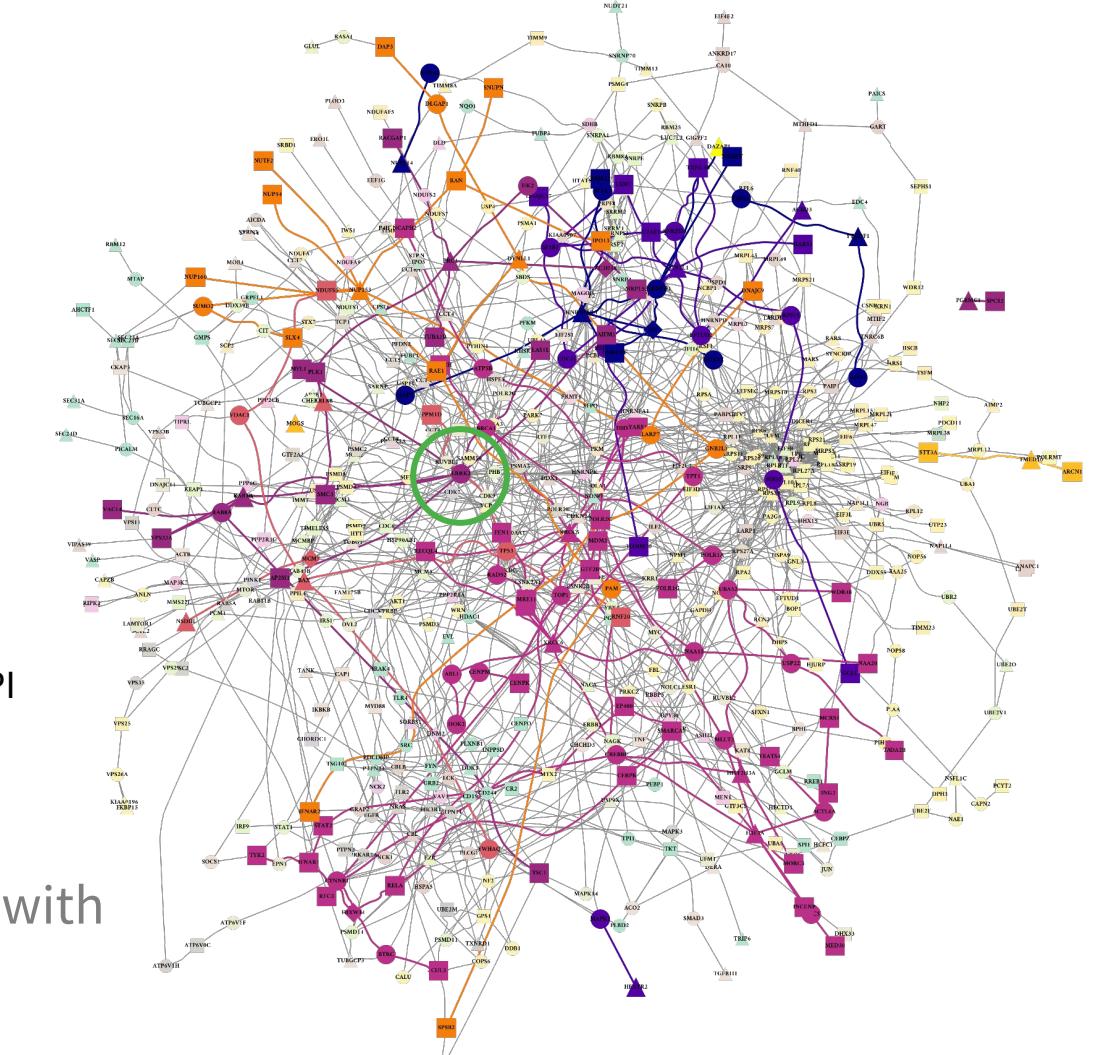


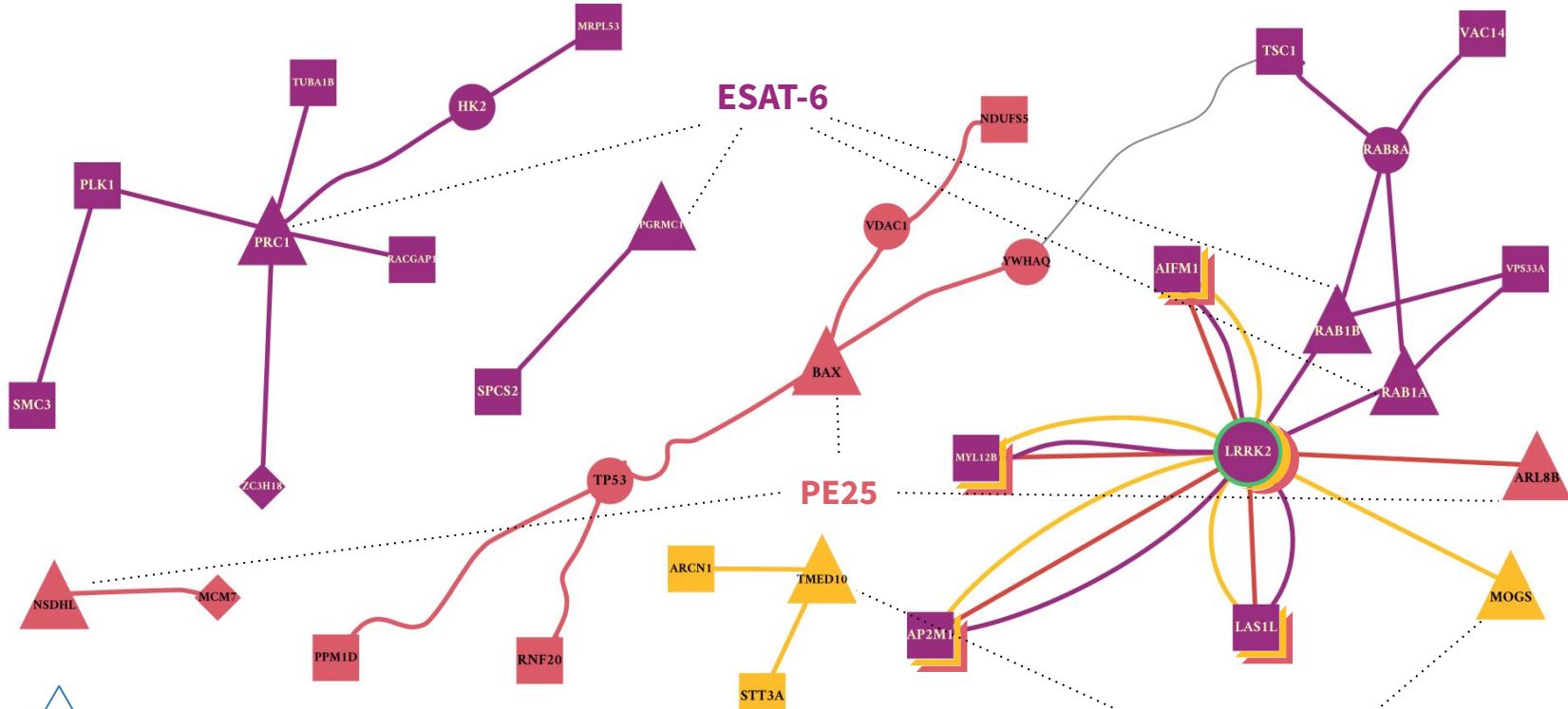
Network colored by interaction with  
TB-Secreted ESX-1 Proteins





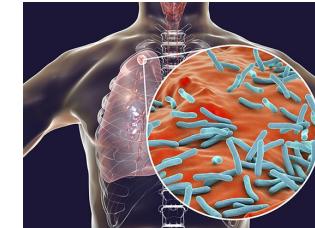
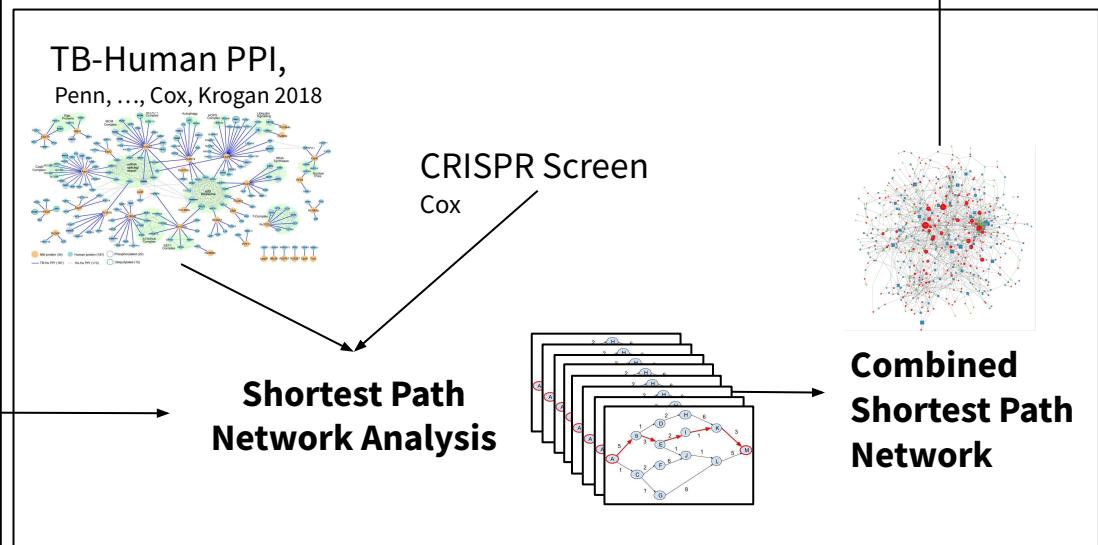
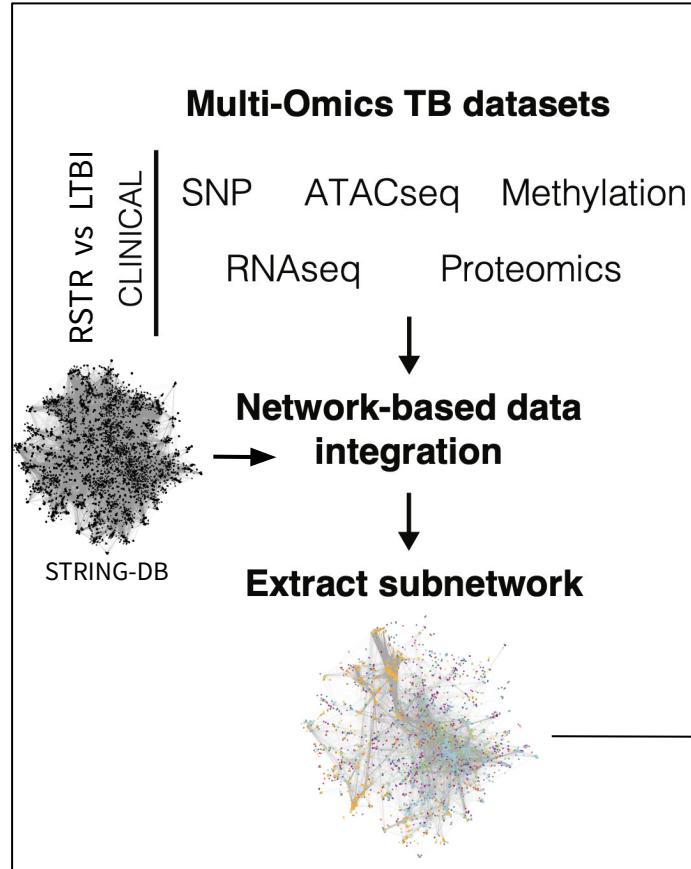
Network colored by interaction with  
TB-Secreted ESX-1 Proteins





**ESAT-6, PE25, and PPE11**  
Interactor Subnetworks

# Overview



**Biological Hypotheses**

# Top Hits - Betweenness

POLR2E	RNA polymerase
BRCA1	DNA repair protein
POLR1A	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates
SF1	nuclear pre-mRNA splicing factor
EIF1AX	eukaryotic translation initiation factor
CENPM	Component of the CENPA-NAC (nucleosome-associated) complex
CENPO	Component of the CENPA-CAD (nucleosome distal) complex
MAPK3	Serine/threonine kinase which acts as an essential component of the MAP kinase signal transduction pathway. MAPK1/ERK2 and MAPK3/ERK1 are the 2 MAPKs which play an important role in the MAPK/ERK cascade
ZC3H18	Enables mRNA cap binding complex binding activity and protein-macromolecule adaptor activity
CDC5L	plays role in pre-mRNA splicing as core component of precatalytic, catalytic and postcatalytic spliceosomal complexes
RTF1	Component of the PAF1 complex (PAF1C) which has multiple functions during transcription by RNA polymerase II and is implicated in regulation of development and maintenance of embryonic stem cell pluripotency
RPL11	component of the ribosome
LRRK2	Serine/threonine-protein kinase which phosphorylates a broad range of proteins involved in multiple processes such as neuronal plasticity, autophagy, and vesicle trafficking
RPL9	component of the ribosome

# Top Hits - number of shortest paths

LRRK2	Serine/threonine-protein kinase which phosphorylates a broad range of proteins involved in multiple processes such as neuronal plasticity, autophagy, and vesicle trafficking
CDC5L	plays role in pre-mRNA splicing as core component of precatalytic, catalytic and postcatalytic spliceosomal complexes
SOD2	mitochondrial protein that forms a homotetramer and binds one manganese ion per subunit.
RPL13	Component of the ribosome
POLR1A	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates
ATP5B	Mitochondrial membrane ATP synthase
UBA52	Ubiquitin A-52 Residue Ribosomal Protein Fusion Product 1
RPS3A	subunit of ribosome
HSP90AA1	Binds bacterial lipopolysaccharide (LPS) and mediates LPS-induced inflammatory response, including TNF secretion by monocytes. Mediates the association of TOMM70 with IRF3 or TBK1 in mitochondrial outer membrane which promotes host antiviral response. Plays a critical role in mitochondrial import
USP22	Histone deubiquitinating component of the transcription regulatory histone acetylation (HAT) complex SAGA.
RPL11	component of the ribosome
HDAC1	Histone deacetylase that catalyzes the deacetylation of lysine residues on the N-terminal part of the core histones
RPL22	component of the ribosome
MRPS10	mitochondrial ribosome component
TP53	In cooperation with mitochondrial PPIF is involved in activating oxidative stress-induced necrosis; the function is largely independent of transcription.
EFTUD2	Component of the spliceosome complex which processes precursor mRNAs to produce mature mRNAs

# HDAC1 - validation

Hit originally published comparing transcriptome of LTBI  
and RSTR macrophages

## Transcriptional networks are associated with resistance to *Mycobacterium tuberculosis* infection

Chetan Seshadri , Nafiseh Sedaghat, Monica Campo, Glenna Peterson, Richard D. Wells, Gregory S. Olson, David R. Sherman, Catherine M. Stein, Harriet Mayanja-Kizza, Ali Shojaie, W. Henry Boom, Thomas R. Hawn, on behalf of the Tuberculosis Research Unit (TBRU) 

# HDAC1 - validation

- HDAC1 Mediates suppression of IL-12B gene expression in macrophages
  - Increased expression in response to infection
  - Increased cleaved caspase-1
  - ESX-1 dependent
  - In our Type 1 IFN - ko → lower IFNb

 | Host-Microbial Interactions | Research Article | 24 February 2021



Transcription Repressor Protein ZBTB25 Associates with HDAC1-Sin3a Complex in *Mycobacterium tuberculosis*-Infected Macrophages, and Its Inhibition Clears Pathogen by Autophagy

# Top Hits - number of shortest paths

LRRK2	Serine/threonine-protein kinase which phosphorylates a broad range of proteins involved in multiple processes such as neuronal plasticity, autophagy, and vesicle trafficking
CDC5L	plays role in pre-mRNA splicing as core component of precatalytic, catalytic and postcatalytic spliceosomal complexes
SOD2	mitochondrial protein that forms a homotetramer and binds one manganese ion per subunit.
RPL13	Component of the ribosome
POLR1A	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates
ATP5B	Mitochondrial membrane ATP synthase
UBA52	Ubiquitin A-52 Residue Ribosomal Protein Fusion Product 1
RPS3A	subunit of ribosome
HSP90AA1	Binds bacterial lipopolysaccharide (LPS) and mediates LPS-induced inflammatory response, including TNF secretion by monocytes. Mediates the association of TOMM70 with IRF3 or TBK1 in mitochondrial outer membrane which promotes host antiviral response. Plays a critical role in mitochondrial import
USP22	Histone deubiquitinating component of the transcription regulatory histone acetylation (HAT) complex SAGA.
RPL11	component of the ribosome
HDAC1	Histone deacetylase that catalyzes the deacetylation of lysine residues on the N-terminal part of the core histones
RPL22	component of the ribosome
MRPS10	mitochondrial ribosome component
TP53	In cooperation with mitochondrial PPIF is involved in activating oxidative stress-induced necrosis; the function is largely independent of transcription.
EFTUD2	Component of the spliceosome complex which processes precursor mRNAs to produce mature mRNAs

# LRRK2

- Among top-scorers in both network measures
- Knockdown was negatively enriched in TNFpos\_iNOSneg
  - Knockout → less TNFa and/or more iNOS

## LRRK2 is a negative regulator of *Mycobacterium tuberculosis* phagosome maturation in macrophages

Anetta Härtlova <sup>1 2</sup>, Susanne Herbst <sup>3 4</sup>, Julien Peltier <sup>1 2</sup>, Angela Rodgers <sup>3</sup>,  
Orsolya Bilkei-Gorzo <sup>1</sup>, Antony Fearn <sup>3</sup>, Brian D Dill <sup>1</sup>, Heyne Lee <sup>5</sup>, Rowan Flynn <sup>5</sup>,  
Sally A Cowley <sup>5</sup>, Paul Davies <sup>1</sup>, Patrick A Lewis <sup>6 7</sup>, Ian G Ganley <sup>1</sup>, Jennifer Martinez <sup>8</sup>,  
Dario R Alessi <sup>1</sup>, Alastair D Reith <sup>9</sup>, Matthias Trost <sup>10 2</sup>, Maximiliano G Gutierrez <sup>11</sup>

## Mitochondrial ROS promotes susceptibility to infection via gasdermin D-mediated necroptosis

Chi G Weindel <sup>1</sup>, Eduardo L Martinez <sup>1</sup>, Xiao Zhao <sup>2</sup>, Cory J Mabry <sup>1</sup>, Samantha L Bell <sup>3</sup>,  
Krystal J Vail <sup>4</sup>, Aja K Coleman <sup>1</sup>, Jordyn J VanPortfliet <sup>1</sup>, Baoyu Zhao <sup>5</sup>, Allison R Wagner <sup>1</sup>,  
Sikandar Azam <sup>1</sup>, Haley M Scott <sup>1</sup>, Pingwei Li <sup>5</sup>, A Phillip West <sup>1</sup>, Jason Karpac <sup>2</sup>, Kristin L Patrick <sup>6</sup>,  
Robert O Watson <sup>7</sup>

# LRRK2

- Associations with susceptibility to *Mycobacterium leprae*
- Increased in expression in response to TB infection
- **Open question - What is the role of functional LRRK2 in TB infection**
  - Phagosome maturation arrest?
  - Oxidative stress?
  - Mitochondrial sensitivity → cell death?

## Genomewide Association Study of Leprosy

Fu-Ren Zhang, M.D., Ph.D., Wei Huang, Ph.D., Shu-Min Chen, M.D., Ph.D., Liang-Dan Sun, M.D., Ph.D.,  
Hong Liu, M.D., Yi Li, Ph.D., Yong Cui, M.D., Ph.D., Xiao-Xiao Yan, M.D., Hai-Tao Yang, M.D., Rong-De Yang,  
M.D., Tong-Sheng Chu, M.D., Chi Zhang, M.D., et al.

**Meta-analysis of human gene expression in response  
to *Mycobacterium tuberculosis* infection reveals  
potential therapeutic targets**

Zhang Wang, Seda Arat, Michal Magid-Slav✉ & James R. Brown✉

# Future Directions

- Network Analysis
  - Keep developing heat propagation method
  - Try different underlying networks such as GIANT or ConsensusPathDB
  - Implement a weighted network
  - Implement k-shortest path algorithm
  - Further narrow shortest path selectivity by CRISPR screen phenotype or TB-secreted protein
- Data Analysis
  - Compare LRRK2 and other mitochondrial hits between the RSTR and LTBI cohorts
    - GWAS, unique snps → lof or gof

# Acknowledgements

**UC San Francisco**

**Krogan Lab** Nevan Krogan, Mehdi Bouhaddou, Ben Polacco, and Max Bennett



**UC Berkeley**

**Cox Lab** Jeff Cox and Ellie Reinhart, CRISPR screens done by Allison Roberts, Nick Garelis, and Bennett Penn



**University of Washington**

**Hawn Lab** Thomas Hawn and Kim Dill-McFarland



**And the entire TBRU,  
especially the data  
collection teams**